European Commission

INSAR

Information Summary on ARchives

EUROPEAN ARCHIVES NEWS

SUPPLEMENT VII

Proceedings

of the DLM-Forum 2002

@ccess and preservation of electronic information: best practices and solutions

DLM-FORUM

Electronic Records

Barcelona, 6–8 May 2002

EUROPEAN
ARCHIVES NEWS
INSAR

- Secretariat-General of the European Commission

- Editorial Committee
  Supplement VII:
  Peter BERNINGER
  Frank BRADY
  Hans HOFMANN
  Jef SCHRAM

- For further information:
  European Commission
  Tel. (32-2) 29-50557
  (32-2) 29-56721
  Fax (32-2) 29-61095
  E-mail: archis@cec.eu.int
  Address: rue de la Loi 200
  Office: JECL 3/35
  B-1049 Brussels

Proceedings

of the

DLM-Forum 2002

@ccess and preservation of electronic information: best practices and solutions

Barcelona, 6–8 May 2002

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server (http://europa.eu.int).

Cataloguing data can be found at the end of this publication.

Luxembourg: Office for Official Publications of the European Communities, 2002

ISBN 92-894-4415-0

Printed in Italy

PRINTED ON WHITE CHLORINE-FREE PAPER

Contents

PREFACE

The DLM-Forum 2002 that took place in Barcelona between 6 and 8 May 2002 was a major event in the interdisciplinary exchange of information on electronic document and records management. The third DLM-Forum overall, it was also the first DLM-Forum of a new generation. The main objective of the DLM-Forum 2002 was to bring forward the most recent expertise and applications in this field, to focus in particular on the access and the preservation of electronic information and to present concrete results and best practices. This time, on a much larger scale than was possible in the previous two DLM-Forums in 1996 and 1999 ([1]).

The Barcelona DLM-Forum 2002 was organised by the Secretariat for Telecommunications and Information Society of the Catalan Government in collaboration with other Catalan institutions and departments of the Spanish central Government. It hosted some 600 experts and decision-makers representing public administration, archives services, research and industry. They came from all EU Member States, candidate countries, and other European countries as well as from Canada, the United States, Japan and Australia.

The Forum's success owed much to the organisational and preparatory work of the DLM-Scientific Committee and the Local Organisation Committee. It also benefited from the support of the EU Member States, especially the EU Presidencies of Sweden, Belgium and Spain (the latter integrated the DLM-Forum 2002 in its official cultural programme), the services concerned of the European Commission (Directorate-General for the Information Society, Secretariat-General) and the ICT industry.

The President of the Catalan Government, Jordi Pujol, opened the DLM-Forum 2002 and welcomed the participants to Barcelona. He noted in particular that the documentary legacy of an institution, body or even a country is part of its history, shapes its present and assures its future. He emphasised the importance of education in schools for the progress of the new technologies and referred to the important role of the Open University of Catalonia which is accessible online for all interested citizens.

In his keynote address, Erkki Liikanen, the European Commissioner for Enterprise and Information Society stressed that public-sector information plays a fundamental role in the proper functioning of the internal market. Serving citizens and businesses well requires, among other things, information-retrieval mechanisms and metadata technology-standardised directories. He emphasised that since the mid-1990s, the DLM-Forum has evolved as a platform for identifying and promoting concrete solutions to many of the problems facing our public administrations.

The message from the International Council on Archives was given by Elisa de Santos, President and Sub-Director of the Spanish States Archives. Reporting on DLM-related activities in Spain and in the International Council on Archives she underlined the contribution of different professions involved in the DLM-Forum 2002, their role in the information society and the support of the Spanish Presidency of the European Union with particular regard to the preservation of digital information.

The plenary opening session provided a first series of high-quality keynote presentations. Given the rapid development and direct impact of modern technologies on archives and also in view of the envisaged enlargement of the European Union, the Director of the Direction des Archives de France, Martine de Boisdeffre, urged the need for greater coordination on archives in Europe. This was also confirmed by the National Archivists of the 15 EU Member States in their Valladolid Declaration and draft resolution of 12 March 2002.

([1]) *Proceedings of the DLM-Forum on electronic records, Brussels, 18–20 December 1996*, Brussels-Luxembourg, 1997; *Proceedings of the DLM-Forum on electronic records, European citizens and electronic information: the memory of the information society, Brussels, 18–19 October 1999*, Luxembourg 2000; see also the DLM web site: http://europa.eu.int/historical_archives/dlm_forum/

(2) For more detail see: http://www.aiim.org/article_aiim.asp?ID=24080

The ICT industry's point of view was introduced by Piero Corsini, Vice-President of IBM Public Sector in Europe, the Middle East and Africa. Stating that electronic information management is in a transitional phase, he referred to the risk of information loss (for example e-mail; web sites), the concept of the entire document life cycle, the capture, management, distribution and access to information and opened perspectives towards a 'next generation Internet' and implications of e-government for the citizen.

The keynote presentation of Ulrich Kampffmeyer, President of Project-Consult GmbH, gives a deeper insight into the partnership between public administration and the ICT industry and should be obligatory reading for everybody interested in document life-cycle management for the European public sector. In addition, he presented, and commented on, the six Industry White Papers (2) which show the ability of the ICT industry to deliver solutions to major requirements and projects in electronic document and records management.

Within the context of the IDA (interchange of data between administrations) programme, Paul E. Murphy from the European Commissions' Directorate-General for Enterprise elaborated on best-practice recommendations for records management with regard to the successful MoReq (model requirements for the management of electronic records) specifications which are already followed in several EU Member States. He also examined the recent developments in metadata for Europe-wide e-government applications.

In his plenary closing speech Erik Ketelaar, Professor of Archivistics at Amsterdam University and former National Archivist of The Netherlands, outlined the broader cultural context of the public-private partnership for the promotion of the preservation and accessibility of the European archival heritage, which has been further reinforced at the DLM-Forum 2002. Who would not be fascinated by 'Ketelaar's law' and the concept of the 'archive as a time machine' covering the present, the past and the future. Another obligatory reading!

The six parallel sessions dealt with main sections of the central theme on access and preservation of electronic information and examined the most recent experiences, best practices and solutions in this field. The first parallel session addressed the memory of the information age, preservation, migration and long-term availability. This topic was approached from different perspectives, bringing forward proven strategies as well as reports on practical archival cases, including major projects in the EU Member States and the European Commission. The second parallel session discussed the use of public information with regard to content, authenticity, protection and intellectual control. The speakers focused on new generation tools provided by 'directory services', the experiences with the Electronic Records Archives Programme of the American National Archives (NARA) and the integrity of content (e.g. through the use of electronic signatures). The third parallel session on content management and delivery concentrated on establishing and maintaining systems, best practices and easy-to-use-access. It gave a precise insight into electronic records management in Europe, covering both public- and private-sector organisations and showing its particular importance for cultural heritage as well as e-commerce. The fourth parallel session looked at one of the top priorities of the DLM-Forum regarding metadata, standardisation and model requirements for organising records and archives. Given the diversity of existing standards, specifications and schemes, the urgent need was stressed to encourage persistently innovative development (e.g. MoReq specifications) and achieve progressively greater coordination by the various actors and disciplines in this field. The fifth parallel session on improving access to knowledge dealt with the impact of modern technologies on new research tools and strategies, content management, training curricula, development of specific skills and competence profiles. Speakers asked for a better understanding between document providers and archive users and to upgrade existing practices. They appreciated in particular the E-TERM (European training programme on electronic records management) which is now operational in different EU Member States and the implications of

e-learning. The sixth parallel session concentrated on the area of information, recognition and retrieval technologies. Speakers dealt with the capture and transformation of information, automatic technologies for indexing and introducing retrieval solutions.

See for more detail the session reports by the three rapporteurs on pages 556, 562 and 569.

In addition, the three workshops gave experts the opportunity to present and discuss practical applications and solutions in more depth. The themes of the workshops concentrated on the use and implementation of metadata standards, the use and implications, and the legal admissibility of digital signatures (pp. 431, 479 and 519).

A new and specific characteristic of the DLM-Forum 2002 was the active involvement of the ICT industry. This has considerably strengthened the common effort towards a lasting partnership in electronic document and records management between the public and private sectors and holds even greater promise for the future. The Barcelona DLM-Forum 2002 clearly indicated that the ICT industry has risen to the challenge issued by the DLM-Forum 1999 in the DLM message to the ICT industry ([3]). The positive reply from the ICT industry and its commitment to address the issues at stake and discuss concrete solutions and best practices with public administrations and archives was one of the main achievements of the Barcelona DLM-Forum 2002. This was stressed in different keynote presentations and discussions during the plenary and parallel sessions, in specific workshops, in demonstrations of available applications in some 40 exhibition stands which complemented and enriched the DLM-Forum.

Furthermore, six Industry White Papers ([4]), prepared by specialist companies and presented at the DLM-Forum 2002, dealt with issues such as user access and protection of information; long-term availability and preservation; education, training and operation. These issues are of direct concern to the preservation and accessibility of archives to the European citizen.

The conclusions of the DLM-Forum 2002 (see page 593) concern the following five main issues.

- Firstly, the continuing reinforcement of the cooperation with the concerned ICT industry to promote best practices and solutions (see also the above-mentioned six Industry White Papers distributed in Barcelona) for electronic document and records management.

- Secondly, the progressive application of specifications on European model requirements for the management of electronic records (MoReq) ([5]) in interested EU Member States and regions.

- Thirdly, the progress on professional education and training on electronic document- and records-management, with particular focus on the European training programme on electronic records management (E-TERM) which is making the first training modules and relevant pedagogic material available (see web site: www.ucl.ac.uk/e-term) to interested EU Member States and candidate countries.

  In this context the ongoing development of competencies, standards and knowledge profiles were stressed. These are aiming towards (a) the enhancement of professional skills for modern archivists in the information society, public administrators and other interested information managers; (b) the implementation of practical solutions; (c) the improvement of recruitment facilities in the EU Member States, regions and institutions.

- Fourthly, the examination of possibilities for the coordination of legal regulations and standards at the regional, national and European level, concerning e-government, e-citizenship and e-commerce which effect the field of archives management.

([3]) *Proceedings DLM-Forum 1999*, pp. 345–348; for the reply from the ICT industry see: *INSAR, European Archives News*, No 8, pp. 1–3.
([4]) For more detail see: http://www.aiim.org/article_aiim.asp?ID=24080
([5]) MoReq, see *INSAR European Archives News*, supplement No VI, Luxembourg, OPOCE, 2002; web site: http://europa.eu.int/ISPO/ida/jsps/index.jsp?fuseAction=showDocument&parent=highlights&documentID=682

- Finally, the launch of the European DLM-network of excellence on electronic archives, which is one of the major results of the DLM-Forum 2002. It is essential to continue to combine interdisciplinary efforts and experiences at a European level. In this way the DLM-Forum has taken on its own dynamic, it has come to maturity, so to say, and shows every sign of being able to stand on its own feet. The newly established DLM-Network Monitoring Committee will further develop and evaluate this DLM-network of excellence, including the application for the sixth framework programme of the European Union.

This preface is signed by the two co-chairpersons of the plenary closing session of the DLM-Forum 2002 in order to stress the above-mentioned DLM-conclusions 2002 and their progressive implementation. The present volume of Proceedings of the DLM-Forum 2002 shall help to inform all those interested in electronic document and records management and to stimulate Europe-wide interdisciplinary discussion and exchange of expertise leading to relevant solutions and best practices in this field. We also wish to take this opportunity to call upon the public sector, the archival profession, ICT industry, all other specialised bodies and interested parties to join the European DLM-network as a new partner/correspondent (for more details please contact e-mail: claes.granstrom@riksarkivet.ra.se; jsymon@aiim.org), to share the benefits of our common DLM enterprise and to contribute actively to the further development of this new European platform on electronic archives.

Eric Norberg
Director-General of the
National Archives of Sweden

Hans Hofmann
Head of the Historical Archives
of the European Commission

DLM-FORUM 2002
Access and preservation of electronic information: Best practices and solutions

Committees

Scientific Committee

The main tasks of the Scientific Committee of the DLM-Forum 2002 were to advise on the quality and substance of the programme for the DLM-Forum 2002, to evaluate the papers and select speakers for the DLM-Forum 2002, to advise on the creation of a European DLM-network of excellence on electronic archives and to detect, evaluate and promote solutions within the context of the DLM-Forum 2002 and its follow-up.

The Scientific Committee was made up of the following members:

Ramon Alberch i Fugueras, Arxiver en Cap, Arxiu Municipal, Ajuntament de Barcelona, Spain
Joan Boadas Rasset, President, Associació d'Arxivers de Catalunya, Arxiu Municial de Girona, Spain
Frank Brady, Head of Unit SG.B.3, European Commission
Nils Brübach, Senior Lecturer, Archivschule Marburg, Germany
María Luisa Conde Villaverde, Directora, Archivo General de la Administración, Spain
Elizabeth Danbury, Senior Lecturer, Director of International Projects and Research, School of Library, Archive and Information Studies, University College London, United Kingdom
Elisa De Santos Canalejo, Subdirector General de los Archivos Estatales, Dirección General del Libro, Archivas y Bibliotecas, Spain
Catherine Dhérent, Conservateur général du patrimoine, Direction des Archives de France
Peter Doorn, Head of Information Technology, Netherlands Institute for Scientific Information Services (NIWI)
Claes Gränström, Deputy Director-General, Swedish National Archives
Cecilia Henriques, Arquivo Nacional da Torre do Tombo (IAN/TT), Instituto dos Arquivos Nacionais, Portugal
Barbara Hoen, Landesarchivdirektion, Baden-Württemberg, Germany
Johannes Hofman, Senior Adviser, Ministry of the Interior, Netherlands
Hans Hofmann, Head of the Historical Archives, European Commission
Per Johansson, Volvo IT, Sweden
Ulrich Kampffmeyer, Geschäftsführer, Project Consult GmbH, Germany
Walter Koch, Professor, Angewandte Informationstechnik Forschungsgesellschaft mbH, Austria
Thijs Laeven, Innogration Management Consultants, The Netherlands
Luis Lopez Corral, Departamento de Ciencias de la Computación, Universidad de Alcalá de Henares, Spain
Philip Lord, Digital Archiving Consultant, United Kingdom
Maria Pia Rinaldi Mariani, Director, Direzioni generali archivi, Ministero beni et attivitá culturali, Italy
Jill Moorcroft, Section Head, Freedom of Information Unit, Information Services Group, Department of Health, United Kingdom
Paul. E. Murphy, Enterprise Directorate-General, European Commission
Tom Quinlan, National Archives of Ireland
Seamus Ross, Director, Advanced Technology and Information Institute, University of Glasgow, United Kingdom
Jef Schram, Secretary of the Scientific Committee of the DLM-Forum 2002, European Commission
Jordi Serra Serra, Arxiu Central Administratiu, Departament de Cultura, Generalitat de Catalunya
Bernard Smith, Head of Unit, Cultural Heritage Applications, Information Society Directorate-General, European Commission
John Symon, Senior Vice-President Europe, AIIM International
Jean-Pierre Teil, Archives Nationales, Centre des Archives Contemporaines, France
Juhani Tikkanen, National Archives of Finland

Heiko Tjalsma, Netherlands Institute for Scientific Information Services (NIWI)
Dan Tørning, National Archives of Denmark
Lucie Verachten, Algemeen Rijksarchief, Belgium
Gerhard K. Wagner, Generalsekretär, Verband für Informationswirtschaft in Österreich (VIW)
Hartmut Weber, President of the Bundesarchiv, Germany
Steve Wells, Departmental Records Officer, Department of Health, United Kingdom

DLM-ICT Working Group

The DLM-ICT Working Group, a sub-group of the DLM-Monitoring Committee, had as its main task to assist the Local Organisation Committee in Barcelona with the preparation and organisation of the DLM-Forum 2002 and to prepare the meetings of the DLM-Scientific Committee.

Claes Gränström, Deputy Director-General Riksarkivet, National Archives of Sweden
Ulrich Kampffmeyer, President, Project-Consult GmbH, Hamburg
John Symon, Senior Vice-President Europe, AIIM International
Josep Manuel Prats Moreno, Secretaria per a la Societat de la Informació, Generalitat de Catalunya
Illana Giner Comin, Secretaria per a la Societat de la Informació, Generalitat de Catalunya
Llanos Gallo Martinez, Secretaria per a la Societat de la Informació, Generalitat de Catalunya
Wim Hamers, Senior Management Consultant Information Services, Unisys Nederland N.V.
Bob Wolfs, Business Group Manager, Agfa-Gevaert N.V.
Jordi Serra Serra, Arxiu Central Administratiu, Departament de Cultura, Generalitat de Catalunya
Luis Lopez Corral, Departamento de Ciencias de la Computación, Universidad de Alcalá de Henares, Spain
Lucy Verachten, Algemeen Rijksarchief, Belgium
Bernard Smith, Head of Unit, Cultural Heritage Applications, Information Society Directorate-General, European Commission
Concepcion Fernandez De La Puente, Information Society Directorate-General, European Commission
Frank Brady, Head of Unit, Secretariat-General, European Commission
Hans Hofmann, Head of the Historical Archives, Secretariat-General, European Commission
Jef Schram, Secretary of the DLM-ICT Working Group, Secretariat-General, European Commission

Local Organisation Committee

The Local Organisation Committee was responsible for the practical organisation of the DLM-Forum 2002 in Barcelona

Generalitat de Catalunya
Ajuntament de Barcelona
AIIM International
Associació d'Arxivers de Catalunya
Collegi Oficial de Bibliotecaris-Documentalistes de Catalunya

With the support of

Spanish Presidency of the European Union
Ministerio de Educación, Cultura y Deporte (Spain)
Ministerio de Administraciones Públicas (Spain)
Ministerio de Cienca y Tecnologia (Spain)
European Commission

DLM-FORUM 2002
@ccess and preservation of electronic information: best practices and solutions

Barcelona, Palau de Congressos de Catalunya, 6–8 May 2002

Pre-conference

Monday 6 May 2002

DLM-Forum 2002 Preliminary Congress of Spanish Archivists

Electronic records, administration and archives

Preparatory meetings

Conference registration

Opening of the exhibition stands

Plenary opening session

Tuesday 7 May 2002

Official welcome and opening of the DLM-Forum 2002	Jordi Pujol, President of the Generalitat of Catalonia
	Erkki Liikanen, European Commissioner for Enterprise and Information Society
	Elisa De Santos Canalejo, President of the International Council on Archives (ICA), Deputy General Director of the State Archives, Ministry of Education, Culture and Sport
Plenary speakers	Martine De Boisdeffre, Director of the Direction des Archives de France
	Piero Corsini, Vice-President of IBM Public Sector Business, Europe, Middle East and Africa
	Ulrich Kampffmeyer, AIIM International European Board Executive Director, President of Project Consult GmbH
	Paul E. Murphy, European Commission, Directorate-General for Enterprise

Parallel sessions

Tuesday 7 May 2002

Parallel session 1	THE MEMORY OF THE INFORMATION AGE Preservation, migration and long-term availability
Parallel session 2	THE USE OF PUBLIC INFORMATION Content, authenticity, protection and intellectual control
Parallel session 3	CONTENT MANAGEMENT AND DELIVERY Establishing and maintaining archival systems, best practices and easy-to-use access

Wednesday 8 May 2002

Parallel session 4	ORGANISING RECORDS AND ARCHIVES Metadata, standardisation and model requirements (MoReq)
Parallel session 5	IMPROVING ACCESS TO KNOWLEDGE Training and education for information experts and users
Parallel session 6	CAPTURING AND TRANSFORMING INFORMATION Automatic indexing technologies and implementing retrieval solutions

Workshops

Tuesday 7 May 2002

Workshop 1	The use and implementation of metadata standards
Workshop 2	The use and implications of digital signatures

Wednesday 8 May 2002

Workshop 3	The legal admissability of digital storage

Plenary closing session

Wedenesday 8 May 2002

Rapporteurs	George Mackenzie, Keeper of the Records of Scotland
	Thekla Kluttig, Home Office of Saxony, Germany
	Catherine Dherent, General curator of national heritage, Direction des Archives de France
Conclusions and follow-up	Erik Norberg, Director-General, National Archives of Sweden
Closing speech	Eric Ketelaar, Professor of archivistics, University of Amsterdam

PLENARY OPENING SESSION

Tuesday 7 May 2002

Plenary opening session of the DLM-Forum 2002

La Sagrada Familia in Barcelona

Official welcome and opening of the DLM-Forum 2002

Opening of the DLM-Forum 2002\ Andreu Mas-Colell, Catalan Minister for Universities, Science and Information Society; Jordi Pujol, President of the Catalan Government; Elisa de Santos, President of the International Council of Archives and Subdirector General de los Archivos Estatales (from left to right)

Jordi Pujol,
President of the Generalitat of Catalonia

Jordi Pujol is president of the Generalitat of Catalonia, a post to which he was elected in 1980 and re-elected in 1984, 1988, 1992, 1995 and 1999. ('Generalitat' is the name of the institutions of the Autonomous Government of Catalonia).

In the first legislative elections to be held in Spain after the restoration of democracy (1977) he won a seat in the Congress in Madrid, where he led his parliamentary group in the interim constitutional period. He was re-elected to the Congress in 1979, but resigned in 1980 in order to stand in the elections to the Parliament of Catalonia as a candidate for the presidency of the Generalitat.

He was born in Barcelona in 1930. From 1946 he was active in the resistance to the regime of General Franco on two counts: in defence of democracy and in defence of the culture and language of Catalonia, as well as of Catalan identity in general, which was harshly persecuted during the dictatorship. Additionally, as early as the 1950s, he started to become active as a pro-European at a time when this stance was rare in Spain. His work was on two fronts, culture and politics.

As a consequence of these activities, he was arrested in May 1960 and accused of organising campaigns in favour of democracy and Catalan nationalism. He was imprisoned until the end of 1962 and then spent a year confined to the city of Girona.

Jordi Pujol's work in government has been crucial for the

Extract from the speech by the President of the Autonomous Government of Catalonia on the occasion of the opening ceremony of the DLM-Forum 2002, Barcelona, 7 May 2002

Jordi Pujol

It is a great pleasure to be here to open the third DLM-Forum, the organisation of which was entrusted to the Secretariat of Telecommunications and Information Society of the Generalitat de Catalonia, in collaboration with the European Commission, local and central government, the archivists' associations and the College of Documentalists.

People from around the world, more than 60 expert speakers, have come together to talk and seek solutions, and to recommend the best practices so that our documentary heritage, which is today kept in electronic form, is not lost and is preserved with the best possible safeguards, yet remaining instantly accessible. Many issues will be dealt with here, from the strictly technical (programmes), to legal problems, and the need to establish a uniform course of action in a global context.

The documentary legacy of an institution, body or even a country is part of its history, shapes its present and assures its future. It is possible to have globalisation while, at the same time, preserving the cultural identity of a nation, thanks to information and communication technologies. Catalonia, a small country where the information society is on the way to being well and truly established, has always been receptive to new ideas, largely due to its geography. It is a country that has embraced and integrated innovation, yet taking care to ensure the Catalan language plays its part in the world of electronic information, and is not perceived as an obstacle for the preservation of, and access to, documents – that it should be seen alongside the other languages within the European Union. The very fact the DLM-Forum is being held in Catalonia seems to me to be an acknowledgement of this fact, and I cannot be anything but pleased that we have all come together here for this event.

I wish you all every success with the Forum and hope you will enjoy your stay in Barcelona, here in Catalonia. It gives me great pleasure to open formally the DLM-Forum 2002.

Extracto del discurso del Presidente de la Generalitat de Catalunya durante la ceremonia de apertura del DLM-Forum 2002, Barcelona, 7 de mayo de 2002

Jordi Pujol

Es para mí un gran placer estar hoy aquí para inaugurar la tercera edición del DLM-Forum, cuya celebración se ha otorgado a la ciudad de Barcelona y cuya organización ha sido concedida a la Secretaría de Telecomunicaciones y Sociedad de la Información de la Generalitat de Catalunya, en estrecha colaboración con la Comisión Europea, el Gobierno local y central, las asociaciones de archiveros y el Colegio de Documentalistas.

Gente de todo el mundo, más de 60 oradores altamente cualificados en la materia, se han reunido para discutir e intentar encontrar soluciones, y para recomendar las mejores prácticas a fin de que la herencia documental, que hoy en día ya se almacena en formato electrónico, no se pierda y se preserve con las máximas garantías de conservación y, por lo tanto,

de accesibilidad. Hoy se tratarán múltiples temas, desde los estrictamente técnicos (de programación), hasta los problemas jurídicos o la necesidad de establecer líneas de actuación uniformes en el contexto de un mundo globalizado.

La herencia documental de una institución, de una entidad o de un país es parte de su historia, forma su presente y es garantía de su futuro. Es posible la globalización y, al mismo tiempo, la preservación de la identidad cultural de una nación a través de las tecnologías de la información y de las telecomunicaciones. En el caso de Cataluña, un pequeño país en el que se está estableciendo plenamente la sociedad de la información; un país siempre abierto a todo tipo de ideas, quizás debido a sus propias características geográficas; un país de acogida e integración, hay que procurar que la lengua catalana esté también presente en la información electrónica, que no sea un obstáculo para la conservación de los documentos o el acceso a ellos, y que la convivencia con otras lenguas de la Unión Europea sea también satisfactoria. El hecho de que el DLM-Forum se celebre aquí, me hace pensar que existe esta voluntad, que hay un reconocimiento, por lo que no puedo sino congratularme por el hecho de que hoy estemos todos aquí reunidos.

Les deseo a todos mucho éxito en el certamen y espero que disfruten de una agradable estancia en Barcelona, en Cataluña. Tengo el gran placer de dar por inaugurado el DLM-Forum 2002.

establishment of Catalan autonomy. In his capacity as president of Catalonia, Jordi Pujol has been extremely active abroad and has throughout received the unfailing support of the Spanish Foreign Ministry. This activity has been political and cultural, but there has also been a significant economic emphasis which has contributed a great deal to the high degree of internationalisation to be witnessed in the Catalan economy today.

In this respect, Jordi Pujol's work in the European context deserves special mention. The Government of Catalonia was concerned as early as 1980 with preparing Catalan society for European integration and it created a channel of information and means of cultivating awareness and study through the Catalan Pro-Europe Association. More recently, the Catalan Institute of Mediterranean Studies was set up, devoted to economic, political and demographic issues concerning the western Mediterranean and the North-west African countries in particular. Jordi Pujol is, moreover, a member of the Action for Europe Committee, successor to the previous Jean Monnet Committee.

Jordi Pujol's pro-Europeanism has always had a strongly regionalist content, which has spanned the founding of the Euro-region with regions neighbouring Catalonia (Midi Pyrenées, Languedoc-Roussillon and Catalonia) to special relationships with regions throughout Europe. His greatest contribution in this respect has been his work with President Edgar Faure, dating from 1985 in the Council of European Regions, later to become the Assembly of European Regions (AER), of which he was vice-president from November 1988 to July 1992. At the General Assembly, held in Santiago de Compostela in July 1992, he was elected President for a two-year term. In Strasbourg, in December 1994, he was re-elected for a second term in office. In accordance with the statutes limiting presidents to two terms in office, he stepped down at the General Assembly held in Basle on 4 December 1996. However, he continues to sit as a member of the bureau in his capacity as out-going President.

Auszug aus der Ansprache des Präsidenten der Autonomen Regierung von Katalonien anlässlich der Eröffnung des DLM-Forums 2002, Barcelona, 7. Mai 2002

Jordi Pujol

Es ist mir eine große Freude, hier mit Ihnen zu sein, um das dritte DLM-Forum zu eröffnen. Die Organisation des Forums war dem Sekretariat für Telekommunikation und Informationsgesellschaft der Generalitat de Catalunya in Zusammenarbeit mit der Europäischen Kommission, der regionalen und zentralen Regierung, den Archivarsvereinen und der Hochschule für Dokumentalisten anvertraut.

Personen aus aller Welt, mehr als 60 sachverständige Redner, sind hier zusammengekommen, um zu diskutieren und Lösungen zu finden. Auf diese Weise sollen die besten Praktiken empfohlen werden, damit unser dokumentarisches Erbe, das heute in elektronischer Form gehalten wird, nicht verloren geht, sondern unter den bestmöglichen Sicherheitsbedingungen dauernd zugänglich und erhalten bleibt. Viele Fragen werden hier behandelt, von technischen Aspekten und Programmen bis hin zu rechtlichen Problemen und der Notwendigkeit, eine einheitliche Handlungsweise in einem globalen Zusammenhang festzulegen.

Das dokumentarische Vermächtnis einer Institution, einer Einrichtung oder auch eines Landes ist Teil seiner Geschichte, formt seine Gegenwart und sichert seine Zukunft. Durch die Informations- und Kommunikationstechniken ist es möglich, sowohl die Globalisierung zu haben als auch zur selben Zeit die kulturelle Identität einer Nation zu bewahren. Katalonien ist ein kleines Land, in dem die Informationsgesellschaft auf gutem Weg ist verwirklicht zu werden; es war stets offen für neue Ideen, vornehmlich auch durch seine geographische Lage. Es ist ein Land, das die Innovation umfasst und integriert hat, jedoch auch

There has obviously been a political component to all this work, but there has also been an intellectual element. Jordi Pujol is the author of various books, some of which focus on the Catalan question, such as, among others: Fer poble, fer Catalunya, published clandestinely in 1965, Des dels turons a l'altra banda del riu (writings from prison), published in 1978, Construir Catalunya (1979), Als joves de Catalunya (1988), and La força serena i constructiva de Catalunya (1991). Among those works that refer to the Spanish question in general are: Los desequilibrios territoriales en España (1978) and Estado y Sociedad (1987). Finally, there are those that look at the theme of Europe and, frequently, the regions of Europe: Afirmación catalana de europeísmo (1985), Apport et rôle des régions dans la construction européenne (1985), The regions of Europe in the perspective of 1992 (1989), El papel de la Europa mediterránea (1990), Pensar Europa (1993), etc.

darum bemüht ist sicherzustellen, dass die katalanische Sprache, die neben den anderen Sprachen in der Europäischen Union gesehen werden sollte, ihre Rolle in der Welt elektronischer Informationen spielt und nicht als ein Hindernis für die Aufbewahrung und den Zugang zu Dokumenten empfunden wird. Die Tatsache, dass das DLM-Forum in Katalonien gehalten wird, scheint mir eine derartige Anerkennung zu sein, und ich bin außerordentlich erfreut darüber, dass wir alle zu diesem Forum hier zusammengekommen sind.

Ich wünsche dem Forum viel Erfolg und hoffe, dass Sie Ihren Aufenthalt in Barcelona, hier in Katalonien, genießen werden. Es ist mir eine große Freude, das DLM-Forum 2002 offiziell zu eröffnen.

Extrait du discours du président du gouvernement autonome de la Catalogne à l'occasion de la cérémonie d'ouverture du Forum DLM 2002 le 7 mai 2002 à Barcelone

Jordi Pujol

C'est un grand plaisir pour moi d'être ici afin d'ouvrir le troisième Forum DLM, dont l'organisation a été confiée au Secrétariat des télécommunications et de la société de l'information de la Generalitat de Catalunya, en collaboration avec la Commission européenne, le gouvernement local et le gouvernement central, les associations d'archivistes et le Collège des documentalistes.

Des participants du monde entier, dont plus de 60 orateurs experts en la matière, sont réunis pour parler, chercher ensemble des solutions et recommander les meilleures pratiques de sorte que notre patrimoine documentaire, aujourd'hui conservé sous format électronique, ne soit pas perdu et soit préservé avec toutes les garanties possibles tout en restant accessible. Beaucoup de questions seront traitées ici, certaines strictement techniques, d'autres portant sur les programmes ou les problèmes juridiques, d'autres encore sur la nécessité d'établir une ligne de conduite uniforme dans un contexte global.

Le legs documentaire d'une institution, d'un organisme voire d'un pays fait partie de son histoire, forme son présent et assure son avenir. Il est possible de réaliser la mondialisation et en même temps de préserver l'identité culturelle d'une nation grâce aux technologies de l'information et des communications. La Catalogne, un petit pays où la société de l'information est bel et bien en train de s'établir, a toujours été réceptive aux nouvelles idées, en grande partie grâce à sa position géographique. C'est une région qui a compris et a intégré l'innovation, tout en s'assurant que la langue catalane joue son rôle dans le monde des données numériques, qu'elle ne soit pas perçue comme un obstacle à la conservation et à l'accès aux documents mais prise en compte à côté des autres langues au sein de l'Union européenne. Le fait même que le Forum DLM ait lieu en Catalogne me semble être une reconnaissance de ce fait, et je me félicite que nous soyons tous ensemble réunis ici pour cet événement.

Je souhaite que ce forum soit un grand succès pour vous et que votre séjour à Barcelone, ici en Catalogne, soit agréable. Cela me donne ainsi le grand plaisir d'ouvrir formellement le Forum DLM 2002.

THE MEMORY OF THE INFORMATION SOCIETY ([1])

Erkki Liikanen

Ladies and gentlemen,

I am very pleased to address you today at the opening session of the DLM-Forum 2002.

Competitiveness of Europe's companies and new knowledge-rich jobs are key to the European digital economy. Equally important, the information society must reinforce the core values of Europe's social and cultural heritage. This means that information and communication technologies must support equality of access, social inclusion and cultural diversity.

Cooperation between the public sector and the private sector, the ICT industry, should help to balance these social and economic imperatives. Over the past years, the European Commission with EU Member States, has promoted such cooperation.

Public administrations aim to become more transparent. They are looking for ways to better inform the citizen and to retain the collective memory of the information society.

Public-sector information plays a fundamental role in the proper functioning of the internal market. Serving citizens and businesses well requires information retrieval mechanisms, metadata technology, standardised directories, etc.

Since the mid-1990s, the DLM-Forum has evolved as a platform for identifying and promoting concrete solutions to many of the problems facing our public administrations.

In 1996 the initial focus of the DLM-Forum was on guidelines for best practices, machine-readable data and electronic documentation.

At the DLM-Forum in Brussels in 1999 the ICT industries were challenged to cooperate with public administrations. They were asked to provide proven and practical solutions for electronic document and content management and for the preservation of electronic information.

The positive response of the ICT industries is reflected in their participation in this conference.

* * *

You will spend the next two days on best practices and solutions for the preservation and long-term accessibility of electronic information.

These will be crucial to preserve the 'memory of the information society'. They are also essential for more effective government.

Solutions are needed that guarantee short- and long-term accessibility and the intelligent retrieval of the knowledge stored in document management and archival systems. These solutions also need to be adaptable to rapid technological advances. Training and educational programmes are needed as well.

The content of public archives is part of the knowledge heritage of mankind, for present and future generations. Safeguarding and ensuring the continued accessibility of the European archival heritage in the information society is of primordial importance. It is therefore the primary concern of the DLM-Forum 2002.

This requires making the right choices between existing standards and practices. We should encourage further cooperation between the concerned parties.

* * *

Erkki Liikanen

Commissioner for Enterprise and Information Society
European Commission

Nationality
Finnish

Personal details
Born in Mikkeli, Finland, on 19 September 1950
Married, two children

Education
Masters degree in political science (economics), University of Helsinki

Political career
1972–90 — Member of Parliament; member of Cultural Affairs Committee (1972–75), Agriculture and Forestry Committee (Vice-Chairman) (1977–79), Foreign Affairs Committee (member 1975–82; Chairman 1983–87)
1976–79 — Member of the Supervisory Board of Televa Oy
1978, 1982, 1988 — Elector of the President of the Republic
1980–89 — Member then Chairman of Supervisory Board of Outokumpu
1981–87 — Secretary-General of the Social Democratic Party
1983–87 — Parliamentary Trustee to the Bank of Finland (Vice-Chairman), Speaker's Council
1987–90 — Minister of Finance
1990–94 — Ambassador Extraordinary and Plenipotentiary, Head of Finnish Mission to the European Union
1995–99 — Member of the Commission for Budget; Personnel and Administration; Translation and In-house Computer Services
1999- Member of the Commission for Enterprise and Information Society

([1]) The keynote address by Commissioner Liikanen was given as a video presentation.

I want to thank the organisers, contributors, exhibitors and the participants for their contribution to the DLM-Forum 2002. Let us together strengthen interdisciplinary cooperation. In this way we will help to safeguard the active and lasting memory of the information society for all citizens in Europe.

Thank you and enjoy your conference.

La memoria de la sociedad de la información ([1])

Erkki Liikanen

Señoras y señores:

Me complace intervenir ante ustedes en la sesión inaugural del Forum DLM 2002.

La competitividad de las empresas europeas y los nuevos empleos basados en el conocimiento son un elemento clave de la economía digital europea. También es importante que la sociedad de la información refuerce los valores principales del patrimonio social y cultural de Europa. Esto significa que las tecnologías de la información y de las comunicaciones (TIC) deben promover la igualdad de acceso, la inclusión social y la diversidad cultural.

La cooperación entre el sector público y el privado, es decir, la industria de las TIC, debería contribuir a **equilibrar estos imperativos sociales y económicos.** La Comisión Europea y los Estados miembros de la Unión Europea han promovido dicha cooperación durante los últimos años.

El objetivo de las administraciones públicas es conseguir una mayor transparencia. Buscan métodos para informar mejor al ciudadano y salvaguardar la memoria colectiva de la sociedad de la información.

La información del sector público desempeña un papel fundamental en el correcto funcionamiento del mercado interno. Para ofrecer un buen servicio tanto a los ciudadanos como a las empresas, es necesario contar con mecanismos de recuperación de la información, tecnología de metadatos, directorios normalizados, etc.

Desde mediados de los años noventa, el Forum DLM ha evolucionado convirtiéndose en una plataforma para identificar y promover soluciones concretas para muchos de los problemas que deben resolver nuestras administraciones públicas.

El enfoque inicial del Forum DLM en 1996 se centraba en conseguir directrices para mejores prácticas, datos de lectura automática y documentación electrónica.

En el Forum DLM de Bruselas en 1999 se instó a las industrias de las TIC a cooperar con las administraciones públicas. Se les pidió que aportaran soluciones prácticas y eficaces para la gestión de contenidos y documentos electrónicos y para preservar la información electrónica.

La respuesta positiva de la industria de las TIC se ve reflejada en su participación en esta conferencia.

* * *

En los próximos dos días se debatirá sobre mejores prácticas y soluciones para la preservación y accesibilidad a largo plazo de la información electrónica.

([1]) El discurso de Comisario Liikanen se dio como presentación en vídeo.

Todo ello será crucial para conservar «la memoria de la sociedad de la información», así como para conseguir un gobierno más eficaz.

Se necesitan soluciones que garanticen la accesibilidad a largo y corto plazo y la recuperación inteligente del conocimiento almacenado en sistemas de archivos y gestión de documentos. Estas soluciones también deben adaptarse a los rápidos avances tecnológicos. Asimismo, se necesitan programas educativos y de formación.

El contenido de los archivos públicos forma parte del patrimonio de conocimiento de la humanidad para las generaciones presentes y futuras. Es sumamente importante salvaguardar y garantizar la accesibilidad continuada del patrimonio archivístico europeo en la sociedad de la información. Por consiguiente, ésta es la principal preocupación del Forum DLM 2002.

Todo ello requiere tomar **buenas decisiones** con respecto a las normas y prácticas existentes. Debemos fomentar una mayor cooperación entre las partes implicadas.

* * *

Agradezco a los organizadores, colaboradores, expositores y participantes su contribución al Forum DLM 2002. Juntos debemos reforzar la cooperación interdisciplinaria. Así contribuiremos a preservar la memoria activa y duradera de la sociedad de la información para todos los ciudadanos europeos.

Gracias y que disfruten de la conferencia.

Das Gedächtnis der Informationsgesellschaft ([1])

Erkki Liikanen

Sehr geehrte Damen und Herren,

ich freue mich, Sie heute zur Eröffnungsveranstaltung des DLM-Forums 2002 willkommen zu heißen.

Die Wettbewerbsfähigkeit der europäischen Unternehmen und neuartige wissensintensive Dienstleistungen sind für die europäische digitale Wirtschaft von entscheidender Bedeutung. Genauso entscheidend ist, dass die Informationsgesellschaft zur Stärkung der Grundwerte des sozialen und kulturellen Erbes Europas beiträgt. Das bedeutet, dass die Informations- und Kommunikationstechniken den gleichberechtigten Zugang zu den IK-Technologien, die soziale Einbeziehung und die kulturelle Vielfalt fördern müssen.

Die Zusammenarbeit zwischen dem öffentlichen Sektor und der Privatwirtschaft, der IKT-Industrie, sollte zu einem **ausgewogenen Verhältnis zwischen diesen sozialen und wirtschaftlichen Geboten beitragen.** Die Europäische Kommission hat sich in den vergangenen Jahren gemeinsam mit den EU-Mitgliedstaaten für diese Zusammenarbeit eingesetzt.

Die öffentlichen Verwaltungen der Mitgliedstaaten streben nach größerer Transparenz. Sie suchen nach Möglichkeiten, den Bürger besser zu informieren und das Kollektivgedächtnis der Informationsgesellschaft zu erhalten.

([1]) Die Ansprache von Kommissar Liikanen erfolgte als Videopräsentation.

Die Informationstätigkeit des öffentlichen Sektors ist für das Funktionieren des Binnenmarktes von grundlegender Bedeutung. Ein effizienter Dienst am Bürger und an Unternehmen setzt allerdings Information-Retrieval-Systeme, Metadatentechnologien, genormte Verzeichnisse usw. voraus.

Seit Mitte der 90er Jahre sind im Rahmen des DLM-Forums für zahlreiche Probleme, denen sich unsere öffentlichen Verwaltungen gegenübersehen, praxisnahe Lösungen erarbeitet und gefördert worden.

1996 lag der Schwerpunkt des DLM-Forums auf der Ausarbeitung von Leitlinien für vorbildliche Verfahren, maschinenlesbare Daten und Dokumentation auf elektronischen Datenträgern.

Auf dem DLM-Forum 1999 in Brüssel wurde die IKT-Industrie zur Zusammenarbeit mit der öffentlichen Verwaltung aufgefordert und gebeten, bewährte und praxisbezogene Lösungen für die Verwaltung der Inhalte sowie für die Verwaltung und Erhaltung elektronisch gespeicherter Dokumente bereitzustellen.

Die positive Antwort der IKT-Industrie kommt in ihrer Teilnahme an der heutigen Konferenz zum Ausdruck.

* * *

Sie werden sich in den kommenden zwei Tagen mit vorbildlichen Verfahren und Lösungen zum Thema „Erhaltung und langfristiger Zugriff auf elektronisch gespeicherte Informationen“ auseinandersetzen.

Diese Überlegungen und Ergebnisse werden für die Erhaltung des „Gedächtnisses der Informationsgesellschaft“ von entscheidender Bedeutung sein. Sie werden auch für eine effizientere Form des Regierens maßgeblich sein.

Wir brauchen Lösungen, die den kurz- und langfristigen Informationszugriff und das Auffinden von Wissen gewährleisten, das in Dokumentenverwaltungs- und Archivsystemen gelagert ist. Diese Lösungen müssen an den raschen technischen Fortschritt angepasst werden können. Darüber hinaus brauchen wir entsprechende Bildungs- und Ausbildungsprogramme.

Der Inhalt öffentlicher Archive ist Teil des Erbes „Wissen der Menschheit“ dieser und künftiger Generationen. Die Erhaltung und Sicherstellung des dauerhaften Zugriffs auf die europäischen Archivalien in der Informationsgesellschaft ist von grundlegender Bedeutung und daher auch das Hauptanliegen des DLM-Forums 2002.

Dies setzt voraus, dass wir die **richtigen Entscheidungen** zwischen den vorhandenen Normen und Verfahren treffen. Wir sollten eine weiter gehende Zusammenarbeit zwischen allen Beteiligten fördern.

* * *

Mein Dank gilt den Veranstaltern, Ausstellern und allen Teilnehmern, die einen Beitrag zu dem DLM-Forum 2002 geleistet haben. Lassen Sie uns die interdisziplinäre Zusammenarbeit gemeinsam ausbauen. So werden wir unseren Beitrag zur Erhaltung des aktiven und unvergänglichen Gedächtnisses der Informationsgesellschaft für alle Bürger in Europa leisten.

Ich danken Ihnen und wünsche Ihnen eine interessante Konferenz.

La mémoire de la société de l'information ([1])

Erkki Liikanen

Mesdames, Messieurs,

C'est avec grand plaisir que je prends la parole devant vous aujourd'hui, à l'occasion de la séance d'ouverture du Forum DLM 2002.

Des entreprises européennes compétitives et de nouveaux emplois faisant intervenir de multiples connaissances sont des éléments essentiels de l'économie numérique européenne. Il est tout aussi important que la société de l'information étaye les valeurs fondamentales du patrimoine social et culturel européen. En d'autres termes, les technologies de l'information et des communications doivent favoriser l'égalité d'accès, l'intégration sociale et la diversité culturelle.

La **coopération entre le secteur public et le secteur privé** — l'industrie des TIC — doit permettre de **trouver un juste milieu entre ces impératifs sociaux et économiques.** Ces dernières années, la Commission européenne s'est efforcée d'encourager cette coopération, en association avec les États membres de l'Union européenne.

Les administrations publiques tendent vers plus de transparence. Elles recherchent des moyens de mieux informer les citoyens et de conserver la mémoire collective de la société de l'information.

Les informations émanant du secteur public jouent un rôle fondamental dans le bon fonctionnement du marché intérieur. Offrir des services de qualité aux citoyens et aux entreprises nécessite des mécanismes de recherche de l'information, des technologies en matière de métadonnées, des répertoires normalisés, etc.

Depuis le milieu des années 90, le Forum DLM est devenu une plate-forme destinée à trouver et à promouvoir des solutions concrètes aux nombreux problèmes auxquels se heurtent nos administrations publiques.

En 1996, le Forum DLM s'est attaché à élaborer des orientations concernant les meilleures pratiques, les données lisibles par machine et les documents électroniques.

Lors du Forum DLM qui s'est tenu à Bruxelles en 1999, l'industrie des TIC a été mise au défi de coopérer avec les administrations publiques. Elle devait fournir des solutions pratiques et avérées dans le domaine de la gestion des documents électroniques et du contenu ainsi que de la conservation de l'information électronique.

La participation de l'industrie des TIC à la conférence d'aujourd'hui concrétise l'accueil favorable que l'industrie a réservé à cette demande.

* * *

Au cours des deux prochains jours, vous allez examiner les meilleures pratiques et solutions en termes de conservation et d'accessibilité à long terme de l'information électronique.

Ces solutions seront décisives, tant pour conserver la «mémoire de la société de l'information» que pour accroître l'efficacité de l'administration.

([1]) Le discours enregistré de Erkki Liikanen a été diffusé en vidéo.

Des solutions doivent être trouvées pour garantir l'accessibilité à court et à long terme et la recherche intelligente des connaissances qui figurent dans des systèmes d'archives et de gestion des documents. Ces solutions doivent pouvoir être adaptées aux progrès rapides de la technologie. Des programmes d'enseignement et de formation sont également nécessaires.

Le contenu des archives publiques fait partie du patrimoine de connaissances de l'humanité, pour les générations actuelles et futures. Il est primordial de veiller à ce que l'héritage archivistique européen soit accessible en permanence dans la société de l'information. Telle est la préoccupation première du Forum DLM 2002.

Cela suppose de faire les **bons choix** entre les normes et les pratiques existantes. Nous devons encourager les parties concernées à coopérer plus étroitement.

* * *

Je tiens à remercier les organisateurs, les auteurs des contributions, les exposants et les participants pour leur présence au Forum DLM 2002. Renforçons ensemble la coopération interdisciplinaire. Nous pourrons ainsi contribuer à préserver la mémoire active et durable de la société de l'information pour l'ensemble des citoyens européens.

Je vous remercie. Bonne conférence à tous.

COMMUNICATION FROM THE INTERNATIONAL COUNCIL ON ARCHIVES

Elisa de Santos Canalejo

Right Honourable President of the Generalitat of Catalonia, Honourable Councillor for Universities, Research and the Information Society,

Ladies and gentlemen, dearest colleagues,

First I would like to thank the institutions who have organised this event for their kind invitation to attend this opening ceremony of the DLM-Forum.

I cannot deny that it gives me great satisfaction to see that archives are producing a growing interest among professionals in highly varied fields, which is a clear reflection of their role in the information society. It is no wonder that the 14th International Congress on Archives held in Seville in the year 2000 had the slogan, 'Archives of the new millennium in the information society.' The International Congress on Archives, of which I have the honour of being President until the year 2004, wanted the Seville Congress to map out the courses of action to be developed during these years.

I would therefore like to highlight that with the Spanish presidency of the ICA coinciding with the Spanish presidency of the European Union, we resolve to lend our support for the new challenges faced by archives in the field of digital information.

The programme over these two days is highly intense and ambitious, covering the two sides to the electronic document: (1) the electronic document itself, as the origin; and (2) the electronic document as an alternative format to the original. And as has often been the case over the last few years, these archiving meetings are not restricted to the regional level, which would have been the European Union in this case, but instead welcome the presence of professionals from other continents. Nor are these meetings limited to archivists, which would make no sense in this day and age: they also include professionals in the new technologies.

The fact that Barcelona is the venue for the DLM-Forum has encouraged the creation of a Spanish Working Group and the organising of a pre-congress for Spain and Latin America, which was held yesterday as a prelude to this meeting. This Spanish Working Group will no doubt contribute to the management of electronic documents and archives in the Public Administrations, but it has also been formed to support the work of the ICA's Electronic Documents Committee, which includes archivists from Spain and Portugal. This committee is currently preparing a 'Guide to electronic documents from the archiving point of view,' an updated version of the guide published in 1997 and which shall be presented at the 15th International Congress on Archives to be held in Vienna in the year 2004.

The President of the Latin-American Archiving Association (ALA) has also been invited to the pre-congress and this DLM-Forum. One of the regional branches of the International Council on Archives, this association devoted much attention to electronic administration in Ibero-America at its annual seminar held a few weeks ago in Santiago de Chile.

The year 2002 has so far seen more initiatives concerning the information society: at the end of this month the Council's Executive Committee will be meeting in Peking and shall use the opportunity to present the Unesco Programme known as 'Information for all' to the regional branches of Asia and the Pacific. A year ago, in my capacity as President of the ICA I attended the presentation of this programme at Unesco headquarters. The Intergovernmental Committee of this programme, of which Spain is a member, met in April and is working on a charter for the preservation of digital heritage, with clearly defined aims: the preservation of information and universal access to it but, furthermore, with the participation of all in the emerging information society.

Elisa Carolina de Santos Canalejo

Elisa Carolina de Santos Canalejo is Deputy General Director of the State Archives of the Ministry of Education, Culture and Sport since 1997. She is a Doctor cum laude of Medieval History and since 1986, member of staff of the Faculty of Archivists, Librarians and Archaeologists (Archives Department),.decorated by the Ministry of Foreign Affairs with the Award of the Order of Civil Merit.

Other posts and positions held:
— President of the International Council on Archives (ICA) from 2000 to 2004.
— One of the Vice-Presidencies of the Latin-American Association of Archives since 1999.
— Member of the Intergovernmental Committee of the Support Programme for the Development of Latin-American Archives.
— Director of the State Archives Network Project.
— Former member of the Committee on Archival Descriptions Standards (ICA/CDS), 1997–2000.
— Director of the General Archives of the Ministry of Foreign Affairs from 1990 to 1997 and formed part of the Group of Archivists of the Diplomatic Archives of European Union Countries and Institutions.
— Head of Public Administration Relations from 1986 to October 1990 in the General Archive of Administration.
— Participated in the 1999 DLM-Forum with work on 'The contribution of the archives to the information society'.
— Teaching experience in archive administration.
Author of many publications including the following:

— Guide to the archive of the Ministry of Foreign Affairs, *Madrid, Ministry of Foreign Affairs, 1997, 229 pp.*
— National archive system design, the Spanish archive system, *ALA, 1997.*
— Archive Thesaurus, *published by the Ministry of Foreign Affairs and the only one existing for the diplomatic archives of European Union and Latin-American countries, (1994, 474 pp.).*
— Indexing in information retrieval, *in the Fifth European Conference on Archives, May 1997, Janus, pp. 116–131.*
— Archives and diplomacy in 'contemporary history', *University of the Basque Country, 1996.*
— *Archive of the Ministry of Foreign Affairs as sources for the history of international relations, in conferences on the history of international relations, Madrid, October, 1994.*
— Considerations on access to diplomatic archives in the 'Boletín de la Anabad' *(Spanish Association of Archivists, Librarians, Curators and Documentary-Makers' Gazette), XLI, (1991) Nos 3–4, pp. 195–200.*
— Spanish local administration (1940–78), *Archive Group of the Ministry of the Interior, Ministry of the Interior, 1990, 209 pp.*

This programme focuses particularly on digital information created for distribution over the web but also forms a starting point for multidisciplinary working groups to tackle the different aspects of digital documentary heritage. Working groups which, indeed, have already been working in the heart of the European Union. I am confident that in Peking the Executive Committee of ICA shall be able to analyse closely this Unesco Programme, which is highly positive from an archiving point of view.

Within this framework, it is no wonder that the theme chosen by the Spanish Presidency of the EU for cultural affairs has been the long-term preservation of digital memory and, in fact, at the end of this very month the Council is to approve a resolution on the preservation of digital memory.

The draft resolution was presented before the national archives directors of the EU and candidate countries at the meeting held in Valladolid on 11 and 12 March. Many of these directors have joined us for this meeting here today and all are members of the regional branch of the ICA, Eurbica. Thus, our ability to discuss these matters here in Europe is assured from different quarters. The DLM-Forum is another step towards tackling the treatment necessary to guarantee the conservation of digitalised documentary heritage.

In short, national and international policies design plans and programmes to bring new technologies closer to citizens and to favour access to information, and there can be no doubt that archives are taking on the role that corresponds to them in electronic administration, since without the archiving point of view, the memory of this new society would doubtlessly be otherwise lost.

Comunicación del Consejo Internacional de Archivos

Elisa de Santos Canalejo

Sr. Presidente de la Generalidad de Cataluña, Sr. Consejero de las Universidades, la Investigación y la Sociedad de la Información,

Señoras y señores, queridos colegas:

En primer lugar, permítanme agradecer a las instituciones que han organizado este acontecimiento su amable invitación para asistir a esta ceremonia de apertura del Foro DLM.

No puedo ocultar la gran satisfacción que me produce ver el creciente interés que suscitan los archivos entre los profesionales de sectores muy diferentes, lo que demuestra muy claramente su papel en la sociedad de la información. Nadie se extrañará de que el XIV Congreso internacional de archivos, celebrado en Sevilla en 2000, tuviera como título: «Los archivos del nuevo milenio en la sociedad de la información». El Consejo Internacional de Archivos (CIA), cuya Presidencia tengo el honor de asumir hasta 2004, deseaba que el congreso de Sevilla esbozase las acciones que deben realizarse durante estos años.

Quisiera destacar que, habida cuenta de la coincidencia de la Presidencia española del CIA con la Presidencia española de la Unión Europea, decidimos dar nuestro apoyo a los nuevos retos que deben afrontar los archivos en el ámbito de la información digital.

El programa de estos dos días, muy intenso y ambicioso, cubre los dos aspectos del documento electrónico: 1) el documento electrónico en sí, como original; y 2) el documento electrónico como ejemplar que sustituye al original. Tal como se ha observado en estos últimos

años, estas conferencias sobre los archivos no se limitan al ámbito regional, que habría sido la Unión Europea en este caso, sino que acogen a profesionales de otros continentes. Del mismo modo, estos encuentros no van dirigidos sólo a archiveros, lo que no tendría ningún sentido en esta era, sino que también están abiertos a profesionales de las nuevas tecnologías.

La elección de Barcelona como lugar de organización del Foro DLM fomentó la creación de un grupo de trabajo español y la organización de un congreso preparatorio destinado a España y América Latina, que tuvo lugar ayer como preparación al encuentro de hoy. El grupo de trabajo español contribuirá sin duda alguna a la gestión de documentos y archivos electrónicos en las administraciones públicas, pero también se constituyó para apoyar los trabajos del Comité de documentos electrónicos del CIA, que cuenta entre sus miembros con archiveros españoles y portugueses. Este Comité prepara actualmente una guía sobre los documentos electrónicos desde el punto de vista de los archivos, guía que constituye una versión actualizada de la publicada en 1997 y que se presentará en el XV Congreso internacional de archivos, que se celebrará en Viena en 2004.

El Presidente de la Asociación de Archivos de América Latina (ALA) ha sido invitado al congreso preparatorio y a este Foro DLM. En su calidad de oficina regional del CIA, esta asociación dedicó gran parte de sus trabajos a la administración electrónica en América Latina en su seminario anual, organizado hace algunas semanas en Santiago de Chile.

En 2002 se han lanzado un gran número de iniciativas relativas a la sociedad de la información: al final de este mes, el Comité ejecutivo del CIA se reunirá en Pekín y aprovechará esta oportunidad para presentar el programa de la Unesco titulado «Información para todos» en las oficinas regionales de Asia y del Pacífico. En mi calidad de Presidenta del CIA, asistí hace un año a la presentación de este programa en la sede de la Unesco. La comisión intergubernamental del programa, de la que España es miembro, se reunió en abril y trabaja actualmente en una carta para la conservación del patrimonio digital, que persigue objetivos claramente definidos: conservación de la información y universalidad del acceso a ésta, con la participación de todos en la sociedad de la información emergente.

Este programa se centra en particular en la información digital creada para su divulgación en la red, pero ofrece también una plataforma de partida a grupos de trabajo pluridisciplinares para abordar los distintos aspectos del patrimonio digital. Estos grupos de trabajo ya funcionan de hecho en la Unión Europea. Estoy segura de que el Comité ejecutivo del CIA estará en condiciones en Pekín de analizar de cerca este programa de la Unesco, que es muy positivo desde el punto de vista de los archivos.

En este contexto, nadie se extrañará de que la conservación a largo plazo de la memoria digital sea el tema escogido por la Presidencia española de la Unión Europea. De hecho, el Consejo deberá adoptar al final de este mes una resolución sobre la conservación de la memoria digital.

El proyecto de resolución se presentó a los directores de los archivos nacionales de los países de la Unión Europea y de los países candidatos en la reunión de 11 y 12 de marzo último en Valladolid. La mayoría de estos directores están hoy presentes con nosotros y son todos ellos miembros de la oficina regional del CIA (EURBICA). Así pues, nuestra capacidad para tratar estas cuestiones en Europa está garantizada en muchos aspectos. El Foro DLM constituye una nueva etapa hacia las medidas necesarias para garantizar la conservación del patrimonio digital.

En conclusión, las políticas nacionales e internacionales conciben proyectos y programas que permiten acercar a los ciudadanos a las nuevas tecnologías y favorecer su acceso a la información. No cabe duda de que los archivos asumen su función en la administración electrónica, ya que, sin el punto de vista del archivero, la memoria de esta nueva sociedad se perdería sin duda.

Mitteilung des Internationalen Archivrats

Elisa de Santos Canalejo

Sehr geehrter Herr Präsident der Generalitat von Katalonien, sehr geehrter Herr Rat für Universitäten, Forschung und Informationsgesellschaft,

meine Damen und Herren, verehrte Kollegen,

zunächst möchte ich den Institutionen, die diese Veranstaltung ausrichten, ganz herzlich für die Einladung zu dieser Eröffnungsfeier des DLM-Forums danken.

Ich muss zugeben, es erfüllt mich mit großer Genugtuung, wenn ich sehe, dass die Archivthematik immer mehr Interesse bei den Vertretern der verschiedensten Fachgebiete findet, was ein eindeutiges Zeichen für die Rolle der Archive in der Informationsgesellschaft ist. So ist es kein Wunder, dass der XIV. Internationale Archivkongress, der im Jahr 2000 in Sevilla stattfand, unter dem Motto „Archive und die Informationsgesellschaft im neuen Jahrtausend" stand. Ziel des Internationalen Archivrats (ICA), dessen Präsidentin bis 2004 zu sein ich die Ehre habe, war es, dass der Kongress von Sevilla den Kurs für die Aktionen der nächsten Jahre festlegt.

Daher möchte ich unterstreichen, dass wir mit dem Zusammenfall des spanischen Vorsitzes im ICA und des spanischen Ratsvorsitzes der Europäischen Union den festen Willen bekunden, den neuen Aufgaben, denen die Archive auf dem Gebiet der digitalen Informationen gegenüberstehen, unsere Unterstützung zukommen zu lassen.

Das Programm für die nächsten zwei Tage ist sehr straff und anspruchsvoll. Es umfasst die beiden Seiten eines elektronischen Dokuments, also zum einen das elektronische Dokument selbst als Ursprung und zum anderen das elektronische Dokument als alternatives Format zum Original. Und wie so oft in den letzten Jahren sind diese Zusammenkünfte des Archivwesens nicht auf die regionale Ebene beschränkt, in diesem Fall also auf die Europäische Union, sondern stehen auch Vertretern anderer Kontinente offen. Zudem werden sie nicht nur von Archivaren besucht, was in der heutigen Zeit nicht sinnvoll wäre, sondern auch von Berufsvertretern der neuen Technologien.

Aus dem Umstand, dass Barcelona als Tagungsort für das DLM-Forum dient, ergab sich die Einrichtung einer spanischen Arbeitsgruppe und die Durchführung eines Kongresses für Spanien und Lateinamerika, der gestern im Vorfeld dieses Treffens stattfand. Die spanische Arbeitsgruppe wird zweifellos einen Beitrag zur Verwaltung elektronischer Dokumente und Archive in der öffentlichen Verwaltung leisten, wurde jedoch auch zur Unterstützung der Arbeit des ICA-Ausschusses für elektronische Dokumente eingerichtet, dem Archivare aus Spanien und Portugal angehören. Dieser Ausschuss erarbeitet zur Zeit einen „Leitfaden für elektronische Dokumente aus archivischer Sicht", eine Aktualisierung der 1997 veröffentlichten Fassung des Leitfadens, die auf dem XV. Internationalen Archivkongress 2004 in Wien vorgelegt werden soll.

Zum Vorfeldkongress und zum DLM-Forum wurde auch der Präsident des Lateinamerikanischen Archivverbands (ALA) eingeladen. Als regionale Zweigorganisation des Internationalen Archivrates widmete dieser Verband auf seinem Jahresseminar vor ein paar Wochen in Santiago de Chile dem Thema „Elektronische Verwaltung in Ibero-Amerika" große Aufmerksamkeit.

Im Jahr 2002 wurden bereits einige weitere Initiativen in Sachen Informationsgesellschaft durchgeführt. Ende dieses Monats tritt der Exekutivausschuss des ICA in Peking zusammen und wird bei dieser Gelegenheit den regionalen Zweigorganisationen Asiens und des

Pazifikraums das Unesco-Programm „Informationen für alle“ vorstellen. Der Zwischenstaatliche Ausschuss dieses Programms, dem auch Spanien angehört, traf sich im April und arbeitet zur Zeit an einer Charta zur Bewahrung des digitalen Erbes mit klar umrissenen Zielsetzungen: Erhaltung von Informationen und des universellen Zugangs zu ihnen, jedoch auch unter Beteiligung aller in der sich herausbildenden Informationsgesellschaft.

Das Programm richtet sich insbesondere auf digitale Informationen, die zur Verbreitung über das Internet erstellt werden, dient aber auch als Ausgangspunkt für fachgebietsübergreifende Arbeitsgruppen, die sich mit den verschiedenen Aspekten des digitalen Schriftguterbes befassen und bereits im Herzen der Europäischen Union tätig sind. Ich bin sicher, dass der Exekutivausschuss des ICA in Peking eine eingehende Analyse dieses Unesco-Programms vorlegen kann, das vom Standpunkt des Archivwesens äußerst positiv zu bewerten ist.

Vor diesem Hintergrund überrascht es nicht, dass das vom spanischen EU-Ratsvorsitz Kultur gewählte Thema „Langzeitsicherung des digitalen Gedächtnisses“ lautet, und Ende dieses Monats wird der Rat zudem einen Beschluss zur Sicherung des digitalen Gedächtnisses fassen.

Der Beschlussentwurf wurde den Direktoren der Staatsarchive der EU und der Beitrittskandidaten auf einem Treffen in Valladolid am 11. und 12. März vorgelegt. Viele dieser Direktoren sind auch heute anwesend und gehören alle der regionalen ICA-Zweigorganisation EURBICA an. Dass wir also diese Angelegenheiten hier in Europa erörtern können, ist von den verschiedensten Seiten abgesichert. Das DLM-Forum ist ein weiterer Schritt auf dem Weg zu Lösungen für die gesicherte Konservierung des digitalisierten Schriftguterbes.

Mit Hilfe nationaler und internationaler Konzepte werden Pläne und Programme erstellt, die neue Technologien den Bürgern näher bringen und den Zugang zu Informationen verbessern sollen. Es besteht kein Zweifel, dass die Archive die ihnen in der elektronischen Verwaltung zukommende Rolle einnehmen werden, denn ohne die Archivkomponente würde diese neue Gesellschaft ohne Zweifel ihr Gedächtnis verlieren.

Communication du Conseil international des archives

Elisa de Santos Canalejo

Monsieur le Président de la généralité de Catalogne, Monsieur le Conseiller aux universités, à la recherche et à la société de l’information,

Mesdames et Messieurs, mes chers collègues,

Tout d’abord, permettez-moi de remercier les institutions qui ont organisé cet événement pour leur aimable invitation à assister à cette cérémonie d’ouverture du Forum DLM!

Je ne peux dissimuler ma très grande satisfaction de voir que les archives suscitent un intérêt croissant parmi les professionnels de secteurs très différents, ce qui témoigne très clairement de leur rôle dans la société de l’information. Nul ne s’étonnera que le XIV[e] Congrès international des archives, qui s’est tenu à Séville en 2000, ait eu pour titre: «Les archives du nouveau millénaire dans la société de l’information». Le Conseil international des archives (CIA), dont j’ai l’honneur d’assurer la présidence jusqu’en 2004, souhaitait que le congrès de Séville esquisse les actions à mener durant ces années.

J'aimerais donc souligner que, compte tenu du fait que la présidence espagnole du CIA coïncide avec la présidence espagnole de l'Union européenne, nous avons décidé de donner notre appui aux nouveaux défis que doivent relever les archives dans le domaine de l'information numérique.

Le programme de ces deux journées, très chargé et ambitieux, couvre les deux aspects du document électronique: 1) le document électronique en soi, en tant qu'original; 2) le document électronique en tant qu'exemplaire se substituant à l'original. Ainsi qu'on l'a souvent vu ces dernières années, ces conférences sur les archives ne se cantonnent pas dans la sphère de la région, laquelle en l'espèce aurait été celle de l'Union européenne, mais accueillent au contraire des professionnels d'autres continents. De même, ces rencontres ne s'adressent pas seulement aux archivistes, ce qui n'aurait aucun sens à notre époque, mais sont également ouvertes aux professionnels des nouvelles technologies.

Le choix de Barcelone comme lieu d'organisation du Forum DLM a suscité la mise en place d'un groupe de travail espagnol et l'organisation d'un congrès préparatoire destiné à l'Espagne et à l'Amérique latine, qui a eu lieu hier en prélude à la rencontre d'aujourd'hui. Le groupe de travail espagnol contribuera sans aucun doute à la gestion de documents et d'archives électroniques dans les administrations publiques, mais il a été également constitué pour soutenir les travaux du comité «Documents électroniques» du CIA, qui compte dans ses rangs des archivistes espagnols et portugais. Ce comité prépare actuellement un guide sur les documents électroniques du point de vue des archives, guide qui constitue une version actualisée de celui publié en 1997 et qui sera présenté au XV^{e} Congrès international des archives, qui se tiendra à Vienne en 2004.

Le président de l'Association des archives d'Amérique latine (ALA) a été invité au congrès préparatoire et à ce Forum DLM. En sa qualité de bureau régional du CIA, cette association a consacré une partie importante de ses travaux à l'administration électronique en Amérique latine lors de son séminaire annuel, organisé il y a quelques semaines à Santiago-du-Chili.

Un grand nombre d'initiatives concernant la société de l'information ont été lancées en 2002: à la fin de ce mois, le comité exécutif du CIA se réunira à Pékin et se saisira de cette occasion pour présenter le programme de l'Unesco intitulé «L'information pour tous» aux bureaux régionaux d'Asie et du Pacifique. En ma qualité de présidente du CIA, j'ai assisté il y a un an à la présentation de ce programme au siège de l'Unesco. La commission intergouvernementale du programme, dont l'Espagne est membre, s'est réunie en avril, et elle travaille actuellement à une charte pour la conservation du patrimoine numérique qui poursuit des objectifs clairement définis: conservation de l'information et universalité de l'accès à cette information, avec en outre la participation de tous à la société de l'information émergente.

Ce programme cible en particulier les informations numériques créées pour être diffusées sur le web, mais offre également une plate-forme de départ à des groupes de travail pluridisciplinaires pour aborder les différents aspects du patrimoine numérique. Ces groupes de travail fonctionnent en fait déjà au cœur de l'Union européenne. Je suis sûre que le comité exécutif du CIA sera en mesure à Pékin d'analyser de près ce programme de l'Unesco, qui est tout à fait positif du point de vue des archives.

Dans ce contexte, nul ne s'étonnera que la conservation à long terme de la mémoire numérique soit le thème choisi par la présidence espagnole de l'Union européenne. En fait, le Conseil devrait adopter à la fin même de ce mois une résolution sur la conservation de la mémoire numérique.

Le projet de résolution a été soumis aux directions des archives nationales des pays de l'UE et des pays candidats, lors de la réunion des 11 et 12 mars dernier à Valladolid. La plupart de ces directeurs sont présents aujourd'hui avec nous et sont tous membres du bureau régional du CIA, Eurbica. Notre capacité à traiter de ces questions ici en Europe est ainsi garantie

à maints égards. Le Forum DLM constitue une nouvelle étape vers les mesures nécessaires pour garantir la conservation du patrimoine numérique.

En conclusion, les politiques nationales et internationales conçoivent des projets et des programmes permettant de rapprocher les nouvelles technologies des citoyens et de favoriser leur accès à l'information. Il ne fait aucun doute que les archives doivent assumer le rôle qui est le leur dans l'administration électronique car, sans le point de vue de l'archiviste, la mémoire de cette nouvelle société serait incontestablement perdue.

Discussion at the plenary opening session Frank Brady, Head of Unit Simplification and modernisation of archival procedures and historical archives at the European Commission; Hans Hofmann, Head of Sector of the Historical Archives of the European Commission; Jordi Pujol, President of the Catalan Government (from left to right)

Communication from the International Council of Archives

One of the parallel sessions of the DLM-Forum

Demonstrations at the exhibitions stands at the DLM-Forum 2002

DLM-audience at the plenary closing session

Plenary speakers

Towards greater coordination in archive cooperation in Europe
Opening of the DLM-Forum 2002

Martine De Boisdeffre

Martine de Boisdeffre

Conseiller d'État
Director of the Archives de France

Personal details
Martine de Boisdeffre, daughter of Colonel René Langlade and Mme Solange Langlade, née Laborde, born in Orthez, Pyrénées-Atlantiques, on 15 August 1957.
Married on 19 April 1986 to M. Christian Néraud le Mouton de Boisdeffre, sub-prefect of l'Haÿ-les-Roses since 1998.
Three children: Marie-Aurore, Bertrand and Vincent.

Education
— École normale supérieure de jeunes filles, Sèvres
— École normale d'administration (class of solidarity 1981–83)

Qualifications
— Master's degree in history
— Diploma from the Institut d'Études politiques in Paris

Career
1983 – Auditor at the Conseil d'État
– Rapporteur at the Special Pensions Appeal Commission
1984–98 Reader at the Institut d'Études politiques, Paris
1985–92 – Secretary-General of the National Ethics Committee
1986 – Counsel
1988 – General rapporteur on the work of the Conseil d'État on Life Sciences
1990 – Policy officer at the Interministerial Mission for Central and Eastern Europe
1990–93 Technical adviser in the private office of Elisabeth Guigou (Minister Delegate attached to the Minister of State, Minister for Foreign Affairs with responsibility for European Affairs)
since 1995
Secretary-General of the Conseil d'État
January 2001 – Appointed Director of the Archives de France

Mr Minister,
Mr Chairman,
Ladies and gentlemen,

Allow me to start by saying how delighted and honoured I am to open this meeting of the 2002 Forum on Machine-Readable Data.

When my friend Hans Hofmann asked me a few months ago to open this meeting, I was both surprised and delighted. I was surprised because I was — and still am — very new to the post of Director of Archives de France. I took up my duties on 22 January 2001, barely 18 months ago.

I was delighted because this meeting concerns electronic archives and the Forum on Machine-Readable Data which, as I have quickly come to realise, plays an essential role in an equally essential field.

I was also very pleased because we are currently in the Spanish Presidency, and as Chairwoman of the European branch of the International Council on Archives, I can only be delighted to be in the country of the Council's Chairwoman. Allow me therefore to salute Elisa de Santos, my colleague and friend.

Lastly, I am delighted because this is my first time in Barcelona, Gaudi's pearl. So thank you, Hans, for your kind invitation.

And thank you to all of you for your warm welcome.

I shall deal with four topics in succession:

- the Forum on Machine-Readable Data and electronic archives;
- archives in France;
- archives and Europe;
- archives in the city.

For each topic, I shall report on progress to date as well as examining future developments.

The DLM-Forum and electronic archives

It is not my intention to trace the origins and development of this forum in the presence of its founders and so many longstanding participants.

I should merely like to emphasise a few points.

First, that this is a forum; hence its ethos and tone are characterised by exchange and debate on the basis of equality, with an absence of constraints or pressure. Dialogue and exchange are enriched by their multidisciplinary character, for this forum brings together members of public bodies, archive departments, representatives of industry and researchers, notably, for obvious reasons, in the field of new information and communication technologies. This range of skills and pro-

fessions is crucially important. In fact, I wonder whether the legal profession is properly represented, for I am increasingly struck by the fact that the application of these technologies to archives by archivists, and indeed by all the other participants, raises a wide range of complex legal issues.

Second, the Forum on Machine-Readable Data did not emerge out of nowhere. It is the result, and I want to emphasise this, of the Council conclusions of 17 June 1994 concerning greater cooperation in the field of archives. It is already eight years ago that this text noted with interest among the actions contemplated the plan to organise a 'multidisciplinary forum to be held in the framework of the Community on the problems of the management, storage, conservation and retrieval of machine-readable data, inviting public administrations and national-archives services, as well as representatives of industry and of research, to take part in the forum'. The text was visionary on this point, for it predicted the vital importance that electronic documents would acquire. It has been put into practice given that all the work you have carried out over the last eight years, and in particular the activities implemented in the course of this year thanks to the hospitality of the Catalan authorities have focused on developing a partnership between the public and private sectors. This has been done by intensifying the debate and promoting best practice and solutions in electronic archiving, with a view to modernisation and greater transparency of public services. This debate is of particular interest at a time when the Spanish presidency is working for the adoption of a resolution on the conservation of all digital cultural data, including archives. It is also of topical interest for my own country and for archives in France, because over the last few years, we have relaunched and expanded our activities and discussions in the electronic-archives field.

Archives in France

Two sets of reasons motivated our decision to proceed. First, the increasing pace of technological change, coupled with the explosion in the use of electronic documents, and second, changes in the regulatory framework.

I should mention that it was now 24 years ago, in 1978, that Archives de France launched its first experiments in archiving electronic documents. These operations were (and still are) limited to a few major national public bodies involved in the production of statistical records and surveys, such as INSEE or INED. They were all implemented by a single national archiving centre, namely the Fontainebleau contemporary archive, which I should like to congratulate on the work completed to date and its continuing efforts in the field.

For about 20 years, this work took place against a background of general indifference because electronic documents were regarded as ephemeral. These 'volatile' archives were not deemed important; only hard copies, it was thought, would ensure the survival of those documents which were felt to be worth keeping in the long term.

However, since 1998, a great deal has changed.

Local/regional authorities, irrespective of their size, have access to electronic hardware or have digitalised, at great cost, data which they deem worth saving, especially in view of the fact that they can sometimes be held liable under the law and in the eyes of the local community.

A number of texts have also been produced. Recommendations have been issued since 1998, first by the European Commission, then by the French Government, to the effect that documents produced electronically should be stored electronically too. Under the Administrative Documents (Access) Act of 12 April 2000, citizens may ask for data to be provided in digital form. The important law of 13 March 2000 on changes to the law of evidence to reflect developments in information technology and on electronic signatures recognises digital data as having probative value, subject to certain conditions. That same law provides for the most important documents governing the lives and rights of citizens, such as legal deeds, including documents showing civil status or notarial minutes, to be stored in paperless form. Lastly, on 2 November, a circular from the Prime Minister was published on the management of central government documents which takes account of their electronic form. This circular extends nationwide standard ISO 15489 of October 2001 on records management.

Thus the need to archive digital data as efficiently as possible, which was previously limited to major national institutions, has filtered down to every mayor and administrative worker in the country. Even the general public are beginning to become aware of it.

Accordingly, since 1998, Archives de France has taken a number of measures to inform and train the managers of the 800 or so archive departments under its scientific and technical control. In March 2001 it organised several highly successful international conferences to raise archivists' awareness of their new duties and responsibilities. We are as actively involved as possible in national and international discussion and research programmes and in various training programmes in France and abroad. Since April 2000 we have published an e-bulletin on the subject and we have produced a practical handbook which will be available shortly.

We also intend to be deeply involved with the first e-document archiving operations carried out by local authorities and not just to share our tried and tested methodology with them but also to experiment with new methods where possible. In a nutshell, our aim is to provide practical help and reassurance.

It was in this spirit that we decided to follow closely, in conjunction with the Ministry of Justice, the archiving of the very rich collection of nouvelle chaîne pénale de Paris et de la région parisienne, known to users by the acronym NCP.

The leading role of Archives de France in archiving electronic documents in general and the French authorities' Internet and intranet sites in particular was decided at interministerial level in accordance with that same spirit.

That concludes my review of recent developments in electronic archiving in France.

From a more general perspective, I should also like to mention four other new developments.

First, the decision to open a new centre for the national archives with a view to resolving the virtual saturation of existing centres and to enabling us to carry out our three key duties of data collection, conservation and dissemination. What is the state of play at the moment?

A decision in principle has been taken. Five sites have been selected and are undergoing an in-depth assessment. The first studies have been commissioned.

Second, the setting-up of the Archives de France Interministerial Committee by decree dated 23 January 2002, published in the *Official Gazette* on 24 January. The committee is chaired by the Prime Minister and consists of ministers. Archives de France takes care of the secretariat. Article 1 of its statute states that its remit is to determine the general direction of State policy on data collection, conservation and dissemination and exploitation of archives. It also coordinates action by ministries in this field.

Third, important decisions have already been taken as part of preparations for the first committee meeting.

In the administrative field, the need was recognised to monitor application of the circular of 2 November 2001 which defined the respective roles of the authorities and archiving departments regarding the management of current and intermediate records.

As regards access to archives, a national observatory has been set up to record all requests for a derogation, i.e. for archives to be released before the statutory period has expired and for statistics on this subject to be made available to the public. It has also been decided to open up local-government records on the Second World War.

In the legislative field, we were instructed to launch a discussion on the comprehensive renewal of the 1979 law, including the issue of political archives.

Fourth, an archives board (Conseil supérieur des archives) with an extended membership and more significant role has been relaunched and specialist committees set up. Some of these subjects are so important and useful to us all that, in my capacity as Chairwoman of Eurbica, I suggested examining them within the context of the European, or should I say greater European, branch.

Archives and Europe

As regards the European branch, Eurbica, I should begin by stressing that our objectives are simple ones: greater reciprocal knowledge, more information and better access for all and developed exchanges. The main policy decisions taken to date are based on what we identified as the priority needs.

First, selection. In his introduction to the 1993 edition, Jean Favier referred to French archiving practice. 'Wanting to keep everything — given that everybody wants above all to keep everything of interest to him – is purely Utopian. But the size of the problem means that individuals cannot be allowed to decide what to keep and what to discard. Although archivists take no pleasure in the fact, what archiving has become is an inexact science of elimination.'

With this in mind, we felt that it would be useful to compile and compare selection standards, criteria and practices in all the Member States of the European branch. Not in order to identify a single solution or to standardise practice, which I am not convinced would be a good thing, but with a view to testing the validity of ideas, comparing them with those of others and, possibly, learning from them.

Second, training. Some countries, such as Germany or France, are fortunate enough to have training institutions of particular renown, such as Marburg University or the Ecole des Chartes. However, in those countries and elsewhere, there are also less well-known institutions. In addition, training courses have had to adapt to the new challenges facing archivists: mass data processing, management of new supports, catering for new users. Initial training and continuing training are increasingly essential in order to maintain archivists' professional standards and, in a way, to ensure their status in society. For that reason we felt that it was crucially important to establish, at the greater European level, a kind of training fair for archivists. Work on this has already begun.

Third, legal data. André Malraux wrote 'The 21st century will be religious or it will not be at all.' My view is that the legal dimension will also be crucially important. I am struck by the fact that archive management, especially as a result of new technology, raises increasingly complex issues. Archivists must acquire some legal knowledge or work closely with lawyers. The profession must take on a legal dimension, and indeed acquire some knowledge of comparative law, because increased contacts, lending and joint operations will create situations in which interaction between legal systems and interpretation, if not outright conflict, are inevitable.

That is why we are interested in collecting information on the laws of each country concerning not just archives but also access to administrative files, the protection of personal data and all other areas which henceforth are interlinked. The presence not just of national legislation but also European legislation produced by the Council of Europe and the European Union, means that a single legal framework for archives, if it ever existed, is no more.

The European Union

Culture, as we know, does not fall within the Community's remit. Nevertheless, the Maastricht Treaty refers to the Community contributing to 'the flowering of the cultures of the Member States, while respecting their national and regional diversity and at the same time bringing the common cultural heritage to the fore.' That is the wording of Article 151 (former Article 128) of the Treaty.

However, Article 151(2) states that action by the Community must be aimed at encouraging cooperation between Member States and, if necessary, supporting and supplementing their action,

inter alia to improve the knowledge and dissemination of the culture and history of the European peoples, and to conserve and safeguard cultural heritage of European significance.

There is therefore scope for more coordination and information.

Various initiatives have been launched to that end:

- INSAR, the Information Summary on Archives, which provides updates on archiving activities in the European Union;
- the setting-up of an official group comprising representatives of national archives and archivists from EU institutions;
- the draft Lund-Brussels resolution endorsed by the Valladolid declaration by national archivists;
- the new edition of the report on archives in the European Union;
- the draft resolution on the storage of digital data.

But perhaps student and archivist exchanges should be developed via Community action programmes. No doubt they already exist, but I feel they are inadequate given that a thorough mutual knowledge cannot be acquired through meetings alone; it entails complete immersion in a department and hands-on involvement. I'd like to look at a few examples of active cooperation along these lines.

Training in electronic archives carried out by French specialist archivists for Belgian archivists; the launch at Lausanne University in Switzerland, at the initiative of the Director of the Swiss Federal Archives, of an archive management course which various European archivists will be involved in teaching; the organisation this summer by the Italian authorities at Dobbiaco of a seminar on document storage attended by experts from all over the European Union, etc.

It should be noted that the 1994 text I have just referred to called for exchanges of this kind. It was ahead of its time in several respects, but perhaps it focused on just one aspect of archives, namely the cultural aspect. For my part, I believe that archives occupy a space at the heart of the community not just for that reason, but for others too. It is their very complexity but also, perhaps, their ambivalence, which makes it so difficult to develop and promote them.

Archives in the city

I am speaking here of the city in the Greek sense of the word, i.e. the relationship between archives and citizens, or citizenship.

Archives must be at the heart of the city, for they are our heritage. Our fellow citizens are becoming aware of this and beginning to understand that heritage is more than just bricks and mortar: there are paper cathedrals as impressive as cathedrals built in stone and charters as precious as the Crown jewels. The community, or city, is starting to rediscover these treasures because they represent a bridge to the past.

Archives are a history resource which shed light not just on major events but also on intimate or anonymous experiences, individuals' stories and collective memories; ultimately, all the rising tide of genealogists has done is to reflect the development of history up to the Annals School and the history of groups rather than individual heroes.

The role of archivists is simultaneously to guard this heritage and to communicate it to an increasingly numerous and varied audience. We have to rise to these new challenges as best we can.

Archivists are, can be and indeed must be involved in preserving and displaying our shared history and common heritage to all the people of Europe.

But there's more to archives than the cultural dimension alone.

They are also an administrative tool functioning simultaneously as a resource and a test of a sound administration; one which has a memory and knows how to use it; one which is democratic because it allows access to its archives while treading the fine line between the need for transparency and the necessity of protecting the legitimate interests of the State and the individual (right to privacy).

It needs to be understood by all that a well-managed administration requires well managed archives. Good archiving means collecting the right documents, storing them properly and ensuring the best possible access to them. That is our rule and our duty but we also need to persuade the authorities, bodies producing archives and decision-makers. Archiving doesn't mean purely working for the future and for other people. It also means working for ourselves, in the here and now.

If we can get this message across, everyone will benefit: public archives will get more resources and respect, and I am sure that there is scope for this in almost all European countries. Governments will gain in terms of efficiency, transparency and recognition from our fellow citizens. This is becoming acutely important for electronic archives, Internet sites, etc. for the risks of loss and dilution are immense, especially since access is neither direct nor immediate but is nevertheless well publicised, as the acronym DLM suggests. Ordinary citizens will benefit in terms of culture and information.

Our countries will benefit, both individually and collectively, from increasing coordination and cohesion in this field.

It is for us to get these ideas and messages across, individually within our own borders and together vis-à-vis the European institutions. Unity is strength, let us not forget. But above all it must enable us in this Europe of ours, ripped apart by conflict for so many centuries, or 'carefully ordered barbarism', as Malraux put it in *The temptation of the West*, to secure peace.

Archives have a part to play in this. As the source of our history and a reflection of it, they must help us to gain understanding and learn lessons from the past so that we do not revert to sterile conflict and 'civil war' as Victor Hugo put it in *The Burgraves*. The poet stated in that play that 'today, there is a European nationality as there was a Greek nationality at the time of Aeschylus, Sophocles and Euripides.' More than a century later, I want to believe that too.

Una mayor coordinación de la cooperación archivística a nivel europeo

Martine de Boisdeffre

1. Introducción

Importancia del DLM-Forum de Barcelona; cooperación interdisciplinaria (administración pública, archivos, industria de las tecnologías de la información y de las comunicaciones [TIC], centros de investigación); debate y promoción de las mejores prácticas y soluciones sobre archivos electrónicos para la modernización y transparencia de los servicios públicos; colaboración entre los sectores público y privado; contexto de la Presidencia española de la Unión Europea, Generalidad de Cataluña.

2. Acciones realizadas en Francia

— Creación del Comité Interministerial de Archivos de Francia, dependiente del Primer Ministro (Decreto n° 2002-95 de 23 de enero del 2002, *Journal Officiel de la République*

Française 1601, de 24 de enero de 2002) y presidido por su representante; participación de los principales ministerios, secretaría a cargo de la Dirección de Archivos de Francia; coordinación para facilitar la modernización de los archivos de la administración central francesa.

Véase también la circular de 2 de noviembre del 2001 sobre gestión de los archivos en los servicios y establecimientos públicos del Estado (*Journal Officiel de la République Française* 17359, de 4 de noviembre de 2001).

— Reactivación del Consejo Superior de Archivos (archiveros, historiadores) para promover una mayor coordinación de los archivos en Francia.

— Instalación inminente de un nuevo centro de archivos (edificio) en Francia, apoyo del Primer Ministro, organización y funcionalidades del nuevo centro; examen de las posibles perspectivas de desarrollo de una política nacional y europea para edificios modernos de archivos (normas, funcionamiento, exigencias técnicas, conservación, acceso, seguridad = elaboración «Duchein II»).

3. Mayor colaboración de los archiveros a nivel europeo

La Sra. De Boisdeffre es Presidenta de la sección europea del Consejo Internacional de Archivos: organizaciones, tareas, acciones prioritarias.

— Creación de un grupo oficial de archiveros nacionales y de las instituciones de la Unión Europea.

— Proyecto de resolución de Lund-Bruselas: acciones prioritarias («puntos de cristalización»).

— Seguimiento en el Comité de Asuntos Culturales del Consejo de la Unión Europea.

— Declaración de Valladolid de los archiveros nacionales.

— Nueva edición del Informe sobre los archivos en la Unión Europea, incluida una mayor coordinación para la gestión de los archivos electrónicos.

4. Conclusiones

— Ampliación de la Unión Europea.

— Evolución tecnológica.

— Profundización de la colaboración interdisciplinaria en materia de archivos electrónicos.

— Mayor coordinación de la cooperación archivística para informar mejor al ciudadano y salvaguardar y promover el patrimonio común de los archivos en Europa.

Für eine bessere Koordinierung der Zusammenarbeit im Archivwesen auf europäischer Ebene

Martine de Boisdeffre

1. Einleitung:

Bedeutung des DLM-Forums von Barcelona; interdisziplinäre Zusammenarbeit [öffentliche Verwaltung, Archivdienste, Industrie der Informations- und Kommunikationstechnologien (IKT), verschiedene Forschungsbereiche]; Diskussion und Förderung der besten Methoden und Lösungen auf dem Gebiet der elektronischen Archive zwecks Modernisierung und Transparenz der staatlichen Dienststellen; Partnerschaft zwischen öffentlichem und privatem Sektor; im Kontext der spanischen Präsidentschaft der Europäischen Union, Generalitat de Catalunya usw.

2. Aktionen in Frankreich, darunter:

— Gründung des Comité interministériel des Archives de France (Interministerieller Ausschuss der Archive Frankreichs) beim Premierminister (s. Dekret Nr. 2002-95 vom 23. Januar 2002, JO de la République française 1601, 24.1.2002) unter dem Vorsitz eines Vertreters des Premierministers, Teilnahme der wichtigsten Ministerien, Sekretariatsführung durch die Direction des Archives de France, Koordinierungsaufgaben, um die Modernisierung der Archive der französischen Zentralverwaltung zu erleichtern usw.

S. auch das Rundschreiben vom 2. November 2001 über die Führung der Archive in den öffentlichen Dienststellen und Einrichtungen des Staates (JO de la République française 17359, 4.11.2001)

— Wiederbelebung des Obersten Archivrates (Archivare, Historiker) zur Förderung einer besseren Koordinierung der Archive in Frankreich usw.

— Bevorstehende Einrichtung eines neuen Archivzentrums (= Gebäude) in Frankreich, Unterstützung durch den Premierminister, Organisation und Funktionen des neuen Archivzentrums; Prüfung der eventuellen Perspektiven für die Entwicklung einer nationalen und/oder europäischen Politik in Bezug auf moderne Archivgebäude (Normen, Funktionsweise, technische Anforderungen, Aufbewahrung, Zugang, Sicherheit = Erarbeitung von „Duchein II").

3. Verstärkte Zusammenarbeit der Archivare auf europäischer Ebene:

— Frau de Boisdeffre ist Vorsitzende des „Europäischen Zweiges" des Internationalen Archivrates: Organisationen, Aufgaben, vorrangige Aktionen;

— Schaffung einer offiziellen Gruppe von nationalen Archivaren und Archivaren der EU-Institutionen;

— Entschließungsentwurf von Lund-Brüssel: vorrangige Aktionen („Kristallisationspunkte");

— Kontrolle im Ausschuss für Kulturfragen beim Rat der Europäischen Union;

— Erklärung von Valladolid der nationalen Archivare;

— Neuausgabe des „Berichts über die Archive in der Europäischen Union", einschließlich einer besseren Koordinierung bei der Führung der elektronischen Archive.

4. Schlussfolgerungen

— Erweiterung der Europäischen Union;

— technologische Entwicklung;

— Vertiefung der Zusammenarbeit im Bereich elektronische Archive;

— für eine bessere Koordinierung der Zusammenarbeit im Archivwesen zur besseren Information des Bürgers und zur Erhaltung und Förderung des gemeinsamen Bestandes der Archive in Europa.

Pour une plus grande coordination de la coopération archivistique au niveau européen

Martine de Boisdeffre

1. Introduction:

Importance du Forum DLM de Barcelone; coopération interdisciplinaire [administration publique, services d'archives, industrie des technologies de l'information et de la communication (TIC), différents milieux de la recherche]; discussion et promotion des meilleures pratiques et solutions en matière d'archives électroniques visant la modernisation et la transparence des services publics; partenariat entre les secteurs public et privé; contexte de la présidence espagnole de l'Union européenne, généralité de Catalogne…

2. Actions menées en France, y compris:

- Création du comité interministériel des Archives de France auprès du Premier ministre (voir décret n° 2002-95 du 23 janvier 2002, *Journal officiel de la République française* n° 20 du 24.1.2002, p. 1601), présidé par un représentant du Premier ministre, participation des principaux ministères, secrétariat assuré par la direction des Archives de France, tâches de coordination pour faciliter la modernisation des archives de l'administration centrale française…

Voir aussi circulaire du 2 novembre 2001 relative à la gestion des archives dans les services et établissements publics de l'État (*Journal officiel de la République française* n° 256 du 4.11.2001, p. 17359).

- Relance du Conseil supérieur des archives (archivistes, historiens) pour promouvoir une plus grande coordination en matière d'archives en France…

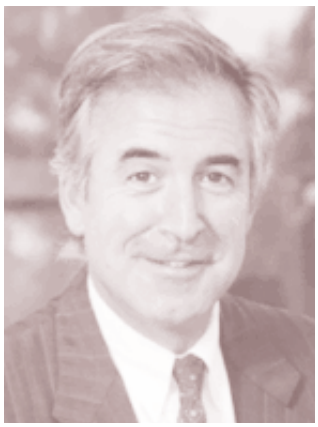
- Installation imminente d'un nouveau centre d'archives (= bâtiment) en France, soutien de la part du Premier ministre, organisation et fonctionnalités du nouveau centre d'archives; examen des perspectives éventuelles pour le développement d'une politique nationale et/ou européenne en matière de bâtiments modernes d'archives (normes, fonctionnement, exigences techniques, conservation, accès, sécurité = élaboration «Duchein II»).

3. Collaboration accrue des archivistes sur le plan européen:

- M[me] de Boisdeffre est présidente de la «branche européenne» du Conseil international des archives: organisations, tâches, actions prioritaires.
- Installation d'un groupe officiel des archivistes nationaux et des archivistes des institutions de l'UE.
- Projet de résolution de Lund-Bruxelles: actions prioritaires («points de cristallisation»).
- Suivi au sein du comité des affaires culturelles (CAC) du Conseil de l'Union européenne.
- Déclaration de Valladolid des archivistes nationaux.
- Nouvelle édition du *Rapport sur les archives dans l'Union européenne,* y compris une plus grande coordination en matière de gestion des archives électroniques.

4. Conclusions

- Élargissement de l'Union européenne.
- Évolution technologique.
- Approfondissement de la collaboration interdisciplinaire en matière d'archives électroniques.
- Vers une plus grande coordination de la coopération archivistique pour mieux informer le citoyen et pour sauvegarder et promouvoir le patrimoine commun des archives en Europe.



Piero Corsini

Piero Corsini, Vice-President, IBM EMEA Public Sector, is responsible for IBM work with governments, educational institutions and the healthcare and pharmaceutical industries in Europe, the Middle East and Africa.
Having gained a university degree in mechanical engineering he joined IBM in 1976 as a systems engineer. From 1983 to 1986 he managed IBM's Academic Independent Business Unit in London, returning to Italy in 1986 as the Director of the Central Government Branch Office in Rome. In 1989 he then moved to Milan to head the Large Industrial Enterprises Unit. In the last 10 years he has run a number of IBM businesses in Italy and internationally and was global manager of IBM's healthcare business in Europe before taking on his wider public sector role in 1999.
He is married with three children.

Access — Archive — Leverage
How 'e' do you want to be?

Piero Corsini

The next generation Internet

We know that the Internet is at the heart of transformations in society, and we have identified seven trends that we believe will characterise the next generation Internet. Each characteristic is vitally important to every government organisation.

(a) Fast
A really significant development on the horizon is a dramatic increase in bandwidth. It took us roughly 15 years to improve bandwidth 10 times. In the next five years we will make a 150 times' improvement. This will make video on the Internet a workable reality; a breakthrough which, for example, will make online training and video education on demand available for everyone.

(b) Always on
Businesses and citizens will have continuous and instant access to the Internet with no need to dial up or log on. There will be a continuous instant access with a variety of access types which will include more and more wireless connectivity.

(c) Everywhere
Sooner than you may think, almost everything you purchase that's worth more than USD 10 or USD 20 — whether it be a bicycle or a refrigerator — will contain a tiny chip that can communicate via a wireless link to the Internet. Imagine you will be able to track down the stolen bicycle using a global positioning system. Meanwhile, the refrigerator will automatically order bread, milk and other foods when they run out.

We call this pervasive computing, and see it approaching a kind of mathematical extreme in which Internet-connected chips will literally disappear into all the things around us.

Here are some interesting figures:

- Today 98 % of Internet access is through traditional PC-based browsers. In five years, that will drop to just 50 %;
- In two years, 35 % of all Internet transactions will be on mobile devices. More than 50 % of Internet access in Japan comes from non-PC devices.

(d) Natural
Today, we adapt to technology ... it doesn't adapt to us. Soon, technology will be much more intuitive and easier to use. In many cases, we won't even realise that we are using it. It will be natural.

(e) Intelligent
As you can see, technology is becoming intelligent. It's sensing me and adjusting to my world.

So, for example, if I am accessing the tax office's web site from a data-enabled cell phone, the infrastructure will detect that and automatically convert my account to a format that can be displayed on a very small screen. If my native tongue is Spanish, the infrastructure should detect that in my user profile and do translations on the fly.

(f) Easy
In the future, when we communicate or transact business using the Internet, the whole experience will flow easily from one step to the next. Software and applications will be able to talk to one another and work across time and distance.

(g) Trusted
And finally, to make the sharing of applications and information not only a possibility, but also a necessity, you need to have a trusted environment. Technology is finding solutions for the security issues ... ensuring that people's identities can be verified over the Internet (through digital IDs), and that information travelling across networks is secured beyond a reasonable doubt.

The complex scenario of e-government

e-government is a journey. It is generally accepted that e-government transformation often progresses in four waves.

In the first wave, governments are extending current services by publishing information online. This is driven at the department level with each individual group doing their own thing. For example, the tax office may publish basic tax information and a 'How to' guide for citizens and businesses on its web site.

In the second wave, basic transactions go online. For example, you may be able to do your tax return online. In fact, tax offices have been leading the way with e-government. In Italy, 80 % of income is taxed over the Internet — some 23 million transactions. Other countries have the facilities but not the take-up. In France, of the millions of taxpayers, only 16 000 went online to do their tax in 2001. In Spain, the figure was 120 000 out of 14 million.

During the second wave, we also get better organisation of information and portals are designed to create a single point of access to all the relevant information and services from a specific agency or department. To support this, back end systems are improved and security, reliability and scalability become a major cause of concern.

The third wave is where things get complicated but much more interesting. This is where real transformation take place, most notably, the one-stop-shop. Where you re-engineer the way things are done to suit the customer as opposed to the department.

Returning to the tax office example. In the third wave you don't have to fill in and send off a form. Instead, the tax office presents you with your tax declaration, filled out with input on your pension, your investments, your property, family and children, etc. You just update, sign and return it (ideally online!) At this stage the system is really helping you in a friendly way.

Once this transformation has taken place, you can begin to extend your service offerings even further. This is the fourth wave. This is where you can start to dream. This is where governments act in a proactive way.

For instance, 10 years before my retirement, a government which knows my situation, could offer me additional financial products to ensure a better pension. Or, going back to my tax situation, which my government knows well, it could help me to pay less taxes.

But as we move up the value chain, the infrastructure will have to be far more robust to deal with this level of complexity. The trick is to preserve the investment in legacy systems while creating a scalable platform for future growth and development.

Information and record management

These days, we are in transition regarding managing information. A few years ago, most information was on paper. In a few years only a small fraction of governmental information will be on paper.

We are used to managing paper-based information — in physical archives, keeping journals, managing document-retention cycles etc.

With e-mail becoming a preferred communication method for many, we have seen a severe loss of control — the responsibility information management has often been left to the person receiv-

ing or sending an e-mail, and access control, journalising, audit trails, retention policies and back-up have been either abandoned or done as a multitude of point solutions. The same issues are important for governmental web sites — how can you document that this vital piece of information was published and generally available on a certain date two years ago?

These issues have been driving forces behind the records-management efforts of the governmental sector all over the world — MoReq from the European Union, the US Department of Defence, the UK Public Records Office and DOMEA in the German-speaking countries being good examples.

We need to manage your information through its entire life cycle. We need to make sure that only authorised persons have access to the information — within the government as well as prohibiting citizens from accessing other citizens' records.

We need intelligent ways of retrieving information from archives in order to leverage the knowledge within the records.

When making a decision we need to be able to retrieve similar decisions in the past to understand issues and established practices.

It is important to retain information but it is just as important to remove the information at the right time if that is your policy — for privacy reasons or to prevent legal action.

In a heterogeneous world of paper-based information and electronic information we need strong and robust links between the two so we can be sure we have all relevant information about a particular topic in front of us.

Let us talk a little more in detail about the various steps in a records management process.

First you need to capture the information. Since information is presented in many forms — from data records through documents and e-mails to video — the technologies for capture are several: electronic capture of already available information through applications, scanning inbound paper mail, capturing inbound faxes electronically, loading video sequences etc.

Indexing is key. This is where you attach information like author, title, date, recipient, topics, and maybe even a summary. Much of this can be done automatically today through intelligent text recognition, text mining and categorisation. There are even tools to cut videos into single takes and provide small pictures of each 'scene' in the video.

Important storage topics are: balancing cost through using different media like hard disk, optical disk and tape, about managing your information by migrating information between media, and by ensuring that you have large information objects close to the users for bandwidth reasons — and of course it is about backup and access control.

Information needs to be distributed — by pushing e.g. reports to certain people or providing them through a portal.

Employees and citizens need to be able to search and retrieve information. The search and retrieve functions should not depend on the nature of the information — a person should be able to search and retrieve all information related to a topic or a case — no matter whether it is e-mails, documents or video.

You need to manage the life cycle of the record — from the very birth of the information — through active life, archived life and finally policy-based deletion.

Throughout the cycle you need to control access. This is not only about managing who can access a particular piece of information at a server or a web site — it is about managing access to the electronic record throughout its life cycle — also when the information is copied and distributed.

Who is allowed to copy and distribute, who is allowed to access the copies and for how long. Let me give you just one example. A citizen is about to receive treatments at a hospital. The citizen's doctor sends X-rays to the hospital — public or private. For personal protection reasons the copy of these X-rays can only be seen by authorised personnel at the hospital — and only for two or four weeks — as long as the citizen is hospitalised. This requires digital asset-management technology at a very detailed level. We have developed technology for this — in the first place to manage distribution of music ensuring intellectual property rights. We now have this digital asset-management technology available in a generic form being able to handle any kind of information, including government records.

And — most important of all: you need to integrate with your legacy applications — in order to capture information from them, and in order to make records available to the users — internal and citizens — of the legacy applications!

Let me just say a few words about standards and guidelines. The previously mentioned examples are only examples — there are many more.

These standards or guidelines do not have the same scope — some are standards to which you can be certified as a software or solution provider, others are guidelines to which you can comply. At IBM we are much in favour of open standards. They can accelerate development, they ease integration across organisations and they are a help to us as an IT provider to document the capabilities of our offerings. Regarding the standards and guidelines you see on the screen we are either certified, or we comply, or we are using the guidelines in our product design and development efforts.

Perspectives

With the ongoing decrease in the cost of storage, it will become attractive to record and use even more information than we are doing today.

A larger part of the information volumes in the future will be digital media like high-resolution images, moving pictures and sound. Speech technology combined with text mining and portals will make it possible to search for information delivered in meetings in an intelligent way.

These trends will change our world profoundly. I have seen projections of 1 000-fold increases in Internet traffic in the next few years. In IBM we talk about 10 times' more connected people, 100 times' more network speed, 1 000 times' more devices and a million times' more data.

Think of what we have achieved managing just a tiny fraction of the world's information existing today: ERP applications, government archives, databases, e-commerce solutions, customer relationship management, e-mail systems, and the entire Internet. Then think of the benefits citizens and governments will have from leveraging information many times richer. e-government is indeed an exciting journey.

Acceso, archivo, apalancamiento: ¿cuánto «e» quieren ser?

Piero Corsini

Estamos en un período de transición por lo que se refiere a la gestión de la información. Hace unos años, la información se presentaba en su mayoría en papel. De aquí a pocos años, sólo una pequeña parte de la información administrativa estará en papel.

Tenemos la costumbre de gestionar información en papel, ya se trate de archivos, diarios, ciclos de conservación documental, etc. Con la llegada del correo electrónico como método de comunicación favorecido por un gran número de personas, asistimos a una grave pérdida de control: la responsabilidad de la gestión de la información se deja a menudo a la persona que envía o que recibe el correo electrónico, y los controles de acceso, la gestión de diarios, los análisis retrospectivos, las medidas de conservación y las protecciones se abandonan o se llevan a cabo como soluciones específicas. Lo mismo sucede con los sitios web de la administración: ¿cómo puede documentarse, por ejemplo, que una información vital se publicó y divulgó ampliamente en una determinada fecha hace dos años?

Estas cuestiones constituyen la dinámica de los esfuerzos realizados por las administraciones del mundo entero en materia de gestión documental. Inspirándose en distintos ejemplos, Piero Corsini mostrará la transformación de la administración electrónica y la transición que se operará en la gestión de la información.

Zugang – Archiv – Einfluss: wie „e“ wollen Sie sein?

Piero Corsini

Derzeit befinden wir uns, was die Verwaltung von Informationen betrifft, in einem Übergangsstadium. Noch vor einigen Jahren standen die meisten Informationen auf Papier. In ein paar Jahren wird nur noch ein Bruchteil der Informationen staatlicher Stellen Papierformat haben.

Wir sind den Umgang mit papiergebundenen Informationen gewöhnt – die Aufbewahrung in räumlichen Archiven, das Führen von Protokollen, die Einhaltung von Aufbewahrungsfristen usw. Seit nunmehr viele Leute die E-Mail als Kommunikationsmethode bevorzugen, haben wir in gravierendem Maße die Kontrolle verloren: Die Verantwortung für die Informationsverwaltung wird oftmals demjenigen überlassen, der eine E-Mail empfängt oder versendet, und Zugangskontrolle, Protokollführung, Audit Trails, Aufbewahrungsfristen sowie Anlage von Sicherungskopien spielen entweder gar keine Rolle mehr oder kommen in Form einer Vielzahl von Punktlösungen zur Anwendung. Die gleichen Fragen spielen auch in Bezug auf staatliche Websites eine Rolle: Wie lässt sich beispielsweise dokumentieren, dass eine wichtige Information an einem bestimmten Tag vor zwei Jahren veröffentlicht wurde und allgemein verfügbar war?

Diese Fragen stehen in aller Welt hinter den Bemühungen im Bereich Datenverwaltung im staatlichen Sektor. Anhand verschiedenster Beispiele wird Piero Corsini den Übergang zu E-Government und den Stand beim Informationsmanagement beleuchten.

Accès, archive, puissance de levier: Comment «e» veut être?

Piero Corsini

Nous sommes dans une période de transition en ce qui concerne la gestion des informations. Il y a quelques années encore, les informations se présentaient pour la plupart sur papier. Dans quelques années, seules une partie infime des informations administratives seront produites sur papier.

Nous avons l'habitude de gérer des informations sur papier, qu'il s'agisse de gérer des archives, des journaux, des cycles de conservation documentaire, etc. Avec l'avènement de l'e-mail comme méthode de communication privilégiée par un grand nombre de gens, nous assistons à une sérieuse perte de contrôle — la responsabilité de la gestion de l'information étant alors souvent laissée à la personne émettant ou recevant l'e-mail, et les contrôles d'accès, la tenue de journaux, les analyses rétrospectives, les mesures de conservation et les sauvegardes étant soit abandonnés, soit effectués comme autant de solutions ponctuelles. Il en va de même concernant les sites web de l'administration: comment rendre compte du fait, par exemple, qu'un élément d'information vital a été publié et largement diffusé à une certaine date, il y a deux ans?

Ces questions constituent la dynamique des efforts déployés par les administrations du monde entier en matière de gestion documentaire. S'inspirant de divers exemples, Piero Corsini montrera la transformation de l'administration électronique et la transition s'opérant dans la gestion des informations.

Ulrich Kampffmeyer

Dr Ulrich Kampffmeyer, born in 1952, is the founder and President of Project Consult GmbH, one of the leading independent management consultancies for business process organisation, knowledge management and workflow in Germany. He is founder and managing partner of Project Consult International Ltd., London.
He has been founder and Chairman of the Board of the German Association for Document Management, VOI Verband Organisations und Informationssysteme e.V., from 1991 to 1998. He served the International Document Management Vendor Association IMC as member of the Board of Directors from 1993 to 1998. Since the merger of IMC and AIIM in 1999 he actively supported AIIM International, the global association of users and suppliers of Enterprise Content Management Solutions. From 1999 to 2002 he co-chaired the European Board of Directors of AIIM International Europe. Since 2002 he is a member of the International Board of AIIM International and chair of several AIIM committees.
Dr Kampffmeyer is a member of the DLM-Monitoring Committee, chaired the DLM-Scientific Committee for the DLM-Forum Conference Barcelona 2002, and is a member of the steering committee for the establishment of the DLM-network of excellence. He is editor of the six DLM/AIIM Industry White Papers on electronic document, records and content management for the public sector in Europe.

Document life-cycle management for the European public sector Industry White Papers

Ulrich Kampffmeyer

The second DLM-Forum on Electronic Records in Brussels 1999 summarised in its conclusions (INSAR, Proceedings of the DLM-Forum on Electronic Records, Brussels 18–19 October 1999, pp. 333 and 339; European Communities, 2000) the urgent need for practical solutions facing the problems of public archives in Europe. Although there have been tremendous improvements on both sides, available technologies and organisational implementations in numerous archives, the growing volume of digital information created an even greater demand than stated at the DLM-Forum 1999 (Ulrich Kampffmeyer, 'Electronic documents management market in Europe: Technologies and solutions', INSAR, Proceedings of the DLM-Forum on Electronic Records, Brussels 18–19 October 1999, pp. 50–65, European Communities, 2000). The third DLM-Forum on Electronic Records in Barcelona 2002 was to address these challenges.

Drowned by the digital flood

The tasks of archivists, records managers, information managers and other related staff in public administrations today includes a growing variety of management and organisational issues to qualify, organise, store, manage, protect, provide access and deliver information in electronic archives. The 'digital flood' adds to the original tasks of public archives and increases the pressure to develop suitable solutions.

The 'traditional' problem

The traditional problems were focused on the management of paper and other physical objects in archives and how to transfer them in digital systems.

- Kilometres of paper documents

The huge amount of paper documents alone, in some archives reaching kilometres of shelves, millions of documents and billions of pages, creates a barrier for digitisation.

- Fragile, fading consistence

Especially in historic archives the consistence of paper, not only from historical periods but as well the late modern paper, demands for urgent initiatives for preserving the records physically.

- Capture

Both the huge amounts of documents and their fading consistence create problems for digital capture with cameras or scanners. Providing documents electronically can be part of a strategy to enable access without endangering the fragile original documents.

- Ordering systems and indexing

Indexing is the access key to documents. The depth and structure defines if a document can be easily retrieved or not. Different needs from archive specialists and future non-specialist users create difficulties in organising electronic archives.

- Quality

Capturing paper documents electronically, especially using high resolution and colour, create huge volumes of storage. Selecting appropriate formats for long-term archival and preservation is a challenge in the rapidly changing information technology market.

- Access

Providing access to specialists on-site in a local archive system environment is the easy part. Allowing access to other archivists at other locations, to third parties and to the citizen is both a challenge for user rights administration and high performance delivery of requested information.

- Funds

Politicians proclaimed open access to public archives. The necessary funds for installing and running systems, capturing and indexing the information, and managing the content are mostly not sufficient. Claim and reality may differ extremely.

- Awareness about the value of information

One of the largest tasks is to communicate the value of the content stored in archives. If the stored information is not used and the value is not recognised, the work of transferring all archives into electronically accessible records management solutions is not worth the effort.

The 'new' problem

Besides the 'traditional' problem of converting paper into electronic records, a new challenge has arisen by the increasing amount of digital information. Most documents today are created electronically with text processors, spreadsheets, business applications, graphic and construction design programmes, digital cameras and video.

- Exponentially growing volumes of digital information

The increase of digital information can no longer be measured. Industry analysts assume that per year today more information is created than in all the years since Gutenberg's invention. Software makes it easy to edit, copy, alter, store, distribute and print information. So, not only the digital flood flows higher but as well the output of paper swells.

- Digital 'only' information with no physical representation

There is an increasing amount of electronic information generated which is no longer designed to have a physical representation in paper. XML documents with separated content, structure and layout elements, which are generated dynamically when viewing them; digital video, multidimensional construction and simulation objects, electronically signed documents where the electronic signature can be only approved in a software system. The list of these electronic objects grows and creates problems for long-term availability in changing software and hardware infrastructures.

- The 'information gap'

We already face today a widening information gap. Once existing information is no longer available. From simple web sites disappearing from the Internet to valuable information of space expeditions on no longer readable tapes, the spectrum and the value of lost information cannot even be estimated. We have to start immediately with the creation of electronic archives for the preservation of the earlier periods of the information age.

- The 'information divide'

On the other hand we recognise a growing divide in those who can easily access any needed information and those who strive for pure survival without any connection to the virtual information world. We face today simultaneously an 'information overflow' and an 'information divide'. Both create risks for the information society.

- Information redundancy

One of the big problems of the digital flood is that most of the information are copies, slightly altered documents, and reformatted content. The new tasks for information and records managers include to discover what was the original document, what was its context, and to select informa-

Dr Kampffmeyer is a well-known keynote speaker, presenter and panellist on the subject of records management, document management, workflow, groupware, business re-engineering and organisational change management. His speaking engagements include major national and international congresses and conferences. He published numerous books and articles, i.e. the German codes of best practice for electronic archival. He is a member of several standardisation bodies in the field of document and records management.

tion which may be valuable for future generations. Storing everything without selection and detention cannot be the strategy.

- Commercialisation

In the early days of information technology digital information was private and only accessible in closed communities like companies or administrations. The age of free information on the web is coming to an end. Valuable information has to be paid. Digital rights management, privacy, media asset-management and other e-business software requirements add to the technological challenge for public archives. How to protect open public information from commercial reuse?

The new quality of electronic documents

As already discussed above the fast-developing software and telecommunications industry not only provides solutions but also new problems.

- Any format from data to digital video or complex virtual containers

An electronic document in the 1980s was an easy to manage object: an ASCII-data file, a scanned black/white image as TIFF. Today a document can be anything in a software system: a video stream, a container combining different files, a file containing references on other files in different systems, complex proprietary digital objects, etc. There is no unique format for every type of information available which could be used for long-term digital preservation, thinking in decades and centuries.

- Different renditions

Information today is provided in different renditions, the same content in different file and object formats. Decisions have to be made, as to what is the original, and which rendition should be stored. Due to the fact that electronic originals are often created with software with a short-term life cycle, renditions of stable representation formats may be more suitable for long-term preservation.

- Separation into content, metadata and structure information

Especially in enterprise content, media asset and web content management solutions with the use of HTML and XML we recognise a separation of the content information from context, structure and layout. The same information is used for different purposes and representations. This leads to the challenge, which form of representation to store and archive if the original software environment is not available.

- Digital signature and time stamps

Today's electronic signature solutions provide both a solution and a problem. The solution is that the author of a document can be authenticated and that the unchanged content of the document is proven. On the other hand personal electronic signatures 'live' only for a certain period and can only be verified in a given software environment. Time stamps, nowadays available as well as certified electronic signature, can be used in addition to personal electronic signatures. By this combination of the authenticity of a document and its originator, together with the certified time when it was generated, a legally valid electronic record can be created.

- Digital rights management

Today, digital rights management regulations and solutions add to the complexity of archiving electronic documents. They are on one hand a solution to protect copyrights and authorship rights, to trade assets and to prove original content. On the other hand, the management of digital rights of different origin and different technical implementation creates challenges for providing open access to electronic archives.

A challenge: open access

One of the goals of all projects for the implementation of electronic archive, document, asset, and records management systems in the public sector, is to allow open access to the citizen.

- Political mission

The political directive is clear — open public archives to provide transparency. Initiatives on the European level like *e*Europe2002 or on the national level like Bund-Online 2005 are to provide digital services to the citizen. Most of these initiatives focus on e-government. To allow open access to archives is often only a sub-topic.

- Democratic user access

The mission, to allow every citizen access to public archives via electronic means, ignores one important question: is there a demand, a need or an interest of 'the' citizen to access information in public archives? And if yes, to what kind of information?

- Professional users and private users

In the past, archives were administered and used by records and information specialists only. Written demands, personal visits were the way to access information by third parties. In the future, archives have to be able to serve different needs. Those of the specialists and historians as well those of journalists, companies and citizens.

- Rights protection

These new user groups lead to the necessity to install systems which allow open access but on the other hand protect information if it is classified, personal rights are involved or time of publication restrictions apply. In the past, the archivist or records manager could handle this individually, in the future intelligent software has to have the same ability.

- Easy use

Most of the records and document management software was designed for use by professionals in private or public administrations. User interfaces, navigation and retrieval functions require deep knowledge about the functionality and organisation of the system. To allow access for the citizen, easy-to-use interfaces have to be developed, not only for web access but in the future possibly as well for home television systems, mobile devices, and other technologies. And it makes no sense if every public archive develops its own propriety interface!

- Availability

The success of any open public archive will depend on the availability of information. Nice interfaces and comfortable navigation will lead to nothing, if the content is not prepared in a way that untrained users will find what they are looking for. And to launch an electronic archive with only a few documents already available is a 'good way' to make every potential user never use it again.

- Cultural dimensions

When discussing open access to public archives on a European level, we have to take into consideration the different information cultures in the Member States. A general approach, which information should be provided by which technologies, seems not to be successful. The use of information technologies, especially in regard to the new applicant States, is very different in Europe.

- Language barriers

Europe is a community of many languages. Language is a barrier in regard to information access. In regard to cross-language access to information, thesauri and automated translation technologies still have to be developed further. The goal to allow access to information independently from the language in which it was generated is still a vision and a major challenge. The European Commission and the European Community have to face this challenge and have to provide suitable solutions.

- Time and budgets

Given all these challenges and unsolved problems, all claims and programmes fall short today, both in regard to published timeframes as well as in regard to provided funds. Funding lots of different initiatives and single projects do not lead forward.

A Babel of acronyms

The challenges cannot be met by archivists, records managers, administrations and public institutions on their own. The ITC information and telecommunication industry has to supply suitable solutions. In the past, the public sector was not regarded as a strong key market: complicated tenders, low rates and budgets, bureaucracy, few chances for organisational change. This has changed significantly in the last three years. Today, the public sector is regarded as one of the drivers for the development of the DRT document-related technologies market. But still there is a lack of suitable, repeatable and affordable solutions to address the needs of long-term archival.

Today it is no easy job to find a suitable solution for the problems of the DLM-community. The ITC industry adds to this on its own with creating new acronyms and marketing slogans for slightly different solutions every year. Shall we look for:

- KM knowledge management?
- RM records management?
- DM document management?
- DLM document life-cycle management?
- DRT document-related technologies?
- CLM content life-cycle management?
- EA electronic archival?
- MAM media asset management?
- ECM enterprise content management?
- EDM electronic document management?
- EIP enterprise information portals?
- xyz next year?

All of these slightly different categories contain valuable components for the solutions needed by public-sector archives. Considering the gigantic tasks of transforming existing archives into electronic repositories and managing the digital flood we need suitable solutions now. We need no individual solutions but standardised systems which fulfil the specific needs of public archives at least by a 80:20 ratio. These solutions must be accessible via web interfaces, support multilingual cross-over retrieval, use automated classification to support indexing of large volumes, and provide formats and storage systems for constant migration and long-term preservation. Vendors always claimed, the needs of archivists and records managers are too specific that there is no chance to create standardised systems, the market for these solutions seemed too small. Today these functionalities are also required by any large private business and international organisation. This trend will increase the market offer of suitable solutions for public sector electronic archives.

The DLM acronym

The DLM community identifies by an own acronym. When created in 1996, 'DLM' was used for 'Données lisibles par machine'. Today, the ICT industry uses the DLM acronym for 'document life-cycle management'. To adapt to the wider tasks, scope and challenges, the DLM community changed the meaning of the acronym respectively. Machine-readable data are only a small part of the solutions necessary to provide the memory of the information society. DLM now stands for managing the life cycle of information from generation and use to long-term availability and preservation.

Best practice

In 1999, as a result of the second DLM-Forum the DLM Monitoring Committee issued a 'DLM-message to the ICT industry' and a 'consultative document' (INSAR, Proceedings of the DLM-Forum on Electronic Records, Brussels 1999, pp. 345 and 349, European Communities, 2000) to address the needs of the DLM community. In 2000 the ICT industry, represented by a group of independent industry specialists from Germany, the Netherlands, Sweden and the United Kingdom, presented the 'Answer to the DLM-message to the ICT industry' (INSAR, Volume 8, pp. 1 and 3, Brussels, 2000). The answer to the DLM-message included several proposals on educational activities supported by the ICT industry.

Based on this initiative, AIIM International, the global association for ECM (enterprise content management), launched a series of six industry White Papers to address the special needs of public-sector archives and to present practical solutions (industry White Papers on Records, Document and Enterprise Content Management, Series, AIIM International Europe, Datchet, United Kingdom, 2002; ISBN 3-936534-00-4).

The industry White Paper series was endorsed by Erkki Liikanen, Member of the Commission for Enterprise and Information Society. Liikanen states:

The importance of providing public access and long-term preservation of electronic information is seen as a crucial requirement to preserve the 'memory of the information society' *as well as improving business processes for more effective government. Solutions need to be developed that are, on the one hand, capable of adapting to rapid technological advances, while on the other hand guaranteeing both short and long-term accessibility and the intelligent retrieval of the knowledge stored in document management and archival systems. Furthermore, training and educational programmes on understanding the technologies and standards used, as well as the identification of best practice examples, need to be addressed.*

The AIIM/DLM White Papers are to demonstrate the ability of the ICT industry to deliver solutions to the major known problems and address the following topics:

1. Capture, indexing and auto-categorisation
2. Conversion and document formats
3. Content management
4. Access and protection
5. Availability and preservation
6. Education, training and operation

The following chapters summarise the content of the six publications, which focus on 'common' and 'best practice', standardisation, and trends.

IWP1 Capture, indexing and auto-categorisation

This White Paper on Intelligent Methods for the Acquisition and Retrieval of Information Stored in Digital Archives was authored by the company SER AG, a leading supplier of knowledge management, automatic classification and archival technologies from Germany. SER delivers as well workflow solutions supporting the DOMEA standard of the German public sector.

The White Paper addresses the ever-increasing overload of information. An individual can read approximately 100 pages per day, but at the same time 15 million new pages are added to the Internet daily. Our limited human capabilities can no longer filter out the information that is relevant to us. We therefore need the support of a machine which facilitates the exchange of knowledge by storing information and enabling personal, associative access to it through the lowest common denominator in human communication: The common human index is natural written and spoken language. All other types of indexing are limited aids which humans must first learn

to use before they can employ them. The standard has already been set and recognised as natural language, but systems which have adapted this natural standard are still missing.

The seven chapters of the White Paper deal with the importance of safe indexing, methods for indexing and auto-categorisation, the role of databases, standards for indexing, best-practice applications and an outlook on citizen portals and natural language-based portals. Special focus topics are high-speed data entry, OCR/ICR recognition technologies, and knowledge-based indexing and access. The two described applications from the Statistical Office of the Free State of Saxony and the computer magazine *CHIP*, give an overview about the efficient use of automated indexing and categorisation of unstructured information.

IWP2 Conversion and document formats

This White Paper on Backfile Conversion and Format Issues for Information Stored in Digital Archives was prepared by Hewlett Packard (HP), a worldwide leading supplier of computer hardware and storage solutions. For long-term archival, HP offers digital optical media and jukeboxes to automatically manage the media.

The White Paper addresses the issues which arise when considering the conversion of existing physical archives, that contain documents of different formats and types, into electronic format. These issues are broad in nature including the logistics of capture involving high volumes; the determination of appropriate strategies and tactics, for both delivering the conversion and maintaining normal business operations in the process; and the adoption of appropriate, reliable and sustainable document formats.

The first chapters describe the 'Bottleneck of conversion', types of information, information capture and standards for formats. Focus is laid on topics like long lasting archive formats, strategies for long-term information management, standards for storage formats and migration paths. The four best-practice applications are from the Department of Forestry, Sanctuary Housing Association, Staffordshire County Council, and Levy Gee. The last chapter gives an insight on conversion strategies and the value of archives.

IWP3 Content management

The third White Paper authored by the leading manufacturer of archival, document management and workflow management software, FileNET, concentrates on the theme 'managing the life cycle of information'. FileNET provides enterprise content-management solutions to the public sector worldwide.

The paper defines content management and the various technologies it embraces. It examines the differences between several content-management architectures and the different types of solutions being deployed today. The White Paper explains the different functionalities included in content-management solutions and outlines the relevant standardisation bodies, definitions and technologies. The mentioned best-practice applications feature examples from both the private and public sector. It forecasts the future of content management and identifies possible trends and developments. Special focus is laid on areas like scalability and availability of systems, Internet technologies, electronic archives as backbone infrastructure of modern information systems, and new standards for electronic documents. The seven chapters of the White Paper cover 'From archival to enterprise content management', architectures and integration issues, necessary functionality, content-management standards, and an outlook on e-government and 'e-merging technologies'. The volume contains several best practice applications including the solutions of the Federal Foreign Office, New Jersey Division of Revenue, and others.

IWP4 Access and protection

The fourth White Paper addresses the problems of managing open access and information protection. It was provided by IBM, the leading systems and solutions provider. IBM has been engaged in numerous e-government projects and has installed a great number of electronic archives in the public and the private sector.

In this paper the key topics for user and information access are addressed. Issues regarding litigation, privacy protection and network attacks need to be addressed in order to provide secure access to citizens. The ability to locate and identify relevant information is becoming key — with the portal as a paradigm for the rich function needed for information access. Planning for any significant IT application requires knowledge about standards — in particular with open application that will interact with many other systems. Protection of public information is not only about how to avoid hacker attacks. Governments need validated audit trails of their information interchange with their citizens, and there is a need for building proof of authenticity into the information infrastructure. The White Paper also describes the main drivers for architectural change. Special focus is given to topics like digital rights protection, audit trails and logging, and standards for access, architectures, security and directory services for user rights management. The chapters of the White Paper cover 'The challenge of open access', methods of accessing and protecting public information, and standards for access and protection. The best practice applications are describing the solutions of the cities of Naestved and Skurup in Denmark, and of the National Danish Art Museum. The outlook gives an overview about technology benefits and critical success factors.

IWP5 Availability and preservation

This fifth industry White Paper, authored by Kodak, covers issues of long-term availability and preservation of digital information. Kodak specialises in both analogue microfilm and digital optical storage technologies for long-term preservation of records, documents and media assets. The company was one of the early providers of these technologies and has numerous reference installations worldwide.

The paper offers Kodak's perspective on the long-term retention and availability of digital information. Digital documents require management just as their paper-based forerunners do. The electronic technologies used to create, distribute, and store them present special problems for archiving this information as time advances. Successive iterations of technology, inevitable media decay, and their inherent editability ill-suits them for long-term keeping in their native formats. A reference archive of permanent document images offers a cost-effective long-term solution. By rendering digital information to microfilm as uncoded, analogue images, organisations may create technology-proof repositories. The information stored has to be made available for decades, even centuries, including issues of migration and secure storage media. The focus topics include migration issues, advice on mix of technologies, standards for formats and the use of reference archive media systems. The White Paper contains chapters on 'The virtual memory' and 'Long-term archival'. The described best practice applications are the United Kingdom Census 2001 and the State of Virginia. The last chapter provides an insight on digital preservation strategy and benefits with reference archives.

IWP6 Education, training and operation

Industry White Paper number six was jointly authored by the UCL University College of London, TRW systems, Austria, and Communicando, Italy. The White Paper addresses changing role of records managers: 'From the traditional archivist to the information manager'. UCL is one of the developers and providers of E-TERM educational programmes. TRW systems and Communicando are providing e-Learning solutions. The idea behind this White Paper is not only to use electronic solutions to manage records, documents and assets, but to use software as well as to provide education to archivists and records managers.

The White Paper looks at issues of education, training and information management in an electronic world. It considers the challenges faced by universities and institutions of higher education, some of the new pedagogic methods under development and the new possibilities for continuing professional development and lifelong learning. It analyses the market drivers and requirements for e-learning, discusses some of the potential benefits of software-based learning and argues that businesses and corporate institutions in the 21st century must have and implement a learning and training vision.

The main topics of the White Paper focus on 'Archivists and information managers', managing and maintaining digital archives, education and training requirements, computer-based training and e-learning. Special respect is given to how to qualify records managers for the digital age, how to train effectively and how to run projects to implement solutions. The best practice examples were

provided by Communicando on business-to-business and business-to-consumer solutions. The final chapter describes possible ways from strategy to implementation and the requirements of change management.

Challenges

'Electronic archives are the memory of the information society' — this headline of Erkki Liikanen's has become most prominent in publications, marketing brochures and conference presentations. It demonstrates clearly the dimension of the challenge the information society is facing. Whilst we are still struggling with the basic technologies to create electronic archives, the development of information technology is heading for future visions like 'Information at your fingertips' or 'The Internet is the global brain and memory of mankind'. Archive solutions are today no mainstream software. Electronic archives and records-management systems are offered as special products for a special purpose. In fact, these systems should belong as infrastructure in every software and should be delivered as part of the basic operating system services. Everybody needs electronic archival. This is not a special problem of 'some archivists or records managers in dusty halls below museum floors'. The preservation of valuable information is a task for everybody dealing with electronic information.

Consolidation

Electronic archival and records management is still a niche market segment. Most of the software manufacturers are middle-sized companies. The crisis of the ITC industry since the beginning of this century, driven by the fallout of dot.coms, hit as well providers of solutions for long-term electronic archival.

- The vendor market is undergoing a consolidation phase.

This development has a special impact for those who already installed electronic archives which are now no longer available or supported. On the other hand there is no reason to panic: thinking in decades and centuries of availability requires strategies for 'constant' or 'continuous' migration. Companies will always rise and fall over the years. Migration strategies should be independent from these natural cycles.

- The shake-down has already hit document management and electronic archival suppliers.

The next wave of consolidation will hit the vendors of content management and portal solutions. There are too many product offers on the marketplace. But good ideas and innovative products will survive if they are absorbed by larger companies who can provide more stability.

- To provide archival software is 'to do the splits'

Customers always ask for the latest technologies and newest features. On the other hand they want from archiving and records management solutions that they keep their information available for years, decades and centuries. To serve both demands is nearly impossible for smaller-sized software companies. They have not enough resources to do both, deliver latest functionality and providing stable solutions for long-term information storage. Archival and records management solutions should stick to the basic requirements.

Standards

Standards are necessary for vendors and users alike. Vendors can streamline their development, add components from other manufacturers, and link to other systems more easily using standards. Users need standards to have test criteria, gain independence from proprietary products and ensure long-term availability of their information.

- Standards on different levels for different purposes

Without standards and interchange formats there will be no document interchange. Without predefined structures and defined metadata there will be no long-term accessibility and no cross-over usage of information.

- Standards are developing, changing and disappearing

Although standards can provide more safety for investments and information availability, they are no final lifeline. Migration has to be implemented as a regular, continuous process. Standards help to make migration easier.

- Standards must be auditable

General standards like document formats and interfaces are normally specified in detail and easy to test for compliance. Complex standards are often restricted to a functional description. Important standards for electronic archival and records management like MoReq Model Requirements or ISO 15489 Records Management have to be enhanced with auditable compliance criteria to enable vendors to deliver compatible solutions and to enable users to test on compliance.

Redundancy

We face redundancy in different areas: development of products, sponsoring of projects, redundancy in information itself. Redundancy in computer systems and system architectures is an attribute of safety and security. Redundancy as mentioned above is a risk.

- Redundancy in product development

Every supplier creates his products individually and independently. This leads to a lack of standardised, multipliable solutions. 'Re-inventing the wheel' costs too much resources, money and time, and endangers standardisation and the implementation of standards.

- Redundancy in information

Uncontrolled renditions, copies and reuse lead to a mountain of information where the original content often gets lost. New techniques are necessary to identify original content and protect it against unlawful use and reuse.

- Redundancy costs resources

Archivists and records managers will spend a lot of their future working time on sorting out which information is valid to be saved for future generations. Intelligent software tools are needed to support and partially automate these processes.

Coordination

A European approach to develop standardised procedures, metadata structures and best practice applications, to create stable solutions for long-term archival of valuable information, and to allow access to electronic archives to the European citizen, makes sense. To make it a success the DLM community needs to join forces. No vendor, no archive, no institution, and even no Member State alone can provide the 'final solution' for the 'memory of the information society' 'out of the box'. To achieve the goal:

- more and effective coordination is necessary;
- redundancy has to be avoided;
- initiatives like DLM have to be transformed into sustainable networks;
- criteria for auditing standards to improve compatible solutions have to defined;
- initiatives like 'e-government', 'Open access to public archives' and other related topics, beginning to overlap more and more, have to be harmonised. The coordination body for this task could be the initiated European DLM-network of excellence.

Cooperation

Cooperation between the ICT industry and the user organisation in the European public sector is essential. Up to now cooperation mostly happened only in individual projects to create a specific solu-

tion. Cooperation on a higher level, between competing vendors, standardisation organisations, and coordination bodies on the European and the Member State level are necessary to develop standards and certification procedures to create compatible solutions. The demand is that the ICT industry:

- cooperates with the public sector not project-by-project but in a continuous process on the European level;
- delivers standardised, affordable, easy to adopt, install and run solutions;
- takes the term 'long-term availability' seriously, and provides strategies and tools to meet the challenge of the DLM community;
- undertakes own efforts to avoid incompatible, individual solutions on the European, national and local level.

A mission?

Information policy, standardisation, coordination, education and qualification issues have been addressed by the DLM-Forum since its creation.

In 1999 the DLM-Forum addressed the challenge:

- to move the traditional archivists from the end of the information chain upwards;
- to become the information manager; and
- to get control over the complete document life cycle.

In 2002 the DLM-Forum has to address politicians and administrations to:

- create more awareness about the value of information and the value of archives and;
- coordinate joint efforts more efficiently to avoid the evolving 'digital gap'.

Three simple statements to conclude this paper:

- electronic archives are the memory of the information society;
- information has a value of its own only if the information is used;
- document life-cycle management is just in the beginning.

La gestión del ciclo de vida del documento para el sector público europeo

Ulrich Kampffmeyer

La utilización de las tecnologías de gestión electrónica de la información, los documentos y sus archivos se han convertido en un componente esencial para la gestión documentaria en las instituciones públicas europeas. El continuo crecimiento de la información digital, los problemas de accesibilidad, preservación y disponibilidad a largo plazo, crean una necesidad cada vez mayor de encontrar soluciones profesionales. Para abordar el más amplio alcance de la iniciativa DLM, sus siglas originales se redefinieron recientemente como «Gestión del ciclo de vida del documento» (DLM = Document Lifecycle Management). La asociación mundial de vendedores y usuarios de estas tecnologías (AIIM Internacional) ha presentado a la comunidad DLM una serie de seis Libros Blancos que recogen las mejores soluciones prácticas. Una descripción de los desafíos, las tendencias y la evolución de los mercados en el tema de las tecnologías relacionadas con la gestión documentaria completa la presente contribución.

Document life-cycle management für den europäischen öffentlichen Sektor

Ulrich Kampffmeyer

Die Verwendung elektronischer Informations-, Dokumenten- und Archivmanagementtechniken wird für Archive und Verwaltungen der Institutionen des europäischen öffentlichen Sektors immer wichtiger. Das extensive Anwachsen digitaler Informationen und die Probleme von Langzeitzugriff, Aufbewahrung und Verfügbarkeit verlangen zunehmend nach professionellen Lösungen. Um dem größeren Umfang der DLM-Initiative gerecht zu werden, wurde das ursprüngliche Akronym in DLM Document Lifecycle Management umdefiniert. Der internationale Anwender- und Herstellerverband AIIM International stellte der DLM-Gemeinde sechs Weißbücher mit bewährten Lösungsverfahren vor. Ein Überblick über die Probleme, Trends und die dem Wandel unterworfenen Märkte für archivgutbezogene Technologien beschließt diesen Beitrag.

La gestion du cycle de vie des documents dans le secteur public européen

Ulrich Kampffmeyer

L'utilisation des technologies de gestion électronique de l'information, des documents et des archives est devenue un aspect essentiel du classement et de la gestion documentaires dans les institutions publiques européennes. La forte augmentation des informations numériques et les problèmes d'accessibilité, de conservation et de disponibilité à long terme de celles-ci rendent de plus en plus nécessaire la définition de solutions professionnelles. Pour donner suite à l'élargissement du champ d'application de l'initiative DLM, l'abréviation originale a été redéfinie et recouvre désormais l'expression «document lifecycle management» (gestion du cycle de vie des documents). L'association mondiale des vendeurs et des utilisateurs de ces technologies (AIIM International) a présenté à la communauté DLM une série de six livres blancs reprenant les meilleures solutions pratiques. Un aperçu des défis, des tendances et de l'évolution des marchés dans le domaine des technologies liées à la gestion documentaire complète la présente contribution.

Paul E. Murphy

Paul E. Murphy works in the IDA (interchange of data between administrations) programme of the Enterprise DG, where, among other things, he was the project officer for the MoReq project. The objective of MoReq was to develop model functional requirements for the management of electronic records.
Paul is also responsible for the new IDA metadata project — MIReG (managing information resources for e-government). The objective of MIReG, which is being developed in close cooperation with the Dublin Core Metadata Initiative, Statens Information (Denmark), the Office of the e-Envoy (UK) and the European Parliament is to define, based on the Dublin Core, a metadata set for the retrieval of government information.
Prior to his secondment to the European Commission, Paul worked in CMOD (the Centre for Management and Organisation Development) of the Ministry of Finance in Ireland. Here, his responsibilities included advising government ministries on records management requirements arising from the adoption of Ireland's Freedom of Information Act. Paul was also Ireland's national coordinator for the EU Green Paper on Public Sector Information.
He is a member of the Scientific Committee of the DLM-Forum 2002, and has a BSc degree and a Masters degree in public administration from University College Dublin.

Metadata standards and model requirements for electronic document and records management

Paul E. Murphy

Introduction

As the main focus of this DLM-Forum is best practice and solutions this presentation will cover best practice for records management and solutions for metadata.

The presentation will cover:

— *information on the IDA programme;*

— *best practice recommendations for records management;*

— *developments in metadata for pan-European e-government applications.*

IDA (interchange of data between administration)

At the request of the DLM-Monitoring Committee, MoReq developed by the IDA programme of the Enterprise DG of the European Commission.

The IDA programme is a strategic EC initiative for 1999–2004.

Its mission is to coordinate the exchange of information between the Member States and the European Commission, as well as between the European Institutions, to support the management of the single market, the Community decision-making process, and the implementation of a wide range of Community policies, in areas such as agriculture, employment, the environment, statistics, and health and consumer protection, that affect Europe's citizens and enterprises.

More information on the IDA programme can be found on the IDA web site at:
http://europa.eu.int/ISPO/ida/.

Records management

Records management best practice is distilled in the IDA deliverable MoReq (Model Requirements for the management of electronic records). For records management best practice, the recommendation is simply to follow the recommendations of MoReq.

This is because MoReq:

— was developed, not in isolation from other records-management specifications, but includes best practice recommendations from 12 foundation reference texts on records management, used in archives and in the public sector;

— has a firm basis in records management theory;

— is aimed at a wide target audience. These are:
  - potential and actual ERMS (electronic records management system) users;
  - training organisations;
  - academic institutions;
  - ERMS suppliers and developers;
  - records management service providers; and
  - potential users of externalised records-management services;

— was developed in a quality model that included validation by:
  - an independent records management expert (UK);

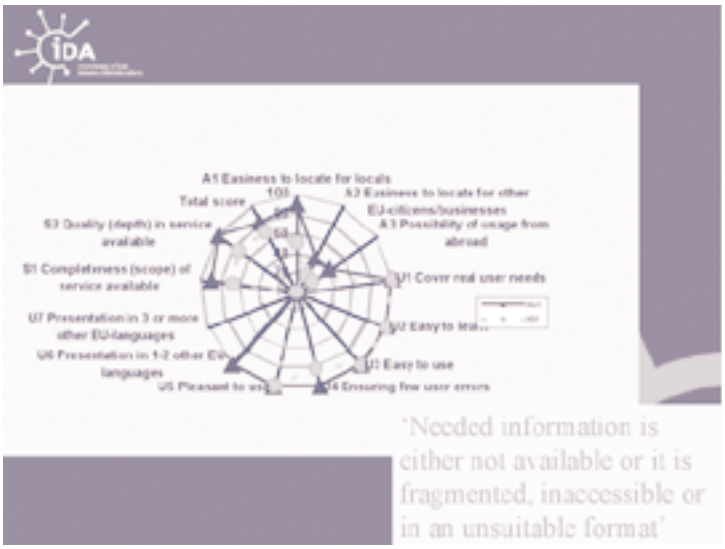
- an international guiding team of experts from records management and archives (CA, D, E, F, I, NL and P);
- validation organisations in:

— the enterprise sector, records management software suppliers, and public administration;

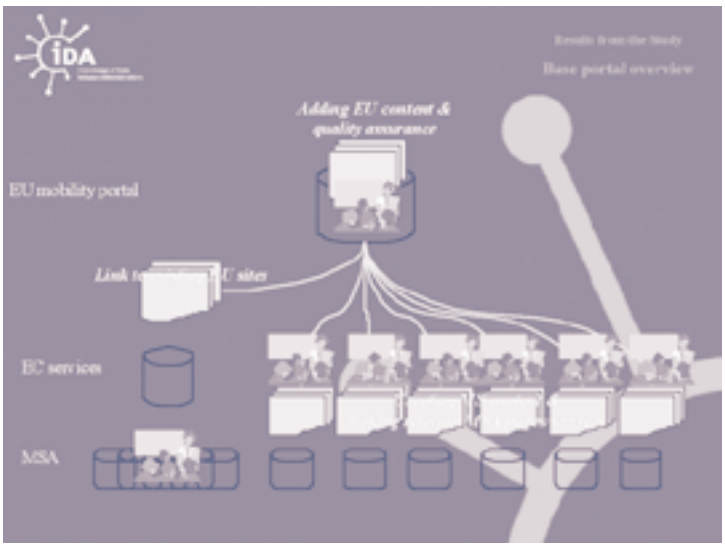
— in three EU Member States (Spain, Italy and the United Kingdom);
- is freely available (download from http://www.europa.eu.int/ispo/ida);
- is being translated from English, on request from the Member States. Translations in Spanish, Portuguese (and soon Italian) are available;
- is not in opposition to but is complementary with national public administration records management guidelines.

Metadata

Studies by the IDA in the area of access to public-sector information conclude that frequently: 'needed information is either not available or it is fragmented, inaccessible or in an unsuitable format'.



In addition, the IDA is building, in cooperation with the Member States, a 'Portal of the EU administration' that will provide a single point of access to public administration information and e-government services in Europe.



The importance of metadata in providing ease of access to the disparate and heterogeneous information resources of Europe's public administrations cannot be overemphasised:

The association of standardised descriptive metadata with networked objects has the potential for substantially improving resource discovery capabilities by enabling field-based (e.g. author, title) searches, permitting indexing of non-textual objects, and allowing access to the surrogate content that is distinct from access to the content of the resource itself. (Weibel and Lagoze, 1997)

MIReG

In this regard, the IDA launched this year a project MIReG (Managing Information Resources for e-government). MIReG is a metadata project, being developed with the guidance of the Dublin Core Metadata Initiative, Statens Infoamation (Denmark), Office of the e-Envoy (United Kingdom), and the European Parliament.

Among its objectives is to examine recommendations made by the Member States for metadata for public-sector information, recommendations and standards, such as the Dublin Core, and recommendations made by public administrations elsewhere. The intention is to produce a metadata model for European public-sector information and a framework for its use. This will result in recommendations for extensions to the Dublin Core that are required for the management of public-sector information.

Member States participation:

MIReG is being developed in cooperation with Member States, the EEA and candidate countries. Examples of national recommendations for public-sector information include:

Denmark
Recommendation on use of Dublin Core:
http://www2.si.dk/netsteder/raad/metadata/index.html
http://www.oio.dk/

Finland
Recommendations on usage of metadata for government documents:
http://www.intermin.fi/juhta/suositukset/jhs143.htm government documents.
Recommendations on usage of metadata for electronic forms.
http://www.intermin.fi/juhta/suositukset/jhs144.htm

Hungary
Kikeres information inventory
http://www.kikeres.hu/

Iceland
Lysigagnahandbók
http://www.menntagatt.is/lysigagnahandbokin.html http://www.menntagatt.is/pdf/handbok.pdf

Ireland:
Recommendation on use of Dublin Core:
http://www.gov.ie/webstandards/metastandards/ipsms_part1.doc

United Kingdom
e-government metadata framework (e-GMF)
http://www.e-envoy.gov.uk/publications/frameworks/metadata/metadata.htm

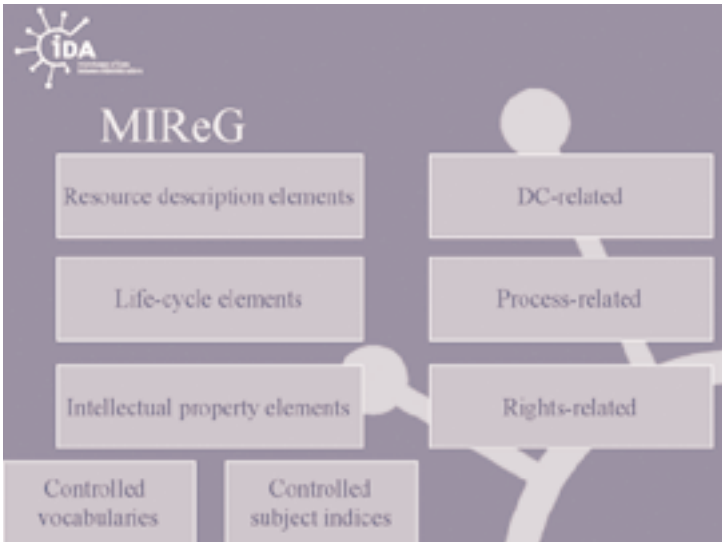
MIReG: work-in-progress

As MIReG is a work-in-progress it would not be appropriate to present an unfinished model to this conference. However, MIReG is developing on the lines of:

— resource description elements (i.e. DC-related);

— life-cycle elements (i.e. process-related elements);

— intellectual property elements (i.e. coverage and rights-related elements)



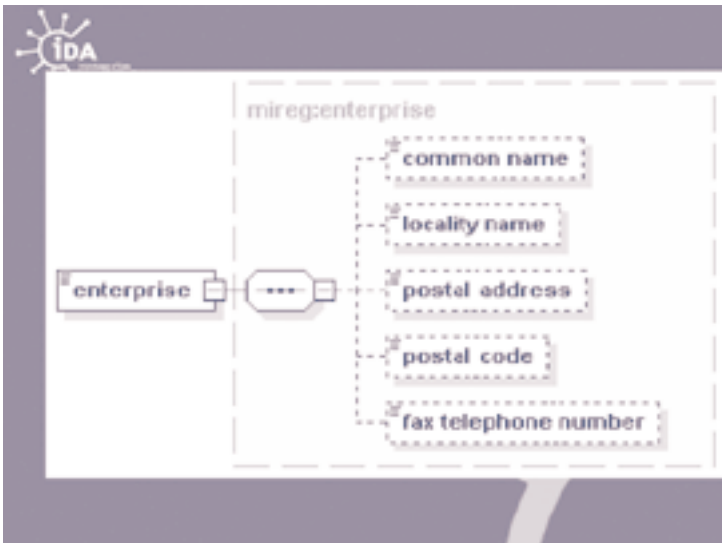
The DC is comprised of content, intellectual property and instantiation elements. The development of MIReG shows the DC as the source of the resource description elements, and the life-cycle and IP elements being developed from the recommendations of the Member States.

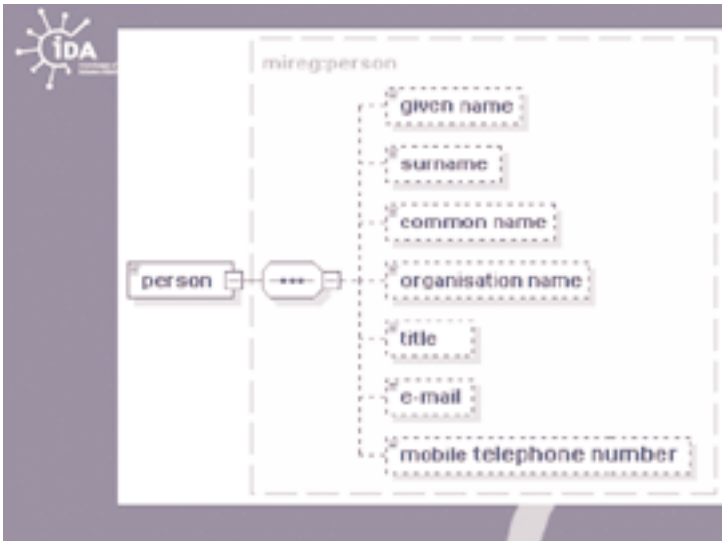
To promote interoperability across Europe, a number of the element descriptions may be associated with controlled vocabularies for the respective element values. Member States would also be free to develop controlled vocabularies for interoperability within their national domains.

In addition, MIReG may be combined with controlled subject indexes, to allow more precise searching and records management.

Current status

Currently the model exists as a set of XML documents (RDF and Schema) that are not very readable. The following slides show the documentation of the model as generated by the development tool (XMLSpy).





Management of metadata

MIReG is concerned not only with the development of a metadata model for European public administrative information, but also for tools for the management of metadata.

Metadata can provide responses to the 'What, where and how' needs of public administration information users. These include:

What:

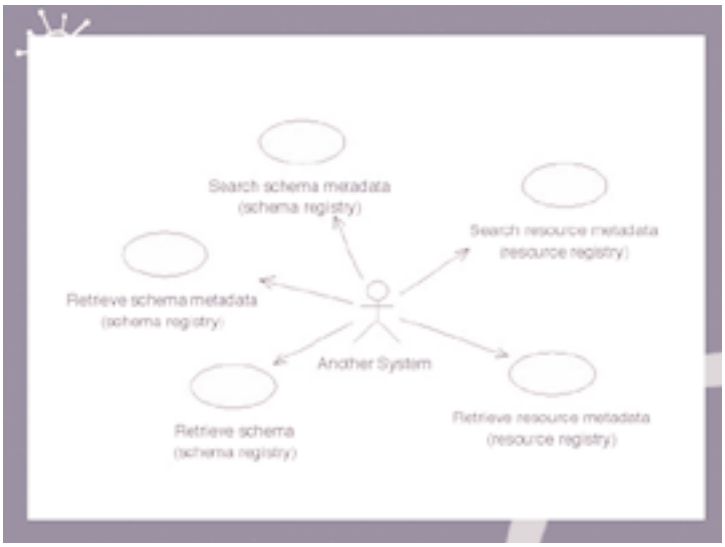
— What is this document about?

— What organisation is responsible for this subject?

— What is its status?

Where

— Where is the document about this subject?



How

— Can I access it?

— How can I contact the person or organisation?

— How is this document processed?

Who

— Which persons and organisations are involved in the process?

Registries

These information use needs can be met by the development of registries for managing 'e-government aware' resource descriptions.

Therefore, in addition to the development of the MIReG metadata model, the IDA is in the process of developing a tool to manage the metadata framework. This will consist of a series of registries designed to meet the above information needs.

Schema registry

A schema registry will answer the 'What' question, by storing definitions of metadata elements, qualifiers, their allowable values and value ranges as well as classification schemes, taxonomies, etc.

Resource registry

A resource registry will contain both content and contact details. It will answer the 'Where' and 'Who' questions by storing descriptions of resources and allowing users to search document metadata. It can also store information about persons, enterprises and organisations and allows users to search for that information.

A process registry would be required to answer the 'How'. But as processes relate primarily to the internal administrative functions of the Member States it is, at present, somewhat outside the scope of the IDA.

These registries will serve as common infrastructure of the e-government web-based information services; services that will be available to all kinds of applications by a simple and clearly defined, public application interface. While accessible by human users their main function would be providing application-to-application services. This will be done by a generic, well-defined and standard interface.

The provision of an open source metadata management tool to Europe's public administration should help those developing e-government applications.

Make metadata gathering as easy as possible

The first objective is to help the user and facilitate the filling of metadata. With this system when a user uploads a document into a content-management system and needs to fill its metadata, the application would go to the Schema Registry and get an RDF template with pre-defined data, for example, a selection list of controlled subject keywords. This would then allow the automatic capture of metadata elements but also high accuracy in metadata recording.

Manage 'e-government aware' resource descriptions

The second objective is to make easy the management of 'e-Government aware' resource descriptions for integration into various applications.

Open source software

The registries and management software are based on open source software:

— UML methodology for use cases and application design;

— Rational Rose as the UML tool;

— Java EJB (J2EE from java.sun.com);

— Apache web application and XML frameworks: Xerces for XML parsing, Xalan for XML transformations, SOAP for messaging, etc.;

— mySQL database (open source);

— Apache web server (open source from http.apache.org);

— Apache JServ servlet engine (open source for apache.org).

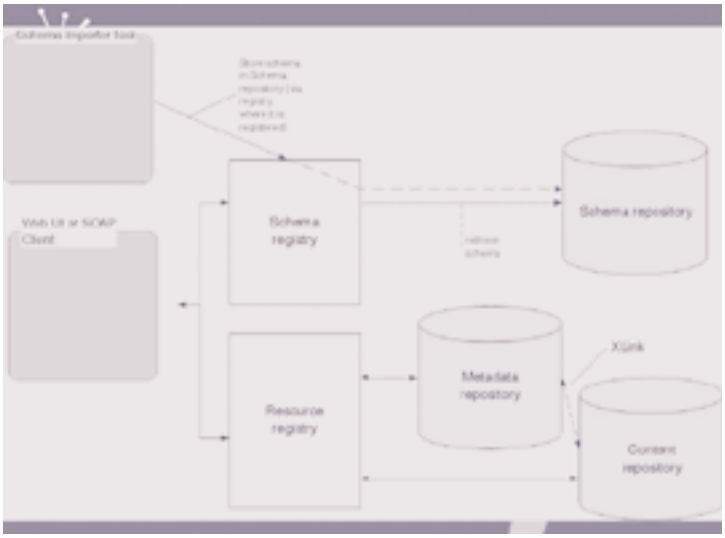
Metadata management tool

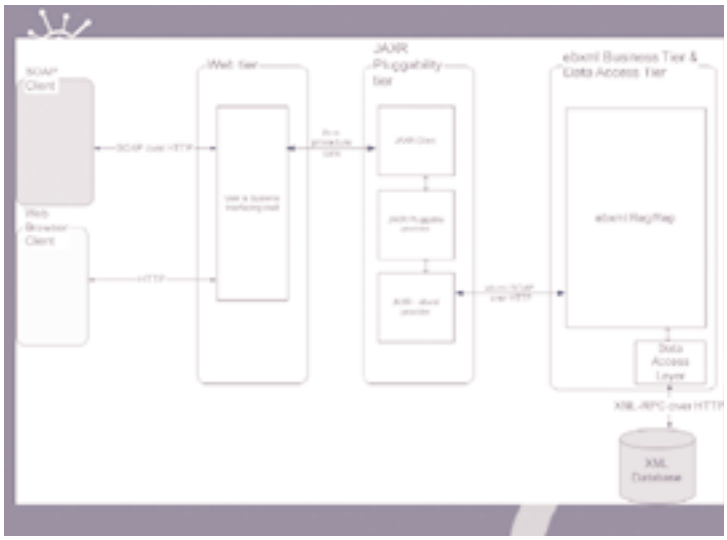
The following provides and idea of what the MIReG metadata management tool will look like:

1. The registry architecture
2. Use case diagram for a human user
3. Use case diagram for an application
4. Login to the system
5. Access the registries
6. Search a registry
7. Retrieve results
8. Display of results. Results can also be exported as a .zip file.

However, most of the interactions with the registries will be from computer applications that need, for example, to obtain schemae, and access will be via a well-defined API.

9. The protocols used in accessing the registries. Communications between the various components is achieved primarily by sending requests using the SOAP protocol over HTTP.





Availability

It is envisaged that the MIReG metadata set and framework will (hopefully) be available before the summer, and that the metadata management tools will be available for testing in the autumn.

The MIReG framework will be available for download from the IDA web site. The metadata tool will be made available to the Member States, EEA countries and candidate countries that participate in the IDA programme via the IDA TAC (telematics between administrations) management committee.

To sum up:

This presentation has covered:

— the IDA programme;

— best practice guidelines for records management — MoReq;

— developments in the area of metadata for pan-European e-government applications.

Normas sobre metadatos y modelo de requisitos para los documentos electrónicos y la gestión de documentos

Paul E. Murphy

Introducción

Dado que el principal objetivo de este Foro DLM son las mejores prácticas y soluciones, esta presentación tratará las mejores prácticas en materia de gestión de documentos y las soluciones para los metadatos.

La presentación tratará los siguientes ámbitos:

— información sobre el programa IDA;

— recomendaciones sobre mejores prácticas para la gestión de documentos;

— progresos en materia de metadatos para aplicaciones paneuropeas de administración electrónica.

IDA (intercambio de datos entre las administraciones)

A petición del Comité de supervisión DLM, el programa IDA de la Dirección General de Empresa de la Comisión Europea desarrolló unos MoReq (modelo de requisitos para la gestión de documentos electrónicos).

El programa IDA es una iniciativa comunitaria estratégica para 1999-2004.

Su misión es coordinar el intercambio de información entre los Estados miembros y la Comisión Europea, así como entre las instituciones europeas; apoyar la gestión del mercado único y el proceso de toma de decisiones comunitario; y la aplicación de una amplia gama de políticas comunitarias, en ámbitos tales como la agricultura, el empleo, el medio ambiente, las estadísticas y la salud y la protección al consumidor, que afectan a los ciudadanos y a las empresas de Europa.

Para más información sobre el programa IDA consúltese el sitio Internet de IDA en la dirección siguiente: http://europa.eu.int/ISPO/ida/.

Gestión de documentos

Las mejores prácticas de gestión de documentos figuran en los MoReq del IDA. Por lo que respecta a las mejores prácticas para la gestión de documentos, se aconseja simplemente seguir las recomendaciones de los MoReq.

Esto se debe no sólo al contenido de los MoReq, sino también a que se desarrollaron teniendo en cuenta a un amplio público. Además, el modelo de calidad bajo el cual se desarrollaron los MoReq incorporó el compromiso de un equipo internacional de expertos en gestión de documentos, así como la validación de los MoReq por organizaciones de los sectores empresarial y público. Ahora se empieza a disponer de traducciones de los MoReq.

Metadatos

Los estudios realizados por el IDA en el ámbito del acceso a la información del sector público concluyen que frecuentemente *«la información necesaria no está disponible o bien está fragmentada, inaccesible o en un formato inadecuado»*.

No puede hacerse demasiado hincapié en la importancia que tienen los metadatos para proporcionar un acceso fácil a los dispares y heterogéneos recursos de información de las administraciones públicas de Europa. Este acceso es necesario no sólo para el eficaz funcionamiento de las administraciones públicas de Europa, sino también para el desarrollo de los servicios de la administración electrónica, particularmente cuando éstos comportan un elemento transfronterizo.

MIReG

A este respecto, el IDA lanzó este año un proyecto MIReG («Managing Information Resources for e-Government» o «Gestión de recursos de información para la administración electrónica»). MIReG es un proyecto sobre metadatos, desarrollado bajo la dirección de Dublin Core Metadata Initiative, Statens Information (Dinamarca), Office of the E-Envoy (Reino Unido) y el Parlamento Europeo.

Entre sus objetivos figura el examen de las recomendaciones realizadas por los Estados miembros sobre metadatos para la información del sector público, de recomendaciones y normas, tales como el Dublin Core, y de recomendaciones realizadas por otras administraciones públicas. La intención es elaborar un modelo de metadatos para la información del sector público europeo y un marco para su uso. Ello dará lugar a recomendaciones para ampliar el Dublin Core, necesarias para la gestión de la información del sector público.

Estado actual

Actualmente el modelo MIReG existe, no como un documento en lengua inglesa, sino como un conjunto de documentos XML (RDF y Schema) que no son muy legibles. Sin embargo, MIReG se está desarrollando en las siguientes líneas:

— elementos de descripción de recursos (esto es, relacionados con el Dublin Core);

— elementos de ciclo de vida (esto es, elementos relacionados con el proceso);

— elementos de propiedad intelectual (esto es, cobertura y elementos relacionados con los derechos).

Gestión de metadatos

MIReG se ocupa no sólo del desarrollo de un modelo de metadatos para la información administrativa pública europea, sino también de las herramientas para la gestión de metadatos. Esto se debe a que los metadatos pueden proporcionar respuestas al «qué, dónde y cómo» de las necesidades de información de los usuarios de la administración pública.

Registros

Estas necesidades de uso de la información pueden cubrirse con el desarrollo de registros para la gestión de descripciones de recursos que tengan en cuenta la administración electrónica. De esta manera, se está desarrollando una herramienta de gestión de metadatos *open source* como parte del proyecto MIReG. Esto servirá para gestionar tanto los registros de recursos como los esquemas.

Estos registros servirán de infraestructura común para los servicios de información de la administración electrónica basados en Internet, servicios que estarán disponibles para todo tipo de aplicaciones mediante una interfaz de aplicación pública sencilla y claramente definida. Si bien estará accesible para usuarios humanos, su principal función será proporcionar servicios de aplicación a aplicación. Esto se realizará a través de una interfaz genérica, bien definida y estándar.

Disponibilidad

Se ha previsto que el marco y el conjunto de metadatos MIReG estén disponibles antes del verano, y que las herramientas de gestión de los metadatos estén disponibles para su prueba en otoño.

El marco MIReG estará disponible para ser descargado del sitio web del IDA. La herramienta de metadatos se pondrá a disposición de los Estados miembros, de los países del Espacio Económico Europeo (EEE) y de los países candidatos que participen en el programa IDA a través del Comité de Gestión IDA TAC (telemática entre las administraciones).

Metadatenstandards und Musteranforderungen für die Verwaltung elektronischer Aufzeichnungen und Archive

Paul E. Murphy

Einleitung

Da der Schwerpunkt dieses DLM-Forums auf bewährten Verfahren und Lösungen liegt, beschäftigt sich dieser Vortrag mit bewährten Verfahren für die Archivverwaltung und Lösungen für Metadaten.

Er umfasst im Einzelnen:

— Informationen über das IDA-Programm;

— Empfehlungen für bewährte Verfahren in der Archivverwaltung;

— Metadaten-Entwicklungen für gesamteuropäische *e*-Government-Anwendungen.

IDA (Datenaustausch zwischen Verwaltungen)

Auf Wunsch des DLM-Begleitausschusses wurden MoReq (**Mo**del **Req**uirements for the management of electronic records – Musteranforderungen für die Verwaltung von elektronischen Aufzeichnungen) im Rahmen des IDA-Programms der Generaldirektion Unternehmen der Europäischen Kommission entwickelt.

Das IDA-Programm ist eine strategische EG-Initiative für 1999–2004.

Ziel ist die Koordinierung des Informationsaustauschs zwischen den Mitgliedstaaten und der Europäischen Kommission sowie zwischen den europäischen Organen zur Unterstützung der Verwaltung des Binnenmarktes, der Beschlussfassung in der Gemeinschaft und der Umsetzung einer Reihe politischer Gemeinschaftsmaßnahmen in Bereichen wie Landwirtschaft, Beschäftigung, Umweltschutz, Statistik sowie Gesundheit und Verbraucherschutz, die für die Bürger und Unternehmen in Europa von Belang sind.

Weitere Informationen zum IDA-Programm finden sich auf der IDA-Webseite unter: http://europa.eu.int/ISPO/ida/

Archivverwaltung

Bewährte Verfahren für die Archivverwaltung sind herausgearbeitet in den lieferbaren IDA-MoReq. Im Hinblick auf bewährte Verfahren für die Archivverwaltung lautet die einfache Empfehlung, den MoReq-Muster-Anforderungen zu folgen.

Dies liegt nicht nur an den MoReq-Inhalten, sondern auch daran, dass sie mit Blick auf eine breite Zielgruppe konzipiert wurden. Ferner beinhaltete das Qualitätsmodell, nach dem die MoReq entwickelt wurden, die Wünsche eines internationalen Leitungsteams von Archivverwaltungsfachleuten und die Validierung der MoReq durch Organisationen im Unternehmens- sowie im öffentlichen Sektor. MoReq-Übersetzungen sind mittlerweile verfügbar.

Metadaten

Studien, die von IDA im Bereich Zugang zu Informationen des öffentlichen Sektors durchgeführt wurden, ergeben, dass häufig *„die benötigten Informationen entweder nicht verfügbar, verstreut oder unzugänglich sind oder aber in einem unpassenden Format vorliegen“*.

Die Bedeutung von Metadaten bei der Bereitstellung eines einfachen Zugangs zu den verstreuten und heterogenen Informationsquellen der öffentlichen Verwaltungen in Europa kann nicht genug betont werden. Dieser Zugang ist nicht nur im Hinblick auf eine effiziente Arbeitsweise der öffentlichen Verwaltungen in Europa erforderlich, sondern auch für die Entwicklung von *e*-Government-Diensten, insbesondere da, wo diese grenzüberschreitende Elemente aufweisen.

Managing Information Resources for e-Government (MIReG)

In diesem Zusammenhang hat IDA in diesem Jahr ein **MIReG** (**M**anaging **I**nformation **R**esources for **e-G**overnment)-Projekt gestartet. MIReG ist ein Metadaten-Projekt, das unter Anleitung der Dublin Core Metadaten-Initiative, Statens Information (DK), Office of the E-Envoy (UK) und des Europäischen Parlaments entwickelt wurde.

Zu seinen Aufgaben gehört die Prüfung der Empfehlungen der Mitgliedstaaten in Bezug auf Metadaten für Informationen des öffentlichen Sektors, Empfehlungen und Standards wie die Dublin Core Initiative sowie Empfehlungen anderer öffentlicher Verwaltungen. Ziel ist die Erstellung eines Metadatenmodells für Informationen des europäischen öffentlichen Sektors und eines Rahmens zu ihrer Nutzung. Dies führt zu Empfehlungen für eine Ausweitung auf die Dublin Core Initiative, die für die Verwaltung von Informationen des öffentlichen Sektors notwendig sind.

Stand der Dinge

Gegenwärtig existiert das MIReG-Modell nicht als englisches Sprachdokument, sondern als eine Reihe von XML-Dokumenten (RDF und Schema), die nicht besonders gut lesbar sind. Allerdings entwickelt sich MIReG entlang von:

— Resource Description-Elementen (d. h. DC–bezogen);

— Life Cycle-Elementen (d. h. prozessbezogene Elementen);

— Elementen des geistigen Eigentums (d. h. Berichterstattungs- und rechtebezogene Elementen).

Verwaltung von Metadaten

MIReG beschäftigt sich nicht nur mit der Entwicklung eines Metadatenmodells für Informationen der öffentlichen Verwaltung in Europa, sondern auch für Instrumente zur

Verwaltung von Metadaten, weil Metadaten Antworten auf die Fragen „Was, Wo und Wie“ von Benutzern in der öffentlichen Verwaltung bieten kann.

Registraturen

Dieser Informationsbedarf kann gedeckt werden durch die Entwicklung von Registraturen zur Verwaltung von „e-government-bewussten“ Resource Descriptions. Daher werden im Rahmen des MIReG-Projekts Instrumente zur Verwaltung von Open Source Metadaten entwickelt, die sowohl Schema- als auch Resource-Registraturen verwalten.

Diese Registraturen dienen als allgemeine Infrastruktur der webgestützten e-government-Informationsdienste, Dienste, die für jede Art von Anwendungen durch eine einfache und eindeutig definierte öffentliche Anwendungsschnittstelle zur Verfügung stehen. Sie ist zwar für menschliche Anwender zugänglich, ihre Hauptfunktion ist jedoch die Bereitstellung von Anwendungs-zu-Anwendungsdiensten durch eine generische, genau definierte und genormte Schnittstelle.

Verfügbarkeit

Vorgesehen ist, dass der MIReG-Metadatensatz und -Rahmen (hoffentlich) vor Mitte des Jahres und die Instrumente zur Verwaltung der Metadaten im Herbst zur Prüfung vorliegen.

Der MIReG-Rahmen wird zum Herunterladen von der IDA-Webseite bereitgestellt. Das Metadaten-Instrument wird den Mitgliedstaaten, EWR-Ländern und Bewerberländern, die über den IDA-TAC (Telematik zwischen Behörden)-Verwaltungsrat am IDA-Programm teilnehmen, bereitgestellt.

Normes et spécifications types pour les métadonnées en matière de gestion des documents et des archives électroniques

Paul E. Murphy

Introduction

Dans la ligne du thème central du présent Forum DLM, nous axerons notre présentation sur les bonnes pratiques en matière de gestion de documents et sur les solutions en matière de métadonnées.

Nous proposerons successivement:

— des informations concernant le programme IDA;

— des recommandations de bonnes pratiques en matière de gestion de documents et d’archives;

— une description des évolutions des métadonnées dans la perspective d’applications pan-européennes dans l’administration électronique.

IDA (Interchange of data between administrations)

À la demande du comité de suivi DLM, des spécifications types (MoReq) ont été développées dans le cadre du programme IDA (échange électronique de données entre administrations) de la Commission européenne (DG Entreprises).

Initiative stratégique communautaire couvrant la période 1999-2004, le programme IDA a pour mission de coordonner les échanges d'informations entre les États membres et la Commission européenne de même qu'entre les institutions européennes, dans le but de faciliter le fonctionnement du marché intérieur, d'optimiser la prise de décision à l'échelon communautaire et d'accélérer la mise en œuvre de la politique européenne dans des domaines aussi divers que l'agriculture, l'emploi, l'environnement, les statistiques, la santé et la protection des consommateurs, qui touchent directement les citoyens et les entreprises d'Europe.

Des informations plus détaillées concernant ce programme peuvent être consultées sur son propre site web (http://europa.eu.int/ISPO/ida/).

Gestion des documents et des archives

Les bonnes pratiques en matière de gestion de documents et d'archives sont incorporées dans le produit MoReq (Model Requirements ou spécifications types pour la gestion des documents et des archives électroniques) du programme IDA. Les recommandations MoReq sont donc la voie à suivre dans ce domaine, non seulement en raison du contenu même de ces spécifications, mais également parce qu'elles ont été conçues à l'intention d'un large public cible. Il convient de préciser, en outre, que les points de vue d'une équipe internationale d'experts en gestion de documents ont été intégrés au modèle de qualité dans le cadre duquel les spécifications MoReq ont été élaborées, et que ces dernières ont été elles-mêmes approuvées par des organisations du secteur privé comme du secteur public. Les spécifications MoReq deviennent progressivement disponibles dans d'autres versions linguistiques.

Métadonnées

Des études réalisées dans le cadre du programme IDA à propos de l'accès aux informations du secteur public concluent que, très fréquemment, les données requises ne sont pas disponibles, ou qu'elles sont fragmentées, inaccessibles ou présentées dans un format inapproprié.

On ne soulignera jamais assez l'importance des métadonnées pour faciliter l'accès aux ressources disparates et hétérogènes des administrations publiques européennes — un accès impératif non seulement pour l'efficacité de fonctionnement de ces administrations publiques, mais également pour le développement de services dans le cadre de l'e-administration, surtout si les services en question revêtent des aspects transfrontaliers.

MIReG

C'est dans ce contexte que le programme IDA a lancé cette année le projet MIReG (Managing Information Resources for e-Government). Il s'agit d'un projet axé sur les métadonnées, dont les orientations proviennent de plusieurs sources: Dublin Core Metadata Initiative, Statens Infoamation (DK), Office of the E-Envoy (UK) et Parlement européen.

Le projet a notamment pour objectif d'examiner les recommandations formulées par les États membres concernant les métadonnées destinées aux informations du secteur public, les recommandations et les normes de type Dublin Core, et les recommandations émanant d'administrations publiques en dehors de l'Union. L'intention est de réaliser un modèle et un cadre d'utilisation pour les métadonnées applicables aux informations du secteur public. Ces travaux devraient aboutir à des recommandations permettant d'élargir le Dublin Core à la gestion des données détenues par le secteur public.

Situation actuelle

Le modèle MIReG existe déjà, mais il ne se présente pas sous la forme d'un document en langue anglaise: il est constitué d'une série de documents XML (RDF et Schéma) assez difficiles à lire. Le modèle MIReG évolue cependant en intégrant:

— des éléments descriptifs (liés au contenu);

— des éléments relatifs au cycle de vie (liés au processus);

— des éléments de propriété intellectuelle (liés à la couverture et aux droits).

Gestion des métadonnées

Le projet MIReG porte non seulement sur le développement d'un modèle pour métadonnées destinées aux informations des administrations publiques européennes, mais également sur la mise au point d'outils pour la gestion de ces métadonnées — ces dernières pouvant répondre aux besoins des utilisateurs publics désireux de savoir quelle information existe, où elle se trouve et comment y accéder.

Registres

L'instauration de registres pour la gestion des descriptions de ressources intégrant l'e-administration permet de répondre à ces besoins d'information. Aussi le projet MIReG veille-t-il également à développer des outils de gestion de métadonnées de source générale, capables de gérer à la fois les schémas et les registres de ressources.

Ces registres serviront d'infrastructure commune aux services d'information de l'e-administration basés sur le web, qui serviront à une très large gamme d'applications grâce à une interface publique simple et clairement définie. Il s'agit essentiellement de services interapplications, utilisant une interface générique standard, mais qui resteront accessibles à des personnes physiques.

Disponibilité

La série de métadonnées MIReG et leur cadre devraient être mis à disposition avant l'été, et les outils de gestion des métadonnées devraient être prêts aux essais dans le courant de l'automne.

Le cadre MIReG pourra être téléchargé à partir du site web du programme IDA. Les outils seront, pour leur part, mis à la disposition des États membres, des pays de l'EEE et des pays candidats qui participent au programme IDA par l'intermédiaire de son comité télématique entre administrations (CTA).

Parallel session 1

Tuesday 7 May 2002

THE MEMORY OF THE INFORMATION AGE
Preservation, migration and long-term availability

Chairperson:	Luciana Duranti (Canada)
Co-chair:	Francisco Barbedo (Portugal)
Rapporteur:	Thekla Kluttig (Germany)

Richard Blake

Richard Blake is Project Manager in the Electronic Records Management Development Unit (ERMDU) at the Public Records Office of the United Kingdom. This is the lead agency responsible for standard-setting in records management across the UK Government, and is very heavily involved in the cross-government target to achieve capability in electronic records management in the UK central Government by 2004.

Stephen Harries

Stephen Harries is head of the Electronic Records Management Development Unit in the Public Record Office of the United Kingdom.
Stephen has a background in information management and information science, and has worked in both public and private sectors, including higher education, software development and supply and as a consultant.

Sustaining electronic records for business purposes

Richard Blake, Stephen Harries

This paper discusses the nature and scope of the concept of sustainable electronic records in the context of government departments and agencies — the need for government organisations to rely assuredly on future access to electronic records for as long as they are needed for business purposes.

Background

Most (if not all) departments and agencies are already creating electronic documents, which may or may not, but probably should, be records. Implementation of electronic document and records management (EDRM) systems — a UK Government target for 2004 — will support organised capture of corporate documents across the enterprise, and from this date conformant government organisations will be relying on the electronic records for their formal corporate memory.

The business justification for EDRM lies primarily in the medium-term benefits which can be gained through continued access to reliable, authentic information, to support more effective use, better and faster access, and innovation in ways of working through continued integration with other enterprise information and line-of-business systems.

These benefits will only be sustained over the medium term if the value of the information is retained through good records management of the information content. While part of this is in secure storage and control of the artefact, which conveys the content, in order to retain authenticity a part is also in maintaining practical access to, and usability of, the content itself for business purposes.

Technology change will tend to depreciate the value of the electronic objects as artefacts: therefore the relationship between technology change and the electronic record must be managed to prevent loss of access, integrity and accountability.

There is a medium- to long-term need to manage for sustainability: sustainable access; sustainable integrity and accountability; sustainable benefits realisation.

The emphasis here is on a managed integration of electronic records with business information, rather than archival preservation of the historic record. This implies managing the information content as record, which may be delivered through different vehicles as technology platforms change — by carefully managing the transformation across platforms — in order to deliver the benefits expected and required.

Designing a good fit

Sustainability of electronic records for continued business use in government organisations is a complex area. The move towards electronic records in general requires the records manager to focus less on the management of the physical artefact — physical storage, circulation and disposal — and much more on design of the procedural, human and technical environment in which records are created, captured, used, maintained and eventually disposed of. The electronic nature of such records brings records management firmly into the central information management processes of the organisation, as one strand interwoven with others in managing and exploiting information and knowledge assets. Success will to a great extent depend on the skill with which this design makes a 'good fit' with, and delivers benefits for, the business and operational context; and the extent to which it supports wider organisational needs and developments.

This is no less true for the later parts of the record life cycle — sustainability — as it is for the immediate operational context of creation and capture. Sustainability is concerned not just with preservation of the physical artefact (which in itself presents difficult issues in the electronic environment) but also with sustaining the record in order to achieve the benefits predicated by an initial investment in the technology and organisational change necessary to capture and store the record

in the first place. These two aspects — sustainability of, and sustainability for — are complementary aspects and must be considered together in a coherent whole: neither is sufficient on its own.

If this perspective is accepted, it follows that the need is not simply to identify a list of categories of records which need to be maintained for specific periods of time, and determine the means to do so, although this is certainly part of the problem; but also to ensure that the means of achieving this is conditioned by, and consistent with, the broader context. Within a government department this broader context, which determines the fundamental reasons for sustaining the record — to continue to realise the business value and benefits in a changing and uncertain environment — has its own dynamic and shaping factors. These broader issues themselves feed back into issues of sustainability.

Sustainability is, therefore, best seen as an integral part of developing and improving the overall electronic-records management infrastructure, each aspect of which influences, and is influenced by, the others; and not as a separate, 'follow-on' activity. Elaboration of the concept involves consideration of:

- the electronic records themselves;
- the departmental context in which they are captured and used;
- the cross-government context, where records are part of an increasingly 'joined-up' information landscape;
- the technical, policy, economic and commercial factors which bring an influence to bear.

The model for sustainability used in this study is shown in Figure 1 on page 5. This consists of:

- a central core set of intrinsic dimensions of the electronic record in use;
- a middle circle of sustainability management objectives, showing the record in context;
- an outer field of broader environmental factors which influence the context of the record.

This is a fairly complicated model, which has the virtue of showing relationships between complex sets of factors; and can be used to begin to segment the issues involved, for further analysis.

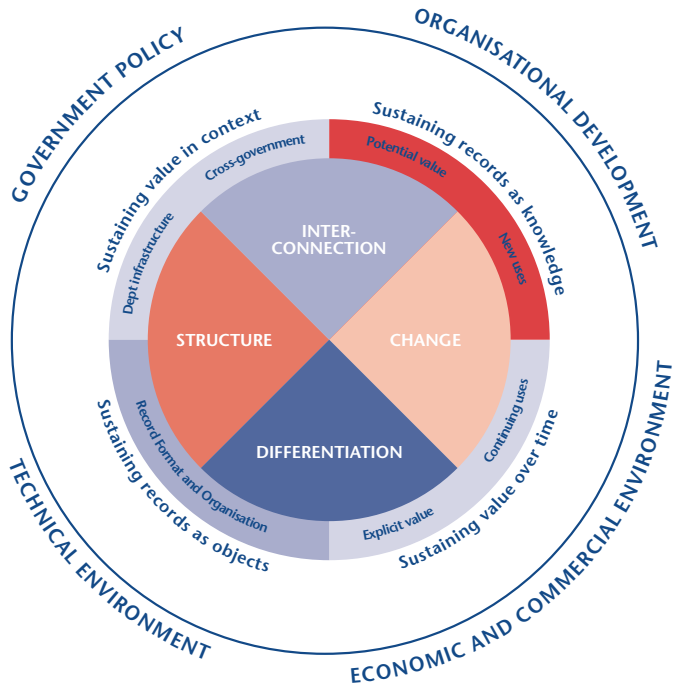


Figure 1: Model for sustainable electronic records

Structure is a dimension most closely associated with the paper record. To function as an element of corporate memory, an electronic record needs to retain both its own internal structure and coherence as an accurate record of what took place, and its place within the context of a set of records which reflect departmental business activities: part of an organised filing structure. Structure reflects the need for stability across time.

Change is a dimension of electronic information, which can be easily fragmented and re-assembled in response to the needs of new and changing circumstances, and support innovation through flexibility and adaptation. Electronic records are a form of electronic information, and much of the continued business benefit stems from an ability to access and reuse the content in a way not feasible with paper records.

There is an inherent tension between structure and change, and any sustainability strategy must aim to find an optimal balance which manages this tension.

The inter-connection dimension expresses the increasing requirement to integrate electronic records with other forms of corporate information asset. This works at both a departmental level — ensuring interoperability between records management and other corporate information and knowledge-management systems — and at a cross-government level: supporting the increasing need for data exchange and 'joined-up' working within and between sectors of government, releasing the potential value of the records.

Differentiation expresses the need to reduce uncertainty by ensuring accurate, complete and reliable sets of records. This is achieved through organising records into categories that make their explicit value plain: the reasons they were created and the uses to which they are put; by ensuring that up-to-date and full information can be retrieved when needed; and that records no longer required are formally and consistently removed according to established business rules.

While inter-connectedness stresses the similarity between different sets of records, differentiation stresses the dissimilarities: both are needed, and here too an appropriate balance must be maintained.

The various management objectives involved in sustainability can be usefully characterised by considering the relationship between adjacent dimensions in the model.

Sustaining records as objects

The relationship between the structure and differentiation dimensions is concerned with characteristics such as record format, component structure, relationship to context of creation and use, categorisation and filing structure. This involves managing records as physical objects to ensure their contents and contexts can continue to be accessed and used.

Sustaining value over time

The formal value of a corporate record lies in its continued integrity and authenticity as an accurate and accountable record of events. The purpose of managing the physical object is to retain this explicit value and support continuing use of the record in a changing environment, and the means of doing so must reflect this. The focus here is on managing the content of records, and on establishing structured retention and disposal.

Sustaining value in context

Essentially the value of a set of electronic records is realised to the extent that it contributes to the achievement of the business objectives to be supported. These objectives not only continue over time — past the point at which the initial record was created — but also span different parts of an organisation, and different organisations that have related business objectives. Different organisations, for example, may have different retention requirements.

Sustaining records as knowledge

To continue to realise the future potential of electronic records as corporate knowledge resources in support of continuing innovation, they will need to be maintained as meaningful content. This is the least certain, and least predictable, area of sustainability.

Segmenting the issues

The model described at Figure 1 above sets out the main aspects of the problem area, and shows their relationships. While complex, it can be used to help structure and segment the range of issues by overlaying with a matrix as displayed in Figure 2 below.

Managing structure/managing content

The left hand side of the model is primarily concerned with management of structure: of records themselves, of their inter-relationships within a departmental system, and of their relationship with other resources in emerging cross-government infrastructure. The issues here tend to be 'harder': for example, technological frameworks and formats, standards for interoperability.

The right-hand side is primarily concerned with managing content, in order to unlock the value which the records carry: both as corporate memory, providing reference, audit and accountability for past events; and as a realisable asset for future uses, for example as a research base to support evidence-based policy-making. The issues here tend to be the 'softer' ones associated with more intangible and user benefits.

Maintaining authenticity/realising benefits

The second axis contrasts the need to maintain integrity and authenticity — conventional record-keeping — with the need to use the investment in electronic records management to realise benefits and the potential for innovation.

Placing these two axes together gives four quadrants, each with an identifiable set of issues. The quadrants are best read in the numbered order shown in the diagram, from (1) to (4).

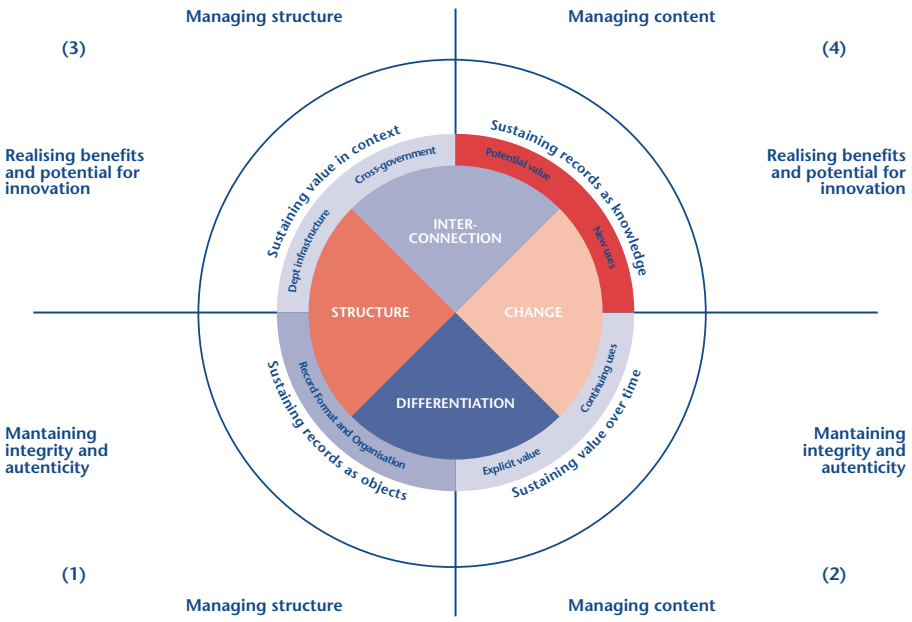


Figure 2: Model for sustainable electronic records: segmenting the issues

(1) Managing structure to maintain integrity and authenticity

Most of the practical activities in managing sustainability are carried out in this quadrant, which is primarily concerned with asking: what are the requirements for managing records across technology change in the medium term, to ensure they remain reliable and accessible? To answer this, we need to investigate:

- What are the objects (records) to be managed?
- In what technical environments are they produced?
- Who may do the managing?

Activities in (1) are carried out, at least in part, in order to deliver (2). Better understanding of these issues will help to size and scale the problem of physical management.

(2) Managing content to maintain integrity and authenticity

The reason for keeping records is to support business needs and responsibilities. From the point of view of the business unit and the end user, content rather than structure is most important. This involves ensuring the authenticity and currency of records, and that they are kept for as long as needed for all relevant purposes, but not longer — managing retention and disposal. This necessitates asking: what are the critical features of the business objectives, and what are the expected benefits for current use? We need to consider:

- In what business environments are records produced?
- In which ways will the records be used?
- For how long will they be required?

Taking (1) and (2) constitute conventional record-keeping requirements. A modernised approach to records management goes beyond 'keeping' to active management, and will need to consider potential uses and benefits, and integration with other corporate and government systems to achieve this. In this view, the activities in (1) and (2) are carried out to support (3) and potentially (4).

(3) Managing structure to realise benefits and potential for innovation

Electronic records and document-management systems are also operational information-management systems, and much of their business value stems from this. Electronic records need to be integrated with other information and knowledge-management resources in the same organisation, and increasingly related to those in other organisations across government, as we move into sharing and exchange of information and joint working. In this area, we need to consider:

- In which ways will the objects be used across government?
- How will the changes in government infrastructure affect the requirements for sustainability?

(4) Managing content to realise benefits and potential for innovation

Little can be said with certainty very far in advance in this area, since it deals with future requirements. The aim is to ensure that the necessary activities can be carried through when identified and needed, through sound management at other levels.

Current content management issues

The need to meet the modernising government target of managing records electronically by 2004 is leading all UK Government departments to consider the acquisition of solutions, which will deliver effective electronic records management. This in turn will lead to the implementation of corporate-wide classification systems for managing electronic objects which will have disposal values pre-assigned to the folders in which they are saved. Within two years, we will see a large growth in these systems and these in turn will generate significant additional sustainability requirements.

Each organisation already creates paper records with varying retention requirements. These requirements reflect statutory, business and operational needs and the periods of retention are generally standard across the whole of central government. Organisations are also working to identify their unstructured electronic record collections to determine how these are to be managed and accessed in the present and also to identify how long these collections need to be retained for business purposes by developing inventories.

Structured records such as enterprise-wide database systems are generally produced as a result of policy, administrative or operational processes. Currently, organisations have not necessarily formally identified the full retention requirement although this may be implicit in the project documentation or the specification, which led to a major database being designed and implemented.

Generic record categories

Records created across government can be divided into a number of categories. Some of these are generic, for example, personnel and finance records have a distinct business focus and reflect the actual purpose and remit of the organisation that created or received them.

The vast majority of the records created within the government fall into one of the 10 categories listed below:

- administrative records
- case records
- estates and accommodation management
- finance records
- health and safety
- personnel records
- policy records
- project records
- scientific and research records
- transaction records.

In addition to these should be added those records created or acquired by public inquiries and evidential records created or acquired through the management of justice.

It is estimated that the annual growth rate for all records in the UK Government is about 9 % per year; and found that about half of departments that may not have adequate information to implement retention schedules effectively. However, of those files that passed through the conventional 'first review' process (about three to five years after closure), some 82 % were destroyed and a further 15 % given a future destruction date. This suggested that a much higher proportion could be given a definitive destruction date by the point the file was closed — in the ERMS scenario by allocation of an electronic disposal schedule.

Each of these categories is explored in greater detail below.

Administrative records

Administrative records are produced in significant volumes and the number of personnel employed normally defines the overall volume. Generally, they have low retention values and are usually disposed of within one to seven years after the date of creation.

Case records

The retention period for case records as opposed to transaction records are usually defined by individuals' age and lifespan unless a statutory or a long-term operational requirement defines a period for their continued retention. For example:

- tax records concerning family trusts may be kept until the trust is wound up
- records relating to criminal cases — long retention up to 75 years.

Departments adopt different approaches to the retention of case files. Tax cases, for example, (approximately 26 million of them) are retained for between four and nine years after closure. Some case files are routinely destroyed after 15 years, others are kept until 18 months after the individual's death: an average life for each file of perhaps 50 years. From the scale, it can be seen that a small change in retention period, or an improved efficiency in applying disposal, can have a very significant effect on the potential costs of managing sustainability.

Estates

Estate records include a substantial amount of administrative type records with a low retention value. However, certain records are retained for much longer periods. Examples are given below:

- legal documents — estate title, leasehold and other contract documentation relating to the building and its surrounding land — retention period defined by occupancy;
- policy type records — these include surveys, evaluation reports, policy studies, etc. — retention varies between 10 to 25 years;
- information resource records — these documents are particularly relevant to the maintenance, repair and reconstruction of buildings. Many of these are kept for as long as the building remains in use.

Personnel records

The volume of personnel records normally reflects the numbers employed by the organisation. Retention periods vary but some records are kept for very long periods indeed. Common examples are listed below:

- documents bearing on pension entitlement should generally be kept for 72 years from date of birth or five years after the last action, whichever is later;
- appointment, staff appraisal or performance-management records are normally kept as separate sub-sets of personal files. If kept in annual sets, they are destroyed on a rolling basis;
- personnel security records are also kept as separate annual sub-sets of personal files. The Guide to Personal Security (GPS) suggests that these records should be kept until five years after leaving, if this is at the normal retirement age, or 10 years after leaving, if before normal retirement age;
- medical records are normally filed as a separate sub-set of individual personal files to allow for separate retention, if necessary. Departmental health and safety issues potentially affecting multiple members of staff, are documented on departments' administrative or subject files (although cross-references to these on personal/medical sub-files of the affected staff may also be required). In some instances where they relate for example to exposure to radiation, these may be kept indefinitely for research purposes.

Finance records

Financial records normally have a short working life. They include bank account records, expenditure records, ledger records, records of receipts and revenue. In most cases a period of between two to seven years is considered sufficient. They are unlikely to pose much of a technical challenge if held in electronic form as most if not all of the records will be deleted before the native software format and operating platform becomes redundant.

Health and safety and environmental records

Health and safety records can have a very long retention requirement partially because some issues involving exposure to radiation or contamination may have implications for safety decades after the event when the contamination occurred. In such cases where an identifiable employee is mentioned 40 years or 50 years may be the required retention period.

Some statutory requirements specify the maintenance of certain inspection records or the provision of certain notification forms. Inspection records include those relating to the maintenance of ventilation systems, the control of substances hazardous to health in the workplace, fire safety checks, and inspection of machinery and equipment.

Statutory notification forms include those required to advise enforcing authorities about certain operations or circumstances — for example notifications of accidents, diseases and dangerous occurrences — and would include certificates issued by enforcing authorities, such as fire certificates or licences for the storage of certain substances

Policy records

Policy records are normally retained for at least 25 years and in cases where the records relate, for example, to the development of primary legislation they are marked for permanent preservation.

The content of these records can range widely but essentially they will record the formulation of policy and legislation and, more selectively, its implementation and interpretation. These may include principal policy papers, for example include those leading to primary or subordinate legislation in which the department took the lead, submissions to ministers, and papers created in the course of preparing material for the Cabinet or a Cabinet committee, including all drafts.

They include papers created during the formulation, implementation or interpretation of major policies, for example, those which reflect significant changes in policy, they throw light on the main functions and programmes of the department, the interaction between the department and other areas of public or private life, or contain information, in particular statistics, returns or digests, not readily found or brought together elsewhere.

Project records

Project records generally tend to follow a distinct form especially where a recognised methodology such as Prince was adopted. Normally they can be retained according to their importance to the organisation for anything from 10 to 25 years.

Public inquiries

General practice is to retain all the records relating to the evidence considered by an inquiry, internal discussion papers, draft and final reports. In practice, although the inquiry that created the records has a finite life usually dictated by the publication date of the final report, the records are either passed to the Public Record Office for permanent preservation or to the institution that commissioned the inquiry until it is convenient or appropriate to hand them to the PRO for permanent retention. In most cases, the need for the creating or owning organisation to maintain the records is very short although the archival preservation requirement is long-term.

Scientific research

Records of the more important aspects of scientific, technological or medical research and development are normally retained as a long-term research resource for other scientific researchers. The format of the record will range from highly personalised and unstructured material such as scientists 'workbooks' in which records of experiments are recorded to formal reports and digests of research programmes, which are developed as specific products of a research programme. The content of the 'workbooks' is highly idiosyncratic whereas the reports and papers will adopt the recognised forms required by academe when preparing to publish results to a community of peers.

Retention periods can differ but scientists' workbooks are often the original source material when studying previous research. Most laboratories will treat them as part of their permanent library along with the published papers and reports. However, in comparison routine case files are often dispensed with after 10 years.

Transaction records

Transaction records as opposed to case records record specific events or transactions, which have a finite life. A simple example would be record of a contract allocated to a named contractor to commission a particular task. Large numbers of these contracts are let each year and once the work is complete and judged of an acceptable quality the file is closed and marked for destruction. Such files can be deleted after as little as 12 months. Depending on the nature of the transaction the retention period may vary but normally these do not exceed six years until the period allowed by the Statute of Limitations for any consequential litigation has expired.

While there is a degree of commonality between government organisations, there are also significant differences determined by the nature and work of the organisations.

Technical and structural management issues

Although all central government departments and agencies rely on the creation and transmission of electronic records to conduct their business to some extent, this conceals some degree of variance. There is a heavy reliance within government on e-mail, and documents transmitted as e-mail attachments, to transact business and record significant policy and procedural decisions. Currently, until departments implement their preferred electronic document and record management solutions the records held on the e-mail servers are unmanaged and there are no mechanisms to sustain e-mail records over time.

Other objects in the form of word-processed documents, spreadsheets, presentations, digital photographs, scanned images and small databases are often not subject to any form of corporate control. Some departments are currently developing inventories to identify these collections and have plans to evaluate their importance to the business and to identify sustainability requirements.

Some thought however, has gone into planning for the future redundancy of large database systems on which organisations rely for their core business. Because these involve large capital invest-

ment, and the risk of future information loss would have serious consequences on the ability of these institutions to discharge their obligations, the design parameters of these databases have to take into account the need to migrate to new platforms. Although these systems are often heavily customised or bespoke, there is usually sufficient funding to support successful transitions, as the core business of the institution would otherwise be imperilled.

Management roles and models

Management of IT within government already varies considerably within departments. Sustainable solutions and maintenance services will vary in size and the manner in which they are delivered. Functional requirements will need to address this issue and define the controls, which will need to be developed. The current models are:

- in-house management and procurement supported by external expert consultancy as required;
- data storage has been outsourced to a data warehouse;
- partial outsourcing where key services have been outsourced to an application service provider;
- fully out-sourced service where an external IT provider or lead contractor provides an integrated service usually under a PFI or PPP contract.

In all cases it will be necessary to identify what are the key management roles and responsibilities, and the key skills, which must be retained and nurtured within the organisation in order to support effective maintenance strategies. This will need a formal reviewing mechanism to confirm the viability of the current skills set.

Departmental and government infrastructure issues

Government in the United Kingdom is in a period of rapid modernisation. The movement towards e-government is now seen as a process of business change, rather than simply the implementation of new technology. A primary aim of e-government is to increase the level of integration between government activities and services, both in presentation to the citizen, and in terms of efficiencies gained from modernisation. The logic of e-business change folds back into internal processes within departments, requiring changes in existing practices in line with modernisation; these changes will affect the way in which electronic records are produced, managed and used.

Emerging cross-government standards, which sustainable electronic records management will have to take account of include:

- interoperability standards defined by the government interoperability framework (e-GIF) for browser access, which determine allowable formats and set a strategic direction towards IP- and XML-based systems;
- initial metadata standards for resource discovery also set by the e-GIF;
- security and authentication standards.

In addition, the PRO has already established generic standards for current electronic records management, and work is ongoing on establishing a government standard for records management metadata. Recent discussions have set out the case for establishment of data-sharing protocols.

While these technical and management standards will help to build more integrated 'hard' systems, there remains much to be established in ensuring consistency and coherence at the level of 'soft' issues concerning relationships and flows between information and record sets across government. The move towards 'joined-up' government, and more sharing of records and data within sectors, pushes us towards a whole-of-government approach to the management of information and records: for example, to ensure that records which are a shared resource between two organisations are maintained as a consistent and accessible whole for the longest length of time required by either stakeholder — whether physically managed by one or both organisations. At present, records are largely managed at the departmental level only.

There are early examples of a new generation of cross-cutting systems and services. As these develop, there will be an increasing disjunction between an information and records mapping of such systems and services across the government information landscape, and the organisational struc-

ture of government. At present, the primary assumption is that, for practical purposes, records are owned and managed by a single department or agency; in the future this will be less and less the case. In line, sustainability for electronic records will have to be implemented more at the whole-of-government level, as well as at the level of the individual organisation.

Potential strategy options

This section outlines some key differences in the balance of costs for sustainability in the move from a paper to an electronic environment, and some implications of this for evaluation, and identifies some of the possible mechanisms which may be available in both the commercial and government sectors.

Cost-balance of sustainability

The balance of costs and values, which is likely to develop for the medium-term management of electronic records, is different from the pattern, which typically applies to paper records. This is summarised in the following table:

	Paper	Electronic
Storage	Medium	Low
Maintenance	Low	High
Disposal	High	Low ([1])/high ([2])

([1]) By retention and disposal schedules.
([2]) By individual review.

Storage costs for electronic information on digital media are considerably lower than the cost of physical space which paper requires. Conversely the maintenance costs are low for paper — as long as storage conditions are adequate, little needs to be done to maintain the information content — but high for electronic material: the obsolescence effect of technological change requires fairly constant attention to both media and file format renewal, probably through some form of migration strategy.

Because migration is, for the foreseeable future, a costly process that will involve the application of scarce human expertise, there is a high per unit cost for each record or set of records. There will be a pressure to migrate only that material which has a demonstrable continuing value, and much less temptation to migrate material 'just in case'.

This discussion suggests that there will be a pressure to reduce the amount of material migrated through time, and to develop more efficient methods of identifying the records of continuing value, which qualify. There will also, of course, be a need to permanently destroy records for which there is no further business need and which have not been selected for permanent preservation. The obvious means of such active management is by the use of disposal schedules: pre-determined business rules which are applied to categories of records, and which can be semi-automated as a management process. This is a relatively cheap per unit method of disposal.

While it is very unlikely that the need for review will be completely eliminated, there will be pressure to minimise it where possible; and to consider alternative methods for carrying it out: where review by content is not mandated, then review by metadata or on a risk-management basis is emerging.

From a departmental perspective, a rolling review process will focus on continuing business need, and will be cognisant of the need to reduce migration costs as an annual budget item. Where a review determines that there is no continuing business need for the department, but that there is a requirement to retain for permanent preservation for historical value, then these economics suggest that there will be pressure to transfer custody to the appropriate place of deposit as soon as possible.

Potential market mechanisms

The first essential point to note from the above discussion is that, because of the explicit and regular costs of maintaining electronic records over time, there will be a continuing pressure for cost

reduction by agile management of the whole life-cycle, through disposal schedules, and review where necessary. Feasible strategies for long-term maintenance will be heavily influenced by two factors in combination:

- the business rules which determine the length of retention, and the eventual disposal action (destruction or transfer)
- the frequency of use for business purposes by stakeholding organisations (both creating department, and other organisations which may use or share the record).

Some types of record may need to be kept for long periods, but will see very infrequent use after the first few years of their life: for example, personnel and military service records must be retained for 70–100 years, but the vast majority will not be referred to once the staff member is no longer in employment with the organisation. Others may be kept for shorter periods, but will see much more frequent use over the period: for example, policy papers on a continuing topic, which are utilised for research and analysis.

Different sustainability strategies will be appropriate for different categories of records. Some of these categories will be present in most or all organisations; others will be particular to a few departments or agencies. The requirement to be cost-efficient makes it unlikely that a single strategy will be optimal for any one organisation.

The second essential point to note is that, for most organisations, there will be a need, in management systems and products supporting sustainability, to provide a mechanism which can implement managed disposal of electronic records, as business rules determine they no longer need to be retained (rather than continue to bear the cost of maintaining them), on a continual basis. This requirement will apply in principle, even though the point of disposal is many years into the future. An implication is that simply archiving to a non-rewriteable media, such as optical disk, probably does not support the level of granularity needed to manage discreet categories.

At the present time, the market strength for electronic records management is focused on systems, which are capable of capturing and storing records in organised structures. The market for products, which support the technical and procedural aspects of sustainability, is very immature, and only general predictions can be made about the ways in which it might develop. Some broad options are sketched out here, but these will require further elaboration.

Departmental archive systems

Some large departments, or those with particular needs, may establish an internal archive management system for sustaining electronic records within their own organisation: for example, the Foreign and Commonwealth Office have established the Minerva archive for maintaining electronic records between three and 30 years of age. It seems unlikely, however, that it will be a cost-effective option for most organisations to have two separate systems: one for current records capture and management, and another for medium term archival management.

Development of current EDRM systems

One option is to encourage the development of electronic document and records management (EDRM) systems to incorporate the functionality required to sustain records over the longer period. Such systems currently contain mechanisms for managing retention and disposal, and controlling access, including management of personal data and closed records. If a department has already made an investment in this technology, it may be more effective to leverage this functionality by requiring its extension to more sophisticated staged life-cycle management than is currently offered: towards integrated and sustainable EDRM.

Commercial archive systems

Existing commercial providers of long-term storage for paper records, such as 'Iron mountain', are known to be developing equivalent services for electronic records. This is probably an attractive option for long-term storage of certain categories (for instance, the personnel example) but many issues of access, security, and standards compliance would need to be addressed, as well as issues related to costs and contractual length and responsibility.

Potentially, commercial IT service providers, such as EDS, may wish to develop a long-term archive service for their government customers. There is no real evidence of such an intention at the present time; most are still struggling with the need to provide current capture and management facilities.

Application service providers

An ASP provides a full service, usually on some form of 'pay per use' basis, which does not require the client to implement and maintain the relevant software products: data and programme are called across a network connection when required, from the service provider. In this arrangement, records would be stored and managed remotely from the owning organisation by the service provider, but remain accessible to the department by, say, browsing to a secure web site. At least one currently active EDRM software supplier in the United Kingdom is known to be considering the development of a market offering of this type.

An ASP service arrangement may take one of two forms:

- currently active records (e.g. 0–3 years) managed by site-specific EDRM system, with archival (3+ years) records managed on a remote ASP system;
- all records, current, semi-active and inactive, managed by remote ASP system (i.e. newly created captured directly to a remote ASP system.

An ASP service arrangement may potentially be shared between a number of organisations, with the capability of bringing together cross-cutting records across a government sector.

Cross-government integrated electronic records management

An alternative to the commercial ASP option is that government should determine to provide this service for itself, for any mixture of current and non-current records across a range of departments and agencies. The primary argument for this approach — which may still aim to transfer some risk to a technical system service provider — is that it would locate control of integration, standards and implementation with government itself, while obtaining economies of scale and scope, and providing a systemic underpinning to whole-of-government innovation and business change. There is, however, little precedent for this approach. Rather than attempting to provide such a service directly, government might aim to coordinate the setting of, and firmly establish the use of, a set of wide-ranging and fully integrated standards: for technical issues, information management issues (for example, data sharing, government-wide retention and disposal rules, minimum implementation standards), coordination of procedural and business rules. This approach stresses the integration of records management with other areas of information management. Standards are, however, easier to set than to ensure compliance with; the practical means by which a service is provided would still need to be determined.

Conclusion

In order to address the issues discussed in this paper we recognise we need to build upon our existing functional requirements for electronic-records management by developing generic requirements for sustainability to ensure required records continue to exist in a reliable form for as long as they are needed. The process of defining these requirements will necessitate consultation between PRO and other UK Government departments and agencies to ensure the requirements address the cross government business need for sustainability.

We also recognise that where the period of sustainability is very long, for example 75 years, some of the technical requirements for the maintenance of these records will have implications for the design of digital archives, which aim to secure the permanent preservation of selected records.

We will also need to consider the definition of a metadata standard for sustainability, which will complement the forthcoming current records management metadata standard. We would hope to then develop appropriate XML schemas for adoption and incorporation within the United Kingdom e-government metadata standard (e-GMS) and we will need to closely liaise with our colleagues in the office of the e-Envoy.

We are also conscious of our need to reference existing and developing European and international standards and guidance such as MoReq and the outputs of the InterPARES projects and we will be monitoring current and planned developments in order to ensure the sustainability requirements are both credible and robust. It is our intention that as our products are developed that these will be published on our web site either as drafts for comment or as finalised documents.

Mantenimiento de documentos electrónicos para uso interno

Richard Blake, Stephen Harries

En el Gobierno británico, la mayoría de los ministerios y organismos (si no todos) ya crean documentos electrónicos, que pueden o no, pero probablemente deberían, constituir archivos. El objetivo de modernización de los Gobiernos, consistente en lograr una gestión electrónica de los documentos antes de 2004, requiere la evaluación/migración de los documentos electrónicos existentes que tienen un valor de cuasiarchivos en un medio bajo control. La instauración de sistemas de gestión de archivos y documentos electrónicos (EDRM), de aquí a 2004 y en lo sucesivo, permitirá la introducción organizada de documentos internos en cualquier organización, y a partir de esta fecha se supone que las organizaciones gubernamentales utilizarán los archivos electrónicos para su memoria interna oficial.

Desde el punto de vista de la administración, la justificación interna de la gestión de archivos y documentos electrónicos radica fundamentalmente en los beneficios a medio plazo que pueden obtenerse a través del acceso continuo a información fiable y auténtica, lo que permitirá un uso más eficaz, un acceso mejor y más rápido y la innovación en los métodos de trabajo, gracias a la integración continua con otros sistemas de información y actividad internos. Estos beneficios solamente se mantendrán a medio plazo si el valor de la información se mantiene mediante una buena gestión de su contenido. Mientras que parte de esta gestión se refiere al almacenamiento seguro y al control del vehículo que transporta el contenido, otra parte, destinada a mantener la autenticidad, se refiere al mantenimiento de un acceso práctico al propio contenido y a su utilidad para los fines de la actividad.

El cambio tecnológico tenderá a depreciar el valor de los objetos electrónicos como vehículos: por tanto, la relación entre el cambio tecnológico y el documento electrónico deberá gestionarse a fin de impedir la pérdida de acceso, integridad y transparencia. Sin embargo, para mantener el valor interno del documento, la gestión de esta relación debe considerarse en el contexto más amplio del desarrollo y del cambio organizativo, garantizando una buena adecuación entre los cuatro aspectos: el documento, la tecnología, los usuarios y la organización. Si los archivos internos y la gestión de la tecnología (y las responsabilidades internas sobrevenidas por falta de gestión de los archivos) se conciben en una dimensión horizontal en que la necesidad de estabilidad es el aspecto clave, la gestión de contenidos y los beneficios para los usuarios se encuentran en una dimensión vertical que exige flexibilidad y cambio. El reto es mantener un equilibrio viable entre estas dos dimensiones, de manera que se optimicen ambas.

Existe una necesidad a largo plazo de gestionar los archivos electrónicos a efectos de la *durabilidad*: acceso duradero para el uso interno; integridad y transparencia duraderas como requisito interno; y obtención duradera de beneficios como lógica de la actividad. Esta perspectiva de la *durabilidad* difiere bastante de la conservación permanente, aunque hay muchas coincidencias relativas a cuestiones específicas. La durabilidad se refiere no solo a la conservación del vehículo físico (que por supuesto presenta dificultades en el medio electrónico), sino también a la conservación del documento a fin de obtener los beneficios previstos por una inversión inicial en el cambio organizativo y tecnológico necesario para capturar y almacenar los documentos. Estos dos aspectos, durabilidad *de* y durabilidad *para,* son

aspectos complementarios y deben considerarse unidos en un conjunto coherente: ninguno es suficiente por sí mismo.

La exposición aborda los principales problemas y presenta un primer enfoque de los factores de éxito críticos, centrándose en los siguientes aspectos:

- cuestiones de gestión técnica y estructural: los medios de gestión;
- cuestiones actuales de la gestión de contenidos: los archivos y sus aplicaciones;
- cuestiones de infraestructura ministerial y administrativa.

Se subrayan algunas diferencias significativas entre la economía de la gestión de los archivos en papel y de los archivos electrónicos, y se establecen diversas estrategias posibles. Los mecanismos de mercado en este ámbito no son por ahora lo suficientemente maduros como para establecer una única estrategia clara, y en todo caso es probable que se requiera un conjunto de estrategias para diversos segmentos del problema y del mercado.

El documento realiza una segmentación inicial de las posibles categorías de archivos y de las posibles opciones de una administración a la hora de dar respuestas; y examina los problemas y las acciones dirigidas a abordarlos, que derivan del movimiento hacia una administración integrada conjunta. El documento también aborda algunas áreas específicas de acción en las que una institución nacional puede intervenir en el desarrollo de esta cuestión, y las tendencias del mercado, que aunque no sean capaces de determinar completamente la dirección, contribuirán a perfilar los acontecimientos de forma beneficiosa para la administración en conjunto.

Stützung elektronischer Aufzeichnungen für Geschäftszwecke

Richard Blake, Stephen Harries

So gut wie alle staatlichen Behörden des Vereinigten Königreichs erstellen bereits elektronische Dokumente, bei denen es sich möglicherweise um Archivgut handelt oder handeln dürfte. Angesichts der für den gesamten Staatsapparat geltenden Modernisierungsvorgabe, bis 2004 eine elektronische Archivgutverwaltung einzuführen, sind die Behörden angehalten, vorhandene elektronische Dokumente von dauerhaftem Wert entsprechend zu bewerten und sie als (Quasi-)Archivgut in eine verwaltete Umgebung zu migrieren. Mit der Einführung der elektronischen Dokumenten- und Archivgutverwaltung (EDRM) bis 2004 wird die organisierte Erfassung von dienstlichen Dokumenten im gesamten Staatsapparat unterstützt, und von den entsprechend eingerichteten staatlichen Stellen wird erwartet, dass sie ihr gesamtes Corporate Memory auf elektronische Unterlagen aufbauen.

Aus behördlicher Sicht liegt die betriebswirtschaftliche Rechtfertigung der EDRM vor allem in den mittelfristigen Nutzeffekten, die sich aus dem ständigen Zugriff auf zuverlässige, authentische Informationen erzielen lassen. Damit wird eine effektivere Nutzung der Informationen sowie ein besserer und schnellerer Zugang zu ihnen gefördert, und aufgrund der ständigen Integration mit anderen behördlichen Informations- und Line-of-Business-Systemen entstehen innovative Arbeitsweisen. Diese Nutzeffekte lassen sich nur dann mittel-

fristig aufrechterhalten, wenn der Wert der Informationen durch gute Verwaltung des Informationsinhalts bewahrt wird. Während dies zum Teil durch sichere Speicherung und Kontrolle des Inhaltsträgers geschieht, so dass die Authentizität gesichert ist, spielt auch die Aufrechterhaltung des praktischen Zugangs zum Inhalt selbst und dessen Nutzbarkeit für dienstliche Zwecke eine Rolle.

Im Zuge des Technologiewandels kommt es natürlich mitunter zu einer Abwertung der elektronischen Objekte als simple Artefakte. Um zu verhindern, dass Zugang, Integrität und Rückverfolgbarkeit verloren gehen, muss daher für ein gutes Verhältnis zwischen technologischen Veränderungen und dem elektronischen Archivgut gesorgt werden. Um den dienstlichen Wert zu bewahren, muss jedoch dieses durch entsprechende Archivgutverwaltung gepflegte Verhältnis im größeren Zusammenhang organisatorischer Veränderungen und Entwicklungen betrachtet werden, nämlich als Garant für die Balance zwischen den vier Aspekten Archivgut und Technologie, Benutzer und Organisation. Wenn man sich die behördliche Archivgut- und Technologieverwaltung (und die behördliche Haftung aufgrund einer fehlenden Archivgutverwaltung) als horizontale Dimension mit dem Stabilitätserfordernis als Hauptaspekt vorstellt, liegen Inhaltsmanagement und Benutzervorteile in einer diese Dimension schneidenden vertikalen Dimension, die Flexibilität und Wandel verlangt. Hier geht es darum, Ausgewogenheit zwischen diesen beiden Dimensionen so zu erzielen, dass beide optimiert werden.

Mittel- bis langfristig sind elektronische Unterlagen aus Gründen der *Nachhaltigkeit* zu verwalten, d. h. für den nachhaltig gewährleisteten Zugang zu dienstlichen Zwecken, für die nachhaltig gewährleistete Integrität und Rückverfolgbarkeit und für die Realisierung nachhaltig gewährleisteter Nutzeffekte aufgrund dienstlicher Vorgaben. Diese Sichtweise der *Nachhaltigkeit* unterscheidet sich weitgehend von der archivischen Konservierung, obgleich es beim Umgang mit konkreten Fragen durchaus viele Überschneidungen gibt. Bei der Nachhaltigkeit geht es nicht nur um die Konservierung physischer Artefakte (was natürlich in der elektronischen Umgebung Probleme aufwirft), sondern auch um die Erhaltung der Unterlagen, also des Archivguts, zur Erzielung der Nutzeffekte, die von einer Investition in die zur Erfassung und Aufbewahrung des Archivguts überhaupt notwendigen Technologien und organisatorischen Veränderungen ausgehen. Diese beiden Aspekte – Nachhaltigkeit des Archivguts und Nachhaltigkeit *für* einen bestimmten Zweck – ergänzen einander und müssen als einheitliches Ganzes betrachtet werden. Ein Aspekt allein reicht nicht aus.

Es werden die vordringlichen Fragen sowie einige entscheidende Erfolgsfaktoren umrissen und erörtert, und zwar mit folgenden Schwerpunkten:

- technische und strukturelle Managementfragen: Mittel des Managements;
- aktuelle Fragen des Inhaltsmanagement: die Unterlagen selbst und deren Anwendungsbereiche;
- behördliche und behördenübergreifende Infrastrukturfragen.

Es werden eine Reihe von signifikanten Unterschieden bei den ökonomischen Aspekten des nachhaltigen Managements von papiergebundenen und elektronischen Unterlagen erläutert sowie einige potentielle Strategieoptionen herausgestellt. Die Marktmechanismen für diesen Bereich sind noch zu unausgereift, um eine eindeutige zukunftsorientierte Strategie auszuweisen, und es ist davon auszugehen, dass ein Bündel von Strategien für die jeweiligen Problem- und Marktsegmente benötigt wird.

Der Beitrag nimmt eine erste Einteilung der potentiellen Kategorien von Unterlagen und der Möglichkeiten einer staatlichen Stelle zur Erarbeitung von Lösungen vor und untersucht die Probleme und entsprechenden Lösungsansätze, wie sie sich aus der Entwicklung hin zu einer Reduzierung des Ressortprinzips („joined-up government") ergeben. Außerdem werden spezifische Aktionsfelder herausgearbeitet, auf denen eine nationale Einrichtung vielleicht in die Entwicklung dieses Problembereichs und der Markttrends eingreifen kann, was – obgleich sich keine grundsätzliche Richtungsbestimmung geben lässt – dazu beiträgt, die Entwicklungen in einer für den Staat insgesamt günstigen Weise zu beeinflussen.

Maintien des documents électroniques à des fins de gestion

Richard Blake, Stephen Harries

Dans l'administration britannique, la plupart des ministères et des organismes (sinon la totalité) créent d'ores et déjà des documents électroniques, qui peuvent ou non constituer des archives — mais le seront probablement. L'objectif de modernisation de l'administration, qui est de mettre en place un système de gestion d'archives électroniques d'ici à 2004, implique d'évaluer et de transférer les documents électroniques actuels dotés d'une valeur permanente comme des (quasi-) archives, dans un environnement sous contrôle. La mise en place de systèmes de gestion de documents et d'archives électroniques, d'ici à 2004 et ultérieurement, permettra la saisie organisée des documents internes dans toute l'organisation; à compter de cette date, les administrations en conformité seront supposées ne plus utiliser que des archives électroniques pour leur mémoire interne officielle.

Du point de vue de l'administration, la justification interne des systèmes de gestion de documents et d'archives électroniques réside en premier lieu dans les avantages à moyen terme qui découleront d'un accès permanent à des informations fiables et authentiques, ce qui permettra une utilisation plus efficace et une consultation plus facile et plus rapide de ces informations, ainsi que dans l'innovation dans les modalités de travail, grâce à une intégration continuelle avec les autres systèmes d'information et d'activité internes. Ces avantages ne se pérenniseront à moyen terme que si l'information conserve toute sa valeur, à travers une bonne gestion de son contenu. Si cette condition réside en partie dans un stockage sécurisé et dans le contrôle des produits qui véhiculent le contenu, afin d'en préserver l'authenticité, il y a lieu également de maintenir un accès pratique au contenu lui-même et à son utilisabilité, aux fins de l'activité.

Bien sûr, l'évolution technologique aura tendance à déprécier la valeur des objets électroniques en tant que créations humaines: aussi le lien entre évolution technologique et document électronique doit-il être géré pour éviter toute perte d'accès, d'intégrité et de transparence. Toutefois, pour que la valeur interne du document soit conservée, la gestion de ce lien doit être vue dans le contexte plus général du changement organisationnel et du développement, en tant qu'il garantit un bon ajustement entre quatre facteurs: document, technologie, chercheurs et organisation. Si les archives internes et la gestion technologique (et les responsabilités internes encourues du fait de lacunes dans la gestion des archives) sont représentées le long d'un axe horizontal sur lequel la nécessité de stabilité constitue l'aspect majeur, la gestion des contenus et les avantages pour les chercheurs se situent sur un axe vertical sécant, qui appelle flexibilité et changement. Toute la difficulté est de préserver un équilibre durable entre ces deux dimensions, de façon que toutes deux soient optimisées.

Il importe à moyen et long terme de gérer des archives électroniques en vue de leur *durabilité:* accès durable pour un usage interne, intégrité et transparence durables en tant qu'obligation interne, concrétisation durable des avantages en tant que raison d'être de l'activité. Cette perspective de la *durabilité* est très différente de celle de la conservation dans des archives, encore qu'il existe certainement plusieurs recoupements dans le traitement de ces questions spécifiques. La durabilité se préoccupe non seulement de la conservation physique du produit (laquelle soulève évidemment des difficultés dans un environnement électronique), mais aussi du maintien du document pour pouvoir concrétiser les avantages escomptés d'un investissement initial dans la technologie et le changement organisationnel nécessaires à la saisie et au stockage du document en première instance. Ces deux aspects, la durabilité *de* et la durabilité *pour,* sont complémentaires et doivent être considérés comme un tout logique: aucun ne se suffit en lui-même.

L'exposé aborde les principaux problèmes et présente une première approche des facteurs de réussite critiques, en abordant plus spécialement:

- les questions de gestion technique et structurelle: les moyens de gestion,
- les questions actuelles de gestion du contenu: les archives elles-mêmes et leur finalité,
- les questions d'infrastructure ministérielle et administrative.

Les différences significatives entre l'économie de la gestion des archives papier et celle des archives électroniques sont soulignées, et un certain nombre d'options stratégiques possibles sont proposées. Les mécanismes du marché dans ce domaine n'ont pas encore atteint un stade de maturité permettant d'identifier une stratégie dynamique claire et nette, et il est probable qu'en tout état de cause tout un éventail de stratégies seront nécessaires pour divers segments du problème et du marché.

L'exposé se livre à une première segmentation des catégories d'archives possibles et des options potentielles à la disposition d'une administration pour élaborer des réponses. Il examine les problèmes et les mesures pour y remédier, lesquelles procèdent d'une évolution vers une administration «engagée» intégrée. L'exposé identifie également certains domaines d'action précis dans lesquels une institution nationale peut intervenir. Bien qu'elles ne permettent pas de déterminer tout à fait une direction, les tendances du marché contribueront à façonner les développements dans un sens qui soit bénéfique à l'administration tout entière.

Practical experiences of the digital preservation testbed

Jacqueline Slats

Jacqueline Slats

After her study in information management, Jacqueline worked for seven years at the computer centre of the Ministry of Transport, Public Works and Water Management. In 1994, she joined the Dutch State Archive Service, where she was responsible for different information technology projects. Since October 2000, she is the programme manager of the digital preservation testbed, which is sponsored by the Dutch State Archive Service and the Ministry of the Interior and Kingdom Relations.

The digital preservation testbed is part of the non-profit organisation ICTU. The ICTU is the Dutch organisation for ICT and government. The ICTU's goal is to contribute to the structural development of e-government. This will result in improving the work processes of government organisations, their service to the community and interaction with the citizens.

The Dutch e-government house

Government institutions, such as ministries, design the policies in the area of e-government, and the ICTU translates these policies into projects. Together, these projects form what we call the e-Government house or ELO-house. In many cases, more than one institution is involved in a single project. They are the principals in the projects and retain control concerning the focus of the project. In the case of the digital preservation testbed the principals are the Ministry of the Interior, Jan Lintsen and the Dutch National Archives, Maarten van Boven. Together with public key infrastructure, digital longevity is the fundament of the ELO-house.

Digital longevity

Under the umbrella of digital longevity, we have several programmes like record-keeping system, quality assurance, testbed etc. It is the objective of digital longevity securing the accessibility of reliable government information, the objective of the digital preservation testbed is securing the sustained accessibility of reliable government information.

According to Dutch law and regulations, the transfer of archival records takes place after 20 years, in a 'good, ordered and accessible state'. Therefore, the target group of the digital preservation testbed is not only archival organisations, but also the whole government.

The current Dutch Cabinet aims to carry out 25 % of its transactions between the government and its citizens through digital means by 2002. Because of this, there is currently a great deal of work going on to develop strategies, methods, techniques and tools to handle the digital produce of the government in a responsible way.

Long-term digital preservation

The most important problem concerning the preservation of authentic digital records is technological obsolescence. Technological change is increasing exponentially. This brings up many questions, such as what to do with files that were made with old hard-and software, which cannot be used anymore? Unless action is taken now, there is no guarantee that current files can be read in the future with future technologies.

The digital preservation testbed is researching three different approaches to long-term digital preservation: migration, emulation and XML. Not only will the effectiveness of each approach be evaluated, but also their limits, costs and application potential.

Experiments are taking place on text documents, spreadsheets, e-mails and databases of different size, complexity and nature.

Experiments

The digital preservation testbed is carrying out experiments according to pre-defined problem-solving research questions to establish the best preservation approach or combination of approaches. The experiment process started with these basic research questions. And each experiment raises new questions.

Not only to control the project, but also to run experiments in a controlled environment, we developed a 12-step experiment process. Here, we make also explicit, mostly by desk research of available publications, if a record type is excluded from a certain preservation approach. These steps are all fully documented in the experiment database of the testbed. Records are monitored during experiments to establish whether (and how) a specific method is suitable for long-term preservation.

This approach requires a multidisciplinary team. The testbed team consists of ICT-expertise, record managers, archivists, national and international experts, etc. Very valuable is the evaluation feedback group, which consists of archivists from various institutions, e.g. the Dutch National Archives, the Archival Inspection, Graphic Industries, Tax Services, etc. The governmental institutions that provide us with copies of records are participating in the team during the experiments.

Experiments on migration

There are many different definitions of migration. Testbed defines migration as the conversion of records from one hardware and/or software environment to another. Migration is currently the most common preservation strategy for digital records, but not always used in a responsible way: when new versions arrive, documents are simply updated into the new versions.

Testbed experimented with migration of text documents:

MS Word 95, 97, 2000, 2002 step by step, and for example directly from MS Word 95 into 2002 and Conversion to Adobe Acrobat PDF 1.2, 1.3, 1.4. Experiments from WordPerfect into MS Word are now taking place.

The experiments give good results, if the documents are created in a responsible way, (e.g. don't use automated date fields) and captured, and if the migration is well prepared. What is remarkable is that migration from MS Word 95 directly into MS Word 2002 gives better results than migration step by step. We don't have a clarification yet as to why Adobe puts the last character on the next line

Still, there remain a few disadvantages of migration. Each record must be migrated every few years, this is only feasible if the process is automated. It still requires manual checking of the results. And eventually, changes to the format of the record can lead to information loss, thereby compromising the records' authenticity.

Experiments on emulation

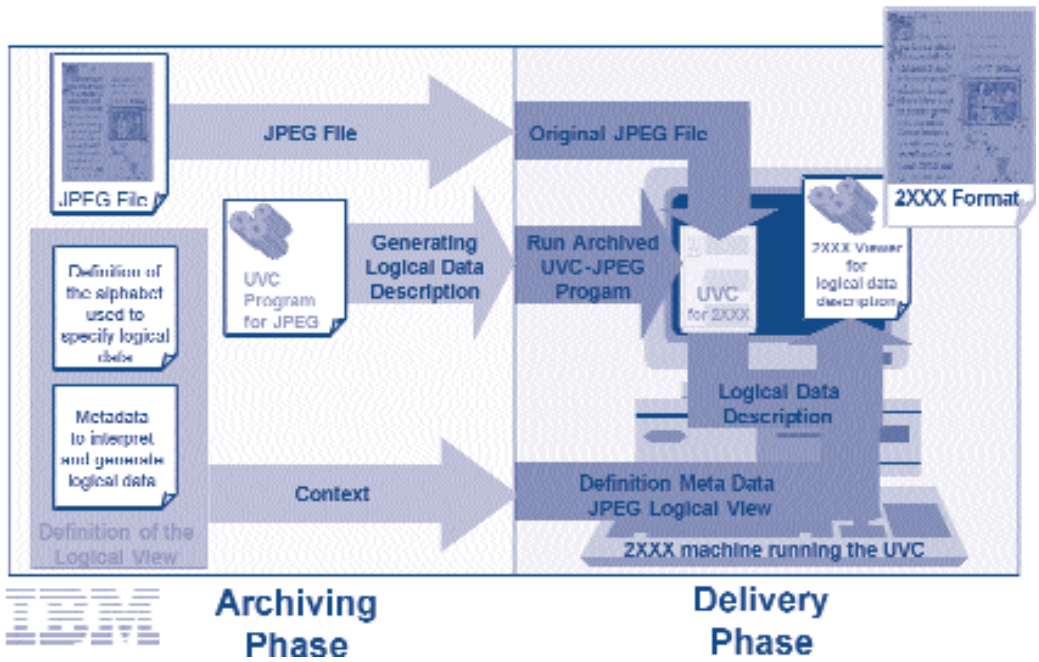
The universal virtual computer-based methodology makes a distinction between preserving data and preserving the behaviour of a programme. For data, it implements a conversion programme able to decode the original form of the data into a logical format that will be much easier to understand in the future. This conversion programme is written in 2000 (for a UVC machine). It can be executed in 2050, on an emulator of the UVC on the 2050 machine. For programmes, the UVC-based methodology will rely on an emulator of the 2000 machine (for a UVC machine), written in 2000, and an emulator of the UVC on the 2050 machine. It clearly differs from the emulation method proposed by for instance Jeff Rothenberg, in that it does not require writing in the future an emulator of a real machine of the past.

With the UVC approach we have a solution that can be applied to support multiple preservation strategies both data and programme preservation. The UVC architecture relies on concepts that have existed since the beginning of the computer era: memory, registers, and a set of low-level instructions. The fact that the computer is virtual and that performance is of secondary importance allows for a simpler, more logical, maybe less optimised, design. This will guarantee its durability along waves of technology changes.

Because the UVC instruction set is so simple, it is relatively straightforward to write a UVC emulator for any given computer. In the context of long-term preservation of digital data, initiated with-

in IBM Research, they considered an approach that relies only partially on emulation. The approach is applicable to digital object types that do not need to maintain the functionality of the application(s) that were used initially to create or manipulate the objects. IBM refers to this approach as data preservation. For data preservation, we propose to save, with the data, a programme that can extract the data from the bit stream and return the information to the caller in an easy to understand, technology-independent way, so that it may be exported to a new system.

Data preservation is the first and simplest mode of operation of the UVC approach



For example the preservation of JPEG files. The UVC JPEG programme is written for a universal virtual computer (UVC). This UVC will be stable across technology changes. All that is needed in the future for executing the UVC JPEG programme is an interpreter (an emulator) of the UVC architecture.

The execution of the UVC JPEG programme in the future will return the data with additional information, according to a logical view — defined by a logical view description or schema, which is also archived. This complete data preservation approach enables organisations to always retrieve a technology independent description (logical data description) of any JPEG file in the future with the aid of three components: UVC JPEG programme, UVC, and the archived logical data definition for JPEG.

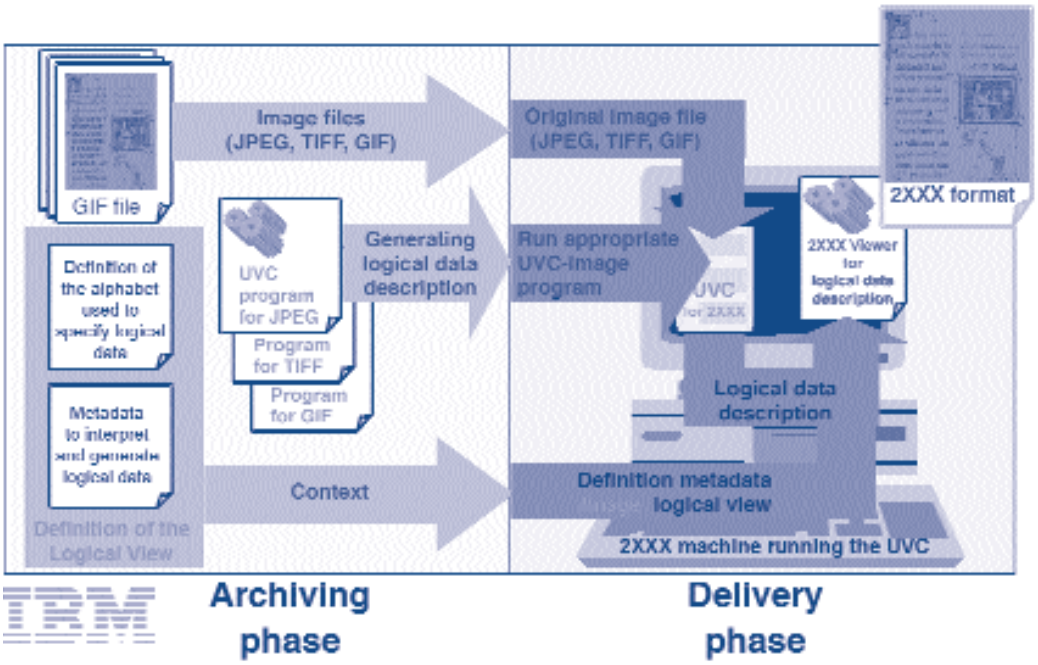
These three archived components enable any person in 2000 to regenerate the information in the current environment, using new data formats. Writing a viewer programme that invokes the emulator, runs the UVC JPEG programme, and processes the returned data as desired. The first proof of concept with this approach was very successful.

In the next step, numerous digital object classes, like images, may be translated to the same logical view, eliminating the need for a future client to implement separate viewers for each original format.

The data preservation approach is not limited to static information types. Sound and video can also be dealt with by data preservation. Even programme-dependent applications like relational databases could in essence be described by a logical data description; basically any information that can be described logically in a static way is a candidate for data preservation.

When the approach reaches critical mass probably a small number of logical data descriptions will remain, like: general text, images, sound, video etc. However, the beauty of the approach is the fact that an organisation doesn't have to wait for this standardisation process to complete.

Multiple digital object types can produce the same logical data reducing the need for multiple viewers



Everybody at this stage can define their own logical data definitions and support it with the aid of the UVC.

Experiments on XML

We all know XML as a format or mark-up language. But because of its characteristics and because it is an open standard, it is promising to use XML as a preservation strategy.

Testbed experimented with XML and different types of e-mail: MS Outlook 98 and 2000, K-mail, Eudora and Hotmail.

The experiments give good results, but development of templates for end-users for internal and external use is needed to create, capture and store the e-mail with required metadata, which makes it possible to interpret the e-mail in the right context. The testbed will develop these templates for MS Outlook within six weeks.

Further experiments

New experiments expected in 2002 are the migration of spreadsheets, conversion of spreadsheets and databases into XML and a proof of concept with the UVC for text documents and spreadsheets.

Products

Eventually at the end of 2003 the testbed project will provide:

- advice on how to deal with current digital records;
- recommendations for an appropriate preservation strategy or a combination of strategies;
- functional requirements for a preservation function;
- cost models of the various preservation strategies;
- a decision model for preservation strategy;
- recommendations concerning guidelines and regulations.

Experimentación práctica del banco de pruebas de conservación digital

Jacqueline Slats

Se ha dicho que la pasada década será la época peor documentada del siglo XX. La información se registrará cada vez más en forma digital, sin que haya métodos fiables para conservar esta información a largo plazo. Así pues, es cuestionable que las generaciones futuras puedan reconstruir lo que sucedió en nuestro tiempo.

El actual gobierno neerlandés aspira a realizar el 25 % de sus transacciones entre la administración y sus ciudadanos por medios digitales antes de 2002. A causa de esto, actualmente hay mucho trabajo consistente en desarrollar estrategias, métodos, técnicas y herramientas para manejar la producción digital de la administración de forma responsable.

El problema más importante relativo a la conservación de documentos digitales auténticos es la obsolescencia tecnológica. El cambio tecnológico aumenta exponencialmente. Esto plantea muchas preguntas, tales como qué hacer con los ficheros elaborados con *hardware* y *software* antiguos, que ya no pueden utilizarse. A menos que se tomen medidas ahora, no hay ninguna garantía de que los ficheros actuales puedan leerse en el futuro con tecnologías futuras.

A fin de hallar soluciones y estrategias para esta situación, el Ministerio del Interior y el Ministerio de Educación, Cultura y Ciencias (Servicio de archivos nacionales neerlandeses) decidieron crear un «banco de pruebas» para adquirir los conocimientos y la experiencia necesarios.

Enfoque

El banco de pruebas de conservación digital está realizando experimentos en función de cuestiones predefinidas con el fin de establecer la mejor estrategia o grupo de estrategias posible en materia de conservación.

El banco de pruebas está investigando tres enfoques distintos de la conservación digital a largo plazo en especial: migración, emulación y XML. No sólo se evaluará la eficacia de cada enfoque, sino también sus límites, costes y potencial de aplicación.

El banco de pruebas de conservación digital está experimentando con documentos, hojas de cálculo, correos electrónicos y bases de datos de diversos tamaños, complejidad y naturaleza.

Los registros se supervisan en los experimentos para establecer si es adecuado (y de qué manera) un método específico de conservación a largo plazo.

Praktische Erfahrungen mit dem Prüffeld für digitale Archivierung

Jacqueline Slats

Wie es heißt, wird das letzte Jahrzehnt die am schlechtesten dokumentierte Periode des 20. Jahrhunderts sein. Immer mehr Informationen werden in digitaler Form aufgezeichnet, ohne dass zuverlässige Methoden für ihre langfristige Aufbewahrung vorhanden sind. Es ist also fraglich, ob künftige Generationen rekonstruieren können, was in unserer Zeit geschehen ist.

Ziel der derzeitigen niederländischen Regierung ist es, bis 2002 etwa 25 % der Vorgänge zwischen dem Staat und seinen Bürgern digital abzuwickeln. Aus diesem Grunde laufen derzeit umfangreiche Arbeiten zur Entwicklung von Strategien, Methoden, Verfahren und Tools für einen verantwortungsvollen Umgang mit den digitalen Erzeugnissen des Staates.

Das größte Problem im Zusammenhang mit der Sicherung authentischer digitaler Unterlagen ist die Veraltung der Technik. Der technologische Wandel vollzieht sich in rasendem Tempo. Dies wirft viele Fragen auf, so z. B. wie mit Dateien zu verfahren ist, die mit alter, mittlerweile nicht mehr nutzbarer Hard- und Software erstellt wurden. Wenn jetzt nichts geschieht, gibt es keine Garantie, dass heutige Dateien in der Zukunft mit künftigen Technologien lesbar sind.

Zur Erforschung von entsprechenden Lösungen und Strategien beschlossen das Ministerium für Inneres und Königreichsbeziehungen sowie das Ministerium für Bildung, Kultur und Wissenschaft (der niederländische Staatsarchivdienst) die Einrichtung eines „Prüffelds“, um so die notwendigen Kenntnisse und Erfahrungen zu gewinnen.

Konzept

Das Prüffeld für digitale Archivierung führt Versuche anhand von vorformulierten Problemlösungsfragen durch, um das beste Konservierungskonzept bzw. die beste Kombination derartiger Konzepte zu ermitteln.

Auf dem Prüffeld werden drei verschiedene Konzepte für die digitale Langzeitarchivierung erforscht, und zwar Migration, Emulation und XML. Dabei wird nicht nur die Wirksamkeit der einzelnen Konzepte bewertet, sondern auch ihre Grenzen, ihre Kosten und ihr Anwendungspotenzial.

Das Prüffeld für digitale Archivierung experimentiert mit Textdokumenten, Tabellenkalkulationen, E-Mails und Datenbanken unterschiedlicher Größe, Komplexität und Art.

Während der Versuche werden die Unterlagen beobachtet, um so festzustellen, ob (und wie) eine bestimmte Methode für die Langzeitaufbewahrung geeignet ist.

Expérimentation pratique du banc d'essai de la conservation numérique

Jacqueline Slats

On a dit que la dernière décennie serait l'époque la plus mal documentée du XX^{e} siècle. Les informations sont enregistrées de plus en plus souvent sous forme numérique, sans qu'il existe de méthodes fiables pour les préserver sur le long terme. Aussi peut-on s'interroger sur la capacité des générations futures à reconstruire ce qui s'est passé à notre époque.

Le gouvernement néerlandais actuel a pour objectif de réaliser 25 % de ses transactions entre l'administration et ses citoyens en recourant aux moyens numériques avant la fin de 2002. Partant, de nombreux travaux sont en cours sur le développement de stratégies, de méthodes, de techniques et d'outils permettant de traiter en toute responsabilité la production numérique de l'administration.

L'obsolescence technique constitue le problème le plus important concernant la préservation de documents numériques authentiques. La technologie progresse à un rythme exponentiel, ce qui soulève des questions sur ce que l'on peut faire des fichiers créés en utilisant d'anciens matériels et logiciels, qui ne sont plus utilisables désormais. À moins de prendre des mesures maintenant, il n'est pas certain que les fichiers actuels puissent être lus à terme par les technologies futures.

Afin de trouver des solutions et des stratégies pour remédier à cette situation, le ministère de l'intérieur et le ministère de l'éducation, de la culture et des sciences (service des Archives nationales néerlandaises) ont décidé de mettre en place un banc d'essai pour acquérir des connaissances et une expérience primordiales.

Stratégie

Le banc d'essai de conservation numérique procède à des expérimentations en fonction de questions prédéfinies afin de mettre en place la meilleure stratégie ou combinaison de stratégies possible.

Le banc d'essai analyse trois approches différentes de la conservation numérique à long terme: transfert, émulation et XML. Il permet non seulement d'évaluer l'efficacité de chaque approche, mais aussi ses limites, son coût et ses applications possibles.

Le banc d'essai peut procéder à des tests sur des textes, des feuilles de calcul, des e-mails et des bases de données, de taille, de complexité et de nature variables.

Les documents sont contrôlés durant l'expérimentation pour voir si une méthode particulière se prête à leur conservation à long terme et, le cas échéant, comment.

Kevin Ashley

Kevin Ashley (K.Ashley@ulcc.ac.uk) works at the University of London Computer Centre, which operates computing services for the United Kingdom and European research, education and public sectors. For the past seven years his work has primarily involved the preservation of large-scale digital resources and their description and access. Most recently these resources have been primarily archival in nature, whether born digital or as digital surrogates, and have involved many types of information (databases, text, images, video and audio) with different access patterns and cataloguing requirements. He is currently Service Manager for NDAD (http://ndad.ulcc.ac.uk/) operated for the Public Record Office of England and Wales, and the National Data Repository at ULCC, which provides digital archiving and distribution services for organisations such as the British Library. He is a board member of the Digital Preservation Coalition and a contributor to electronic records management training provided by the Archive Skills Consultancy (www.archive-skills.com). His career has previously involved pattern recognition in medical image analysis, network protocol development, standards development and numerical software tools.

Producing practical preservation procedures

The best-laid schemes o' mice an' men
Gang aft agley
Robert Burns — 'To a mouse'; 1786

Kevin Ashley

An introduction and apology

The quote from Robert Burns with which I opened my presentation and this paper is intended as an apology and, to some extent, an excuse. (For those not familiar with the Lalland Scots in which Burns wrote, an imperfect translation would probably be useful: 'The most well-made plans of mice and men will often go wrong.') This paper does not contain the content which I intended when I offered it to the DLM-Forum in late 2001. I and some of my colleagues have been involved in reviewing digital preservation practices in another organisation in the UK academic sector, and following this review are due to produce preservation procedures for that organisation and assist in their implementation. By now we had intended that this process would be well advanced and that I could report on our findings. For reasons for which I must take the blame, the timetable is some way behind this.

Many of the lessons I had hoped to illustrate can still be drawn by looking at preservation procedures in a wider context. In this paper, I will take account of the development of preservation procedures in NDAD ([1]) (http://ndad.ulcc.ac.uk/), a service managed by my own organisation, and of my experiences with testing digital preservation procedures within the context of training in electronic records management. I will also look at what we have learned about the problems of developing procedures for a distributed, multi-organisational environment and attempt to draw some conclusions as to what, in my own view, needs to be done to improve the current situation.

Digital isn't that different

When speaking to archivists, records managers or librarians about digital preservation, or about the wider problems presented by digital materials, I often repeat the phrase 'digital isn't that different.' It's my intent in saying this to remind them that the knowledge and skills they already possess for dealing with traditional materials on paper are still generally applicable in the world of computer-generated objects. Whilst these deserve special concern, for reasons made abundantly clear at this and earlier DLM-Forums, it is still worth reminding people of this simple fact. Otherwise, they will often find that fear of the unknown leads them to lose any confidence that their existing training and experience is adequate to the task they face. The result is often paralysis, or worse, action taken in defiance of all the best practice of which they are actually aware. By remembering that 'digital isn't that different' they should instead have confidence that their existing knowledge and skills are relevant, until it is shown that they are not.

The experiences of my partner at the time I began this work perhaps show that the phrase has wider relevance — 'digital preservation procedures aren't that different' does not have quite the same catchy ring, but it has illustrated for me that many of the issues facing us in the development of simple, practical and reusable procedures are not due to the special nature of digital materials but due to far more fundamental issues of human nature and organisational change.

My partner works as a hospital consultant (the senior grade in UK hospital medicine) in an Accident and Emergency (A&E) department of a district general hospital in north-east London. They took the decision some time ago to create a separate paediatric A&E to allow urgent problems with children to be treated entirely separately from those of adults. This would involve a number of changes in practice. One key change involved staffing: traditional A&E departments are staffed by people whose speciality is A&E. They bring in other specialist staff as necessary for particular patients, or refer the patients on to other specialists once their immediate needs are dealt with and their condition stabilised. The paediatric A&E, by contrast, was to be staffed and run

([1]) UK National Digital Archive of Datasets (NDAD).

jointly by A&E staff and paediatric staff. They would each bring their own specialist knowledge to the provision of the service. It was realised that for this to work effectively, guidelines and protocols would need to be developed for use by all staff, and that these guidelines and protocols would need to be agreed beforehand by both specialities. Both 'guidelines' and 'protocols' would translate as 'procedures' in the world I work in. The difference between them is that protocols are absolute instructions as to what must be done in certain situations; guidelines are statements of best practice which should generally be followed, but it is recognised that one must sometimes deviate from them. But it is a characteristic of both that they must be concise and address a clearly-identifiable problem, and be of use in the time-critical environment of A&E. One attribute of the documents is that they must deal with problems as they present in A&E rather than with an illness as identified in a textbook. A patient may appear with severe breathing difficulties; only after some investigation and tests may a diagnosis as to their illness be made. But in the meantime, some action must be taken to prevent the breathing difficulties from becoming worse or threatening the patient's life. It is more important, therefore, that there is a procedure whose title is 'patients with breathing difficulties' rather than a number with titles such as 'patients with asthma.' The latter diagnosis may take some time to make; until it is made, what procedure do you follow?

Paediatric A&E departments are not a totally new development in the United Kingdom, although only a minority of hospitals have them as yet. Nonetheless, the hospitals that do have them are generally held up as leading centres in the United Kingdom in all aspects of medical practice. They are the teaching hospitals, in which most doctors are trained and in which many of the leading practioners work, and the major research hospitals which contain leading experts and research centres working at the cutting edge of medical practice. Working in a fairly average district hospital, my partner reasoned that they should begin their own procedures by examining those already in use in these leading centres, or anywhere else.

She was rapidly disappointed by what she found. A survey of what was available and in use at a wide range of centres in the United Kingdom found very few documents that fitted the description of 'procedures' or 'guidelines' given above. A great deal of written material was found. Much of it was more like a textbook essay than a procedure to be followed in the often fraught environment of a busy A&E department. These textbook essays, which describe not only what to do but why, and examine all the alternative approaches which might have been taken but, on reflection, were rejected, are useful documents in their place. For those who wish to study the subject, or those who wish to understand why medical practice does what it does, they are essential reading. But they are not what the busy nurse or doctor needs at 2 am, faced with a very sick child, very anxious parents and a long queue of other patients. They need a set of simple steps, ideally a single page of A4, that are clearly applicable to the problem in hand and which do not require extended intellectual effort at the time of use.

My partner and her colleagues had to produce their own procedures ([1]) from scratch, although they built on much of the knowledge contained in the material they found. They are now confident that they have something which is ideal for their working environment. It is unlikely that any other hospital could adopt them directly without some adaptation to local working environments. But they are still useful for others to examine and build on because it is clear what they are intended to achieve, and clear to any other hospital which parts require amendment to be useful to them.

What relevance does this have to my topic? Unfortunately, as I stated earlier, 'digital isn't that different.' Our own efforts to survey available procedures for digital preservation have produced results not dissimilar to my partner's. We have found a great deal of informed and learned writing on what forms such procedures might take, on why some steps should or should not be taken, and a lot of very general material which is educational and informative, but difficult to apply in practice. We found very little in the way of practical procedures that were being actively used in the day-to-day task of digital preservation.

In fact, we found precisely two examples. The first was my own group's procedures for all aspects of the operation of NDAD. These are not published or available in any way to the outside world at present, a topic I shall return to later. The other example was from the service providers of the Arts and Humanities Data Service (AHDS) — the very organisation for which we were carrying out the survey of existing procedures([2]).

Procedures at NDAD

I will begin, then, by looking at the nature and development of procedures at NDAD, with which I am most familiar. NDAD's role has been described extensively elsewhere and I will not repeat the detail here. It is sufficient for present purposes to understand that its role is to act as an agent of the Public Record Office (PRO), the UK's national archives, in dealing with structured computer-readable data from UK Government departments. NDAD's functions encompass the transfer of records and their preservation, assessment and cataloguing. It also provides public access to open records, support for the users of the archive and undertakes publicity and marketing roles for the archive's material. Many of these areas are covered by internal procedures. I am concerned in this paper only with those which touch on preservation.

NDAD's procedures were originally developed at the same time as the entire system, in the latter half of 1997 and early 1998. For much of this time, most of the staff of the service were not in post as early-system development was conducted in parallel with the recruitment process. Many aspects of the detailed function of the system were unclear at this time. Nonethless, the nature of our contract with the PRO, arising as it did from a competitive tender, required us to produce documented procedures for all areas of activity, and to have those procedures approved by PRO, before we could take a single step further and have access to actual government records. The procedures were also to form part of our goal of gaining ISO 9000 accreditation for all aspects of our work.

Producing procedures in this way, without the input of the staff who must carry them out and without any practical work environment in which to test them, is not an ideal path to follow. I can understand why the PRO found it necessary to proceed in this way. They wished to establish that we understood what we needed to do before they took the risk of placing real, and sometimes confidential, records into our custody. But as a result the quality of those procedures was very variable. Some were extremely detailed and covered all eventualities that might be encountered. Some were extremely vague and it is questionable whether they deserved the name 'procedure' at all.

One piece of guidance I received from Susan Healy at the PRO, then managing the contractual relationship with us, was to prove vital then and in later developments. 'Do not let the best be the enemy of the good' is a cliché which is worth periodic re-consideration. It was imperative to us then to move to a stage where we could do real work and demonstrate real results, and it would have been easy to be sidetracked by the wish to develop perfection. It is particularly applicable today in a world in which many organisations have an urgent need for procedures for dealing with digital materials. What is often available to them is 'best practice'. Best practice is useful advice to have, and for many organisations it represents an ideal to which they ought to be aspiring. But many do not seem to realise that we cannot all hope to achieve, much less begin, with best practice. We must hope we start with average practice and hope rapidly to move towards good practice. We must avoid bad practice and hope that we never slip from average to poor. This is all very obvious from an analytical perspective but it is forgotten by almost every organisation we have examined when attempting to implement their own guidelines and procedures, and it was an error we nearly made ourselves.

Our procedures have since undergone significant adaptation and enhancement since their initial creation. The reasons are numerous. Technology changes, and hence procedures dependent on it must also change. The skills and knowledge of staff alter, and these changes should and will often lead to changes in the procedures which they must follow. External factors, such as the ability of government departments to transfer material via a network, will also alter and have an influence on some elements of the procedures.

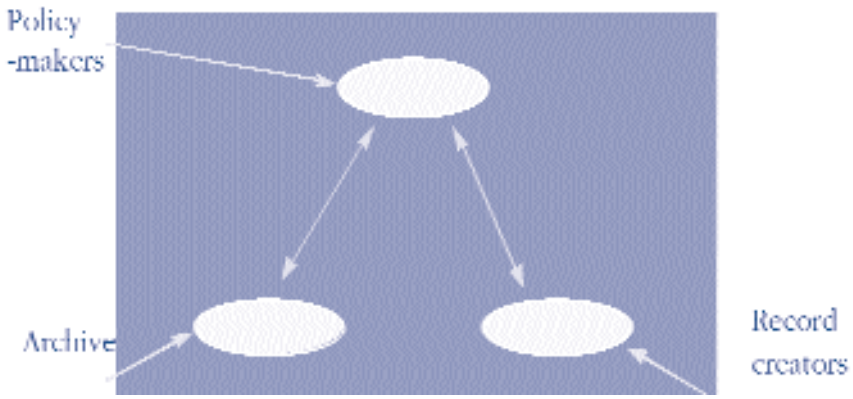
The key lesson which I think can be drawn by others from this is to realise that, for a host of reasons, any procedures you try to develop for yourself will never be completely right the first time. It is important to produce something workable quickly and revise it in the light of real practical experience, rather than spending a great deal of time trying to produce an unattainable perfection. And no matter what approach you take, you will never be totally satisfied with your procedures — you will always see room for improvement.

The influence of organisational models

Organisational complexity has a significant influence on how easy it is to develop procedures and ensure compliance with them. In particular, the model offered by ISO 9000 seems to be applicable only within a context where a single organisation is the home for all the actors in a procedure. We have found difficulties with external compliance auditors in gaining acceptance for a procedure which involves activities undertaken by a third party outside our organisational control.

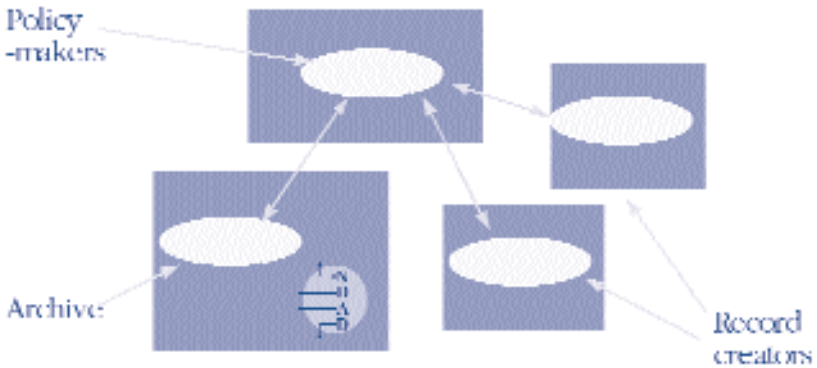
The classical organisational view of the ISO 9000 world is one much like Figure 1. The external box represents the organisation as a whole and the ovals, units within it. The archive, which creates the procedures, can work through policy-makers to influence record creators.

Figure 1



The slightly more complex world in which NDAD operaters is illustrated by Figure 2. There are multiple record creators each in their own organisation. The policy-makers are in PRO, a separate organisation from ourselves or the record creators. Getting compliance from all these actors to procedures which we develop is not simple — in fact, we simply cannot assume it will happen. Most processes involve actors from at least two organisations, but compliance and auditing of procedures can only be conducted in a single organisation. Hence, our procedures must be made resilient to variances in external behaviour. They must attempt to ensure that the correct result is achieved regardless of those variances, and they must somehow attempt to encourage compliance by the external actors.

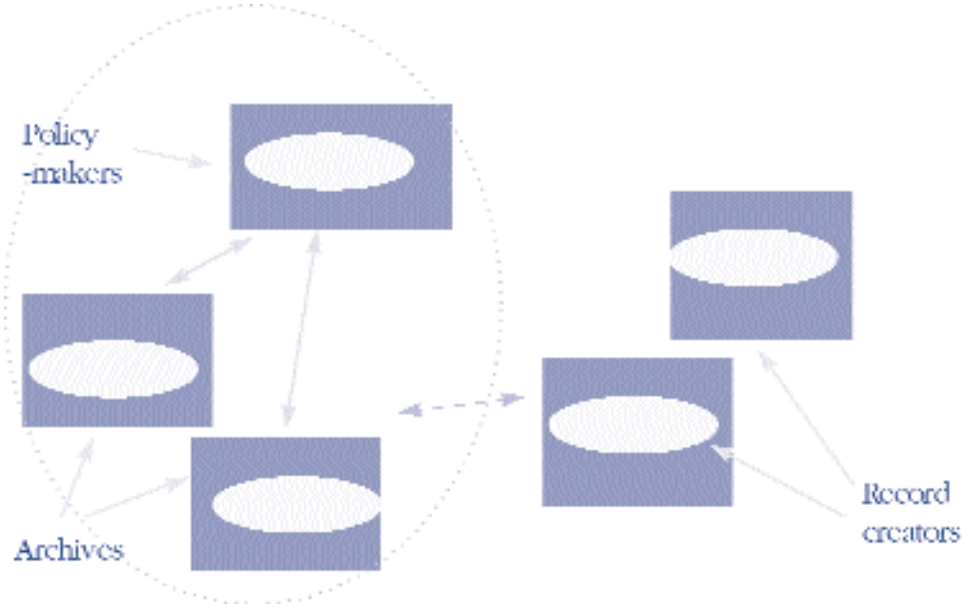
Figure 2



The world in which AHDS operates is more complex still and analagous to Figure 3. There are multiple archives operating in separate institutions with a single executive body overseeing them all. Each archive is embedded in a host organisation with its own requirements and procedures, yet the archives must be capable of interchanging material and metadata with each other in a straightforward fashion. Each of the archives is dealing with multiple sets of record creators and each of those is in a separate organisation. The influence of any archival bodies over the record-

creating organisations is minimal or non-existent. Methods of encouraging compliant behaviour, and mechanisms to provide tolerance of non-compliant behaviour, are thus even more important in this world.

Figure 3



There are a number of ways to achieve our aims in these multi-organisational scenarios. One is to devise procedures which minimise interactions with external actors. To some extent, this is the approach we have taken at NDAD. Another is to embed positive incentives for compliance by external actors. One excellent example of this was provided in a paper at the previous DLM-Forum ([3]) by staff of the central archives of the department of culture of the Catalan regional Government. Although concerned with records creation and management rather than preservation, I was struck by the use of simple processes and cheap technology combined with clear user benefits to create a system which people were actually happy to use and which achieved its objectives.

Practical examples needed

More practical examples of real-world procedures are needed to enable organisations to develop their own. Our work and training programmes have looked to existing examples and guidelines both to help others develop their own procedures and to use for workshops within a training session. We have taken examples such as the 'Preservation management of digital materials' ([4]) (particularly the fold-out decision tree), the DLM guidelines and those of the PRO in the United Kingdom. Each is too general and lacking in specifics to enable an organisation to employ them directly. They are excellent tools for awareness raising and for education, and we would all be poorer for the lack of them.

Why should this be a concern? It is partly because of my experience that many people and organisations seem to believe that these existing guidelines are the only answer that is needed — that because they exist, we are all safe and no more needs to be done. This is a very dangerous form of complacency for any organisation that has not yet tackled the task of digital preservation in a realistic manner. The documents currently available are simply starting points, tools which will help those new to the task to grasp the general principles and frame a plan of action as to what needs to be done. You cannot buy a copy of these guidelines today and start doing preservation tomorrow. Yet some people behave as if that is the case.

Two concrete examples of the need for real-world procedures relate to two specific problems in preservation: naming for digital objects, and formats for preservation. Naming relates to the need for every object to have some unique identifier which enables us to track it through migrations, refreshing, rendition changes and movements between repositories. Preservation formats are

those which we would choose to use to store objects for the long term. The principles behind both of these problems are well-understood and widely described in the literature. But it is extremely difficult to find any particular example of a naming or identification scheme, or a list of preservation formats used by one organisation. As a result, each of us active in this area today ends up re-inventing schemes which meet the widely-accepted principles, but which do not interact with each other. This lack of interaction will become a problem as the preservation field matures, as a market develops and as the need to move material between repositories becomes more and more widespread.

One organisation at least deserves some credit in this respect: the Archaeology and History Data Services in the United Kingdom have the best examples we could find of definitive statements on the preservation formats they employ, and how they relate to ingest formats. These can be found on their web sites: ads.ahds.ac.uk and hds.essex.ac.uk.

Conclusions, and a look forward

The work we have done convinces me that it is necessary for more of those actively involved in digital preservation to publish our procedures. I know that there are many other organisations who have moved beyond the pilot or exploratory phase into performing digital preservation as a part of their everyday tasks. Like NDAD, they must have developed internal procedures to enable this to happen. And like NDAD, for whatever reason, they have not chosen to publish them. We are all suffering from this lack of openness and the problem will get worse. I am determined that we will make our procedures available, even if it opens us to justified criticism where those procedures are seen not to be the best possible. I encourage all those involved in areas similar to ours to do the same.

There is a message also for the funding agencies. Much of the work funded so far has been closer to academic research than industrial development. This research was necessary and I do not wish to argue that it should not have been funded. But there is a need now to move the focus towards practical application of that academic research. Most digital preservation at present is being undertaken by experts in national-level institutions. I have no doubt that they are all doing an excellent job. But the problem we face is to scale this in a number of different dimensions. We need to move what were pilot, laboratory-type preservation regimes into production-line systems dealing with millions or billions of objects per year. And we also need to create systems and practices that can be used by a county, town or village archive as easily as they can be used by national government. Helen Shenton alluded to this problem in her paper presented to 'Preservation 2000' ([5]). She drew the analogy with traditional preservation and conservation, in which a small number of highly-qualified academics study the detailed chemistry and mechanics necessary to develop the skills and techniques of conservationists, but the bulk of the work is carried out by skilled craftspeople many of whom stopped their formal education at the age of 16. Digital preservation procedures are not yet capable of implementation in such an environment. Yet they must be if we are to deal with the scale of the problem that the promised information society and e-government will present.

Acknowledgements

This paper would not have been possible without the views, hard work and support of colleagues from my own and other organisations over a number of years. Not all of them may have been conscious at the time of the assistance or insights they were providing, but their assistance is gratefully acknowledged nonetheless.

The staff at NDAD, past and present, have performed sterling work in turning ideas into practical, working reality and in developing and refining our own preservation and accession procedures during the past four to five years. They have taught me more than they realise.

Others to who particular thanks are due are:

Marc Fresko and Mike Rice of Cornwell Management Consultants (www.cornwell.co.uk) for their collaboration on the work with AHDS; All the staff at AHDS (www.ahds.ac.uk) and their service

providers; Margaret Crockett and Janet Foster of the Archive Skills Consultancy (www.archive-skills.com); The staff of the Estonian National Archives.

References

(1) Brown, N., Solebo, S., 'King George Hospital paediatric A&E guidelines' — Version 2.1, CD-ROM and paper, Redbridge Healthcare, 2001.

(2) An example is the History Data Service procedures at: http://hds.essex.ac.uk/publications.asp

(3) Serra, J., Canela Garayoa, M., Campos Gonzalez, I., Domingo Basara, J., 'The integration of electronic records into a records management system: the process of cost optimisation', INSAR Supplement IV, 2000.

(4) Jones, M., Beagrie, N., *Preservation management of digital materials*, British Library, 2000, and also online at: www.dpconline.org

(5) Shenton, H., 'From talking to doing: digital preservation at the British Library', *The new review of academic librarianship*, Volume 6, 2000.

Elaboración de procedimientos prácticos de conservación

Kevin Ashley

En el otoño de 2001, el Servicio de Datos sobre las Artes y las Humanidades británico (AHDS) solicitó ayuda para elaborar y aplicar procedimientos coherentes y viables de conservación digital en sus cinco proveedores de servicios, situados en diversos puntos del Reino Unido. Para este trabajo fuimos seleccionados compañeros del ULCC (University of London Computer Centre) y yo mismo, junto con Marc Fresko y miembros de Cornwell Affiliates. Los resultados del proyecto serán: una taxonomía de los tipos de datos tratados, un manual de conservación de metadatos, un estudio de las prácticas actuales del AHDS y un conjunto de procedimientos que puedan adoptarse en el AHDS a fin de garantizar la aplicación de prácticas coherentes de conservación en el futuro.

El AHDS (http://www.ahds.ac.uk/) es un servicio financiado por la enseñanza superior británica, cuyo fin es recopilar, catalogar, conservar y proporcionar una gama de recursos archivísticos en el ámbito de las artes y las humanidades. En la actualidad consta de un ejecutivo central, que administra los fondos, determina la política y supervisa los resultados de todo el servicio, y de cinco proveedores de servicios: el Servicio de Datos sobre Artes Interpretativas (PADS), el Servicio de Datos sobre Arqueología (ADS), el Servicio de Datos sobre Historia (HDS), el Servicio de Datos sobre Artes Visuales (VADS) y el Archivo de Textos de Oxford (OTA). Los proveedores de servicios están distribuidos por todo el país, en distintas instituciones académicas. Se centran en la recogida de recursos de interés para grupos de investigación concretos. Algunos proveedores de servicios existían antes de la creación del servicio global, y conservan datos desde hace muchos años; otros solamente han existido los cinco años de vida del AHDS. A medida que fue creciendo el servicio, cada uno desarrolló su propio método de trabajo, utilizando los distintos niveles de conocimientos técnicos de que disponían a escala local.

La gama de información que recogen cubre casi todas las formas conocidas de información digital. Además de los tipos más conocidos, como imágenes digitales fijas y textos estructurados, también recogen un gran volumen de datos variables en el tiempo, como datos de audio y vídeo digitales. Las bases de datos y los sistemas de información geográfica consisten, al igual que los ficheros de diseño automatizado, en sistemas de datos de realidad virtual, sitios web enteros, estructuras de datos híbridas complejas y algunos tipos de datos

muy especializados como información de modelos geofísicos. Para la viabilidad a largo plazo del servicio, es importante que esta información se describa y se gestione de manera coherente. Por ejemplo, puede ser deseable conservar o proporcionar acceso a todo ello desde una única infraestructura de servicios centralizada, incluso aunque los cinco proveedores de servicios sigan estando dispersos. Los proveedores de servicios podrán estar fusionados, o podrán crearse nuevos proveedores que puedan heredar parte de los datos recopilados por los proveedores existentes. Esto requiere unos metadatos coherentes, unos identificadores únicos coherentes para cada objeto y grupo de objetos conservados, y una cierta lógica en el uso de formatos de conservación y las prácticas de migración.

Nuestro objetivo era estudiar hasta qué punto el servicio se alejaba de este ideal, y desarrollar procedimientos, taxonomías de tipos de datos y normas sobre metadatos que contribuyan a alcanzar el ideal. También debíamos asegurarnos de que exista un plan que permita al servicio y a los proveedores de servicios evolucionar desde su posición actual de práctica muy variable a una posición de uniformidad, cuando ello sea práctico y deseable. Se ha efectuado mucho trabajo teórico sobre lo que debería hacerse en un mundo ideal con recursos ilimitados. Sin embargo, nuestros procedimientos tenían que poder aplicarse en relativamente poco tiempo, con recursos limitados y en lo que eran de hecho seis organizaciones distintas.

Creemos que, a consecuencia de esto, nuestros procedimientos pueden utilizarse con mayor amplitud y pueden beneficiar a otras instituciones y servicios, particularmente las que actúan en el ámbito del patrimonio cultural en Europa. Estos procedimientos se publicarán como parte de las guías de buenas prácticas del AHDS, y estarán disponibles en su sitio Internet y en formato impreso por parte de sus editores asociados.

Erarbeitung praktischer Archivierungsverfahren

Kevin Ashley

Im Herbst 2001 bat der britische Geisteswissenschaftliche Datendienst (Arts and Humanities Data Service, AHDS) um Unterstützung bei der Entwicklung und Umsetzung einheitlicher und praktikabler Verfahren für die digitale Archivierung bei seinen fünf Dienstleistern, die an verschiedenen Standorten im Vereinigten Königreich angesiedelt sind. Für diese Arbeiten wurden ich und Kollegen vom ULCC sowie Marc Fresko und Kollegen von Cornwell Affiliates ausgewählt. Als Ergebnisse sind eine Taxonomie der behandelten Datenarten, ein Handbuch der Konservierung von Metadaten, ein Überblick über die derzeitige Praxis im AHDS sowie eine Zusammenstellung von Verfahren vorgesehen, die sich im gesamten AHDS anwenden lassen und dafür sorgen, dass künftig einheitliche Archivierungsmethoden verwendet werden.

Der AHDS (http://www.ahds.ac.uk/) ist ein vom britischen Hochschulwesen finanzierter Dienst für die Erfassung, Katalogisierung, Konservierung und Bereitstellung von Archivgut im Bereich Kunst und Geisteswissenschaften. Er besteht zur Zeit aus einer Zentrale, die die Finanzmittel verwaltet, den Kurs festlegt und die Arbeitsleistung des gesamten Dienstes überwacht, sowie fünf Dienstleistern: dem Datendienst für darstellende Kunst (Performing Arts Data Service, PADS), dem Archäologie-Datendienst (Archaeology Data Service, ADS), dem Geschichtsdatendienst (History Data Service, HDS), dem Datendienst für bildende Kunst (Visual Arts Data Service, VADS) und dem Oxford Text Archive (OTA). Die Dienstleister sind über das gesamte Land verteilt und verschiedenen Hochschuleinrichtungen angeschlossen. Sie konzentrieren sich jeweils auf die Erfassung von Ressourcen für bestimmte Forschungsrichtungen. Einige Dienstleister bestanden bereits vor der Bildung des zentralen

Dienstes, und ihre Sammlungen reichen viele Jahre zurück. Andere wurden erst mit dem AHDS vor fünf Jahren gegründet. Zunächst wurde es jedem Dienstleister gestattet, unter Nutzung der in seinem Umkreis vorhandenen unterschiedlichen Fachkenntnisse eigene Arbeitsmethoden zu entwickeln.

Die von ihnen erfassten Daten erstrecken sich auf nahezu sämtliche bekannten Formen digitaler Informationen. Neben den gängigeren Arten wie unbewegten digitalen Bildern und strukturiertem Text sammeln sie auch zeitabhängige großmengige Datenarten wie digitale Video- und Audiodaten. Datenbanken und geografische Informationssysteme spielen ebenso eine Rolle wie CAD-Dateien, Datensysteme der virtuellen Realität, ganze Websites, komplexe Hybriddatenstrukturen und einige hochspezielle Datenarten wie geophysikalische Modellinformationen. Für die langfristige Funktionsfähigkeit des AHDS ist es wichtig, dass diese Informationen einheitlich verzeichnet und verwaltet werden. So ist es unter Umständen sinnvoll, alle Informationen von einer zentralen Dienstinfrastruktur aus zu konservieren oder zugänglich zu machen, auch wenn die fünf Dienstleister weiterhin dezentral arbeiten. Auch können sich Dienstleister möglicherweise zusammenschließen, oder es können neue gebildet werden, die einen Teil der Sammlungen bestehender Dienstleister übernehmen. Dazu bedarf es einheitlicher Metadaten, einheitlicher Kennungen für die einzelnen aufbewahrten Objekte und Objektgruppen sowie einer einheitlichen Anwendung von Archivierungsformaten und Migrationspraktiken.

Unser Anliegen war es zu untersuchen, wie weit der AHDS von diesem Idealzustand entfernt ist, sowie Verfahrensweisen, Datentyptaxonomien und Metadatenstandards zu entwickeln, die den Idealzustand mit herbeiführen sollen. Darüber hinaus gilt es sicherzustellen, dass ein Plan vorliegt, anhand dessen der AHDS und seine Dienstleister ihre derzeitige äußerst unterschiedliche Handhabung so weit vereinheitlichen, wie es angebracht und sinnvoll ist. Aus vielen theoretischen Vorarbeiten können wir erfahren, was in einer idealen Welt mit unbegrenzten Mitteln getan werden sollte. Unsere Verfahren mussten jedoch innerhalb relativ kurzer Zeit, mit begrenzten Mitteln und in praktisch sechs verschiedenen Organisationen umsetzbar sein.

Daher lassen sie sich unserer Ansicht nach umfassender einsetzen und können auch anderen Einrichtungen und Diensten zugute kommen, insbesondere im Bereich des kulturellen Erbes in Europa. Die Verfahren werden demnächst im Rahmen der Leitfäden für gute Praxis des AHDS auf der Website des Dienstes sowie in gedruckter und gebundener Form von seinen Verlagspartnern veröffentlicht.

Élaboration de procédures de conservation pratiques

Kevin Ashley

À l'automne 2001, le service britannique AHDS (Arts and Humanities Data Service) a sollicité une aide pour développer et mettre en œuvre des procédures de conservation numérique cohérentes et viables chez ses cinq prestataires de services, implantés en divers lieux du Royaume-Uni. Mes collègues de l'ULCC et moi-même, ainsi que Marc Fresko et mes collègues de Cornwell Affiliates, ont été retenus pour réaliser ce travail. Le projet aura pour résultats: une classification des types de données traitées, un manuel de conservation des métadonnées, une analyse des pratiques actuelles au sein de l'AHDS et un ensemble de procédures qui pourront être adoptées dans l'ensemble de ce service pour assurer des pratiques conservatoires cohérentes à l'avenir.

L'AHDS (http://www.ahds.ac.uk/) est un service financé par l'enseignement supérieur britannique, qui a pour but de collecter, de faire l'inventaire, de conserver et de mettre à disposition tout un éventail de ressources archivistiques dans le domaine des arts et des lettres. Ce service consiste actuellement en une direction centrale, qui gère les financements, définit les politiques et supervise les résultats de l'ensemble du service et des cinq prestataires: le PADS (Performing Arts Data Service), l'ADS (Archaeology Data Service), le HDS (History Data Service), le VADS (Visual Arts Data Service) et enfin l'OTA (Oxford Text Archive). Ces différents prestataires sont répartis un peu partout dans le pays et sont hébergés chacun auprès d'un institut universitaire différent. Chaque service est spécialisé dans le recensement de ressources intéressant des groupes de chercheurs particuliers. Certains de ces prestataires existaient avant la création du service global et possèdent des collections qui remontent à plusieurs années, tandis que d'autres n'existent que depuis cinq ans, c'est-à-dire depuis la création de l'AHDS. Aux premières heures de ce service, chacun avait la possibilité de développer ses propres méthodes de travail et de faire appel aux différents niveaux d'expertise présents à l'échelon local.

L'éventail des types de données collectées couvre la plupart de toutes les formes d'information numérique connues. Outre les types les plus familiers, images numériques et texte structuré, par exemple, ces services collectent également de gros volumes de données variables dans le temps, notamment des données vidéo et du son numériques. Les bases de données et les systèmes d'information géographique incluent, à l'instar des fichiers de CAO, des systèmes d'information virtuels, des sites web entiers, des structures de données hybrides complexes et certains types de données hautement spécialisées, telles que des informations de modélisation géophysique. Il importe pour la viabilité à long terme du service que cette information soit décrite et gérée de façon cohérente. Il pourrait être souhaitable, par exemple, de conserver cette information ou d'en assurer l'accès à partir d'une infrastructure de services centralisée unique, même si les cinq prestataires sont toujours répartis sur le territoire. Des fusions entre les prestataires pourraient intervenir, ou de nouveaux pourraient être constitués et hériter d'une partie des collections des anciens. Pour que ces évolutions puissent avoir lieu, il faut donc concevoir des métadonnées cohérentes, des identifiants uniques cohérents pour chaque objet et groupe d'objets conservés, et garantir une certaine logique dans l'utilisation des formats de conservation et des pratiques de transfert et de conversion.

Notre objet était de voir dans quelle mesure le service était loin de cet idéal et de développer des procédures, des classifications de types de données et des normes de métadonnées pour contribuer à atteindre cet idéal. Il nous incombe aussi de veiller à ce qu'il existe un plan permettant à l'AHDS et à ses prestataires d'abandonner la multitude des pratiques actuelles au profit de pratiques homogènes, chaque fois que cette uniformisation est indiquée et souhaitable. De nombreuses études théoriques ont été réalisées, qui nous renseignent sur ce qui doit être fait dans un monde idéal aux ressources illimitées. Toutefois, nos procédures devaient aussi pouvoir être mises en œuvre en un laps de temps relativement court, avec des ressources limitées et dans un environnement qui réunissait en fait six organisations différentes.

Nous estimons qu'au bout du compte ces procédures pourront être plus largement utilisées et pourront bénéficier à d'autres institutions et services, notamment ceux opérant dans le domaine du patrimoine culturel en Europe. Ces procédures seront ultérieurement publiées en tant que partie intégrante des guides de bonnes pratiques de l'AHDS, et seront proposées sur son site web et sous forme reliée et imprimée par les éditeurs partenaires.

Amanda Blunden

Amanda Blunden is Processing Contract Project Manager at the Office for National Statistics. She has worked at the Office for National Statistics since 1985, working on the 1991 Census from 1989 before starting work in Research and Development for the 2001 Census in 1993. Originally working in the information technology division as a Cobol programmer, she moved to the business area in 1995 to lead the work on researching alternative data capture and coding approaches for the 2001 Census. As part of a team, Amanda worked on the procurement of the contract with Lockheed Martin to outsource the capture and coding services and led the team to produce specifications for processing the 2001 forms. Since October 2000 she has acted as Project Manager, ensuring that Lockheed Martin delivers to the contract and that the Office for National Statistics meets its contractual obligations.

Digital preservation: The 2001 UK Census

Amanda Blunden

I am delighted to be able to speak to you today about how technology, and specifically document management technology, has assisted the Office for National Statistics (ONS), General Register Office for Scotland (GROS) and Northern Ireland Statistical Research Agency (NISRA) carry out the 2001 population census in a more efficient and effective way than ever before.

As the Project Manager for the 2001 Census processing contract, I have a day-to-day responsibility for ensuring that the contract details are carried out by our contractors as we move towards concluding the data extraction and image-retention phase of the 2001 Census. Before taking on this responsibility I was heavily involved in preparation required to design the forms and award the contract.

In this presentation I will show how we have used document management in the context of the census; give an overview of the data processing function; and identify the benefits that this approach delivered with particular emphasis on the impact of changes since the 1991 Census.

The United Kingdom Census of Population and Housing, to give it its full title, took place on 29 April 2001, the latest in a series of censuses that have taken place every 10 years since the late 19th century, with just one or two exceptions for national emergencies.

For the 2001 Census we needed to take into account that changes, in society and especially in technology, are taking place at an ever-increasing pace and that the 2001 Census needed to rise to this challenge. Particularly compelling was the way in which advances in technology would allow for a totally fresh approach to both document management and operational activities.

Performing a census is a huge undertaking. Indeed, it is probably the biggest peacetime exercise that a country conducts, involving many different organisations and people. The logistics for the 2001 UK Census are impressive: we had to deal with in excess of 30 million forms in 20 different variants to take account of regional and language differences. These forms themselves were not uniform with a range of between a single page up to 20 pages long. In total, the amount of paper created would fill filing shelves 37 miles long. For the document scanning and character reading part of the census we would have to scan over 400 million pages and store over 28 terabytes of image and data.

These figures mean that good, solid planning is essential. In fact, planning for the 2001 Census started before the 1991 Census was complete and already our thoughts are turning towards the 2011 Census.

For the 2001 Census we started by conducting research and development: research into the type of census required, what information needed to be collected, what our customers expected from the census; as well as research into the technology available, and into what would be available as the date of the census grew closer. We knew that the technology available at the turn of the millennium would be vastly different from the technology we could view in the early 1990s. In fact, the development of the system in terms of technology would continue right through until 2001.

A prototype system was designed and tested in order to prove the concept for the technology to be used. This led to an EU procurement process with the contract being awarded to Lockheed Martin in 1998. A full rehearsal started in September 1999 for the 'go live' date in June 2001. The processing of the forms covered the processes to extract, collate and distribute the data — which will initially begin this year — and to ensure that the forms themselves would be available to be viewed in 100 years' time.

The 2001 Census involved many large and small providers. The main ones were Lockheed Martin/Polestar Petty for the printing and processing of the forms; Fujitsu (formerly ICL) who man-

aged the warehousing for forms storage and all associated operational functions; and Kodak who provided all of the digital imaging and microfilming equipment. In addition, there were a host of independent business and topic experts who contributed on a number of fronts.

The business benefits that this provided were considerable.

The Census offices did not have the scanning, automatic capture, and workflow expertise for a capture and coding exercise of this scale and complexity yet were keen to use modern methods of data capture for efficiency gains — previous censuses had involved processing using paper and then keying the end result (there were over 550 keyers employed in the 1991 Census keying much less data). Work had been contracted-out to a very limited degree in the past. By contracting out we were using the contractors' expertise and reducing our risk.

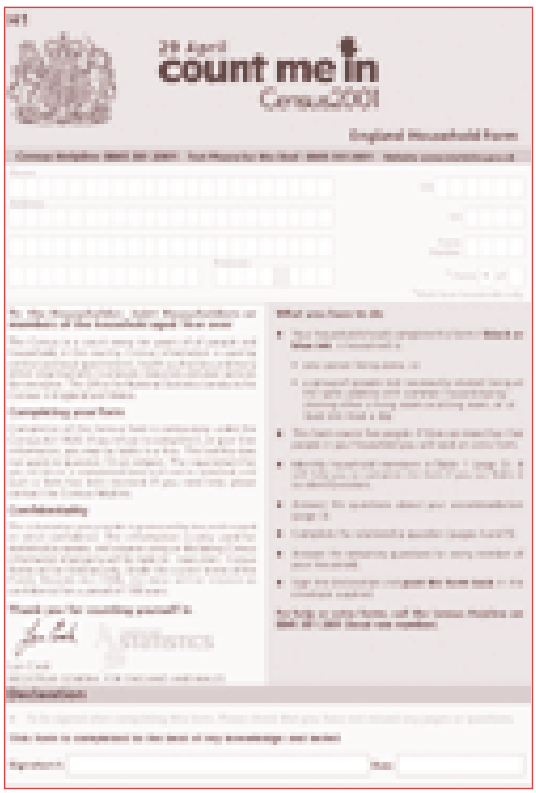
One hundred percent processing allows us to meet many more requirements of our customers and improves the quality of the output. Previously, a 10 % sample of households were processed for 'difficult to code' questions.

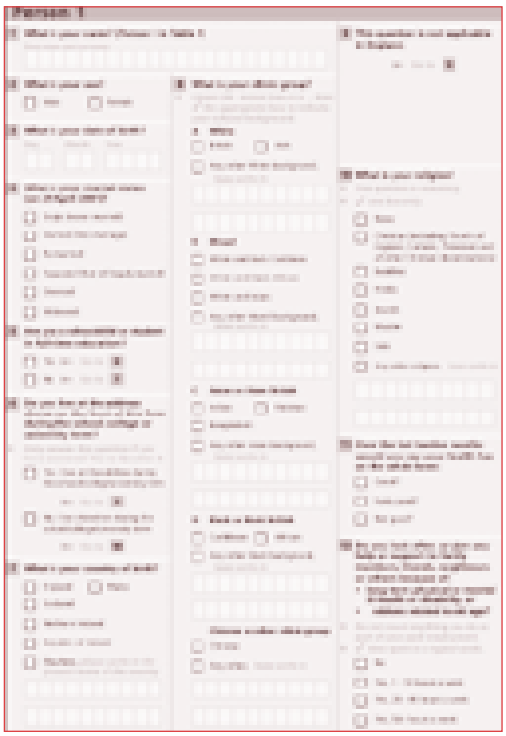
These gains meant that the 2001 Census was longer than the 1991 form, meaning that more of our customers' requirements could be met by capturing more data.

Producing digital and microfilm images reduces the overheads associated with paper handling and long-term maintenance. In addition, the storage of the forms would have required in excess of 37 miles of shelving. Now a set of microfilms needs just 300 square feet.

The form design had a number of important considerations. Any content that was to be included needed to have a strong business case as it would essentially be a public burden. The form would need to provide good-quality data that could not be found from alternative sources. With the adoption of new technology there would also be a need to strike a balance between public acceptability and the scanning requirements for the forms. Independent print and form design experts were consulted as well as taking input from business and technical partnerships. The end result was forms that were very different from the 1991 Census. Once the design was agreed thorough testing took place — both question wording cognitive tests as well as scanning/capture trials — and then the massive print run was made using a high-tech 16-jet barcoder.

As a large part of the data was to be captured automatically, certain design characteristics were added to assist this: barcodes were used to identify both form type and page number as well as to give each form a unique ID — this also ensured form integrity was preserved throughout the whole process; tick and character boxes were used to facilitate optical mark reading or optical character reading; colour drop ink was used to ensure that as much 'noise' on the forms was eliminated as possible (to reduce image size) although the actual images stored were to be of the complete forms; and to assist with automatic character recognition capture zones were pre-defined although this had to be balanced out with the need to consider question layout.

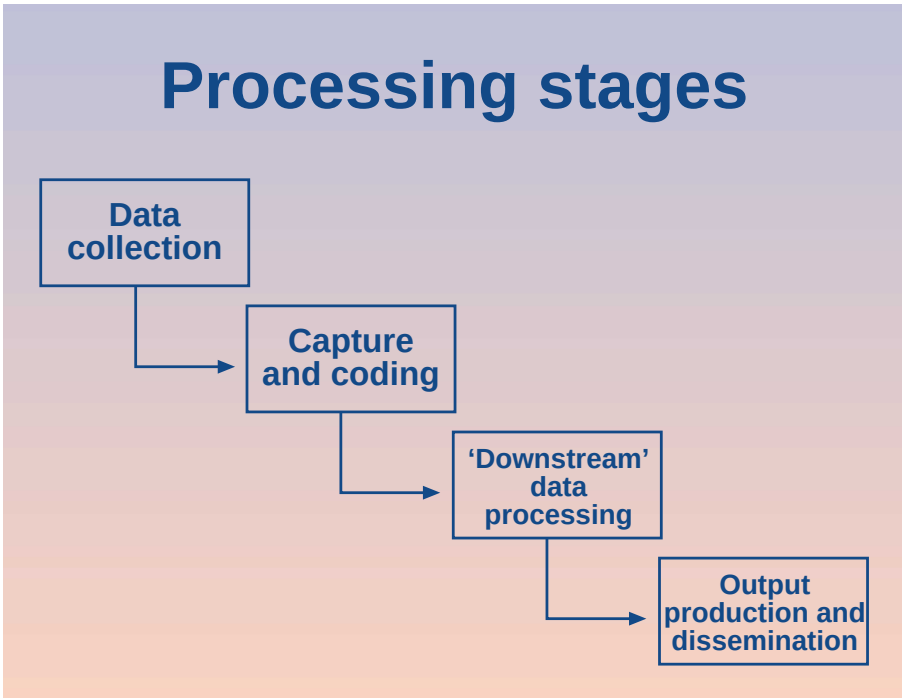




This is the front page of the Census 'Household' form for England. As you will notice the majority of the data on this page is either alpha or alphanumeric information.

This is the first page on the questions relating to the person. In total there were 32 questions (33 in Wales) relating to the individual, although not all applied to everyone. There were an additional nine questions on housing. As much use as possible is made of tick boxes with the ability, for certain questions, for write-in responses to be added.

Now I'd like to take you briefly through a high level overview of the processing system, focusing on capture and coding, workflow, and organisation.



A brief outline of the processing stages

'Data collection' refers to the field exercise that has been described earlier. This provides the main input to processing — the forms themselves, enumerators' record books and counts for reconciliation purposes.

'Capture and coding' covers the capture of the forms as images and the capture and coding of the information on those forms. This work has been contracted out to Lockheed Martin, who were also responsible for delivering the capture system for the US Census in 2000. This is the area that will be described in more detail in the remainder of this presentation.

'Downstream' processing has been developed primarily by ONS for all three Census offices. It covers a number of highly complex and automatic processes that make the data consistent and com-

plete. For example, item imputation fills missing or invalid responses with valid values from donors. Complete household and person imputation also takes place. Another example is disclosure control which is applied to the data to preserve confidentiality.

'Output production and dissemination' is the final process of tabulating the data and making it available to customers (i.e. via the Internet, on CD-ROM, on paper etc.), including laying certain reports before the relevant parliament.

Moving specifically to capture and coding, here are some of the facts and figures:

There was an 11-month processing schedule (taking into account ramp-up) which meant that up to 1 million forms per week were being handled at peak production. In total, there were just under 30 million forms representing 400 million double-page sheets to be scanned. There were over 20 different types of forms to account for regional variances (for example, the Northern Ireland household forms had 24 pages rather than 20). The forms themselves were transferred daily between the warehouse and the processing centre. The buildings were acquired and run by Fujitsu. The warehouse was a refurbished building whilst the processing centre was a new building. The staff were mainly recruited, trained and managed by Fujitsu.

In previous censuses, complex coding has been the most labour intensive process. This is where an industry and occupation code is derived from the text written for every potentially economically-active person. Only 10 % of cases were coded in previous censuses but the technology employed here allowed for 100 % coding.

The key deliverables gained were the census data; the audit data; the digital images; and the microfilm images.

At the end of the processing cycle the forms were destroyed leaving just the digital images for the short term and the microfilm images for the long-term retrieval.

A rough idea of the timeline involved for capture and coding is as follows:

Capture and coding timetable

- **Development & testing started 1999**
- **Operational readiness testing March/April 2001**
- **Dry-run - May (using 'live' forms)**
- **Census Day - 29 April 2001**
- **Service started 18 June 2001**
- **Service ands 17 May 2002**
- **Main results from March 2003**

Development and testing commenced in 1999. During the Autumn of 2001 system functionality testing (SFT) took place to ensure that the system requirements were fully met. In March and April 2002 operational readiness testing took place at the warehouse and processing centre. This checked that all services (equipment and staff) were in place and ready to go.

In May a 'dry run' using 100 000 live forms, collected early from selected areas, took place. This provided further confirmation that the service was ready to commence properly. As an aside, these forms would subsequently be re-processed at their planned point in the processing schedule.

For Census Day itself we had issued a plea to the country — 'Please complete your form and return it on time'. This was to allow us to start the service on 18 June 2001. The current planned end date is 17 May 2002 (original end date was end-March 2002). In fact the last data deliveries were made on 8 May 2002. The main results from the Census will be available in 2003 although figures for the population estimates should be available in the summer of this year.

In order to facilitate processing, the forms were divided between seven parallel streams, or 'clusters'. These clusters would then process the forms in the following way.

The first thing to happen was document preparation. Here, any unwanted items were removed and the forms were allowed to acclimatise for 24 hours. Then they were put into batches of 10–15 forms. The edges of the forms were guillotined to remove staples.

Each cluster was then equipped with three high-speed Kodak 9520 document scanners outputting the images in 200dpi TIF files. A document analyst randomly checks samples for quality whilst the scanner checks, using the barcodes, that no page has been missed. To allow for contingency the planned throughput for the scanners is 80 double-sided pages per minute, which is actually half of the theoretical maximum throughput speed.

The data is then passed through optical character recognition (OCR) and optical mark recognition (OMR) software to automatically capture the data. Any data which fails the confidence thresholds is sent to operators to key the data from images.

The production/quality control section monitors scanner and keyer rejections. Any form that fails at these stages is pulled from its original box placed in a new exception box ready for exception scanning and keying from paper. In each cluster the exception scanning uses one Kodak 1500D low-volume flatbed scanner. This was used to get images from forms that were too damaged, had been completed in red or another drop-out colour or had been pulled at production control. What became clear during this exercise is that poor-quality forms meant poor-quality recognition and the need for some keying from paper as back-up.

Once all data had been captured, the textual responses were passed through automatic coding systems to translate the responses into numeric codes in accordance with coding indexes. Responses, which could not be coded automatically, were referred to manual coders who had access to the images, data and reference material to carry out the coding operation.

The prime data export process was to use RAID storage with magnetic tape back-up. All data and images were staged together by output area (there were 112 geographical areas across England, Wales, Scotland and Northern Ireland). Once organised into the logical sequence they were exported either by DLT or CD.

It was decided that for the 2001 UK Census all of the digital images would be output to microfilm to ensure that in 100 years an eye-readable copy, totally independent from technology changes, would be available to be viewed. This was achieved by using 12 Kodak digital archive writers situated on site. The digital images were pre-sorted by form identity within a geographical area (to match the data delivery), blank images were deleted in accordance with pre-specified rules and then output onto 16mm roll of microfilm at 48x reduction with the front and backs of the forms next to each other on the film. In total, around 300 million images will be output to microfilm in an operation that took just six months and which is due to be completed on 30 June 2002.

To recap on some of the deliverables that the capture and coding process enabled: we had digital data for all output areas; approximately 300 million digital images loaded on an in-house image management system; a similar number of 16mm roll microfilm frames for long term archiving and delivery to the Public Records Office; and the ability to run a number of management reports.

As this process is still ongoing, I thought I would just detail some of the checks that we are currently going through. The digital images are received on DLT or CD and are then logged and listed on receipt. These are then loaded onto the in-house image management system and we carry out selective quality checks for coverage, legibility and other defects. So far the quality has been very good.

A similar process takes place for the microfilm images. These are also logged and listed and then they are loaded onto a motorised digital reader-printer, which can scan the microfilm image, turning it back into a digital format. The images are selectively checked for coverage, legibility, scratching and other defects and to ensure that they comply with British Industry standards. The quality has also been very good to date.

To finish the presentation, I would just like to highlight the role that the Census offices take in this type of process, which is split into two areas: the role taken during development and the role taken during the live operation.

In the development phase, the Census offices have specified numerous capture and coding rules that have been implemented by Lockheed Martin (e.g. from simple handling of multiple ticks to complex filtering rules (economically active)).

The Census offices are responsible for the preparation and delivery to the contractor of the coding indexes and geography database used in the capture and coding operations. We have shared our expertise to ensure that the systems meet requirements and that coders are properly trained.

Forms and data from the 1991 Census, 1997 Census test and the 1999 rehearsal are provided to allow Lockheed Martin to tune their system.

Specific test cases are supplied to Lockheed Martin to gain further assurance that all our rules have been implemented correctly and to verify that our specification is correct.

We have agreed how quality should be measured, both within the service by Lockheed Martin and independently by the ONS.

The Census offices' role in the live environment include:

— monitoring the performance of Lockheed Martin to ensure that they are meeting their contractual obligations;

— receiving reports on throughput and quality. These are reviewed daily by our management team at Widnes;

— we will introduce processes to check independently the quality of the data and images received;

— we will resolve queries where Lockheed Martin experts require the Census offices' judgment or where we have specifically specified that certain cases should be referred;

— as capture and coding progresses omissions from indexes may be discovered (e.g. a new term for an occupation), we update our master indexes accordingly and will provide Lockheed Martin with copies.

The security and confidentiality of census information is critical. We have worked with Lockheed Martin and ICL to introduce measures that will protect the information that we gather. We have involved independent security experts in this process (HEDRA), and will continually monitor procedures until the service is ended.

Conclusion

The Census has moved from a paper-based environment in 1991 to an automated image-based system in 2001. This has been through the use of new technologies to automate and streamline the processes.

This has enabled the Census offices to process more data than in 1991.

Image and microfilm outputs have eliminated the need for costly paper storage.

Production of microfilm from the digital images has speeded up the creation of the archival copy of the Census.

Conservación digital: el censo británico de 2001

Amande Blunden

La elaboración de un censo de población es una tarea ingente; de hecho, es la mayor labor que puede realizar un país en tiempos de paz.

La planificación del censo británico de 2001 empezó ya en 1992, incluso antes de que finalizara el trabajo sobre el censo de 1991. Durante más de tres años se realizó una investigación detallada sobre la mejor manera de llevar a cabo la operación. Desde el principio se quiso aprovechar los últimos cambios en la tecnología de gestión de documentos, a fin de aumentar la eficacia operativa. Ya en 1995 se diseñó un sistema prototipo que permitió validar el concepto para la utilización de la tecnología. A ello siguió un proceso de licitación de la Unión Europea que culminó con la adjudicación del contrato en 1998. En septiembre de 1999 se inició un ensayo general, previéndose la fecha de inicio en abril de 2001.

La operación global se dividió esencialmente en tres partes distintas:

La primera fue la impresión de los formularios. Éstos se diseñaron específicamente para ser escaneados, leyéndose los datos con tecnologías de lectura óptica de caracteres alfanuméricos y de lectura de marcas ópticas (casillas de punteo). En total se necesitaron 34 tipos de formularios distintos, con un total de más de 30 millones de formularios, lo que supuso más de 400 millones de páginas impresas.

En un segundo momento se abordó el almacenamiento de los formularios, así como las necesidades de mano de obra para manejar la cantidad de documentos utilizados. Se acondicionó un almacén y se adquirió un nuevo centro de tratamiento para manejar el gran número de documentos utilizados. Los formularios se trasladaron diariamente del almacén al centro de tratamiento para el escaneo y la extracción de datos.

El último elemento fue el tratamiento de los formularios y la creación de imágenes digitales y analógicas. Este proceso completo debía durar un año y en ese tiempo se escanearon y leyeron casi 300 millones de páginas. Ello fue realizado por siete equipos, cada uno equipado con tres escáneres de gran volumen y un escáner de poco volumen para escaneados excepcionales. Los datos se extrajeron de la imagen escaneada y se almacenaron en sistemas RAID, mientras se verificaba la extracción. Los 300 millones de imágenes se pasaron entonces a un microfilm de 16 milímetros utilizando doce grabadores de archivo digitales durante un período de seis meses. Como el microfilm puede verse y escanearse fácilmente, serán estas imágenes las que se utilicen para reconstruir los formularios una vez estén disponibles al público, lo que conforme a la legislación británica será el 1 de enero de 2102.

Las ventajas de este procedimiento son numerosas: el tratamiento se realiza íntegramente, incluso por lo que respecta a la delicada cuestión de la codificación de las preguntas, que antes sólo se trataban al 10 %, y ello a pesar de que el formulario de 2001 es más largo que el de 1991; la creación de imágenes suprimió los gastos generales asociados al almacenamiento de papel a largo plazo; y se redujo el espacio de almacenamiento necesario de más de 60 000 metros lineales de estanterías a poco menos de 1 000 metros cuadrados de armarios de microfilms.

El censo británico se realizó el 29 de abril de 2001 y el tratamiento empezó en junio. El objetivo actual es finalizar en mayo de 2002 para tener los principales resultados en marzo de 2003.

Digitale Sicherung: Die britische Volkszählung 2001

Amande Blunden

Die Durchführung einer Volkszählung ist eine gewaltige Aufgabe. Es handelt sich dabei um die größte Aktion überhaupt, die ein Land in Friedenszeiten durchführen kann.

Die Planungen für den britischen Zensus 2001 liefen bereits 1992 an, also noch vor dem Abschluss der Arbeiten zur Volkszählung von 1991. Mehr als drei Jahre lang wurde detailliert erforscht, wie die Maßnahme am besten durchgeführt werden kann. Von Anfang an bestand der Wunsch, aus Gründen der Effektivierung die jüngsten Veränderungen in der Technologie der Dokumentenverwaltung zu nutzen. Bereits 1995 erfolgte die Entwicklung des Prototyps eines Systems, mit dem das Konzept für die einzusetzende Technologie nachgewiesen wurde. Im Anschluss an ein EU-weites Ausschreibungsverfahren wurde schließlich im Jahre 1998 der Auftrag erteilt. Eine Gesamtprobe begann im September 1999, und der eigentliche Zensus wurde im April 2001 eingeleitet.

Im Wesentlichen wurde die Maßnahme in drei Komponenten unterteilt:

Zunächst ging es um den Druck der Formulare. Die Vordrucke wurden so gestaltet, dass eine Auswertung der Einträge mit Hilfe von Technologien der alphanumerischen optischen Zeichenerkennung sowie der Erkennung von Ankreuzkästchen möglich war. Insgesamt wurden 34 verschiedene Formulararten benötigt, so dass mehr als 30 Millionen Formulare mit über 400 Millionen Blättern zu drucken waren.

Das zweite Element war die Lagerung der Formulare sowie der Arbeitskräftebedarf für die Bearbeitung der Riesenmenge von Dokumenten. Hierzu wurde ein Lagerhaus renoviert, und es wurde eine neues Bearbeitungszentrum für die Dokumente erworben. Jeden Tag wurden dann Formulare vom Lager zum Bearbeitungszentrum gebracht, wo das Scannen und das Datenauslesen erfolgte.

Letzte Komponente war die Verarbeitung der Formulare und die Erstellung digitaler und analoger Faksimiles. Dieser gesamte Prozess dauerte ein Jahr, und in dieser Zeit wurden nahezu 300 Millionen Seiten gescannt und gelesen. Dazu waren sieben Teams im Einsatz, die jeweils mit drei Dokumentenscannern für große Mengen und einem Spezialscanner für Abweichungen ausgestattet waren. Die Daten wurden aus den gescannten Faksimiles extrahiert, die während der Datenprüfung auf RAID gespeichert waren. Anschließend wurden alle 300 Millionen Faksimiles innerhalb von sechs Monaten unter Verwendung von zwölf digitalen Archive Writern auf 16-mm-Mikrofilm kopiert. Da sich Mikrofilm leicht auswerten und scannen lässt, werden diese Bilder dann zur Rekonstruktion der Formulare verwendet, wenn sie der Öffentlichkeit zur Einsicht bereitzustellen sind, was nach britischem Recht am 1. Januar 2102 der Fall sein wird.

Der Nutzen einer derartigen Arbeitsweise war immens: Die Verarbeitung konnte nunmehr zu 100 % erfolgen, auch bei den schwer zu codierenden Fragen, die zuvor nur eine Verarbeitungsrate von 10 % aufwiesen, und das, obwohl das Formular von 2001 länger war als das von 1991. Durch die Bildsynthese entfielen die indirekten Kosten für die Langzeitaufbewahrung von Papier, und der Lagerraumbedarf wurde von mehr als 60 Regal-Kilometern auf lediglich 28 m^2 Mikrofilmschrankplatz reduziert.

Die britische Volkszählung fand am 29. April 2001 statt. Die Verarbeitung begann im Juni 2001 und soll nach derzeitigem Stand im Mai 2002 abgeschlossen sein, so dass die ersten Hauptergebnisse im März 2003 zur Verfügung stehen.

Conservation numérique: le recensement de 2001 au Royaume-Uni

Amande Blunden

Procéder au recensement démographique est une tâche d'ampleur considérable: en fait, il s'agit du plus vaste exercice qu'un pays puisse réaliser.

Les préparatifs du recensement de 2001 au Royaume-Uni remontent à 1992, voire avant la fin des travaux sur le recensement de 1991. Pendant trois ans, des études précises ont été menées pour voir quelle était la meilleure manière de conduire l'opération. Dès le départ, l'intention était de tirer parti des toutes dernières évolutions des technologies de gestion documentaire afin d'améliorer l'efficacité opérationnelle. Dès 1995, un prototype avait été conçu, qui avait permis de valider le concept pour la technologie à appliquer. Une procédure de passation des marchés dans l'UE s'en est suivie, qui a débouché sur la signature de l'ensemble des contrats en 1998. Une répétition intégrale a débuté en septembre 1999 pour une date de lancement prévue en avril 2001.

Les opérations ont été subdivisées en trois séquences différentes:

- La première a consisté à imprimer les formulaires. Ceux-ci ont été conçus spécialement pour être scannés, la lecture des données les concernant faisant appel aux techniques à caractères optiques alphanumériques et de marquage optique (cases à cocher). Au total, il a fallu mettre au point 34 types de questionnaires différents, pour un total de plus de 30 millions de formulaires représentant plus de 400 millions de pages imprimées.
- Les besoins d'entreposage, de stockage des formulaires et de main-d'œuvre nécessaires pour traiter la quantité de documents concernée ont été étudiés dans un deuxième temps. Un entrepôt a été rénové, et l'on a investi dans un nouveau centre de traitement destiné spécialement à la manipulation de ce gigantesque volume de documents. Les formulaires étaient transmis tous les jours depuis l'entrepôt au centre de traitement pour y être scannés et en extraire les données.
- Le traitement des formulaires et la création d'images numériques et analogiques a constitué la dernière étape. L'opération complète devait durer un an, et presque 300 millions de pages ont été scannées et analysées durant cette période. Sept équipes, dotées chacune de trois scanners pour gros volumes et d'un scanner d'appoint pour les applications exceptionnelles, ont effectué le travail. Les données étaient extraites des images scannées qui étaient mémorisées sur des systèmes RAID, tandis que l'on vérifiait l'extraction. Les 300 millions d'images ont été ensuite converties dans leur intégralité sur des microfilms 16 mm, en utilisant une douzaine de traceurs de films numériques pendant une période de six mois. Les microfilms pouvant être facilement visualisés et scannés, ce sont ces images qui seront utilisées pour recréer les formulaires lorsque ceux-ci pourront être consultés par le public, c'est-à-dire à partir du 1er janvier 2102, conformément à la législation britannique.

Les avantages d'une telle procédure sont nombreux: le traitement est achevé intégralement, même en ce qui concerne la délicate question de la codification des questions, qui n'avait été traitée qu'à un taux de 10 % au même stade. Cela en dépit du fait que le formulaire 2001 est plus long que celui de 1991! La production d'images permet de supprimer les frais liés au stockage permanent d'archives papier et de ramener les besoins en espace de stockage de près de 60 000 mètres linéaires de rayonnage à un peu moins de 1 000 m^2 d'armoires à microfilms.

Le recensement a été organisé le 29 avril 2001 et les opérations de traitement ont débuté en juin. Elles devraient être achevées en mai 2002, les principaux résultats devant être connus d'ici à mars 2003.

Towards a modern archiving system for the European Commission

Frank Brady

Frank Brady

Frank Brady has worked in the field of printing and documentation for 30 years. After 12 years in private industry in the United Kingdom, he came to the European Commission in 1982. He has a vast experience in this field. While he was head of the printing division and always with an eye for the environment and improving efficiency, he reduced the Commission's internal printing runs from 1 billion pages (A4) per year to 400 million pages. This was done by a better use of the Commission's intranet and a rationalisation of procedures. In 1999 he was appointed Head of the Simplification and Modernisation of Archive Systems and the Historical Archives Unit of the European Commission. His mandate here is to design and implement a Commission-wide electronic-records management system that will cover 25 000 officials, in 11 languages and over 60 buildings.
He has a Bachelor of Arts degree in politics and economics.

The mandate

Simplification and modernisation of the filing system

The Commission does not have a standard filing system applicable in every directorate-general. Having a suitable modern filing procedure is a matter of the utmost importance, firstly to ensure the effectiveness and coherence of the day-to-day work of all those concerned and secondly to safeguard the institution's records and thereby guarantee effective access to the documents held by the Commission.

How effectively the Commission meets its present and future obligations in terms of openness and public access to documents very much depends on the filing systems it uses.

Context

One of the important objectives of the reform of the European Commission is 'simplification and modernisation'. One area of specific importance is how we handle information stored in our documents. Not only has the European Parliament been critical of our treatment of documents, in fact the establishment of a modern archive system was one of the conditions for the budgetary discharge of 1999. There is a real need for modernising and simplifying rules and guidance in this area. To this end, we have started a project which we call Action 9 dealing with the treatment and archiving of information in the Commission.

Current situation

The current situation is quite complex. Although new archiving rules were adopted in 1986, these have not been applied everywhere and the last Commission-wide filing plan dates from 1972. At the moment, many directorates-general treat their documents in their own fashion and as such each DG has different rules, different standards and many different levels of competence in this field. Given the scope of the new transparency and access to documents regulations, we need to have a standardised approach that will facilitate the finding and retrieval of documents. Most services do have a filing plan of some description but they at best can be considered 'ad hoc'.

Equally, the Commission has many information sources and the formats used are as equally diverse. Many services have developed their own IT-based information systems for their own specific needs, some of which are interoperable and many which are not. In order to satisfy the new access to document regulations (the Commission's 'Freedom of Information Act' — FOI Regulation (EC) No 1049/2001 of the European Parliament and of the Council of 30 May 2001 regarding public access to the European Parliament, Council and Commission documents), this is a situation which must change and is therefore one the mainstays of Action 9.

Objectives

With the support of our hierarchy, our first task was to set up a legal framework endorsed by President Prodi, Vice-President Kinnock and the college itself. This was in the form of communication and decision, which is both a description of the whys hows, whens and wherefores as well as a legal framework within which we will work. We need to be able to preserve our present and past documents in a way that ensures the accuracy of the document, that is proof of our activities, that these documents are accessible (where allowed) not only internally but that this access is enabled in such a way to ensure the citizens of Europe their rights vis-à-vis documents access.

One of our first tasks was to set up a temporary working group which we called Tafdom (taskforce for document management). This taskforce comprised staff from many directorates-general to scrutinise which areas we needed to examine. This taskforce looked at many factors, legal issues, best practice, e-mail, IT issues, archiving, specifying what is a document, etc.

One of the approaches decided was to set up a permanent network of document management officers (DMO) whose main task will be to make sure as many as possible of our systems are standardised and interoperable. One of his/her first tasks will be to establish a typology of documents that already exist. The DMO will oversee the document administration centres (CA, in French) that most DG's already have in some form or another, to ensure that our documents are registered, attributed and archived in a professional and modern manner as befits the European Commission.

One of the DMO's tasks will be an evaluation of what is actually happening here in the Commission. All documents for example will be required to have a minimum set of metadata and a unique identifier in order to aid identification, registering, filing and archiving of each document. Some of the systems in the Commission already comply with these standards, but many do not and will have to be brought into line. The new global system will encompass the whole life cycle of documents and rather than totally re-inventing the wheel, will hopefully cover the many different supports that are used in-house.

There will also be a high-level committee at Director-General level, whose job will be to provide the political engine and direction for the whole project. Their role will be to ensure the progress of the project and to oversee the whole project. This hierarchical support is one of the most crucial guarantees that a project can have.

We will set up training mechanisms that will target all officials at many different levels using many diverse media. These will include traditional training such as courses in classrooms, leaflets and brochures all the way through to modern web-based interactive systems, FAQs etc. Our first test web pages are up and running and are being evaluated.

Evidently, at the moment we use both electronic and paper records. The use of paper is inevitable but our aim is to bring online more and more electronic document management. However, the rules will define the standard for document registration whatever the support. The paperless office is probably difficult to achieve but that is definitely the direction that the communication and the decision will take us.

This is a long-term project in line with the aims of the e-Commission. (The e-Commission has three main tenets, the first of which, directly related to this project, is 'modernising the Commission's internal administration and management'. The other two deal with relations with the public and other administrations and authorities). There are still many unknown factors especially in the archiving of documents i.e. web pages, databases etc. Equally, which is the best long-term electronic standard. Could this standard be guaranteed in the way that paper can?

Outputs

What are the main practical objectives of Action 9? A decision and communication were the first steps towards a standardisation of the rules and practices for document administration. These will be rules, which set the context for a long-term, standardised filing and archive system for the whole house. These rules are designed to ensure that they can be applied to any document whether electronic or paper and that there will be certain minimum standards applied no matter in which system this document is held. The decision and communication, which were the result of a long and often difficult consultation process with many departments, are the foundation to our project. In the initial phase we looked at many aspects. We looked at the 'best' practices in Member States and in-house. We had to consider our legal obligations as well as the practicalities of such a diverse multilingual environment. We looked at organisational procedures, operational procedures and the whole administrative process that would enable us to complete our tasks. This included the very thorny issue of resources which was much complicated by the different states of progress (or lack of) of good document management practice in the many services of the commission. We also outlined our recognition of the need for information in its many forms including

training, manuals and a general need to raise the awareness of staff in the whole process. There will be a new manual explaining all the procedures necessary for good document management from initial registration through to documents being archived in the historical archives of the Commission.

To give them more weight, the regulations on document management were attached to our internal rules of procedure.

Thus, the decision and communication we prepared are consensual documents even though there were some disagreements. We have the full backing of our hierarchy right up to the President and the Vice-President. This opportunity to have a legal framework gives us the necessary weight to push the project through. The project phase will be steered by a committee of directors-general, chaired subsequently by the heads of the most advanced services in the field. The secretariat of this high-ranking group is provided by my service.

What's new

Basically, four new tenets are being re-enforced, if not being introduced.

(a) There will be, for the first time, a house-wide structure to support the correct management of document management. The Secretariat-General of the Commission will be the driving motor and the guide. There will be regular meetings with the services of the commission via the DMOs, whose task will be to harmonise our treatment of documents.

(b) An annual rolling plan will be produced to chart the progress of the project and to ensure the best use of resources available.

(c) Implementation support — a Commission-wide filing plan will be produced based upon the activities of the house. Training will be given in many forms and to different levels of users online information and extensive manuals will support this. An intranet web site is launched to support all activities.

(d) Commonality — a dictionary-like referential defines common nomenclature as the state of sharing features or attributes, this is exactly what our registration system shall do. Although we recognise that each service has it's own specific wants and needs, there will be, in fact there has to be, this interoperability between registration data and systems.

Assumptions and risks

In a public administration with tasks as diverse as the European Commission, many assumptions have been made in the development of Action 9. Directorates-general are at different stages when it comes to where, how and when they can implement certain aspects of the plan. Nevertheless, we have assumed that our standardisation approach is something that is required by all directorates-general and that in the long term it will benefit the individual directorates-general and thus the house itself.

The risks involved are many and can be broken down into three main areas of known risks, predictable risks and unknown risks.

Known risks include the management of the project itself (especially given the differences of opinion as to what is best, how resources should be used etc. of some of the main actors, i.e. the 'our IT' directorate and the IRMs. These are the IT 'leaders' in each directorate-general). These actors might want to defend their present systems and there may well be resistance from some of the departments that are a long way from our overall vision as to what is or is not 'best' practice. There is a real risk that this will cause major difficulties for Action 9, we must be open and able to listen but on the other hand when necessary, be firm. Unfortunately, not everybody's vision of Action 9 is the same and these disagreements are a potential for delays in the project. The project leader must have the political support to ensure that progress is possible. The complexity of this plan is in itself a risk and the planning stage though important must not become so bogged down in detail that nothing happens. The IRMs (an internal resource management system which is used to

report the use of financial and human resources) example is a good one of how a project should advance and how their expertise and help will be invaluable in the implementation of this project.

Other predictable risks include:

(a) the need for constant hierarchical support — without this, Action 9 will fail;

(b) Action 9 is a good idea which must be sold to all actors involved, otherwise we will meet much resistance;

(c) project staffing — the core group is too small and in certain areas lacks the skills necessary to complete these tasks. Appropriate expertise must be brought or bought in. To this end, outside tenders are being prepared for outside contractors and in the long term the Commission will organise a competition for archivists;

(d) time — the project should not be rushed but must be given a realistic deadline which will deliver a product that satisfies all the requirements of the decision and the communication.

Summary of Action 9

Action 9 of the PAI of October 2000 sets out the framework for a system of document archiving and registry to allow the retrieval of any document tied to the official mission of the Commission, whatever the support may be. It will facilitate the Commission's commitment to transparency, conforming with present rules for public access to documents and in accordance with Action 78 of the White Paper which ensures the establishment of minimum standards for internal control in many diverse areas.

This project can be viewed as the following three groups (some of these can be seen as sub-projects and can be classified as such.)

(1) Organisational:

- DMOs (document management officers)
- CAD (document administration centres)
- implementation teams
- the action plan itself.

(2) Operational:

- manual
- filing plan
- training
- definition of systems
- long-term archiving
- information to actors
- support – helpdesk — network.

(3) Administrative:

- decision and communication
- calls for tender
- competition for archivists
- transfer rules.

Conclusion

This is an ambitious project that for many reasons as stated above must succeed. A body such as the European Commission is a large and complex organisation. This complexity however cannot enable it to disregard its duty vis-à-vis the sound organisation of its document management. From the initial production until final archiving, this organisation is a fundamental aspect of a good administration. If the correct rules are put into place and the project is well managed there is no reason why the European citizen cannot have confidence in this correct application of document administration.

Modernización del archivo y la gestión documental en la Comisión Europea

Frank Brady

Como la mayoría de las administraciones públicas de todo el mundo, la Comisión Europea lleva tiempo sintiendo la necesidad de emprender una reforma. La reforma administrativa interna se convirtió en una prioridad política de la Comisión Prodi. Desde entonces, en virtud de la simplificación y la modernización, casi todos los procedimientos internos de la Comisión han sido objeto de una evaluación crítica, y la mayoría han cambiado o están cambiando. Desde el código de conducta a la auditoría financiera, se han introducido procedimientos modernos y sencillos, y el trabajo de la Comisión es ahora mucho más transparente.

Todos estos distintos ámbitos de reforma figuran en documentos: la información se comunica mediante documentos, los resultados se almacenan en documentos, etc. La información es probablemente el bien más valioso que produce la Comisión. Para garantizar la transparencia dentro de la Comisión y con respecto a los ciudadanos europeos, la modernización y la simplificación de la gestión de documentos en la Comisión se ha convertido en uno de los pilares importantes de la reforma de la Comisión.

El primer paso fue la instauración del marco jurídico. La política de «libertad de información» de la Comisión se vio reforzada en el Tratado de Amsterdam, y en diciembre de 2001 entró en vigor una nueva directiva. El acceso de los ciudadanos europeos a los documentos se ve garantizado ahora por una base jurídica estable y normas claras. Un registro de documentos proporcionará acceso central a todo el mundo.

El marco organizativo y estructural de la gestión de documentos quedó fijado por una decisión y una comunicación de enero de 2002. Estos textos fueron elaborados por mi servicio. Mientras que la decisión establece normas estrictas sobre cómo garantizar una gestión adecuada de los documentos, la comunicación esboza los trabajos futuros.

La comunicación establece seis áreas principales para los trabajos futuros:

- definición de los principios horizontales aplicables a todos los servicios de la Comisión;
- introducción de la estructura organizativa necesaria;
- definición de las normas horizontales y de las disposiciones de aplicación;
- definición, para cada Dirección General o servicio equivalente, de las normas y procedimientos específicos de sus ámbitos de actividad;
- introducción de herramientas de tecnologías de la información (TI) compartidas y específicas;
- medidas auxiliares, formación y apoyo.

La estructura organizativa se basará en una red de personas responsables de la gestión de documentos en su propia Dirección General. Estos funcionarios gestores de documentos (FGD) garantizarán que los principios y la tecnología estén armonizados y sean utilizables en toda la Comisión Europea. La interoperabilidad permitirá el acceso central a todos los documentos de la Comisión. Obviamente, la clave del proyecto será la TI. Nuestro objetivo es poder elaborar un sistema de gestión de todo el ciclo de vida de los documentos, gestionado por la TI, donde toda la información, documentos, ficheros, bases de datos, páginas Internet, etc. existan en formato digital, lo que no sólo proporcionará información, sino una herramienta de trabajo desde la creación inicial de la información hasta el archivo digital a largo plazo. Esto no sólo garantizará que conservemos la memoria de la institución como herramienta de trabajo para los funcionarios de la Comisión, sino también como registro de documentos y archivo a largo plazo para el público en general.

Los documentos son el principal vehículo de transporte de la información. El nuevo sistema gestionará cualquier clase de información. Nuestro trabajo estará regido por varios principios:

- La gestión de la información no se limita a una sola disciplina, sino que incluye programas, bibliotecas, gestión de archivos y otras actividades de la organización.
- La información es un recurso colectivo que debe compartirse y enriquecerse.
- Hacer las cosas bien una sola vez y reexplotarlas.
- Trabajar en el momento más racional, en la fase más temprana del proceso.
- El proceso de gestión de registros/documentos/información es lo más invisible y transparente posible para el usuario.
- Se mantiene la integridad (totalidad, exactitud y fiabilidad) del objeto original.

Se trata de un gran proyecto, pero la Comisión Europea no va a reinventar la rueda. Nos beneficiaremos de la experiencia existente en todo el mundo. Aún no hay una solución única, pero estamos convencidos de que lograremos un sistema moderno y sencillo para gestionar la información de la Comisión a fin de hacer nuestro trabajo lo más transparente posible.

Für eine Modernisierung der Dokumentenverwaltung und Archivierung bei den Dienstellen der Europäischen Kommission

Frank Brady

Wie die meisten anderen öffentlichen Verwaltungen in aller Welt verspürt auch die Europäische Kommission seit geraumer Zeit einen Reformbedarf. Die interne Verwaltungsreform wurde zu einem politischen Schwerpunkt der Prodi-Kommission. Seither wurden nahezu alle internen Verfahren der Kommission unter der Maßgabe der Vereinfachung und Modernisierung einer kritischen Bewertung unterzogen, und die meisten wurden oder werden derzeit geändert. Vom Verhaltenskodex bis hin zur Finanzprüfung wurden moderne und einfache Verfahren eingeführt, und die Arbeit der Kommission ist weitaus transparenter geworden.

Alle diese verschiedenen Reformbereiche werden in Dokumenten festgehalten. So werden Informationen durch Dokumente weitergegeben, Ergebnisse werden in Dokumenten gespeichert usw. Informationen sind wahrscheinlich das wertvollste Gut, das die Kommission hervorbringt. Mit dem Ziel, Transparenz innerhalb der Kommission und gegenüber dem europäischen Bürger zu gewährleisten, wurde die Modernisierung und Vereinfachung der Verwaltung von Dokumenten in der Kommission zu einer wichtigen Säule der Kommissionsreform.

Als erster Schritt erfolgte die Errichtung des entsprechenden Rechtsrahmens. Die Kommissionspolitik der Informationsfreiheit wurde im Vertrag von Amsterdam bekräftigt, und im Dezember 2001 trat eine neue Richtlinie in Kraft. Nunmehr ist für den europäischen Bürger der Zugriff auf Dokumente durch eine stabile Rechtsgrundlage und eindeutige Anwendungsbestimmungen gewährleistet. Dabei dient ein Dokumentenregister als zentraler Zugang für alle.

Der organisatorische und strukturelle Rahmen der Dokumentenverwaltung wurde mit einem Beschluss und einer Mitteilung im Januar 2002 abgesteckt. Beide Texte wurden von meiner Dienststelle erarbeitet. Während der Beschluss strenge Regeln für die Sicherung einer angemessenen Verwaltung von Dokumenten enthält, vermittelt die Mitteilung einen Blick auf die künftigen Arbeiten.

In der Mitteilung werden sechs anstehende Arbeitsbereiche aufgeführt:

- Festlegung der auf alle Kommissionsdienststellen anwendbaren horizontalen Grundsätze,
- Einführung der erforderlichen Organisationsstruktur,
- Festlegung der horizontalen Regeln und Anwendungsmodalitäten,
- Festlegung der für die jeweiligen Aufgabenfelder spezifischen Vorschriften und Verfahren durch die einzelnen Generaldirektionen oder gleichgestellten Dienste,
- Einführung gemeinsamer und spezifischer IT-Tools,
- begleitende Maßnahmen, Schulung und Support.

Die Organisationsstruktur wird auf einem Netz von Verantwortlichen für die Dokumentenverwaltung in den jeweiligen Generaldirektionen beruhen. Diese „Beauftragten für die Verwaltung von Dokumenten“ werden dafür sorgen, dass die entsprechenden Grundsätze und Technologien harmonisiert werden und in der gesamten Europäischen Kommission interoperabel sind. Ausgehend von dieser Interoperabilität wird ein zentraler Zugang zu allen Kommissionsdokumenten möglich sein. Es liegt auf der Hand, dass die Informationstechnik den Schlüssel zum gesamten Projekt darstellt. Unser Ziel besteht darin, ein IT-gestütztes Managementsystem für den gesamten Lebenszyklus von Unterlagen zu errichten, bei dem sämtliche Informationen, Dokumente, Akten, Datenbanken, Internetseiten usw. in digitaler Form existieren, so dass nicht nur Informationen vermittelt werden, sondern auch ein Arbeitsmittel von der ursprünglichen Erstellung der Information bis zur digitalen Langzeitarchivierung bereitsteht. Damit dürfte sichergestellt sein, dass wir das Gedächtnis des Organs nicht nur als Arbeitsmittel für Kommissionsbedienstete, sondern auch als Dokumentenregister und Langzeitarchiv für die Öffentlichkeit bewahren.

Dokumente sind der Grundpfeiler für die Übermittlung von Informationen. Mit dem neuen System werden Informationen aller Art verwaltet. In unserer Arbeit werden wir uns von sieben Grundsätzen leiten lassen:

- Das Informationsmanagement ist nicht auf ein Fachgebiet beschränkt, sondern umfasst Programmbereiche, Bibliotheken, die Verwaltung von Unterlagen sowie weitere Tätigkeitsbereiche der Organisation.
- Informationen sind eine institutionseigene Ressource, die es gemeinsam zu nutzen und zu bereichern gilt.

- Was einmal richtig und gründlich bearbeitet wurde, lässt sich immer wieder verwenden.
- Notwendige Arbeiten sollten so früh wie möglich und dort ausgeführt werden, wo sie am sinnvollsten sind.
- Der Prozess der Verwaltung von Unterlagen/Dokumenten/Informationen sollte für den Benutzer so unsichtbar und unaufdringlich wie möglich ablaufen.
- Die Integrität (Vollständigkeit, Genauigkeit und Zuverlässigkeit) des ursprünglichen Objekts wird bewahrt.

Zwar handelt es sich um ein gewaltiges Vorhaben, doch wird die Europäische Kommission das Rad nicht neu erfinden. Wir werden auf Erfahrungen aus aller Welt zurückgreifen können. Es gibt noch keine allumfassende Lösung, doch sind wir davon überzeugt, dass wir – wenn wir alle Einzelteile zusammengesetzt haben – zu einem modernen und einfachen System für die Verwaltung der Informationen der Kommission gelangen, das unsere Arbeit so transparent wie möglich macht.

Vers une modernisation de la gestion documentaire et de l'archivage à la Commission européenne

Frank Brady

À l'instar des autres administrations publiques de par le monde, la Commission européenne a ressenti la nécessité de réformes il y a quelque temps déjà. La réforme de l'administration interne est devenue une priorité politique de la Commission Prodi. Depuis lors, la quasi-totalité des procédures internes de la Commission ont été soumises à une évaluation critique et la plupart ont été modifiées, ou sont en voie de l'être, en vue de leur simplification et de leur modernisation. Des procédures modernes et simples ont donc été mises en place, depuis le code de conduite jusqu'à l'audit financier, et le travail de la Commission est désormais beaucoup plus transparent.

Tous les domaines de réforme sont consignés dans des documents: l'information est communiquée par des documents, les résultats sont mémorisés dans des documents… L'information est probablement le bien le plus précieux que la Commission produise. La modernisation et la simplification de la gestion documentaire dans cette institution est donc devenue l'un des piliers majeurs de la réforme de la Commission, pour garantir la transparence en interne et vis-à-vis du citoyen européen.

La première étape a été de fixer le cadre juridique. L'action de la Commission en faveur de la liberté d'information a été renforcée dans le traité d'Amsterdam, et une nouvelle directive est entrée en vigueur en décembre 2001. Un fondement juridique stable et des règles d'application claires régissent désormais l'accès du citoyen européen aux documents. Un registre de documents fournira un accès centralisé à chaque individu.

Une décision et une communication en janvier 2002 ont fixé le cadre organisationnel et structurel de la gestion documentaire. Mon service est à l'origine de ces textes. Si la décision fixe des règles strictes sur la manière de garantir une gestion documentaire adéquate, la communication expose une vision des futurs travaux à mener.

La communication définit six domaines majeurs de travail:

- définition des principes horizontaux applicables à tous les services de la Commission,
- introduction de la structure organisationnelle nécessaire,
- définition des règles horizontales et des modalités d'application,
- définition par chaque direction générale ou service équivalent des règles et des procédures propres à ses domaines d'activité,
- introduction d'outils informatiques communs et spécifiques,
- mesures annexes, formation et support.

La structure organisationnelle reposera sur un réseau de personnes chargées de la gestion documentaire dans leur propre direction générale. Ces responsables de la gestion documentaire veilleront à l'harmonisation des principes et des technologies et à leur interopérabilité dans toute la Commission européenne. L'interopérabilité permettra de centraliser l'accès à tous les documents de la Commission. Bien évidemment, les technologies de l'information sont au cœur du projet: notre objectif est d'être en mesure d'élaborer un système de gestion de l'ensemble du cycle de vie des documents, reposant sur les TIC, dans lequel l'ensemble des informations, des documents, des fichiers, des bases de données, des pages web, etc., se présenteront sous forme numérique, ce qui permettra de disposer non seulement d'informations, mais aussi d'un outil de travail, depuis la création originale de l'information jusqu'à son archivage permanent. Ce système devrait donner la garantie que nous conservons la mémoire de l'institution non seulement sous la forme d'un outil de travail pour les responsables de la Commission, mais aussi sous celle d'un registre de documents et d'archives à long terme pour le grand public.

Les documents sont l'épine dorsale de l'acheminement de l'information. Le nouveau système gérera n'importe quel type d'information. Sept principes ont guidé nos travaux:

- la gestion documentaire ne se borne pas à une discipline unique, mais elle implique les domaines de programmation, les bibliothèques, la gestion d'archives et d'autres activités au sein de l'organisation;
- l'information est une ressource interne qui doit être partagée et enrichie;
- ne faire les choses qu'une seule fois, bien les faire et les réexploiter;
- faire le travail là où c'est le plus rationnel, le plus tôt possible dans le processus;
- le processus de gestion d'archives/documents/informations est le plus invisible et le plus transparent possible pour l'utilisateur;
- l'intégrité (exhaustivité, exactitude et fiabilité) de l'objet original est préservée.

Il s'agit là d'un vaste projet, mais la Commission européenne ne va pas réinventer la roue. Nous bénéficierons de l'expérience déjà acquise dans ce domaine de par le monde. Il n'existe pas de panacée, mais nous sommes convaincus que c'est en assemblant différents éléments entre eux que nous parviendrons à concevoir un système moderne et simple pour gérer l'information à la Commission, afin que notre travail soit le plus transparent possible.

Birgit Hansen

Since 1998, Birgit Hansen has held the post as head of the IT department at the Danish State Archives with the responsibility of long-term preservation of electronic archives, systems development and IT operations and IT security. Birgit has been and is currently involved in a number of committees under the Ministry of Culture in Denmark, including the preservation of the digital cultural heritage, the Digitisation Expert Committee and the Steering Committee to establish and implement a common IT strategy covering all institutions under the ministry, including the establishment of common policies for IT infrastructure, IT procurement and IT operations. Birgit is also working closely with the Ministry of IT and Research and the Danish e-government taskforce concerning subjects such as e-government, electronic archiving and the consequences for long term preservation. Prior to her present employment at the Danish State Archives, she was IT executive of the National Survey and Cadastre of Denmark (1993–98) and chief IT consultant at the Danish Technological Institute (1982–93). Birgit holds a Masters degree in electronic engineering and a PhD in the design of cad-systems from the University of Manchester Institute of Science and Technology at which she also conducted research and laboratory work from 1980–82.

Preservation of the digital cultural heritage

Birgit Hansen, Jan Danielsen

The challenge

Is it at all possible in a satisfactory and adequate manner to secure the long-term preservation of data in a broad sense? In connection with e-government, data is not 'just' text documents, but encompasses all types of data including e-mails, web-based multimedia documents including active links to be preserved, extracts of topographical maps, spreadsheets, voice mails, video extracts etc. The preservation of digital data over a long period of time requires a long-term preservation strategy to ensure that all processes, procedures and technologies are determined and tested for its life-cycle validity.

Demands and requirements

When data is created at the authority, special arrangements have to be made to secure that data can be preserved over time. This means that the necessary requirements and demands have to be put forward and complied with before the IT system creating the data is taken into use. The systems must be approved by the Danish State Archives before taken into use by the authorities (see paper by Jan Dalsten Sørensen elsewhere).

The demands have an essential contributory factor in securing that the electronic archive system:

- is capable of preserving data and other relevant information in an authentic manner;
- can create and store data in such a manner that the subsequent long-term preservation at the Danish State Archives can take place without loss of data etc.;
- can be delivered to the State Archive system as independently as possible.

The Danish State Archives set only the minimum possible number of regulations with the sole purpose of securing preservation of electronic archives. The aim is to apply the fewest possible restrictions to the authorities in order to give them ample opportunity to develop their own administrative systems in accordance with their needs.

Guiding principles

The most important guiding principles for the preservation strategy is the following:

- long-term preservation of digital information requires a continued proactive effort. Just 10 years of neglect may be crucial;
- preserve only what is highly valuable;
- paramount to the preservation success is the specification and control of the input to the archive i.e. if we can control the input we can control the output encompassing both the issue of preservation and access;
- make the preservation strategy as technology independent as possible thereby reducing the risks of technological obsolescence;
- keep it simple! Complexity at all levels and in all aspects of the preservation life cycle is a huge risk. Make sure to standardise and simplify as extensibly as possible to enable automation of all processes. Ensure scalability of the strategy.

The preservation strategy in Denmark — a holistic view

The Danish State Archives see preservation of electronic archives in a holistic view shown in Figure 1. The preservation strategy should take into account a multitude of issues, all of which have severe consequences if they are not handled appropriately and seen as connected with the other issues.

Central to the preservation strategy is the choice between the four well-known strategies thinking in all aspects from acquisition to preservation and access:

- the migration or conversion strategy — the chosen strategy at the Danish State Archives as the strategy is the most viable, simple and economic preservation strategy giving ample results over time;
- the emulation strategy — is not chosen due to its technical complexity requiring many highly skilled staff. It is therefore not an economically sound strategy and is judged to be unsustainable over time;
- the museum strategy — is not sustainable over a long time due to the lack of working equipment and the lack of technical expertise to support outdated equipment and software;
- the custodial strategy — is not in accordance with the Danish archival legislation.

It is likewise central to preservation to ensure the continued monitoring of the holdings. We have developed a spot-test concept based on statistical theory and practical experience with longevity of media such as magnetic tape and CD-ROM. Choice of technology and standards like metadata standards, formats and media is likewise paramount to our success.

Strategic political issues such as archival legislation, political goodwill and funding over time is important. These aspects should be seen in close connection with the total cost of ownership (TCO) i.e. the life-cycle cost including workload over time, conversion rates and amounts, investments in technology, education etc. If your choice of strategy is not sufficiently simple, you may not over time be able to acquire and maintain staff with the right skills. The preservation of electronic archives or indeed the digital cultural heritage requires a suitable mix of competent IT specialists and archivists. If you do not have staff who understand the technical issues and who are capable of defining/choosing the appropriate standards, formats etc., you will not be able to preserve your holdings over time. The preservation strategy must be well defined, described and understood at all three levels i.e. the strategic, tactical and operational level.

The preservation strategy should also encompass issues regarding organisation, location and facilities. IT operations and IT security known from the world of IT should be of a very high standard as the preservation of the digital cultural heritage must be seen as a highly important, if not the most important, security issue. These issues include daily operations with back-up routines, installation procedures etc., secluded network with no connection to the Internet, physical access control etc. Quality assurance using the ISO 9000 methods is paramount to the preservation success.

Jan Danielsen

Since 1995, Jan Danielsen has participated in the theoretical and practical handling of electronic archives. Since 1998, he has worked as a special consultant and head of section in the IT department, with the responsibility of the long-term preservation of electronic archives and the development of standards and regulations for delivery to the Danish State Archives.
Jan Danielsen has through the years participated in a number of working groups and committees in the Danish administration, on electronic administration and archiving. Most of his work is available on the homepage of the Ministry of Science, Technology and Innovation.
Jan is co-editor of the journal Arkiv and has published a number of articles about creation, scanning and archiving of information. He is currently a member of the Danish initiative on a national XML standard, the Danish ISO 15849 working group and is working with the Danish digital taskforce in the Ministry of Finance on implementation of digital administration in Denmark. Jan Danielsen has a Master's degree in computer science and geography.

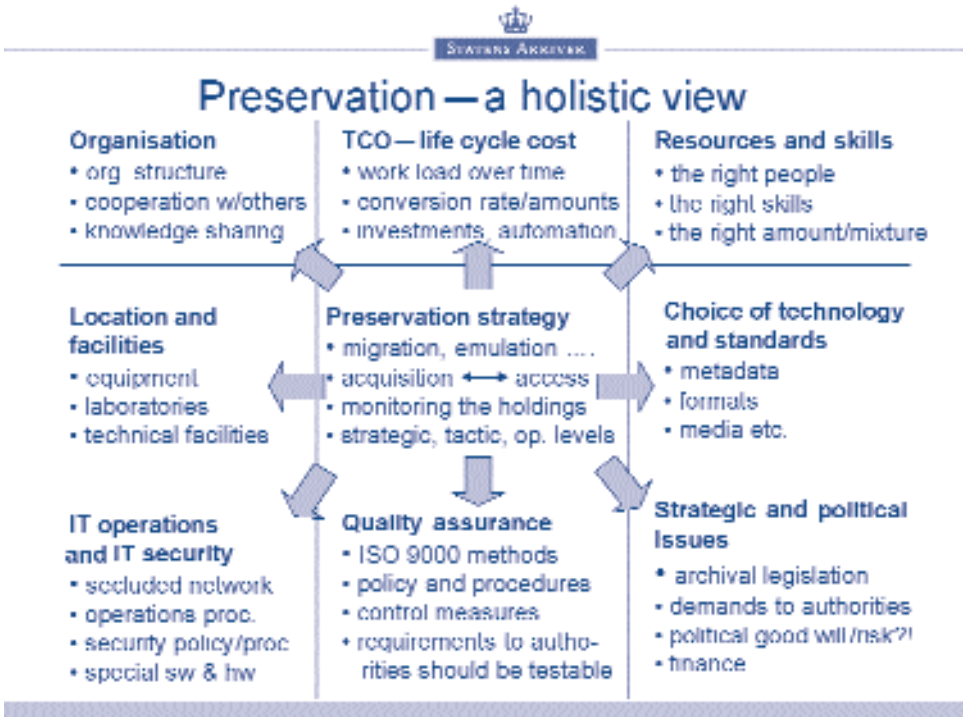


Figure 1. Preservation of electronic archives — a holistic view.

It is important to have policies, procedures and control mechanisms in place. Procedures and instructions must be very precise and well defined, leaving no ambiguity to the staff. Regular follow-up and evaluation of the procedures and instruction in order to refine and improve them is important. Quality assurance procedures will never be perfect but continued gradual improvements are much better than not doing anything at all!

A very important aspect of quality assurance is testing. The requirements to the authorities should be such that you can test the acquisitions automatically. This means that each requirement should be electronically testable i.e. the requirements must be so specific and well defined that you can construct a programme to test the acquisition automatically. It you cannot do that, then don´t put forward the requirement! If all requirements can be tested automatically on acquisition, they can also automatically be converted from one metadata standard and format to another. Scalability is very important. If your procedures or choice of technology are too cumbersome for small volumes of electronic archives, they will never work when e-government takes off and the archives get huge volumes of data to preserve.

Key elements in the Danish preservation strategy

The strategy for long-term preservation of electronic archives consists of regular conversion or migration of both electronic media and formats i.e. conversion from one medium or format to a new one. We want to preserve the content and the semantics of the converted data and not the technology by which they are generated or stored on. Therefore, the formats should:

- be standardised — preferably international standards;
- be widespread i.e. have a great market share and be low cost;
- have a long lifetime;
- secure against loss of data or loss of integrity;
- and be suitable for subsequent electronic search.

Saving only the data and not the technology applies a loss of information when the data is extracted from the original system. This loss of information is compensated for by the documentation delivered with the system, and the metadata created on the time of delivery.

If no formats comply with all the criteria above it may be necessary to appoint a three-tier concept for formats to be complied with: (1) a delivery format; (2) a preservation format; and (3) a distribution format. It is essential to allow as few formats as possible to keep the running cost of conversion as low as possible.

The selected formats

From these criteria the Danish State Archives has selected the following formats:

- text documents: TIFF (presentation) and ISO 8859 Latin 1 (searching)
- video: MPEG-2
- audio: MP3

There are at present no formats for spreadsheet, GIS or CAD which live up to the above criteria. However, up and coming XML standards may be the solution to all of the above types of data. These kinds of data can only partly be stored by converting them into TIFF, thereby saving the visual image but losing any dynamics. The MPEG-2 and MP3 formats are delivery formats, as they do not 'secure against loss of data'. Both formats are based on compression technology that loses data. When continuously converting these formats into new formats, the risk of losing information is high. As a consequence, the Danish State Archives will appoint preservation formats for sound and video, but no one has yet delivered sound or video, so the decision will be taken when needed.

The selected media

The Danish National Archives have chosen CDR as both delivery and preservation media formats The CDR is widely used, requires very little special knowledge, has a long life expectancy, and most importantly, the quality of the media can be tested.

Different formats for different purposes

The different criteria have different impacts on the three different purposes as shown in Figure 2.

- For delivery purposes it is important that the formats are standardised and widely used, so that the delivery can be tested against an available and fixed format description without adding unnecessary cost to the authorities.
- For preservation purposes it is important that the formats are standardised to secure that a conversion tool can be developed at any time, without being dependent on knowledge or assistance from any vendors. The life expectancy is important to minimise the number of necessary conversions. Likewise, securing of integrity is important.
- Distribution is very dependent on the method of distribution, but typically the size of the files and fast search will be essential issues.

Almost identical principles have been set for selecting the storage media.

Delivery of electronic archives

When an authority is required to deliver to the Danish State Archives they have to comply with a certain set of requirements which are specified on the State Archive web site www.sa.dk. Data has to be submitted in a system-independent archiving version with specific requirements to library structure, name giving etc. From the defined metadata specification the delivering authority describes the structure of the IT system and the content in a standardised manner, which unambiguously can be machine-interpreted. The State Archive also requires more general information about the IT system, all delivered on one or more CDRs (two copies) to ensure a self- documenting electronic archive.

The metadata are structured by XML tags, describing the structure within one or more flat files and the connections between the files of the electronic archive. Thereby, we can describe any database that can be transformed into a relational structure. In addition to the normal machine-extractable technical information, some tags have been added to both the field and table descriptions. They contain contextual information about the given table or field, making it possible to understand the function of the field or table without a human being reading through the documentation thus making it easier to make one single efficient standard search and access programme.

The XML tagging of metadata allows authorities (or there consultants) to use standard XML tools to generate and test the syntax of the metadata before they are delivered, thereby heightening the quality of the deliveries to the Danish State Archives.

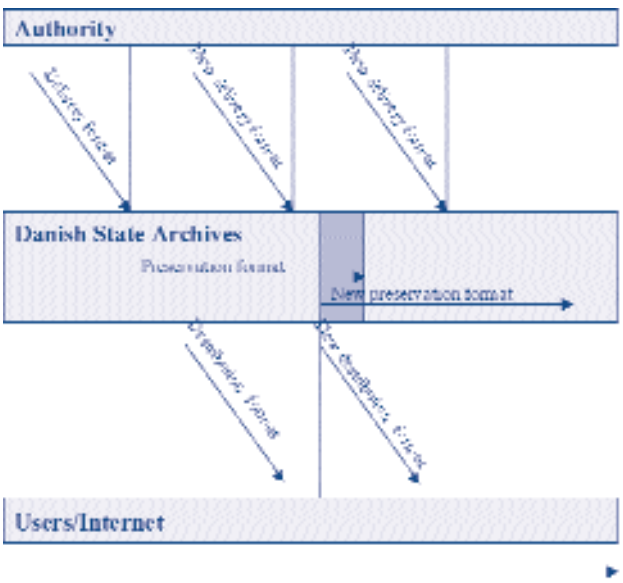


Figure 2. Different formats for different purposes

To support the delivery of electronic archives form the authorities, the Danish State Archives have developed a set of tools making it easier to create a correct delivery. The tool set encompasses a character set analyser and conversion tool, a tool for date conversion, a tool that analyses and removes any separator characters and so on.

Testing CDRs

When the Danish State Archives have received the electronic archival version on a number of CDRs from a public authority, the CDRs are tested using the test tool TEA (test of electronic archives). The TEA is developed by the Danish State Archives as the electronic implementation of the technical delivery requirements.

The TEA automatically tests that all files listed in the metadata are present and that they are named correctly. The next step is to test that the XML syntax is correct and that the semantics is correct. When the metadata is tested, a simple diagram of the delivery is generated to give an overview of the structure of the archive. Data is subsequently tested for data type and validated if there are any limits for the values of a field. The accuracy of this test can be configured separately for each test. Concluding the test is a visual spot check of the attached files. As part of the quality assurance measures the quality of the CDRs are tested to ensure that defective CDRs or CDRs with technical problems that might shorten the lifetime of the preserved CDR are reburned before storage in the archive.

Long-term preservation means regular test and conversion

Having tested and approved a delivery from a public authority, the long-term preservation at the State Archives begin. Apart from the test tool, the TEA and the delivery tools, we have also developed a series of conversion tools and a set of quality-assurance methods and procedures for test, conversion and documentation of electronic archives. As part of a comprehensive set of IT security regulations we have established a secluded workshop for the reception, test and continued conversion of electronic archives. We are presently occupied with media conversion of magnetic tapes (dating back to the 1970s) to the present CDR medium. In the year 2002 the format conversion will be prepared and put into action.

The structure and the description of the delivery is designed so that the electronic archive can be moved from media to media without altering the original tables. Only the surrounding description files will have to be manipulated. The original tables and documents will only be handled if a character set conversion or other file format conversions are needed.

At present, electronic archives delivered before 1998 do not have the necessary metadata and description files. In order to be handled automatically by future conversion or access programmes this information will have to be extracted and created from the paper documentation that followed the original delivery.

The circulars regulating the authorities' handling and delivery of electronic archives to the Danish State Archives are evaluated each year, so that any change in technology or administrative practice can be adapted in the circulars. As a consequence, the TEA is modified and the procedures for testing are equally altered. Thereby, we are convinced that the Danish administration can develop the administration practise using the newest technology without any major risk of losing the ability of preserving documentation for the future.

The preservation over time — a question of IT security and quality assurance

The State Archives are among the more experienced countries in the world with respect to the long-term preservation of digital data. We have created an integrated set of technical and organisational requirements, processes and procedures which ensures a satisfactory preservation and accessibility over time, thereby preserving the digital heritage for the benefit of historic researchers and our descendants.

The presentation concluded with a short demonstration of the test tool TEA, which demonstrated that automatic testing and quality assurance of electronic archives is possible.

Conservación del patrimonio digital

Birgit Hansen, Jan Danielson

¿Es posible, de manera satisfactoria y adecuada, garantizar la conservación a largo plazo de los datos en sentido amplio? Por lo que respecta a la administración electrónica, los datos no son meros documentos de texto, sino que abarcan todos los tipos de datos, incluidos correos electrónicos, documentos multimedia de la red y los vínculos activos que deben mantenerse, extractos de mapas topográficos, hojas de cálculo, correos vocales, extractos de vídeo, etc. La conservación de datos digitales durante mucho tiempo requiere una estrategia de conservación a largo plazo que garantice que todos los procesos, procedimientos y tecnologías se definan y se prueben para asegurar su validez durante todo el ciclo de vida de la información.

Estrategia para la conservación a largo plazo de datos digitales

Cuando se crean los datos, deben establecerse disposiciones especiales para garantizar que éstos puedan conservarse a lo largo del tiempo. Esto significa que deben establecerse y cumplirse los requisitos necesarios *antes* de que empiece a utilizarse el sistema de tecnologías de la información (TI) que crea los datos.

Los requisitos desempeñan una función esencial para garantizar que el sistema de archivo electrónico:

- es capaz de conservar datos y otra información pertinente conservando su autenticidad;
- puede crear y almacenar datos de manera que la conservación posterior a largo plazo en los archivos estatales daneses pueda realizarse sin pérdida de datos;
- sea lo más independiente posible.

La estrategia para la conservación a largo plazo de los archivos electrónicos consiste en la conversión o la migración periódica de los medios electrónicos y de los formatos, es decir, la conversión de un medio o formato a otro nuevo. Queremos conservar el contenido y la semántica de los datos convertidos y no la tecnología mediante la que se generan o donde se almacenan. Por tanto, los formatos deberían:

- estar normalizados, preferiblemente con normas internacionales;
- estar muy extendidos, es decir, tener una gran cuota de mercado y un bajo coste;
- tener un ciclo de vida largo;
- estar protegidos contra la pérdida de datos o la pérdida de integridad;
- ser adecuados para su posterior consulta electrónica.

Los formatos seleccionados

En virtud de estos criterios, los archivos estatales daneses han seleccionado los siguientes formatos:

- documentos de texto: TIFF (presentación) y ASCII (consulta);
- vídeo: MPEG-2;
- audio: MP3.

Actualmente no existen formatos para hojas de cálculo, SIG o CAO que cumplan los criterios anteriormente mencionados. Sin embargo, las próximas normas XML pueden ser la solución para todos los tipos de datos anteriormente mencionados.

Envío de archivos electrónicos

Cuando una autoridad envía a los archivos estatales daneses un archivo, debe cumplir un conjunto de requisitos que se especifican en el sitio de Internet del archivo estatal (http://www.sa.dk). Los datos deben presentarse en una versión independiente del sistema, con unos requisitos específicos respecto a la estructura de biblioteca, nombre, etc. A partir de la especificación definida de los metadatos, la autoridad que envía los datos describe la estructura del sistema de TI y el contenido de manera normalizada, que pueda ser leída por la máquina sin ambigüedades. El archivo estatal también requiere información de carácter general sobre el sistema de TI, todo ello entregado en uno o más CD-R (dos ejemplares) para garantizar la autodocumentación de los archivos electrónicos.

Prueba de los CD-R

Cuando los archivos estatales daneses reciben la versión electrónica para archivo en varios CD-R procedentes de una autoridad pública, estos CD-R se comprueban utilizando la herramienta de prueba TEA (Test of Electronic Archives). La TEA ha sido desarrollada por los archivos estatales daneses como la aplicación electrónica de los requisitos técnicos aplicables al envío de archivos. La TEA comprueba automáticamente que todos los metadatos son correctos, que la estructura y los nombres son correctos, que todos los formatos son correctos, etc. Como fin de la prueba se realiza un control visual de los ficheros adjuntos.

La conservación a largo plazo supone pruebas y conversiones periódicas

Una vez comprobado y aprobado un envío de archivos procedentes de una administración, comienza la conservación a largo plazo en los archivos estatales. Aparte de la herramienta de prueba TEA, también hemos desarrollado una serie de herramientas de conversión y un conjunto de métodos de garantía de calidad y procedimientos para la prueba, la conversión y la documentación de los archivos electrónicos. Como parte de un amplio conjunto de normas de seguridad de TI, hemos creado un taller independiente para la recepción, la prueba y la conversión continua de archivos electrónicos. En la actualidad nos ocupamos de la conversión de cintas magnéticas que datan de la década de los setenta a CD-R. La conversión de formatos se preparará y se pondrá en marcha en el año 2002.

La conservación en el tiempo: una cuestión de seguridad de las TI y de garantía de calidad

Los archivos estatales daneses figuran entre los más experimentados del mundo en cuanto a conservación a largo plazo de datos digitales. Hemos creado un grupo integrado de requisitos, procesos y procedimientos técnicos y organizativos que garantizan una conservación y una accesibilidad satisfactorias en el tiempo, conservando así la herencia digital en beneficio de los investigadores y de las generaciones futuras.

Bewahrung des digitalen Erbes

Birgit Hansen, Jan Danielson

Ist es überhaupt möglich, die Langzeitaufbewahrung von Daten im weitesten Sinne zufrieden stellend und angemessen zu gewährleisten? Im Zusammenhang mit e-Government umfasst der Datenbegriff nicht nur Textdokumente, sondern alle Arten von Daten wie E-Mails, zu konservierende internetbasierte Multimedia-Dokumente einschließlich aktiver Links, Auszüge topografischer Karten, Tabellenkalkulationen, Voicemails, Videoauszüge usw. Zur Aufbewahrung digitaler Daten über einen langen Zeitraum bedarf es einer Langzeitkonservierungsstrategie, die dafür sorgt, dass alle Prozesse, Verfahren und Technologien daraufhin ermittelt und geprüft werden, ob sie für den gesamten Lebenszyklus der Daten nutzbar sind.

Strategie für die Langzeitaufbewahrung von Daten

Bei der Erstellung von Daten in einer Behörde ist mit speziellen Vorkehrungen dafür zu sorgen, dass sie konserviert werden können. Das bedeutet, dass die notwendigen Anforderungen zu stellen und einzuhalten sind, *bevor* das die Daten erstellende IT-System in Betrieb genommen wird.

Die Anforderungen tragen in wesentlichem Maße dazu bei, dass gesichert ist, dass das elektronische Archivierungssystem

- Daten und andere relevante Informationen authentisch konservieren kann,
- Daten so erstellen und speichern kann, dass die anschließende Langzeitaufbewahrung im Dänischen Staatsarchiv ohne Datenverlust u. Ä. erfolgt,
- dem Staatsarchiv so systemunabhängig wie möglich bereitgestellt werden kann.

Die Strategie für die Langzeitaufbewahrung von elektronischem Archivgut besteht aus der regelmäßigen Konvertierung oder Migration von elektronischen Datenträgern und Formaten, d. h. aus der Übertragung von einem Träger oder Format zu einem neuen. Dabei wollen wir den Inhalt und die Bedeutung der konvertierten Daten konservieren und nicht die Technologie, mit der sie erzeugt oder gespeichert wurden. Deshalb sollten die Formate

- genormt sein, vorzugsweise durch internationale Normen,
- weit verbreitet sein, also einen großen Marktanteil besitzen und kostengünstig sein,
- eine lange Lebensdauer haben,
- gegen Datenverlust und Verlust der Datenintegrität geschützt sein und
- für eine anschließende elektronische Recherche geeignet sein.

Die ausgewählten Formate

Anhand dieser Kriterien hat das Dänische Staatsarchiv folgende Formate ausgewählt:

- Textdokumente: TIFF (Darstellung) und ASCII (Suche),
- Video: MPEG-2,
- Audio: MP3.

Derzeit gibt es keine Formate für Tabellenkalkulationen, GIS oder CAD, die den oben genannten Kriterien genügen. Für diese Datenarten könnten jedoch die derzeit entstehenden XML-Standards die Lösung darstellen.

Übergabe von elektronischem Archivgut

Wenn eine Behörde zur Übergabe an das Dänische Staatsarchiv verpflichtet ist, muss sie bestimmte Anforderungen erfüllen, die auf der Website des Staatsarchivs www.sa.dk aufgeführt sind. Die Daten müssen in einer systemunabhängigen Archivierungsversion mit konkreten Anforderungen an Fundstellenstruktur, Benennung usw. übermittelt werden. Anhand der festgelegten Metadatenspezifikation beschreibt die übergebende Behörde die Struktur des IT-Systems und den Inhalt in einer standardisierten Weise, die eindeutig maschinenlesbar ist. Darüber hinaus verlangt das Staatsarchiv noch allgemeinere Angaben zum IT-System, die alle auf einer oder mehreren CD-R (zwei Exemplare) bereitzustellen sind, so dass ein selbstdokumentierendes elektronisches Archiv vorhanden ist.

Prüfung von CD-R

Wenn das Dänische Staatsarchiv von einer staatlichen Behörde die elektronische Archivversion auf einer Reihe von CD-R erhalten hat, werden die CDs mit Hilfe des Tools TEA (Test of Electronic Archives) geprüft. TEA wurde vom Dänischen Staatsarchiv als elektronische Umsetzung der technischen Übergabeanforderungen entwickelt. Das Tool kontrolliert automatisch, dass alle Metadaten, Struktur und Namen sowie alle Formate usw. korrekt sind. Zum Abschluss erfolgt eine visuelle Stichprobe der angefügten Dateien.

Langzeitaufbewahrung bedeutet regelmäßige Prüfung und Konvertierung

Nachdem die Lieferung von einer staatlichen Behörde geprüft und genehmigt worden ist, beginnt die Langzeitaufbewahrung im Staatsarchiv. Neben dem Prüftool TEA haben wir auch eine Reihe von Konvertierungstools sowie Qualitätssicherungsverfahren und Prozeduren zur Prüfung, Konvertierung und Dokumentation von elektronischem Archivgut entwickelt. Im Rahmen umfangreicher IT-Sicherheitsregelungen haben wir eine gesonderte Stätte für den Empfang, die Prüfung und die ständige Konvertierung von elektronischem Archivgut eingerichtet. Derzeit befassen wir uns mit dem Umkopieren von Magnetbändern (aus den 70er Jahren) auf die jetzt aktuellen CD-R. Im Jahr 2002 wird die Formatkonvertierung vorbereitet und umgesetzt.

Langfristige Konservierung – eine Frage der IT-Sicherheit und der Qualitätssicherung

Das Dänische Staatsarchiv gehört zu den Archiven der Welt, die bereits größere Erfahrungen mit der Langzeitaufbewahrung von digitalen Daten gesammelt haben. Wir haben einen integrierten Komplex technischer und organisatorischer Anforderungen, Prozesse und Verfahren eingeführt, die für eine zufrieden stellende langfristige Konservierung und Zugänglichkeit sorgen, so dass das digitale Erbe im Interesse der Geschichtsforscher und unserer Nachfahren bewahrt werden kann.

Conservation du patrimoine numérique

Birgit Hansen, Jan Danielson

La conservation à long terme de données (au sens large), de façon satisfaisante et appropriée, est-elle tout simplement possible? Dans le cadre de l'administration électronique, les données ne sont pas simplement des textes, mais peuvent être également d'un autre type: e-mails, documents multimédias du web et leurs liens actifs à conserver, extraits de relevés topographiques, feuilles de calcul, messages vocaux, extraits vidéo, etc. La conservation d'informations numériques sur une longue période implique une stratégie de préservation à long terme, qui garantisse que tous les processus, les procédures et les technologies soient définis et testés pour assurer leur validité sur tout le cycle de vie des informations.

Stratégie de conservation à long terme des données numériques

Lorsque des données sont créées dans un organisme créateur, des dispositions particulières doivent être prises pour garantir leur conservation dans le temps. Cela implique de proposer un cahier des charges et de veiller à son respect *avant* que le système informatique de création des données soit utilisé.

Le cahier des charges joue un rôle essentiel pour veiller à ce que le système d'archives électroniques:

- soit capable de conserver des données et toute autre information pertinente en préservant leur authenticité,
- puisse créer et stocker des données de façon qu'elles puissent être conservées ensuite à long terme aux Archives nationales danoises sans aucune perte,
- soit aussi indépendant que possible des systèmes.

La stratégie de conservation à long terme des archives électroniques consiste à convertir ou à transférer périodiquement des supports électroniques et des formats, c'est-à-dire à réaliser des transferts d'un support ou d'un format à un autre. Ce qu'il nous faut préserver, c'est le contenu et la sémantique des données converties, non la technologie qui a permis de les générer ou de les stocker. Aussi les formats doivent-ils:

- être uniformisés (pour satisfaire de préférence aux normes internationales),
- être de grande diffusion, c'est-à-dire disposer d'une grande part de marché et être d'un faible coût,
- présenter une grande longévité,
- être sécurisés contre la perte de données ou d'intégrité,
- se prêter à une consultation électronique ultérieure.

Formats sélectionnés

À partir des critères ci-dessus, les Archives nationales danoises ont sélectionné les formats suivants:

- documents textuels: TIFF (présentation) et ASCII (consultation),
- vidéos: MPEG-2,
- documents sonores: MP3.

Actuellement, aucun format ne respecte les critères ci-dessus en ce qui concerne les feuilles de calcul, les SIG et la CAO. Toutefois, les normes XML à venir pourraient être la solution pour tous les types de données ci-dessus.

Versement des archives électroniques

Lorsqu'un organisme créateur est invité à verser des pièces aux Archives nationales danoises, il doit observer certaines conditions spécifiées sur le site web des Archives nationales (http://www.sa.dk). Les données doivent être versées dans une version indépendante du système, respectant certains critères de structure de bibliothèque, de nom, etc. À partir des métadonnées spécifiées, l'organisme versant décrit la structure du système d'information et son contenu suivant un mode normalisé, qui peut être interprété sans ambiguïté par la machine. Les Archives demandent également des informations plus générales sur le système d'information, toutes étant livrées sur un ou plusieurs CD-R (en deux exemplaires), qui permettent d'assurer l'autodocumentation des archives électroniques.

Validation des CD-R

Une fois que les Archives nationales ont reçu de la part d'un organisme public versant la version électronique de documents d'archives sur des CD-R, ces derniers sont testés à l'aide de l'outil TEA (Test of Electronic Archives). Cet outil est développé par les Archives nationales en tant qu'implémentation électronique des spécifications techniques applicables au versement. TEA vérifie automatiquement que toutes les métadonnées sont correctes, de même que les structures et les noms, les formats, etc. Une vérification visuelle des fichiers attachés est effectuée en fin de test.

La conservation à long terme implique des tests et des conversions périodiques

La conservation permanente aux Archives nationales commence dès lors que les archives versées par une administration ont été testées et approuvées. Outre l'outil TEA, nous avons développé également divers outils de conversion, des méthodes d'assurance qualité et des procédures de test, de conversion et de documentation d'archives électroniques. Nous avons mis en place un atelier séparé pour la réception, les tests et la conversion permanente d'archives électroniques, en tant que partie intégrante d'une série complète de réglementations sur la sécurité des TI. Nous sommes actuellement occupés à la conversion de bandes magnétiques (datant des années 70) sur CD-R. La conversion des formats sera préparée et lancée en 2002.

La conservation dans le temps: sécurité des TI et assurance qualité

Les Archives nationales danoises sont parmi les plus expérimentées au monde en ce qui concerne la conservation permanente des données numériques. Nous avons créé une série intégrée de spécifications, procédures et processus techniques et organisationnels qui garantit une bonne conservation et l'accessibilité dans le temps, préservant ainsi le patrimoine numérique au bénéfice des chercheurs en histoire et des générations futures.

Parallel session 2

Tuesday 7 May 2002

THE USE OF PUBLIC INFORMATION
Content, authenticity, protection and intellectual control

Chairperson:	Miquel Á. Amutio Gómez (Spain)
Co-chair:	Salvatore Italia (Italy)
Rapporteur:	Catherine Dhérent (France)

Hermann Weidemann

Municipal Director
Born: 1940
Student: 1961
Student of Economy at Aarhus University: 1968–71
Municipal Director, Municipality of Midtjursland: 1976
Municipal Director, Municipality of Qaqortoq, Greenland: 1978
Municipal Director, Municipality of Hasle: 1983
Municipal Director, Municipality of Gundsoe: 1985
Municipal Director, Municipality of Morsoe: 1986
Municipal Director, Municipality of Naestved: 1989
Chairman of the Board, Naestved Info-Society 2000 A/S: 1995
Chairman of the Board, IVCeDK A/S: 2001

End-user driven e-content management
A view from a manager

Hermann Weidemann

In Denmark, the journey towards the information society started officially late 1994 using the Delors White Book as a guidance. From the outset the objective in Denmark was to give the information society a human face and user-driven content.

In our community — Naestved — we started at the same time by forming a public-private partnership with the Telecompanies TDC and Telia et. al. early in 1995 in order to get the whole region, both the public, the private sector and the citizens on to the information society as fast as possible.

Our first task was to provide a good infrastructure based on CATV and Euro-ISDN.

This was accomplished in 1996 when our local private web site NaestvedNet (thought of as a virtual representation of the whole region in cooperation with the municipalities' homepage).

Providing information to the public and offering some interactive solutions were among the initial services of NaestvedNet. The knowledge society (KS) was on our minds already at that time too, as we in 1994 began a profound change of our organisation, establishing the first physical one-stop service centre for citizens and enterprises.

At the same time we started cooperation with Lotus and IBM, as we had chosen Lotus Notes as our main collaboration tool — both internally and externally. Based on this we set out to work in a matrix-like organisation. Our way of working developed with and by the Internet, driven by content and competence from the very beginning.

In 1997/98, the first 800 employees were educated using e-learning, mostly at their homes. In the private sector only the financial institutions followed suit.

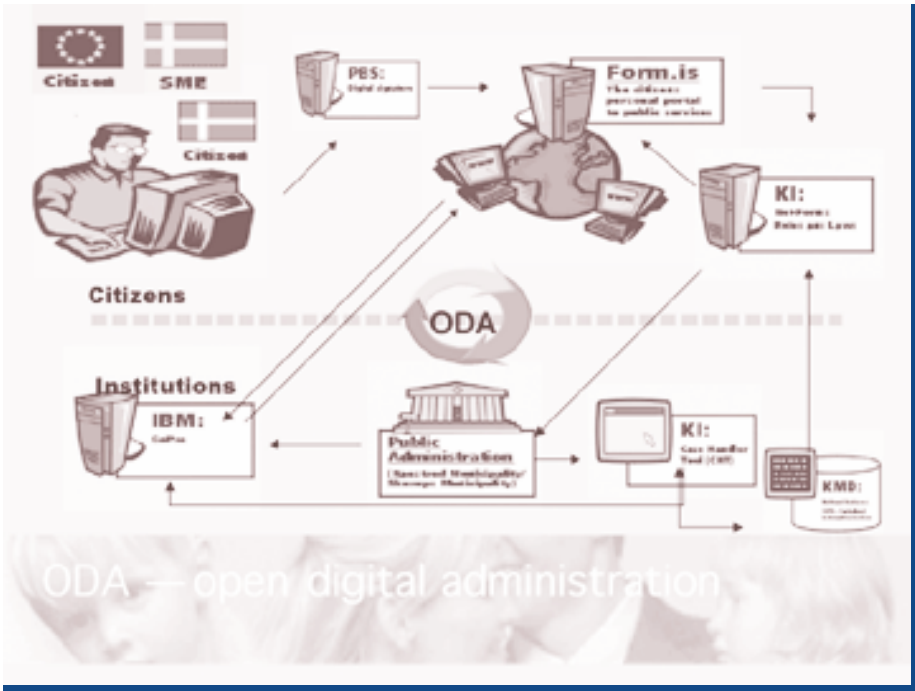
It was at that point we got convinced that the way to the knowledge society would be very long and cumbersome, and that a change in culture and understanding in all organisations was necessary.

Therefore, in 1999 we altered our private-public partnership to focus the work on content and competencies. Lotus and IBM joined officially as well as all the local educational institutions and partners.

The ideas about ODA (open digital administration) were born — and, slowly and painfully, our organisation started to prepare for KS, while the local private sector was undisturbed.

ODA is based on an international public-private partnership, and internally on Notes, GoPro (our universal central and decentralised archiving system), with Internet-thinking and digital signatures, high security and new ways of working, where possible automating the processes using IT tools wherever it made sense to avoid manual interference.

ODA is, well, a technical solution — but first of all it is a 'standard solution' which can be applied to transform administrative tasks both in the public and private sectors, and offers services and administrative experiences to the digital world, which will result in more value — savings and increased service at the same time for enterprises, the administration as well as for the citizens.



Marketing plan – focusing

- services – and related services
- potentital user groups (pendlers and companies with a high amount of employees)
- strategy of user groups to be invited
- sickness benefits
- building permits
- geographic dimension – pendlers ...
- cross border services ...

ODA – open digital administration

Geographicdimension — pendlers ... cross-border services

We soon realised that ODA was only the first small step on our way to the KS.

Therefore, in collaboration with private partners, last but not least IBM, we dared to think about ComKeen — common knowledge for everybody everywhere (anytime) on the Internet.

The vision is the same as in 1994/95, now the term 'information society' changed to 'knowledge society' with a clear 'global focus': local and international, general and individual at the same time.

1. Deliver structured and reusable (documented) access to a wide range of both public and private, local and worldwide information 24 hours a day.
2. Enterprises and the public to be able to apply for and receive both public and private service offfers online.
3. To replace industry-society like intensive manual and standard administrative solutions by individually chosen electronic (net-based) services.
4. Prevent social and others form of exclusion (e.g. handicap, language).
5. To fulfill different needs e.g. develop the same services in different ways depending on what elderly, young, working people or enterprises require.
6. To deliver different channels and offers to further the user's ability to use e-government and to create confidence in e-solutions and the new general way of living and working in the knowledge society and to reap the gains of effective, competitive e-services inside agencies, organisations and enterprises and for the public.

The main tasks are to:

— define, install and test a common ICT-platform to provide access for citizens, enterprises and employees alike to content, services, communication channels for:

individual education;

*e*Learning;

entertainment;

work;

cooperation;

community building;

political participation;

services from public and private organisations;

economical transactions of many kinds (from micro-payment to e-trade);

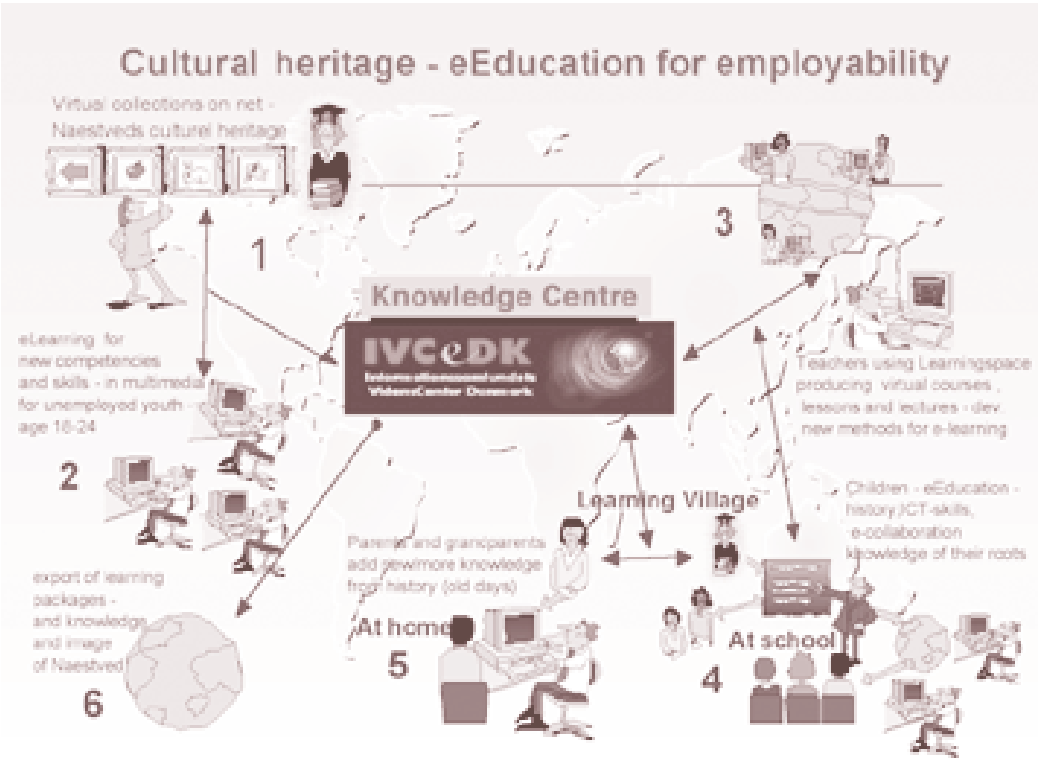
to ensure quality of life, freedom of choice from all over the world's offers and at the same time build and preserve an individual, a local (or regional) identity on the Net with the first main application (starter) packages.

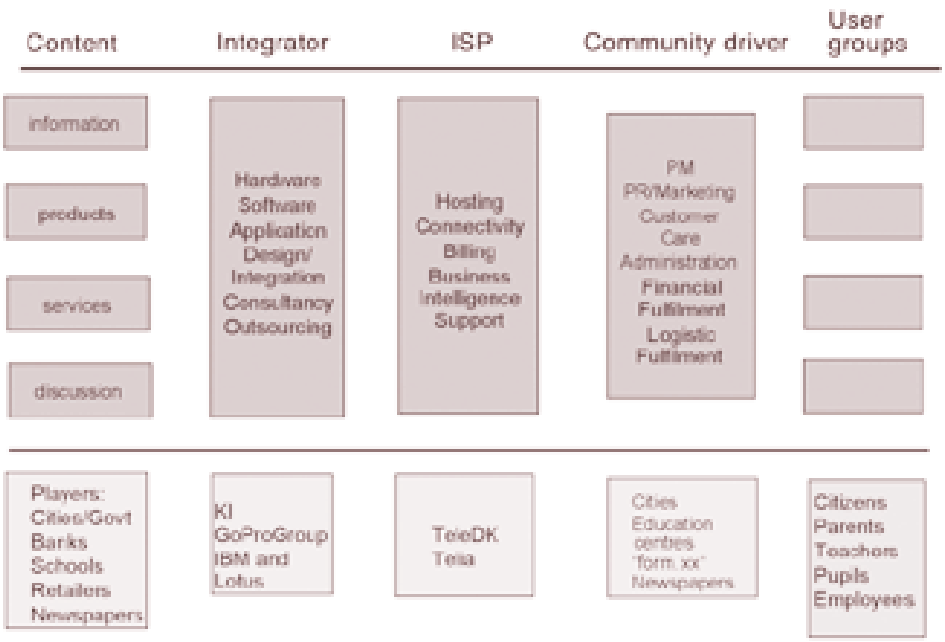
'Learning Village' (school communities):

- The local schools seen as the local community centre
- Teachers, parents and pupils working together and exchanging information on the Internet inside the Learning Village, both in public and in private
- The collaboration platform for teachers as professional cooperators
- QuickPlace and SameTime as an alternative standard platform for the same tasks.

Lifelong learning for all:

- Lifelong learning on the Internet for all with the public and private sectors as providers and the established educational institutions when they change to net-services
- Fighting social and other forms of exclusion (e.g. language, handicaps)





Citizens' self-services (life events, holistic solutions)

Technical services (health, energy, security ...)

- Access to health-security services, such as:

  — delivering data about pollution online or by mobile;
  — surveillance of citizens with special risks (heart failure …);

- surveillance of energy consumption in houses and factories;
- deliver security services — preventing theft;

Care for the elderly and come what may, and what the customers crave for

- decentralise e-business at home, e.g. by home-helpers;
- providing networking, securing social networks.

And at the same time to change our cultural and working surroundings in order to:

— develop significantly the civil servants' and citizens' ability and aptitude to use ICT;

— provide more resources, obligations and possibilities to users to secure IT-enabled change seen as powered by organisational and individual ongoing development;

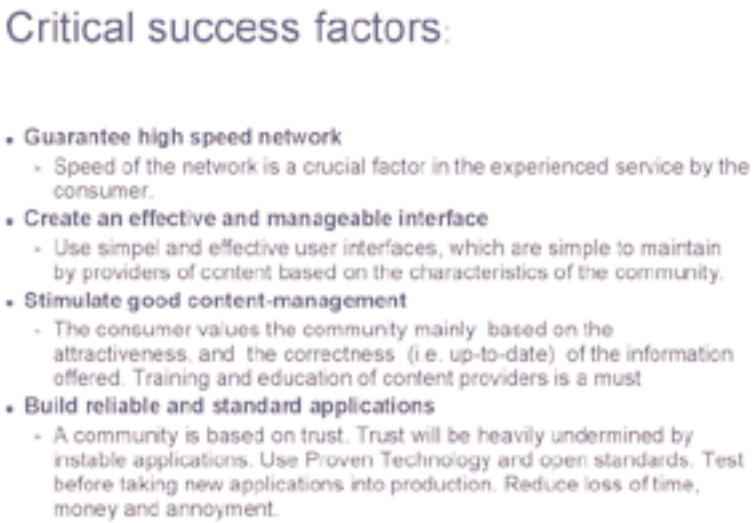
— implement a collaboration process to encompass further technological improvements and to updating existing ICT systems as the normal procedure in every agency, enterprise and for all citizens — on their own terms;

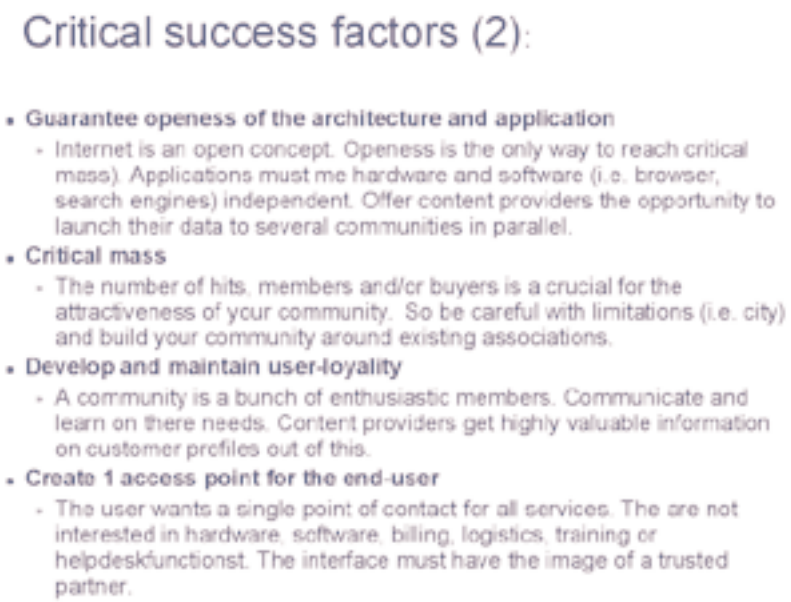
— secure the reliable assessment of costs and benefits both in an agency, in an enterprise and for the public as an integrated part of the daily work and cross-sectorial;

— build and maintain partnerships with other organisations, both public and private and with the public as the drivers of change og growing overall competence and competitional strength in a region or nation etc.;

— build understanding for exercise management both top-down and from bottom to top, as a general collaboration process.

There are no longer solutions, only ongoing processes, collaboration and the nursing of competences in individuals, organisations and society — the way to the knowledge society.





From then on content is defined to be everything available on the Internet and no longer limited by the individual provider of knowledge. Likewise, the available services will extend to what is accessible. In this knowledge-based society, the power resides only in processes, in constant change and development and, last but not least, with the citizen, the customer.

So the knowledge society inherently is also providing truly intellectual freedom. It has to!

Gestión de contenidos electrónicos basada en el usuario final

Hermann Weidemann

En Dinamarca, el viaje hacia la sociedad de la información comenzó oficialmente a finales de 1994, apoyándose en el Libro Blanco de Delors. Desde el principio, el objetivo en Dinamarca fue dar a la sociedad de la información una cara humana y unos contenidos basados en las necesidades de los usuarios.

En nuestra comunidad, Næstved, comenzamos formando una asociación pública-privada con las empresas de telecomunicaciones TDC y Telia y otras a principios de 1995, a fin de llevar toda la región, tanto el sector público como el sector privado y los ciudadanos, a la sociedad de la información lo más rápidamente posible.

Nuestra primera tarea consistió en proporcionar una infraestructura de calidad basada en la televisión por cable y en la norma EURO-ISDN.

Esto se logró en 1996, con nuestro sitio web privado local NæstvedNet (concebido como una representación virtual de toda la región en cooperación con las páginas web de los municipios).

Entre los servicios iniciales de NæstvedNet figuraban proporcionar información al público y ofrecer soluciones interactivas. En esta época, ya estaba en nuestras mentes la sociedad del conocimiento, pues en 1994 iniciamos un cambio profundo en nuestra organización creando el primer centro físico de servicios de ventanilla única para ciudadanos y empresas.

Al mismo tiempo empezamos a cooperar con Lotus e IBM, pues habíamos escogido Lotus Notes como nuestra principal herramienta de trabajo, tanto interna como externamente. Sobre esta base, comenzamos a trabajar en una organización matricial. Nuestro método de trabajo se desarrolló con y a través de Internet, y estuvo dirigido desde el principio por el contenido y por la competencia.

En 1997/98 se formó a los primeros 800 empleados utilizando el aprendizaje informático, sobre todo en sus hogares. En el sector privado, solamente las instituciones financieras siguieron esta tendencia.

En este momento nos dimos cuenta de que el camino hacia la sociedad del conocimiento sería muy largo y difícil.

En 1999 modificamos nuestra asociación privada-pública para centrar el trabajo en el contenido y las competencias. Lotus e IBM se adhirieron oficialmente, así como todos los socios e instituciones educativas locales.

Surgió la idea de la ADA (administración digital abierta) y, lenta y dificultosamente, nuestra organización comenzó a prepararse para la sociedad del conocimiento, mientras que el sector privado local se mantenía al margen.

La ADA se basa en la asociación pública-privada internacional, e internamente utiliza Lotus Notes y GoPro (nuestro sistema de archivo universal central y descentralizado), utilizando Internet y las firmas digitales, alta seguridad y nuevas maneras de trabajar, automatizando en la medida de lo posible los procesos mediante herramientas de tecnologías de la información (TI), cuando sea razonable evitar interferencias manuales.

La ADA es una solución técnica, pero en primer lugar es una «solución estándar» que puede aplicarse para transformar tareas administrativas tanto en los sectores público como privado, y ofrece servicios y experiencias administrativas al mundo digital, lo que dará lugar a ahorros y a un mayor servicio para las empresas, la administración y los ciudadanos.

Pronto nos dimos cuenta de que la ADA era solamente un modesto primer paso en el camino hacia la sociedad del conocimiento.

Por tanto, en colaboración con socios privados, entre los que se encontraba IBM, comenzamos a concebir ComKeen: conocimiento para todos en todas partes (en todo momento) en la red.

El concepto es el mismo que en 1994/95; ahora la sociedad de la información ha cambiado a la sociedad del conocimiento, con un claro foco global: local e internacional al mismo tiempo. Las principales tareas son las siguientes:

— Definir, instalar y probar una plataforma de TIC común para proporcionar acceso por igual a los ciudadanos, las empresas y los empleados, a contenidos, servicios, canales de comunicaciones para formación individual, aprendizaje por vía informática, entretenimiento, trabajo, cooperación, desarrollo de comunidades, participación política y servicios de organizaciones públicas y privadas.

— Garantizar la calidad de vida, la libertad de elección entre todas las ofertas mundiales y al mismo tiempo construir y conservar una identidad individual, local (o regional) en la red con las primeras aplicaciones principales:

  - Aldea de aprendizaje (comunidades escolares).
  - Formación continua para todos.
  - Autoservicios para los ciudadanos (acontecimientos de la vida, soluciones integrales).

- Servicios técnicos (salud, energía, seguridad).
- Cuidados para los mayores y... todo lo que los clientes puedan solicitar.

A partir de aquí, el contenido se define como cualquier objeto disponible en la red y no limitado por los proveedores individuales de contenidos. Del mismo modo, los servicios disponibles se ampliarán a lo que esté accesible. En esta sociedad basada en el conocimiento, el poder reside en los ciudadanos y los clientes.

De esta manera, la sociedad del conocimiento también proporciona intrínsecamente una auténtica libertad intelectual.

Am Endbenutzer orientiertes Management elektronischer Inhalte

Hermann Weidemann

In Dänemark setzte sich der Zug in Richtung Informationsgesellschaft offiziell Ende 1994 in Bewegung, wobei das Delors-Weißbuch als „Fahrplan" diente. Von Anfang an hatte Dänemark das Ziel, der Informationsgesellschaft ein menschliches Gesicht zu geben und Inhalte an den Erfordernissen der Benutzer auszurichten.

In unserer Gemeinde Naestved gründeten wir damals Anfang 1995 eine öffentlich-private Partnerschaft mit den Unternehmen TDC und Telia, um die gesamte Region – d. h. den öffentlichen und den privaten Sektor sowie die Bürger – so schnell wie möglich in die Informationsgesellschaft zu führen.

Unsere erste Aufgabe bestand darin, eine gute Infrastruktur auf der Basis von Kabelfernsehen und Euro-ISDN bereitzustellen. Das wurde 1996 erreicht, als unsere örtliche private Website NaestvedNet ins Netz ging (gedacht als virtuelle Vertretung der gesamten Region in Zusammenarbeit mit der Homepage der Gemeinden).

Zu den ersten Dienstleistungen von NaestvedNet gehörte die Vermittlung von Informationen für die Öffentlichkeit und ein Angebot interaktiver Lösungen. Bereits damals hatten wir eine Wissensgesellschaft vor Augen, als wir 1994 unsere Organisation von Grund auf umwandelten und die erste zentrale Anlaufstelle für Bürger und Unternehmen einrichteten.

Gleichzeitig begannen wir eine Kooperation mit Lotus und IBM, da wir uns für Lotus Notes als Haupt-Tool der internen und externen Zusammenarbeit entschieden hatten. Auf dieser Grundlage arbeiteten wir in einer matrixartigen Organisation. Unsere Arbeitsweise entwikkelte sich mit dem und durch das Internet und von Anfang an auf der Basis von Inhalten und Kompetenz.

In den Jahren 1997/98 wurden die ersten 800 Beschäftigten mittels E-Learning ausgebildet, größtenteils zu Hause. Hier zogen im Privatsektor nur die Finanzinstitute nach. An diesem Punkt gelangten wir zu der Erkenntnis, dass der Weg in die Wissensgesellschaft sehr lang und beschwerlich sein würde.

1999 verlagerten wir den Schwerpunkt unserer öffentlich-privaten Partnerschaft auf die Arbeit zu Inhalten und Kompetenzen. Lotus und IBM traten der Partnerschaft ebenso wie alle lokalen Bildungseinrichtungen und Partner offiziell bei.

Es entstand die Idee einer ODA (Open Digital Administration), und langsam wie mühselig begann unsere Organisation die Vorbereitung auf die Wissensgesellschaft, während dies den lokalen privaten Sektor völlig gleichgültig ließ.

ODA basiert auf einer internationalen öffentlich-privaten Partnerschaft, und intern auf Notes, GoPro (unserem universellen zentralen und dezentralen Archivierungssystem), mit Internetausrichtung und digitalen Signaturen, einem hohen Sicherheitsniveau und neuen Arbeitsweisen, wobei die Prozesse unter Verwendung von IT-Tools möglichst immer dann automatisiert werden, wenn die Vermeidung eines Eingriffs von Hand sinnvoll ist.

ODA ist natürlich eine technische Lösung. Doch in erster Linie handelt es sich um eine „Standardlösung", mit der sich administrative Aufgaben sowohl im öffentlichen als auch im privaten Sektor umgestalten lassen. ODA bietet Dienste und administrative Erfahrungen mit der digitalen Welt, woraus sich für Unternehmen, die Verwaltung und für die Bürger Einsparungen und gleichzeitig ein besserer Service ergeben.

Wir erkannten bald, dass ODA nur einen ersten kleinen Schritt auf unserem Weg zur Wissensgesellschaft darstellt.

So entstanden in Zusammenarbeit mit privaten Partnern, nicht zuletzt IBM, erste Vorstellungen von ComKeen [*Com*mon *K*nowledge for *E*verybody *E*verywhere (anytime) on the *N*et], also jederzeit für jedermann im Internet zugängliches allgemeines Wissen.

Die Vision ist die gleiche wie 1994/95, nur dass sich die Informationsgesellschaft zur Wissensgesellschaft mit eindeutigem „glokalen", also lokalen und internationalen zugleich, Schwerpunkt gewandelt hat. Dabei stehen folgende Hauptaufgaben an:

— Festlegung, Installation und Prüfung einer gemeinsamen IKT-Plattform, die Bürgern, Unternehmen und Mitarbeitern gleichermaßen Zugang zu Inhalten, Diensten, Kommunikationskanälen für individuelles Lernen, E-Learning, Unterhaltung, Arbeit, Kooperation, sozialem Engagement, politischer Mitbestimmung und Dienstleistungen staatlicher und privater Organisationen bietet;

— Sicherung von Lebensqualität und Wahlfreiheit für Angebote aus aller Welt bei gleichzeitiger Errichtung und Bewahrung einer individuellen, lokalen (oder regionalen) Identität im Netz mit den ersten wichtigen Anwendungspaketen:

  - lernendes Dorf (Schulgemeinschaften);
  - lebenslanges Lernen für alle;
  - Bürgerselbsthilfe (Veranstaltungen, ganzheitliche Lösungen);
  - technische Dienste (Gesundheit, Energie, Sicherheit);
  - Betreuung älterer Bürger;
  - und... weitere Angebote ganz nach Wunsch der Benutzer.

Künftig gilt alles als Inhalt, was im Internet verfügbar ist, ohne auf den jeweiligen individuellen Anbieter von Wissen begrenzt zu sein. Ebenso werden sich die verfügbaren Dienste auf alles erstrecken, was zugänglich ist. In dieser wissensbasierten Gesellschaft gehört die Macht dem Bürger, also dem Kunden.

Somit sorgt die Wissensgesellschaft ihrem Wesen nach auch für eine wirkliche geistige Freiheit.

Gestion des contenus électroniques axée sur l'utilisateur final

Hermann Weidemann

Le voyage vers la société de l'information a débuté officiellement au Danemark vers la fin de 1994, en prenant appui sur le livre blanc de Jacques Delors. Dès le départ, l'objectif des Danois était de donner à la société de l'information un visage humain et des contenus orientés sur les besoins des utilisateurs.

Au sein de Naestved, nous avons commencé alors à constituer un partenariat public-privé avec les télésociétés TDC et Telia et al., au début de 1995, afin de conduire toute la région, c'est-à-dire le secteur privé et public et les citoyens, vers la société de l'information et ce, le plus rapidement possible.

Notre première mission a consisté à mettre en place une infrastructure de qualité, reposant sur la télévision par câble et sur la norme Euro-RNIS. Ce travail a été mené à terme en 1996, avec notre site web privé local NaestvedNet (conçu comme une représentation virtuelle de toute la région, en coopération avec les pages d'accueil des municipalités).

Les premiers services assurés par NaestvedNet ont été de fournir des informations au public et de proposer certaines solutions interactives. La société de l'information était déjà dans notre esprit à cette époque, puisqu'en 1994 nous avions entamé une transformation en profondeur de notre organisation, en mettant en place le premier centre physique de services à guichet unique destiné aux citoyens et aux entreprises.

Parallèlement, nous avions entamé une coopération avec Lotus et IBM, puisque nous avions choisi Lotus Notes comme principal outil de travail — au niveau interne et externe. À partir de cette plate-forme, nous avons défini une organisation matricielle. Notre méthode de travail a été élaborée à l'aide de et par l'internet, et a été dictée par le contenu et par la compétence, dès le début du projet.

En 1997/1998, les 800 premiers salariés ont bénéficié d'une formation par e-apprentissage, la plupart depuis leur domicile. Dans le secteur privé, seules les institutions financières nous ont emboîté le pas. C'est à ce moment que nous avons acquis la conviction que le chemin vers la société de l'information serait très long et difficile.

En 1999, nous avons modifié notre partenariat privé-public pour concentrer les travaux sur les contenus et les compétences. Lotus et IBM nous ont rejoints officiellement, ainsi que la totalité des établissements d'enseignement et des partenaires locaux.

Les idées sur l'administration numérique ouverte (ODA) sont nées — et notre organisation a commencé, lentement et dans la douleur, à se préparer à la société de l'information tandis que les acteurs privés locaux ne bronchaient pas.

L'administration numérique ouverte repose sur un partenariat public-privé international; en interne, elle repose sur une plate-forme Lotus Notes et GoPro (notre système d'archivage universel central et décentralisé) dans un environnement internet, et met en œuvre des signatures numériques, un haut niveau de sécurité et de nouvelles modalités de travail qui consistent à automatiser les processus à l'aide des technologies de l'information chaque fois qu'il est rationnel d'éviter une interférence manuelle.

L'ODA est certes une solution technique, mais c'est d'abord et avant tout une "solution standard", qui peut être appliquée pour transformer les tâches administratives dans le public comme dans le privé et qui offre une expérience des services et de l'administration au monde du numérique, ce qui créera à terme de la valeur — en permettant des économies

et en assurant un meilleur service simultanément aux entreprises, à l'administration et aux citoyens.

Nous nous sommes rendu compte très vite que l'ODA n'était qu'une première étape modeste vers la société de l'information.

Aussi, en collaboration avec des partenaires privés dont IBM n'est pas le moindre, nous avons sauté le pas et commencé à réfléchir au projet ComKeen (Common Knowledge for Everybody Everywhere), système de connaissances accessible à tous (n'importe quand) sur le net.

Le tableau est le même qu'en 1994/1995, maintenant que la société de l'information évolue vers la société de la connaissance, avec un centrage "glocal", c'est-à-dire à la fois global et local. Les principales tâches sont de:

- Définir, installer et tester une plate-forme TIC commune pour permettre aux citoyens, aux entreprises et aux salariés d'accéder à des contenus, des services et des canaux de communication à des fins multiples: enseignement individuel, apprentissage électronique, loisirs, travail, coopération, constitution d'un forum, participation politique, services des organismes publics et privés.

- Veiller à la qualité de vie et à la liberté de choix dans toutes les offres mondiales et, parallèlement, constituer et protéger une identité individuelle et locale (ou régionale) sur le net, avec comme premières applications:
  - le village de la connaissance (milieux scolaires),
  - l'éducation et la formation tout au long de la vie pour tous,
  - le libre-service citoyen (événements de la vie, solutions intégrées, etc.),
  - les services techniques (santé, énergie, sécurité, etc.),
  - la prise en charge des personnes âgées et… tout ce que le client peut souhaiter ardemment.

À partir de là, le contenu se définit comme tout ce qui est proposé sur le net, et non plus simplement par chaque fournisseur de connaissance. De la même façon, les services disponibles s'étendent à ce qui est accessible. Dans cette société basée sur la connaissance, le pouvoir est entre les mains du citoyen, *c'est-à-dire* du client.

La société de la connaissance est donc intrinsèquement pourvoyeuse d'une véritable liberté intellectuelle.

On the use of digital signatures in the preservation of electronic objects

Andrew Waugh

Andrew Waugh is a senior scientist within CSIRO Mathematical and Information Sciences. For the last four years, he has been working on the Victorian electronic record strategy (VERS) (http://www.prov.vic.gov.au/vers/) The VERS project is funded by the Victoria State Government to develop a strategy that will ensure availability of electronic records for the forseeable future. Currently, Andrew is acting as a consulting engineer for an implementation of the VERS strategy within a medium-sized Victorian State Government department.

Introduction

A major concern with the preservation of electronic objects is ensuring that the objects retain their authenticity over time. The concept of authenticity means different things to different groups of people (for recent views see [Bearman, CLIR]), but one aspect is ensuring the 'integrity' of the object, that is, that no unauthorised alterations have been made to the object.

One technology that can be used to detect modifications to digital objects is public key digital signatures. Over recent years some preservation literature has suggested the use of digital signatures to prevent objects from being undetectably modified. Hedstrom [Hedstrom], for example, suggests that digital signatures are part of a set of tools that can 'maintain the physical and intellectual integrity of the records'. Lynch [Lynch] has pointed out that using a digital signature is equivalent proving the authenticity of your copy of an object by comparing your copy with a master copy — the master copy being the object at the point of time when it was signed, and your copy either being the same object at a later time, or a different copy. The OAIS reference model [OAIS] discusses 'fixity' information and technology amongst which can be included digital signatures.

When we were developing the Victorian electronic record strategy (VERS) [VERS1] in the late 1990s, we used digital signatures to detect modifications to the preserved objects. Subsequent implementation experience with VERS within the Victorian Department of Infrastructure (DoI), however, has highlighted a number of practical challenges in using digital signatures in preservation work. Digital signatures can be used, but if care is not taken in design and implementation the signatures can prove useless.

Fundamentally, the problem is that using digital signatures to secure an object over a long period of time is subtly different to conventional applications of digital signatures. Conventional applications protect messages, such as web pages, as they are being transmitted across a network. A characteristic of this usage is that the messages are checked for corruption very shortly after they are signed. If corruption occurred the message can be re-sent, and there is no need to keep information around for long periods. Compare this with the preservation usage where it is not possible to 're-send' the message and the information required to verify the signature must be kept for as long as the item is preserved.

In the remainder of this paper we describe a number of challenges in applying digital signatures in a preservation context, and outline the approaches we have taken to surmount these challenges.

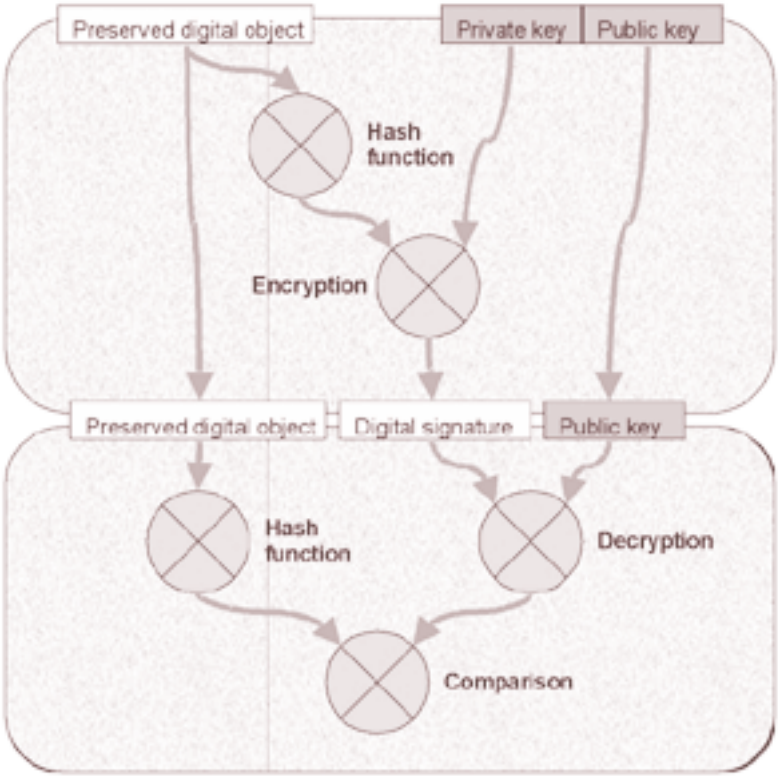
Before commencing, however, it should be noted that the ability to detect modifications to a digital object addresses only one aspect of the authenticity of the object. Other techniques must be used to address the other aspects and these techniques are beyond the scope of this paper.

Public key digital signatures

For those readers unfamiliar with public key digital signatures, we present the following simplified introduction. For a more detailed discussion see [VERS2].

A digital signature has little in common with a physical handwritten signature. In particular, it is not a scanned image of a handwritten signature. Instead, it is the result of a mathematical calculation which takes as input the digital object to be signed and a secret known only to the signer. The calculation has the property that the digital signature changes if the digital object is changed (even in a minor way) or a different secret is used.

Diagram of the process of signing a preserved digital object and subsequently verifying the digital signature.



Conceptually, generating and verifying digital signatures is simple. Signing commences by 'hashing' the digital object to be signed. This simply means running the object through a mathematical function (known as a hash function) to produce a number. The hash function has the property that even tiny changes in the message will result in a different hash value; this is the core of detecting alterations in the message. The output of the hash function is then encrypted using a private key known only by the signer to produce the digital signature.

Verifying a digital signature starts by decrypting the signature to give the original hash value. The digital object is then rehashed using the same hash function and the two hash values are compared. If they are the same the digital object has not been changed since it was signed.

Encryption of the hash value is by public key cryptography. This type of cryptography uses keys that come in pairs: the public key and the private key. The private key is used to encrypt the hash value and the matching public key is used to decrypt the hash value. The use of two matched keys allows the signer to keep the private key secret while publishing the public key.

It is worth noting that the term 'digital seal' would have been a more accurate name. Like a conventional seal, a digital signature can be mechanically applied by anyone who has access to it. It follows that the security of a digital signature is dependent on keeping the private key secret, just as the security of a conventional seal is dependent on keeping it physically secure to prevent misuse.

Public key digital signatures have the following characteristics:

- access to the public key is necessary to verify the digital signature. The public key must consequently be accessible for as long as the signed object is kept. The implications of this will be considered in Section 4;
- any change in the preserved object (even one bit) will result in a different hash value, and hence will cause the digital signature to fail. This implications of this will be discussed in Section 5.

Authentication of the public key

Challenge: Validation of a digital signature requires the authentic public key of the signer. In conventional digital signature applications this is obtained from a public key infrastructure consisting of certificates issued by trusted certificate authorities. This is acceptable where the digital signatures have a short life, but certificate authorities are unlikely to keep certificates for the life of a preserved object.

Validation of a digital signature requires the public key of the signer. If the public key has been lost or discarded the integrity of the preserved object cannot be verified. Further, verification depends on being sure that the stored public key actually belonged to the purported signer (otherwise the preserved object could be modified, resigned, and the public key replaced). Public keys must consequently be securely stored for the lifespan of the signed objects; this could be for a century or more.

In a conventional digital signature application, public keys are obtained from certificates produced and stored by certificate authorities (a discussion of certificates and certificate authorities can be found in [VERS2]). However, it is open to question whether a certificate authority can (or should) be trusted to store the certificates it produced for the length of time required for preservation activities. Certificate authorities are usually commercial organisations and there is no guarantee that if the organisation fails or exits the business that the certificate store will be retained. How many commercial organisations are still in existence after 100 years? Note that there is little commercial pressure to provide cast iron guarantees of long-term access to certificates as most digital signatures have a relatively short life.

The solution to this challenge requires an organisation holding preserved digital objects also to store the necessary public keys to verify the preserved objects. The public keys would normally be held within certificates. This should not be an onerous requirement as certificates are simply digital objects and can be preserved within the same archive system that manages the actual preserved objects.

Care needs to be taken that the necessary certificates are actually captured into the system. Custom verification software will need to be written to obtain the certificates from the archive system rather than from conventional certificate authorities. If preserved objects are moved from one system to another the relevant certificates must be identified and moved with the preserved object. Finally, it should be appreciated that holding certificates requires careful security. If a forger can replace the certificates in the local system then they can modify any digital object in the collection.

A second option is to hold the necessary certificates within the preserved object itself. A particularly attractive option within VERS as a key assumption of VERS was that preserved objects would outlive the archive system that held them, so the preserved objects should stand alone from the archive system. Including the certificates within the preserved object reduces the dependency of the preserved object on other objects, ensures that the certificates are captured when the digital object is preserved, and transferred with the preserved object. There are two problems, however. The minor problem is the inefficiency involved in storing multiple copies of certificates, though this is not serious as certificates are quite small. The major problem is that it is not secure. A little thought reveals the circular argument that you are validating the contents of a preserved object by means of a signature which, in turn, is verified by the contents of the object.

A solution to this circular argument is to discard the conventional concept of digital signature verification by means of a certificate chain. Fundamentally, the problem we are trying to solve is to show that a particular public key belonged to a particular signer at the time a preserved object was signed. One way of doing this is to use the suspect public key to verify other digital objects created by the same user around the same time. If these verifications succeed, then either the public key is valid, or all the digital objects are forgeries. Clearly this is a probabilistic approach, but with a sufficiently large number of digital objects there would be strong evidence that the records have not been tampered with. This method is exactly analogous to verifying handwritten signatures in a traditional archive. When it is necessary to verify a handwritten signature, the suspect

signature is compared with other examples of the signature in the archive. If they match, the handwritten signature is treated as valid, otherwise it is considered suspect. In both cases the evidence can be made stronger if the preserved object is signed several times and each of these signatures is compared with other contemporary signatures.

In practice, the VERS at DoI implementation uses a combination of approaches to preserve public keys. A preserved object in VERS contains all of the certificates necessary to verify the digital signature and the archive system also holds the root certificates. Validation of the digital signature can either proceed conventionally by validating the chain of certificates, or by the probabilistic method described in the previous paragraph.

Changing an unchangeable object

Challenge: Digital signatures detect any change to a signed object; but it is not possible to discover what has been changed. This means that when an object has been locked by a signature, any change (even valid changes) will break the digital signature and it will be henceforward useless for protection.

A digital signature locks the signed object. The signature will detect any modification to the signed object — be it a change to one bit or to thousands of bits. When signature verification fails it is impossible to determine what has been changed or how much has been changed; all that is known is that something has changed.

This would be irrelevant if it was never desired to modify a preserved digital object. However, we have found that it can be necessary to modify a preserved object. A preserved object within VERS, for example, contains a complete collection of metadata for the record (this supports our goal of being independent of the archive system that holds the object). It is quite legitimate to change some of this metadata, for example, to correct a spelling mistake, or to add additional descriptive information. This metadata forms the context of the preserved object (i.e. what it is and how it relates to other preserved objects), and in archival theory the context of a preserved object is as important as the content of the preserved object. For example, changing the date a record was created may be more important for a forger than changing the content.

Many archival systems manage the metadata about a preserved object separately from the content of the preserved object. The metadata is directly held by the system; this allows it to be modified and protects it from unauthorised modification. However, apart from making the preserved object dependent on the archive system, this raises the questions: if the archive system can be trusted to protect the metadata from modification, why can it not be trusted to protect the content as well, and if it can protect the content directly, why do we need digital signatures at all?

In VERS we made the choice that the archive system should not be trusted to protect the integrity of either the content or the metadata. We consequently faced a complex challenge. We wished to use a digital signature to detect unauthorised modifications to both the metadata and the content of the record. On the other hand, the digital signature prevented users from performing legitimate modifications to the metadata.

The technique we used to resolve this contradiction is always to retain the original preserved object intact with its digital signature. Thus, it is always possible to verify the integrity of the preserved object. Modifications are added outside the original preserved object and protected by their own digital signature. The original object and the collection of modifications are integrated into a single object by digital signatures.

We came up with two practical implementations of this general technique: the onion and the skewer.

In an onion implementation the modified object is wrapped around the original object. Modifications can be repeatedly applied and the resulting object resembles an onion with the original object at the centre surrounded by successive layers of modifications. The digital signature at each level protects the modifications at that level from change and also protects the relationship

between this version of the object and the previous version of the object. One problem with this approach is that nothing prevents a forger from peeling off and discarding the outermost layer of the onion (i.e. the most recent changes).

In a skewer implementation, the modification is appended to the end of the original object and protected by a digital signature. A final digital signature is then applied to 'lock' the original object and the modifications into one object. When it is necessary to add a new modification the archival system checks the lock digital signature (to ensure the integrity of the object) and discards it. The new modification is added and the resulting collection relocked by a new digital signature. This approach prevents a forger from discarding modifications, but requires the archival system to be trusted while the object is 'unlocked'. It is possible to come up with subtle modifications of this scheme to improve utility: for example the lock signature might only cover the digital signatures of the components of the object. This provides equivalent protection, but is much faster to calculate as much smaller amounts of data need to be signed.

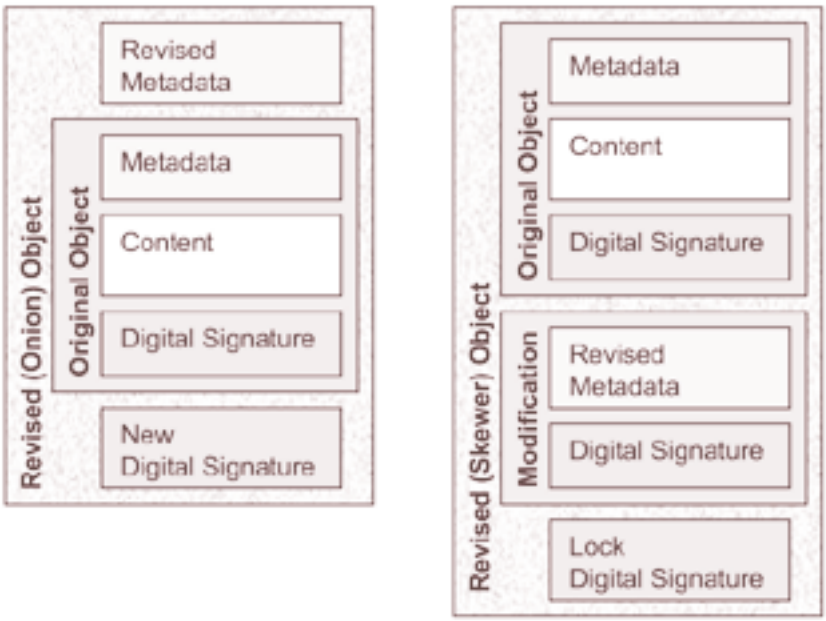


Figure 2. Revising the metadata associated with a preserved digital signature without invalidating the original signature. Left, the 'onion' approach where the original object is wrapped inside a new object with the revised metadata and new digital signature. Right, the 'skewer' model where the revised metadata is appended to the original object and all of the parts 'locked' with a third 'lock' signature.

In either the onion or skewer implementations, thought needs to be given as to whether the modifications contain everything (including information that is unchanged) or just the changes. The latter is much more efficient in storage, but requires a complex merge to apply the modifications when the object is accessed.

The second implementation question is how often changes result in modifications to the preserved object. One extreme is, of course, that each change results in a modification to the object. In practice, this is likely to result in very large objects, particularly if the modification contains all information. The alternative is to store changes in the archive system and to flush periodically the changes through to the preserved object. This flush could be done periodically, upon demand by a system administrator, when a sufficient quantity of changes had accumulated, or when important changes have occurred.

It is interesting to trace the development of our thinking in VERS. Initially we were not going to allow any changes at all. When this proved impractical, we moved to the onion model with each layer containing all of the metadata. This is the current situation, but when we moved into implementation it was pointed out that the outermost layer could be undetectably discarded. If we were designing VERS today, we would use the skewer model and we might switch to only storing the changes. The lesson we learnt is to think carefully about what is protected by the digital signature and how to efficiently manage changes.

Implementation validation

Challenge: Applications that use digital signature libraries can be difficult to implement correctly. It is possible to create signatures that another implementation cannot verify. In a conventional net-

working application this problem will become quickly apparent due to interworking with other implementations. In a digital preservation application, however, it might be years or decades before the signature is checked by an independent application and the problem comes to light.

This challenge covers two separate problems. The first is that the implementation of the archival application may be incorrect and calculates the digital signature over the wrong object. The second is that the implementation of the underlying digital signature software may be incorrect.

A digital signature will detect any change to the object it protects. This object is ultimately represented as a sequence of octets. To verify the signature it is necessary to use exactly the same sequence of octets that was used when signing. Although this may seem easy, our experience has shown that it is surprisingly difficult to achieve. Octets that have no significance in normal processing (e.g. spaces, line feeds, null bytes) do have a significance when calculating the digital signature and must be carefully processed. This is particularly true when the digital signature only covers a part of the preserved object.

Even when the archive software correctly verifies the digital signatures it calculates, it is worth thinking about whether a future implementator can determine exactly what octets are to be fed into the verification software. Some form of specification is required, and in VERS we include this specification in each preserved object to ensure that it is available to a future implementor. The specification must be unambiguous and the implementation must implement it correctly.

The second problem is buggy implementations of the digital signature algorithm itself. We did not expect this problem; there are only a few digital signature implementations and they must all interwork in conventional digital signature applications. However, when we tested the signatures generated by the VERS implementation for DoI we could not verify them using an independent implementation (and vice versa). After several weeks' work, we discovered that the signature generated by the digital signature software used by the VERS implementation did not conform to the digital signature standard.

These two problems (buggy cryptographic software, and buggy use of the software) are particularly dangerous for digital preservation. Simply testing that an implementation can verify the signatures that it produces does not show that it has implemented the specification correctly, nor that the underlying cryptographic software is correct. This can only be shown by testing against an independent implementation. However, it may be years before an independent implementation exists and until that happens all the signatures you generate may be incorrect and hence worthless. Part of the acceptance testing for any archival system that uses digital signatures must be a verification of the signatures using an independent implementation. This independent implementation should be implemented by a different software engineer from a specification; this will help ensure that the specification is adequate and that it is correctly implemented. The independent implementation should use a completely different cryptographic implementation to ensure that the implementation you are using is correct.

Conclusions

Digital signatures can be used to preserve the integrity of preserved digital objects, but care needs to be taken in the design and implementation of such a system. Preservation applications have different characteristics to conventional digital signature applications which usually verify signatures shortly after the signature has been applied. Preservation applications require the signature to be verified long after the signature has been applied (possibly centuries later) by software that may not be written for years into the future.

A key design issue of the preserved object is whether it will be necessary to modify parts of the preserved object after it has been signed. If so, then these parts must either be outside the protection of the digital signature (and protected by other means) or the preserved object must be designed to allow modification of the preserved object. We have suggested two possible designs: the 'onion' and the 'skewer' which allow modifications to the preserved object. Neither are ideal, and we believe that the requirement to modify a preserved object is the most difficult aspect of using digital signatures in preserving digital signatures.

A second key design issue is to ensure that the necessary public keys to verify the signature are preserved for as long as it is required to preserve the signed objects. Essentially the only method of ensuring this is for the preservation organisation to preserve the necessary public keys. We store the public keys in the preserved objects, and use a probabilistic approach to verifying these keys where the purported public key is compared with the public key of other objects signed by the same signer around the same time.

The resulting implementation must be carefully tested, ideally by a completely independent test programme. Independent means written by separate developers from a specification using different cryptographic software. If the signature is not tested by means of an independent test programme it is quite possible that the digital signature will be incorrect.

References

Bearman, D., Trant, J., *Authenticity of digital resources*, D-Lib, June 1998, http://www.dlib.org/dlib/june98/06bearman.html, visited 4 March 2002.

Council on Library and Information Resources, *Authenticity in a digital environment*, May 2000, ISBN 1-887334-77-7, http://www.clir.org/pubs/reports/pub92/pub92.pdf, visited 18 March 2002.

Hedstrom, M., 'Building record-keeping systems: archivists are not alone on the wild frontier', *Archivaria*, No 44, 1997, pp. 44–71.

Lynch, C., 'Authenticity and integrity in the digital environment', *Authenticity in a digital environment*, Council on Library and Information Resources, May 2000, ISBN 1-887334-77-7, http://www.clir.org/pubs/reports/pub92/pub92.pdf, visited 18 March 2002.

McCullagh, A., Caelli, W., 'Non-repudiation in the digital environment', *First Monday*, Vol. 5 No 8, August 2000, http://firstmonday.org/issues5_8/mccullagh/index.html, visited 8 August 2000.

'Reference model for an open archival information system (OAIS)', Consultative Committee for Space Data Systems, *Red book,* Issue 2, July 2001, http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html visited 23 March 2002

Victorian electronic record strategy web page, http://www.prov.vic.gov.au/vers/, visited 18 March 2002.

'Standard for the management of electronic records', PROS 99/007 (Version 1), Public Record Office, Victoria, April 2000, http://www.prov.vic.gov.au/vers/published/publcns.htm#PROStandards, visited 18 March 2002.

Utilización de las firmas digitales en la conservación de los objetos electrónicos

Andrew Waugh

Las firmas digitales pueden utilizarse para detectar modificaciones realizadas en objetos protegidos. En algunas aplicaciones (por ejemplo, registro de datos), la autenticidad del objeto es un aspecto crítico. En este documento se discutirán algunas de las lecciones que aprendimos en el proyecto VERS sobre utilización de firmas digitales para probar la autenticidad de los documentos electrónicos.

La longevidad de las firmas digitales

Reto: La vida de un objeto firmado digitalmente debe ser más corta que el tiempo empleado en descodificar la firma digital. La vida útil de un objeto protegido puede medirse en siglos. En este tiempo, los algoritmos de descodificación mejorarán y funcionarán en ordenadores más rápidos.

Soluciones: Elegir la tecnología de firma digital más segura disponible en la actualidad. Periódicamente, volver a firmar el objeto con la tecnología actual. Realizar análisis de riesgo: ¿cuánto esfuerzo (coste) emplearía alguien en falsificar un objeto?

Protección de la clave privada

Reto: Para probar la identidad del firmante es necesario que éste conserve en secreto su clave privada. Actualmente no se dispone de forma generalizada de una tecnología adecuada a estos efectos; las claves se almacenan generalmente en ordenadores y están protegidas solamente con contraseñas. Cualquier fallo de seguridad (accidental o deliberado) en el ordenador permite que otros tengan acceso a la clave privada y tengan por lo tanto la capacidad para falsificar firmas.

Soluciones: No confiar excesivamente en la firma digital para probar la identidad del firmante. Las tarjetas inteligentes y la bioinformática pueden ser la solución del futuro, pero todavía no es rentable desarrollarlas.

Autentificación de la clave pública

Reto: La validación de una firma digital requiere la clave pública auténtica del firmante. Ésta suele obtenerse de una infraestructura de clave pública consistente en certificados expedidos por autoridades de certificación acreditadas. Esto es aceptable en el caso de que las firmas digitales tengan una vida corta, pero es improbable que las autoridades de certificación conserven los certificados a lo largo de la vida de un objeto protegido.

Soluciones: Primera: Gestione su propio organismo de certificación. Aquí hay cuestiones de seguridad, porque si un falsificador es capaz de añadir un certificado, es capaz de falsificar firmas. Segunda: Almacene los certificados en el objeto protegido. Esto no soluciona completamente el problema, pues no puede verificar independientemente el certificado raíz. Tercera: Conserve la clave pública en el objeto firmado. Para verificarlo, compare las claves públicas de un gran número de objetos presuntamente firmados por el mismo firmante. Si las claves públicas son iguales, se puede confiar razonablemente en que el documento no está falsificado; cuantos más objetos se examinen, mayor será la seguridad. Esto es parecido al método que se utiliza para verificar las firmas manuscritas cuando no se puede obtener una muestra de la firma.

Modificación de un objeto no modificable

Reto: Las firmas digitales detectan cualquier cambio en un objeto firmado; pero no es posible descubrir lo que se ha cambiado. Esto significa que cuando un objeto esté protegido por una firma, cualquier cambio (incluso los cambios válidos) romperá la firma digital, que en adelante no servirá para protección; un cambio válido de poca importancia en un objeto rompe la firma y abre la vía a cambios desautorizados masivos. Los cambios válidos pueden incluir cambios en los metadatos incluidos en el objeto protegido (por ejemplo, corrección de un error ortográfico en los metadatos o adición de metadatos).

Solución: Ser cuidadoso en la definición de lo que está y no está protegido por una firma digital. Verificar y volver a firmar los objetos cuando sea necesario cambiarlos (esto equivale a la ruptura del sello y al nuevo sellado de los objetos físicos). La desventaja de este procedimiento es que la firma solamente protege desde el momento de la última firma; hay que confiar en que el sistema haya verificado correctamente el objeto y lo haya vuelto a firmar, y que no haya cambiado el objeto. Pueden proporcionarse capas de firmas digitales de modo que las firmas internas permanezcan intactas, mientras que las firmas externas puedan volverse a calcular cuando sea necesario.

Validación de la aplicación

Reto: Las aplicaciones que utilizan bibliotecas de firma digital pueden ser difíciles de ejecutar correctamente. Es posible crear firmas que no pueda verificar otra aplicación. En una aplicación de red, esto se pondrá de manifiesto muy rápidamente debido a la interactuación con otras aplicaciones. Esto sucedió con la aplicación de VERS; la versión criptográfica RSA de Microsoft que utilizamos produce firmas en un formato no válido. Esto se descubrió porque comprobamos la firma con un programa independiente. En una aplicación de conservación digital, pueden pasar años o incluso décadas antes que la firma sea comprobada por un programa independiente y el problema salga a la luz.

Solución: Contar siempre con un sistema de validación independiente para verificar las firmas antes de comenzar a utilizar un sistema de conservación. El sistema independiente debe utilizar una aplicación de firma digital diferente a la del sistema de producción, e, idealmente, debería ser elaborada de forma independiente. Desconfiar de la utilización de productos que parezcan independientes, pero que en realidad utilicen el mismo motor de encriptación. Puede ser difícil y frustrante detectar los errores en las firmas digitales.

Zur Verwendung digitaler Signaturen bei der Archivierung elektronischer Objekte

Andrew Waugh

Digitale Signaturen können verwendet werden, um Änderungen an konservierten Objekten festzustellen. Bei einigen Anwendungen (z. B. Archivgutverwaltung) stellt die Authentizität des Objekts einen kritischen Aspekt dar. In diesem Beitrag werden eine Reihe von Erfahrungen erörtert, die wir im Rahmen des VERS-Projekts bei der Nutzung digitaler Signaturen zum Nachweis der Authentizität elektronischer Unterlagen gesammelt haben.

Die Langlebigkeit digitaler Signaturen

Problem: Die Lebensdauer eines mit digitaler Signatur versehenen Objekts muss kürzer sein als die Zeit, die erforderlich ist, um die Signatur zu „knacken“. Die Lebensdauer eines konservierten Objekts lässt sich in Jahrhunderten messen. In dieser Zeit verbessern sich die Entschlüsselungsalgorithmen und laufen auf schnelleren Computern.

Lösung: Die sicherste heute verfügbare Technologie für digitale Signaturen auswählen; das Objekt in regelmäßigen Abständen mit aktueller Technologie neu signieren; Risikoanalyse durchführen: Was würde jemand (kostenmäßig) aufwenden, um ein Objekt zu fälschen?

Geheimhaltung des privaten Schlüssels

Problem: Zum Nachweis der Identität des Unterzeichners muss dieser seinen privaten Schlüssel geheim halten. Eine geeignete Technologie dafür ist derzeit noch nicht allgemein verbreitet. In der Regel werden die Schlüssel auf Computern gespeichert und sind nur durch Passwörter geschützt. Jede (zufällige oder absichtliche) Sicherheitslücke auf dem Computer

eröffnet Dritten den Zugang zum privaten Schlüssel und somit die Möglichkeit, Signaturen zu fälschen.

Lösung: Man sollte sich nicht zu stark auf digitale Signaturen als Identitätsnachweis des Unterzeichners verlassen. Smart Cards und Bioinformatikanwendungen gehört möglicherweise die Zukunft, doch lassen sie sich bisher noch nicht kostengünstig einführen.

Authentifizierung des öffentlichen Schlüssels

Problem: Zur Validierung einer digitalen Signatur ist der authentische öffentliche Schlüssel des Unterzeichners erforderlich. Dieser wird in der Regel über eine Public-Key-Infrastruktur eingeholt, die aus von vertrauenswürdigen Stellen ausgestellten Zertifikaten besteht. Das ist dann sinnvoll, wenn die digitalen Signaturen von kurzer Lebensdauer sind. Es ist jedoch unwahrscheinlich, dass Zertifizierungsstellen die Zertifikate während der gesamten Lebensdauer eines konservierten Objekts aufbewahren.

Lösungen: Erstens: Selbst als Zertifizierungsstelle fungieren. Allerdings gibt es auch hier Sicherheitsprobleme: Wenn es einem Fälscher gelingt, ein Zertifikat hinzuzufügen, kann er Signaturen erfolgreich fälschen. Zweitens: Zertifikate im konservierten Objekt speichern. Dadurch wird das Problem jedoch nicht vollständig gelöst, da sich das Root Certificate nicht separat verifizieren lässt. Drittens: Öffentlichen Schlüssel im signierten Objekt speichern. Zur Verifizierung werden die öffentlichen Schlüssel in einer Vielzahl von vorgeblich vom Unterzeichner signierten Objekten verglichen. Sind sie gleich, kann man hinlänglich davon ausgehen, dass das Dokument nicht gefälscht ist. Je mehr Objekte untersucht werden, desto höher ist das Vertrauensniveau. Dieses Verfahren entspricht dem Vergleich handgeschriebener Unterschriften, wenn kein Unterschriftsmuster zur Verfügung steht.

Änderung eines nicht veränderbaren Objekts

Problem: Mit digitalen Unterschriften lassen sich zwar jegliche Änderungen an einem signierten Objekt nachweisen, doch ist es nicht möglich festzustellen, was genau verändert wurde. Wenn ein Objekt also durch eine Signatur gesperrt ist, wird die Signatur durch jede Art von Änderung (auch zulässige Änderungen) erbrochen und ist nicht mehr zum Schutz verwendbar. Eine geringe zulässige Änderung an einem Objekt erbricht die Signatur und öffnet damit vielen unzulässigen Änderungen die Tür. Zulässig sind beispielsweise Veränderungen der mit dem konservierten Objekt gespeicherten Metadaten (z. B. die Korrektur eines Schreibfehlers oder die Hinzufügung weiterer Metadaten).

Lösung: Von der Anlage her darauf achten, was durch eine digitale Signatur geschützt ist und was nicht. Verifizieren und Neusignieren von Objekten bei erforderlicher Änderung ermöglichen (analog zum Öffnen und Wiederverschließen physischer Objekte). Nachteil: Signatur gewährt nur seit dem letzten Neusignieren Schutz; man muss darauf vertrauen, dass das System das Objekt korrekt verifiziert und neusigniert und nicht verändert hat. Anlegen mehrerer Schichten von digitalen Signaturen, so dass die inneren Signaturen unangetastet bleiben, während die äußeren gegebenenfalls neuberechnet werden können.

Validierung der Implementierung

Problem: Anwendungen, die digitale Signaturbibliotheken verwenden, lassen sich unter Umständen nur schwer korrekt implementieren. Es ist möglich, Signaturen zu erstellen, die von einer anderen Implementierung nicht verifiziert werden können. Bei einer Netzwerkanwendung wird dies aufgrund des Zusammenwirkens mit anderen Implementierungen sehr schnell spürbar. So geschah es beispielsweise bei der VERS-Implementierung. Die von uns genutzte Verschlüsselungsimplementierung Microsoft RSA erzeugt Signaturen mit ungültigem Format. Bemerkt wurde dies nur, weil wir die Signatur mit einer unabhängigen Implementierung überprüft haben. Bei einer Anwendung zur digi-

talen Archivierung könnten Jahre oder Jahrzehnte vergehen, bis eine Signatur durch eine unabhängige Anwendung überprüft wird und das Problem zutage tritt.

Lösung: Vor Inbetriebnahme eines Archivierungssystems immer ein unabhängiges Validierungssystem zur Verifizierung der Signaturen schreiben lassen. Das unabhängige System muss eine andere digitale Signaturimplementierung verwenden als das zu prüfende System und sollte idealerweise eine unabhängige Entwicklung sein. Vorsicht bei Produkten, die scheinbar unabhängig sind, denen aber die gleiche Schlüsselmaschine zugrunde liegt. Die Fehlersuche bei digitalen Signaturen kann schwierig und frustrierend sein.

Utilisation des signatures numériques dans la conservation des objets électroniques

Andrew Waugh

Les signatures numériques peuvent être utilisées pour détecter des modifications apportées à des objets conservés. L'authenticité de l'objet est un aspect critique dans certaines applications (archivage, par exemple). L'exposé présentera certains des enseignements tirés du projet VERS sur l'utilisation des signatures numériques pour prouver l'authenticité de documents électroniques.

Longévité des signatures numériques

Défi: la vie de l'objet numérique signé doit être plus courte que le temps nécessaire pour décrypter la signature numérique. La durée de vie d'un objet conservé se mesure en siècles. Durant cette période, les algorithmes de déchiffrement iront s'améliorant et tourneront sur des ordinateurs plus rapides.

Solutions: choisir la technologie la plus sûre disponible aujourd'hui. Procéder à de nouvelles signatures périodiques avec la technologie courante. Analyser les risques: quel investissement un individu est-il prêt à consacrer pour percer le code d'un objet?

Protection de la clé privée

Défi: prouver l'identité du signataire implique que ce dernier conserve sa clé privée secrète. Les technologies appropriées à cet effet ne sont pas encore très diffusées; les clés sont généralement mémorisées dans des ordinateurs et ne sont protégées que par des mots de passe. Toute lacune (accidentelle ou volontaire) dans la sécurité de l'ordinateur permet à n'importe qui d'accéder à la clé privée et, partant, lui donne la capacité de déchiffrer la signature.

Solutions: ne pas miser de façon trop excessive sur la signature numérique pour prouver l'identité du signataire. Les cartes à puce et la bio-informatique seront peut-être la voie de demain, mais elles ne sont pas encore assez rentables pour être déployées.

Authentification de la clé publique

Défi: valider la signature numérique implique d'authentifier la clé publique du signataire. Cette validation s'obtient généralement à partir d'une infrastructure de clé publique, qui consiste en une série de certificats émis par des organismes de confiance. Cette procédure est acceptable lorsque les signatures ont une durée de vie courte, mais il n'est guère proba-

ble que les autorités de certification puissent conserver les certificats pendant toute la durée de vie de l'objet conservé.

Solutions: tout d'abord, créer votre propre organisme de certification. Il existe en effet des implications pour la sécurité: si un faussaire peut ajouter un certificat, il peut réussir à déchiffrer des signatures. Deuxièmement, mémoriser le certificat dans l'objet conservé. Le problème ne s'en trouvera pas complètement résolu, puisque vous ne pouvez pas vérifier le certificat racine indépendamment. Troisièmement, mémoriser la clé publique dans l'objet signé. Pour vérifier, comparez les clés publiques d'un grand nombre d'objets réputés signés par le même signataire. Si les clés publiques sont identiques, vous pouvez raisonnablement estimer que le document n'a pas été décrypté. Plus les objets examinés seront nombreux, plus votre confiance sera grande. C'est la même méthode que l'on utilise pour vérifier des signatures manuscrites, lorsqu'il est impossible d'obtenir un original de la signature.

Modification d'un objet non modifiable

Défi: les signatures numériques permettent de détecter toute modification apportée à un objet signé, mais il n'est pas possible de découvrir ce qui a été modifié. Cela signifie que chaque fois qu'un objet est verrouillé par une signature, toute modification (même valide) casse la signature numérique, qui devient alors inopérante pour la protection de l'objet. Une modification mineure valide apportée à un objet casse la signature et ouvre la voie à des modifications non autorisées nombreuses. Des modifications valides peuvent impliquer de modifier les métadonnées incluses dans l'objet conservé (correction d'une erreur d'orthographe dans les métadonnées, ajout de nouvelles métadonnées, etc.).

Solution: étudier soigneusement ce qui est protégé par signature numérique, et ce qui ne l'est pas. Vérifier et signer à nouveau les objets lorsqu'il est nécessaire de les modifier (c'est la même chose lorsqu'on décachette et que l'on recachette un objet matériel). Cette solution a l'inconvénient que la signature ne protège que depuis la dernière opération de signature; elle oblige à supposer que le système a vérifié et resigné correctement l'objet, sans y apporter de modifications. Cette solution permet de superposer des couches de signatures, de sorte que les signatures internes demeurent intactes alors que les signatures externes peuvent être recalculées chaque fois que nécessaire.

Validation de l'implémentation

Défi: les applications utilisant des bibliothèques de signatures numériques peuvent être difficiles à implémenter correctement. Des signatures peuvent être créées, qu'une autre implémentation ne pourra pas vérifier. Dans une application réseau, cette impossibilité deviendra vite apparente du fait de l'interaction avec d'autres implémentations. C'est ce qui s'est produit avec l'implémentation VERS: la version cryptographique Microsoft RSA, que nous avions utilisée, a produit des signatures dans un format non valide. Nous n'avons pu le constater que lorsque nous avons vérifié la signature avec une version indépendante. Dans une application de conservation numérique, des années ou des décennies pourront s'écouler avant que la signature soit vérifiée par une application indépendante et que le problème devienne apparent.

Solution: utiliser toujours un système de validation indépendant conçu pour vérifier les signatures, avant de mettre en place un système de conservation. Ce système indépendant doit utiliser une implémentation de signatures différente du système de production et, dans l'idéal, doit avoir été développé par un développeur indépendant. Veiller à ce que des produits qui en apparence sont indépendants fassent en fait appel au même moteur de chiffrement. Il peut être difficile et frustrant de rechercher des erreurs dans des signatures numériques.

Promoting and supporting open archives in Europe: The Open Archives Forum project

Donatella Castelli

1. Introduction

The 'open archives' approach is gaining popularity in several communities such as the cultural heritage, scientific and learning communities. The technology adopted makes it possible to ensure cross-archive access, and low-cost dissemination and access to content that is currently 'hidden' to the wider public at a very low cost. If the full potential of this approach is fulfilled, it will have a significant impact on the building of new service provision. It is essential that European data and service providers be prepared to participate in this innovation.

The building of a cross-archive access service from distributed archives has been firstly made feasible by the Z39.50 community [1]. The Z39.50 protocol allows complying systems to search on multiple distributed databases in a homogeneous manner, thus permitting to view the selected information sources as a unique library. Although a Z39.50 profile for access to digital collections has been developed, the Z39.50 protocol has been mainly used for bibliographic databases.

Recently, the scientific communities that publish their pre-prints on electronic archives have raised again the problem of building an infrastructure for supporting cross-archive services. In a meeting held in Santa Fe (New Mexico) in July 1999, they set up the open archives initiative (OAI) [2] whose objective is to develop and promote interoperability solutions that aim to facilitate the efficient dissemination of content. OAI proposed a simple protocol, the open archives initiative metadata harvesting protocol (OAI-PMH) that, once implemented by an archive, allows third parties services to harvest its records. At present, more than 60 large archives have adhered to this initiative implementing the suggested protocol and more than 10 organisations have implemented cross-archive services relying on this protocol.

Other general solutions are currently being experimented by different communities, mainly scientific communities [3]. These solutions often reflect the different requirements exposed by these communities, in particular the services they need and the kind of information stored in their archives e.g. large numeric/scientific datasets, medical and satellite images, moving images and sound, molecular structures and engineering records.

A lot of work is also underway in the educational environment [4,5]. The opening of the archives offers to the teachers the possibility of training their students by giving them access to a great amount of different information. This offers unexpected possibilities and it is going to revolutionise strongly education at all levels.

The Open Archives Forum (OA-Forum) [6], is a recent EU fifth framework accompanying measure (IST-2001-32015) ([1]) which has been set up for disseminating the open archives-based approach and for supporting the coordination and cooperation among those European and national initiatives that adopt, or are willing to adopt, this approach. The OA-Forum aims to build a framework where both data providers (archives) and service providers can share their experiences and coordinate the development of software tools and infrastructures. Special attention will be initially dedicated to those initiatives, which are implementing or using the OAI-PMH protocol but the plan is to be able to extend the scope also to the other application areas where the need of cross-archives services is emerging.

This paper presents the OA-Forum project by describing its objectives, its context and the activity planned to create a framework for supporting a wider cooperation among the actors that are involved, or are willing to be involved, in the open archive approach.

The rest of the paper is organised as follows. Section 2 briefly introduces the OAI, which is the initiative that inspired the OA-Forum project proposal; Section 3 presents the project activities; and Section 4 describes the next steps of the project.

Donatella Castelli

Donatella Castelli graduated in computer science at the University of Pisa in 1983. From 1983 to 1985 she worked at the Computer Science Department of the University of Pisa on conceptual modelling. In 1986, she was invited as Visiting Researcher to the Computer Science Department of Arizona State University (United States) where she worked on formal aspects of databases and knowledge representation. Since 1987 she has been a member of the research staff of the Istituto di Elaborazione dell'Informazione (IEI) of the Italian National Research Council (CNR). She has been involved in several EU projects on digital libraries. Among these included 'ERCIM digital library — DELOS' (Esprit — long-term research working group No 21057), 'A network of excellence on digital libraries — DELOS' (RTD-V framework network of excellence No IST 1999-12262), European chronicles online — ECHO (RTD-V Framework Project No IST-1999-11994), Coordinator of 'A digital library testbed to support networked scholarly communities — Scholnet', (RTD-V framework project No IST-1999-20664), 'An open collaborative virtual archive environment — Cyclades' (RTD-V framework project No IST-2000-25456), 'Open Archives Forum', (V framework project No IST-2001-32015).
Her current research interests include formal methods, metadata models and digital libraries.

([1]) The participants to the Open Archives Forum project are: UKOLN-University of Bath, Bath (United Kingdom); Humboldt University, Berlin (Germany); Istituto della Elaborazione della Informazione-CNR, Pisa (Italy).

2. The open archives initiative

There are currently several e-print archives that serve large communities of scientists. The most widely know is the Los Alamos e-print archives, called ArXiv.org [7], which at present maintains non peer-reviewed research papers, not only in the physics research areas but also in the mathematics, non-linear systems and computer science. This archive stores more than 150 000 documents, it has a monthly submission rate of 2 000 documents and typically process more than 100 000 connections per day. Other significant archives are NDLTD [8], a digital library that maintains electronic theses and dissertations published in more than 100 different countries and ETRDL [9], a European collection of computer science research reports.

Each of these archives has been set up by a specific community which use them as an instrument for supporting collaboration among its members. Their services address the peculiar needs of their own specific community and are accessible through ad-hoc interfaces. This specificity prevents the exploitation of the full power of these archives since it precludes the possibility of serving cross-domain communities and, when there are more archives, also limit their full utilisation even within the same community. The open archives initiative (OAI) was set up to discuss and solve this problem of interoperability between author self-archiving solutions.

OAI originated by a meeting held on July 1999. The Council of Library and Information Resources (CLIR), the Digital Library Federation (DLF), the Scholarly Publishing and Academic Resources Coalition (SPARC), the Association of Research Libraries (ARL) and the Los Alamos National Laboratory (LANL) sponsored the meeting. The participants were representatives of different scholarly e-print archives. The aim of the meeting was to explore the cooperation among these archives as a way to contribute in a concrete manner to the transformation of the scholarly communication. The participants to the meeting wanted a very simple, low-barrier to entry interface that could shift implementation complexity and operational processing load away from the repositories. They also wanted to minimise the interdependency between the quality of application service as viewed by the user and the behaviour of repositories that supplied data to the application services [10]. Thus, instead of a distributed search model, as that adopted by Z39.50, they chose a 'harvesting-based' approach. According to this approach, service providers take the responsibility for collecting the metadata stored in the repositories and normalising these metadata to support functions as federated searches. To support this approach the OAI designed a protocol, known as 'open archives initiative metadata harvesting protocol' (OA-PMH) that defines the set of requests to which each archive must respond. In designing this protocol, a lot of effort was put in defining requests that could be implemented by each repository with a small effort (one or two days). This requirement was deemed essential for a wide adoption of the interoperability solution.

A lot of discussion about the OAI and its solution for interoperability had taken place over the past years. Many other groups that have similar problems to those faced by the e-print community, i.e. libraries, museum, commercial journal publishers, and communities of scholars who need to share distributed resources, had discussed the possibility of adopting the OAI approach and had raised their own specific requirements. Several archives had implemented the protocol and a number of services had been built on top of it. As a result of this experimentation a number of open questions has been identified. Some of them will get an answer in the revised version of the protocol that will be officially released in June 2002, others, which still need considerable research and experimentation have been left out at the moment. These include very important issues like property issues for metadata, registries, granularity, and metadata schemes.

3. The Open Archives Forum

The Open Archives Forum (OA-Forum) project proposal was motivated by the increasing discussion about the open archives approach raised by the OAI. The aim of the project is to set up a framework for discussing the issues related to open-archive approach in general, i.e. not necessarily related to the OAI technical proposal, and for sharing experiences and solutions. The ultimate goal of the project is to promote this approach in all the areas where its application can provide a framework for the construction of more powerful user services.

The activity of the project is articulated in a number of tasks. Each of them focuses on a specific issue related to the open archives-based approach. The main tasks are the following:

- **Creation of a repository of information about activities related to the open archive approach**

The different European experiences in exposing the archive data to the outside services and in implementing services that use these data have been collected, described and registered on a web-accessible database that supports public search. The database maintains descriptive information about each experience, like the application or research project, the tools used, the people involved and, when possible, existing documentation about it. Particular attention is dedicated to the collection of developed services, like document delivery, searching and browsing harvesting, summarisation, linking, etc. The purpose of this database is to provide a picture of the European open archive research and developments that can be exploited in setting up new projects or reusing software solutions. The OA-Forum consortium is currently exploring the possibility of sharing this database with other organisations, e.g. OAI, in order to create a more complete source of information.

- **Validation of the European experience in implementing and using the OAI-PMH protocol and other similar approaches to interoperability**

The OA-Forum project monitors the ongoing activities that implement mechanisms for supporting the interoperability among open archives in order to highlight limitations, difficulties and emerging software needs. In particular, it records evaluations of these mechanisms in terms of the potential they offer for building new services and the resources they need. It also collects requirements from implementers and users, both in terms of functionality and supporting tools. This information is made public, available on the web site of the project and can be used by anyone who is approaching the development of new open archives and services.

- **Review and analysis of a variety of possible business models for the implementation of open archives**

The added value that the open archives model might offer in different application communities is being explored. This information will lead to an understanding of the issues necessary to ensure that European organisations benefit from the potential of the open-archive technology. Particular attention is dedicated to the issues relating to intellectual property rights, terms and conditions of use, and quality assurance.

- **Dissemination of the open archives approach**

Four European workshops will be organised over the two years' life of the project to increase awareness of the value to the user of open archives, encourage the sharing of experiences and propagation of good practice in opening-up archives. They will be targeted towards different domains, e.g. cultural institutions, research organisations, public sectors, community services and commercial organisations, and towards different classes of users e.g. potential implementers, service providers, archivists. Particular attention will be dedicated to involve communities that may be potentially interested in open archives but which have not yet developed the technology necessary to do it.

The OA-Forum project started its activity in October 2001. The first six months of the project has been mainly dedicated to the set-up of the instruments required to support its activities (web site, database) and to acquire an initial panoramic of the European activity around open archives. In particular, this first part of the project focused mainly on those initiatives that are related to the OAI approach. The project's initial investigation outlined an emerging need of across-archives services in several sectors all around Europe. As a result of this preliminary investigation we found that the situation is very different from northern to southern Europe. The northern countries, e.g. Germany, United Kingdom, Denmark, are deeply involved in the discussion about the open-archives approach. They have implemented the OAI-PMH and participate actively in the experimentation set up by the OAI. On the contrary, the Mediterranean and eastern Europe, i.e. Italy,

France, Greece, Portugal, Spain, Czech Republic, Hungary, etc. with very few exceptions [11, 12, 13] are just now starting to discuss about the open archives approach and the OAI. One of the reasons for such mismatch is certainly related to the presence, in northern countries, (absence in all the others,) of a national programme for supporting the digitalisation and the creation of digital libraries. The national programmes have in many cases favoured the construction of large digitalised archives and have feed the research and experimentation of new solutions to their problems. As a result, in these countries there are many large digitalised archives and the problem of making them cross accessible has been widely discussed. More details about the European map on the open archives activity in Europe can be found by visiting the OA-Forum web site [6].

4. Open Archives Forum future activity

The collection of information about initiatives and projects related to the open archive approach will continue. The scope of the monitoring activity, which started by focusing on the OAI, will be enlarged to consider other communities and other technical approaches. This will involve coming face-to-face with different types of archives and different services. The OA-Forum database will be updated and, according to the requests received, its design may be re-thought to add further information. The organisational and validation activities will widen their scope by evaluating and comparing the different requirements and solutions.

A first step towards the dissemination of the open archives approach to other communities will be achieved at the first open archives workshop, titled 'Creating a European forum for open archives activities', which will be held on 13–14 May 2002 at IEI-CNR in Pisa (Italy). This two-day workshop seeks to bring together researchers, technical implementers and project managers who are experimenting, or are willing to experiment, the open archives approach. The goal of the workshop is to set up the basis of a European forum for sharing experiences and solutions and encouraging networking among projects. The first day of the workshop will be dedicated to an overview of the OAI activity in Europe. Invited speakers from different European countries will illustrate their experiences in opening their e-print archives and libraries in accordance with the OAI-PMH protocol and in implementing cross-archive services based on it. The second day will see the involvement of other communities, in particular those that use scientific data archives. The coexistence of these different communities will allow discussion and the opportunity to compare incentives, and technical and organisational challenges faced by each community. It will also offer the possibility of understanding which aspects might be generalised to serve other realities.

The project plans to continue this process of enlarging the set of communities in the future, moving from those that are strictly related to the scientific world to those more closed to the real life, such as cultural, educational, public administration, etc.

The project will also provide reviews prepared by experts on main topics related to the open archives approach. An expert review report on property rights has already been planned. In the world proposed by the OAI-PMH, metadata resources become migratory and recombinant. This creates the problem of expressing and guaranteeing metadata ownership and permissible uses and reuses. The representation and handling of property rights is one of the open questions for most of the archives that are planning to open themselves to cross-archive services.

Both the results of the projects workshops and the experts reviews reports will be made publicly available on the project web site (http://www.oaforum.org).

References

[1] Z39.50 Maintenance Agency, http://lcweb.loc.gov/z3950/agency

[2] Open archives initiative, http://www.openarchives.org

[3] Astrophysical Virtual Observatory, http://www.eso.org/projects/avo/intro.html

[4] Re-inventing scholarly information dissemination and use, http://elib.cs.berkeley.edu/

[5] National science, mathematics, engineering, and technologies education digital library, http://www.nsdl.nsf.gov/tourl.html

[6] Open Archives Forum, http://www.oaforum.org

[7] arXiv.org e-print archive, http://arXiv.org

[8] Networked digital library of theses and dissertations, http://www.ndltd.org

[9] Ercim technical reference digital library, http://dienst.iei.pi.cnr.it/

[10] Lynch, C., 'Metadata harvesting and the open archives initiative', ARL bi-monthly report 217, http://www.arl.org/newsltr/217/mph.html

[11] Cyclades, *An open collaborative virtual archive environment*, http://www.ercim.org/cyclades/

[12] Torii, 'An open portal over open archives', High Energy Physics Libraries Webzine, http://library.cern.ch/HEPLW/4/papers/4/

[13] Budapest open access initiative, http://www.spros.org/openaccess/read.shtml

Promoción y apoyo a los archivos abiertos en Europa: el proyecto «Foro de Archivos Abiertos»

Donatella Castelli

El enfoque de los «archivos abiertos» está adquiriendo renombre en el ámbito del patrimonio cultural, de las comunidades basadas en el conocimiento y de las comunidades de enseñanza. La tecnología adoptada permite garantizar el acceso a todos los archivos, así como la difusión a bajo coste y el acceso a contenidos que actualmente están «ocultos» al público, a un coste muy bajo. Si se explota plenamente este enfoque, tendrá un impacto significativo en el establecimiento de nuevas prestaciones de servicios. Es esencial que los prestatarios europeos de servicios y datos estén dispuestos a participar en esta innovación.

El Foro de Archivos Abiertos (Foro AA) (http://www.oa-forum.org) es una medida de acompañamiento reciente del Quinto Programa Marco de Investigación de la Unión Europea (IST-2001-32015) ([1]), que se creó para apoyar las iniciativas europeas y nacionales que adoptan, o están dispuestas a adoptar, un enfoque basado en archivos abiertos. El Foro AA creará un marco en el que tanto los proveedores de datos (archivos) como los proveedores de servicios puedan compartir sus experiencias y coordinar el desarrollo de infraestructuras y herramientas de *software*. Se prestará especial atención a las iniciativas que apliquen o utilicen el protocolo de recogida de metadatos (OAI-MHP) (http://www.openarchives.org).

Las principales actividades de este proyecto tendrán como objetivo lo siguiente:

- Divulgación de los conceptos básicos que subyacen en la filosofía de archivos abiertos y la Iniciativa de Archivos Abiertos (IAA). El Foro AA proporcionará un vocabulario común para la aplicación de las IAA en Europa, creando un glosario y definiendo los antecedentes y el contexto de conceptos tales como recogida, uso aceptable, derechos, propiedad, resolución del identificador, uso común de metadatos, etc.

- Fomento del uso del enfoque de archivos abiertos, proporcionando acceso a una gama más amplia de recursos digitales. Se organizarán cuatro talleres europeos a lo largo de dos años, destinados a lograr una mayor sensibilización respecto del valor para el usuario de los archivos abiertos, a fomentar las experiencias compartidas y a propagar las buenas prácticas en materia de apertura de archivos. Estos seminarios irán dirigidos a distintos sectores, como por ejemplo instituciones culturales, organizaciones de investigación, sectores públicos, servicios comunitarios y organizaciones comerciales, y a distintos tipos de usuarios, como posibles ejecutores, proveedores de servicios o archiveros.

([1]) Los participantes en el proyecto del Foro de Archivos Abiertos son los siguientes: UKOLN-Universidad de Bath (Reino Unido); Universidad de Humboldt, Berlín (Alemania); e Istituto della Elaborazione della Informazione-CNR, Pisa (Italia).

- Validación de la experiencia europea en cuanto a la aplicación y la utilización del protocolo de recogida de la IAA y otros enfoques similares de interoperabilidad. La información sobre las distintas experiencias se recogerá y se pondrá a disposición del público. Estos conocimientos supondrán información útil para proyectos y actividades futuras basados en archivos abiertos e interoperables. El proyecto también estimulará a las iniciativas europeas a compartir y reutilizar soluciones de *software*. A fin de contribuir a estos procesos, el proyecto creará un inventario accesible vía Internet de productos informáticos en uso y en desarrollo en Europa y otros puntos. Cada producto del inventario irá acompañado de una descripción de los metadatos y, cuando sea posible, de una evaluación del producto, de la documentación existente, etc.

- Revisión y análisis de distintos modelos empresariales posibles para la aplicación de los archivos abiertos y exploración del valor añadido que podría ofrecer el modelo de archivos abiertos en distintas comunidades que lo apliquen. Esto llevará a una comprensión de las cuestiones necesarias para garantizar que las organizaciones europeas se beneficien del potencial de la tecnología de archivos abiertos. Se prestará especial atención a las cuestiones relativas a los derechos de propiedad intelectual, las condiciones de utilización y la garantía de calidad, en un sistema basado en la IAA.

El documento presentará una descripción de la actividad europea en materia de archivos abiertos e informará acerca de los resultados obtenidos por cada una de las actividades del proyecto anteriormente mencionadas.

Förderung und Unterstützung offener Archive in Europa: das Projekt „Open Archives Forum“

Donatella Castelli

Das Konzept offener Archive findet in den Bereichen kulturelles Erbe, wissensbasierte Gesellschaft und Lernen immer größeren Anklang. Mit der eingeführten Technologie ist es möglich, für einen Zugang zu mehreren Archiven gleichzeitig und eine kostengünstige Verbreitung zu sorgen und zu sehr geringen Kosten auf Inhalte zuzugreifen, die der größeren Öffentlichkeit bisher „vorenthalten“ werden. Wird das Potenzial dieses Ansatzes voll ausgeschöpft, ergeben sich bedeutende Auswirkungen auf die Errichtung neuer Dienste. Für die europäischen Daten- und Diensteanbieter gilt es, sich auf die Teilnahme an dieser Neuerung vorzubereiten.

Das Open Archives Forum (OA-Forum) (http://www.oa-forum.org) ist eine Begleitende Maßnahme des 5. EU-Rahmenprogramms (IST-2001-32015) ([1]). Es wurde eingerichtet, um europäische und einzelstaatliche Initiativen zu unterstützen, die sich für die Einführung eines auf offenen Archiven beruhenden Konzepts einsetzen. Das OA-Forum wird einen Rahmen bilden, in dem Datenanbieter (Archive) und Diensteanbieter ihre Erfahrungen austauschen und die Entwicklung von Software-Tools und Infrastrukturen koordinieren können. Besondere Aufmerksamkeit erhalten dabei Initiativen, die das Metadata Harvesting Protocol der Open Archives Initiative (http://www.openarchives.org) implementieren oder anwenden.

([1]) Am Projekt Open Archives Forum sind beteiligt: University of Bath, Bath (Vereinigtes Königreich); Humboldt-Universität Berlin (Deutschland); Istituto della Elaborazione della Informazione-CNR, Pisa (Italien).

Die wichtigsten Aktivitäten dieses Projekts haben folgende Ziele:

- Verbreitung der Grundbegriffe der Philosophie offener Archive und der Open Archives Initiative (OAI). Das OA-Forum wird einen gemeinsamen Fachwortschatz für OAI-Implementierungen in Europa bereitstellen und dazu ein Glossar mit Definitionen zu Geschichte und Hintergrund von Begriffen wie Harvesting, Acceptable Use, Rechte, Ownership, Identifier-Auflösung, gemeinsame Nutzung von Metadaten usw. erstellen.

- Förderung der Nutzung des Konzepts offener Archive bei der Zugänglichmachung einer breiteren Palette von digitalen Ressourcen. Im Zeitraum von zwei Jahren werden vier europäische Workshops veranstaltet, um die Vorteile offener Archive für die Benutzer zu propagieren, den Erfahrungsaustausch anzuregen und bewährte Verfahren der Öffnung von Archiven bekannt zu machen. Die Workshops werden sich an verschiedene Bereiche wenden, z. B. Kultureinrichtungen, Forschungseinrichtungen, den öffentlichen Sektor, soziale Dienste und kommerzielle Organisationen, sowie an verschiedene Kategorien von Benutzern, wie z. B. potenzielle Anwender, Diensteanbieter und Archivare.

- Validierung der europäischen Erfahrungen bei der Implementierung und Nutzung des OAI Harvesting Procol und ähnlicher Ansätze für die Interoperabilität. Informationen über die verschiedenen Erfahrungen werden eingeholt und der Öffentlichkeit zur Verfügung gestellt. Diese Erkenntnisse werden einen wertvollen Beitrag zu künftigen Projekten und Aktivitäten auf der Basis offener und interoperabler Archive leisten. Außerdem wird das Projekt europäische Initiativen zur gemeinsamen Nutzung und Wiederverwendung von Softwarelösungen fördern. Um die entsprechenden Prozesse mit Informationen zu versorgen, wird das Projekt ein über das Internet zugängliches Verzeichnis einrichten, das Angaben zu in Gebrauch und in der Entwicklung befindlichen Softwareprodukten in Europa und andernorts enthält. Jedem Produkt im Verzeichnis sind eine Metadatenbeschreibung und, sofern möglich, eine Produktbewertung, Dokumentation usw. beigefügt.

- Überprüfung und Analyse der verschiedensten möglichen Geschäftsmodelle für die Implementierung offener Archive und Erkundung des Zugewinns, den das offene Archivmodell den unterschiedlichen Anwendergruppen bieten kann. Auf diese Weise lässt sich besser verstehen, was zu tun ist, damit europäische Organisationen von den Potenzialen der Technologie offener Archive profitieren können. Besondere Beachtung erhalten die Fragen im Zusammenhang mit geistigen Eigentumsrechten, Benutzungsbedingungen und Qualitätssicherung in einem OAI-basierten System.

Der Beitrag vermittelt einen Überblick über die europäischen Aktivitäten im Bereich offene Archive und berichtet über die bisherigen Resultate der einzelnen Projektfelder.

Promotion et soutien aux archives ouvertes en Europe: projet «Open Archives Forum»

Donatella Castelli

Les “archives ouvertes” gagnent de plus en plus en popularité dans le monde du patrimoine culturel, de la connaissance et de la formation. Les technologies adoptées permettent de garantir un accès à toutes les archives, une diffusion à faible coût et un accès à des contenus actuellement “cachés” au grand public à un coût très modique. Si le potentiel de cette approche est pleinement exploité, il aura un impact significatif sur la constitution de nou-

veaux services. Il est primordial que les prestataires de données et de services européens se préparent à prendre part à cette innovation.

Le Forum sur les archives ouvertes (OA-Forum) (http://www.oa-forum.org) est une mesure d'accompagnement récente du 5e programme-cadre de recherche communautaire (TSI-2001-32015) ([1]); il a été mis en place pour soutenir les initiatives européennes et nationales qui visent à adopter une stratégie reposant sur les archives ouvertes, ou qui souhaitent le faire. Le projet OA-Forum mettra en place un cadre dans lequel les fournisseurs de données (archives) et les prestataires de services pourront partager leur expérience et coordonner le développement d'outils logiciels et d'infrastructures. Les initiatives qui implémentent ou utilisent le protocole de recherche de métadonnées OAI-MHP (Open Archives Initiative-Metadata Harvesting Protocol) (http://www.openarchives.org) bénéficieront d'une attention particulière.

Les principales activités de ce projet viseront à:

- Diffuser les concepts de base sous-tendant la stratégie des archives ouvertes et de l'initiative pour les archives ouvertes (OAI). L'OA-Forum mettra au point un vocabulaire commun pour les implémentations OAI en Europe, en créant un glossaire et en définissant l'historique et le contexte de concepts tels que *harvesting* (recherche), utilisation acceptable, droits, propriété, résolution d'identifiant, partage de métadonnées, etc.

- Promouvoir l'utilisation de la stratégie des archives ouvertes dans la fourniture d'accès à un large éventail de ressources numériques. Quatre ateliers européens seront organisés sur deux ans pour mieux faire connaître la valeur des archives ouvertes aux chercheurs, encourager le partage des expériences et la diffusion des bonnes pratiques dans l'ouverture des archives. Ces ateliers cibleront différents publics: organismes culturels, centres de recherche, secteur public, services locaux, organisations commerciales, etc. Ils cibleront également différentes catégories d'utilisateurs: implémenteurs potentiels, prestataires de services, archivistes, etc.

- Valider l'expérience européenne dans la mise en œuvre et l'utilisation du protocole de recherche OAI et dans d'autres approches de l'interopérabilité. L'information sur les différentes expériences sera recensée et mise à la disposition du public. Ces connaissances contribueront utilement aux projets et aux activités futurs sur les archives ouvertes et interopérables. Le projet encouragera également les initiatives européennes à partager et à réutiliser les solutions logicielles. Pour faciliter ce partage, le projet mettra en place sur le net un inventaire des logiciels utilisés et en cours de développement, en Europe et ailleurs. Chaque produit de l'inventaire sera accompagné d'une description des métadonnées et, si possible, d'une évaluation du produit, de la documentation existante, etc.

- Examiner et analyser divers modèles de gestion possibles pour la mise en œuvre des archives ouvertes, et voir la valeur ajoutée que le modèle des archives ouvertes pourrait offrir dans différents milieux d'application. Cette analyse permettra de comprendre les aspects essentiels à prendre en compte pour que les organisations européennes bénéficient du potentiel de cette technologie. Les questions concernant les droits de propriété industrielle, les conditions d'utilisation et l'assurance qualité dans un environnement OAI bénéficieront d'une attention spéciale.

L'exposé présentera un aperçu de l'activité menée en Europe sur les archives ouvertes et rendra compte des résultats obtenus dans chacune des activités du projet décrites ci-dessus.

([1]) Les participants au projet "Open Archives Forum" sont: UKOLN-University of Bath, Bath (Royaume-Uni); université Humboldt, Berlin (Allemagne); Istituto della Elaborazione della Informazione-CNR, Pise (Italie).

Knowledge and action for digital preservation: Progress in the US Government

Kenneth Thibodeau

Kenneth Thibodeau

Kenneth Thibodeau is a widely recognised expert in the management and preservation of electronic records. With 26 years experience in the field, he has spoken at over 120 conferences worldwide and has published more than 20 articles and chapters, in two languages, on these topics. He is Director of the Electronic Records Archives Programme at the National Archives and Records Administration in Washington and chairs the Preservation Task Force of the Interpares Project. Among other accomplishments, he directed the US Department of Defence Task Force that developed the DoD Standard 5015.2 for records management applications and served as secretary of the Committee on Electronic Records of the International Council on Archives and principal editor for the Committee's Guide to managing electronic records from an archival perspective. A fellow of the Society of American Archivists, he holds a Ph.D. in the history and sociology of science.

In 1998, the National Archives and Records Administration (NARA) of the United States launched a new initiative to tackle the daunting challenges posed by electronic records. Although NARA had over 30 years of experience in this area, it was clear that the functional capabilities and quantitative capacity it had developed in that time were insufficient to cope with the ever-increasing variety and complexity, and the exponentially growing quantities of electronic records being produced by the US Government. The situation elsewhere was not particularly encouraging. NARA had extensive knowledge of what had been accomplished and was being explored by other institutions around the world, gained through participation in activities such as the International Council on Archives, the Interpares project, and the development of the open archival information system (OAIS) reference model. The state of affairs in 1998 could easily be summarised:

- proven methods for preserving and providing sustained access to electronic records were limited to the simplest forms of digital objects;
- even in those areas, proven methods were incapable of being scaled to a level sufficient to cope with the expected growth of electronic records; and
- archival science had not responded to the challenge of electronic records sufficiently to provide a sound intellectual foundation for articulating archival policies, strategies, and standards for electronic records.

In this environment, it seemed prudent to pursue a strategy of divide and conquer, partitioning the challenge of electronic records into manageable segments. The problem of the increasing diversity of electronic records was due to the different formats or data types in which digital data is encoded. To cope with this problem, NARA articulated a 10-year plan addressing different classes of data types sequentially in a three-step process. First, the characteristics of each class would be analysed. Then, options for preserving those characteristics would be evaluated. Finally, the best option would be chosen as the basis for developing the capability for preserving records in that class.

Parallel to this, NARA initiated a special project to address the worst-case scenario, which we had already encountered several times; namely, where the archives receive important bodies of digital objects which had not been managed as records, or had been poorly managed. In several of these cases, NARA had received obsolete computer systems and backup media from organisations which had ceased to function. In such cases, there is a basic and difficult problem of culling the records from the mass of other files, such as operating systems, application software, tutorials, help files, and temporary system files. To attack this problem, NARA sponsored a research project in collaboration with the US Army Research Laboratory and Georgia Tech Research Institute. This project focused on the contents of over 550 PC hard drives from the White House during the Administration of the former President Bush. It has developed a pilot system — called the Presidential Electronic Records PilOt System (Perpos) — of automated tools for identifying the data type of each file in a system or stored on media. On this basis, user-created files can be distinguished from software and related files. The project is investigating the application of advanced technologies to characterise the records and identify significant contents.([1])

To improve the knowledge base for digital preservation, NARA continued its support for development of the OAIS standard, and became one of the principal supporters of the International Research on Preservation of Authentic Records in Electronic Systems (Interpares) project. The OAIS effort had the clear advantage of bringing experts from many, independent disciplines and from many countries together to articulate the generic requirements for a system capable of maintaining and delivering information over time. Furthermore, because it was intended as an international standard, it had the potential for influencing the development of information technology products suitable for implementing an OAIS.([2]) The Interpares project had the advantage of focusing the efforts of ex-

([1]) http://perpos.gtri.gatech.edu/

perts from a variety of disciplines and countries on the specific problem of guaranteeing the authenticity of electronic records over time. Since 1998, this project has articulated clearly and coherently the specific requirements for assessing and certifying the authenticity of electronic records, and produced a richer understanding of the relationships between the archival and the digital properties of the records. In its process models of selecting and preserving electronic records, the Interpares project has delineated the decision processes involved in these activities, providing archives and other institutions responsible for digital preservation a firmer and broader basis for evaluating and selecting information technology products on the basis of archival criteria.([3])

The problem of rapidly growing quantities of electronic records demanded a frontal assault in its own right. It combined an engineering problem of scaling input/output operations to cope with enormous numbers of physical files with the archival challenge of identifying and controlling the records and aggregates of records contained in those files. NARA attacked this problem by joining the distributed object computation testbed, a partnership of the Defence Advanced Research Projects Agency, the US Patent and Trademark Office, with the San Diego Supercomputer Centre (SDSC) at the University of California, San Diego, as the leading research centre. The specific challenge NARA added to this testbed was to find a way to preserve exponentially growing quantities of electronic records while respecting the archival principles of provenance and original order, and without raising unnecessary barriers to access to the records.([4])

This project took those initial constraints, of respecting provenance and original order and optimising accessibility, and recast them as the guiding principles of a comprehensive method for digital preservation, initially termed 'persistent object preservation.' This method reverses the focus that had characterised previous efforts to develop techniques for digital preservation. Techniques like emulation and migration focus on overcoming technological obsolescence. In contrast, persistent object preservation focuses on the essential properties of the objects that are to be preserved and insulates those properties from the effects of continuing change in information technology. For archives, the objects to be preserved are records and archival aggregates of records, rather than data types. Options for preservation should be selected on the basis that they maintain the content and documentary form of individual records and the organic relationships among records in files, series, and archival fonds. Given that most electronic records and aggregates are created in data types which are subject to obsolescence, the persistent object method entails transforming the digital objects into formats that are more suited for long-term maintenance and access. Superficially, this looks like migration, but there is a basic difference. Migration techniques typically move objects from older to newer formats, where the control on the process is at the level of data types. But data types do not necessarily correlate with the essential characteristics of records. Hence, migration can only guarantee that old data remains accessible in new technology. It cannot guarantee that records remain authentic. Rather than seeking to keep old software working or translating stored information to new data formats, the persistent object method makes the essential features of the objects to be preserved explicit in formal models and encapsulates the objects in metadata defined in those models, thus differentiating between data (encapsulated objects stored as bit sequences) and information (syntactic and semantic relationships among data articulated in models). It enables systems to evolve independently at the data and information levels by using software mediators between them. Similarly, a system can be designed to provide the basic functions of ingest, management, and dissemination, prescribed by the OAIS model, using software mediators to make each function relatively independent of the others. Thus, over time, the hardware and software used to implement any function, at any level, can be replaced without requiring changes in other parts of the architecture or in the collections being preserved. Only the mediators that provide the interfaces need to be adjusted in response to such change.([5])

The relationship between NARA staff and the scientists and engineers at SDSC developed into a mutually reinforcing collaboration. The results of successive rounds of research defined for archivists unprecedented possibilities for approaching the challenge of preserving electronic records from a systematic, open-ended perspective. Conversely, as the computer scientists and engineers conducting the research progressively enriched their understanding of the archival challenge, they were articulating a comprehensive, knowledge-based information management architecture for digital preservation which not only incorporated the full range of archival requirements from accession to dissemination and from individual records to archival fonds, but also offered clear potential for application in other fields, including not only digital libraries, but also research in natural sciences such as astronomy, physics, and neurosciences.([6])

([2]) Consultative Committee for Space Data Systems, 'Reference model for an open archival information system, CCSDS 650.0-R-2, *Red book*, July 2001. The model has been adopted and is being published by the ISO. Available at: http://www.ccsds.org/documents/pdf/CCSDS-650.0-R-2.pdf

([3]) http://www.interpares.org/index.htm

([4]) http://www.darpa.mil/ipto/psum1999/d642-0.html See also: http://www.sdsc.edu/DOCT/

([5]) Thibodeau, K., Moore, R., Baru, C., 'Persistent object preservation: advanced computing infrastructure for digital preservation', *DLM-Forum — European citizens and electronic information: the memory of the information society, cooperation Europe-wide*, Brussels, 18–19 October 1999, pp. 113–119, available at: http://europa.eu.int/ISPO/dlm/fulltext/full_thib_en.htm

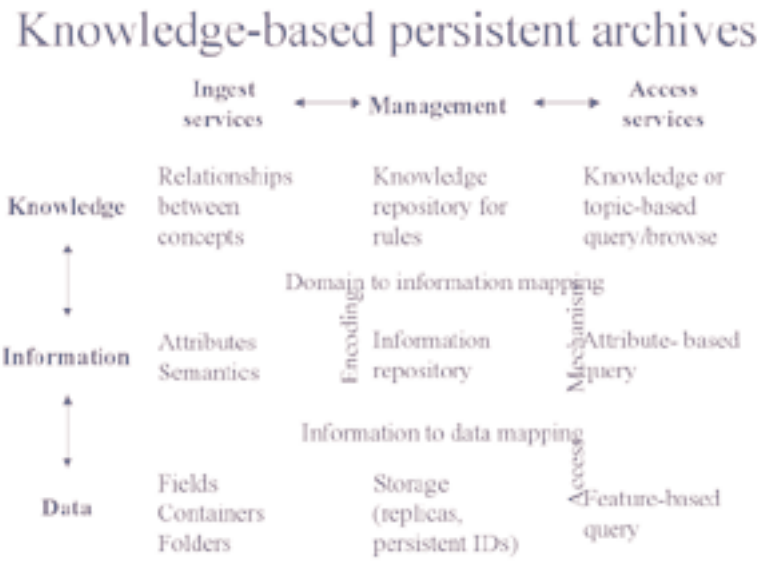
([6]) Rajasekar, A., Moore, R., 'Data and metadata collections for scientific applications', European high performance computing conference, Amsterdam, Holland, 26 June 2001, available at: http://www.npaci.edu/DICE/Pubs/Data-management_moore.pdf

The fertility of this interaction between archivists, on the one hand, and computer scientists and engineers, on the other, is most evident in the maturation of the initial 'persistent object' preservation method into the 'persistent archives' architecture. The persistent archives concept adds a knowledge layer to the architecture, on top of the data and information layers. This addition is at once simple in its basic formulation and prolific in its implications. At bottom, all digital information is stored as data (byte sequences). These byte sequences can only be processed and interpreted correctly given syntactic and semantic context at the information level. Information, then, is defined as valid and meaningful combinations of data. In turn, information can only be understood and exploited in the context of domain-specific knowledge. Knowledge, then, is defined as combinations of pieces of information or knowledge.([7]) The following figure depicts the persistent archives architecture. It is constructed of three rows corresponding to the knowledge, information and data layers, from the top down, and three columns corresponding to the OAIS functions of ingest, maintenance and dissemination, from left to right. The principal entities or processes involved in each function at each level are shown in the individual cells, and the types of mediation services are indicated between the rows and columns.

([7]) Ludäscher, B., Marciano, R., Moore, R., 'Preservation of digital data with self-validating, self-instantiating knowledge-based archive', *ACM Sigmod Record*, 2001, 30(3), pp. 54–63.

To illustrate the differentiation over levels, consider the types of queries that could be executed. At the data level, regardless of how information or knowledge are managed, simple queries such as string searches for character data or pattern matching for other data types could be executed against the stored data. The process would be similar to searches run on the world wide web today. If data were stored with tags that identified attributes, one could incorporate this additional information in slightly more sophisticated searches, such as where tag = 'income',or where tag = 'author' and the data value = 'John Mark'. Moving up to the information level would enable us to apply syntactic and semantic criteria. For example, where a data model indicated that a database contained elements for year, location, earned income, and occupation, one could execute queries asking for the income of electronics engineers in a specified city in 1999. If the system could materialise the structure of the database as indicated in the data model, and populate that structure with the stored data, it would not matter whether these data elements were stored in the same flat file or were scattered in different tables of a complex relational database. In order to reinstantiate the database on a target platform at some time in the future, the system would need a software mediator which could translate the metadata stored in the information repository into terms recognisable by the target technology. The maintenance of the logical model apart from the stored data would provide independence between information management and data storage. The software mediator would achieve independence between the management of information about stored objects and collections and access to the preserved records.

Providing an appropriate response to a question such as, 'Find all diplomatic correspondence concerning the North-American Free Trade Agreement', would require greater sophistication than the examples cited so far. For example, a string search of the web on the phrase, 'North-American Free Trade Agreement', identified over 160 000 relevant items. Adding the phrase 'diplomatic corre-



spondence' reduced the result set to six items, but none of them contained any diplomatic correspondence. An appropriate response to this type of query would benefit from the application of knowledge from several different domains, including foreign affairs: that NAFTA is a treaty among Canada, Mexico, and the United States; government: which agencies of those three governments are likely to engage in diplomacy regarding trade; archival science: what types of records comprise correspondence; what distinguishes diplomatic from other correspondence, and where in the record-keeping systems of the relevant governmental entities would diplomatic correspondence be found. To execute such a query, an archival system would need a knowledge base for topics, such as 'North America' and 'free trade' and another for rules, such as those concerning the conduct of foreign affairs in the three nations and those concerning record-keeping in the government entities. It would also entail mapping the relevant topics and rules to objects and the information level and from there to the data level. Ensuring that the outcome of the query was access to authentic copies of the relevant correspondence would also require application of knowledge about archival requirements for preservation to the information about the management of information about the records, the data that constitute the records, and the processes for maintaining and reproducing them. Separating knowledge management from information management has obvious benefits. It enables tracking of significant changes in records systems over time. For example, diplomatic correspondence within the US Department of State was moved from paper to digital form in the 1970s, and in the intervening decades the electronic system underwent several migrations. Some of the technology changes entailed significant changes in the records, although such changes might only be implicit or not apparent to someone reading only one message at a time. A knowledge base can enable such changes to be made explicit and readily avialable. But it would also enable maintaining coherency across lower-level variability. For example, each of the State Department systems would be represented at the information level by a different schema; however, the diplomatic messages constitute a single records series, one that has been in continuous existence for over a century. A knowledge level concept of the series enables it to be managed as a single archival aggregate, but also makes it possible to translate a query to execute appropriately against difference instantiations of the records system.

The persistent archives approach reflects the conclusion of technologists that it is impossible to satisfy archival requirements for preservation and access to electronic records without systematic infusion of archival knowledge. This domain knowledge is essential to distinguish archives of records from other types of information collections in order to ensure records and archival aggregates of records. There are several areas where archival knowledge needs to be applied. Most obvious is the need to ensure the continuing value of records as evidence of the activities in which they are produced and used. This entails application of the archival principles of provenance and respect for original order. In addition, archival knowledge of requirements for preservation dictates what information is required to assess and assert the authenticity of the preserved records.

Domain knowledge is also needed to understand records. Unlike other types of information objects, such as publications, records are created within a universe of discourse where there is often a high degree of shared information and expectations among participants. This common knowledge includes both specific empirical information about prior steps in a multi-step process, generic knowledge about the process, and expectations about both subsequent steps and the norms for recording and communicating information about the process. In such contexts, important information is often conveyed by form, as well as by substance. In processes involving businesses and government agencies, for example, participants expect certain forms to be used for certain types of transactions, and certain signatures to be required for authorisation of certain types of actions. Common knowledge enables processes to be carried out efficiently and provides a systemic check on their validity and on the reliability of their records. For example, participants in a recurring process are likely to note deviations from prescribed documentary forms or absences of required authorisations. To enable parties who were not participants in a process to understand the records of that activity, often long after the fact, an archival system should contain and convey information about the types of records typically produced, the elements of intrinsic and extrinsic form of each type, the relationships between processes and records, and also the implied knowledge that was either common to participants or would have resulted from recognised discrepancies between generic knowledge and expectations on the one hand and specific instances on the other.

The persistent archives architecture came to fruition thanks to a new partnership NARA formed in March 2000 with the US National Science Foundation's National Partnership for Advanced Computational Infrastructure.([8]) This partnership offered the tantalising prospect of simultaneously fine

([8]) National Archives and Records Administration, 'National archives announces plan for collaboration with National Science Foundation to create an electronic records archives', press release, 28 March 2000, available at: http://www.archives.gov/media_desk/press_releases/nr00-58.html

tuning the preservation method to address specific archival requirements while finding solutions capable of addressing these requirements in mainstream technologies being developed to enable electronic commerce and electronic government. Archives and advanced computation are, in one sense, at opposite ends of an intellectual spectrum: the one seeks to preserve the past the other to create the future. But it would be more accurate to say that, while they focus on different axes, they both face a common problem. To preserve electronic records, archives need in effect to develop interoperability among disparate systems, chronologically spanning generations of information technology. To enable e-commerce, e-government, scientific research, and education in the digital era, those responsible for developing information and communications technologies must provide for interoperability across disparate administrative and political domains, simultaneously spanning several generations of technologies implemented in these domains at any given time. The common need for interoperability in turn entails a need for persistence of the information assets created and used in these domains. These assets have to be carried — intact and authentic — across space, time, technologies, and human (political, institutional, social and cultural) domains. These common interests have generated a sustained and growing partnership between NARA and NSF, and a widening range of research initiatives at the University of California, Berkeley, the University of Maryland, Ohio State University, and the University of Urbino, Italy, as well as at SDSC. Further expansion in the range of active researchers is expected in the near future.([9])

The common interests of archives and institutions promoting advanced technologies are especially important in the government arena. By its nature, a democratic government is responsible to its citizens. Records are an essential instrument of that accountability. If government archives — at national and other levels — cannot find ways to preserve and deliver authentic electronic records over time, governments will not be able to fulfill a fundamental responsibility. The importance of this challenge has been recognised at the highest levels of the US Government. Each year, the President sends to the Congress a supplement to the President's budget, which sets out the overall picture of the government's support for research and development in computer and networking technologies. In the Supplement for the fiscal year 2002, the Interagency Working Group on Information Technology Research and Development identified 'managing and enabling worlds of knowledge' as one of the major challenges facing the nation. Within this context, it asserted, 'Strategies to assure long-term preservation of digital records constitute another particularly pressing issue for research. As storage technologies evolve with increasing speed to cope with the growing demand for storage space, the obsolescence of older storage hardware and software threatens to cut us off from the electronically stored past.' It also specified, 'How to determine, collect, and preserve what is of value in the world's dizzying new digital output now joins older questions of how and what to digitise from humanity's pre-digital knowledge stores as issues for archivists.'([10]) These assertions are not simply abstract statements of principle. The Congress translated this direction into real support for the preservation of electronic records in the NARA appropriation for fiscal year 2002. The current NARA budget includes USD 22 302 000 for preserving electronic records. While some of this money is being used for current operations, the majority is devoted to undertaking the development of a new archival system for electronic records. According to John Carlin, the Archivist of the United States, the Electronic Records Archives 'is NARA's strategic response to the challenge of preserving, managing, and accessing electronic records. Our goal is to build a nationwide digital archive that preserves and provides access to virtually any type of electronic record created anywhere in the Federal Government at any time.'([11])

([9]) Thibodeau, K., 'Building the archives of the future — advances in preserving electronic records at the National Archives and Records Administration', D-Lib Magazine, February 2001, Volume 7, Number 2. http://www.dlib.org/dlib/february01/thibodeau/02thibodeau.html

([10]) National Science and Technology Council, 'Networking and information technology research and development', Supplement to the President's Budget For FY 2002, a report by the Interagency Working Group on Information Technology Research and Development, Washington, July 2001, pp. 22–23, available at: http://www.ccic.gov/pubs/blue02/index.html

([11]) National Archives and Records Administration, 'ERA vision statement', NARA notice: NARA 2002-135, 10 May 2002.

Progresos realizados en la administración americana por lo que respecta al conocimiento y a la acción en el ámbito de la conservación digital

Keneth Thibodeau

Desde 1998, la Administración Nacional de Archivos y Documentos (NARA) de Estados Unidos ha patrocinado y colaborado en varias iniciativas relativas a diversas cuestiones, desde infraestructura informática avanzada hasta los requisitos para garantizar la autenticidad de los archivos. Estos proyectos han supuesto importantes asociaciones con otros organismos del Gobierno estadounidense, tales como National Science Foundation, Defense Advanced Research Projects Agency y Army Research Laboratory; con universidades y organismos de investigación, incluidas las universidades de California, Maryland y Urbino y el Georgia Tech Research Institute, y colaboraciones multinacionales y multidisciplinares tales como el proyecto InterPARES y la Global Electronic Records Association. Esta diversidad de iniciativas se ha centrado en los objetivos básicos de desarrollar los conocimientos necesarios para comprender los requisitos y evaluar las opciones para la conservación a largo plazo de los documentos electrónicos y el acceso continuo a ellos, así como de fomentar el desarrollo o la transferencia de las tecnologías necesarias para realizar con éxito la conservación y el acceso.

Los resultados de estas actividades son los siguientes: 1) un cambio fundamental en el concepto de los retos de la conservación y del acceso continuo a la información digital; 2) grandes progresos en las áreas de la informática y de la ingeniería, que son críticas para tener éxito al abordar estos retos; 3) la concentración de la dirección estratégica de la NARA en construir soluciones sólidas para la conservación y el acceso a los documentos electrónicos, y 4) grandes beneficios en cuanto al reconocimiento de la importancia y el valor de estos esfuerzos y su aplicabilidad a una amplia gama de importantes actividades.

Nuestro concepto de los retos relativos a la conservación de los archivos electrónicos ha variado desde el limitado objetivo de encontrar medios para superar problemas tecnológicos tales como la obsolescencia del *hardware* y del *software* y la fragilidad de los medios digitales, hasta la articulación de un modelo abstracto en el que la conservación y el acceso continuo se consideran íntimamente ligados entre sí y con la infraestructura de la tecnología de la información necesaria para apoyar el comercio electrónico y la administración electrónica.

Esta visión conceptual se ha articulado paralelamente al progreso del desarrollo de una arquitectura de gestión de la información para archivos permanentes basada en el principio de que la conservación de la información digital debe ser esencialmente independiente de la tecnología de la información específica utilizada para aplicar la solución. Este principio de independencia de la infraestructura tiene dos facetas esenciales. En primer lugar, aísla los objetos digitales y cualquier recogida arbitraria de objetos digitales, incluso en el nivel de los fondos de archivo, de las vicisitudes de los cambios tecnológicos, a fin de reforzar la autenticidad continua de la información conservada. En segundo lugar, permite que las soluciones se adapten de forma dinámica, tanto para abarcar nuevos tipos de objetos digitales que puedan surgir en el futuro como para poder hacer frente a las demandas de envío de los objetos conservados, realizadas por los usuarios, a través de una variedad grande y cambiante de plataformas informáticas. La arquitectura de los archivos persistentes se ha desarrollado hasta el punto de incluir como componente básico esencial la introducción del conocimiento temático sobre los datos conservados. La necesidad y el valor de incluir el conocimiento del dominio se han demostrado en casos que van desde los registros de la actividad legislativa en el Congreso estadounidense hasta los datos experimentales de las neurociencias.

Hace unos años, frente a los retos desalentadores y crecientes que planteaban los registros electrónicos, la NARA sólo podía articular una estrategia consistente en dividir los problemas

y resolverlos uno por uno. Los proyectos elaborados a finales de los años noventa clasificaron las dificultades en función de las categorías de aplicaciones y los tipos de datos, con un plan subsidiario para abordar cada categoría, esencialmente en forma secuencial. En la actualidad, la arquitectura de los archivos permanentes proporciona la base para una estrategia completa y coherente en la que diferentes categorías de aplicaciones se consideran casos especiales que suponen diferencias marginales, y no básicas, en cuanto a las soluciones. De conformidad con esta estrategia, la NARA ha puesto en marcha el programa de documentos y archivos electrónicos (Electronic Records Archives Program) para construir los archivos del futuro, previstos como un sistema federado distribuido por Internet y capaz de absorber, conservar y expedir virtualmente cualquier tipo de archivos electrónicos.

El éxito de estos esfuerzos se ha reconocido ampliamente, de forma significativa para la NARA, en los más altos niveles de la administración americana. El Gabinete del Presidente, en su plan gubernamental de investigación y desarrollo en el ámbito de las tecnologías de información y la creación de redes, considera la conservación digital en general y la selección y conservación de archivos en particular como elementos importantes para "facilitar la emergencia de mundos de conocimiento". En el presupuesto actual, el Congreso estadounidense ha aprobado la petición de la Casa Blanca de mayor financiación, asignando más de 2 millones de dólares a apoyar los esfuerzos de la NARA para hacer frente a los retos en el ámbito de los archivos electrónicos.

Kenntnisse und Maßnahmen im Bereich digitale Archivierung: der aktuelle Stand bei der US-Regierung

Keneth Thibodeau

Seit 1998 finanziert und beteiligt sich die National Archives and Records Administration (NARA) der Vereinigten Staaten an einer Reihe von Initiativen zur Erforschung der verschiedensten Fragen von der modernen EDV-Infrastruktur bis hin zu Authentizitätsanforderungen bei der Archivierung. Diese Projekte erfolgen auf der Basis einer breit angelegten Partnerschaft mit anderen Stellen der US-Regierung wie der National Science Foundation, der zentralen FuE-Einrichtung des Verteidigungsministeriums (DARPA) und dem Forschungslabor der US-Army, mit Universitäten und Forschungseinrichtungen wie den Universities of California, Maryland und Urbino und dem Georgia Tech Research Institute, sowie im Rahmen multinationaler und multidisziplinärer Vorhaben wie dem Projekt InterPARES und der Global Electronic Records Association. Im Mittelpunkt dieser breiten Palette von Initiativen stehen folgende Kernziele: Erweiterung der erforderlichen Kenntnisse über die Anforderungen der Langzeitarchivierung elektronischer Unterlagen und ihrer ständigen Zugänglichkeit sowie die Bewertung entsprechender Optionen, und Förderung der Entwicklung oder des Transfers der für die Archivierung und den Zugang notwendigen Technologien.

Die bisherigen Ergebnisse dieser Aktivitäten lassen sich folgendermaßen zusammenfassen: (1) ein grundsätzliches Umdenken im Hinblick auf die Aufgabenstellungen der Konservierung und der ständigen Zugänglichkeit von digitalen Informationen, (2) wesentliche Fortschritte im Bereich Informatik, die für Erfolge bei der Erfüllung dieser Aufgabenstellungen entscheidend sind, (3) strategische Ausrichtung der NARA auf die Erarbeitung tragfähiger Lösungen für die Konservierung und Zugänglichkeit elektronischer Unterlagen und (4) große Fortschritte bei der Anerkennung der Bedeutung und des Werts dieser Bemühungen sowie deren Anwendbarkeit auf ein breites Spektrum von Aktivitäten.

Unsere Vorstellung von den mit der Archivierung elektronischer Unterlagen verbundenen Herausforderungen hat sich verlagert und richtet sich nicht mehr allein auf die Suche nach Mitteln zur Überwindung technischer Probleme wie dem raschen Veralten von Hard- und Software und der Unsicherheit digitaler Datenträger. Vielmehr geht es jetzt um die Formulierung eines abstrakten Modells, bei dem Konservierung und ständiger Zugang als untrennbar miteinander verknüpft und eingebunden in die informationstechnische Infrastruktur betrachtet werden, die zur Unterstützung von e-Commerce und e-Government erforderlich ist.

Diese neue Vorstellung entwickelte sich bei gleichzeitigen Fortschritten bei der Herausbildung einer Informationsmanagement-Architektur für dauerhafte Archive, die auf dem Grundsatz beruht, dass die Konservierung digitaler Informationen im Wesentlichen unabhängig von der zur Implementierung verwandten konkreten Informationstechnologie erfolgen muss. Dabei spielen zwei maßgebliche Gesichtspunkte eine Rolle. Zum einen werden die digitalen Objekte und jede willkürliche Sammlung digitaler Objekte – bis hin zur Ebene von Archivbeständen – von den Wechselfällen des technologischen Wandels isoliert, um für die anhaltende Authentizität der konservierten Informationen zu sorgen. Zum anderen lassen sich auf diese Weise Lösungen dynamisch so anpassen, dass neue Arten digitaler Objekte und Sammlungen, die in der Zukunft entstehen können, aufgenommen werden können und dass den Forderungen der Benutzer nach Ausgabe der konservierten Objekte auf den verschiedensten und sich verändernden Rechnerplattformen Rechnung getragen wird. Die Architektur dauerhafter Archive hat sich so weit entwickelt, dass nunmehr Domainkenntnisse zu den konservierten Sammlungen als wesentlicher Bestandteil mit enthalten sind. Notwendigkeit und Wert der Aufnahme von Domainkenntnissen wurden an konkreten Fällen nachgewiesen, die von Unterlagen über gesetzgeberische Tätigkeiten des Kongresses bis hin zu Experimentaldaten der Neurowissenschaften reichen.

Vor ein paar Jahren konnte die NARA angesichts der gewaltigen und rasant zunehmenden Herausforderungen, die elektronische Unterlagen mit sich brachten, lediglich eine Strategie des „Teile und herrsche“ formulieren. Ende der 90er Jahre erstellte Pläne teilten die Herausforderungen nach Anwendungskategorie und Datentyp, wobei anhand von Teilplänen die einzelnen Kategorien im Wesentlichen der Reihenfolge nach zu behandeln waren. Heute bildet die Architektur für dauerhafte Archive die Grundlage für eine umfassende und einheitliche Strategie, bei der unterschiedliche Anwendungskategorien als Spezialfälle betrachtet werden, die geringfügig, aber nicht grundsätzlich unterschiedlicher Lösungen bedürfen. Entsprechend dieser Strategie hat die NARA ein Elektronisches Archivprogramm aufgelegt, das dem Aufbau eines Archivs der Zukunft in Form eines über das Internet verteilten föderalen Systems dient, mit dem praktisch alle Arten von elektronischen Aufzeichnungen aufgenommen, konserviert und ausgegeben werden können.

Der Erfolg dieser Bemühungen findet breite Anerkennung, vor allem auch auf den höchsten Ebenen der amerikanischen Regierung, was für die NARA von großer Bedeutung ist. In dem alle Regierungsstellen umfassenden Plan für Forschung und Entwicklung im Bereich Informations- und Netzwerktechnik hat das Präsidialamt die digitale Archivierung im Allgemeinen und die archivische Bewertung und Konservierung von Unterlagen im Besonderen als wichtige Elemente beim „Aufbau von Welten des Wissens“ anerkannt. Im laufenden Haushalt hat der Kongress dem Antrag des Weißen Hauses für eine erhöhte Mittelausstattung stattgegeben und mehr als 2 Mio. US-Dollar für die Erfüllung der Aufgaben im Bereich elektronische Unterlagen durch die NARA bereitgestellt.

Sur la connaissance et l'action dans le domaine de la conservation numérique: le point sur les progrès réalisés dans l'administration américaine

Keneth Thibodeau

La NARA (National Archives and Records Administration) des États-Unis parraine et collabore à un certain nombre d'initiatives depuis 1998 qui se proposent d'analyser diverses questions, de l'infrastructure informatique avancée aux spécifications requises pour l'authenticité des archives. Ces projets ont impliqué de larges partenariats avec d'autres organismes de l'administration américaine, notamment avec la fondation NSF (National Science Foundation), l'agence Defense Advanced Research Projects Agency, l'Army Research Laboratory, avec des universités et des instituts de recherche, parmi lesquels les universités de Californie, du Mariland et d'Urbino et le Georgia Tech Research Institute, ainsi que des collaborations multinationales et pluridisciplinaires, à l'instar des projets InterPARES et Global Electronic Records Association. Ces initiatives ont ciblé des objectifs de base: faire progresser l'état des connaissances nécessaires pour connaître les conditions requises de la conservation à long terme des documents électroniques et évaluer les solutions possibles, assurer un accès durable à ces archives et encourager le développement ou le transfert des technologies nécessaires pour réussir la conservation des archives et leur accessibilité.

Toutes ces activités ont débouché sur les résultats cumulés suivants: 1) un changement fondamental dans la conception des défis que posent la conservation et un accès durable à l'information numérique; 2) des progrès substantiels dans le domaine de la science et du génie informatiques, qui sont décisifs pour réussir à relever ces défis; 3) une orientation stratégique de la NARA, qui cible la mise en place de solutions robustes pour la conservation des documents électroniques et leur accessibilité et 4) des avancées majeures dans la reconnaissance de l'importance et de la valeur des efforts déployés et dans leurs possibilités d'application à un large éventail d'autres activités.

Notre conception des défis concernant la conservation des archives électroniques a évolué d'une vision étroite, qui consistait à trouver des moyens pour surmonter des problèmes technologiques comme l'obsolescence des matériels et des logiciels et la fragilité des supports numériques, vers la formulation d'un modèle abstrait, dans lequel la conservation et l'accès permanent sont vus comme inextricablement liés l'un à l'autre et comme intimement mêlés aux technologies de l'information nécessaires pour supporter le commerce et l'administration électroniques.

Cette vue conceptuelle s'est trouvée formulée parallèlement aux progrès réalisés dans le développement d'une architecture de gestion documentaire destinée aux archives permanentes, reposant sur le principe selon lequel la conservation d'informations numériques doit être pour l'essentiel indépendante des technologies de l'information spécifiques utilisées pour implémenter la solution. Ce principe de l'indépendance de l'infrastructure revêt deux aspects primordiaux: tout d'abord, il protège les objets numériques et toute autre collection arbitraire d'objets numériques — voire un fonds d'archives — des caprices du changement technologique afin d'assurer l'authenticité constante de l'information conservée. Ensuite, il permet une adaptation dynamique des solutions à la fois pour qu'elles puissent s'étendre aux nouveaux types d'objets numériques et de collections susceptibles d'émerger à terme et pour qu'elles répondent aux demandes des chercheurs souhaitant que les objets conservés leur soient communiqués sur un large éventail de plates-formes informatiques. L'architecture des archives permanentes a évolué jusqu'au stade où elle inclut des connaissances thématiques sur les collections archivées en tant que composante essentielle. La nécessité et l'importance de cette incorporation des connaissances thématiques ont été démontrées dans de nombreux exemples, qui vont des archives de l'activité législative du Congrès américain aux données expérimentales des neurosciences.

Face aux difficultés décourageantes et croissantes que posaient les archives électroniques il y a quelques années, la NARA ne pouvait que formuler une stratégie consistant à scinder les problèmes et à les résoudre un à un. Les projets élaborés à la fin des années 90 classaient les difficultés en fonction des catégories d'applications et des types de données, avec un projet subsidiaire permettant d'aborder chaque catégorie, essentiellement sous forme séquentielle. Aujourd'hui, l'architecture des archives permanentes jette la base d'une stratégie globale et cohérente, dans laquelle différentes catégories d'applications sont perçues comme des cas particuliers, entraînant des différences marginales, et non fondamentales, dans les solutions. Conformément à cette stratégie, la NARA a lancé le programme en faveur des documents et des archives électroniques afin de mettre en place les archives du futur, vues comme un système fédéré réparti sur l'internet et capable d'admettre, de conserver et de mettre à disposition presque n'importe quel type d'archives électroniques.

Le succès de ces efforts est largement reconnu, même aux niveaux les plus élevés de l'administration américaine, ce qui est encore plus important pour la NARA. Dans son plan gouvernemental pour la recherche et le développement dans les technologies de l'information et des réseaux, le cabinet du président a reconnu que la conservation numérique en règle générale, et le tri et la conservation des archives en particulier, constituaient des éléments majeurs pour "faciliter l'émergence des mondes de la connaissance". Dans le budget actuel, le Congrès américain a accédé à la demande d'augmentation de budget de la Maison blanche, et a ainsi fourni plus de 2 millions de dollars pour soutenir les initiatives de la NARA dans le domaine des archives électroniques.

The use of standardised directory services for secure access to information sources

Felix von Bredow

Felix von Bredow

Felix von Bredow (industrial engineer), born 1971, is project leader at Project Consult Unternehmensberatung GmbH, a leading product-neutral consultancy for document and records management, electronic archiving, office automation, groupware, intranet und workflow in Germany.
Since 1996 Felix von Bredow is involved in projects in the areas of archiving, digital signatures, directory services, automatic classification, research, SAP/R3, ASP (application service providing). His main focus at Project Consult is consultancy, project management and the analysis of technological trends. Over the last several years he has published numerous articles in the area of document-related technologies in different specialised periodicals.

Introduction

Today, anyone wishing to make information available to various user groups will have to read, write and manage various user information. Access to defined contents, user authentification, and checking which user has used which information must all be considered in this context. Particular security requirements for protecting one's own information must also usually be taken into account.

Directory services have not been the subject of much public discussion up to now but they may be used for this very purpose. Basically, directory services are standardised systems for recording any (type of) information in a controlled environment. Standardised directory services originated from international telecommunications enterprises. These businesses originally got together in order to build up an international electronic telephone directory.

Direct information about individuals is not, of course, the main object. However, the idea of making public information accessible to various user groups is not too far off.

Directory services may be operated centrally, or on a decentralised or shared basis. A standardised data model, taking into account all information-relevant aspects, such as MoReq, makes it possible to provide users who have been listed once with standardised controlled access to the most varied information sources in different locations.

A directory — which logically needs to be central — of users and user groups gives rise to considerable potential for cost reduction in capturing, managing and integrating information into the most varied information systems and, above all, preventing security problems which may arise from inconsistent or out-of-date directory entries.

Directory services are specialised data bank systems and thus are organised on the principles of universal databases. Unlike these, however, directory services have been developed especially for accessing information from relatively short and simple applications.

Thus, directory services are specialised data banks with limited functions. Because the support of complex transactions can be dispensed with, a directory service can work well even where there is a very large number of entries. Because of its characteristics, a directory service today is suitable for providing controlled access to information, for example over the Internet, with a very high security level.

The use of directory services is imperative for providing comprehensive and controlled access to data collections in public electronic archives.

Function areas

A directory service provides a basic service within a computer network system. As a rule, its services are made accessible to the user via the client system's user interface. Administrators can use the service administrative tools. A directory service has four main function areas:

- Security

In order to protect information from unauthorised use, directory services have mechanisms for the allocation and application of protection rights, for authentification, authorisation and coding.

- Partitioning

For managing large amounts of information, directories are split technically over various repositories or computers in order to achieve optimal performance.

- Distribution

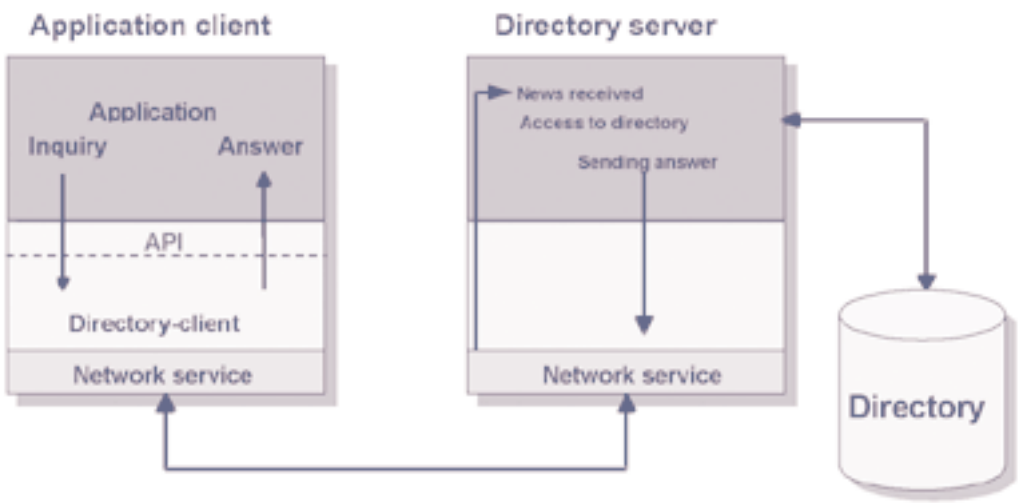
Moreover, directories are distributed, geographically, over a network of computers. These computers work together as a single directory. Search inquiries are transferred from one directory to other directory servers within the network.

- Replication

Directories whose contents are wholly or partially identical are kept up to date by the use of replication mechanisms i.e. information is copied onto other directory servers in the network. These directories may be accessed quickly and securely even by remote users who will not be exposed to inconsistencies.

In directories the client/server model is normally used. For reading or writing information, an application will use a function or 'application programming interface' (API), in order to transfer the corresponding inquiry to a server process which processes the inquiry and reports back the results.

Diagram 1: How a directory service works



Metadirectories — new and old worlds

Many organisations wishing to set up a directory service today are already using other directories in their system e.g. address data banks, network services or mail systems. All of these systems already contain information which are needed for access to e.g. public, electronic archives. In planning a directory service or a network of several systems of this kind, one must decide how to organise it conceptually so that the interests of all participants are taken into account. There is also often the matter of how pre-existing information may be used without having to input it again.

The aim of so-called metadirectories is to synchronise existing information. A metadirectory consists of a conventional directory and tools for filling and synchronising it. Metadirectories can provide applications with consistent information from diverse data sources. A directory service serves as both a consolidation point and a 'normal' directory for providing applications. Integration tools carry out the tasks of data extraction, modification and reformatting. There are particular requirements for this:

- Triggers

In order to make changes in a directory, an impulse must reach the metadirectory so that the changes may be carried out and also made accessible to other applications (event notification).

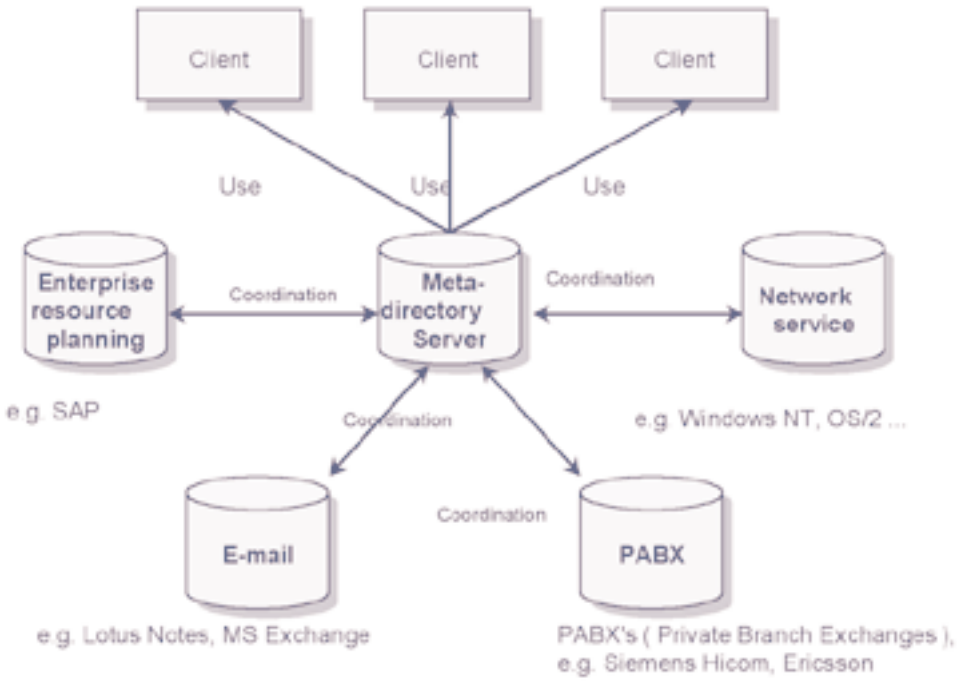


Diagram 2: How a metadirectory works

- Join process rules

The join process rules lay down which objects and attributes of a particular directory are definitive for other directories.

- Filtering

Filtering mechanisms can make the information listed accessible to other directories, so that these may be processed by the most varied applications.

- Scheduling

Not all applications provide the necessary impulse for triggers. In order to coordinate the different directories, times must be fixed at which coordination will occur.

The use of metadirectories does not only allow the most flexible integration of older applications into a directory service but is also the best way of consolidating information from other directory services.

Integration is differentiated according to the information flow and the manner in which information is provided.

Information flow paths:

- One-way synchronisation

  Directory information is transferred into a metadirectory from the source directories at sufficient intervals.

- Two-way synchronisation

  Information flows in both directions, e.g. in order to coordinate mail directories with the address database directory. This may be done on a discontinuous basis, using import and export mechanisms or semi-continuous, using replication mechanisms.

- N-way synchronisation

  Several directories are coordinated with each other using replication or import/export mechanisms. Highly sophisticated replication mechanisms are used for this.

Ways of providing information:

- Reference

  Directory information remains in local directories. The metadirectory only contains a reference to the local directory.

- Accessing

  Data from directories of pre-existing applications are not physically copied into another directory but are accessed in transparently and in real time. The metadirectory forms a virtual directory.

- Import/export

  The information is completely re-entered, discontinuously and at sizeable intervals. Transport is only possible in one direction (distribution by central point).

- Replication

  Two or more directories are coordinated with each other at fixed intervals or when a particular change arises. Only the altered entries or even the altered parts of the entries are transferred. The history of the changes are followed and assessed on both sides so that only the latest change will be recorded as the valid one.

Potential uses

The information in the directory can be used:

- To authenticate users so that they can access various systems with a name and password;
- To personalise web pages so that contents corresponding to a person's interests will be pre-selected;
- It will be possible to recognise which group a user belongs to when they log on.

The possibilities for use and potential benefits are summarised in the following list.

- Data from various applications or information providers are maintained by a directory service so that they are accessible and consistent. This includes, when used internally, all workers and, when used externally, all listed citizens. Only in this way is it possible to avoid an inconsistent and redundant storage of user information. The directory service determines the meaning of a service for all types of root data in addresses both internally and externally. It is no longer necessary to input all the information independently for each user. Rather it can be uniformly provided through this service in a consolidated form.

- By observing international standards such as LDAP or X.500 a controlled communication with all participants is possible. The greatest possible flexibility for adapting to regional characteristics or later migration is ensured. In an overall plan, local central solutions, as well as international solutions and solutions distributed at European level may be realised by applying these international standards. Unified formats and protocols are used for a harmonious exchange and replication of information. The use and provision of information can be controlled by randomly configured views and rights profiles.

- The most varied public archives and further applications can be logically, centrally managed and configured using directory services. This consists of supporting any roles, groups or other organisational forms. In this way a directory service becomes a central repository for individual user-management systems. Such a directory service must be seen as a uniform maintenance component which can considerably reduce efforts for system administration even across national borders.

Different user groups

When dealing with public information, three main user groups must be distinguished.

- Users in one's own organisation

Internal users have high requirements when using their own information system. What is important here is that users are not only able to read information but are also able to write when accessing certain information. This user group has a small number of members compared to those where users are able to read public information. It is also the most complex group. This group of individuals reflects the entire organisational structure. When implementing a directory service for providing access to public information it is often advantageous to consider the users in one's own organisation where planning the data model, but to implement them at a later stage. Efficient mechanisms for user administration have not always been established in organisations. But where the primary goal is making one's information accessible to the public, internal implementation on the basis of the experiences collected may be viewed more as a bonus.

- Users of other known organisations

Users from known organisations can only be regarded as in a relationship of trust to a certain degree. Here, one must distinguish between users who can read and/or write on their own information systems and users who can read, change or coordinate their own directory entries. Since the user group which is entitled to write is usually very small, the requirements in terms of data modelling are negligible here too. Users who are only able to read should basically be treated like the anonymous users described below.

- Anonymous users

Basically, users who are unknown to the organisation only have reading rights to certain information. But this group can also be further divided into groups which have access to different data collections. When an electronic archive is opened to the public, this is the group which must be examined closely. Not many groups need to be distinguished here, as their members must all be provided with uniform attributes. The group of individuals should be prioritised using definite considerations, since for access to the most varied public information only relatively few, well demarcated characteristics need to be distinguished.

It is clear from this first rough division of the main user groups that no great difficulties are to be expected in terms of the requirements for a uniform data model. Such models may be extended on a modular basis, and it is possible to start with the largest user group. In further developments, care must be taken to avoid a considerable overlap between the members in the consideration of other user groups. Careful preparation and stage planning within the framework of a European committee is at least recommended.

Standardisation

To be able to provide blanket coverage of a network of directory services, a universal basic conception must be implemented at three different levels. This basic definition constitutes the starting point for European standardisation.

- Standardisation of the basic data model

As already mentioned above, the group of anonymous users should be considered first. In order to be able to extend the basic data model to other user groups, it must be left sufficiently open so that future requirements of the solution will not impede further operation. In relation to the group of anonymous users, basic access groups must be defined. Thus, a definition must be reached which distinguishes which information in a public archive is accessible to which user groups. This classification should be as uniform as possible for all possible information sources. Local specifications must be defined so that they can be processed by corresponding metadirectory mechanisms.

- Reproducing uniform entitlement structures in the public archives

It should be possible for any relevant archive systems to reproduce and process the components described in terms of the standardisation of the basic data model. The class definition for the information described must correspondingly be reproduced in the document management of archive

systems. This is clearly a precondition for such a scenario. Archive systems must be able to work with a directory, at least with the intelligent mechanisms of metadirectories. As well as being closely connected to access classes, the user groups used in the directory service must be known to the access classes. Only when both these components are known in the directory service and in the archive system, is it possible to provide secure and controlled access to information in public archives.

- Concept of a standardised technological platform

The above two points are primarily concerned with the standardisation of the data structure to be reproduced. As well as this consideration, there is the matter of fixing technical standards. For directory services the main options are LDAP, lightweight directory access protocol and adjoining standards, and the X.500 standard group. X.500 is the most comprehensive standard. It is managed by ITU, International Telecommunication Union. However, its claim of completeness has also revealed that a certain overhead has been defined. This overhead was due to the considerable expenditure for the development of X.500 compatible systems as well as the incalculable costs of operating such systems. Because of frequent political committee work, the extra results required could not always be processed on time. In order to avoid the weaknesses of X.500 a group has been formed with a view to producing a lighter system. LDAP, managed by the internet engineering task force, concentrates on the components needed for a directory of information on the Internet. At its core are the original X.500 definitions. In the very democratic set-up of the IETF — anyone can send in their suggestions — the LDAP has developed far more dynamically and is in many areas comparable to and compatible with X.500. To summarise the foregoing, it is clear that the technical planning of a network of directory services should draw on existing standards. After some basic investigation into which standard should be adopted at its core, the technical details for setting up public archives may be elaborated.

Summary

The use and diffusion of public information through setting up electronic archives will not be possible in the long term without taking into account standardised directory services and metadirectories. At a technical level, comprehensive and widely used standards such as X.500 and LDAP have established themselves. This makes it possible to define an infrastructure which will be of long-term use and a topology, without incurring enormous expense. Many solutions have already proved the usefulness of such services. Through this opportunity to meet individual, national or local requirements, in terms of operation and data models, without having to forego a system which functions well at a European level, it is possible to offer controlled and speedy general access to public information while ensuring authenticity, the observance of high security requirements and the control of copyright.

Utilización de servicios de anuarios normalizados para garantizar el acceso a las bases de datos

Felix von Bredow

Cualquiera que desee comunicar información a distintos círculos de usuarios se encuentra actualmente en la obligación de leer, escribir y gestionar distintas informaciones de usuarios. En este contexto, procede tener en cuenta no sólo el acceso a contenidos definidos y la autenticación de los usuarios, sino también comprobar qué contenidos pueden ser consultados por qué usuarios. Además, deben respetarse algunos criterios de seguridad relativos a la protección de la información interna.

Conviene por tanto considerar en principio distintos grupos de usuarios:

- Usuarios internos

  Los usuarios «a domicilio» tienen unas exigencias en cuanto a la consulta de los sistemas de información internos. Es primordial en este caso garantizar no un mero acceso de lectura, sino también y sobre todo un acceso de escritura a determinada información.

- Usuarios de otras organizaciones conocidas

  A los usuarios de otras organizaciones conocidas se les puede conceder una confianza ciega sólo hasta cierto punto. Es por tanto necesario establecer una distinción entre unas categorías de usuarios, que disponen de acceso de lectura y/o de escritura a los sistemas de información internos, y otras categorías, que pueden leer, modificar o armonizar los datos de un registro interno.

- Usuarios anónimos

  En principio, los usuarios anónimos sólo tienen un derecho de acceso de lectura a determinada información interna. No obstante, esta categoría puede también implicar subcategorías, que se benefician de un derecho de acceso a distintos niveles de información.

Los servicios de anuarios (Directory Services) apenas han suscitado un debate público por ahora, pero se prestan plenamente a este tipo de utilización. Los servicios de anuarios son en realidad sistemas normalizados, que permiten registrar cualquier información en un medio controlado. Estos servicios normalizados surgieron en los operadores internacionales de telecomunicaciones. Esta comunidad se agrupó al principio para crear un anuario telefónico electrónico internacional.

Evidentemente, la prioridad no es dar información directa sobre personas concretas. Sin embargo, la intención de comunicar información pública a distintos grupos de usuarios no es ajena.

Los servicios de anuarios se prestan a distintos métodos de explotación: centralizada, descentralizada y distribuida. Gracias a un modelo de datos normalizado, y teniendo en cuenta todos los aspectos relativos a la información, como, por ejemplo, las especificaciones modelo MoReq, es posible asignar un derecho de acceso controlado y uniforme a las fuentes de información más diversas en lugares diferentes a usuarios previamente registrados.

El registro centralizado de usuarios y categorías de usuarios permite realizar considerables ahorros en los costes de introducción, gestión e integración en los sistemas de información más diversos y, sobre todo, evitar los problemas de seguridad vinculados a inscripciones en el anuario incoherentes u obsoletas.

Los servicios de anuarios son bases de datos especializadas y, en principio, están estructuradas como bancos de datos universales. No obstante, a diferencia de estos últimos, los servicios de anuarios se han optimizado esencialmente para el acceso de lectura, debido a sus aplicaciones relativamente pequeñas y sencillas.

Los servicios de anuarios son bases de datos especiales con funciones limitadas. Dado que cabe renunciar a soportes de transacción complejos, un servicio de anuarios funciona de manera potente, incluso con un gran volumen de inscripciones. Debido a sus propiedades, un servicio de anuarios sirve actualmente para la comunicación controlada de información vía Internet, por ejemplo, pero con un alto nivel de seguridad.

La utilización de los servicios de anuarios es indispensable para establecer un acceso generalizado controlado a los datos de los archivos electrónicos públicos. La exposición presentará los distintos tipos de utilización con ayuda de ejemplos.

Die Verwendung standardisierter Directory Services für einen sicheren Zugang zu Informationsquellen

Felix von Bredow

Jeder, der Informationen unterschiedlichen Benutzerkreisen zur Verfügung stellen möchte, steht heute vor der Aufgabe, verschiedene Benutzerinformationen lesen, schreiben und verwalten zu müssen. In diesem Kontext sind sowohl der Zugang zu definierten Inhalten, die Authentifizierung von Benutzern, als auch die Kontrolle, welcher Benutzer welche Inhalte verwendet hat, zu betrachten. Weiterhin sind gewöhnlich besondere Sicherheitsanforderungen an den Schutz der eigenen Informationen zu berücksichtigen.

Es müssen also prinzipiell unterschiedliche Benutzerkreise betrachtet werden:

- Benutzer der eigenen Organisation

  Benutzer der eigenen Organisation habe hohe Anforderungen an den Umgang mit den eigenen Informationssystemen. Wesentlich ist hier, dass nicht ausschließlich lesende Zugriffe benötigt werden, sondern vorrangig schreibende Zugriffe auf bestimmte Informationen zu berücksichtigen sind.

- Benutzer anderer bekannter Organisationen

  Benutzer bekannter Organisationen können nur in einem gewissen Maß als vertauenswürdig eingestuft werden. Hier ist zwischen Benutzergruppen zu unterscheiden, die lesenden und/oder schreiben Zugriff auf die eigenen Informationssysteme haben und zwischen Benutzern, die die eigenen Verzeichniseinträge lesen, ändern oder abgleichen können.

- Anonyme Benutzer

  Für die eigene Organisation anonyme Benutzer haben prinzipiell nur lesende Rechte auf bestimmte Informationen. Aber auch dieser Kreis kann sich durch weitere Gruppen unterscheiden, die auf unterschiedliche Informationsstände zugreifen dürfen.

Bisher eigentlich wenig in der Öffentlichkeit diskutiert, aber genau für diesem Einsatzzweck verwendbar, sind die so genannten Directory Services. Dabei sind Directory Services im Grunde standardisierte Systeme, um beliebige Informationen in einer kontrollierten Umgebung zu verzeichnen. Der Ursprung für standardisierte Directory Services kann auf die internationalen Telekommunikationsunternehmen zurückverfolgt werden. Dieser Kreis hat sich ursprünglich zusammengesetzt, um eine internationales elektronisches Telefonverzeichnis aufzubauen.

Nun steht natürlich nicht die direkte Auskunft über einzelne Personen im Vordergrund. Dennoch liegt der Gedanke, öffentliche Informationen unterschiedlichen Benutzerkreisen zur Verfügung zu stellen, nicht allzu weit entfernt.

Directory Services können zentral, dezentral oder verteilt betrieben werden. Durch ein standardisiertes Datenmodell unter Einbeziehung aller informationsrelevanter Aspekte, wie z. B. MoReq, ergibt sich die Möglichkeit, einmal verzeichneten Benutzern einheitliche kontrollierte Zugriffe auf die unterschiedlichsten Informationsquellen an verschiedenen Standorten zu gewähren.

Durch eine logisch zentrale Verzeichnung von Benutzern und Benutzerkreisen ergeben sich erhebliche Potentiale zur Kostensenkung für die Erfassung, Verwaltung und die Integration in die unterschiedlichsten Auskunftssysteme und vor allem die Vermeidung von Sicherheitsproblemen durch inkonsistente oder nicht aktuelle Directoryeinträge.

Directory Services sind spezialisierte Datenbanksysteme und daher prinzipiell wie universelle Datenbanken organisiert. Im Unterschied zu diesen sind Directory Services jedoch vorrangig für den lesenden Zugriff aus relativ kleinen und einfachen Anwendungen heraus optimiert worden.

Directory Services sind also Spezialdatenbanken mit limitierter Funktionalität. Weil auf die Unterstützung komplexer Transaktionen verzichtet werden kann, arbeitet ein Directory Service auch bei einer sehr großen Anzahl von Einträgen noch performant. Auf Grund seiner Eigenschaften ist ein Directory Service heute dafür geeignet, mit einem sehr hohem Sicherheitsniveau Informationen z. B. über das Internet kontrolliert zugänglich zu machen.

Der Einsatz von Directory Services ist unerlässlich für die Realisierung eines übergreifenden kontrollierten Zugriffs auf Datenbestände in öffentlichen elektronischen Archiven. Die unterschiedlichen Nutzungstypen werden beispielhaft im Referat aufgezeigt.

Utilisation de services d'annuaires normalisés pour sécuriser l'accès aux bases de données

Felix von Bredow

Quiconque souhaite communiquer des informations à différents cercles d'utilisateurs se trouve actuellement dans l'obligation de lire, d'écrire et de gérer différentes informations d'utilisateurs. Dans ce contexte, il y a lieu de prendre en compte non seulement l'accès à des contenus définis et l'authentification des utilisateurs, mais aussi de vérifier quels contenus peuvent être consultés par quels utilisateurs. En outre, certains critères de sécurité relatifs à la protection des informations internes doivent ordinairement être respectés.

Il convient donc de considérer par principe différents groupes d'utilisateurs:

- Utilisateurs internes

Les utilisateurs «maison» ont des exigences quant à la consultation des systèmes d'information internes. Il est primordial dans ce cas d'assurer non pas un simple accès en lecture, mais aussi et surtout un accès en écriture à certaines informations.

- Utilisateurs d'autres organisations connues

On ne peut accorder une confiance aveugle aux utilisateurs d'autres organisations connues que dans une certaine mesure. Il est donc nécessaire de faire une distinction entre les catégories d'utilisateurs qui disposent d'un accès en lecture et/ou en écriture aux systèmes d'information internes et celles qui peuvent lire, modifier ou harmoniser les entrées d'un registre interne.

- Utilisateurs anonymes

En principe, les utilisateurs anonymes ne disposent que d'un droit d'accès en lecture à certaines informations internes. Toutefois, cette catégorie peut également comporter des sous-catégories, qui bénéficient d'un droit d'accès à différents niveaux d'information.

Les services d'annuaires *(directory services)* n'ont guère suscité de débat public pour l'instant, mais se prêtent tout à fait à ce type d'utilisation. Les services d'annuaires sont en fait des systèmes normalisés qui permettent d'enregistrer n'importe quelle information dans un environnement contrôlé. Ces services normalisés ont vu le jour chez les opérateurs internationaux de télécommunications. Cette communauté s'était regroupée à l'origine pour constituer un annuaire téléphonique électronique international.

Bien entendu, la priorité n'est pas de donner des renseignements directs sur des personnes en particulier. Néanmoins, l'intention de communiquer des informations publiques à différents groupes d'utilisateurs n'est jamais très loin.

Les services d'annuaires se prêtent à différents modes d'exploitation: centralisée, décentralisée et distribuée. Grâce à un modèle de données normalisé et en prenant en compte tous les aspects relatifs à l'information, comme par exemple les spécifications types MoReq, il est possible d'attribuer un droit d'accès contrôlé et uniforme aux sources d'information les plus diverses en des lieux différents à des utilisateurs préenregistrés.

L'enregistrement centralisé des utilisateurs et des catégories d'utilisateurs permet des économies considérables sur les coûts de saisie, de gestion et d'intégration dans les systèmes d'information les plus divers et, surtout, d'éviter les problèmes de sécurité liés à des inscriptions dans l'annuaire incohérentes ou périmées.

Les services d'annuaires sont des bases de données spécialisées et sont donc en principe structurées comme des banques de données universelles. Toutefois, à la différence de ces dernières, les services d'annuaires ont été optimisés essentiellement pour l'accès en lecture, de par leurs applications relativement petites et simples.

Les services d'annuaires sont des bases de données spéciales aux fonctionnalités limitées. Comme il est possible de renoncer à des supports de transaction complexes, un service d'annuaires fonctionne encore de façon performante, même avec un gros volume d'inscriptions. Du fait de ses propriétés, un service d'annuaires se prête actuellement à la communication contrôlée d'informations via l'internet, par exemple, moyennant un haut niveau de sécurité.

L'utilisation des services d'annuaires est indispensable pour mettre en œuvre un accès généralisé contrôlé aux stocks de données d'archives électroniques publiques. L'exposé présentera les différents types d'utilisation à l'aide d'exemples.

Parallel session 3

Tuesday, 7 May 2002

CONTENT MANAGEMENT AND DELIVERY
Establishing and maintaining archival systems, best practices and easy-to-use-access

Chairperson:	Dan Tørning (Denmark)
Co-chair:	Ramon Alberch i Figueras (Spain)
Rapporteur:	George Mackenzie (United Kingdom)

Celsa Poza Serrano

Celsa Poza began her professional career in the area of information systems in 1985 in an international consultancy company working in a team for system introduction projects, principally in the sectors of industry and banking, for a period of three years.
Within the same company she went on to manage large projects, taking part in the project for the Olympic Games of Barcelona. After this, she participated in international projects in France (Games of the Mediterranean 1993), the United States (Olympic Games of Atlanta) and Argentina (directing a project to set up the commercial Argentine Water system).
After working on international projects, Celsa Poza moved on to direct the project to adapt and extend (technological update) the geographic information system on water in Barcelona. In 1998 she joined the computer centre of the Autonomous Government of Catalonia, currently Tsystems, to direct the organisation and installation of the control centre of the autonomous police.
In 2000 she joined the management team of the Agencia Catalana del Agua (Catalan Water Agency) where she is the director of information systems.

Paperless public administration

Celsa Poza

What is the Catalan Water Agency?

In Catalonia, as in all the Mediterranean countries, water is a scarce and irregularly available resource. The relief, climate and rainfall, together with the needs of residents, tourism, farming and industry, make it necessary to manage water as a strategic resource.

Awareness of the unsustainability of certain models of development and their impact on the quality of water resources led to the drafting of new legislation on water in Catalonia. As a result, the current rules have been placed in a more modern context, the new requirements relating to sustainable development have come into play and greater emphasis is now placed on environmental considerations.

The new water regulation (Law No 25/1998 of 31 December and Presidential Decree No 125/1999 of 4 May) has two main features:

- it provides a basis for an integrated approach to water management (water supply and wastewater treatment);
- a new water management entity, the Agencia Catalana del Agua (Catalan Water Agency), a public body operating under the auspices of the Catalan Department of the Environment, has been set up.

All the water management powers of the Autonomous Community of Catalonia fall within the remit of the Catalan Water Agency, set up in January 2000, which thus covers all aspects of the water cycle.

What is paperless administration?

Since its inception, the Agency has sought to be a modern, efficient and flexible public body which is open to the public and makes full use of information technology in providing a high-quality service to the public, while at the same time rationalising and optimising in-house procedures.

To meet this ambitious objective, a systems management plan was drawn up in 2000 which sets out the Agency's new needs and the strategy to be followed in order to meet them over the next two years.

Projects to be carried out under the plan are grouped together under the following initiatives:

- paperless administration: this is a measure to modernise, standardise and optimise the Agency's procedures;
- Agency on the net: the main goal here is to open the Agency up to the general public by providing a variety of means of access;
- environmental know-how: this encompasses all projects dealing with the management of environmental matters with implications for water.

In early 2001 DMR Consulting was commissioned through open competition to set up the first two initiatives, 'paperless administration' and the 'Agency on the net'.

In the proposal selected, the suggested approach to paperless administration includes the use of FileNet as a tool to manage all the Agency's documents.

What objectives can FileNet achieve?

Whilst the main purpose of a document manager is to cut down on the use of hard copies and promote the use of electronic documents, FileNet facilitates integration with the other systems being put in place, and this will enable us to:

- set up a central electronic archive, which will avoid papers going from one desk to another within the Agency, thus reducing the risk of losing or mislaying documents;
- register all incoming and outgoing documents. By registering incoming and outgoing documents and scanning, we can ensure that all the Agency's documentation is registered in a single system, which will make it easier to search for information;
- ensure that all documents created within the Agency are always linked with the file which generates them. FileNet is a single management system for files at all stages, with documents becoming one more part of the information in a file;
- unify the management of different types of document by using a single IT processing and storage system. This renders work processes independent of the source of incoming or outgoing documents. The user always works in the same way, whatever the channel by which the document enters (a request on paper or in electronic format, an Internet transaction, etc);
- create, input, store, edit, print and process as text documents and files from different sources and in different formats. This will enable us to be more flexible when it comes to accepting different working formats and will also increase possibilities for integrating all the Agency's information;
- reduce and make the best possible use of the time spent searching for information. The fact that all the Agency's sources of information — both incoming documents and those produced in-house — are stored in the same system will enable more secure searches to be carried out. Moreover, it will be possible for individuals without expertise in the procedure to carry out searches.

Introducing a document manager will also enable us to:

- check the archives and access documents from any of the Agency's workstations;
- establish different levels of security and confidentiality for documents, introducing functions to block documents according to the procedure to which they are related and the user concerned;
- control the different versions and track documents throughout their life cycle. It will be possible at all times to ascertain the version being handled and the document's history and to add to all ongoing versions notes that can be consulted by users with access to the document;
- structure the information and distribute documents to multiple users. Each item of information will be automatically incorporated in a document that is stored electronically and accessible to all authorised employees;
- manage templates and forms in different formats, facilitating the transition to a single format for all the templates and forms used by the Agency.

For what work processes is FileNet used?

The document manager is used in all the Agency's work processes, from the outset and in all the stages involving document creation or modification.

In the Agency's register of incoming and outgoing documents, each document is recorded in the system and given an initial classification. It is then passed on to the team which scans and files it. The following tasks are carried out at the same time as document scanning:

- registration of the electronic document in the document manager and entry in the incoming register;

- input of the information contained in the document into the database (only where the document is a request form). This is done using an intelligent OCR;
- assignment of a registration number to the electronic document and filing of the paper document in the appropriate place in the archive.

From this moment on, any authorised user can consult the documents stored and continue with the corresponding tasks.

In the case of requests, the procedure for processing the application is started.

Each document generated in handling files is automatically registered in the document manager.

The document manager is also used for internal purposes, making it possible to keep track of documentation generated in-house in relation to the tasks with which it is associated.

In addition to performing the above functions, using FileNet has helped us to:

- redesign and standardise existing templates and application forms. New standard templates and forms have been drawn up, introducing criteria for classifying and identifying types of document;
- assign a registration number to each document in accordance with a single standard system for the entire Agency;
- monitor the documentation being used in different work processes, both inside and outside the Agency.

Implementation

The approach taken towards implementation has been to start using the document manager gradually as and when new work processes, and therefore new IT applications, are introduced in the Agency.

The plan was initially to replace paper as the new systems were implemented, but this has proved impossible for legal reasons. Electronic storage is not recognised by law, and we have therefore had to continue to duplicate storage of both the hard copy and electronic versions of documents.

The Autonomous Community has set up a committee which is currently examining the validity of electronic document storage and is drafting standards for the recognition of electronic archives. We hope that it will settle the issue, enabling us to implement the system completely and replace paper with electronic storage.

Benefits obtained

Despite the legal difficulties we have encountered, we have already reaped a number of benefits, for example:

- creation of a single centralised electronic archive, making for easier consultation and greater reliability of the documentation handled by the Agency;
- integration of the documentation generated by a file with the processing of that file, enabling all the information on a file to be obtained from the same system, namely the one in which the file is handled;
- possibility of consulting documentation without knowledge of the file or procedure which has generated it. Any authorised person can make searches in the system using different criteria, without expertise in the procedure;
- reduction of time spent searching for documentation. Since there is a single system in which searches can be made using different criteria, including keywords, less time is spent looking for information.

La Administración Pública sin Papeles

Celsa Poza

Uno de los objetivos de la Agencia Catalana del Agua desde su creación es ser una administración pública moderna, eficaz, ágil, abierta al ciudadano, con una decidida apuesta por las tecnologías de la información como instrumento para dar un servicio de alta calidad a los ciudadanos y al mismo tiempo racionalizar y optimizar los procesos internos. Para abordar tan ambicioso objetivo, durante el año 2000 se llevó a cabo la definición de un plan director de sistemas, donde se identifican las nuevas necesidades de la Agencia y la estrategia que se debe seguir para darles cobertura en un plazo de dos años.

Esta definición ha permitido reagrupar los proyectos, en especial la Administración sin Papeles, cuya iniciativa es la de modernizar, normalizar y optimizar los procedimientos de la Agencia.

La solución para abordar la iniciativa de la Administración sin Papeles, incluye la utilización de FileNet como herramienta para la gestión documental de toda la Agencia. La implantación de un gestor documental tiene como objetivo principal la reducción de la documentación en papel y la utilización de documentos en formato electrónico. Con FileNet conseguimos una integración total con el resto de los sistemas.

El método seguido para su implantación ha sido su paulatina introducción a medida que se ponían en marcha los nuevos procesos, y las nuevas aplicaciones, de la Agencia.

Inicialmente se pretendía sustituir el papel a medida que se iban implantando los nuevos sistemas, pero esto no ha sido posible. Legalmente no está reconocido el archivo electrónico. Por ello, nos hemos visto obligados a mantener el archivo por duplicado, en papel y electrónico.

Esperamos que esta cuestión se solucione y que se pueda realizar una implantación total del sistema con el fin de sustituir totalmente el papel por los documentos electrónicos. A pesar de las dificultades legales que nos hemos encontrado, ya hemos obtenido una serie de resultados, tales como:

- La creación de un archivo electrónico único y centralizado que facilite la consulta y fiabilidad de la documentación de la Agencia.

- La integración de la documentación generada en un expediente con su tramitación. Ello posibilita la obtención de toda la información sobre un expediente desde un mismo sistema, el de tramitación.

- La posibilidad de consultar la documentación sin conocimientos previos de la tramitación o del expediente asociado. Cualquier persona autorizada puede realizar consultas al sistema, por diferentes criterios y sin ser un especialista en la tramitación.

- Disminución del tiempo de búsqueda de documentación. Al ser un único sistema y permitir la consulta por diferentes criterios, incluso palabras, reduce el tiempo de búsqueda de la información.

Die papierlose öffentliche Verwaltung

Celsa Poza

Seit seiner Gründung gehört zu den Zielen der **Agencia Catalana del Agua** (das Katalonische Wasseramt), eine moderne, effiziente, dynamische und bürgernahe Behörde zu sein, die entschlossen auf die Informationstechnologie als Instrument für einen hochwertigen Dienst am Bürger setzt und gleichzeitig die internen Arbeitsabläufe rationalisiert und optimiert. Um dieses ehrgeizige Ziel zu erreichen, wurde im Jahr 2000 ein Systemplan entwickelt, in dem der neue Bedarf des Wasseramtes und die Strategie zur Deckung dieses Bedarfs für den Zeitraum der nächsten zwei Jahre festgelegt wurden.

Auf dieser Grundlage wurden die durchzuführenden Maßnahmen konzipiert, insbesondere die papierlose Verwaltung: Ziel ist die Modernisierung, Standardisierung und Optimierung der internen Verfahren des Amtes.

Der ausgewählte Vorschlag sieht als Lösung für die **papierlose Verwaltung** die Verwendung des Dokumentenverwaltungssystems FileNet vor. FileNet dient nicht nur der Reduzierung des Papierverbrauchs und der Verwendung von Dokumenten in elektronischem Format, sondern der vollständigen Verknüpfung mit den übrigen schon vorhandenen Systemen.

Die Einführung des Systems erfolgt stufenweise; die Dokumentenverwaltung wird entsprechend der Umsetzung der neuen Arbeitsabläufe und Einführung der neuen Anwendungen in Betrieb genommen.

Ursprünglich war geplant, die Papierunterlagen schrittweise entsprechend der Einführung des neuen Systems zu ersetzen; dies hat sich jedoch im juristischen Bereich als undurchführbar erwiesen. Elektronische Archive sich rechtlich nicht anerkannt. Daher war es unumgänglich, zwei Archive, ein Papier- und ein elektronisches Archiv, beizubehalten.

Es steht zu hoffen, dass die Angelegenheit eine Lösung findet und die vollständige Einführung des Systems möglich wird, damit das Papierarchiv ersetzt werden kann. Trotz der aufgetretenen rechtlichen Probleme sind bereits verschiedene Vorteile festzustellen:

- Einführung eines einheitlichen zentralen elektronischen Archivs, das das zuverlässige Auffinden von Arbeitsunterlagen des Amtes erleichtert.
- Verknüpfung der zu einem Dossier gehörigen Dokumente mit der Bearbeitung des Dossiers, so dass über dasselbe System auf sämtliche Informationen einer Akte zugegriffen werden kann.
- Möglichkeit des Zugriffs auf Dokumente durch hierzu berechtigte Personen, ohne dass diese das Dossier oder den jeweiligen Bearbeitungsvorgang kennen müssten.
- Zeitersparnis bei der Suche nach Dokumenten durch die Verwendung eines einheitlichen Systems.

L'administration publique sans papier

Celas Poza

Un des objectifs que s'est fixé l'**Agencia Catalana del Agua** (Agence), depuis sa création, est d'être une administration publique moderne, efficace, souple, ouverte au citoyen, résolument tournée vers les technologies de l'information qui permettent d'offrir au citoyen un service de haute qualité tout en rationalisant et en optimalisant les processus internes. Pour se donner les moyens d'un objectif si ambitieux, l'Agence a défini, au cours de l'année 2000, un plan directeur de système visant à identifier ses nouveaux besoins et la stratégie à adopter pour répondre à ceux-ci dans un délai de deux ans.

Cette définition a permis de regrouper les projets à implanter, dont notamment l'administration sans papier, initiative destinée à moderniser, normaliser et optimaliser les processus de l'Agence.

La solution pour aborder l'initiative de l'**administration sans papier** comprend l'utilisation de FileNet comme outil de gestion documentaire de l'ensemble de l'Agence. Si l'implantation d'un gestionnaire de documents vise principalement à réduire la documentation sur papier et à permettre l'utilisation de documents en format électronique, FileNet, lui, assure une intégration totale avec les autres systèmes implantés.

La méthode choisie pour l'implantation a été de procéder progressivement; le gestionnaire de documents est utilisé au fur et à mesure de la mise en place des nouveaux processus et, partant, des nouvelles applications de l'Agence.

À l'origine, il s'agissait de remplacer le papier à mesure que les nouveaux systèmes étaient installés, mais cela n'a pas été possible pour des motifs juridiques. Légalement, l'archivage électronique n'est pas reconnu. C'est pourquoi il a fallu conserver une double archive, l'une sur papier et l'autre électronique.

Il faut espérer que cette question sera résolue, ce qui permettra de réaliser une implantation totale du système de manière à remplacer entièrement le papier par l'électronique. Malgré les difficultés juridiques rencontrées, quelques résultats ont déjà été obtenus, à savoir:

- La création d'une archive électronique unique et centralisée, qui facilite la consultation des documents et qui en assure la fiabilité.
- L'intégration de la documentation générée dans un dossier et le traitement de ce dernier. Cela permet d'obtenir toutes les informations d'un dossier depuis un seul et même système, celui qui est utilisé pour l'instruction.
- La possibilité pour toute personne autorisée de consulter des documents sans connaître la procédure ou le dossier y afférents.
- La réduction du temps de recherche des documents, étant donné qu'il s'agit d'un système unique.

Philip Lord

Philip Lord studied mathematics at the Universities of Reading and London, afterwards joining the UK Medical Research Council's Scientific Staff to research the medical effects of air pollution. He later took a post as Head of Computing at Excerpta Medica, within Elsevier Science Publishers in the Netherlands. Here he introduced new technologies for the production of scientific electronic publications. He moved on within Elsevier to a position as internal consultant to the directors on new publishing technologies and their promotion, and in this role he served as Vice-Chairman of the ANSI standardisation committee for CD-ROM. He joined SmithKline Beecham Pharmaceutical's R & D division in 1991 to introduce document management systems and publishing solutions for regulatory documentation, and was responsible for the introduction of SGML/XML standards for documentation in the company. Over the past six years he has been engaged full-time in scientific and technical data archiving issues and has led an international project in GlaxoSmithKline plc to introduce an electronic archive into their R & D division.
In 2002 Philip Lord left GlaxoSmithKline in order to establish his own consultancy business, providing advice to businesses and other organisations on the business, technical and organisational aspects of archiving and preserving their digital assets.

Preserving scientific and technical records — strategies and solutions

Philip Lord

Introduction

There are few true digital archives in the private, commercial sector. Where they do exist they tend to be in the science-based sectors, such as pharmaceuticals. Paradoxically, these sectors have the most difficult issues to confront in digital archiving. Why are there so few solutions, why are the problems difficult?

Electronic archiving is an area where practical progress in the public sector outstrips that in private industry; development work in academic institutions is also in advance of privately funded research. In contrast to the publicly funded area (i.e. government funded), progress is poor in the private sector. This is no accident. Though private industry has the same need to preserve electronic records (but under different business drivers) a number of factors have inhibited progress.

In my view the lack of private industry progress is an important issue. Relatively poor progress has led to reduced interest and investment by potential suppliers of systems and services. This narrowing of the market for products results in fewer products available on the market, which in turn increases their cost and leads to an impoverishment of choice and reduced functionality. Secondly, by failing to come to grips with the problem we may expect inefficiencies, at the least, in the performance of the private sector, which now relies so much on digital information.

A bright spot is the science-based industries, and particularly pharmaceuticals and related sectors, where they are driven by regulatory necessity. Consequently, this sector has some solid business drivers on the one hand, but a particularly difficult set of requirements to meet on the other.

This paper reviews, briefly, the business drivers in these industries and then examines some of the data characteristics and solutions which are available — and draws attention to areas where there is more work to be done.

Business drivers — and the FDA's CFR21 Part 11 regulation

The drivers to establish archives, whether digital or otherwise, generally fall into the following two broad categories: either the retention of information is mandated, or it is a voluntary act driven by hard-headed business reasons or sentiment.

In the category of enforced retention the drivers are:

1. Regulatory requirements from government agencies which define rules that say records must be retained. Failure to comply results in punitive action. The most significant example here is the requirements which flow from the United State's Food and Drug Administration's 21 CFR Part 11 regulation. (See: http://www.fda.gov/ora/compliance_ref/part11/).
2. Legal requirement, to comply with a law (for example, to keep records related to taxation issues).

In industry, in general, action will only result from these drivers, but sometimes voluntary action is also seen, as in the following circumstances:

3. Contractual: retention may be a condition of entering into a contractual relationship with a third party. The choice of entering the contract may be voluntary, but once entered into, an obligation arises with penalties for default. An example is the retention on design information on a major civil engineering project.

4. Intellectual property protection: companies may wish to preserve records which encapsulate intellectual property, such as supporting evidence behind patents.
5. Legal protection: to be able to establish evidence in the future to support offensive or defensive litigation; an example in this category might be the defence of a product liability suit. (There is tension here between having data to support one's own position, and having information retained and therefore also liable to legal discovery activities by opposing litigants.)
6. Preservation: data which may have value in the future through their reuse, or as future reference materials, or have value for data mining and prospecting.

What is rarely seen to any degree is retention merely to preserve the historical or cultural record, or for reasons of sentiment (as you might see in a personal photograph archive).

It will be seen that these criteria are focused on the mitigation of risk (which, rephrased, may be characterised as the minimisation of unwanted costs) or the maximisation of value from existing assets. What is absent in general is the motivation to retain the cultural record of the company and its history. The establishment of archives, and digital archives in particular, is viewed as a cost rather than an aid to revenues and profits, and this attitude can reinforce inertia.

The regulation 21CFR Part 11 referred to above is particularly important, since eventually its impact may extend well beyond the pharmaceuticals sector. This regulation became law in the United States in August 1997, and came from the US Federal Government's Food and Drug Administration (FDA). The regulation is titled 'Electronic records; electronic signatures' and it is directed at all companies who sell medicinal products in the United States (or contribute to companies who sell these products). Many of these are of course European, and they too need to comply with the regulation to operate in that market. It is worth noting that the Internet is drawn very wide, and includes the medical devices and cosmetics industries as well as those making pharmaceutical drug products. As its name implies this regulation sets out the good records management practices with which computer systems must comply and under which they should be managed, and specifies conditions for acceptable electronic signatures as equivalent to ink-on-paper signatures. These regulations are applicable to all systems which generate records used in the development and manufacturing of products falling under the FDA's purview. Without going into the details of the regulation (there is now a whole industry devoted to interpretation and advice on compliance with this), what is of interest here is the definition of a record and the requirements for electronic-records retention.

Electronic records are defined in the regulation as:

> *any combination of text, graphics, data, audio, pictorial information or any other information representation in digital form, that is created, modified, maintained, archived, retrieved or distributed by a computer system.*

This is a very wide definition — and note that it includes reference to processes which information is subject to, including archiving.

The rule states that where a record is created electronically it must be maintained electronically throughout its whole life cycle, which will include its long-term retention. There is no option to transfer a record to print or microfilm for this purpose. This point is of the greatest significance, forcing attention on long-term preservation questions.

It should be noted that various 'predicate' rules from the regulatory authorities also mandate the archiving of these data, whether electronic or paper-based. Thus the good laboratory practice (GLP) regulations mandate that all records falling under their umbrella must be archived, and they go on to define conditions for a satisfactory archive. In summary these are:

- they require the responsibility of an assigned, identified individual;
- only authorised personnel can enter the archive;
- materials must be indexed to expedite retrieval;

- there must be logging of materials removed and returned;
- they must provide appropriate storage conditions to minimise deterioration.

All these apply to electronic archives as well as archives for paper and physical objects, though it is for the latter they were first formulated.

Taken together, 21CFR Part 11 and the GLP regulation (plus similar regulations covering clinical practice and manufacturing practice), mandate that the industries affected must create electronic archives to remain compliant. Failure to do so can lead to big fines and/or exclusion of the relevant products from the American market. Industry is slowly coming to terms with this regulation; it is almost certainly true that no one is yet fully compliant with it, and meanwhile the FDA is increasing its vigilance, as the initial period of tolerance to allow time for industry to adjust passes. It is now recruiting and training many new inspectors. The implications for data management and IT practices and consequences for costs are substantial.

For the purposes of this paper, the important points are that:

- a significant business sector is mandated to create electronic archives (in Europe, and elsewhere, as well as in the United States) if they wish to exploit the American market;
- the regulation is likely to be used as a model for other US Government regulatory agencies, such as the Environmental Protection Agency (EPA), thus drawing in an even wider circle of the private sector; and
- it may be a model from which government agencies outside the United States may draw ideas.

Technical characteristics in science-based industries

The foregoing sets out the regulatory context which is forcing one significant part of private industry to set up electronic archives. The following section describes the technical problems which pharmaceuticals, and other scientific and technologically-based industries have when addressing the establishment of electronic archives. A later section will discuss some organisational obstacles.

The technical and scientific information produced by the manufacturing and R & D organisations in industries such as pharmaceuticals, chemicals, aircraft, and manufacturing is very different from the office and business data that these companies also produce. The challenges these data present fall into a number of areas. For instance:

1. Heterogeneity of data types, applications and systems.
2. Complex data-record structures.
3. High-data volumes (and more generally heterogeneity of data volumes).
4. System configuration dependencies in the data.
5. Geographical dispersion.

Let us look a little more closely at these:

1. When one looks at the systems inventory of a large R & D organisation it would not be unusual to find the number of different systems, generating data to be numbered in thousands. This applies to hardware, operating systems and applications: there will be many architectures of computer in use, and they may be running under a variety of operating systems (and different versions of these). Similarly, there will be many data creation software applications present, and in a number of different versions.

   To take pharmaceuticals as an example again, there will be very many systems, often highly specialised and often attached to (or part of) an instrument of some kind, like a gas chromatograph or an infrared spectrometer. There will be a great variety of these from different manufacturers. In the absence of any agreed format standards there are likely to be as many

file formats as there are suppliers. These suppliers will often be quite small companies, subject to takeover or short life spans.

When devising archiving solutions for this market no assumptions can be made about systems architectures, and they must capture as much 'systems metadata' as possible to give a key for future accessibility.

2. The data itself raises issues. In many ways it demonstrates, in an exaggerated form, some of the characteristics of 'difficult' office data. Again, the key problem is one of heterogeneity. There is great diversity in data types, data structures and record structures. Very often the data is quite like multimedia — the content can conform to a variety of formats and forms, in much the same way as a consumer CD-ROM product may contain interlinked video, sound and text.

   The collection of these data into meaningful records may be complex. A meaningful record may be built around a single file at one extreme (possibly with a complex but proprietary internal format), and at the other may consist of many thousands of files, possibly interdependent and organised in a way (such as in a specific file hierarchy) which makes sense to the originating software, and to the software which presents the results. It has also been noted that it is not sufficient just to keep final printed human readable forms of the information. Sometimes this is simply not possible because printing loses or obscures information (e.g. a three-dimensional image).

   Solutions here are problematic. Neutral, well-defined standards should be used if possible. XML is a good candidate, but finding a suitable standard is often simply not possible. It is also necessary to collect as much information about the data — metadata — as possible. This represents a significant barrier. The technological obsolescence problem is essentially one without a satisfactory solution at the moment (the choices are repeated migrations to new systems, emulation, computer museums and adopting long-term neutral standards). All of these have their drawbacks and uncertainties. Leaving aside neutral long-term formats they all represent significant long-term cost commitments. This is one of the major impediments to progress. In sectors such as pharmaceuticals where action is demanded by law it presents managements with an unwelcome added cost of doing business.

3. The heterogeneity of file formats is matched by the variety in file sizes. Files may vary from a few kilobytes (perhaps from some simple weighing machine or word processor) to many gigabytes (for high-resolution, dynamic imaging applications).

   Archiving systems should make no assumptions about record sizes.

4. As noted above, in this environment, information often comes from some instrument or machine. This can mean that the interpretation of the data and the way results are presented to the user will have hardware dependencies. Thus, for example, the configuration of machine switch settings, environmental variables and calibrations will need to be preserved (if possible) to create reproducible results. Any archiving system must in some way be able to capture and retain these dependencies.

   This is a demanding requirement, and it is fair to say there does not exist any satisfactory solution yet.

5. The R & D organisations in large companies are geographically dispersed (though they may be managed as one unit). In these environments records may be assembled from sources which are located over different time zones and possibly in different countries.

   Lastly, the computing infrastructures which support the R & D and manufacturing activities may be distributed. That is, the data records, rather than being found on one machine, may be spread over different systems (possibly separated geographically). While this may not be significant so long as reliable networking infrastructures are in place it raises issues on authenticity about such things as the 'correct' time zone to employ and possible issues of where legal responsibilities are to be placed.

On the positive side there are some requirements that would not be expected in this environment — thankfully. These are requirements such as demanding retrieval times (e.g. there is no demand for real-time access, as there might be in an online video archive). Another example is public access with charging.

Organisational characteristics and issues

Lastly I will look at some management and structural issues to which this sector is subject. These include:

1. Lack of suitable systems and services.
2. Management problems.
3. Structural problems.
4. Costs.

A few comments on these:

1. Lack of systems and services. An immediate problem is the lack of suitable systems and services. There are few (any?) systems which are available and which can cope with the demands of gathering, indexing, storing and preserving records from the environments typical in R & D organisations and in manufacturing. In pharmaceuticals, for example, there are a few very new systems of a specialised character, which tackle some of the problems noted above — but they are not commercially available. Alongside these there are some specialised suppliers of laboratory software who can put information in well-defined locations and, since they know the internal structure of the files they collect, they can extract metadata. Some also attempt conversion to standard, neutral formats (some of which are proprietary).

2. Management issues: To give a few of these affecting the establishment of digital archives:
   - In general there are fewer records managers in private companies; where they are employed, they generally do not have the IT skills which enable them to ask the right questions of their IT groups.
   - Lack of expertise and understanding: The level of understanding among decision makers in commercial environments is generally poor. The problem lies between well-known existing categories: often it may be seen as merely an IT problem (which it isn't).
   - Decentralisation of the management of departments and of responsibilities can create a barrier, since the problem is usually beyond local managements' reach or ability to influence the corporate centre, though they understand the issues.
   - Geographical spread may also be a factor, independently of the decentralisation or otherwise of management responsibility. When the spread crosses legislative boundaries, the company may encounter complications accommodating competing laws and regulations, and which may militate against introducing uniform solutions.

3. Structural issues: At a practical level there are organisational, structural difficulties to add alongside the more conceptual issues raised above. For example, in general it may not be clear how long-term responsibility is to be bequeathed within the organisation. This operates at two levels: who 'owns' specific records as personnel change (and who knows how to interpret the information as personnel changes, and memory fades); secondly how is responsibility to be bequeathed for the organisational structures which will manage the archive?

4. Costs: As noted above, retention of electronic records is seen as a cost which provides no assured or tangible return. The costs can be substantial. Archiving paper (and similar) records requires investment in storage space and people, utilities (and possibly indexing software) and consumables to service and run it. Electronic information imposes additional costs since it requires more maintenance of the stored artefacts themselves, with more sophisticated (and larger) computer systems needed to manage the archived information.

Conclusions

In the private sector, progress with electronic archiving is slow. However, because of regulations originating in the United States, the healthcare industries in particular are beginning to consider the

problem. While this is welcome, it is clear that these industries face a difficult set of technical, managerial and structural obstacles, imposed by the nature of their business and also by the fact that they are businesses and not public-sector entities, which are assured of a long-term framework.

Conservación de documentos científicos y técnicos: estrategias y soluciones

Philip Lord

Este documento presenta y analiza las estrategias y soluciones que las organizaciones que generan documentos científicos y técnicos (así como documentos internos «normales») pueden adoptar para su gestión y archivo a largo plazo. Los sectores empresariales que se ven particularmente afectados por este problema son el farmacéutico, el de la salud, el sector químico y la fabricación, que son económicamente vitales en Europa. En estos sectores, la protección de los activos digitales puede ser primordial para el mantenimiento de su eficacia comercial e incluso puede afectar a su viabilidad: muy a menudo los activos digitales son el núcleo de su éxito y su rentabilidad continua. Otros sectores y organizaciones comerciales también dependen cada vez más de la buena gestión a largo plazo de activos digitales de todas clases, y pueden crear registros semejantes a los procedentes de los sectores científico y técnico, particularmente cuando se utiliza el multimedia.

Los documentos técnicos y científicos tienen características que presentan dificultades especiales. Muy a menudo deben gestionarse en un marco regulador estricto, que exige a su vez requisitos especiales. Entre los problemas que se abordan figuran la falta de sistemas adecuados disponibles comercialmente para apoyar la conservación a largo plazo de este tipo de información, las necesidades especiales de los metadatos que presentan los archivos científicos y técnicos (y las dificultades para recabar estos metadatos) y la muy diversa naturaleza de fuentes de información en un medio en el que hay muchos centenares de vendedores, cada uno con sus propios formatos y estructuras de ficheros. En general, los datos pueden compararse con el multimedia en cuanto a la forma del contenido, y los volúmenes de información generados para un solo registro pueden ser enormes (posiblemente del orden de gigabites). Las empresas y las organizaciones que crean estos registros suelen operar internacionalmente, y esto añade otra capa de complejidad técnica y normativa a la búsqueda de una solución. En este análisis se incluirá el examen de las opciones para adquirir sistemas de archivo y para superar el problema de la obsolescencia de los datos en un medio con una gran rotación de las tecnologías informáticas.

Junto con los problemas técnicos que deben solucionarse, se encuentran los problemas comerciales de la adquisición de un sistema y de los procesos necesarios para hacerlo funcionar y mantenerlo de forma rentable, y es quizás este aspecto el que a largo plazo será el más crítico para las empresas. Desde el punto de vista de la alta dirección, éste será uno de los aspectos más importantes que deben tenerse en cuenta al decidir sobre las estrategias y tácticas que se adopten. Estos aspectos también se discutirán en las opciones de estrategia que se presentan. El presente documento contará con la aportación de la experiencia en materia de diseño, envío y funcionamiento de un archivo de datos electrónicos de una gran empresa multinacional de productos farmacéuticos establecida en el Reino Unido, y discutirá todos los tipos de datos, incluidos los generados en el ámbito de la investigación clínica y científica.

Gran parte del revolucionario trabajo en materia de archivo digital se ha realizado en el sector público, pero la cuestión reviste la misma importancia para las empresas del sector privado, que deberán extender las mejores prácticas a su medio. En resumen, en la esfera de los documentos electrónicos técnicos y científicos, las mejores prácticas y soluciones que son objeto de esta reunión apenas están surgiendo, y se espera que el presente documento contribuya a su establecimiento tanto para el sector público como para el sector privado.

Die Archivierung wissenschaftlicher und technischer Aufzeichnungen – Strategien und Lösungen

Philip Lord

In diesem Beitrag werden die Strategien und Lösungen vorgestellt und analysiert, die Organisationen, die wissenschaftliche und technische Aufzeichnungen (sowie „normale" Geschäftsunterlagen) erzeugen, für deren Langzeitverwaltung und -archivierung anwenden können. Von diesem Problem besonders betroffene Wirtschaftszweige sind die Pharmabranche, das Gesundheitswesen, die Chemieindustrie und das verarbeitende Gewerbe; sie alle spielen in der europäischen Wirtschaft eine maßgebende Rolle. In diesen Branchen kann der Schutz digitaler Wirtschaftsgüter von größter Bedeutung für die Sicherung ihrer wirtschaftlichen Effektivität sein – und unter Umständen sogar ihre Lebensfähigkeit betreffen: Sehr häufig bilden digitale Güter die Grundlage ihres Erfolgs und ihrer dauerhaften Rentabilität. Auch andere Wirtschaftssektoren und -organisationen sind in immer stärkerem Maße abhängig von der guten Langzeitverwaltung digitaler Güter aller Art, und dies kann zur Entstehung von Aufzeichnungen führen, die denen aus Wissenschaft und Technik gleichen, vor allem wenn Multimedia zum Einsatz kommt.

Technische und wissenschaftliche Aufzeichnungen weisen Merkmale auf, die ihre Handhabung besonders schwierig machen. Sehr oft müssen sie innerhalb eines strengen Rechtsrahmens verwaltet werden, was weitere spezielle Anforderungen und Probleme mit sich bringt. Bei den Problemen wird u. a. eingegangen auf das Fehlen geeigneter im Handel erhältlicher Systeme für die Langzeitarchivierung dieser Art von Informationen, die besonderen Metadatenanforderungen bei wissenschaftlichen und technischen Aufzeichnungen (und Schwierigkeiten bei der Erfassung dieser Metadaten), und den überaus vielgestaltigen Charakter von Informationsquellen angesichts Hunderter Anbieter mit jeweils eigenen Dateiformaten und -strukturen. Allgemein lassen sich die Daten am besten mit Multimedia in Content-Form vergleichen, und schon für einen einzigen Datensatz können die erzeugten Informationsmengen gewaltig sein (bis in den Gigabyte-Bereich). Sehr häufig sind die Unternehmen und Organisationen, die diese Aufzeichnungen erstellen, international tätig, was die Lösungssuche aus technischer und rechtlicher Sicht noch zusätzlich erschwert. Ferner befasst sich diese Analyse mit Wahlmöglichkeiten für die Anschaffung von Archivierungssystemen und der Überwindung des Problems der Datenveraltung vor dem Hintergrund sehr kurzer Zeitabstände für die Erneuerung der EDV-Technologie.

Neben den technischen Problemen geht es auch um die betriebswirtschaftlichen Fragen des Erwerbs eines Systems und der Verfahren, die benötigt werden, um es kostengünstig betreiben und pflegen zu können. Gerade dieser Gesichtspunkt ist vielleicht auf längere Sicht ausschlaggebend für die Unternehmen. Aus Sicht der Unternehmensführung handelt es sich um einen der wichtigsten zu bedenkenden Aspekte, wenn darüber zu entscheiden ist, welche Strategien und Taktiken angewendet werden sollen. Diese Fragen werden auch im Rahmen der vorgestellten Strategiealternativen diskutiert. In dem Beitrag werden Erfahrungen bei der Konzipierung, Lieferung und Inbetriebnahme eines elektronischen Datenarchivs in einem großen, im Vereinigten Königreich ansässigen multinationalen Pharmaunternehmen herangezogen und alle Arten von Daten erörtert, auch solche, die im Bereich der klinischen und wissenschaftlichen Forschung erzeugt werden.

Ein erheblicher Teil der Pionierarbeit zur digitalen Archivierung wurde im öffentlichen Sektor geleistet, doch ist die Thematik oft genauso wichtig für Unternehmen in der privaten Wirtschaft, die beispielhafte Praktiken in ihre Umgebungen übernehmen werden müssen. Zusammengefasst lässt sich sagen, dass auf dem Gebiet der wissenschaftlich-technischen elektronischen Aufzeichnungen die besten Vorgehensweisen und Lösungen, mit denen wir uns auf dieser Zusammenkunft beschäftigen, gerade erst im Entstehen begriffen sind. Wir hoffen, dass dieser Beitrag zu ihrer Etablierung im öffentlichen wie im privaten Sektor beiträgt.

Conservation des archives scientifiques et techniques: stratégies et solutions

Philip Lord

Le présent exposé décrit et analyse les stratégies et les solutions que les organisations créatrices d'archives scientifiques et techniques (ainsi que d'archives d'entreprise «normales») peuvent adopter pour la gestion et l'archivage à long terme de leurs documents. Les secteurs d'activité particulièrement concernés par ce problème sont l'industrie pharmaceutique, la santé, la chimie et l'industrie manufacturière, autant de secteurs qui sont d'une importance vitale sur le plan économique en Europe. Dans tous ces secteurs, la protection des actifs numériques peut être de la plus extrême importance pour leur efficacité économique — et peut même toucher directement leur viabilité: très souvent, les actifs numériques sont au cœur de leur réussite et de leur rentabilité permanente. D'autres secteurs économiques et d'autres organisations sont également de plus en plus tributaires d'une bonne gestion à long terme de leur patrimoine numérique, quelle qu'en soit la nature, et peuvent même créer des documents semblables à ceux en provenance des milieux scientifiques et techniques, notamment lorsqu'ils emploient les techniques du multimédia.

Les archives techniques et scientifiques possèdent des caractéristiques qui présentent des difficultés particulières. Très souvent, elles doivent être gérées dans un cadre réglementaire strict, qui présente des critères et des aspects particuliers. Les problèmes évoqués seront les suivants: absence sur le marché de systèmes capables de supporter la conservation à long terme de ce type d'information, besoins particuliers des archives scientifiques et techniques en matière de métadonnées (et difficultés à collecter celles-ci) et, enfin, diversité des sources d'information dans un environnement rassemblant des centaines d'entités versantes, chacune possédant ses propres formats et structures de fichiers. En règle générale, les données s'apparentent, de par leur contenu, au multimédia, et les volumes d'information ainsi générés pour un document seulement peuvent être énormes (éventuellement de l'ordre de plusieurs gigaoctets). Les sociétés et organisations créatrices de ces documents opèrent souvent à l'international, ce qui superpose une nouvelle couche de complexité technique et réglementaire dans la recherche d'une solution. Cette analyse intégrera un examen des options possibles pour acquérir des systèmes d'archivage et pour remédier au problème de l'obsolescence des données dans un environnement informatique très changeant.

Parallèlement aux aspects techniques à résoudre, nous aborderons les questions commerciales liées à l'acquisition d'un système et les opérations nécessaires pour le faire fonctionner et assurer sa rentabilité, cet aspect étant sans doute à long terme le plus critique pour les entreprises. Du point de vue des dirigeants d'entreprise, ce sera là l'un des aspects majeurs à prendre en considération au moment de décider des stratégies et des tactiques à adopter. Ces divers aspects seront également évoqués dans le cadre des options stratégiques proposées. L'exposé s'inspirera de l'expérience acquise dans la conception, la livraison et la mise en route d'un système d'archivage de données électroniques dans une grande entreprise pharmaceutique multinationale basée au Royaume-Uni, et abordera tous les types de données, y compris celles générées dans le milieu scientifique et hospitalier.

L'essentiel des travaux d'avant-garde sur l'archivage numérique est réalisé dans le secteur public, mais la question est tout aussi importante pour les entreprises privées, qui auront besoin de développer de bonnes pratiques dans leur environnement propre. En conclusion, les bonnes pratiques et solutions qui font l'objet de cette conférence ne font que voir le jour dans le domaine des archives électroniques techniques et scientifiques, et nous espérons que le présent exposé contribuera à leur mise en place à la fois pour le public et le privé.

Jan Dalsten Sørensen

- *Born 1970*
- *MA (history and Latin), University of Aarhus, 1999*
- *Archivist at the National Archives of Denmark, Department of Appraisal and Transfer, Section for Electronic Records, 1999*
- *Head of Section for Electronic Records, 2001*

Standards for metadata and documentation: practical experiences

Jan Dalsten Sørensen

The intention of this paper is to present how we today handle the metadata description and documentation of the electronic records that are transferred to the Danish National Archives in order to make the records available for future use. After a few introductory remarks on the terminology that I use and the way the work with electronic records is organised at the Danish National Archives, my paper will outline the principles of the description and documentation of the electronic records that are transferred to the archives.

1. Terminology and organisation

Basically, we consider all electronic systems to fall into one of these three categories:

1. Databases. These can contain almost any kind of information gathered by the authorities in connection with their business, but they are not used for document-management purposes.
2. Traditional document management systems that digitally store metadata about documents and — at least in the Danish administrative tradition — case files but where the actual documents are physical paper documents and thus not part of the electronic system.
3. Electronic document management systems where both the documents themselves and the relevant metadata are stored digitally.

The Danish National Archives has an IT department that sets the standards and formats for the delivery of electronic records (see paper by Birgit Hansen and Jan Danielsen for more information). It is, however, the section for electronic records within the Department of Appraisal and Transfer, that handles the day-to-day business with the authorities concerning appraisal, approval of new systems, preparation of the transfer of electronic records as well as the test of the transferred records. We deal mainly with State authorities since they, under the archival law, are obliged to deliver electronic records worthy of preservation to the National Archives according to a schedule determined by the National Archives. The same obligation does not apply to regional and local governments.

Appraisal and approval

For a long time, it has been mandatory for State authorities to notify the National Archives on new document management systems. Document-management systems have to go through a process of approval in order to make sure that archival issues are considered before the system has been taken into use. Since February 2000, it has also been mandatory for State authorities to notify the National Archives on new databases. Databases are not subject to approval but the notification takes place so that the databases can be appraised as early as possible. At the time of approval and/or appraisal the date for the first delivery of data from the system concerned is set. Usually, the first transfer takes place after five years of operation.

Transfer

Approximately six months before the first transfer from a database and approximately one year before the first transfer from a document management system, we write to the authority responsible for the system asking for a meeting. At the meeting we go through our regulations and try to identify any specific problems the authority might have in that respect. After this meeting, the National Archives issue a specification of the delivery with the basic information that the authority needs in order to make a correct archival version of the system concerned. The concept of the archival version will be dealt with below.

2. The archival version

The IT department has chosen the migration strategy for the long-term preservation of electronic records. This means that when data are transferred to the National Archives it must be migrated to a system-independent format. The consequence is loss of the functionality of the software application that created data. Therefore, it is a main concern when we are preparing the transfer of a particular electronic system, to ensure that the electronic records eventually can be made available to future users. The task is to transfer data from the systems with sufficient metadata and context information. Those elements — the data itself, the metadata description, the context information and the specific archival information — are put together in a package called the 'archival version'.

Metadata

An archival version of an electronic system consists of the tables of the original system stored as individual, sequential files in a system-independent format. System independence requires a well-defined metadata standard that describes the content of the tables and other information needed to reuse data in the future. As mentioned above, metadata can be interpreted as data about documents and case files. In the following, however, the term will be used to denote data about the system itself (its tables, fields, etc.)

Our metadata notation is described using EBNF (Extended Bachus-Naur Form). Thus, we make sure that the tables of any archival version, no matter which system generated them, are described in exactly the same way. This is, of course, necessary in order to be able to make some sort of a standardised retrieval tool for all the archived data.

The machine-readable metadata file that describes the tables includes:

- the title of each table;
- whether the table contains records of variable or fixed length;
- definitions of each field in the table, its data type and length as well as plain text that describes the field;
- identification of fields that have a specific function in document management systems. These functions include:
  - — unique identification of titles of case files;
  - — unique identification, title, date and sender/recipient of documents as well as information about the form in which the document is stored (on paper, digitally, or both);
  - — if two or more authorities share the same document management system it is also necessary that there is a way to identify which authority has been responsible for the registration of a particular case or document. The field that contains this information must also be identified;
- key definitions (identification of primary and foreign keys) in order to be able to recreate the relations between the tables;
- definitions of encoded fields. Many, or even most, databases and document management systems contain numeric fields where the value is some sort of code. Evidently, future reuse of data will be impossible if these codes are not defined in the archival version.

For document management systems, 3 to 10 SQL queries must be defined. These queries should be based upon the searches most commonly performed by the authority. The queries will give future users an idea of how the authority has used the system while in operation and thus make up for the loss of functionality.

General information

A metadata description of the tables is not enough to guarantee the future reuse of the data. It is necessary to ensure additional documentation and a selection of general information must therefore be included in the archival version. It has not been possible to make an actual standard for

this since we receive data from all kinds of systems where the types of documentation available and necessary are equally diverse. However, we try to make sure that the general information includes an administrative description of the system covering topics such as: 'Why was this system created and used?' and 'How was the system used?'.

Obviously, the types of general information vary with the type of system concerned. If it is a document management system, the 'why' is easily answered. What is more important here is the 'how' because it will be important for the future user to be able to understand the process in which data has been created. To answer that, we include users' manuals as well as instructions that have been written specifically for the authority that has used the system.

If the system concerned is a database, it is of great importance to have access to documents which illustrate the purpose of the system. If the database for instance has been created to facilitate an authority's administration of a particular law, the law should be included as general information. If a form has been used in connection with the gathering of information for the database, obviously an example of that form should be included.

We also want technical documentation, which illustrates the structure and functionality of the system, e.g. ER diagrams and screen dumps. All this information must be stored as TIFF files.

Usually we are satisfied with the material the authority already has. Only in rare cases do we ask them to create new documents for the sole purpose of adequate general information.

It should be mentioned that we do not make a description ourselves. The basic explanation is lack of time. Our main priority is to get archival versions of all the systems which we find valuable for future research in due time before technological obsolescence sets in or before the authorities have to erase information according to the Data Protection Act.

Information for archival purposes

Finally, additional information is needed specifically for archival purposes. This information is found in the form of descriptive tables with data about e.g. provenance. The Danish National Archives are based upon the principle of provenance. That is, however, not always an easy principle to deal with because it in some cases can be hard to determine a single records creator. Especially in the field of electronic records, it is fairly easy for two, three or even several authorities to share one electronic system. What is the provenance of the archival version of such a system? In connection with the development of a new archival database, DAISY (Danish Archival Information System), a new descriptive method has been invented. The basis of this method is to focus on the records themselves, whether paper based or digital, and then attach any number of record creators to the records concerned. In the case of electronic records, the information about provenance — the authority or authorities that created the data — is stored in a very simple descriptive table with three columns: one for the name of the record creator, one for the date when that particular authority started typing data into the system and one for the date when it stopped. The table includes a row for each record creator and will eventually be included in the DAISY for search purposes.

Other descriptive tables specify the period of time in which the data was generated and the name of the system. There is also a table that gives an index of the documents providing general information. The National Archives, in cooperation with the authority or authorities involved, decide the content of these descriptive tables. It is also necessary that every archival version is identified by a unique identification number, just as each CD in the archival version will get a media identification number derived from the id number of the archival version.

3. Problems and challenges

When an archival version is transferred it is tested to make sure that it meets our standards. Although we usually have a fairly good cooperation with the authorities, hardly any process of transfer goes without problems concerning the creation of an archival version that meets our standards.

Authorities who are asked to produce archival versions of electronic systems often have a hard time generating correct metadata. Common errors are the results of careless mistakes, e.g. missing mark-up signs (< or >). Missing code definitions or problems describing the keys are also often seen. A main reason for this is probably that the authorities (or the company that produces the archival version for the authority) have not been able to check the metadata in an efficient way. In March 2002 we published a slightly revised circular on the delivery of electronic records. One of the revisions that we made was adapting our metadata notation to XML simply by adding an XML header. In that way, the authorities will be able to parse and test metadata on their own using a standard XML tool.

It ensures first of all, that any authority is able to test the metadata description of the electronic system before the actual delivery to the National Archives. This will hopefully reduce the numbers of careless mistakes as well as errors, which are based upon misunderstandings of the regulations.

Code definitions have posed another problem rather than just the minor error where a few codes have been forgotten when producing the metadata file. Several databases rely heavily on codes that in one way or another can be described as global. This applies for instance, to the databases from the National Board of Health where numerous systems contain thousands of codes according to WHO classifications of diseases. Obviously, it would be a very difficult task to type all those definitions into the metadata file — and so far we have accepted that an edition of the classifications is scanned and stored as part of the general information instead of an actual definition of each code.

According to Danish law, records in the archives are not publicly accessible until after 30 years or, in most cases when we talk about electronic records, 80 years (because they usually contain information concerning identifiable persons). The rules that set the previously-mentioned standard for machine-readable metadata have only been in effect since February 2000. This means that we have had very little experience with the actual use of data stored in our archives according to the present standards. An important task in the coming years will be to gain more practical experience with the actual reuse of archived data and thus get a better basis for improving the description of electronic records.

Normas sobre metadatos y documentación. Experiencias prácticas

Jan Dalsten Sorensen

El presente documento resumirá las soluciones relativas a metadatos y documentación adoptadas por el actual conjunto de normativas referentes al envío de documentos electrónicos a los archivos nacionales, y su manejo en la práctica. El documento presentará la estructura y el contenido de una «versión de archivo» de un sistema electrónico y la descripción de los metadatos en particular. También abordará brevemente cuestiones tales como la organización de la cooperación entre los archivos nacionales y los organismos estatales, y los problemas que hemos detectado hasta ahora. Este extracto es un resumen de los puntos más importantes.

La «versión de archivo»

Una versión de archivo de un sistema electrónico está compuesta por los cuadros del sistema original almacenados como ficheros individuales y secuenciales en un formato independiente del sistema. La estrategia subyacente es la estrategia de migración, que se escoge para garantizar que no se necesita un *software* concreto para recuperar los datos.

La independencia del sistema requiere una norma bien definida de metadatos que describa el contenido de los cuadros, así como cualquier otra información necesaria para reutilizar los datos en el futuro. Nuestra notación de los metadatos se describe utilizando el sistema EBNF (Extended Bachus-Naur Form). Así pues, nos aseguramos de que los cuadros de cualquier versión de archivo, independientemente del sistema que los haya generado, se describan exactamente de la misma manera. Esto es, por supuesto, necesario para poder elaborar una herramienta normalizada de recuperación para todos los datos archivados. El fichero de metadatos, legible por máquina, describe los campos de los cuadros (el tipo de datos, su longitud y su función) e identifica las claves y las relaciones entre los cuadros. También proporciona las definiciones de los campos codificados, etc.

Sin embargo, una descripción de los metadatos de los cuadros no basta para garantizar la reutilización futura de los datos. Por tanto, en la versión de archivo debe incluirse una selección de la información general. No ha sido posible establecer una norma a estos efectos, puesto que recibimos datos de todo tipo de sistemas. Sin embargo, intentamos garantizar que la información general incluya una descripción administrativa del sistema, que cubra cuestiones como «¿por qué se creó y utilizó este sistema?» o «¿cómo se utilizaba el sistema?». También necesitamos documentación técnica que ilustre la estructura y funcionalidad del sistema, como por ejemplo diagramas ER y vaciados de pantalla. Toda esta información debe almacenarse como ficheros TIFF.

Por último, la información adicional es necesaria específicamente a efectos de archivo. Es necesario que cada versión de archivo se identifique con un código de identificación único; al igual que cada CD de la versión de archivo, obtendrá un código de identificación derivado del número de identificación de la versión de archivo. La información sobre la procedencia (el organismo o los organismos que crearon los datos), el período de tiempo en que se generaron los datos, el nombre del sistema, así como un índice de los documentos que proporcionarán información general figurarán en forma de cuadros auxiliares cuyo contenido lo deciden, en gran parte, los archivos nacionales.

¿Cómo se organiza nuestra cooperación con los organismos?

Al evaluar un sistema electrónico, ponemos una fecha para el primer envío a los archivos nacionales con procedencia del sistema, generalmente transcurridos 5 años. Aproximadamente 6 meses antes de ese momento, escribimos al organismo responsable del sistema solicitando una reunión en la que podamos presentar nuestra normativa. También tratamos de detectar los problemas específicos que el organismo pueda tener con respecto a la descripción de metadatos del sistema.

¿Qué problemas hemos encontrado?

Los organismos a quienes se solicita que elaboren versiones de archivo de los sistemas electrónicos suelen tener dificultades para generar metadatos correctos. Los errores más comunes son resultado de descuidos, por ejemplo falta de caracteres de señalamiento (< o >). También suelen darse falta de definiciones de códigos o problemas con la descripción de las claves. Esperamos solucionar estos problemas adaptando nuestra notación de metadatos a XML. De esta manera, los organismos podrán analizar y probar metadatos por su cuenta utilizando una herramienta estándar XML.

Según la legislación danesa, los registros de los archivos no son accesibles al público hasta después de 30 años o, en la mayoría de los casos, por lo que se refiere a los registros electrónicos, 80 años (pues generalmente contienen información relativa a personas identificables). La normativa por la que se estableció la norma previamente mencionada para metadatos legibles por máquina está vigente sólo desde febrero de 2000. Esto significa que contamos con muy poca experiencia en el uso de datos almacenados en nuestros archivos según las normas actuales. Por tanto, no podemos estar seguros de si nuestras normas sobre metadatos y otra documentación resultarán adecuadas. Sólo el tiempo lo dirá.

Standards für Metadaten und Dokumentation: praktische Erfahrungen

Jan Dalsten Sørensen

In meinem Beitrag werde ich die Lösungen zu Metadaten und zur Dokumentation des bestehenden Katalogs von Regelungen für die Überführung elektronischer Aufzeichnungen an das Nationalarchiv darlegen und zeigen, wie wir diese Lösungen in der Praxis handhaben. Ich werde die Struktur und den Inhalt der „Archivversion" eines elektronischen Systems und insbesondere die Metadatenbeschreibung vorstellen. Außerdem werde ich kurz auf Fragen wie die Organisation der Zusammenarbeit zwischen Nationalarchiv und Außenstellen und die Probleme eingehen, mit denen wir es dort bisher zu tun hatten. In dieser Zusammenfassung sollten die wichtigsten Punkte kurz umrissen werden.

Die „Archivversion"

Die Archivversion eines elektronisches Systems besteht aus den Tabellen des Originalsystems, die als einzelne sequenzielle Dateien in einem systemunabhängigen Format gespeichert werden. Dieses Vorgehen basiert auf der Migrationsstrategie, die gewählt wird, um sicherzustellen, dass für das Wiederauffinden der Daten keine bestimmte (und proprietäre) Software notwendig ist.

Systemunabhängigkeit verlangt einen hinreichend definierten Metadatenstandard, der den Inhalt der Tabellen und andere Informationen beschreibt, die für die Wiederverwendung von Daten in der Zukunft benötigt werden. Unsere Metadatennotation wird mittels der EBNF (Erweiterte Backus-Naur-Form) beschrieben. Auf diese Weise stellen wir sicher, dass die Tabellen jeder Archivversion unabhängig von dem System, mit dem sie erzeugt wurden, auf exakt die gleiche Weise beschrieben werden. Dies ist natürlich notwendig, damit ein einheitliches Retrieval-Hilfsmittel für alle archivierten Daten zur Verfügung steht. Die maschinell lesbare Metadatendatei beschreibt die Felder der Tabellen (Typ, Länge und Funktion ihrer Daten) und gibt die Ordnungsbegriffe und Beziehungen zwischen den Tabellen an. Außerdem enthält sie die Definitionen von codierten Feldern u. a. m.

Eine Metadatenbeschreibung der Tabellen reicht allerdings nicht aus, um die künftige Nutzung der Daten zu gewährleisten. Deshalb muss die Archivversion auch eine Auswahl allgemeiner Informationen beinhalten. Bisher ist es noch nicht gelungen, einen echten Standard dafür zu erstellen, da wir Daten von ganz unterschiedlichen Systemen erhalten. Wir versuchen jedoch sicherzustellen, dass die allgemeinen Informationen auch eine administrative Beschreibung des Systems enthalten, etwa zu Fragen wie „Warum wurde dieses System erstellt und benutzt?" und „Wie wurde das System benutzt?". Außerdem planen wir eine technische Dokumentation zur Veranschaulichung von Struktur und Funktionalität des Systems, z. B. ER-Diagramme und Bildschirmabzüge. All diese Informationen müssen als TIFF-Dateien abgespeichert sein.

Benötigt werden nicht zuletzt auch zusätzliche Informationen speziell für Archivierungszwecke. Jede Archivversion ist durch eine eindeutige Identifizierungsnummer gekennzeichnet, und jede CD in der Archivversion erhält eine Datenträger-Identifizierungsnummer, die aus der ID-Nummer der Archivversion abgeleitet ist. Informationen über die Provenienz – die Behörde bzw. Behörden, die die Daten erstellt haben, der Zeitraum, in dem die Daten erzeugt wurden, der Name des Systems sowie ein Index der Dokumente mit allgemeinen Informationen – sind in Gestalt von Hilfstabellen zu finden, über deren Inhalt weitgehend das Nationalarchiv entscheidet.

Wie ist unsere Zusammenarbeit mit den Behörden organisiert?

Bei der Bewertung eines elektronischen Systems legen wir ein Datum für die erste Übergabe vom System an das Nationalarchiv in der Regel nach fünf Jahren fest. Ungefähr sechs Monate vor diesem Zeitpunkt schreiben wir die für das System zuständige Behörde an und bitten um ein Treffen, bei dem wir unsere Regelungen durchgehen können. Wir versuchen außerdem herauszufinden, welche Probleme die Behörde eventuell bezüglich der Metadatenbeschreibung des Systems hat.

Auf welche Probleme sind wir gestoßen?

Behörden, die gebeten werden, Archivversionen von elektronischen Systemen zu erzeugen, haben oft Schwierigkeiten, korrekte Metadaten zu generieren. Häufig sind Fehler auf Nachlässigkeiten zurückzuführen, z. B. fehlende Markup-Zeichen (< oder >). Oft anzutreffen sind auch fehlende Code-Definitionen oder Probleme mit der Beschreibung der Ordnungsbegriffe. Wir hoffen, diese Probleme durch Anpassung unserer Metadatennotation an XML lösen zu können. Auf diese Weise können die Behörden dann Metadaten mit einem XML-Standardtool selbst prüfen und testen.

Nach dänischem Recht ist Archivgut erst nach 30 Jahren öffentlich zugänglich oder in den meisten Fällen, wenn es um elektronische Unterlagen geht, nach 80 Jahren (da sie für gewöhnlich Informationen enthalten, die identifizierbare Personen betreffen). Die Vorschriften, mit denen der zuvor erwähnte Standard für maschinell lesbare Metadaten festgeschrieben wird, sind erst seit Februar 2000 in Kraft. Aus diesem Grund haben wir bisher kaum Erfahrungen mit der eigentlichen Nutzung von Daten, die in unseren Archiven nach den jetzt geltenden Standards aufbewahrt werden. So können wir nicht sicher sein, ob sich unsere Standards für Metadaten und andere Dokumentationen als geeignet erweisen. Dies wird sich erst nach einiger Zeit zeigen.

Normalisation des métadonnées et de la documentation: expériences pratiques

Jan Dalsten Sørensen

Mon exposé présentera les solutions adoptées en matière de métadonnées et de documentation dans la réglementation actuelle applicable au versement d'archives électroniques dans les Archives nationales et la manière dont nous traitons cette question dans la pratique. L'exposé présentera la structure et le contenu de la «version d'archives» d'un système électronique ainsi que la description des métadonnées. Il abordera également certaines questions, telles que l'organisation de la coopération entre les Archives nationales et les organismes versants, et les problèmes rencontrés jusqu'à présent. Le résumé présenté ci-après en récapitule les principaux points.

«Version d'archives»

La version d'archives d'un système électronique est constituée des tables du système original, mémorisées sous forme de fichiers séquentiels individuels dans un format indépendant du système. La stratégie retenue est une stratégie de transfert, choisie pour veiller à ce qu'aucun logiciel particulier (et de marque déposée) ne soit nécessaire pour récupérer les données.

L'indépendance par rapport au système implique une norme parfaitement définie pour les métadonnées qui décrive le contenu des tables et les autres informations nécessaires à la

réutilisation à terme des données. La notation retenue pour les métadonnées fait appel à une version étendue de la forme de Bachus-Naur (EBNF). De la sorte, les tables de n'importe quelle version d'archives, quel que soit le système qui les a générées, sont décrites exactement de la même manière. Cette uniformité est bien entendu nécessaire pour pouvoir constituer un outil de recherche standard pour l'ensemble des informations archivées. Le fichier des métadonnées lisibles par machine décrit les différents champs des tables (le type de données, la longueur des blocs et leur fonction) et identifie les clés et les relations entre les tables. Il donne également la définition des fichiers encodés, etc.

Toutefois, la description des métadonnées des tables ne garantit pas en soi la possibilité de réutiliser à terme les données. Une sélection d'informations générales doit donc être incorporée dans la version d'archives. Il ne nous a pas été possible de normaliser cette procédure car nous recevons des données d'une multitude de systèmes. Cependant, nous essayons de vérifier que les informations générales incluent une description administrative du système, répondant à plusieurs questions telles que: «Pourquoi ce système a-t-il été créé et est-il utilisé?», «Comment le système a-t-il été utilisé?», etc. Nous demandons également certains renseignements techniques précisant la structure et les fonctionnalités du système, notamment les graphes ER (entités-relations) et les vidages d'écran. Toutes ces informations doivent être stockées sous la forme de fichiers TIFF.

Enfin, d'autres informations sont nécessaires aux fins spécifiques de l'archivage. Ainsi, il est impératif que chaque version d'archives soit identifiée par un code unique, de même que chaque CD de la version d'archives recevra un code d'identification support, obtenu à partir du code d'identification de la version d'archives. D'autres renseignements sur l'origine — nom de l'organisme créateur des données, période de création des données, nom du système — ainsi qu'un index des documents fournissant des renseignements à caractère général figureront sous la forme de tables annexes, dont le contenu est fixé pour une large part par les Archives nationales.

Comment est organisée notre coopération avec les organismes créateurs?

Lorsque nous procédons à l'évaluation d'un système électronique, nous fixons la date du premier versement aux Archives nationales à un horizon qui est généralement de cinq ans. Environ six mois avant cette échéance, nous écrivons à l'organisme responsable du système pour le rencontrer et lui présenter nos règles de fonctionnement. Nous nous efforçons aussi d'identifier les problèmes particuliers que l'organisme peut rencontrer dans la description des métadonnées du système.

Quels sont les problèmes que nous avons rencontrés?

Les organismes qui sont invités à élaborer des versions archivistiques de leurs systèmes électroniques rencontrent souvent des difficultés à générer des métadonnées correctes. Les erreurs les plus fréquentes sont le résultat d'étourderies, par exemple l'oubli de balises de marquage (< ou >). On constate également souvent des omissions dans la définition des codes ou des problèmes de description des clés. Nous espérons pouvoir résoudre ces problèmes en adaptant notre notation des métadonnées au format XML. De la sorte, les organismes seront en mesure d'analyser et de tester par eux-mêmes leurs métadonnées à l'aide d'un outil XML standard.

En vertu de la législation danoise, les archives ne peuvent être consultées par le public qu'à l'expiration d'un délai de 30 ans ou, s'agissant d'archives électroniques, d'un délai de 80 ans (car elles renferment des informations concernant des personnes identifiables). Les dispositions fixant la norme dont il est question plus haut, applicable aux métadonnées lisibles par machine, ne sont en vigueur que depuis février 2000. Cela signifie que nous ne possédons que peu d'expérience sur l'utilisation effective des données stockées dans nos archives selon les normes actuelles. En conséquence, nous ne savons pas encore si les normes que nous avons définies pour les métadonnées et la documentation se révéleront adéquates. Seul le temps nous le dira.

Margaret Hedstrom

Dr Margaret L. Hedstrom is an Associate Professor at the School of Information, University of Michigan. Before joining the faculty at Michigan in 1995, she was Chief of State Records Advisory Services and Director of the Centre for Electronic Records at the New York State Archives and Records Administration (1985–95). Dr Hedstrom earned Master's degrees in library science (1977) and history (1979) and a PhD in history (1988) from the University of Wisconsin-Madison. Research: Dr Hedstrom has been conducting research on the management and preservation of electronic records for nearly 20 years. She is Project Director and Co-Principal Investigator of the Camileon project, a joint effort by the University of Michigan and the University of Leeds to investigate the feasibility of emulation as a digital preservation strategy. She is also director of a study of electronic record keeping in collaborative work environments, funded by the National Historical Publications and Records Commission. Her current research interests include strategies for long-term preservation of digital information, the impact of electronic communications on organisational memory and documentation, and remote access to archival materials. Teaching: Dr Hedstrom teaches graduate courses at the School of Information on access systems for archival materials, social systems

Significant properties of digital objects: definitions, applications, implications ([1])

Margaret Hedstrom, Christopher A. Lee

Introduction

Significant properties are those properties of digital objects that affect their quality, usability, rendering, and behaviour. In an ideal world, free from technical and economic constraints, libraries and archives would preserve their physical and digital collections in their original form with all significant properties intact. Resource constraints and technical limitations, however, often require libraries, archivists, and curators to preserve objects as facsimiles or surrogates in lieu of original materials and to weigh decisions about which properties to preserve against institutional priorities, available resources, limitations of available preservation methods and technologies, and anticipated use. Digital preservation usually entails converting or migrating digital information to current computing platforms, normalising objects and collections so that they are less dependent on specific hardware and software, developing tools to emulate obsolete platforms, or some combination of these approaches in order to cope with rapid obsolescence of computing technology. Any of these approaches may alter the structure, appearance, and behaviour of the original objects. The purpose of our research is to identify the significant properties of digital objects that affect their quality, functionality, and look-and-feel so that custodians can select appropriate methods which preserve those significant properties of digital objects that are deemed important by designated user communities. A formal articulation of significant properties also provides a framework for documenting the impact of transformations of digital objects that may be necessary to ensure continuing access to them and to explain the reasons for such actions.

The importance of retaining the significant properties of documents and artefacts has been raised in numerous contexts. Librarians and archivists created microfilm surrogates of millions of books, office documents, and newspapers to overcome physical limitations of highly acidic paper and to conserve storage space, even though the microfilm surrogates were not capable of preserving all of the significant properties of the original hard copy materials, such as accurate colour representation or the exact physical dimensions of the originals. Similar problems may occur when digital objects are converted from one file format to another or when they are normalised into a standard format upon ingest into a repository. Format conversions, for example, may change document structures, stylistic features, navigation, and appearance. Decisions about significant properties of digital objects are much more complex than decisions regarding physical items because properties are expressed at several different levels of abstraction and because there are many more options available for creating surrogates, facsimiles, and derivatives. Digital documents have many features and exhibit behaviors that may or may not be important to preserve.

Guidelines for selecting preservation formats for physical materials provide a starting point for identifying potentially significant attributes of digital objects. Previous discussions of significant properties have invoked needs such as authenticity, legal admissibility, artefactual value, ease or convenience of use, and aesthetic quality as reasons for preserving physical artefacts in their original physical form. According to a recent report on the role of the physical artefact in library collections, several cardinal principles guide decisions about preserving significant properties of physical artefacts, such as age, evidential value, aesthetic value, scarcity, associational value, market value, and exhibition value ([2]). Courts, for example, might challenge the admissibility of copies of original documents as evidence in legal proceedings if the reasons for and processes used to produce surrogates were not well documented. Scholars studying the evolution of book publishing glean much information from physical evidence of paper, typographical methods, and binding techniques. Concerns about significant properties are not limited to written or printed works. Photograph and film curators have to decide when it is necessary to preserve original negatives and which medium and format best captures significant features such as resolution and tonal qualities ([3]). Curators of recorded sound collections are concerned about fidelity, sampling rate, volume, and dynamic range when sound recordings are transferred to new analogue media or digitised.

([1]) The research for this project is part of the Camileon project, an international digital library initiative, at the University of Michigan and the University of Leeds. This work is funded by the National Science Foundation, award number 9905935 and by the Joint Information Systems Committee (JISC) in the UK. Additional information about the Camileon project is available at: www.si.umich.edu/CAMILEON/

([2]) *The evidence in hand: Report of the task force on the artefact in library collections*, Washington, D.C., Council on Library and Information Resources (November 2001), p. 9.

([3]) Schwartz, J., 'We make our tools and our tools make us': lessons from 'Photographs for the practice, politics, and poetics of diplomatics', *Archivaria* 40 (1995), pp. 40–74.

Currently, we do not have a formal method for identifying which properties of digital objects are considered significant or for describing significant properties. We build on the concept of significant properties from the Cedars project that treats decisions about which significant properties to preserve as collection-management decisions ([4]). Our goal is to provide an empirical foundation for making collection-management decisions where choices of preservation strategies may eliminate or alter some of the properties of original objects. Our model is modular and extensible. We treat common data types, such as numeric data, text, and images, as separate modules in the model so that collection managers can direct their attention to those types of data that are present in particularly complex digital objects or collections. The model is extensible in that particular communities may develop much more fully articulated definitions of significant properties that are relevant to the types of materials they preserve or to the needs of particular designated communities. Our work is in a formative stage, not only because of the technical complexity of digital objects, but also because we lack the accumulated social experience to guide decisions about which properties matter and to whom they matter.

Methodology

To develop an inventory and formal expression of potentially significant properties, we first explored the literature from information and library science, archival science, human-computer interaction, and various forms of criticism to evaluate the various properties of preserved information that scholars and other users have considered significant in the past. We found the literature on legal admissibility, authenticity of documents and artefacts, and intrinsic value, especially useful developing an inventory of potentially significant properties ([5]). Several studies have investigated the role of specific characteristics of documents, such as colour of text, layout, annotations, and even the appearance of book jackets ([6]). From the literature on human computer-interaction we drew on the concept of 'affordances' which refers to the functions that various types of digital objects support ([7]). Features such as browsing, viewing, annotating, and visualising digital content are examples of affordances. When encountering a preserved digital object, users may find the original affordances of the object important for accessing and making sense of the data it contains as well as for understanding the functional context in which it was used. We also found guidelines for converting documents and analogue recordings to digital form useful because they specify attributes such as resolution, colour matching, and fidelity that may represent transcendent properties that are significant for print, analogue, and digital information ([8]).

We also explored several common technical guidelines and standards for defining the properties or attributes of digital objects. We considered the Internet protocol MIME (multipurpose Internet mail extensions) which provides for the use of file headers and file name extensions to specify common encoding formats. We also extracted some properties from format-specific specifications, such as the Moving Picture Experts Group (MPEG) standards for digital video and the draft NISO data dictionary of technical metadata for digital still images ([9]). The work within the World Wide Web Consortium on style sheets — cascading style sheets (CSS) and extensible stylesheet language (XSL) — has delineated hundreds of properties that can be used to format and render structured and semi-structured digital objects ([10]). We also reviewed several metadata standards and specifications developed for the management and preservation of digital collections, such the tag library from the Text Encoding Initiative, the Cedars digital preservation metadata model, and the recent work of the OCLC/RLG Working Group on Preservation Metadata ([11]). Our review of numerous technical specifications and descriptive guidelines yielded an inventory of more than 800 specific property references.

A model for expressing significant properties

We developed a conceptual model of complex digital objects and their components which identifies common data types that can be assembled to create digital objects such as multi-level, multimedia web sites. Drawing on the definitions from the OAIS reference model, we defined a digital object as an object composed of a set of bit sequences ([12]). A complex digital object is a digital object composed of more than one type of component. We then mapped significant properties to common types of components, including numeric data, text, images, graphics, audio, video, and executables. All complex digital objects have internal compositional properties that manage how the components are assembled to compose a complex object. Complex digital objects may also have explicit external relationships to other digital objects, such hyperlinks that

([4]) Cedars project, www.leeds.ac.uk/cedars/

([5]) Menne-Haritz, A. and Brübach, N., 'The intrinsic value of archive and library material', Marburg: Archivschule, 1997; *Authenticity in a digital environment*, Washington D.C.: Council on Library and Information Resources, May 2000, National Archives and Records Service, *Intrinsic value in archival material*, Staff Information Paper 21, Washington D.C., NARS, 1982; Bearman, D., and Trant, J., Authenticity of digital resources: Towards a statement of requirements in the research process, *D-Lib Magazine* (June 1998). Available http://www.dlib.org/dlib/june98/06bearman.html; and MacNeil, H., *Trusting records: Legal, historical, and diplomatic perspectives*, Dordrecht: Kluwer Academic Publishers, 2000.

([6]) Baker, N., *Doublefold: libraries and the assault on paper*, New York: Random House, 2001; Marshall, C.C., Annotation: from paper books to the digital library, *Second ACM international conference on digital libraries*, Philadelphia: ACM Press (1997), pp. 131–140, O'Toole, J. M., 'On the idea of uniqueness', *American Archivist* 57 (1994), pp. 632–658; Pettersson, E., 'Automatic information processes in document reading: A study of information handling in two intensive care units', in *European conference on computer-supported cooperative work*, London (1989), pp. 63–72; Cadiz, J. J., Gupta, A., and Grudin, J., 'Using web annotations for asynchronous collaboration around documents', *Computer-supported cooperative work*, Philadelphia: ACM Press (2000), pp. 309–318; Toms, E. G. and Campbell, D. G., 'Genre as interface metaphor: exploiting form and function in digital environments', *32nd Hawaii International Conference on System Sciences*, IEEE Computer Society,1999; O'Connor, B. C., and O'Connor, M. K., 'Book jacket as access mechanism: An attribute-rich source for functional access to academic books. *First Monday*, 3:9 (September 1998).

([7]) Norman, D. A., *The design of everyday things*, New York, Doubleday, 1990, Phelps, T. A., and Wilensky, R., 'Multivalent documents: Inducing structure and behaviours in online digital documents', *Proceedings of Hawaii International Conference on System Sciences '96*, IEEE, 1996. Levy, D. M., 'Fixed or fluid? Document stability and new media', *European Conference on Hypertext Technology (ECHT '94)*, Edinburgh: ACM (1994), pp. 24–31.

provide a means for navigating from one object to another. We have not attempted to define all possible component types or all possible methods of internal composition. Rather, we have developed a decomposable model with a tree structure that defines common properties for each component type. In developing the model, we have discovered that certain properties apply to all types of components, some properties are unique to particular component types, and some properties are applicable to multiple components but not necessarily to all types.

A number of properties that are common across all component types need to be considered when selecting preservation strategies so that the technical approach is documented sufficiently to support access, retrieval, and low-level data management.

These include the definition of the basic data unit, byte-level encoding, data typing, and the logical structure or data model. Although these attributes often are specified or implied for specific file formats we have attempted to separate these properties from both the ways that data are stored on physical media and from file formats. These basic elements are similar to the concept of an underlying abstract form (UAF) which Holdsworth and Sergeant use to 'encapsulate the recognition that the data has an existence and a content separate from the medium upon which it is written. This underlying abstract form contains all the significant properties of the data and is independent of the medium upon which the data is written.' ([13]) Although low-level properties, such basic data unit, byte-level encoding, data type, and logical schema are germane to all types of components, the way these elements are expressed varies considerably from one component type to the next.

The underlying abstract form contains all of the significant properties of data necessary to access a digital document where access means 'realising the UAF on the technology appropriate to the time of access in such a way that the desired form of access (which may not necessarily be viewing) can be achieved.' ([14]) While the UAF is a useful starting point for identifying preservation formats which, at the highest level of abstraction, do not discard significant information, our concept of significant attributes is broader in that it takes into account the various types of stylistic and aesthetic features discussed in the literature above. For example, features such as font, spacing, and layout may be important to preserve textual components, especially if variations in these features alter the intended meaning of the text. Our formal expression of significant properties is recursive because many of the properties can be decomposed further into sub-properties. The property 'font,' for example, can be further decomposed into sub-properties such as font type, style, family, size, or colour. If changes in any of these sub-properties alter the appearance or the meaning of a digital object, they might be considered significant and worth replicating as they were specified in the original object. Some sub-properties are relevant to several different component types. Representations of colour, for example, might apply to images, graphics, video, and text, but not to audio or numeric data. We express colour properties, such as colour space, bit depth and colour model, in such a way that a common formal expression is applicable to all types of components. Definitions of significant properties that affect the aesthetics, implied meaning, and affordances of digital objects tend to be much more subjective and tied to the context of creation and use. Whether or not colour, for example, is a significant property of a given digital object or collection will depend on the extent to which colour features affect the quality and usability of the preserved object for a designated community.

Our model also includes a means for expressing internal composition and external relations between complex objects. By internal composition we mean the methods used to assemble components into complex objects and the means for associating various components with each other. Specifying how a complex object is assembled is important because there are several viable options for preserving complex objects, provided that their internal composition is specified. For example, a repository could preserve all of the components of a complex object along with the methods or procedures that manage the relationships among components, or it could preserve the components as separate files along with specifications for how to reconstruct those relationships. Our model also specifies the types of external relationships between different components. These relationships are important in choosing preservation strategies, especially when components of digital objects are accessed by way of link or pointer to another digital object. The problem of linked content is pervasive for web documents, and collection managers are grappling with policies and methods for dealing with linked objects. A variety of strategies have been proposed for addressing the persistence of links, but most include some sort of string to serve as an identi-

([8]) Kenney, A. R., and Reiger, O., *Moving theory into practice: Digital imaging for libraries and archives*, Mountain View, CA: RLG, Inc., 2000; and US Library of Congress, American Memory Programme, technical information and background papers, available at: http://memory.loc.gov/ammem/ftpfiles.html

([9]) Martínez, J. M., ed., *Overview of the MPEG7 Standard (Version 5.0)*, Singapore: International Organisation for Standardisation, 2001, available at: http://mpeg.telecomitalialab.com/standards/mpeg-7/mpeg-7.htm; MPEG Requirements Group. MPEG-21 Overview. Sydney, Australia, 2001, available at: http://mpeg.telecomitalialab.com/standards/mpeg-21/mpeg-21.htm; and *Draft data dictionary: Technical metadata for digital still images,* Version 1.0. National Information Standards Organisation, 2000.

([10]) World Wide Web Consortium, 'Cascading style sheets' http://www.w3.org/Style/ CSS/; and World Wide Web Consortium, 'eXtensible style sheet language', http://www.w3.org/TR/xsl/.

([11]) Cedars project, 'Metadata for digital preservation: The Cedars outline specification, 2000', available at: http://www.leeds.ac.uk/cedars/OutlineSpec.htm; Phillips, M., Woodyard, D., Bradley, K., and Webb, C., 'Preservation metadata for digital collections', National Library of Australia, 1999, available at: http://www.nla.gov.au/preserve/pmeta.html; Lupovici, C. and Masanès, M., 'Metadata for long term-preservation', Bibliothèque nationale de France, 2000, available: http://www.kb.nl/coop/nedlib/results/D4.2/D4.2.htm; Sperberg-McQueen, C. M., and Burnard, L., eds. *Guidelines for electronic text encoding and interchange: XML-compatible edition (TEI P4)*: TEI Consortium, 2001, available at: http://www.tei-c.org/P4X/; and 'A recommendation for content information. A report by the OCLC/RLG Working Group on Preservation Metadata', (October 2001), available at: www.aclc.org/ research/pmwg/.

([12]) Consultative Committee for Space Data Systems, 'Reference model for an open archival information system (OAIS)', *Red book*, Issue 1.2, Mountain View, CA, June 2001.

([13]) Holdsworth, D. and. Sergeant, D. M., 'A blueprint for representation information in the OAIS model', IEEE Mass Storage Conference (March 2000), available at: www.leeds.ac.uk/cedars/.

([14]) Holdsworth and Sergeant, *A blueprint for representation information in the OAIS model: 416.*

fier, a specification of a domain within which the identifier is meant to apply, and some mechanism for resolving the identifier as a reference to a specific location within a storage system.

Our approach simply provides a means to express the attributes of external relations. In our model, the properties of external relations include such factors as:

- cardinality (one to one, one to many, many to many)
- direction (outbound, inbound, third party, bi-directional)
- semantic relation
- locality (specified by a value, specified by a location)
- access type (ftp, afs, local file)
- target type (pointer, character string, binary object, calculated value, bounded area),

etc.

By specifying the properties of external relations, we provide a basis for breaking down the functions and behaviours of hyperlinks and pointers so that collection managers can decide which properties of the external relations are important to preserve. It is worth noting that the properties of external relations are also applicable to the internal composition of digital objects, if their components are associated with each other through pointers or links. Whether a link or pointer supports internal composition or external relations depends on where a repository draws the boundaries of a given digital object. For example, if a web page links to other web pages within the same site and that site is preserved as a single digital object within a repository, then the links would be considered properties of that digital object's internal composition. Links from those pages to other pages not preserved in the repository could be treated as external relations.

Significant properties and digital preservation strategies: applications

A formal expression of significant properties of complex digital objects has many general and practical applications. Such a model can be applied to appraisal and selection of digital materials, to assessing the risk of information loss associated with various preservation strategies, to the development of preservation metadata, to documenting the basis for preservation decisions, and to the automated management of complex digital objects. In future research, we will also test the model as a basis for developing cost comparisons and trade-offs associated with various digital preservation strategies.

Appraisal and collection management decisions about complex digital objects entail choices about which significant properties of objects and their components merit preservation. For example, a repository might decide to preserve the text, images, and graphics of a web page along with procedures to reproduce its original composition, but to forego the video segments. The repository may decide to preserve the images at a lower resolution than the original. No single set of decision-making rules can or should apply to all such choices. Rather, decisions about which significant properties to maintain will depend on institutional priorities, anticipated use, knowledge of the designated community, the types of materials involved, and the financial and technical resources available to the repository. Nevertheless, a formal means for expressing significant properties provides the basis for more clearly articulating the range of available options and for documenting the trade-offs among them.

Collection managers can also use a formal expression of significant properties to select preservation techniques and to guide choices about the timing of various preservation actions. Anne Kenney and her colleagues at Cornell University have completed two studies that applied risk management methods to digital preservation strategies ([15]). Their multi-stage process includes risk identification, risk classification, risk assessment, risk analysis, and risk management implementation. Our formal model of significant properties could be used for identification, classification, and risk analysis. A repository using a risk management approach could evaluate the risks of technology obsolescence for various component types, the risks of changes to objects that are externally linked, and the risk that a preservation action might degrade the object beyond the point of use-

and collections, electronic records management, and archives as memory institutions. She has also taught numerous continuing education workshops in the United States, Canada, Australia, New Zealand, Sweden, and South Africa.
Professional activities and service: Hedstrom has participated in numerous elected, appointed, and voluntary professional and service activities. She is currently a member of the National Digital Strategy Advisory Board which is advising the Library of Congress on development of a national infrastructure for digital preservation. She is a member of an expert reference group on IT questions for the Swedish Government. She served on the National Research Council Study Committee on the Digital Future of the Library of Congress, 1999–2000; the Commission on Preservation and Access/Research Libraries Group Task Force on Archiving Digital Information, 1994–96; on the Council of the Society of American Archivists, 1992–95; and as Secretary, Vice-Chair, and Chair of the New York State Forum for Information Resources Management, 1992–95. She has been a consultant to more than a dozen government archival programs, the World Bank, the International Council on Archives, and recently the University of Fort Hare in South Africa. Hedstrom is a fellow of the Society of American Archivists and she was the first recipient of the annual Award for Excellence in New York State Government Information Services.

Christopher Lee

Christopher Lee is currently pursuing a Ph.D. at the University of Michigan School of Information after previously obtaining a Master of Science

([15]) Kenney, A. R., et al, 'Preservation risk management for web resources: Virtual remote control in Cornell's project prism, *D-Lib Magazine* 8:1 (January 2002); available at: http://www.dlib.org/dlib/january02/kenney/01kenney.html; and Lawrence, G. W., et al, *Risk management of digital information. A file format investigation*, Washington D.C.: CLIR (June 2000), available: http://www.clir.org/pubs/reports/reports.html.

degree in information with a concentration in archives and records management there in 1999 and a Bachelor of Arts in Philosophy (summa cum laude with honours) with a concentration in public service at Albion College in 1995.
He has been a graduate research assistant in the Camileon (creative archiving at Michigan and Leeds: emulating the old on the new) project since August 2000. Prior to that he worked as an electronic records project archivist at the Kansas State Historical Society between May 1999 and August 2000. Other professional activities include: project member, LUIGUI (Linux/UNIX Independent Group for Usability Information), February–November 1999; co-investigator, Online Archival Description Project, University of Michigan School of Information, January–May 1999; graduate research assistant for Dr Michael Cohen, University of Michigan School of Information, September 1998–May 1999; project team member, Formshare, University of Michigan School of Information — September 1997–September 1998; webmaster, technical advisor and archival processor, University of Fort Hare, National Heritage Cultural Studies Centre, Alice, South Africa, May–June 1998.
Recent publications and presentations include: Guerilla ERM: Lessons learned from some time in the trenches, Ohio Archivist, Spring 2001; Open-source software: A promising piece of the digital preservation puzzle, electronic currents, Midwest Archives Conference (MAC) Newsletter, Volume 29, Number 2 (113), October 2001; Electronic records and government accountability:

fulness. As repositories accumulate experience with such decisions, a common vocabulary for expressing significant properties would provide the basis for sharing data and assessments about the impact of various preservation strategies.

The formal model can assist in selecting appropriate preservation strategies based on acceptable levels of risk defined by each custodial institution. In the Camileon project, we are investigating the feasibility of using emulation as a digital preservation strategy. The results of our research show that emulation is technically feasible, but it is not always most effective or cost-effective strategy. If the primary goal of preservation is to preserve the content and structure of a collection of textual documents, then converting the materials to structured text using a mark-up language such as SGML or XML, or to a page description format, such as PDF, may be more cost-effective and easier technically than emulating the original computer environment. If, on the other hand, properties such as navigational aids, optional viewing methods, macro-enabled procedures, or internal links were considered significant to the designated community, then emulation might be a better choice for preserving this range of functionality.

The model can be applied at several points in the digital preservation process. The open archival information system reference model (OAIS) provides a common set of terms for describing the processes and information flow in an archival repository. Using the OAIS terminology, the formal expression of significant properties is applicable to transformations that might be necessary when resources are extracted from their original creation environment and used to create a submission information package (SIP); when the SIP is transformed into an archival information package (AIP); when it is necessary to repackage or transform an AIP; or when a dissemination information package (DIP) is produced for a consumer. Our model for expressing significant properties is also closely related to the concept of a representation network in the OAIS reference model. The OAIS model defines a category of information called representation information that accompanies a digital object and is used to convert bit sequences into more meaningful information. Representation information is further divided into two types: structure information and semantic information. Structure information describes the format or underlying data structure, similar to the lowest-level properties in our model that are common across all data types, such as the basic data unit, byte-level encoding, data typing, and logical structure. Semantic information is information contained in or associated with an object that is necessary to interpret its meaning. In the case of textual information, for example, an indication of the language or languages represented in the text would be important for future interpretation of the content. Similarly, meaningful analysis of scientific data requires additional semantic information that describes how meaning is ascribed to data elements and how different pieces of data are related ([16]).

Although we found the distinction between structure information and semantic information useful, we agree with the model's developers that in some implementations, that distinction is subjective. Depending on how digital objects are created and managed, significant properties can be imbedded in the structure of an object or managed with additional semantic information, or both. For example, structured text uses mark-up languages and document type definitions to formally describe the meaning associated with various text formatting and layout practices. Well-designed mark-up makes a distinction between the structure of a document and its semantics. In proprietary word processing applications, such a formal distinction may or may not be present. Most contemporary word processing systems include a facility for using style sheets that manage the document structure and its associated semantics. If the stylistic features constitute significant properties of the digital objects, then exporting the object from its native software environment to another platform could eliminate or alter the way the significant properties are represented.

A formal model for significant properties also has many implications for preservation metadata. Our work is an extension to a number of existing metadata models rather than a new approach that requires new concepts, schema, or implementation methods. Both the OCLC/RLG Working Group on Preservation Metadata and the Cedars project recommend that metadata schema to support preservation should include elements that describe the significant properties of digital objects, but neither of these proposals includes a vocabulary for expressing significant properties ([17]). The metadata encoding and transmission standard (METS) provides a framework for exchanging metadata among repositories ([18]). The administrative metadata portion of METS, for example, provides a structure for maintaining technical metadata about the creation, format, and use characteristics of a file or object; intellectual property rights metadata; metadata about the

([16]) Consultative Committee on Space Data Systems, 4-20–4-21.
([17]) Cedars, 'Metadata for digital preservation: The Cedars project outline specification, draft for public consultation', available at: http://www.leeds.ac.uk/cedars/colman/metadata/metadataspec.html; and the OCLC/RLG Working Group on Preservation Metadata, 'A recommendation for content information'.
([18]) Metadata encoding and transmission standard (METS), 2001, available at: www.loc.gov/standards/mets/.

source of the object; and digital provenance for documenting migrations or transformations employed on files. A formal language for expressing significant properties could augment this model by indicating which significant properties were retained which were altered or eliminated in at ingest or through subsequent migrations.

The formal model has at least one other potential application for developing cost models and conducting cost benefit analyses of various digital preservation options. The field of digital preservation in general suffers from the lack of cost models that allow collection managers and decision makers to estimate immediate and long-term costs of digital preservation. We do know, however, that different technical strategies impact long-term preservation costs. A formal means for determining which significant properties are lost or retained using various preservation strategies could assist custodians in selecting strategies and in justifying them to users in the future.

Our work is in a formative stage and it offers only a general framework for expressing significant properties. Further articulation of significant properties is needed for specific types of digital objects and collections for particular designated communities. This work will require input from domain experts who are familiar with the semantics, use practices, and potential users within particular designated communities. We hope that specific communities will populate this model with context-specific cases of significant attributes.

Further work on significant properties also must contend with the role of interactions between digital objects, as they are structured and stored for access, and the user environment. Increasingly, digital objects are stored in some underlying format, but the way the objects appear when they are 'served' or delivered to users depends on a variety of factors in the user's environment. Choices of viewers or browsers as well as user-defined preferences can influence the appearance and behaviour of digital objects. Digital objects may have dynamic formatting effects, where user actions influence the behavior or representation of the content. These affordances raise new challenges for defining what constitutes a digital object and which of its properties are significant to whom.

Present practice and future prospects (with David Wallace) Computer professionals for social responsibility (CPSR) Annual Conference, October 20, 2001; Emulation, migration and long-term preservation of electronic records — ECURE 2001: Preservation and Access for Electronic College and University Records, 12–13 October, 2001; Beyond the byte stream: What to preserve, Society of American Archivists (SAA) annual meeting, 31 August 2001.

Propiedades significativas de los documentos digitales: definiciones, aplicaciones e implicaciones

Margaret Hedstrom, Christopher Lee

La conservación debería implicar la retención de las propiedades más significativas de los objetos digitales desde una generación tecnológica a la próxima. Desafortunadamente, en la actualidad no contamos con ningún medio formal para describir estas propiedades significativas.

El presente documento defiende que las decisiones sobre los métodos de conservación digital deben estar ilustradas por un análisis de las propiedades significativas de los objetos digitales, a fin de conservar las propiedades que justifiquen un uso futuro de los objetos digitales. Presentaremos la investigación inicial sobre los métodos destinados a identificar y expresar formalmente las propiedades significativas. Estos métodos formales contribuirán a una toma de decisiones eficaz sobre los métodos de conservación más adecuados para las distintas clases de objetos digitales que se conservan para comunidades concretas.

La importancia de retener las características originales de los objetos se ha planteado en diversos contextos. Cuando se cambian de formato los objetos físicos utilizando técnicas como la microfilmación, parte de las cualidades físicas de los documentos no son capturadas por la técnica de cambio de formato. Problemas similares pueden ocurrir cuando se con-

vierten los objetos digitales de un formato de fichero a otro, o cuando se normalizan en un formato estándar para su introducción en un depósito de archivos. Las conversiones de formato, por ejemplo, pueden dar lugar a la pérdida de la estructura del documento, de sus rasgos estilísticos y de sus posibilidades de navegación y de reproducción. Se han propuesto varios métodos para abordar este problema, incluido un cuidadoso análisis de los formatos de origen y de destino durante la conversión de formato; el uso de la emulación para conservar el aspecto original de los objetos digitales, y el uso de definiciones de tipo de documento (DTD) y de hojas modelo para producir objetos duraderos que conserven las propiedades significativas de los objetos originales. El modelo de referencia del sistema de información de archivos abiertos (OAIS) utiliza el concepto de informaciones de representación para convertir secuencias de bits en información significativa que retenga la estructura y la semántica del objeto original. Sin embargo, no existe un consenso respecto a qué propiedades de los objetos digitales o qué tipos de objetos digitales constituyen cualidades estructurales y semánticas significativas que deban retenerse.

En este documento presentaremos la investigación inicial realizada como parte del proyecto Camileon. Estamos desarrollando un método para identificar las propiedades significativas de distintas clases de objetos digitales. Nuestro método implica la evaluación de orientaciones sobre las propiedades significativas en la documentación existente, un análisis de las normas para representar las cualidades estilísticas y estructurales de la información digital, el desarrollo de un inventario de propiedades potencialmente significativas, y la agrupación de propiedades significativas en categorías basadas en tipos de objetos, así como el objeto de su conservación. Presentaremos un modelo formal para representar propiedades significativas, como un esquema XML que define cada una de las propiedades significativas y, en su caso, incluye los valores o los calificadores que indican el grado de importancia de cada propiedad o la variabilidad aceptada para una comunidad concreta. Por último, informaremos sobre los resultados de una prueba del sistema con un pequeño grupo de objetos digitales que se están conservando para una comunidad específica.

En conclusión, discutiremos cómo pueden aplicarse a la conservación digital los métodos formales para expresar las propiedades significativas de los objetos digitales. Los resultados de esta investigación son prometedores para varias aplicaciones en el ámbito de la conservación digital. Los archiveros pueden utilizar el modelo formal para elegir entre la emulación, la migración, la normalización u otro método de conservación basado en el análisis de las cualidades significativas que deben conservarse para que la información siga siendo útil para una comunidad concreta. El método también puede apoyar el análisis de coste/beneficio cuando las decisiones sobre el grado de conservación se basen en un análisis de los costes y beneficios de conservar la totalidad o solamente cualidades seleccionadas de los objetos digitales. Las definiciones formales de las cualidades significativas también pueden utilizarse para formular requisitos específicos del formato, la estructura y los metadatos que acompañan a los objetos digitales como parte de los acuerdos entre los productores y el depósito digital y para las descripciones normalizadas de los contenidos. Por último, las definiciones y expresiones formales de las propiedades significativas pueden utilizarse para explicar a los futuros usuarios qué propiedades significativas se conservaron y para justificar y documentar las decisiones de conservación.

Wesentliche Merkmale digitaler Dokumente: Definitionen, Anwendungen, Implikationen

Margaret Hedstrom, Christopher Lee

Bei der Archivierung sollten die wichtigsten Merkmale digitaler Objekte von einer Technologiegeneration zur nächsten bewahrt werden. Leider verfügen wir gegenwärtig über keinerlei formelle Hilfsmittel zur Beschreibung dieser wesentlichen Merkmale.

In diesem Beitrag vertreten wir die Auffassung, dass Entscheidungen über Methoden der digitalen Archivierung eine Analyse der wesentlichen Merkmale digitaler Objekte vorausgehen sollte, damit die Merkmale, die eine künftige Nutzung digitaler Objekte unterstützen, bewahrt werden. Wir werden erste Untersuchungen zu Methoden der Identifizierung und formellen Darstellung wesentlicher Merkmale vorstellen. Formelle Methoden zur Identifizierung und Darstellung wesentlicher Merkmale werden dazu beitragen, dass effektive Entscheidungen darüber getroffen werden können, welche Archivierungsmethoden für die einzelnen Kategorien digitaler Objekte, die für bestimmte Nutzerkreise konserviert werden, am geeignetsten sind.

Die Bedeutung der Bewahrung der ursprünglichen Eigenschaften von Artefakten ist in zahlreichen Zusammenhängen zur Sprache gekommen. Bei der Umformatierung physischer Objekte mittels Verfahren wie Mikroverfilmung werden einige der physischen Attribute von Dokumenten von der Umformatierungstechnik nicht erfasst. Ähnliche Probleme können auftreten, wenn digitale Objekte von einem Dateiformat in ein anderes umgewandelt werden oder wenn sie bei Eingang in eine Ablage nicht in ein Standardformat gebracht werden. So kann es z. B. bei Formatkonvertierungen zum Verlust von Dokumentenstruktur, stilistischen Merkmalen, Navigationsmöglichkeiten und Rendering-Funktionen kommen. Zur Lösung dieses Problems sind verschiedene Methoden vorgeschlagen worden, u. a. eine sorgfältige Analyse von Ausgangs- und Zielformaten während der Formatkonvertierung, die Anwendung von Emulation zur Erhaltung des ursprünglichen „Erscheinungsbilds“ digitaler Objekte und die Verwendung von Strukturbeschreibungen von Dokumenten (DTD) und Formatvorlagen zur Erzeugung von beständigen Objekten, bei denen die wesentlichen Merkmale der ursprünglichen Objekte erhalten sind. Beim Referenzmodell für ein offenes archivisches Informationssystem (Open Archival Information System Reference Model, OAIS) wird das Konzept zur Darstellung von Informationen verwendet, um Bitfolgen in bedeutungstragende Informationen umzuwandeln, bei denen die Struktur und Semantik des Originalobjekts erhalten bleiben. Es besteht jedoch keine Einigkeit darüber, welche Merkmale von digitalen Objekten oder von Klassen digitaler Objekte wesentliche strukturelle und semantische Attribute darstellen, die zu erhalten sind.

In unserem Beitrag werden wir erste Untersuchungen vorstellen, die im Rahmen des Camileon-Projekts durchgeführt wurden. Wir sind dabei, ein Verfahren zu entwickeln, mit dem sich die wesentlichen Merkmale von verschiedenen Klassen digitaler Objekte identifizieren lassen. Unser Verfahren umfasst eine bewertende Einführung zu den wesentlichen Merkmalen in der Fachliteratur, eine Analyse von Standards zur Darstellung der stilistischen und strukturellen Attribute digitaler Informationen, die Aufstellung eines Inventars von potentiell wesentlichen Merkmalen sowie die Einordnung wesentlicher Merkmale in Kategorien je nach Objektart und dem geplanten Zweck der Archivierung. Wir werden ein formelles Modell für die Darstellung wichtiger Merkmale in einem XML-Schema vorstellen, das jede wichtige Eigenschaft definiert und – gegebenenfalls – Werte oder Kennzeichner beinhaltet, die den Grad der Bedeutung der jeweiligen Eigenschaft oder akzeptierten Veränderlichkeit für eine bestimmte Nutzergruppe angeben. Nicht zuletzt werden wir auch über die Ergebnisse einer Erprobung des Schemas mit einer kleinen Gruppe digitaler Objekte berichten, die für einen speziellen Nutzerkreis archiviert werden.

Zum Abschluss werden wir erörtern, wie formelle Verfahren zur Darstellung der wesentlichen Merkmale digitaler Objekte auf die digitale Archivierung angewandt werden können. Die Ergebnisse dieser Untersuchung sind für eine Reihe von Anwendungsmöglichkeiten bei der digitalen Archivierung vielversprechend. Archivare können das formelle Modell verwenden, um Emulation, Migration, Standardisierung oder ein anderes Archivierungsverfahren ausgehend davon zu wählen, welche wesentlichen Attribute bewahrt werden müssen, wenn die Informationen für eine bestimmte Gruppe von Nutzen bleiben soll. Das Verfahren kann auch bei Kosten-Nutzen-Analysen zur Anwendung kommen, wo Entscheidungen über den Umfang der Archivierung nach einer Abwägung der Kosten und des Nutzens einer Bewahrung aller oder nur ausgewählter Attribute digitaler Objekte getroffen werden. Formale Definitionen wesentlicher Attribute können auch dazu dienen, spezielle Anforderungen für Format, Struktur und Metadaten zu formulieren, die digitale Objekte nach Maßgabe von Übergabevereinbarungen zwischen Urhebern und digitaler Ablage und zur Vereinheitlichung von Inhaltsbeschreibungen begleiten müssen. Nicht zuletzt können die formalen Definitionen und Ausdrücke wesentlicher Merkmale genutzt werden, um künftigen Nutzern zu erklären, welche wesentlichen Merkmale archiviert wurden, und um die Archivierungsentscheidungen zu begründen.

Caractéristiques significatives des documents numériques: définitions, applications et implications

Margaret Hedstrom, Christopher Lee

La conservation implique de préserver les caractéristiques significatives des objets numériques d'une génération technologique à l'autre. Malheureusement, nous ne disposons pas actuellement de moyens formels pour décrire ces caractéristiques significatives. Dans cet exposé, nous avançons l'idée que toute décision sur des méthodes de conservation numérique doit être précédée d'une analyse des caractéristiques significatives des objets numériques, de façon à préserver les caractéristiques qui permettront l'utilisation future de ces objets. Nous présenterons les premières études sur les méthodes d'identification et de description formelles de ces caractéristiques. Ces méthodes formelles d'identification et de description contribueront à des décisions efficaces sur le choix des méthodes de conservation les plus propres à différentes catégories d'objets numériques qui sont conservés pour des communautés particulières.

On a souligné à maintes reprises l'importance de préserver les caractéristiques originales des objets créés par l'homme. Lorsque des objets matériels sont reformatés à l'aide de techniques comme le microfilmage, il arrive que la technique choisie ne retienne pas certains des attributs physiques des documents. Il en va de même lorsque des objets numériques sont convertis d'un format de fichier à un autre, ou lorsqu'ils sont uniformisés dans un format type au moment de leur versement dans un dépôt d'archives. Ainsi, les changements de format peuvent occasionner une perte de la structure du document, de ses caractéristiques de style et de ses possibilités de navigation et de rendu. Diverses méthodes ont été proposées pour remédier à ce problème, parmi lesquelles une analyse minutieuse des formats source et cible lors de la conversion, le recours à l'émulation pour conserver l'aspect et le caractère original des objets numériques et, enfin, l'emploi de définitions types de documents et de feuilles de style pour élaborer des objets permanents, préservant les caractéristiques significatives des objets originaux. Le modèle de référence pour un système ouvert d'informations d'archives (OAIS) applique la notion d'informations de représentation pour convertir des séquences binaires en informations significatives qui conservent la structure et

la sémantique de l'objet original. Il n'en demeure pas moins qu'aucun consensus n'existe sur les caractéristiques des objets numériques ou des catégories d'objets qui constituent des attributs structurels et sémantiques significatifs devant nécessairement être conservés.

Nous présenterons dans notre exposé les premières études actuellement menées dans le cadre du projet Camileon. Nous développerons une méthode d'identification des caractéristiques significatives des diverses catégories d'objets numériques. Notre méthode consiste à évaluer les orientations données à ce sujet par les auteurs, à analyser les normes de représentation des attributs stylistiques et structurels de l'information numérique, à élaborer une liste des caractéristiques significatives possibles et, enfin, à regrouper celles-ci dans des catégories définies selon le type d'objet et la finalité de sa conservation. Nous présenterons un modèle formel de représentation des caractéristiques significatives sous la forme d'un schéma XML, qui définit chacune de ces caractéristiques et, s'il y a lieu, incorpore des valeurs ou des indicateurs permettant d'évaluer le degré d'importance de chaque caractéristique ou sa variabilité, telle qu'admise par une communauté désignée. Enfin, nous nous ferons l'écho des résultats de test du schéma avec une série limitée d'objets numériques conservés pour une communauté particulière.

Pour conclure, nous discuterons de la manière dont des méthodes formelles de description des caractéristiques significatives des objets numériques peuvent être appliquées à la conservation numérique. Les résultats de ces travaux laissent espérer un certain nombre d'applications dans ce domaine. Les archivistes pourront utiliser ce modèle formel pour choisir l'émulation, le transfert, la normalisation ou n'importe quelle autre méthode de conservation, à partir d'une analyse des attributs significatifs devant être conservés si l'information doit demeurer utile à une communauté désignée. La méthode peut également s'appuyer sur une analyse coûts/avantages qui permet de fonder les choix sur le degré de conservation à partir d'une analyse des coûts et des avantages de la conservation de tout ou partie des attributs des objets numériques. Les définitions formelles des attributs significatifs peuvent également être appliquées pour formuler les spécifications requises du format, de la structure et des métadonnées qui doivent accompagner les objets numériques, en tant que partie intégrante des conventions de versement passées entre les producteurs de documents et un dépôt d'archives numériques, ainsi que pour uniformiser les descriptions de contenus. Enfin, les définitions et les descriptions formelles des caractéristiques significatives peuvent servir à expliquer aux utilisateurs futurs celles de ces caractéristiques qui ont été conservées et à justifier les choix de conservation documentaire.

Ulrich Thiel

Dr Ulrich Thiel is a senior researcher at Fraunhofer IPSI, the Integrated Publication and Information Systems Institute in Darmstadt, Germany. The goal of Fraunhofer IPSI is to develop concepts, foundations and systems for the next generation of distributed and cooperative multimedia information systems. Dr Thiel received his diploma in computer science from the University of Dortmund, and his PhD in Information Science from the University of Konstanz. Until 1988 he was a researcher and lecturer at the Information Science department at the University of Konstanz. Since 1990 he has been a researcher and manager of several projects and research groups within IPSI. His primary research interests are in intelligent multimedia information retrieval, logic-based retrieval mechanisms, intelligent user interfaces, and dialogue and user modeling. He has been coordinator of several European R & D projects (currently Cogito and Collate), and is programme committee member of several national and international conferences and workshops. Holger Brocks, Andrea Dirsch-Weigand and Adelheit Stein also work at the Fraunhofer Integrated Information and Publication Systems Institute in Darmstadt, Germany. Jürgen Keiper works at the Deutsches Filminstitut (DIF) in Frankfurt am Main, Germany.

A collaborative archive supporting research on European historic films — the Collate project

Ulrich Thiel, Holger Brocks, Andrea Dirsch-Weigand, Jürgen Keiper, Adelheit Stein

Introduction

Various web-based collaboratories have been employed since the early 1990s mainly in the natural sciences, but we find so far only few similar efforts in art and humanities and only systems with very limited functionality. Whereas preservation and organisation of historical knowledge may be comparable, work processes in interpreting sciences are somewhat different and should be supported by appropriate technologies. Especially the process of compiling arguments, counter-arguments, examples, definitions, and references to historical source material — which is the prevailing method in the humanities — may profit from an electronic environment that improves the capacity and reach of the individual knowledge worker.

Another aspect of collaboratories is their capability to enable virtual teams to work together almost as if they were at the same location. Although many (informal) contacts between cultural archives constitute specific professional communities, they so far lack effective technology support for collaborative knowledge working. The World Wide Web can serve both as communication platform for such communities and as gateway for document-centred work in such digital libraries and archives.

The EU-funded project 'Collate — collaboratory for annotation, indexing and retrieval of digitised historical archive material' (http://www.collate.de) started in Autumn 2000 (IST-1999-20882) and runs for three years. An international team of content providers, film-domain experts and technology providers work together to develop a new type of collaboratory in the domain of cultural heritage. The implemented system offers access to a digital repository of historic text archive material documenting film censorship practices for several thousands of European films from the 1920s and 1930s. For a subset of significant films it provides enriched context documentation including selected press articles, film advertising material, digitised photos and some film fragments. Major film archives from Germany, Austria and the Czech Republic provide the sources and work as pilot users with the Collate system.

Collate's goals and approach

Designed as a content and context-based knowledge working environment for distributed user groups, the Collate system supports both individual work and collaboration of domain experts who are analysing, evaluating, indexing and annotating the material in the data repository. The example application focuses on historic film documentation but the developed tools are designed to be generic and as such adaptable to other content domains and application types.

The system provides appropriate task-based interfaces for indexing/annotation and collaborative activities such as preparing a joint multimedia publication or assembling and creating material for a (virtual) exhibition, contributing unpublished parts of their work in the form of extended annotations and commentaries. Appropriate knowledge-management tools, e.g. indexing aids and domain-specific controlled vocabularies, have been developed jointly by the system developers and the film domain experts, thus exploiting the benefits of a participatory design. Using the tools for manual cataloguing and indexing, the users create a growing body of metadata with a special focus on subject indexing and content-based annotations of documents. Automatic indexing of textual and pictorial parts of a document can be invoked to receive suggestions for index terms from the system. In addition, the users can rely on the support from automatic lay-out analysis for scanned documents which allows her to annotate individual segments. Annotations are a central concept in Collate. As a multifunctional means of in-depth analysis, annotations can be made individually but also collaboratively, for example in the form of annotation of annotations, collaborative evaluation and comparison of documents.

The system exploits the user-generated metadata and annotations by advanced XML-based content management and retrieval methods. The final version of the online collaboratory will integrate cutting-edge document pre-processing and management facilities, e.g., XML-based document handling and semi-automatic segmentation, categorisation and indexing of digitised text documents and pictorial material. Combining results from the manual and automatic indexing procedures, elaborate content and context-based information retrieval mechanisms can be applied.

Collate system features

Collaboration support in the Collate working environment makes use of some contemporary groupware products and additional system functions based on an explicit model of collaborative indexing and annotation. Through interrelated free-text annotations users can enter a — direct or indirect — discourse on the interpretation of documents and document passages, e.g. adding information, interpretations, arguments, etc. Possible relations between annotations can be predefined or inferred by the system in order to represent the discourse structure. Additionally, explicit communication about the interpretation of contents and interrelation of annotation are supported by an in-built discussion forum, and in the final system version by an intelligent dialogue/collaboration manager.

The Collate collaboratory is a multi-functional software package integrating a large variety of functionalities, which are realised by cooperating software modules. It comprises several databases and document representation schemata. XML is used as the uniform internal representation language for the documents in the repository and the associated metadata as well as for the implementation of the communication protocol among its system modules. An XML-based content manager is responsible for the integration of knowledge processing methodology and retrieval functionality in the system.

Main modules of the Collate system architecture are:

- three document pre-processing modules for digital watermarking of the documents (copyright and integrity watermarks), intelligent, automatic document structure analysis and classification, and automatic, concept-based picture indexing and retrieval;
- a distributed multimedia data repository comprising digitised text material, pictorial material like photos and posters, and digital video fragments;
- tools for the representation and management of the metadata, the XML-based content manager incorporating an ontology manager and a retrieval engine;
- a collaborative task manager for complex individual and collaborative tasks, such as indexing, annotation, comparison, interlinking and information retrieval, including tools for online communication and collaborative discourse between the domain experts and other system users;
- the web-based user interface of Collate comprises several workspaces for different tasks performed by distributed user groups and user types allowing for different access rights and offered interface functions. The final system version will be generated semi-automatically by exploiting knowledge from the underlying task model and the user-specific dialogue history.

Collaboration in the Collate environment

In Collate, we go beyond the mere replication of traditional domain-specific workflows by providing a comprehensive model of the various Collate domain objects and their potential interrelations. Our notion of task-guided collaboration includes the recognition of structures as well as relations between different types of annotations. By taking the users' roles, tasks and goals into account we aim to provide comprehensive support for the various levels of indexing.

The Collate system supports asynchronous collaboration in indexing for non-technical users. In our understanding the domain objects (scanned documents, metadata) represent the main focus of collaborative work, i.e. collaboration is performed through annotating the digitized artefacts or their associated metadata objects.

The interrelations between the various domain objects can either be unspecified or it can be modeled in a more explicit way by defining specific types of admissible relations. In addition, certain communicative acts on the meta-level (e.g., request for clarification) are part of the Collate collaboration model.

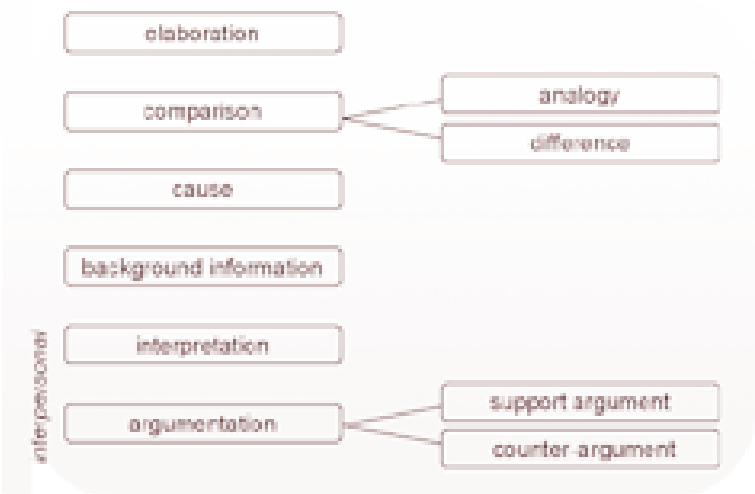
In addition, the users' individual tasks and goals have to be taken into account for modeling a collaborative system. Content-based indexing of a specific document — in this sense — can be considered as a global task, which can be decomposed into partial tasks. In the Collate context, the result of these partial tasks, which are to be performed by various users, is the value-added information in form of metadata objects associated with the original document. But these partial tasks are only rarely performed in isolation. On the contrary, in most cases a specific annotation will be part of a thematic thread, e.g., some newsgroup-like discussion about a certain topic. Digital signatures are employed to ensure authenticity of the individual contributions as well as the chronological order of the annotations.

But in order to represent properly scientific discourses, especially in the arts and humanities, annotations have to consist of more than unstructured and uncontrolled text which comments on another domain object (binary image file, XML metadata object). For this reason we have devised a comprehensive model of discourse structure relations between (a) binary image versions of the original document and annotations and (b) discussion threads realised as annotations on annotations.

Our document-centered discourse model is loosely based on the theory of discourse structure relations. Even though it has been developed for monologues in the linguistic context of text coherence we think that discourse structure relations can also be adapted to describe admissible relations between various data and metadata objects in the Collate context, especially annotations or comments.

In particular we employ a specific subset of relations ranging from factual to more interpersonal levels, i.e. focusing on certain qualities of the participants of a discourse. Figure 1 illustrates the discourse structure relations as used in the Collate system.

Figure 1: Discourse structure relations



In the following we just briefly paraphrase the discourse structure relations used in Collate:

- Elaboration — Providing additional, more detailed information (e.g., '… it's Paris in the United States, and not in France …').
- Comparison — Comparative relations can be further sub-structured to emphasise semantic similarities or contrasts between two elements of a discourse.
- Cause — To state a specific cause for a certain circumstance.

- Background information — Using information about the background of the author of the other annotation (e.g., '… As a lay-person the author does not take psychological aspects into account …').
- Interpretation — (Subjective) interpretation of a statement being referred to (e.g. '… the author actually means …').
- Argumentation — The statement or argument of the other author is either supported, or a counterargument/antithesis is formulated here.

The seamless transition from factual to interpersonal discourse structure relations depicted in Figure 1 also corresponds to the illocutionary aspects of an annotation, i.e. the specific communicative intention its author had in mind at the time of creation (e.g., from stating factual information towards active participation in a discussion thread).

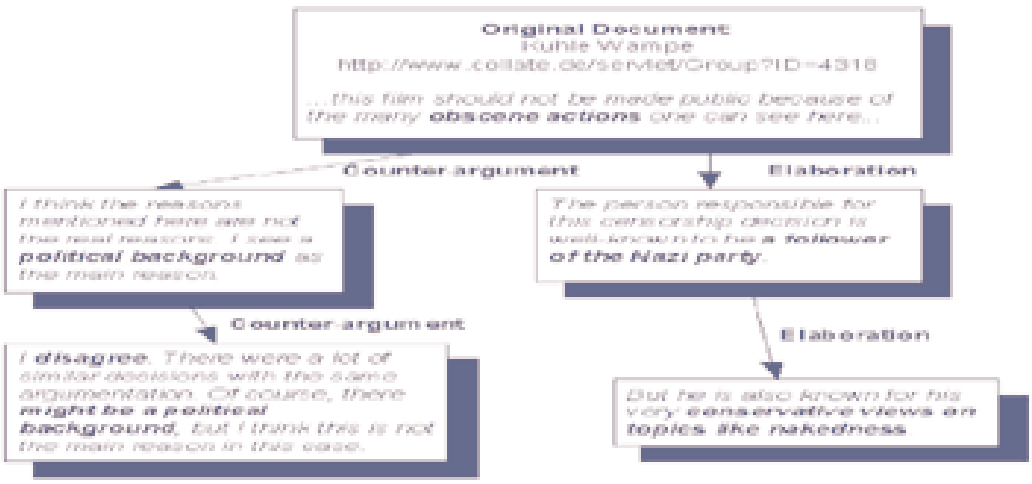


Figure 2: Example discourse

Figure 2 displays a fictious example discourse about the partial ban of the movie 'Kuhle Wampe' by Berthold Brecht.

Even though discourse structure relations proved adequate for modeling the interrelations between annotations it turned out, however, that there are some relevant pragmatic aspects of collaborative indexing work, which are not yet covered. In the next section we describe how discourse structure relations can be complemented by communicative acts to introduce meta-communication, i.e. explicit communication about domain objects, in a seamless way.

Conceptually, discourse structure relations and communicative acts can be considered as complementary: communicative acts focus on illocutionary aspects of a specific dialogue situation, whereas discourse structure relations describe characteristic relationships between assertive acts, e.g. annotations or comments.

But on closer inspection it becomes evident that some communicative acts might invoke certain types of discourse structure relations between the corresponding annotations. In our view, the set of discourse structure relations adopted for Collate can be considered as specific instances of comments, i.e. they are treated as assertive communicative acts.

From this perspective, we can regard explicit collaboration in the context of the Collate project as the combination of specified relation types between annotations, i.e. discourse structure relations, which are complemented by a certain set of admissible COR acts for meta-communication (on the dialogue level) referring to the various types of Collate domain objects (e.g., annotations, cataloguing information).

At its current stage of development, the Collate prototype system already supports various tasks like cataloguing, indexing (structured, free) and free comments. Furthermore, it offers simple search options based on filmographic information as well as on transient information recorded as part of the scanning process:

Context-based retrieval of documents

Appropriate search and retrieval functionality represents a fundamental requirement for enabling the user community to access a cultural digital library in a reasonable way. To allow for advanced content- and context-based search the documents in the digital collection must be indexed by content and subject matter.

In context-based retrieval, a document does not stand for its own, but also the actual context of this document is considered. For the Collate case, this means that we are dealing with the discourse context. The RDF statements, which are used to interrelate the various domain objects, are typed according to the discourse structure relation they represent. With this information, we then know the specific type of an annotation with respect to its context, e.g. an elaboration or an example.

Having a second look at Figure 2 one can see that only the inclusion of the associated annotations would yield it as relevant for a query like 'political background'. But it also becomes evident that the annotations within a certain discourse context cannot be treated in isolation, e.g., the second counterargument weakens the statement it comments upon in this context.

The introduction of discourse structure relations allows for novel retrieval options with respect to the discourse context. They can be used to create a ranking of relevant documents according to the users' queries. Depending on the specific type of its connecting relation, an annotation can possibly raise or lower the overall relevance weight of its discourse context.

Conclusion

The Collate system represents a new type of collaboratory supporting content and concept based work with digital document sources. Innovative task-based interfaces support professional domain experts in their individual and collaborative scholarly work, i.e. analysing, interpreting, indexing and annotating the sources. The hereby provided metadata are managed by an advanced XML-based content manager and an intelligent content and context based retrieval system.

Un sistema de archivos electrónicos colectivo para la búsqueda de películas históricas

Ulrich Thiel, Holger Brocks, Andrea Dirsch-Weigand, Jürgen Keiper, Adelheit Stein

Los archivos y las bibliotecas ocultan una amplia fuente de textos de referencias y de imágenes históricas importantes y únicas que se destruyen o cuyo contenido no se explota. El sistema Collate, fruto de un proyecto de investigación y desarrollo apoyado por la Comisión Europea, tiene por objeto crear un medio de trabajo y de conocimiento virtual en la red en el que científicos y neófitos interesados puedan trabajar en colaboración sobre documentos digitales, con el fin de facilitar el acceso a todo este material del patrimonio cultural y su explotación. La documentación de películas históricas de los años veinte y treinta, en particular con textos digitalizados y también con fotografías, carteles de películas y secuencias de vídeo, proporciona un posible ejemplo de explotación.

El sistema Collate ofrece a este respecto un medio de trabajo cómodo, que permite un trabajo documental y la cooperación entre distintas categorías de usuarios distribuidos. Técnicas específicas de gestión de los conocimientos, como los métodos de consulta por contenidos, las comparaciones de estructuras, la explotación vertical *(Tiefenerschließung)* y las anotaciones dinámicas, ayudan al usuario en su trabajo sobre documentos digitales. Como medio de conocimiento interactivo, Collate permite el acceso en red a documentos culturales e históricos. Los investigadores desempeñan un papel activo en el desarrollo del sistema, dado que enriquecen los documentos con sus anotaciones, es decir, sus observaciones y comentarios sucesivos, y que crean índices. Como método polivalente de explotación vertical, las anotaciones se suelen realizar en equipo; de ahí deriva que el sistema permita la explotación colectiva de documentos digitales. Además, al recurrir a filigranas digitales, a análisis de estructura de textos y a clasificaciones de documentos (semi)automáticas, el sistema Collate aplica principios importantes de gestión de los conocimientos, en particular la autenticidad y la clasificación temática. La aplicación técnica se efectúa utilizando archivos con la norma XML para los metadatos, las anotaciones y los modelos. Además, el lenguaje XML constituye la base de una integración flexible de los métodos de gestión de los conocimientos y funciones de investigación documental en el sistema.

El sistema Collate reviste otro aspecto esencial, en el sentido de que fomenta la aceptación de un trabajo colectivo en el ámbito de la historia. La acogida que tenga tal sistema dependerá de forma determinante de su capacidad para facilitar los procesos complejos de la documentación filmográfica a los investigadores. A tal efecto, se elaboran modelos de interfaces de usuario específicas para una tarea con vistas a la explotación del contenido y las anotaciones. Estos procesos de trabajo constituyen a su vez el punto de partida de una evaluación empírica del sistema.

En resumen, el sistema Collate debería permitir concebir, aplicar y evaluar una colaboración en la red, destinada a ayudar al usuario final en su trabajo de anotación, creación de índices e investigación en archivos históricos, en diferentes soportes y formatos de datos, gracias a una interfaz de usuario cómoda. El grueso del trabajo que debe realizarse consiste en la descripción de un modelo de proceso de trabajo para generar interfaces de usuario específicas de una tarea, es decir, que se procurará describir la manera en que el sistema tiene en cuenta los distintos tipos de usuarios y sus prácticas de trabajo específicas, generando interfaces adaptadas a los distintos criterios de tareas.

Ein kollaboratives elektronisches Archivsystem für die historische Filmforschung

Ulrich Thiel, Holger Brocks, Andrea Dirsch-Weigand, Jürgen Keiper, Adelheit Stein

Viele wichtige und einzigartige historische Quellentexte und Bilddokumente lagern zerstreut und inhaltlich unerschlossen in Archiven und Bibliotheken. Um den Zugriff auf diese Materialien des Kulturerbes und ihre inhaltliche Erschließung zu ermöglichen, zielt Collate, ein von der Europäischen Kommission gefördertes Forschungs- und Entwicklungsprojekt, darauf ab, eine virtuelle Arbeits- und Wissensumgebung im WWW zu schaffen, in der Wissenschaftler und interessierte Laien kooperativ mit digitalisierten Dokumenten arbeiten können. Als exemplarisches Fallbeispiel dient die Dokumentation historischer Filme aus den 20er und 30er Jahren, z. B. durch digitalisierte Textdokumente wie auch Fotos, Filmplakate und Videofragmente.

Das Collate-System stellt in diesem Kontext eine komfortable Arbeitsumgebung dar, welche ein dokumentzentriertes Arbeiten und die gemeinschaftliche Arbeit verschiedener, verteilter Nutzergruppen realisiert. Dedizierte Techniken des Wissensmanagements wie inhaltsbezogene Zugriffsmethoden, Strukturvergleiche, Tiefenerschließung und dynamische Annotationen unterstützen die Nutzer bei ihrer Arbeit mit den digitalen Dokumenten. Collate als interaktive Wissensumgebung ermöglicht somit den vernetzten Zugang zu kulturhistorischen Dokumenten. Die Benutzer sind aktiv in die Systementwicklung eingebunden, da sie die Dokumente durch ihre Annotationen, also Anmerkungen und Kommentare, sukzessive anreichern und indexieren. Die Annotation als multifunktionale Methode der Tiefenerschließung findet häufig im Teamwork statt, infolgedessen wird das System die kollaborative Erschließung von digitalen Dokumenten ermöglichen. Des Weiteren implementiert Collate mit digitalen Wasserzeichen, (halb-)automatischen Textstrukturanalysen und Dokumentklassifikationen wichtige Standards des Wissensmanagements wie Authentizität und thematische Einordnung. Die technische Umsetzung erfolgt unter Verwendung eines XML-basierten Repositories für Metadaten, Annotationen und Templates. XML bildet ferner die Grundlage für eine flexible Integration der Methoden des Wissensmanagements und der Retrievalfunktionalität in das System.

Ein weiterer zentraler Aspekt von Collate ist die Förderung der Akzeptanz eines Kollaboratiums für den kulturhistorischen Wissenschaftsbereich. Die Akzeptanz eines derartigen Systems hängt entscheidend davon ab, ob es gelingt, die komplexen Arbeitsabläufe der Filmdokumentation für die Benutzer zu erleichtern. Zu diesem Zweck werden Modelle für aufgabenorientierte Benutzerschnittstellen für die inhaltliche Tiefenerschließung und Annotation erarbeitet. Diese realistischen Arbeitsabläufe wiederum stellen den Ausgangspunkt für eine empirische Evaluation des Systems dar.

Insgesamt soll mit Collate ein WWW-basiertes Kollaboratium entworfen, implementiert und evaluiert werden, welches den Endnutzer bei der Annotation, Indexierung und dem Retrieval von historischem Archivmaterial verschiedener Medien- und Dateiformate mit einer komfortablen Benutzeroberfläche unterstützt.

Der Fokus der einzureichenden Arbeit wird in der Beschreibung eines Arbeitsprozessmodells für die Generierung von aufgabenorientierten Benutzeroberflächen liegen, d. h., es wird beschrieben, wie das System den verschiedenen Nutzertypen mit ihren spezifischen Arbeitsgewohnheiten durch eine flexible Generierung von entsprechenden Interfaces für die unterschiedlichen Anforderungen und Aufgaben Rechnung trägt.

Un système d'archives électroniques collectif pour la recherche de films historiques

Ulrich Thiel, coauteurs: Holger Brocks, Andrea Dirsch-Weigand, Jürgen Keiper, Adelheit Stein

Les archives et les bibliothèques recèlent un vaste gisement de textes de référence et d'images historiques importants et uniques, qui sont détruits ou dont le contenu est inexploité. Le système Collate, fruit d'un projet de recherche et de développement soutenu par la Commission européenne, vise à créer un environnement de travail et de connaissance virtuel sur le web dans lequel scientifiques et néophytes intéressés peuvent travailler en collaboration sur des documents numériques, afin de faciliter l'accès à tout ce matériel du patrimoine culturel et à son exploitation. La documentation de films historiques des années 20 et 30, notamment par des textes numérisés, mais aussi par des photographies, des affiches de films et des séquences vidéo, fournit un exemple d'exploitation possible.

Le système Collate offre à cet égard un environnement de travail confortable, qui permet un travail documentaire et la coopération entre diverses catégories d'utilisateurs répartis. Des techniques dédiées de la gestion des connaissances, telles que les méthodes de consultation par le contenu, les comparaisons de structures, l'exploitation verticale *(Tiefenerschließung)* et les annotations dynamiques, aident l'utilisateur dans son travail sur des documents numériques. En tant qu'environnement de connaissance interactif, Collate permet ainsi l'accès en réseau à des documents culturels et historiques. Les chercheurs jouent un rôle actif dans le développement du système, étant donné qu'ils enrichissent les documents par leurs annotations, c'est-à-dire par leurs remarques et commentaires successifs, et qu'ils créent des index. En tant que méthode polyvalente d'exploitation verticale, l'annotation est réalisée souvent en équipe, d'où il découle que le système permet l'exploitation collective de documents numériques. En outre, en ayant recours à des filigranes numériques, à des analyses de structure de texte et à des classifications de documents (semi-) automatiques, le système Collate applique des principes importants de la gestion des connaissances, notamment l'authenticité et le classement thématique. La mise en œuvre technique s'effectue en utilisant des archives à la norme XML pour les métadonnées, les annotations et les modèles. En outre, le langage XML constitue la base d'une intégration en souplesse des méthodes de la gestion des connaissances et des fonctionnalités de recherche documentaire dans le système.

Le système Collate revêt un autre aspect majeur, en ce sens qu'il encourage l'acceptation d'un travail collectif dans le domaine de l'histoire. L'accueil réservé à un tel système dépend de façon déterminante de sa capacité à faciliter les processus complexes de la documentation filmographique aux chercheurs. À cet effet, des modèles d'interfaces utilisateur spécifiques à une tâche sont élaborés pour l'exploitation du contenu et les annotations. Ces processus de travail constituent à leur tour le point de départ d'une évaluation empirique du système.

En somme, le système Collate devrait permettre de concevoir, de mettre en œuvre et d'évaluer une collaboration sur le web, qui est destinée à aider l'utilisateur final dans son travail d'annotation, de création d'index et de recherche dans des archives historiques sur des supports et dans des formats de données différents, grâce à une interface utilisateur confortable. L'essentiel du travail à fournir réside dans la description d'un modèle de processus de travail pour générer des interfaces utilisateur spécifiques à une tâche, c'est-à-dire que l'on s'attachera à décrire la manière dont le système prend en compte les différents types d'utilisateurs et leurs habitudes de travail particulières en générant des interfaces adaptées aux différents critères de tâche.

Parallel session 4

Wednesday, 8 May 2002

ORGANISING RECORDS AND ARCHIVES
Metadata, standardisation and Model Requirements (MoReq)

Chairperson:	Hartmut Weber (Germany)
Co-chair:	Paul E. Murphy (European Commission)
Rapporteur:	George Mackenzie (United Kingdom)

The electronic records management in Catalonia: the current state of affairs

Ferran Agelet, Lluís-Esteve Casellas, Lluís Cermeno, Betlem Martínez, Remei Perpinyà

Ferran Agelet

Ferran Agelet has a degree in psychology and a Master in archivist studies, he graduated in management and cultural politics, and also in interactive systems. He works at the Barcelona Municipality Archives in information projects, in programmes for records management, in document digitalisation, quality planning and in European projects. He is professor of programmatic and author's systems at the Open University of Catalunya.

Lluís-Esteve Casellas

Lluís-Esteve Casellas is archivist at the Town Council of Girona, where he is responsible for implantation and development of the records management system. He has worked as a private consultant for several public administrations on records management, and he published many works on those subjects and he participated on many formative and propagation meetings. He is a member of the National Commission on Records Access and Appraisal and Selection (CNAATD) of Catalunya, and coordinator of the Working Group for the Local Economic Records.

The objective of this paper is to discuss the electronic records management in Catalonia, that is; who is responsible for their management, how are they managed and what role archivists perform in the implementation of systems which make this possible. With this aim, the work is based on the study of experiences in several public institutions and private bodies in Catalonia ([1]). In selecting the sample, we have looked at those bodies, which because of their importance or because of the level of interest of their experiences, allow us to understand the state of affairs in Catalonia, both from a point of view of the identification of projects already up and running, and in order to gain an overall vision of the policies and practices observed in the electronic records management.

From the archival point of view we understand the electronic records management as an activity which encompasses the entire life cycle of records: from the design stage of the formats of these records, their creation as authentic, complete and reliable records as testimonies of a specific procedure, to the maintenance of requirements which guarantee their evidential value in the future. For this reason, in the study of each experience special emphasis has been given to fields in which the electronic records management is applied: production of records, administrative management and preservation, with the objective of checking to what extent the criteria and archival methodology has been taken into account throughout the process.

1. Context

The analysis of the context takes as an essential starting point the objectives defined in the documents which represent, at European level, the conceptual basis of what has been called the Information Society. In this sense, the report by the Vice-president of the European Union, Martin Bangemann, in the 1994 *Cities and regions info way to Europe*, better known as the Bangemann Report, noted 10 key areas of action for its development which, directly or not, involved the electronic circulation of records. The following year, the G7 coined the term 'online government', stressing the firm commitment of the public administrations to the information society. At a later date, several initiatives have been aimed at promoting the implementation of information and communication technologies in European society, such as for example the *e*Europe 2002 action plan or the different framework programmes.

At the level of the Spanish State, the *Libro Blanco para la mejora de los Servicios Públicos* ([2]) follows in the same line as the above European reference points. This is supported by the Spanish Plan de Acción INFO XXI ([3]), equivalent to *e*Europe, which has implemented existing projects and developed new ones, such as Ventanilla Única (single window) ([4]), with the interconnection of records registers of the different public administrations, PISTA (Promoción e Identificación de Servicios Emergentes de Telecomunicaciones Avanzadas — promotion and identification of emergent services for advanced technologies) ([5]), the Portal Único de las Administraciones (single portal of the administrations) ([6]), CERES ([7]), aimed at providing certification services from the Fábrica Nacional de Moneda y Timbre, etc.

Localret ([8]) was founded in Catalonia in 1996 in order to develop the telecommunications networks and the use of information and communication technologies. The consortium is made up of 778 Catalan local councils, bringing together 99 % of the population of Catalonia, and the two Catalan municipal authorities: Federació de Municipis de Catalunya and Associació Catalana de Municipis. Localret, together with the Comissionat per a la Societat de la Informació, set up the following year by the Generalitat de Catalunya (Catalan Government), drew up the reference document *Catalunya en Xarxa: Pla estratègic per a la societat de la informació (1999–2003 Catalonia online: strategic plan for the information society)*, which proposes a series of initiatives and actions in six key fields for Catalan society.

([1]) The institutions and companies analysed, directly or indirectly, are: Aigües de Barcelona, Ajuntament de Barcelona, Ajuntament de Blanes, Ajuntament de Girona, Ajuntament de Terrassa, Banc Sabadell, Caixa d'Estalvis i Pensions de Barcelona, Col·legi Oficial d'Arquitectes de Catalunya, Diputació de Barcelona, Fecsa-Endesa, Generalitat de Catalunya, Parlament de Catalunya, Port de Barcelona (Autoritat Portuària de Barcelona and the company Pòrtic), Universitat Autònoma de Barcelona and Universitat Oberta de Catalunya.

([2]) Available at: http://www.map.es/libro/portada.htm

([3]) Information at: http://www.infoxxi.es

([4]) Information at http://www.igsap.map.es/sgpro/ventanilla/ ventunica.htm

([5]) Information at http://www.setsi.mcyt.es/sat/pista/index.htm

([6]) Information at http://www.infoxxi.es/ciuda.htm

([7]) Available at http://www.cert.fnmt.es/

([8]) Information at http://www.localret.es/

Finally, in July 2001 the *Pacte per a la promoció i desenvolupament de la societat de la informació a les administracions públiques catalanes (Pact for the promotion and development of the information society in the Catalan public administrations)* ([9]) was approved, which lays down the objectives of electronic government in Catalonia for the end of 2003 and proposes coordinated lines of action. Among the most interesting proposals is the foundation of the Catalan Certification Agency (ACC) and the Portal of the Public Administrations in Catalonia, which incorporates the *Pla d'implantació de serveis públics electrònics* ([10]). The first consequence of this has been the constitution of the Consorci Administració Oberta de Catalunya ([11]), made up of the Generalitat de Catalunya and Localret, aimed at promoting the objectives laid down in the abovementioned *Pacte*, for the implementation of electronic systems which allow services to be offered to Catalan citizens, companies and institutions through the use of information and communication technologies.

At the same time, all of these projects have been accompanied by the legal framework necessary for the regulation of electronic information exchange. Thus, the law dated 30/1992, of the legal system of the public administrations and of common governmental procedures, already considered the use of IT means for keeping records and in governmental procedures, but it was not until 1996 that the use of IT and telematic means was regulated in State Government ([12]), and not until 1999 that the validity of electronic signatures was approved, which is to be modified or replaced in the near future by a new law ([13]). Furthermore, the explicit regulation of the system of notifying governmental resolutions did not take place until December 2001 ([14]).

In reference to Catalonia, it is necessary to note, firstly, the regulation approved in 2001 on the relations between citizens and the Generalitat de Catalunya through the use of the Internet, which is the first time that governmental procedures can be carried out by telematic means in Catalan autonomous Government ([15]); and secondly, the passing in the same year of the new law on archives and records of Catalonia ([16]), to wit all Catalan public administrations must integrate a records management system in their administrative management. Therefore, it seems to be clear that, as the use of information and communication technologies in the public Administration increases, it will become necessary to integrate the archival criteria of records management within the framework of corporate management. A first indispensable tool in order to achieve this has been the creation of the working group on assessment criteria for electronic records, which answers to the Comissió Nacional d'Avaluació i Tria Documental (CNAATD) ([17]), whose mission is to put forward policy and procedural proposals for the management and conservation of electronic records of the Catalan public administrations.

Analysis and assessment of the use of information and communication technologies

Although all these initiatives affect in one way or another the production of electronic records, the indicators available ([18]) which enable the correct assessment of the use of information and communication technologies and their application in the management of electronic records are rather deficient or, at the most, only allow the following partial extrapolations:

The studies do not put directly forward uses for applications of electronic records management. Nevertheless, these uses can be deduced through those connected to banking and e-commerce and, also, within the area of the shared management of intranet processes and records.

- 45 % of Spanish companies use intranets, and among these, those with more than the 50 employees represent more than 90 %.
- 30 % of Catalan citizens aged over 15 use the Internet at least once a week. Connection is chiefly made from home (55.9 %), but Internet use from their workplace (33.6 %) and study centres (21.4 %) is also significant.
- In Catalonia, 50 % of companies with 10 or more employees have their own web page. 83.8 % are connected to the Internet, 82.7 % use e-mail and 38.4 % have some form of intranet or have a fully integrated in-house IT management system.
- 62.5 % of Internet users in Catalonia visit the web pages of the public administration, but only 19.8 % have ever carried out any kind of administrative application process. The reason probably lies in the scant presence of on-line application processes available and the lack of security perceived by users.

Lluís Cermeno

Lluís Cermeno is archivist, Chief of the Technical Inspection Section of the Archives Service, of the organisation for the coordination of Records Management, of the Government of Generalitat of Catalunya and for the planning and resources of investment at the network of Regional Archives of the Generalitat de Catalunya. Since 1991 he has been Secretary of the National Commission on Records Access and Appraisal and Selection (CNAATD) of Catalunya, where he assumes management.

Betlem Martínez

Betlem Martínez is responsible for the Central Administrative Archives of the Economic and Financial Department of the Generalitat de Catalunya. She has been technician archivist at the National Archives of

([9]) Available at http://www.gencat.es/sial/noticies/noti_localret.htm

([10]) Information at http://www.gencat.es/nova_administracio/egovern/serveis.htm

([11]) Resolution PRE/606/2002, dated 21 February, which makes explicit the Acord del Govern dated 4 December 2001, on the constitution of the Consorci Administració Oberta Electrònica de Catalunya and the approval of its statutes.

([12]) Spanish Reial Decret 263/1996, dated 16 February, which regulates the use of electronic information technology and telematic techniques on the part of Administració General de l'Estat (Spanish State Government).

([13]) Spanish Reial Decret-Llei 14/1999, dated 17 September, on electronic signatures. Information on the draft paper of the new law at http://www.infoxxi.es/dni.htm

([14]) Catalan Llei 24/2002, dated 27 December, on fiscals, administrative social order measures.

([15]) Catalan Decret 324/2001, dated 4 December, on the relations between citizens and the Government of the Generalitat de Catalunya through the Internet.

([16]) Catalan Llei 10 /2001, dated 13 July, on archives and records in Catalonia.

([17]) Comissió Nacional d'Avaluació i Tria de Documentació, agreement of 22 June 2001.

([18]) See the reports and studies in the bibliography at the end of this paper.

Catalunya. She is a member of the Working Group of National Commission on Records Access and Appraisal and Selection (CNAATD) of Catalunya. In 1995 she promoted the Working Group 'Archives and Internet', of the Archivists Association of Catalunya, and in 1998 she also promoted the list of distribution, Arxiforum, where she is administrator. She published many articles and she participated in formative actions related to the Internet and archives. She is a member of the group of professors for records management at the Public Administration School of Catalunya.

Remei Perpinyà

Remei Perpinyà is a doctor in history and Professor of documentation and archival sciences at the Universitat Autònoma of Barcelona. She worked as a private consultant in programmes for cultural management for several firms and institutions, between them the Informatitzation of the Episcopal Archives of Vic. Since 1997 to 2001 she has been board member of the Archivists Association of Catalunya. She has published many works related to information technology and archives.

- 67 % of the web sites of the public administration offer information on official exams, 63 % on subsidies, 34 % on procedures and 32 % on taxes. In contrast, 90 % of the use of CIT within the administration itself is centred on areas of fiscal management.

As a set of overall conclusions from the studies and reports looked at, the following points can be established:

- almost all companies are connected to the Internet and use e-mail, while half have their own web page;
- there is still a great deal to be done in e-commerce, particularly because of the lack of confidence in Internet security. The number of purchases and sales is low, but payments are even lower;
- the main uses of the Internet are information searches (83 %) and data transmission (62 %), followed at some distance by financial transactions (41 %), the study of the competition (21 %) and personnel recruitment (7 %);
- the main reason why companies use CITs lies in their competitiveness. 54 % consider them as an essential element in competitiveness;
- the public administration still uses CIT mainly as a tool to provide information and not as a means of rendering services.

Therefore, the results of the surveys are very positive in terms of equipment and networks but show that the pending question in terms of Internet use lies in the development of moves which are more based on the management and exploitation of business, particularly in terms of relations with suppliers and clients, and in the management and provision of online services on the part of the Public Administration.

2. Analysis of the electronic records management in Catalonia

In general, it can be said that many Catalan institutions and companies have replaced paper with electronic means in the production of records in some application processes. This replacement has essentially occurred in intermediary administrative processes, that is, in internal processes of the organisations. However, whereas this phenomenon has developed as a logical consequence of the widespread use of information and communication technologies, this has been carried out without the existence of specific projects aimed at the management of electronic records. Despite the existence of future projects which forecast their probable massive use, the projects already developed or under development are based either on general approaches to the overall management of records, or on the computerisation of a management process, which inevitably leads to the production of electronic records. Therefore, it is of no surprise that despite the fact that these projects affect the organisation as a whole, the different bodies are not supported by specific regulations on their management, with the exception, particularly in the private sector, of those requirements included in the administrative management regulations of certain processes and, in a few cases (not applicable to the rest), of recommendations on e-mail management.

In the case of public administrations, this has been the case, despite the participation of the respective archival services in the designing of a significant amount of these projects and, to a lesser extent, of in-house departments. Naturally, and independent of the type of organisation, it is necessary to note the significant specific weight of IT departments and the fact that until recently there were no records management systems defined with archival criteria (or at least in the case of the Catalan public administrations they were not sufficiently established). This has resulted, as a rule, in electronic-records management which is precarious and media-driven by very short-term objectives.

In this sense, it can be said that the electronic records management in the Catalan public administration has been conditioned by a clear dissociation between what the adoption of information and communication technologies represents for the production of records, and what is meant by accepting the consequences inherent in their management. It is not necessary to note that the dynamics generated around technological evolution itself has often favoured the setting of political objectives clearly orientated at an immediate result, whose approach has been developed as a result of a media-based focus.

Replacement of paper by electronic format?

As a result of a first assessment, it is necessary to note that while companies have sufficiently decidedly opted for the electronic format, the administration continues in general to base its management on paper. Evidently, this is a consequence of the administrative concept of what is a record, necessarily linked to the condition of authenticity and to the presence of a certified electronic signature. Nevertheless, in general terms it can be seen that the expansion in the use of e-mail, the complexity of some corporate databases and the development of web-enabled remote procedure tools are all leading to an inevitable change towards new visions of administrative management. In this sense, while office-automation documents are still considered as secondary copies which are always connected to the existence of the original on paper, other types of records in an electronic format can be considered *de facto* as originals.

This is the case of some corporate databases generally associated with the accounts management of any organisation, along with the fiscal obligations of tax payers, the management of the population census in the case of local government, and enrolment processes and management of academic reports in the university.

Furthermore, the setting up of online application process services both in companies and in the public administration involves the acceptance of these records as originals. This statement should be clarified in the case of the public administration, as the available operations are limited, in most cases, to obtaining information on the application process and the downloading of forms. In some special cases it is possible to print certificates and carry out the application process when this does not involve a resolution issued by a government body.

Non-text documents produced by an organisation can also be considered as originals, be they in audio, fixed image or image in movement. In the same way, cartographic documents generated directly by means of graphic-design tools are also often considered as original records, particularly those intended for support information upon which a geographical information system is based. Generally, these information systems are capable of doing away with a great deal of records on paper, but typically they are only managed from an electronic format. An example of this is the census on homes, buildings and business premises.

Another type of electronic record to be considered here is that obtained through the digitalisation of paper-based records. It is necessary to note that, in general terms, these are not corporate tools, but rather are of a departmental type, and in no case is the record digitised in order to replace the original, but rather just to facilitate circulation and consultation.

It is also necessary to note the fact that some institutions are, at a departmental level, beginning to have web-authoring tools, with the consequent production of electronic records in this format.

Therefore, we are faced with a situation of technology in full swing, in which it can no longer be taken for granted that all records are available on paper. In contrast, a more detailed analysis demonstrates that electronic records coexist with paper in organisations, even in the Administration, and that there is a clear upward trend in this sense. However, it is necessary to assess their legal validity and/or their consideration as an archival record.

With respect to the legal validity of this type of records as has already been mentioned the private sector has more room for manoeuvre through the use of validation tools agreed upon by both parties, either through certified electronic signatures or through any other method of identification, such as, for example, PINs. In the public sphere, there is also some experience in the distribution of PINs among citizens for the fulfilment of online application processes, although not to any significant extent.

As a general rule, the fact that the public administration does not use electronic signatures conditions the validity of any electronic record which it receives or issues. Nevertheless, some administrations can recognise a record's validity only at an internal level. This is the case of e-mail messages connected to internal procedures without a resolution on the part of any government body and, also, of certain corporate databases, cartographic documents and, even, web-enabled application process systems. Therefore, these documents should be considered as archival records,

both because of the interest of their contents and because of their condition as originals. In some cases, such as for example certain accounts databases, the condition of being an original is supported by context information which is not registered in the same database, but rather in others related to it, so that together they form an authentic network of interrelated data, which cannot be partially modified.

The information exchanged between different organisations deserves special attention. In this sense, one of the few examples in which telematic transmission takes place under the validation of an electronic signature is in the management of the employment situation for workers with respect to national social security. In other exchanges, the transfer of data is generally accompanied by the prior establishment of private protocols between both parties, which is what confers their validity. This is the case, for example, of the standardised transmission of records in the EDI format, widely used in the private sector, and very particularly by banks, where all these procedures are fully regulated in accordance with the requirements laid down by the Spanish Banking Association (AEB). Another example is that of the company Pòrtic, which is in charge of transport management at the port of Barcelona. Here, the company has implemented this format throughout the electronic records management.

However, at present this essential standardisation process is not employed in the data exchanges between the different local administrations and the State Government. Some examples of this include data transmission for the tax on economic activities, the tax on real assets, the tax on mechanical traction vehicles or the population census. In these cases, and despite the fact that often the format has been regulated by sectors, there is still a lack of overall standardisation which would simplify their management.

Degree of implantation of records management systems

In this paper we refer to records management systems (RMS) in the same sense as that defined by the law on archives and records of Catalonia: 'the set of operations and techniques, integrated in general administrative management, based on the analysis of production, application processing and the values of the records which are intended for planning, control, use, conservation and elimination or the transfer of the records to an archive, with the aim of rationalising and unifying treatment and achieving an efficient and cost-effective management' (Law 10/2001, dated 13 July, on archives and records, Article 2). Therefore, the objective of this section is to analyse to what extent organisations have defined records management systems and to what extent these RMS include the management of electronic records. It is necessary to mention that the final aim is to demonstrate the prior need for a records management system so that it may be correctly developed.

In the process of writing this paper we have noted that the implementation of RMS is unequal depending on the sector and, also, depending on its extent. Thus, while in the public administrations analysed, RMS is used in a more or less integrated manner and is computerised to a varying extent, the initiatives in the records management in the private sector stem, above all, from the management of business processes.

A significant number of public institutions have been endowed with the essential tools needed to guarantee a system of this type: classification, description, conservation and destruction systems and regulations on the transfer, destruction and access to the records. Almost all the experiences analysed employ custom-built corporate IT tools, mainly centred on the description of the records. Moreover, the majority of analysed systems of the public administration, despite being aimed at covering the entire range of records, are not in reality prepared for the use of electronic records. Despite this, it is necessary to recognise that in some cases work is being carried out on enabling the telematic recovery of electronic records, as is the case of the Parlament de Catalunya.

The private sector in general does not explicitly use this type of tool. In contrast, it employs exemplary regulations on administrative and management procedures, despite the fact that archival criteria have practically never been taken into account in the design of these processes. However, it is necessary to note that the arrival of the electronic records within organisations has demonstrated the need for the use of archival criteria in its management. The fact of also setting quality in internal work as a priority objective, that is, efficacy and efficiency, has in some cases produced results which we could consider as almost equivalent to an electronic records management sys-

tem (ERMS). In this sense, the use of corporate tools which integrate the management of any type of record throughout all its life cycle is particularly significant in the bank and financial sector, despite the fact that in the case of office-automated records, applications closer to an electronic document management system (EDMS) are used.

In relation to the criteria for the appraisal, selection and destruction of electronic records, the procedures and regulations intended to regulate and document the destruction of records are practically non-existent, with the exception of a few very specific examples. Generally, the criteria applied is conditioned by the capacity of online management and off-line storage systems, that is, they combine the application of a short term at a global level with the selective elimination by management units. In order to download information online, both the entire conservation of all records and its total destruction is used, along with the combination of both criteria. It should be noted that, in certain cases, the entire conservation involves the transfer of any type of electronic records, be it generated through messenger services, office-automation or through databases. However, in these cases the conservation of records in a web format is excluded.

The responsibility of determining the destruction of the documentation varies according to the type of record. Thus, it can be seen that the responsibility for mail messaging and office-automated records is almost always on the final user or the management unit, generally based on a previously-assigned memory quota. Other types of electronic records are eliminated or conserved by the joint decision of the management unit and IT services, or by the sole decision of the latter. The existence of interdisciplinary teams with the presence of archival services is still uncommon in the public administration and non-existent in the private sector.

Preservation and access

The type of electronic records considered as archival records by the majority of organisations analysed here are mainly those contained in databases and cartographic records. The rest, depending on the circumstances, can be considered as such by the full copy system in use, but their 'natural' format of conservation is on paper. There are also some examples in which the electronic messenger service is considered and treated as an archival record, such as the cases of the Parlament de Catalunya and the Banc Sabadell. Moreover, some databases and cartographic records could be classed as originals, despite the fact that they are not being archived regularly. It is also necessary to note as a non-existent practice, the consideration of the institutional webs and organisation's management as archival records, although some administrations intend to introduce this.

However, the fact that an electronic record is considered (or not) as an archival record does not imply that it must be classified for permanent conservation. Moreover, in general, the permanent conservation of electronic records is not considered, except in some very specific examples such as the Collegi Oficial d'Arquitectes de Catalunya, the Universitat Oberta de Catalunya, and the Ajuntament de Girona (Girona Town Council). Evidently, this is particularly serious in those organisations which recognise the existence of electronic records as originals and, therefore, unique. In this sense, the parameters of the temporary conservation of electronic records are between four months and 12 years.

With respect to the storage system, the non-existence (or almost) of large data warehouse systems is clear. The most usual systems involve the transfer of data by full copy to optic and magneto-optic means for permanent or medium-term temporary conservation. With short-term conservation the same systems can be used, though solutions of a more departmental nature, such as CD or DVD, are more common.

In reference to the guarantees of preservation, it is necessary to note the indiscriminate use of re-recordable systems, even for those records which are to be permanently conserved. In general, the downloading of online systems and copies responds to security criteria and, above all, to technological obsolescence, be it because of machines, programming or copy accessories, but in no case has data migration to preservation formats been considered. In a very few cases a standardisation of formats has been introduced, such as for example in the Parlament de Catalunya or in the company Pòrtic of the port of Barcelona. Even more exceptional is the agreement reached between

the different management units and IT services of an institution in order to allow transferred files to still be readable, as is the case of the Universitat Autònoma de Barcelona.

It is necessary to add that the practice of full copying does not permit the necessary distinction between permanently- and temporarily-conserved records and, therefore, despite the fact that these are correctly identified, migration represents a higher and higher cost and, at the same time, a risk for its long-term preservation. In this sense, it is necessary to caution that the availability of storage systems and devices with a surprisingly-growing capacity makes it even more difficult to adopt changes in this respect. It is also interesting to note that no example of the application of quality testing for stored data has been detected and that, according to the information supplied by the different organisations, the losses recorded have been caused by specific problems in the copying system and not because of the volatility of the data.

Furthermore, it is also necessary to consider the aspects linked to the security of records. With respect to physical security, the majority of institutions with a certain degree of organisational complexity outsource the custody of electronic records. In some cases, this type of measure is complemented (in order to achieve an improved security in management data) with the existence of external backup centres, particularly in banking and financial fields. With respect to the security of online records, the use of permits and levels of access is normal in all organisations, some of which have explicit regulations on the periodic and systematic change of passwords of each user, such as for example the Ajuntament de Terrassa (Terrassa Town Council).

Obviously, another key matter lies in guaranteeing the criteria of authenticity, reliability and integrity of the records. As a result of the study carried out, no organisation has been identified which has adopted these criteria in data migration. In fact, the authenticity conditions set at the time of the production of records are conditioned by the use of electronic signatures and by the incorporation of metadata of origin. From the data obtained, it can be seen that the use of electronic signatures does not always imply the adoption of security systems in long-term records preservation. Moreover, the registering of more or less standard context data and its upkeep in each of the different migrations can be considered as merely testimonial, being limited to a single organisation out of those analysed as a whole.

3. Conclusions

It is undeniable that information and communication technology is firmly rooted in almost all Catalan companies and institutions as they are considered an essential factor of competitiveness for the improvement in the efficiency of administrative management. In consequence, the majority of organisations produce records in an electronic format and a significant proportion of them have set up projects for the e-management of records. Some of these projects were begun in the 1990s, but the majority of them are still in the study and design stage.

The result of the study shows that there are different levels in the electronic records management. Firstly, in the majority of the cases their management is not considered at a corporate level, but rather sector-based projects are set up, which do, however, affect key aspects of the economic activity of organisations. Secondly, in the majority of cases the electronic records management is set up in order to computerise intermediate administrative and commercial procedures. Despite this, some innovative experiences are currently in operation in the private sector where electronic certification is being introduced and where electronic records are considered as originals. Thirdly, all the experiences analysed consider the electronic records management during the stage of administrative procedures and in no cases have strategic policies been defined for the long-term preservation of records. Finally, experiences demonstrate that in the electronic records management, archival methodology has not been taken into account in establishing identification criteria for records and their appraisal and selection, systematic classification, description and long-term preservation policies.

Differences have also been observed between the practices in the public administration and those in private companies. In private companies, the electronic records management stems from the implementation of computerised management systems for business processes in order to achieve an improved business efficiency, which inevitably forces them to redesign their records management systems and to establish conservation policies for electronic records. In contrast, in the pub-

lic administration the implementation of electronic records management is slower than in private companies, given that the level of legal exigency is higher. It is for this reason that although it is in the public administration where records management systems are most present, it is also in this area where the existence of systems for the electronic records management is lower.

Nevertheless, the general perception is that the electronic records management is inevitable and that it is only a matter of time before their use becomes widespread. Their implementation depends on four key points:

- promote the innovation and transformation of administrative processes in order to systemise and rationalise the flow of records;
- real implementation of electronic certification which affords electronic records with legal security in order to guarantee their reliability and authenticity;
- technological resources in order to guarantee the circulation of records, sufficient memory capacity, management applications and the durability of supports;
- methodology of records management enabling the identification of electronic records, their appraisal and selection at the time of creation, the establishment of classification systems and conservation policies for formats and supports.

In all the existing and future projects analysed in the writing of this paper, the first three elements have been taken into account, though not the fourth. To this end, it is very important for archivists to be involved in the design of electronic records management systems, allowing us to contribute our knowledge and our methodology, so enabling these projects to be improved in the following areas:

- identification of electronic records. This is indispensable to delimit the definition criteria of what is a record in the IT field;
- the appraisal of electronic records which should allow the value of records to be determined during the different records management stages;
- the systematic classification of records as a basic tool for the identification and organisation of the institution's records;
- description systems in order to incorporate in the metadata the description criteria allowing the records to be efficiently recovered at a later date;
- conservation policies of formats and supports, making it possible to guarantee their long-term preservation.

All of the above should be done in a fully-integrated environment, which means setting up corporate models for the records management, where the decisions cannot be taken partially, but rather by means of commissions made up of the different agents involved and in which the different points of view are equally considered.

Bibliography

Arnau, J.; Costa, À.; de la Hoz, J., 'Les tecnologies de la informació i la comunicació (TIC) i l'empresa catalana' in *Nota d'Economia, Revista d'economia catalana i de sector públic.* (Nos 69–70, 1st and 2nd quarters), Barcelona: Generalitat de Catalunya, Departament d'Economia i Finances, 2001, pp. 75–96.

Castells, M.; Díaz de la Isla, M. I.; IN3-UOC, *Diffusion and uses of Internet in Catalonia and in Spain. A commented summary of available evidence, as of 2001* (online) (consulted 20.3.2002).
Available at: http://www.uoc.es/in3/wp/picwp1201/

Comissionat per a la Societat de la Informació. *Catalunya en Xarxa: Pla estratègic per a la Societat de la Informació.* Barcelona: Comissionat per a la Societat de la Informació, 1999.
Available at: http://dursi.gencat.es/ca/de/pla_estrategic.htm, and also at: http://www.localret.es/

Generalitat de Catalunya. *Guia pràctica d'implantació d'un sistema de gestió ambiental.* Barcelona: Generalitat de Catalunya. Departament de Medi Ambient, Col·lecció Manuals de gestió, 2, 2001.

DMR Consulting-SEDISI (Asociación Española de Empresas de Tecnologías de la Información). *Las tecnologías de la Sociedad de la Información en la empresa española, 2000* (on-line) (consulted: 25.11.2001).
Available at: http://www.spain.dmr.com/content/estudio.pdf

FIGUEROLA, Eduard. 'El Comerç Electrònic entre empreses (B2B)' in *Revist@ d'Indústria, Comerç i Turisme* (No 24, September–October) 2001 (online).
Available at: http://www.gencat.es/dict/revista/24/inici/index.htm (consulted: 25.11.2001)

GARCIA, Maribel; JORDÀ, Daniel. 'Enquesta d'avaluació de la penetració de les TIC a la societat catalana, 2001' in *Nota d'Economia, Revista d'economia catalana i de sector públic* (No 69–70, 1st and 2nd quarters) Barcelona: Generalitat de Catalunya, Departament d'Economia i Finances, 2001, pp. 97–115.

IN3 INTERNET INTERDISCIPLINARY INSTITUTE. *La informació, recurs rendible i mesurable dins la nova economia. Suplement de l'IN3* (online) November 2001, (consulted 25.11.2001)
Available at: http://campus.uoc.es/in3/suplement.htm

MEYA, Montserrat; MELÉNDEZ, Vicenç. 'L'ocupació de personal TIC a l'empresa catalana' in *Nota d'Economia, Revista d'economia catalana i de sector* (Nos 69–70, 1st and 2nd quarters) Barcelona: Generalitat de Catalunya, Departament d'Economia i Finances, 2001, pp. 117–128.

MINISTERIO DE ADMINISTRACIONES PÚBLICAS; COAXI (Cooperación entre las Administraciones Públicas en el campo de los Sistemas y Tecnologías de la Información). *Informe IRIA 2000: Informe de Recursos Informáticos de las Administraciones Públicas* (on-line). Madrid: 2001 (consulted: 5.12.2001).
Available at: http://www.map.es:80/csi/iria2000/pdf.htm

SECRETARIA PER A LA SOCIETAT DE LA INFORMACIÓ; IDESCAT: *Estadístiques de la Societat de la Informació. Catalunya 2001* (online). 2001 (consulted 27.11.2001).
Available at: http://dursi.gencat.es/jsps/documentssi.jsp?area=3&est=true&sub1=11&sub2=1&sub3=0

VILLASOL, L.; MARSAL, J., R., 'Las TIC como instrumento de mejora de las relaciones de la Administración con los ciudadanos. El proyecto Administración Abierta electrónica de Cataluña', in *Actes de les Ves. Jornades d'Organització i Noves Tecnologies en l'Administració Local*. Blanes: 2001, pp. 16–17.

La gestión electrónica de documentos en Cataluña: situación actual

Ferran Agelet, Lluís-Esteve Casellas, Lluís Cermeno, Betlem Martínez, Remei Perpinyà

El objetivo de este documento es abordar la cuestión de la gestión electrónica de documentos en Cataluña, esto es, quién es responsable de su gestión, cómo se gestionan y cuál es la función de los archiveros en cuanto a la aplicación de los sistemas que lo hacen posible. Con este objetivo, el trabajo se basa en el estudio de las experiencias en varias instituciones públicas y organismos privados en Cataluña. Al seleccionar la muestra, nos hemos fijado en estos organismos dado que, por su importancia o por el nivel de interés de sus experiencias, nos permite comprender la situación de Cataluña.

Desde el punto de vista del archivo entendemos la gestión electrónica de los archivos como una actividad que abarca todo el ciclo de la vida de los documentos: desde la fase de diseño de los formatos de estos archivos, su creación como archivos auténticos, completos y fia-

bles como testimonios de un procedimiento específico, hasta el mantenimiento de los requisitos que garantizan su valor probatorio en el futuro. Por esta razón, en el estudio de cada experiencia se ha hecho especial hincapié en los ámbitos en que se aplica la gestión electrónica de los registros: elaboración de registros, gestión administrativa y conservación, con el objetivo de comprobar hasta qué punto se han tenido en cuenta en el proceso los criterios y la metodología de archivo.

El resultado del estudio pone de manifiesto que existen distintos niveles en la gestión de los documentos electrónicos. En primer lugar, en la mayoría de los casos, la gestión no se prevé a nivel global de la empresa, sino que consiste más bien en establecer proyectos sectoriales que, no obstante, afectan a aspectos básicos de la actividad económica de las organizaciones. En segundo lugar, a menudo se realiza la gestión de los documentos electrónicos con el fin de informatizar los procedimientos administrativos y comerciales intermedios. A pesar de ello, en el sector privado se están llevando a cabo algunas experiencias innovadoras, con la introducción de la certificación electrónica y la consideración de algunos documentos electrónicos como originales. En tercer lugar, todas las experiencias analizadas consideran la gestión de los documentos electrónicos en la fase de los procedimientos administrativos y en ningún caso se han definido estrategias políticas para su conservación a largo plazo. Por último, las experiencias demuestran que, en la gestión de los documentos electrónicos, los métodos de archivo no se han tenido en cuenta a la hora de establecer criterios de identificación para los documentos, su evaluación y su selección, la clasificación sistemática, la descripción y las medidas de conservación a largo plazo.

También se han observado diferencias entre las prácticas de la administración pública y las de las empresas privadas. En las empresas privadas, la gestión de los documentos electrónicos deriva de la aplicación de sistemas de gestión informatizados de los procesos de trabajo, con el fin de mejorar la eficacia de la empresa, lo que inevitablemente les obliga a revisar sus sistemas de gestión de archivos y a establecer medidas de conservación de los documentos electrónicos. En cambio, la aplicación de la gestión de los documentos electrónicos en la administración pública es más lenta que en las empresas privadas, dado que los requisitos jurídicos son más rigurosos.

El sentimiento general es que la gestión de los documentos electrónicos es inevitable y que su generalización es sólo cuestión de tiempo. Su aplicación depende de cuatro elementos fundamentales:

- promover la innovación y la transformación de los procesos administrativos con el fin de sistematizar y racionalizar los flujos de documentos;

- aplicar realmente la certificación electrónica, que confiere seguridad jurídica a los documentos electrónicos, con el fin de garantizar su fiabilidad y su autenticidad;

- movilizar los recursos tecnológicos que permiten garantizar la circulación de los documentos, la existencia de capacidad de almacenamiento suficiente, aplicaciones de gestión y la duración de los soportes;

- desarrollar métodos que permitan identificar los documentos electrónicos, clasificarlos y seleccionarlos en el momento de su creación, establecer sistemas de clasificación y medidas de conservación de los formatos y soportes.

Los tres primeros elementos expuestos se tuvieron en cuenta en los proyectos actuales y futuros sujetos a nuestro análisis, pero no así el cuarto. A tal efecto, es primordial que los archiveros participen en la concepción de los sistemas de gestión de los documentos electrónicos para aportar sus conocimientos y su metodología, de tal modo que mejore la definición, la selección y la clasificación sistemática de los documentos electrónicos y la instauración de sistemas de descripción y medidas de conservación de los formatos y apoyos, haciendo posible garantizar su conservación a largo plazo.

Elektronische Schriftgutverwaltung in Katalonien: aktueller Stand

Ferran Agelet, Lluís-Esteve Casellas, Lluís Cermeno, Betlem Martínez, Remei Perpinyà

Anliegen dieses Beitrags ist es, die Verwaltung elektronischer Aufzeichnungen in Katalonien zu erläutern, d. h.: Wer ist dafür zuständig, wie läuft sie ab und welche Rolle spielen Archivare bei der Anwendung von Systemen, die dies ermöglichen. Zu diesem Zweck werden Erfahrungen in verschiedenen öffentlichen Einrichtungen und privaten Körperschaften in Katalonien beleuchtet. Die Beispiele wurden unter Organisationen ausgewählt, die es uns aufgrund ihrer Bedeutung oder wegen der Relevanz ihrer Erfahrungen ermöglichen, den aktuellen Stand in Katalonien darzustellen.

Vom Standpunkt der Archivierung aus verstehen wir die elektronische Schriftgutverwaltung als eine Tätigkeit, die den gesamten Lebenszyklus von Unterlagen umfasst: von der Konzipierung der Formate dieser Unterlagen, ihrer Erstellung in Form authentischer, vollständiger und zuverlässiger Aufzeichnungen als Zeugnisse eines speziellen Vorgangs bis hin zur Einhaltung von Anforderungen, die ihre künftige Beweiskraft sichern. Aus diesem Grund ist bei der Analyse der einzelnen Erfahrungen besonders den Bereichen Aufmerksamkeit geschenkt worden, in denen die elektronische Aktenführung zur Anwendung kommt: Erstellung von Aufzeichnungen, administrative Verwaltung und Aufbewahrung sowie Überprüfung, in welchem Umfang die Kriterien und Archivierungsmethoden während des gesamten Verfahrens berücksichtigt wurden.

Das Ergebnis der Untersuchung zeigt, dass es bei der elektronischen Schriftgutverwaltung verschiedene Ebenen gibt. Erstens wird sie in der Mehrzahl der Fälle nicht auf Unternehmensebene betrachtet, sondern es werden sektorbezogene Projekte eingerichtet, die jedoch zentrale Aspekte der Wirtschaftstätigkeit von Organisationen betreffen. Zweitens wird eine Verwaltung elektronischer Unterlagen meist eingerichtet, um die unmittelbaren Verwaltungs- und Geschäftsvorgänge per Computer abzuwickeln. Davon abgesehen laufen derzeit im privaten Sektor einige innovative Projekte, bei denen eine elektronische Zertifizierung eingeführt wird und elektronische Aufzeichnungen als Originale gelten. Drittens betrachten alle untersuchten Vorhaben die elektronische Aktenführung in der Phase der administrativen Prozesse, und in keinem Fall sind strategische Konzepte für die Langzeitaufbewahrung von Unterlagen festgelegt worden. Nicht zuletzt zeigen die Beispiele auch, dass bei der elektronischen Schriftgutverwaltung keine Archivierungsmethode berücksichtigt wurde, um Identifizierungskriterien für Aufzeichnungen und deren Bewertung und Auswahl, die systematische Klassifizierung, die Beschreibung und Maßnahmen für die Langzeitaufbewahrung aufzustellen.

Unterschiede traten auch zwischen der öffentlichen Verwaltung und privaten Unternehmen zutage. In privaten Unternehmen ergibt sich die elektronische Schriftgutverwaltung aus der Einführung EDV-gestützter Systeme für Geschäftsprozesse zur Verbesserung der Wirtschaftlichkeit, was zwangsläufig eine Neugestaltung der Aktenführungssysteme und die Festlegung von Konservierungsmaßnahmen für elektronische Aufzeichnungen erfordert. Demgegenüber vollzieht sich die Einführung elektronischer Aktenführung in der öffentlichen Verwaltung langsamer, da die gesetzlichen Anforderungen höher sind.

Die allgemeine Erkenntnis lautet, dass eine elektronische Schriftgutverwaltung unumgänglich und ihre breite Anwendung nur eine Frage der Zeit ist. Maßgeblich für ihre Umsetzung sind vier wesentliche Punkte:

- Förderung der Erneuerung und Umgestaltung von Verwaltungsvorgängen zur Systematisierung und Rationalisierung des Datenflusses;

- wirkliche Einführung einer elektronischen Zertifizierung, die elektronischen Aufzeichnungen Rechtssicherheit verleiht, um ihre Zuverlässigkeit und Authentizität zu gewährleisten;

- technische Ressourcen, um den Umlauf der Aufzeichnungen, eine ausreichende Speicherkapazität und die Haltbarkeit von Datenträgern sicherzustellen;

- Methodik der Unterlagenverwaltung, die die Identifizierung elektronischer Aufzeichnungen, deren Bewertung und Auswahl zum Zeitpunkt der Erstellung, die Aufstellung von Klassifizierungssystemen und Konservierungsmaßnahmen für Formate und Datenträger ermöglichen.

Bis auf den vierten sind all diese Faktoren bei sämtlichen bestehenden und künftigen Projekten, die zur Erarbeitung dieses Beitrags analysiert wurden, berücksichtigt worden. Deshalb ist es für Archivare sehr wichtig, in die Konzipierung von Verwaltungssystemen für elektronische Aufzeichnungen einbezogen zu werden, so dass wir mit unserem Wissen und unserer Methodik dazu beitragen können, bei diesen Projekten die Identifizierung, Bewertung und systematische Klassifizierung elektronischer Aufzeichnungen, die Umsetzung von Verzeichnungssystemen und Konservierungsmaßnahmen für Formate und Datenträger zu verbessern und so mit dafür zu sorgen, dass ihre Langzeitaufbewahrung gesichert ist.

État actuel de la gestion des archives électroniques en Catalogne

Ferran Agelet, Lluís-Esteve Casellas, Lluís Cermeno, Betlem Martínez, Remei Perpinyà

L'objectif du présent exposé est d'aborder la gestion des archives électroniques en Catalogne, c'est-à-dire de répondre aux questions suivantes: qui est responsable de la gestion, comment ces archives sont-elles gérées et quel rôle les archivistes jouent-ils dans la mise en œuvre des systèmes qui rendent cet archivage possible. À cette fin, les travaux reposent sur l'analyse de l'expérience acquise dans plusieurs institutions publiques et organismes privés catalans. Nous les avons retenus dans notre échantillon soit en raison de leur importance, soit parce que le degré d'intérêt de leur expérience nous permettait de comprendre l'état actuel de la situation en Catalogne.

Du point de vue archivistique, nous concevons la gestion des archives électroniques comme une activité qui s'étend sur l'ensemble du cycle de vie des documents, à savoir: depuis le stade de l'étude des formats de ces documents, leur création sous la forme de documents authentiques, complets et fiables, témoignant d'une procédure spécifique, jusqu'à la maintenance des conditions qui garantiront leur valeur probatoire à terme. Pour cette raison, l'analyse de chaque expérience a porté en particulier sur les domaines dans lesquels la gestion des documents électroniques est appliquée (production des documents, gestion administrative et conservation), le tout dans le but de voir dans quelle mesure les critères et les méthodes archivistiques ont été pris en compte tout au long du processus.

Le résultat de l'étude montre qu'il existe différents niveaux dans la gestion des documents électroniques. Premièrement, dans la majorité des cas, leur gestion n'est pas envisagée au niveau global de l'entreprise, mais consiste plutôt à mettre en place des projets sectoriels qui, toutefois, ont bel et bien un impact sur certains aspects majeurs de l'activité économique des organisations. Deuxièmement, la gestion des documents électroniques est mise en place très

souvent afin d'informatiser les procédures administratives et commerciales intermédiaires. En dépit de cette tendance, certaines expériences innovantes sont actuellement en cours dans le secteur privé, où l'on assiste à l'introduction de la certification électronique et où des documents électroniques sont considérés comme des originaux. Troisièmement, toutes les expériences soumises à analyse considèrent la gestion des documents électroniques au stade des procédures administratives, et jamais aucune stratégie politique n'est définie pour leur conservation à long terme. Enfin, les expériences prouvent que, en ce qui concerne les documents électroniques, les méthodes de l'archivistique ne sont pas prises en compte pour mettre en place les critères d'identification des documents, leur tri et leur sélection, le classement systématique, la description et les mesures de conservation à long terme.

On note également des différences entre les pratiques de l'administration publique et celles des sociétés privées. Dans les entreprises privées, la gestion des documents électroniques découle de la mise en œuvre de systèmes de gestion informatisés des processus de travail afin d'améliorer l'efficacité de l'entreprise, ce qui les contraint inévitablement à revoir leurs systèmes de gestion d'archives et à mettre en place des mesures de conservation des documents électroniques. En revanche, la mise en œuvre de la gestion des documents électroniques dans l'administration publique est plus lente que dans les entreprises privées, compte tenu du fait que les contraintes juridiques y sont plus rigoureuses.

Le sentiment général est que la gestion des documents électroniques est inévitable et que sa généralisation n'est plus qu'une question de temps. Sa mise en œuvre dépend de quatre conditions:

— promouvoir l'innovation et la transformation des processus administratifs afin de systématiser et de rationaliser les flux de documents;

— mettre en œuvre réellement la certification électronique, qui confère la sécurité juridique aux documents électroniques, afin de garantir leur fiabilité et leur authenticité;

— mobiliser les ressources technologiques permettant de garantir la circulation des documents, l'existence de capacités de stockage suffisantes, des applications de gestion et la pérennité des supports;

— développer des méthodes permettant d'identifier les documents électroniques, de les trier et de les sélectionner au moment de leur création, de mettre en place des systèmes de classement et des mesures de conservation des formats et des supports.

Les trois premières conditions ci-dessus ont été prises en compte dans les projets actuels et futurs soumis à notre analyse, mais pas la quatrième. À cet effet, il est primordial que les archivistes participent à la conception des systèmes de gestion des documents électroniques pour y apporter leurs connaissances et leur méthodologie, de façon à améliorer l'identification, le tri et le classement systématique des documents électroniques et la mise en place de systèmes de description et de mesures de conservation des formats et des supports, pour garantir leur conservation à long terme.

Guidelines on developing an e-business based corporate electronic records management strategy — from business case through to implementation

Martin Waldron

There is a history of management not being sure of the business imperative for investing in records management, and they generally perceive it as a downstream administration system for archiving.

There is a chasm that has to be crossed to convince management of the key role of ERMS in implementing their e-government/e-business plans, and its vital role in an organisation's information management strategy.

This paper addresses this credibility gap by providing a framework with which to develop a business case for electronic document record management system — EDRMS and its vital role in deploying e-government. The framework is based on work we have undertaken in assisting central government groups such as the Scottish Parliament and the Department for Enterprise, Trade and Investment Northern Ireland, and regional government such as Fife Council, Argyle and Bute and Bedfordshire Police Authority, in developing their e-government strategy.

The approach draws on our experience in both commercial and public sector, and uses as a base the UK Government Treasury Greenbook which outlines an appraisal and evaluation approach and report format for system appraisal.

Background

There has suddenly been a realisation by public organisations that their conventional paper based records management systems aren't working. Most business processes both outward facing to the citizen, customer and supplier, and in the back office, are increasingly being undertaken through electronic transactions and documents. Many organisations large and small do not have in place the controls and infrastructure to manage effectively electronic records which are increasingly being received via e-mail and accessed from the web, as well as being generated internally. The result is that documents are printed and managed as records using existing paper-based systems with resulting inefficiencies and delays in access.

Arising at the same time is the fact that records management, previously perceived as a tool for meeting statutory legal/audit demands, has now been more widely recognised as generic to an organisations information/knowledge management strategy. Suddenly, metadata and taxonomy are on the lips of organisations' information-management directors. The result is that records management, traditionally seen as a backroom (basement) activity that has traditionally been an activity at the end of the business process, has had its profile raised and forms part of an organisation's overall information management strategy.

Are e-government initiatives delivering?

There are enormous challenges to take a public sector organisation from currently manually intensive service delivery to an e-business citizen- (customer-) driven operation. There are change management issues, financial constraints and the need to embrace social inclusion including maintaining traditional service channels.

Major strides have however been made in the public sector to improve accessibility of services and information. The web has been a major contributor with public access to service information, forms, local information etc. and the ability to e-mail enquiries. The UK modernising government initiative that is requiring all services to be offered electronically by 2005 has seen the UK public sector shovel information onto the web. This has improved availability of information but not tack-

Martin Waldron

Martin Waldron has a BSc Hons degree in physics from Bristol University and a Diploma in education from Oxford University. He is a member of the Records Management Society and a visiting fellow at Richmond College, International American University in London.
His professional career has been in the application of information technology to business applications. He has held a range of senior posts with system houses, software development companies and in-house IT systems departments including the Canadian Government and ICL. He has been a management consultant since 1987, when he founded In-Form Consult management consultancy (www.informsystems.co.uk), which specialises in the application of electronic document, record and work management systems and web technology to improve business processes.
Martin has been instrumental in the introduction of innovative technology for business improvement. He led the project in 1967 to provide the first UK computer to computer link over a telecommunications line between different systems,
project-managed the first UK high-speed links between disparate mainframes, minicomputers and peripherals at ULCC, responsible in 1984 as business development manager for the introduction of the first UK commercial PC-based imaging system at Xionics.
He is the managing editor of Infobulletin a web-based EDRMS newsletter www.infobulletin.co.uk Recent projects he has been involved in include: co-author of a European Commission project to develop an electronic records management model/functional specification — MoReq — that can be used by organisations in assisting in the selection of ERMS/EDMS systems, assess effectiveness of current electronic records management practices and for educational use. He is Project Manager for an ERMS scoping study for the Scottish Parliament to develop their ERMS/EDMS strategy, define their functional requirements and implementation plan and in developing their ITT and supplier selection.Currently, he is working

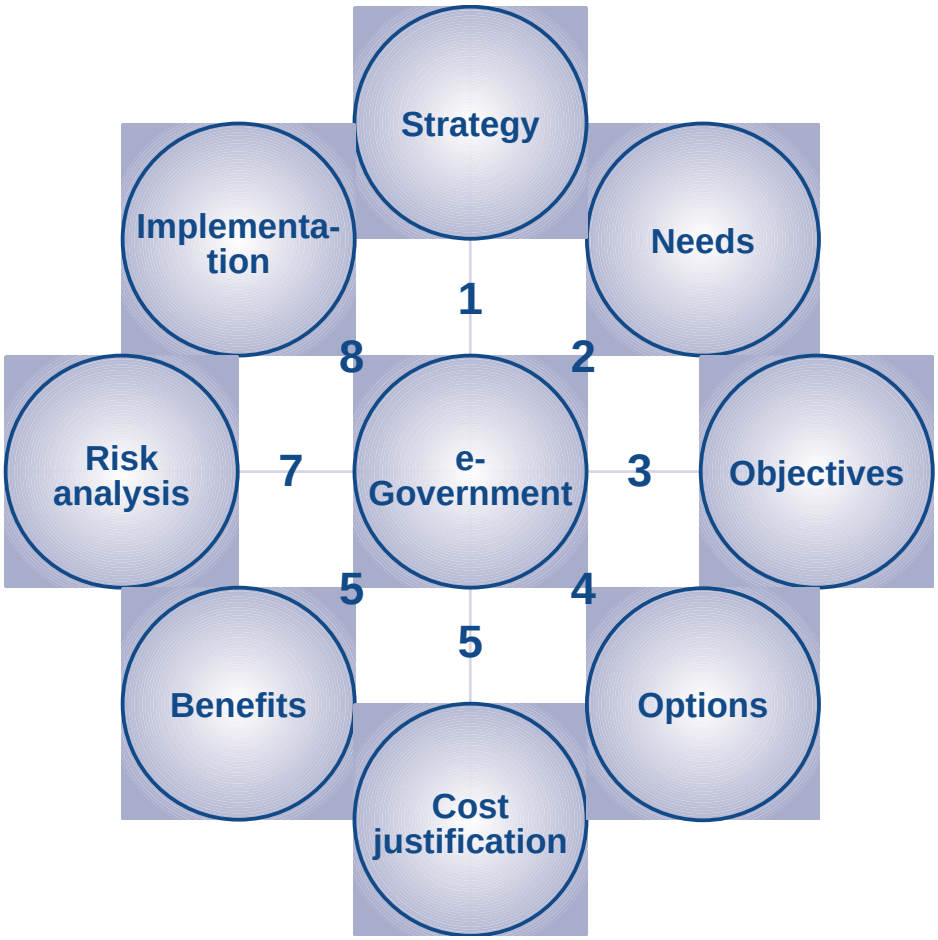
with the Department of Enterprise, Trade and Investment, Northern Ireland, where he is leading a team that have developed their business case for EDRMS and undertook a feasibility study for the implementation of DETI e-business applications. He also leads a team to assist Fife Local Authority, Scotland, in defining their requirements for EDRMS and in the development of a cross-department statement of requirement for EDRMS. He has developed a European White Paper for Hewlett Packard on legislative support and best practice justification of the of the need for WORM technology for long-term electronic archives. He is currently assisting Bedfordshire Police in developing a statement of requirement for a force-wide registry administrations system. He led the project to support the state of Jersey in the selection of a replacement for their text retrieval system, the conversion of their data and procurement of an EDMS/knowledge management system.

led the core business issues of improved service levels for core processes such as benefit claims or delivered internal efficiencies. A cosmetic approach to meeting e-government targets is however rampant. This is understandable as there are major cultural and skills issues in shifting to delivering services electronically. A recent report by the UK Audit Commission 'Councils and e-Government' cited from a survey of 64 Councils that the three main barriers to delivering e-services were the capacity to manage the change process and lack of ICT skills and knowledge of staff. This report also identified councils' 'lack of a robust cost information on their proposed e-activities to inform these choices ...'. The report also stated 40 % of chief executives found the requirements too broad to tackle alone and 61 % that it is too costly.

The following sections outline an approach that has been successfully undertaken by a number of public organisations in cost-justifying and developing an EDRMS strategy to support e-government delivery.

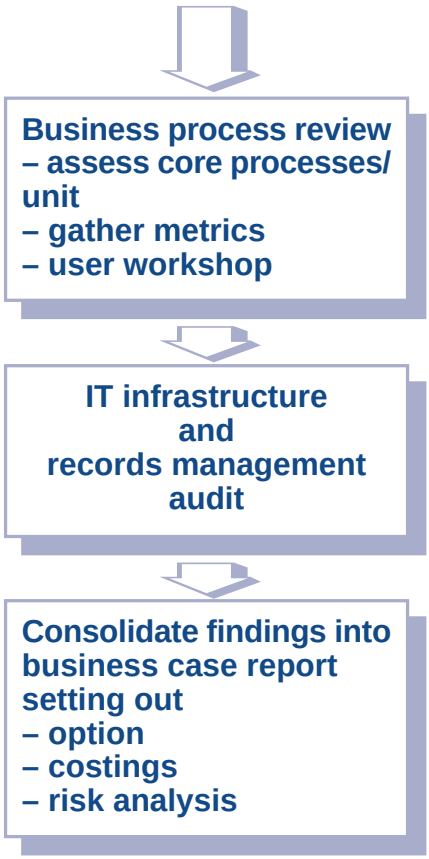
EDRMS corporate system appraisal

The diagram below sets out the steps in progressing an e-government development plan from strategic business drivers through to implementation. Each step requires consultation with all stakeholders to agree the outputs from the step before progressing to the next step. The stakeholders will include senior management, service delivery staff ('the user'), strategic partners and the citizen. Consultation and involvement by these parties will vary according to the organisations 'business' services.



Each of these stages is outlined below illustrated by case histories.

The foundation for this work will be a high-level process and filing/records review of each business unit or service function to provide a base for the metrics and volumes of the business model. An assessment of current IT infrastructure and policies will also need to be undertaken. This stage should include workshops with users to give them an appreciation of the benefits and features of EDRMS.



Step 1 — Strategic drivers for corporate EDRMS

This initial work identifies the key strategic areas that the proposed EDRMS project supports within the departmental e-government strategy plus related business drivers and regulatory requirements. Both central and local government in the United Kingdom have been required to develop an e-business strategy plan. Identifying the vital support for this plan by EDRMS is a major objective of this section.

This extract from DETI ([1])'s e-business plan that illustrates this point:

'Utilise e-business concepts, processes and technologies to modernise and improve the services delivered by DETI in order to:

- encourage growth of the economy by promoting knowledge based business competitiveness and an enterprise culture for NI; and
- improve service delivery and choice to customers whilst developing and maintaining the equality, policy and regulatory environment necessary to achieve high levels of enterprise and fairness;
- to achieve the proposed electronic service delivery (ESD) targets set for the Department and Agency, e.g. delivery of 25 % of key services electronically by 2002 and 100 % by 2005.'

The e-business plan will set out the stages of developing the e-business strategy and the new operational model as outlined below.

These plans represent a radical transformation — both in the way the department will interact with customers and internally within Business Units. EDRMS technologies are key components in meeting these goals by providing an infrastructure for ESD — electronic service delivery with:

- all records and files intelligently held in a electronic repository;

([1]) DETI — Department of Enterprise, Trade and Investment — Northern Ireland.

- the automation and streamlining of processes to improve service levels;
- the use of the Web for publishing and interaction with customers.

Other key strategic drivers will be legislation e.g. Freedom of Information, Data Protection Act, The Public Records Act and the need to develop a corporate information/knowledge management strategy.

Step 2 — Establish the need and current deficiencies

The next phase will be from the analysis of current operations and the demands on organisations to meet e-government electronic service delivery plans articulating the critical role of EDRMS to support these plans.

ERMS is at the heart of the UK Government's 'modernising government' agenda. Implementation of an electronic records management system has to be in place across the department so that all new records are created and managed electronically by 2004.

Current departmental records management at DETI was identified as needing radical change to support these plans:

The following extracts from the report support this:

'Despite an IT infrastructure that allows for all staff to generate, store, receive and send documents electronically, operate almost exclusively paper-based processes for all its core business and administrative functions and with the filing of records and documents in paper form. In order to achieve the e-business objectives laid out above they will require wide-scale adoption of electronic processing and the storing and retrieving of records and documents electronically.'

At DETI currently all records registered by department are in paper format with conversion of e-mails and other electronic files to paper records. The management of these records is manually based with a limited use of technology. The UK Government with their modernising government agenda also set two key drivers for the public sector to develop and implement their 'e-business' plans. Public service from 2004 can only create records in electronic form and all services which are capable of electronic delivery must be available as an e-service by 2005.

For DETI to achieve the 2004 goal will require a very demanding schedule not only to put in place the systems for ERMS but to educate/train staff and convert existing files. The DETI EDRMS project report identified that the current methods of information access and control will need a radical change.

'Given the current processes and systems within the department, there is a high reliance on paper-based processes to enable communication between business units, this leads to:

- limited access to information for users;
- duplication of effort;
- creation of 'silos of information';
- difficulty in obtaining up-to-date information and obtaining a corporate wide view of events;
- inconsistency in the application and completion of common activities throughout the department as these have been developed for each individual business unit, for example, the processing of grant applications from initial request through to payment which is undertaken by a number of business units;
- difficulties in accessing, manipulating and extracting data;
- the existence of information 'silos' due to the fact that systems have evolved over time and for each BU in isolation;
- limited integration of systems therefore limiting DETI's ability to share information across BUs.'

Step 3 — Objectives and constraints of project

The project then identified the objectives of EDRMS project to support the key strategic drivers from Step 1.

These would be such things as:

- meet the modernising government target dates for electronic records;
- standardise the management of electronic and paper records across the organisation;
- support knowledge management developments;
- provide a framework for EDRMS across all departments;
- improve service levels to citizens and businesses;
- improvements in staff efficiency.

This section would also identify and define any constraints to the project in areas such as:

- technical
- regulation and guidelines
- staff education
- customer acceptance
- relationship with other projects.

Step 4 — Assessment of options

The next phase of the analysis was to identify options for introducing EDRMS and make an assessment. This would progress from an assessment of the pros and cons of the current system and then identify number of approaches, each with different levels of sophistication. Examples considered by the project were:

Option 1: Base case describes the current methods of working, identifying the current strengths and weaknesses of record management and processes.

Option 2: Implement a records management programme in each business unit/division

Option 3: Implement a corporate records-management programme

Option 4: Introduce an ERMS by business unit

Option 5: Introduce an ERMS corporately

Option 6: Introduces an integrated corporate EDRMS with document/records management, workflow, web publishing and forms processing

Option 7: Outsource/PFI ERMS Archive

These options were presented to the project board and three options selected for a more detailed costing and benefit analysis.

Step 5 — Identify, quantify and where possible value the costs and benefits of each option

Each shortlisted option was assessed under the following categories:

- software and hardware costs;
- manpower resource requirements;
  - — training and advice
  - — project management
  - — systems integration
  - — system support staff
  - — business analyst
  - — consultancy;
- recurring costs;
- benefits;
  - — infrastructure savings
  - — staff costs.

Also included would be other conversion and service costs

Summary of costs

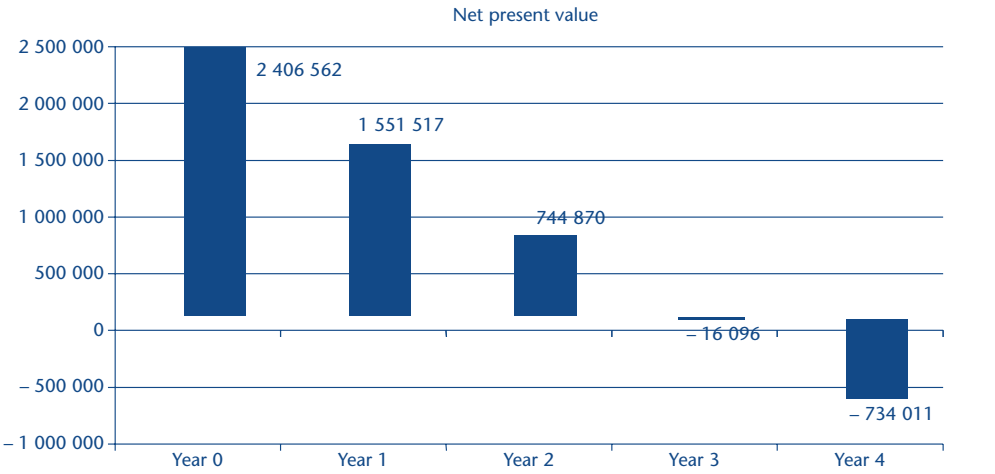
Item	Option 3	Option 5	Option 6
Systems costs GBP K			
Software	87	398	978
Hardware		40	140
File conversion		120	120
Back-file conversion of open paper files		480	480
Implementation	10	1 000	1 440
Total system costs	**97**	**2 035**	**3 158**
Annual costs			
Software maintenance	8	72	176
Hardware maintenance		7	25
Outsource paper files management and storage		118	118
Records-management administrator	25	56	56
Project management		90	90
Technical		90	90
Education/training	35	45	45
Total annual costs (000s)	**78**	**478**	**600**

The above costings were based on prior project experience and software licences based on the average costs of four EDRMS suppliers. Account was taken of the ongoing central support requirements, and training and the education programme for each business unit.

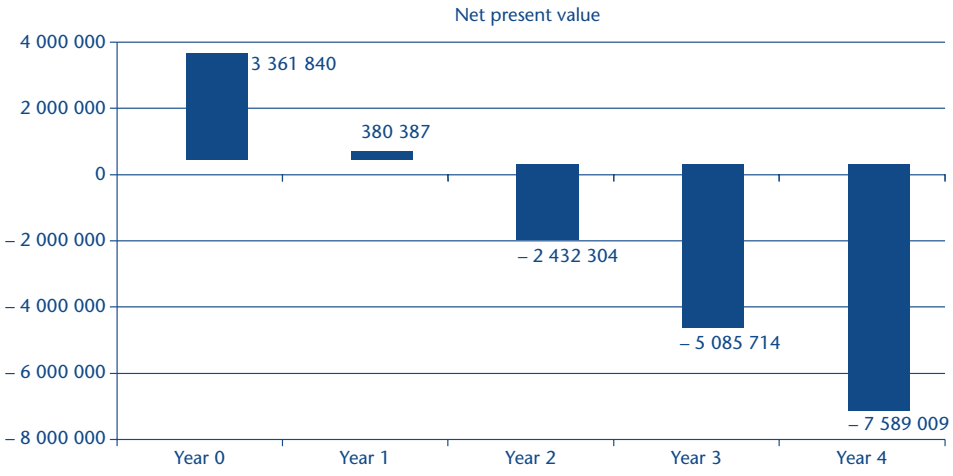
Two options were then selected for a more detailed financial modelling over five years with detailed spreadsheet setting out the costing sources and assumptions.

Financial models for Option 5 and 6 were constructed using the UK HM Treasury Greenbook guidelines using Net Present Value calculation over five years. The main cost savings used were productivity improvements based on current process metrics, other project experience and market studies, allowances were made for releasing floor space. Comparing the two models showed that Option 6 gave an earlier return on investment and a 10 fold in cost savings over the five years compared with Option 5.

Option 5 NPV model



Option 6 NPV model



Option 6 NPV detailed figures

Net present value	Year 0	Year 1	Year 2	Year 3	Year 4
A. Total system costs annual	2 845 654	0	0	0	0
B. Total cost cumulative	2 845 654	2 845 654	2 845 654	2 845 654	2 845 654
C. Total revenue costs (annual)	516 185	516 185	516 185	516 185	516 185
D. Total revenue costs (cumulative)	516 185	1 032 371	1 548 556	2 064 742	2 580 927
E. Total costs (annual)					
(= total system + revenue)	3 361 840	516 185	516 185	516 185	516 185
F. Total costs (cumulative)	3 361 840	3 878 025	4 394 211	4 910 396	5 426 582
G. Total benefits (annual)	0	3 676 513	3 676 513	3 676 513	3 676 513
H. Total benefits (cumulative)	0	3 676 513	7 353 025	11 029 538	14 706 050
I. Net undiscounted cost* (= e-g)	3 361 840	–3 160 327	–3 160 327	–3 160 327	– 3 160 327
J. Discount factor	1.0000	0.9434	0.8900	0.8396	0.7921
K. Net present cost* (annual) (= i*j)	3 361 840	– 2 981 453	– 2 812 691	– 2 653 411	– 2 503 295
L. Net present cost* (cumulative)	3 361 840	380 387	– 2 432 304	– 5 085 714	– 7 589 009

Step 6 — Non-monetary benefits

This section sets out the non-monetary business benefits associated with the implementation of the EDRMS system. Some of these benefits will be generic across all of the options under consideration; others may be associated with one option more than another. It is necessary to identify and describe these benefits in detail and profile them across the options in order to see the degree to which certain benefits apply to different options. This is particularly necessary when the principle justification for expenditure on the project is in the realisation of conformance to regulatory requirements or government targets.

EDRMS adoption across DETI for example will provide support and benefit to a wide range of existing services and processes as well as support government directives and legislation and other complementary strategies and developments.

A list of high level non-monetary benefits were developed based on discussion with the EDRMS project team, experiences from the EDRMS pilot, discussions with e-business strategy team and a review of their reports plus discussions with the knowledge management project manager.

The non-monetary benefit areas were grouped into eight areas:

- strategic
- e-business support
- modernising government
- resource skills and staff morale
- audit
- customers
- management information
- technical.

EDRMS' most powerful justification is the key role it plays in two of DETI's major development programmes. EDRMS has been recognised in DETI e-business and knowledge-management reports as a key component in developing their plans for e-business applications and knowledge management strategy.

The wider role is the support for the potential to establish NICS wide contact centres e.g. the establishment of a 'one-stop' shop for enquiries to the public sector which is already being considered by the central government business development services and complements the 'self-service' element of the DETI e-business strategy.

The other two key strategic areas ERMS supports is in addressing the modernising government requirements and the need to develop an information asset register required by the Freedom of Information Act.

Illustrated below are requirements for two benefit areas, how they will be measured and a weighting of the importance of EDRMS to these requirements.

(2) The weighting is the relative importance of EDRMS to achieving the requirement: 5 – key role to 1 — limited support.
(3) Target date for the implementation of ERM has not yet been clearly defined for NI.

Benefit area		Requirement	Measurement steps	Weighting (2) 1 to 5
Strategic	1.1.	Provide the prototype for NICS EDRMS	Regular consultation with the NICS ERMS Working Party	4
	1.2.	A major contributor to DETI knowledge-management strategy	Regular review with KN project and CITU	3
	1.3.	Assist in providing citizens with controlled access to information as required by Freedom of Information Act	Monitor against Freedom of Information DETI programme	3
Modernising government	3.1.	Assist in meeting ESD by 2005	As above	5
	3.2.	Provide support to meet electronic records deadline of 2004 (3)	As above	5

Each option is then assessed against each requirement in how well it supports the requirement in a scale of 3 — fully meets requirements to 0 — does not meet the requirement.

Benefit area		Requirement	Weighting	Option 3		Option 5		Option 6	
				Benefit level/weighted value for each option					
Strategic	1.1.	Provide the prototype for NICS EDRMS	4	1	**4**	3	**12**	3	**12**
	1.2.	A major contributor to DETI knowledge management strategy	3	0	**0**	2	**6**	3	**9**
	1.3.	Assist in providing citizens with controlled access to information as required by the Freedom of Information Act	3	1	**3**	3	**9**	3	**9**
Total score for all benefit areas					**34**		**136**		**228**
Strategic and e-government requirements							**88**		**175**

(4) Shown in green.
(5) Shown in red.

Sample extracts above give samples of the areas assessed and the total score shows Option 6 scores 70 % higher than Option 5 and for corporate/strategic areas, relating to e-Government and improving citizen access, that Option 6 was the clear favourite by a 2:1 ratio over Option 5.

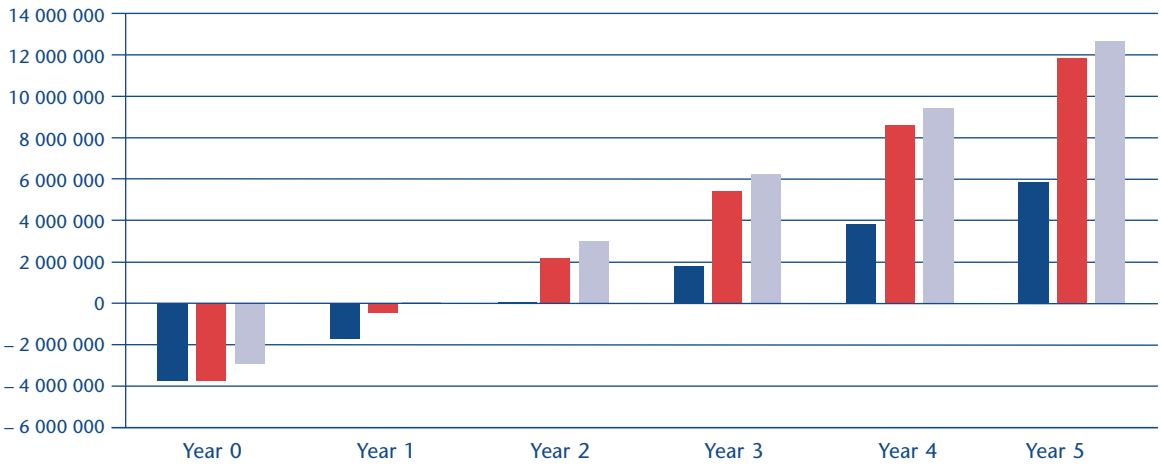
Step 7 — Economic assessment of risk and uncertainty

Risk analysis exercise was undertaken covering key areas including project management, resources, user acceptance, change of requirement. Extract table is shown below.

Risk area	Id	Detail	Probability	Severity	Countermeasure
Project management	1.1.	The handling of change management is very demanding	.7	.8	Invest in an intensive user education and training programme
	1.2.	Delivery dependencies can delay project. Implementation of the system is dependent on a range of preparatory activities including RM policy, housekeeping and conversion	.4	.9	Careful monitoring of these pre-implementation tasks need to be undertaken with proper investment in manpower and qualified resource

Sensitivity analysis

Sensitivity analysis was undertaken on the return on investment to assess the impact of overrun on costs and not achieving the level of estimated productivity improvements.



	Year 0	Year 1	Year 2	Year 3	Year 4	Year 5
■ Hardware, software and implementation +30 % Average productivity 10 %	– 3 997 786	– 1 980 746	36 293	2 053 333	4 070 372	6 087 412
■ Hardware, software and implementation +30 % Average productivity 15 %	– 3 997 786	– 775 746	246 293	5 668 333	8 890 372	12 112 412
■ Hardware, software and implementation +0 % Average productivity saving 15 %	–3 194 340	27 700	3 249 740	6 471 779	9 693 819	12 915 858

The ROI over five years shows a return on investment by the second year (4) and by the third year (5) if system costs were to overrun by 30 % and productivity reduced to 15 %.

The overall cost of implementing and running the system related to costing GBP 900 per person per annum.

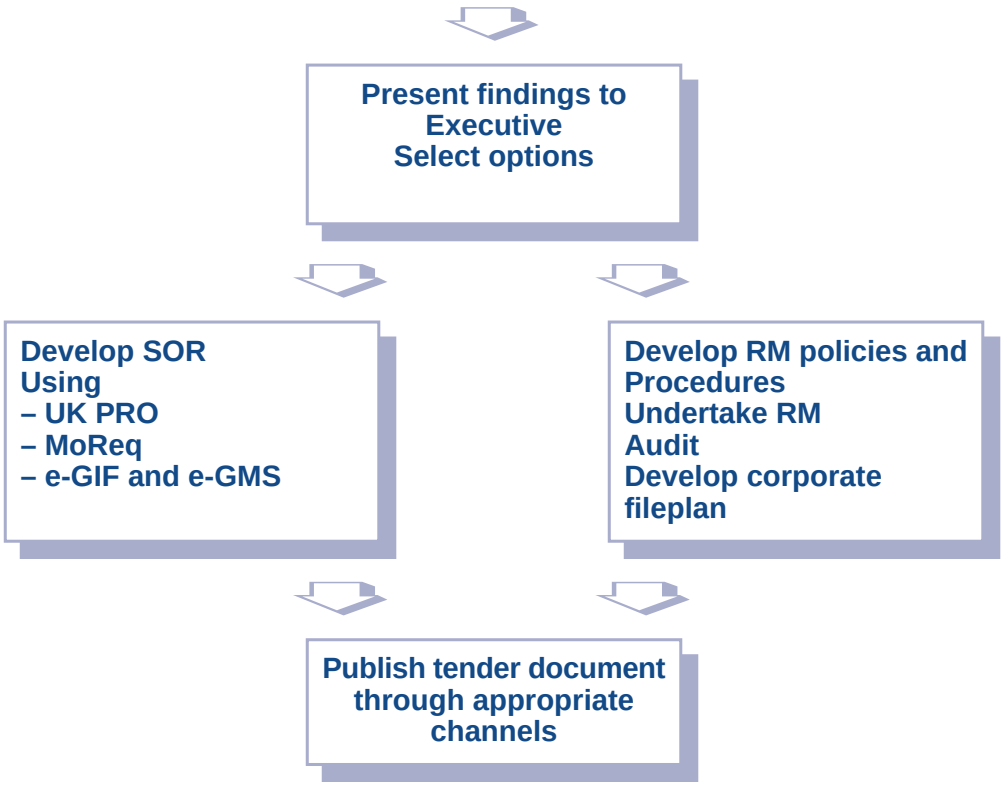
Allowing for the savings shown by the model plus departmental and NI budgets for FOI (freedom of information) and e-government the project has moved forward to project definition phase.

Implementation plan

This section outlines the first two phases of the EDRMS project and recommendations on the organisation of the project and support team for EDRMS having selected the preferred option.

The two parallel phases of the project were to:

1. Initial work that requires to be undertaken in order to develop an overall corporate records management policies and infrastructure.
2. The development of the ITT/statement of requirement could be progressed at the same time. An outline for this work and the procurement process are set out below.



Phase 1 RM implementation programme

Listed below were the steps to assist the DETI in gaining greater control over their recorded information. These elements needed to be addressed regardless of which option is chosen, as they are essential to providing structured access to information. Implementing any type of automated solution without seeking to order the existing systems would result in unsuccessful implementation and greater costs to clean up and reorganise the records and information stored in the new system.

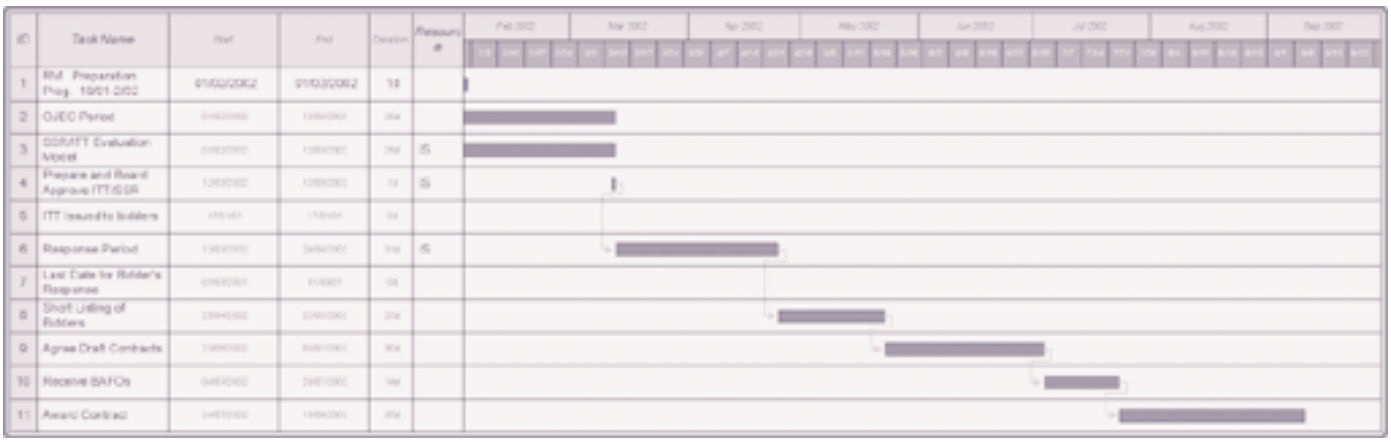
1. Conduct information audit to list and quantify all records and information systems held by DETI and its divisions.
2. Assess information inventory and other related materials (divisional organisation structure, functional activities etc.) to develop a consistent classification system.

3. Develop retention and disposal schedules.
4. Decide whether to retroactively convert files both paper and electronic to new system and file plan or simply start creating in the new system from a specified date. If the former, a significant conversion exercise will have to occur and this is likely to be fairly labour intensive.
5. At a minimum replicate the classification scheme in file directory structure so that there is some control over common desktop electronic records.
6. Develop a records management policy with elements as outlined in UK PRO guidance.
7. Develop standard procedures and guidance material for creating, registering, filing and retrieving paper records. This should include at least centralised registration and tracking of records if not centralised physical control.
8. Train administrative staff and records users on new manual procedures.
9. Conduct an appraisal exercise to clear out all older files, as there is no value in expending resources to manage inactive and unnecessary files. The aim should be to reduce the number of closed files held by 60 %.

DETI EDRMS
Phase 1 Corporate RM
Development

Phase 2 Procurement of EDRMS

The chart below sets out the phases of the project from publishing the ITT through to procurement of the system. OJEC route for procurement is allowed for.



Manpower and project organisation

The diagram below shows the relationship between the three groups that provide the management, implementation and support for the EDRMS project.

Implementation team

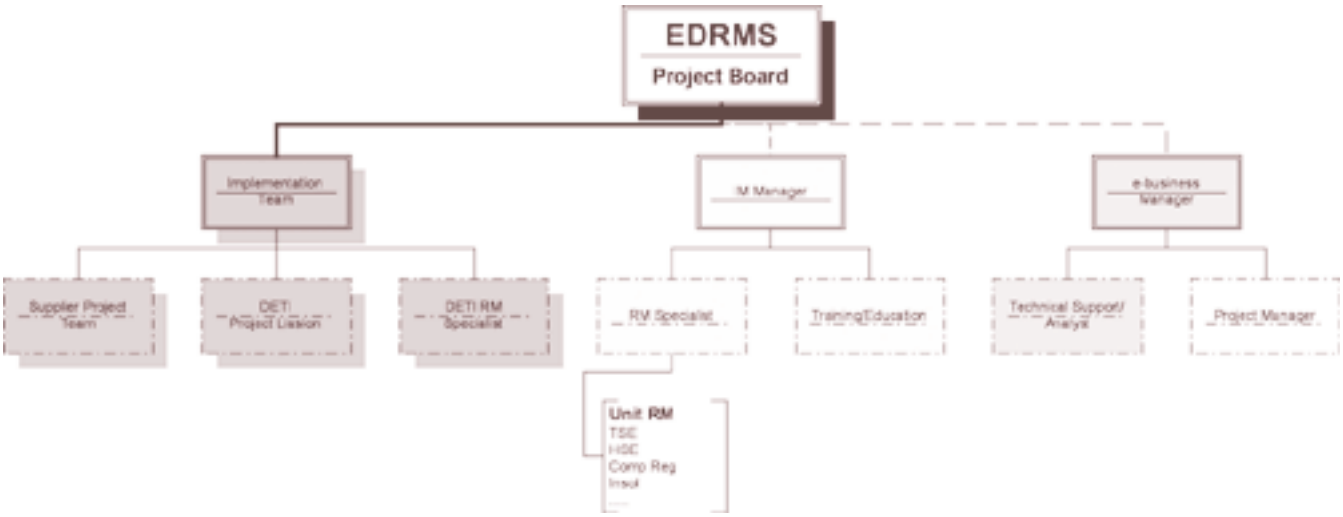
This will be a team made up of supplier project staff undertaking work as set out in the EDRMS implementation plan plus DETI project manager and records management/technical specialist.

DETI RM support and development

Two staff will be dedicated to supporting the staff and developing and administrating the EDRMS system. These were a RM professional responsible for the administration of the system and its development and a training post for user training and refresher courses.

E-business team

This team will be working closely with the supplier in developing specific e-business applications such as the company registry system.

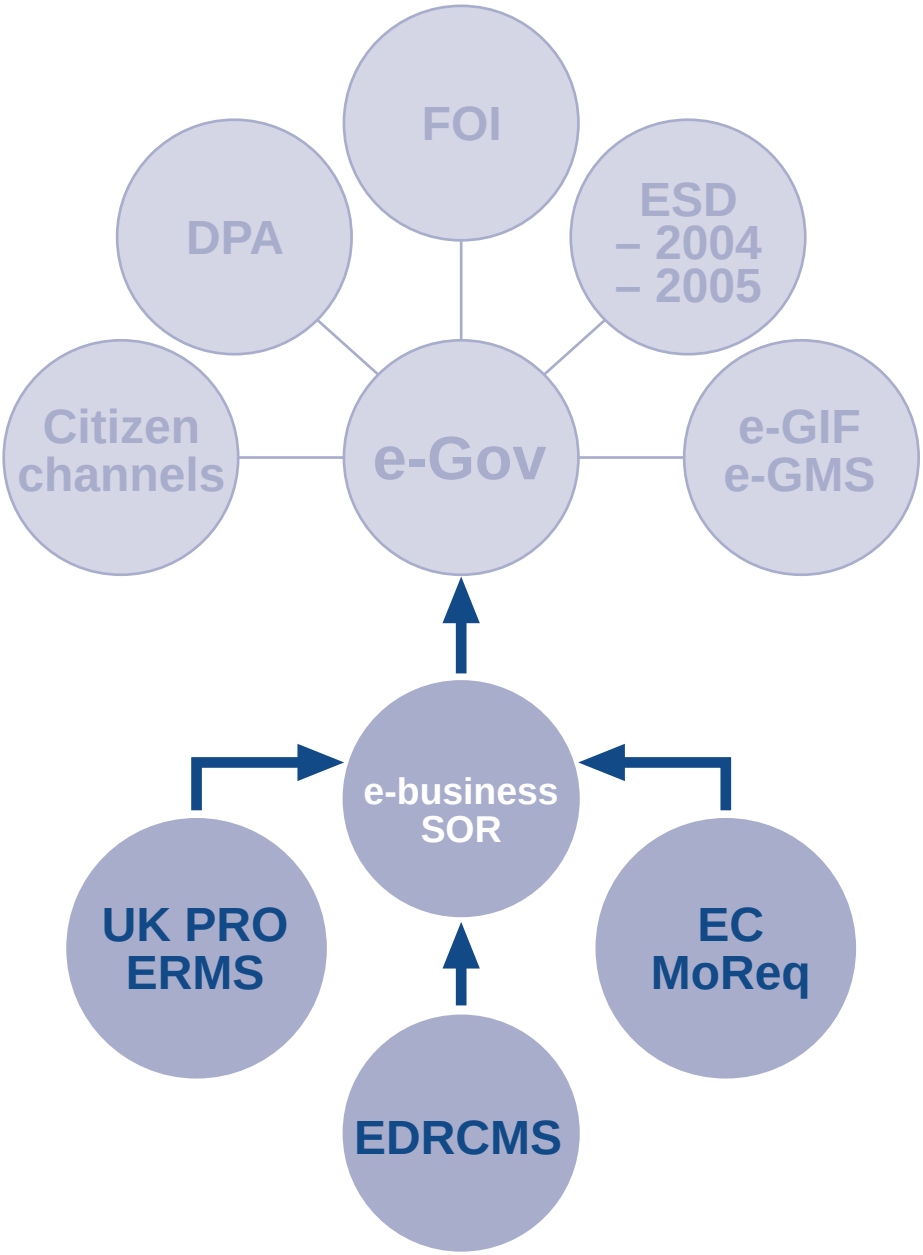


Next steps

The business case report once approved leads onto developing an adoption strategy for implementation. The diagram above summaries the relationship between the next step of developing the corporate e-business statement of requirement — SOR and the dependencies and constraints that need to be taken into consideration.

The SOR would be used to produce a project initiation document setting out the project plan and developing the tender document to support the organisation's e-government strategy. The UK PRO provide a range of documents to assist in developing the SOR (http://www.pro.gov.uk/recordsmanagement/eros/guidelines/) and MoReq provides similar information but also contains useful requirement functions and descriptions for electronic document record content management system — EDRCMS technology.

The top half of the diagram shows the constraints which will influence the specification of the system. These will have been taken into consideration in developing the business case such as FOI, DPA and the electronic service delivery — ESD dates set by the UK Government. The UK Government have all also stipulated interoperability standards for computer systems (e-GIF — gov-



(6) e-GIF.

ernment interoperability standard (6)) and a government metadata framework — e-GMF. Both these will need to be adhered in developing the requirement. The need to embrace not only new innovative channels for delivering services to the citizen but retain and improve current service approaches will need to be addressed. Also a marketing communications programme will need to be undertaken to promote and create awareness of new service channels, as well as consult with the public for their opinion on changes to services.

Summary

The following summary identifies guidelines and useful sources of information for moving the EDRMS business case forward to implementing electronic service delivery to support an organisations e-government strategy. These guidelines are based on projects which we have undertaken in the United Kingdom for both central and local government organisations.

- ERMS implementation must not be considered in isolation from corporate EDMS/workflow and knowledge-management strategy or e-business initiatives and national e-government requirements.
- ERMS will need a communications programme to assist the user in using and adopting ERMS practices. The less computer-literate will need support into not only using the systems but also

changing the mindset in accepting electronic records and not paper as the norm. This transformation will need to be addressed at board as well as operational level. The board will need to be briefed and convinced on legal admissibility of electronic records and that a corporate and not service/departmental approach is essential.

- The European Commission's report 'Model requirements for electronic records — MoReq (published in May 2001) document gives good guidelines on the functional requirements required for EDRMS covering both the generic EDRM requirements but also operational and management considerations
- Other useful guidelines ([7]) are:
  - — BS ISO 15489-1:2001 Information and documentation: records management. General — published in November 2001, this provides useful guidelines on policies and procedures.
  - — BS ISO/TR 15489-2:2001 Information and documentation. Guidelines — published in November 2001, this provides useful guidelines developing and implementing good record management practices.
  - — PRO — Guidelines on the management, appraisal and preservation of electronic records (1999).
  - — BSI-DISC PD 0008:1999 — Code of Practice for legal admissibility and evidential weight of information stored electronically.
- The Data Protection Act and FOI plus industry statutory requirements in relation to electronic records will be key drivers in initiating an ERMS project, and these will influence senior management support to a corporate approach.
- The EDRMS project is about transformation — joined up government, one-stops shops, multi-channel access to services, automating backend processes. There are major risks in such undertakings and an investment in a comprehensive business case will assist in reducing the risk not only in delivery of new services but also their take-up by the citizen.

([7]) See annex for details on documents.

Case study

DETI are now undertaking radical changes to how services are being delivered and to the organisations structure.

Restructuring of the organisation has been undertaken by DETI to reflect the vital role of ERM in support of e-Government and FOI. A new senior post has been created of an Information Manager reporting into the Chief Executive plus a major overall of how services are to be delivered.

DETI are already developing their first e-business application on the back of the EDRMS business case. Also Northern Ireland Civil Service as a result of this work are now evaluating a NI wide approach to EDRMS

Annex 1 — References

[1] Records Management, ISO 15489 (International Organisation for Standardisation)
Published by British Standards Institution (www.bsi-global.com)
As ISO 15489-1/2 – 2001E

[2] Model requirement for the management of electronic records — MoReq — published by the European Commission ISBN 92 – 894 – 1290 — 9
http://europa.eu.int/ISPO/ida/jsps/

[3] UK Government
e-GIF — government interoperability framework standard
e-GMF — government metadata framework
www.govtalk.gov.uk

[4] Functional requirements for electronic records management systems (Great Britain Public Record Office) and other guidelines on records management http://www.pro.gov.uk/recordsmanagement/eros/

[5] Code of Practice for legal admissibility and evidential weight of information stored electronically (British Standards Institution) Published by British Standards Institution (www.bsi-global.com) as BSI DISC PD 0008

[6] UK Treasury Greenbook Reference www.dfpni.gov.uk/economics_division/greenbook

Annexe 2 Abbreviations

Abbrev.	Full name
DETI	Department of Enterprise Trade and Investment Northern Ireland
DPA	Data Protection Act
EDRMS	Electronic document record management system
ERMS	Electronic record management system
ESD	Electronic service delivery
FOI	Freedom of Information Act

Orientaciones sobre la elaboración de una estrategia de gestión de los archivos electrónicos de empresa basada en la gestión electrónica: del modelo a la aplicación

Martin Waldron

Contexto

Las organizaciones han tomado repentinamente conciencia de que sus sistemas clásicos de gestión de archivos en papel ya no funcionaban. En el mundo del «.com», la mayoría de los procesos de trabajo, tanto frente a clientes y proveedores como internamente, se efectúan con creciente frecuencia por medio de transacciones y documentos electrónicos. Al fin estamos evolucionando hacia una oficina «sin papel».

Numerosas organizaciones grandes y pequeñas, del sector privado y público, siguen sin haber establecido mecanismos de control e infraestructuras que permitan gestionar eficazmente los documentos electrónicos que llegan con creciente frecuencia por correo electrónico o a los que se accede en la red, o incluso que se generan internamente. Por tanto, estos documentos se imprimen y gestionan como archivos clásicos utilizando procedimientos «papel», causando así una pérdida de eficacia y retrasos en su consulta.

Sistemas de gestión de los documentos electrónicos

Aunque los principios de la gestión documental están perfectamente definidos, su extensión a la gestión de los documentos electrónicos está aún en pañales. Los usuarios potenciales tratan de aplicar soluciones electrónicas que permiten tratar correctamente este tipo de documentos. Para colmo de males, hasta hace poco tiempo no existía una definición reconocida sobre la manera de gestionar los documentos electrónicos. Esta falta de buenas prácticas en el ámbito de la gestión de documentos electrónicos en muchas organizaciones se ha abordado en los tres últimos años a nivel internacional, europeo y nacional, traduciéndose en recomendaciones sobre buenas prácticas y normalización.

Necesidad de legislar en materia de sistemas de gestión de documentos electrónicos

Las disciplinas de la gestión documental deben cumplir los requisitos legales dictados por las directivas europeas, el Derecho nacional y los organismos reguladores. Las principales obligaciones a este respecto son organizar el archivo de los ficheros de carácter personal con el fin de ajustarse a la legislación sobre protección de datos y derechos de los consumidores, con la obligación suplementaria para el sector público de respetar la legislación sobre los archivos públicos nacionales y la legislación sobre la libertad de información.

Las iniciativas de modernización de la administración son los principales motores del desarrollo de los sistemas de gestión de documentos electrónicos, tal como pone de manifiesto la administración británica, que impone la informatización de todos los principales servicios públicos de aquí a 2005, y la obligación para las administraciones públicas de crear sólo documentos electrónicos de aquí a 2004.

Resumen

El presente documento se basa en la experiencia del autor, que ayudó a una serie de organizaciones públicas de la administración central y local, incluidos el Parlamento escocés, el Ministerio de Empresa, Comercio e Inversión de Irlanda del Norte (DETI), el Consejo de la región de Fife y la administración policial del condado de Bedforshire, a desarrollar un modelo de sistema de gestión de documentos electrónicos para apoyar su estrategia de administración electrónica.

El presente documento propone a las administraciones públicas un marco que les permita elaborar y adoptar una estrategia de gestión de sus documentos y archivos electrónicos, en el contexto de la instauración de procedimientos de administración electrónica destinados a mejorar y a prestar servicios más eficaces al público y a ajustarse a las disposiciones fijadas por el legislador.

La exposición presenta:

- las fases de desarrollo del modelo de sistema de gestión de documentos electrónicos, según las orientaciones fijadas por el Ministerio de Hacienda (Treasury Book Guidelines), ilustradas por la relación determinante entre la gestión de los documentos electrónicos y la estrategia de gestión empresarial electrónica y gestión de los conocimientos elaborada por el DETI;

- una descripción de las especificaciones tipo de los documentos electrónicos (MoReq), fijadas por la Comisión Europea, y de la manera en que estas especificaciones y las especificaciones funcionales de la Public Record Office (PRO) británica pueden utilizarse para desarrollar unos requisitos funcionales;

- una lista de las tareas preparatorias tales como la conversión de ficheros, con ayuda de un inventario de información que debe realizarse antes de la aplicación.

Leitlinien zur Entwicklung einer Unternehmensstrategie für die elektronische Aktenführung auf der Basis von E-Business: von der Geschäftsbewertung bis zur Umsetzung

Martin Waldron

Hintergrund

Bei Unternehmen und Organisationen hat sich plötzlich die Erkenntnis durchgesetzt, dass ihre herkömmlichen Systeme für Papierdokumente für die Verwaltung von elektronischen Aufzeichnungen nicht taugen. In der Dot-com-Geschäftswelt von heute werden die meisten Geschäftsvorgänge sowohl nach außen mit Kunden und Zulieferern als auch im eigenen Büro zunehmend über elektronische Transaktionen und Dokumente abgewickelt. Endlich sind wir wirklich auf dem Weg zum „papierlosen" Büro.

Viele Organisationen – große wie kleine, im privaten wie im öffentlichen Sektor – verfügen gegenwärtig noch nicht über die Kontrollmöglichkeiten und Infrastrukturen für eine effektive Verwaltung elektronischer Unterlagen, die in wachsendem Umfang per E-Mail eingehen, aus dem Internet kommen und intern erstellt werden. Daher werden Dokumente ausgedruckt und als Akten mit den vorhandenen papiergebundenen Systemen und den sich daraus ergebenden Unzulänglichkeiten und Verzögerungen beim Zugriff verwaltet.

Systeme zur Verwaltung elektronischer Aufzeichnungen

Obwohl die Grundzüge der elektronischen Aktenführung umfassend definiert sind, steckt deren Erweiterung auf die Verwaltung elektronischer Aufzeichnungen noch in den Kinderschuhen. Potentielle Nutzer kämpfen um die Umsetzung elektronischer „Lösungen", die eine sachgerechte elektronische Aktenführung ermöglichen. Erschwerend kommt noch hinzu, dass es bis vor kurzem keine anerkannte Festlegung für die Art und Weise der Verwaltung elektronischer Aufzeichnungen gab. Mit diesem Mangel an beispielhaften Praktiken für die Verwaltung elektronischer Aufzeichnungen (Electronic Records Management, ERM) durch viele Organisationen befasst man sich seit nunmehr drei Jahren auf internationaler, europäischer und nationaler Ebene und erteilt Empfehlungen zu besten Verfahrensweisen und Standards.

Notwendigkeit von Rechtsvorschriften für ERM-Systeme

Die ERM-Fachbereiche müssen gesetzliche Anforderungen erfüllen, die von europäischen Richtlinien, einzelstaatlichem Recht und Vorschriften von Regulierungsbehörden vorgegeben werden. Die wichtigsten Anforderungen betreffen die Einhaltung des Datenschutzgesetzes bei der Aufbewahrung persönlicher Unterlagen, Verbraucherrechte im Handel und – mit der zusätzlichen Anforderung der Einhaltung des National Public Records Act (Gesetz über öffentliche Unterlagen) im staatlichen Sektor – die gesetzlichen Bestimmungen zur Informationsfreiheit.

Die wichtigsten Impulse für ERM-Systeme gehen von den Initiativen zur Modernisierung der öffentlichen Verwaltung aus. Ein Beispiel dafür ist die britische Regierung, die die elektronische Erbringung der wichtigsten öffentlichen Dienstleistungen bis 2005 verlangt; staatliche Stellen dürfen ab 2004 Aufzeichnungen nur noch elektronisch erstellen.

Zusammenfassung

Dieser Beitrag beruht auf den Erfahrungen des Verfassers, die er gesammelt hat, als er einer Reihe von nationalen und kommunalen staatlichen Stellen (darunter das Schottische

Parlament, das Ministerium für Handel, Unternehmen und Investitionen von Nordirland – DETI, den Fife Council, die Polizei von Bedforshire) bei der Erarbeitung einer Geschäftsbewertung für ERM-Systeme zur Unterstützung ihrer e-Government-Strategie half.

In dem Beitrag wird ein Rahmen für öffentliche Organisationen vorgestellt, innerhalb dessen sie ein Konzept für die Verwaltung elektronischer Dokumente und Aufzeichnungen entwikkeln und einführen können, wie sie bei einer *e*-Government-Strategie für die Verbesserung und effektivere Erbringung von Dienstleistungen für die Öffentlichkeit sowie zur Erfüllung gesetzlicher Anforderungen notwendig ist.

So enthält der Beitrag

- die einzelnen Phasen der Geschäftsbewertung für ERM-Systeme auf der Grundlage der Treasury Book Guidelines, veranschaulicht anhand des Verhältnisses von ERM-Systemen zur E-Business- und Wissensmanagementstrategie des DETI;
- eine Darstellung der Musteranforderungen der Europäischen Kommission für elektronische Unterlagen (MoReq) und wie sie sowie die funktionsbezogenen Anforderungen des britischen PRO bei der Entwicklung von Funktionsanforderungen genutzt werden können;
- eine Checkliste der vor der Umsetzung zu erledigenden Aufgaben wie die Konvertierung von Akten oder die Durchführung einer Informationsinventur.

Guidelines sur l'élaboration d'une stratégie de gestion des archives électroniques d'entreprise basée sur l'e-business: du modèle à la mise en œuvre

Martin Waldron

Contexte

Les organisations ont soudainement pris conscience que leurs systèmes classiques de gestion des archives sur papier ne fonctionnaient plus. Dans le monde des «dot com», la plupart des processus de travail, aussi bien vis-à-vis du client et du fournisseur qu'en interne, s'effectuent de plus en plus souvent par le biais de transactions et de documents électroniques. Nous évoluons enfin vers le bureau «sans papier».

De nombreuses organisations, grandes ou petites, du secteur privé et public n'ont toujours pas mis en place les mécanismes de contrôle et les infrastructures permettant de gérer efficacement les documents électroniques qui parviennent de plus en plus souvent par courrier électronique ou qui sont récupérés sur le web, voire qui sont générés en interne. Il s'ensuit que ces documents sont imprimés et gérés comme des archives classiques utilisant des procédures «papier», occasionnant ainsi une perte d'efficacité et des retards dans leur consultation.

Systèmes de gestion des documents électroniques

Alors que les principes de la gestion documentaire sont parfaitement définis, leur extension à la gestion des documents électroniques n'en est qu'à ses balbutiements. Les utilisateurs

potentiels se débattent pour mettre en place des «solutions» électroniques permettant de traiter correctement ce type de document. Pire, il y a peu de temps encore, il n'existait aucune définition reconnue sur la manière de gérer les documents électroniques. Cette absence de bonnes pratiques dans le domaine a été abordée au cours des trois dernières années aux niveaux international, européen et national, se traduisant par des recommandations sur les bonnes pratiques et sur la normalisation.

Nécessité de légiférer sur les systèmes de gestion de documents électroniques

Les disciplines de la gestion documentaire doivent adhérer aux obligations légales dictées par les directives européennes, le droit national et les organismes de régulation. Les principales obligations en la matière sont d'organiser l'archivage des fichiers à caractère personnel en vue de se conformer à la législation sur la protection des données et les droits des consommateurs, avec l'obligation supplémentaire pour le secteur public de respecter la législation sur les archives publiques nationales et la législation sur la liberté d'information.

Les initiatives de modernisation de l'administration sont les principaux moteurs du développement des systèmes de gestion de documents électroniques, ainsi qu'en témoigne l'administration britannique, qui impose l'électronisation de tous les principaux services publics d'ici à 2005 et l'obligation pour les administrations publiques de ne plus créer que des documents numériques d'ici à 2004.

Résumé

Le présent exposé se fonde sur l'expérience de l'auteur, qui a aidé un certain nombre d'organisations publiques de l'administration centrale et locale, dont le Parlement écossais, le ministère de l'entreprise, du commerce et de l'investissement d'Irlande du Nord (DETI), le Conseil de la région de Fife et l'administration policière du comté du Bedfordshire, à développer un modèle de système de gestion de documents électroniques venant en appui de leur stratégie d'e-administration.

L'exposé propose aux administrations publiques un cadre leur permettant d'élaborer et d'adopter une stratégie de gestion de leurs documents et archives électroniques, dans le contexte de la mise en place de procédures d'e-administration visant à améliorer et à rendre plus efficaces les services rendus au public et à se conformer aux dispositions fixées par le législateur.

L'exposé présente:

— les étapes de développement du modèle de système de gestion de documents électroniques, d'après les orientations fixées par le ministère des finances (Treasury Book Guidelines), illustrées par la relation déterminante entre la gestion des documents électroniques et la stratégie d'e-business et de gestion des connaissances élaborée par le DETI;

— une description des spécifications types des documents électroniques fixées par la Commission européenne (MoReq), et de la manière dont ces spécifications et les spécifications fonctionnelles du Public Record Office britannique peuvent être utilisées pour développer un cahier des charges fonctionnel;

— une liste des tâches préparatoires, telles que la conversion de fichiers, à l'aide d'un inventaire des informations à réaliser avant la mise en œuvre.

An expanding universe — metadata and accessibility of digital information

Johannes Hofman

Introduction

The increasing use of information technology (IT) has changed our way of dealing with information. On the one hand IT offers us new ways of creating, using, and making available information and on the other hand it also requires new approaches just because of that, and because of the different nature of digital information. Nonetheless, to date, a huge mass of digital information resources is available e.g. on the world wide web, and is growing every minute. Accessibility in this respect is a crucial feature of this digital information. How can that be achieved and above all maintained? To what extent are traditional tools and approaches still sufficient? That are questions that many organisations are facing to date.

In this respect the issue of metadata is all over the place. Everybody seems to have discovered this subject. Especially in the world of the world wide web people are getting increasingly concerned in information resource discovery and in better organising the overwhelming amount of information that is available. It shows the growing importance of the world wide web, but we have to be aware that this is not the only domain where information is managed and maintained. Business companies, memory organisations, such as libraries, archives and museums, government organisations and so on, they all create and manage huge information sources and they all have to deal with issues, such as how to maintain them and how to keep them accessible and understandable. The world wide web is in this respect in most cases 'just' a channel for distribution and dissemination, be it an essential one.

Two perspectives can be distinguished here and that is the perspective of the creator and custodian or preserver on the one hand and the perspective of the searcher or user on the other. Both perspectives have to be satisfied in order to be effective in the area of accessibility. Metadata plays a crucial role in it.

The approaches concerning metadata that are taken can be small or very broad in scope. They range from specific sets of metadata in a specific domain (e.g. geographical data) to a (de facto) standard for information resource discovery, such as the Dublin Core metadata set, to frameworks that help organisations to organise and manage their information sources and make them available and accessible as in the case of record-keeping metadata standards. These different perspectives and approaches show us the scope and also the underlying complexity of the issue. Another approach in this respect is the industrial or technical view. It takes the possibilities of IT as a starting point, and provides tools like automatic indexing, full-text retrieval, fuzzy logic, artificial intelligence etc. The issue is then how do they fit into the picture, how do they contribute, or what problems do they solve?

There is a plethora of initiatives, projects and ongoing research in different domains that are dealing with the issues described and that creates another problem. How to coordinate all these efforts, that do not only take place in one domain, but at the same time in many different domains and communities and from different viewpoints? The capability of Internet and e-mail in connecting people may be a resourceful instrument in overcoming that problem, but still action is needed to cope with it. Some cross-domain collaborative projects are already emerging and will help in exchanging information about new or other initiatives worldwide and in stimulating new research.

Nonetheless, it is necessary to identify how all these initiatives and their underlying questions and answers relate to each other. In the end it has to be the human being, in the quality of user, that should benefit from the results of all the work being done.

Hans Hofman

Hans Hofman (1948) has studied history, archival theory and informatics. Since 2000, he has been working as senior advisor for the government programme 'Digital longevity' on information management at the Ministry of the Interior (Department of Public-Sector Information Policy). This programme was initiated in 1996 by both the Ministry of the Interior and the Ministry of Education, Culture and Science, with the objective of creating appropriate conditions for electronic record-keeping within government organisations and archival institutions. In this position he is, among other things, involved in developing guidelines for digital record-keeping, metadata sets and in formulating strategies for digital preservation.
On the international scene he is (since 1993) a member of the Committee on Electronic Records of the International Council on Archives.
Within the European Union he is a member of the DLM-Monitoring Group and as such involved in activities to stimulate cooperation in the field of electronic records. He is also co-director of the recently started European project Erpanet (electronic resource preservation and access network) on digital preservation.
Furthermore, he is investigator and representative of the National Archives of the Netherlands, participating in an international research project initiated by the University of British Columbia (UBC), called the InterPARES research project (1998–2001), which has the objective to investigate the long-term preservation of the integrity of electronic records.
He is since 2000 representative for the Netherlands in the TC46/SC11, working on ISO Records Management Standard 15489.

In this paper I will try to identify the main issues relating to metadata, the different perspectives taken in different domains and related projects, to some extent discuss how they relate to each other, and what possible approach can be taken. The challenge is to achieve better coordination and to identify ways forward.

What are the issues? Perspectives and purposes

In order to be able to understand the complex area of metadata and the role they play, it is necessary to identify what the (main) issues are and what is being pursued in all these different projects. In general one could say that depending on their point of view these efforts aim at discovering, disclosing and retrieving information, at enabling understanding and interpretation, or thirdly at enabling management and preservation.

The essence of metadata is that they provide us with the necessary information to understand and use information. It starts for example already with the need for communicating. An e-mail message needs information to whom it has to be sent or from whom it is coming, and what the subject is. We may also want to send a copy to somebody else. As soon as we receive a message we want to keep information about when it was received (date and time), and whether it was a reply or just a first message. Other information that may need to be kept is, where we have stored it and under what number and/or name, what happened since then (did it stay unchanged), how it is related to other messages or other documents, and so on. It is all metadata.

Some people say that a reason for collecting metadata about digital objects is, that it is easier to handle a small set of metadata than the objects themselves. That is one argument, but there are other more important principles.

To be able to use information resources it is necessary to satisfy at least three basic requirements. These requirements are: to be able to find the information, subsequently to be able to interpret it and finally to know whether the information can be trusted or not ([1]). Depending on the domain in which it is created and used these requirements may differ in strictness. In an organisation that is vulnerable to lawsuits requirements for authenticity and reliability of information or records will be severe, but in case information is retrieved for more informal reasons the requirements of authenticity and integrity will be less, but those for retrievability, interpretability and meaningfulness will still be important.

It is clear that in relation to the world wide web much attention is being paid to information discovery and retrieval and with reason. After all one of the main goals is to find and use information. Retrievability however only makes sense if all the basic requirements are satisfied. In the following section the basic requirements are further characterised.

Retrievability

The main thing is to be able to find information, but above all information that is relevant to you, or that answers your question. That is the ultimate goal of the work being done in all kinds of organisations that make information resources available.

Searching information on the world wide web is enabled through search engines, which allow people to enter keywords representing subjects they are interested in, hoping they will find information resources that will satisfy their needs. The well-known issue here is the huge amount of hits that are in most cases returned or provided to the user. Although there are mechanisms that value the information resources to what extent they comply with the keywords and arrange the information sources accordingly, there are also many mechanisms that arrange the sources according to other criteria, such as the amount of links or the occurrences of keywords on web pages. The amount of tricks and deceptive methods used here is huge however and makes one suspicious about the retrieval results. Some people applaud or rely on the possibility to search on the content of documents (full-text retrieval) as an easy and useful way of finding the right information. They easily forget however that the use of words in documents is not controlled by any mechanism, except for the human mind, and that has proven to be very unreliable and inconsistent in this respect. So using this method offers hardly any consolation, but probably more frustration. Perhaps in the best case it may be used as an additional tool for retrieving information resources.

([1]) In a document, written in 1999 by the *Description and classification of government records Working Group* of the *Information Management Forum,* in Canada these requirements are also mentioned as the aims for archival description. The document is titled: 'Approach to the description and classification of government records'. (www.imforumgi.gc.ca/new_docs/draft_e.html).

As such this is not a really effective way of searching. Which requirements in relation to retrievability should be satisfied? Essential is that objects have to be identifiable and to be located. A very important issue in this respect is the unique identifier that has to be persistently linked to the object. The uniqueness is dependent on the domain in which objects will be used and can or should be extended to all domains, certainly if we talk about the world wide web. Mostly an identifier will be unique only within the domain in which the document or information object has been created, not for other domains. The well known problem with the current identifier, i.e. URLs is that they are not reliable for finding information, since they are location-oriented (e.g. web domain name). When that location changes, and that happens often, the web address changes. Efforts are being made to solve this problem. Examples are URN and DOI initiatives. Whether they will succeed, is still a question.

Apart from being locatable it is necessary to structure or organise and describe the information objects. Different methods are available for doing so. Examples can be found in libraries (mostly subject-based, e.g. UDC) and archives (mostly function- or activity-based, e.g. business classification schemes). The structure establishes intermediate levels which allow better navigation and make retrievability easier.

Understandability

As soon as the information object has been found and presented, one has to be able to understand and interpret it. Otherwise we cannot value the content in relation to the question we have. That means that information about the origin has to be available, e.g. why and how has it once been created and used, does it have a relationship with other documents or information objects and so on. It might be that not all of this information is necessary for each question, but it has to be available. This meta-information (metadata) has to be described and linked to an object. Issues that may play a role here are, is it written in a language we understand, do we understand enough about the background of the object — which issue may become very important if the object has been created a long time ago — or is the information resource coming out of another domain of knowledge or activity, that we are not familiar with. It all requires metadata that has been captured at its creation or may be added afterwards. These metadata can be embedded in the object itself or can be external or both.

Trust

Finally if we have found an object and are able to interpret it, we still need to have a feeling about the trustworthiness of the presented information. In a digital world this has become even more necessary, because everything in cyberspace is information originating from all kinds of sources, and worse, because digital information has no fixed form and as such is very volatile and vulnerable to alteration or mutilation. In the case of the world wide web this is aggravated by the fact that it is not a controlled environment. So the information has to be taken care of continuously behind 'the screens'. If not, it will be difficult to prove that it has not been tampered with. In some cases that may not be an issue, but in many cases it will, especially in the case of lawsuits or research for instance.

The history of how the object has been managed (can be e.g. an audit trail) is essential for proving that the information presented, is not corrupted and therefore trustworthy. Proper management requires all kinds of procedures, methods and measures that ensure a safe environment for digital resources. In the InterPARES project efforts have been made to identify the requirements that contribute to or provide trust and how these should be implemented in preservation systems in order to achieve the desired outcome: trust in the results presented based on the search and reproduction methods. Tools as digital signatures can be used as additional mechanisms that help to ensure this trustworthiness. Several other initiatives are under way to identify attributes and functionality for ensuring reliability of information, such as the ISO Records Management standard and the RLG report on 'Attributes of a trusted digital repository' ([2]).

The basic requirements as discussed do not fundamentally differ between different domains. The emphasis may be different or the strictness of the requirements, but in fact all three sets of basic requirements have to be accomplished always. Where in records management or archival com-

([2]) RLG/OCLC: 'Attributes of a trusted digital repository. Meeting the needs of research resources', draft for public comment (August 2001). See www.rlg.org/longterm. ISO Records Management Standard 15489 represents in itself requirements for such a trustworthy environment. The InterPARES project as mentioned has gone into its second project (2002–07) to push issues around authenticity and integrity especially of records further and to provide guidance (www.interpares.org).

munities these goals or compliance to these requirements are being pursued in carrying out records and archival management, in other domains such as libraries this management area is mostly called (digital) 'preservation'. The difference is that the latter focuses on the continuing usability and availability of digital information objects, being mostly electronic publications and more and more web resources. Record-keeping or records management is the discipline that deals with records from their creation on for as long as they are needed, which can be for 'eternity'. As such it can be considered to have a broader scope. Appraisal and describing the interrelationships between records are core activities.

At the moment these approaches seem to be more or less competing with each other, because communication between the two communities does not really exist, only on ad hoc basis. There is a predominance of the library community. Whatever the reason for this may be, in order to improve information exchange or cross domain searching it is necessary to find ways to identify the commonalities, to achieve more synergy by using skills from different domains, and see how these can be used as a basis for further research and development. Besides, there are other communities, such as research institutions with scientific data, cultural heritage or industry (e.g. pharmaceutical industry) trying within their own domain to find solutions as well. More openness and information exchange in this area is needed, in order to learn from each other. That is also necessary for another reason, because apart from the above mentioned basic but essential requirements there is another relevant issue, called interoperability. In the openness of a networked environment such as the world wide web it is necessary to coordinate the efforts to improve communication and exchange of information between domains. However there is the issue of different semantics. A publisher in a library environment uses a different terminology than people in public administration or e-government for instance. It is the unavoidable problem of similar terms in different contexts for different concepts or of different terms for similar concepts. These different domains have each their own perspective and domain bound terminology and that will lead to different metadata sets. How to reconcile the different interests or perspectives? What solutions or approaches are available?

In the different domains where this issue is addressed, there is a growing awareness of this interdependence. So initiatives are developed to learn about what is happening elsewhere and to develop instruments that enable interoperability. Examples are for instance the Harmony project that aims at achieving semantic interoperability between different sets of metadata in e-publishing so users can search (electronic) publications on the web uniformly, and the European Schemas project that tries to establish an information service that provides information about the different metadata schemas that are developed and how they relate to each other. It also tries to map them ([3]).

Both projects build heavily on Dublin Core developments. It is remarkable, however, that developments in the area of records management or record-keeping until now hardly have been taken into account. That might have two reasons: there is nothing to report on in this area and/or what is there does not fit the needs of the community or at least that is what people think. A third reason may be the fact that this area is not well known to the outside world. I guess it is a combination of the latter two.

Longevity or preservation

There is another dimension, if we extend interoperability in time. It will mean that we have to take care of the ongoing (technical) readability and the meaningfulness of the information objects and that concerns the area of preservation or management of information resources over time. Both aspects of preservation, intellectual and technical, have to be addressed and need permanent maintenance in order to keep the information resources involved accessible and understandable. The same goes for the meta-information about them.

Theory, methods and practices

The next step will be how to implement the abovementioned requirements. Summarising the above paragraph the following aspects and activities can be distinguished:

(1) description of the information resources (either publications or records or data sets);

([3]) See for Harmony project www.ilrt.bris.ac.uk/discovery/harmony and for Schemas www.schemas-forum.org and www.cultivate-int.org/issue3/schemas.

(2) persistent identification through time and across domains;

(3) ongoing contextualisation to provide meaning to information sources through time;

(4) interoperability between different sets of metadata used in different domains;

(5) standardisation at different levels (as regards to e.g. structure, semantics, value and/or content).

Although these aspects and areas of research are not mutually exclusive and the overview is not exhaustive, it shows the complexity of it.

Instead of discussing all these issues in detail I like to take a slightly different approach and try to identify a core set of activities to make and keep information accessible. Basically the following instruments or methods can be distinguished:

- Appraisal and selection of what will be preserved. This is always necessary, in order to preserve no information that is not needed and has no value. Inherent to this is the necessity of disposing of information as soon as it is no longer needed (clearing up).
- Structuring or clustering of information resources according to certain criteria. In the case of records this structuring activity is done mostly based on business needs by using for instance classification schemes and has two reasons:
  - — to cluster the documents or any other information entities that are interrelated (in business processes these documents are evidence how a certain case has been handled for instance). That can be done almost automatically, since the documents reflect a business process and the way a case is processed. The objective is to articulate or express the documentary context and to maintain the coherence between documents ([4]);
  - — to cluster information around subject/business activities in an organisation that created the information. This is mostly done at a higher level. In libraries structuring takes place mostly based on subject classification.
- the third main instrument is 'describing' the information objects by 'labelling' them, e.g. through a classification scheme, and by making them identifiable. For records this description can be largely derived from the activity that creates them. As indicated one of the main areas for adding metadata is also to provide contextual information on information resources as in the case of records. That makes it possible to understand and interpret them. These metadata provide information about the origin or provenance, nature, state, content, structure and access of an object so we can establish or assess the value, reliability, authenticity etc. Essential is to keep the resources meaningful through time.

In both the record-keeping and archiving community and the library community these methods are well known and used since long. The application of them will be different because of the different nature of the material in custody.

Different approaches are under way in order to deal with these issues. It is hard to distinguish the many different initiatives in different domains and with different perspectives and to keep up with new developments. As already indicated they take place mainly in the area of information-resource discovery. Some of these approaches are object oriented, some are process-oriented, some are function oriented or even a combination of these.

Examples of object-oriented approaches are for instance information resource discovery initiatives such as Dublin Core, as well as the European MiReG project (managing information resources for e-government) ([5]). Function-oriented approaches are the Open Archives Initiative (OAI), trying to achieve interoperability in the publishing area, or Cedars, the research libraries group (RLG), and the open archival information system (OAIS) in the preservation area ([6]). Others are the ISO Records Management Standard 15489 and its consequences for metadata, and the archiving metadata forum (AMF) in the area of record-keeping, though they may also be called (business) process oriented ([7]).

One of the instruments that is being introduced to serve as a describing mechanism, is the Dublin Core metadata set which offers some control on retrieval of information resources. It allows to add information to a document, publication or web page that is not within the document itself for

([4]) This is called the 'archival bond'. See Luciana Duranti, www.slais.ubc.ca/users/duranti.
([5]) MiReG: http://ag.idaprog.org/Indis35prod/doc/312 and http://dublincore.org/groups/government/mireg-metadata-20010828.shtml.
([6]) See for OAIS: http://ssdoo.gsfc.nasa.gov/nost/isoas; Cedars: www.leeds.ac.uk/cedars; RLG: www.rlg.org/longterm.
([7]) ISO RMS 15489 is produced by ISO TC46/SC11. Archiving metadata forum: see www.archiefschool.nl/amf.

instance. In the world of the Internet it seems as if the Dublin Core metadata set is the only tool available. With this 'hammer' some people try to make a 'nail' of everything. Notwithstanding its merits and the fact that it has been adopted worldwide as a *de facto* standard for information resource discovery, it has however its limitations as a publication and resource discovery-based instrument. That is shown for instance in attempts to extend this standard with new elements to satisfy other requirements, such as in the case of the Australian government locator system (AGLS), where four elements are added to the original 15 DC elements. Unfortunately there exist also ideas to extend this set with record-keeping elements. This approach is completely denying the complex and different nature of record-keeping metadata.

A rather recent and interesting development is the emergence of the semantic web and with it ontology. Ontologies describe and structure terms of a certain domain of knowledge into a controlled vocabulary, and as such structure the metadata into hierarchical structures/classes, subclasses etc. It provides a tool for identifying the relationships between terms in that domain. This approach of adding semantics to information resources will enable intelligent agents to search much more efficiently and effectively.

Apart from these more sophisticated approaches industry provides as indicated tools such as automatic indexing. They promise to make things easier, but the question is do they really? Such techniques are dependent on the use of words and terms in documents and as can be expected that is not consistent. The result of automatic indexing, how intelligently it may be done, can never be as good as deliberate and structured metadata creation, capture and management.

An important issue in this respect is to ensure that adequate metadata is generated at the moment of creation of the document or source, that is captured and maintained in a useful way and persistently linked to it. A proactive approach is much more efficient than a retrospective one, in which case information resources are labelled or 'manipulated' when they already exist. The proactive way might be 'easier' or perhaps more natural to achieve in a business activity environment than a more open environment as for instance that of publishing (books or web pages), but it is essential from a cost-effective point of view.

If the requirements for metadata are clearly identified, it will be possible to develop software tools that in the case of records will enable automatic capture of these metadata from the business system with which the records were created, itself or from closely related systems, such as workflow systems. By integrating metadata capture in software applications the cumbersome task of gathering metadata that will give meaning to information sources of whatever kind will be more easily accomplished.

In this respect it is interesting to see what is happening in practice, for example in e-government initiatives. Several governments are now trying to establish standards and frameworks for metadata, in order to make interoperability between government organisations possible. These standards so far focus mainly on information resource discovery. However there are also some interesting examples to establish and include recordkeeping metadata sets for government agencies, such as in Canada, Australia and the United Kingdom. Each of them tries to identify a minimum set that should guide government organisations in managing and maintaining their records.

The Canadian 'record-keeping metadata requirements' are produced in January 2001 and consist of a minimum set of 26 metadata elements that allows organisations to describe and share information and meanwhile facilitates interoperability ([8]). Eleven of these elements overlap with the Dublin Core metadata set. The Australian Record-keeping Metadata Standard for Commonwealth agencies (RKMS) consists of 20 elements of which eight are mandatory. These sets of elements however are very high level and need further refinement and explanation with sub-elements and qualifiers to be implemented and used.

Government organisations are free to add these specific elements or sub-elements. In general the Australian minimum set consists of three parts, describing the organisation, the document or record and the management history of the record respectively. The set however is focusing on records, and not describing the full context of it ([9]). With this set though agencies at least know what metadata they should capture in their record-keeping systems.

([8]) See for the Australian RKMS: www.naa.gov.au/recordkeeping/control/rkms/summary.html and for the Canadian set www.im-forum.ca.

([9]) The set is currently being review in order to improve and adapt it based on the experiences so far.

Although the RKMS intends to describe the records and is not focusing on retrieval of records it is the idea that it should be in line with information resource discovery metadata sets as much as possible. In the case of the Australian RKMS a strong relationship and overlap exists with the abovementioned Australian Government Locator Service (AGLS). This set is accompanied by the Australian Government Interactive Functions Thesaurus (AGIFT), that is an addition to the function element of the AGLS standard and provides a controlled vocabulary for describing government functions, which creates a strong link between information resource discovery and record-keeping ([10]). This approach reflects awareness that it is necessary to link both worlds, a view which sees resource discovery metadata as a subset of recordkeeping metadata.

([10]) See also Cunningham, A., 'Six degrees of separation: Australian metadata initiatives and their relationships with international standards', *Archival Science* (Vol.1, No3, 2001) pp. 271–283.

([11]) See Lagoze, C., Hunter, J., Brickley, D.,'An event-aware model for metadata interoperability', 2000.

The third example is the 'E-government framework for metadata' in the United Kingdom as published in 2001. It recognises the need for standards to ensure consistency in effective information management and intends to provide a framework for government organisations for dealing with resource discovery and records management. The main objectives are to enable effective search through metadata instead of the resources themselves and to make people confident that the retrieved source being presented is the best one. The framework introduces the Dublin Core as the accepted standard, though admitting that this will not be sufficient, and envisages also the development of a rather ambitious pan-government thesaurus.

These examples show the increasing awareness that a broad metadata framework is necessary, which includes both information resource discovery and management metadata. It will not only make better communication between government organisations possible, but also if properly implemented compliance with requirements on trust and understandability.

The fact that metadata tags, such as provided by the Dublin Core set, are meant for information resource discovery, makes it possible to identify the possible overlap with record-keeping metadata, which are mainly focusing on the enduring interpretability, authenticity and integrity of a specific set of information resources, namely records.

Towards a common framework for metadata

The basic notion that can be derived from the previous paragraphs is that two different, seemingly contradictory viewpoints can be distinguished:

(1) the needs of a user or (re)searcher, including reuse of information resources for other purposes than they were created; and

(2) the need of managing and maintaining information sources in order to keep them trustworthy and understandable.

It also seems as if different communities are taking care of each of these perspectives, i.e. the library community on the one hand, and the records management and archives community on the other. This is a general picture of course, which is not completely justified by practice. Apart from that within these and other communities many different metadata sets or standards exist, which make it even more complicated. One of the things necessary is to map these sets and see how they can be connected. It adds another meta-level, but it is an illusion to think that there will be one common metadata set shared by all communities. The consequence is to identify at what level the existing sets can communicate with each other and develop a conceptual framework for it. An example of such an initiative is the ABC-model that is being developed in the Harmony project ([11]). It identifies a set of entities that is common in many metadata sets in different domains and intends to provide a general logical model that describes them and their interrelationships. These entities regard people, organisations, places, events etc. The model focuses on events and the basic idea behind it is that information resources can evolve or transform over time by events, e.g. a translation into another language of an information resource. That event then influences the description of the resource, because one or more properties of the resource have changed. Events are connected to agents, dates, places etc.

This approach is interesting because there is a parallel to the creation of records. Records are the results of activities carried out by agencies in doing business. A metadata model based on that notion can be found in the SPIRT-model, as developed by Monash University. That high-level

model identifies three basic entities, agents, business and records, that can have all kinds of relationships ([12]).

The basic scheme that follows out of the previous models and remarks is that people carry out activities or do business which results in information resources (records or publications). This perspective is taken especially by the SPIRT model. The perspective taken by ABC-model is that information resources can or will be transformed by events, carried out by agents. Despite these different viewpoints there is a strong overlap in entities. And there is a third perspective, the viewpoint of the researcher or user. The question asked by the user is mostly based on who did what when, where and/or why? As such this question or part of it can easily be related to information resource creating activities.

The following diagram shows at a high and simplified level the relationships between the above mentioned entities or elements.

What we need to know in order to be able to fulfil the requirements of retrievability on the one hand and of interpretability on the other hand can be identified as rather simple: who, what, when, where and in some cases why. The 'who did what why' provides us with the information about provenance and identity. At the level of the information sources itself (publications, web pages, or records) metadata on their management (activities such as appraisal, maintenance, description, access etc.), preservation and use have to be captured. The elements of when and where or time and space are applicable to the whole, because agents or organisations and activities including their interrelationships will change over time.

A similar approach is for instance used in the electronic resource citation (ERC) ([13]). This citation idea is based on Dublin Core elements and intends to provide a metadata kernel with a very simple format (who, what, when, and where), that should support the permanence of network discoverable objects. The approach though is rather static and does not take into account the dynamics of metadata description.

Figure 1: Basic entity model

Nonetheless there seems to be a common basic scheme that can be used for different purposes. If we are able to build a model around these elements, we will have a solid core set and the most important needs can be satisfied. This can in principle be applied to all kinds of information sources.

Concluding remarks

The globalisation of the information world has a strong impact on the way how we manage it. The emergence of the world wide web requires new approaches and methods in order to enable easy access and accessibility. Metadata play a key role in this. At the moment there is a predomi-

([12]) See for the SPIRT model, for instance, www.sims.monash.edu.au/rcrg and McKemmish, S., Acland, G., Ward, N., and Reed, B., 'Describing records in context in the continuum: the Australian record-keeping metadata scheme', *Archivaria* 48 (Autumn 1999).

([13]) Kunze, John A., *A metadata kernel for electronic permanence*, JoDI (Journal of Digital Information), A special issue on metadata: Selected papers from the Dublin Core 2001 conference, (Vol. 2, Issue 2, January 2002). See: http://jodi.ecs.soton.ac.uk/articles/v02/i02/.

nance of approaches from the library community. The tendency or even the need, based on the globalising effects and the openness of the world wide web, towards better coordination and collaboration between different information providers and communities requires other attitudes in these communities. This is especially in the archival community achieved rather slowly.

In order to be present, active and effective in this new virtual world it is necessary to be aware of the characteristics of it. In the area of metadata the world of information resource discovery has to be better linked to the world of management and preservation of information resources. This means that sets of metadata have to be mapped with each other, but also that there has to be a better understanding of the different perspectives that exist. Only then will it be possible to achieve the required interoperability. As indicated there are some promising initiatives in this respect.

One question in making information resources accessible in a digital environment is, do we need metadata or should we make use of the possibilities of IT or software to search the content of these sources? Isn't one of the big benefits of IT that computers made search on content of documents possible, while before it was not? Obviously, looking at what is happening at the moment as regards all existing metadata initiatives that seems not to be sufficient at all. The arguments used are the need for reliability, interpretability and interoperability of information resources of all kinds. In this respect the idea of automatic indexing is not relevant, because it does not address these requirements. Moreover, this tool is insufficient in order to deal with the different sets of words, the different semantics, the inconsistency in the use of words etc. in information resources and as such it does not contribute to or solve the issues mentioned. At the most it may provide an additional help. The same goes for other tools as full text retrieval or fuzzy logic tools.

Guidance is needed to find one's way through the dense forest of existing projects and initiatives around metadata in all its forms, and to understand how they relate. In this paper I have tried to identify and describe a possible common concept that could be a solid basis to build on. This concept is in line with the needs of both the creation and the use of information resources.

It is also clear that there are more needs to be served than information discovery and more approaches possible or available that may play a role on the scene of retrieving and managing information. Especially one discipline may bring relevant and useful experience and approaches to other communities and that is the archival community, which for centuries has been and still is very familiar with managing, preserving and making information sources accessible. Although common knowledge among records managers and archivists, it becomes more and more obvious that outside this specialist and small community not many people are aware of that.

So on the one hand there is a community that focuses on information resource discovery and is seeking for a common instrument for making information resources retrievable and searchable by establishing a limited set of tags. On the other hand there is the archival community (including records management) that captures, organises and manages a specific category of information, being created in the course of doing business, called records.

They are complementary, and bring different but relevant skills to the floor. Both perspectives are necessary to help people to retrieve information easily, to assess what the information is about, and whether they can trust it and interpret it. They meet in the domain of the world wide web, but they still have to connect properly, so synergy can be achieved from what they each are trying to do. Coordination can be improved. The same goes for collaboration with software suppliers that can provide useful tools based on the identified requirements.

This paper I hope also makes clear that adding metadata is not a useless burden, though it has to be cost-effective and user-friendly. One could see metadata as the 'value added tax' (VAT) of information. It may be a costly thing and it may be experienced as a burden, but more importantly, it has or should certainly provide added value. So it would be better to speak of metadata as 'value adding tags'. It is up to the specialists, both information and IT, to make it easier e.g. by the use of IT.

Un universo en expansión: metadatos y accesibilidad de la información digital

Johannes Hofman

Al desarrollar un mundo empresarial y de administración electrónica, creamos, comunicamos y utilizamos cada vez más información digital. La ofimática, el correo electrónico e Internet son herramientas potentes a este respecto, pero recogiendo las palabras de Tim Berners-Lee: «[...] las herramientas potentes pueden muchas veces utilizarse con fines constructivos y destructivos [...]». Estas herramientas pueden pues utilizarse para bien y para mal, y se trata de saber cómo obtener el máximo beneficio de ellas.

Otro aspecto interesante que debe tenerse en cuenta es que las actividades vinculadas a la información convergen en uno único mundo virtual, la red, a la que se accede desde la oficina (es decir, desde la pantalla del ordenador). En otros términos, las tecnologías de la información ofrecen medios radicalmente diferentes de registrar, manipular y utilizar la información, pero por otro lado exigen nuevos enfoques para gestionar y tratar esta información. Ahora bien, no siempre administramos adecuadamente la información digital. La pérdida de información es uno de los riesgos de la evolución actual, no sólo como tal, sino también porque no somos conscientes de la existencia de este riesgo. En consecuencia, no sabemos detectarlo cuando es necesario. Debemos por tanto gestionar la información digital y su accesibilidad.

Los medios tradicionales de investigación y acceso a la información no son suficientes en un mundo digital, habida cuenta de la naturaleza diferente de esta información. Contrariamente a los documentos clásicos de papel, los documentos o la información digital ya no son entidades materiales fijas, sino que son intangibles y volátiles. Necesitamos pues nuevos mecanismos adaptados y complejos para hacerlos visibles, recuperables, accesibles y comprensibles.

Varios proyectos y enfoques se esfuerzan actualmente en cumplir estas condiciones. Algunas profesiones, en particular los bibliotecarios y los archiveros, se ocupan básicamente de la utilización de metadatos de todas clases, mientras que los fabricantes de *software* proponen herramientas, tales como la indexación automática y la búsqueda en el texto íntegro. Otros instrumentos son los diccionarios y las normas. La cuestión radica en saber qué enfoque es el más útil, en qué circunstancias y con qué objetivo.

Para poder utilizar las herramientas, es necesario entender la finalidad de su utilización, por ejemplo: el suministro de servicios electrónicos, el acceso a la información, a la comunicación, a la conservación, a la gestión, etc. Todas estas actividades tendrán requisitos diferentes, eventualmente con algunos solapamientos. Cabe distinguir al menos dos perspectivas: la perspectiva de la creación y del mantenimiento de la información y la perspectiva de su utilización. Si el creador y el usuario no son una misma persona, su visión será diferente y por tanto también lo serán sus necesidades.

La red implica también la utilización de una lengua común de investigación y comunicación. La interoperatividad de la información en los distintos ámbitos y competencias es cada vez más importante. Aparte de la interoperatividad, la posibilidad de recuperar datos en sí no es suficiente. La información recuperada y encontrada debe presentarse y probarse en cuanto a su fiabilidad, y como último aspecto, pero no menos importante, debe comprenderse para poder interpretarse. Todos estos aspectos influirán en la gestión de la información digital y en su accesibilidad.

Existen numerosas iniciativas que tratan de encontrar soluciones para buscar información en la red de manera adecuada. Estas iniciativas están dirigidas en su mayoría por las necesida-

des de una comunidad, como es el caso de la iniciativa IAA (Iniciativa de Archivos Abiertos) o del proyecto DCMI (Dublin Core Metadata Initiative). El desarrollo de la «red semántica» constituye otro ejemplo reciente.

Por lo que respecta a la evolución de la comunidad archivística, la codificación de la descripción (EAD) y del contexto (EAC) de los archivos son otros medios de presentar información archivada a comunidades de usuarios. Estas soluciones se basan en series de metadatos de archivo existentes.

La exposición mencionará los aspectos que rodean a la recuperación y la accesibilidad de la información digital, en particular la creación de metadatos, la interoperatividad y los principales aspectos de la evaluación de los enfoques, las iniciativas y las herramientas actualmente disponibles.

Ein Universum im Wachstum: Metadaten und Zugänglichkeit digitaler Informationen

Johannes Hofman

Im Zuge der Entwicklung einer Welt mit elektronischem Geschäftsverkehr (E-Business) und elektronischer öffentlicher Verwaltung (e-Government) schaffen, verbreiten und nutzen wir eine immer größer werdende Menge digitaler Informationen. Büroautomatisierung, E-Mail und Internet sind dabei starke Werkzeuge, doch wie Tim Berners-Lee sagte, lassen sich starke Werkzeuge oft sowohl für konstruktive als auch für destruktive Zwecke benutzen. So können sie entweder zum Guten oder zum Schlechten dienen, und die Frage ist, wie wir das Beste aus ihnen machen können.

Ein anderes relevantes Problem besteht darin, dass alle informationsbezogenen Aktivitäten in einer virtuellen Welt zusammenlaufen, und zwar im Internet, zu dem man vom Schreibtisch (d. h. vom Computerbildschirm) aus Zugang hat. Anders gesagt, ermöglicht die Informationstechnologie vollkommen andere Wege der Aufzeichnung, Manipulierung und Nutzung von Informationen, erfordert andererseits aber auch neue Konzepte zur Verwaltung und Handhabung von Informationen. Es sieht doch so aus, dass wir noch immer nicht in der Lage sind, digitale Informationen sachgerecht zu verwalten. Eine der Gefahren dieser aktuellen Entwicklungen ist der Verlust von Informationen, und zwar nicht nur der Informationen an sich, sondern auch, weil wir nicht wissen, dass sie vorhanden sind. Dadurch können wir sie nicht finden, wenn wir sie brauchen. Es geht also um das Management digitaler Informationen und ihrer Zugänglichkeit.

In einer digitalen Welt wird die herkömmliche Vorgehensweise beim Retrieval und Zugriff auf digitale Informationen wegen des andersartigen Charakters dieser Informationen nicht ausreichen. Im Gegensatz zu traditionellen Papierdokumenten sind digitale Dokumente bzw. Informationen keine festen physischen Einheiten mehr und daher flüchtig und nicht greifbar. Daher benötigen wir neue, zweckentsprechende und moderne Mechanismen, um sie sichtbar, recherchierbar, zugänglich und verständlich zu machen.

Momentan versucht man mit vielen Projekten und Konzepten, diesen Erfordernissen gerecht zu werden. Fachbereiche wie das Bibliothekswesen und das Dokumenten- und Archivwesen beispielsweise befassen sich in erster Linie mit Metadaten aller Art, die Softwareindustrie hingegen bietet Tools wie automatisches Indexieren und Volltextsuche an. Andere Instrumente sind Thesauri und Normen. Dabei stellt sich nun die Frage, unter welchen Umständen und zu welchem Zweck welches Konzept – wenn überhaupt – am hilfreichsten ist.

Um Tools nutzen zu können, muss klar sein, wofür sie verwendet werden, z. B. elektronische Dienstleistungserbringung, Zugriff auf Informationen, Kommunikation, Archivierung, Verwaltung usw. Jede dieser Aktivitäten ist mit anderen, sich vielleicht zum Teil überschneidenden Anforderungen verbunden. Es lassen sich mindestens zwei Sichtweisen unterscheiden: die Sichtweise der Erstellung und Pflege und die Sichtweise der Nutzung von Informationen. Handelt es sich beim Urheber und beim Benutzer nicht um dieselbe Person, sind ihre Sichtweisen und damit auch ihre Anforderungen unterschiedlich.

Das Internet erfordert auch eine gemeinsame Sprache für die Recherche und die Kommunikation. Daher gewinnt die Interoperabilität von Informationen in den verschiedensten Zuständigkeitsbereichen und Domänen zunehmend an Bedeutung. Abgesehen davon wird Recherchierbarkeit allein nicht ausreichen. Die gesuchten und gefundenen Informationen müssen auch dargestellt und auf ihre Vertrauenswürdigkeit hin geprüft werden, und nicht zuletzt muss man sie verstehen, um sie interpretieren zu können. All diese verschiedenen Gesichtspunkte machen es erforderlich, dass im Hinblick auf die Verwaltung digitaler Informationen und die Zugänglichkeit etwas unternommen wird.

Es gibt zahlreiche Initiativen, die sich um Lösungen für das Auffinden sinnvoller Informationen im Internet bemühen. Dahinter stehen zumeist breite Benutzergruppen, wie beispielsweise im Falle der Open Archives Initiative oder des Dublin Core-Standards. Ein anderes interessantes aktuelles Beispiel ist die Entwicklung des Semantic Web.

Im Dokumenten- und Archivwesen sind Entwicklungen wie die Encoded Archival Description (EAD) und der Encoded Archival Context (EAC) weitere Mittel zur Darstellung aufgezeichneter Informationen für Nutzergruppen. Sie basieren auf vorhandenen Metadaten der Aktenführung.

In diesem Beitrag werden Retrieval- und Zugriffsmöglichkeiten bei digitalen Informationen erörtert, so etwa Erstellung von Metadaten, Interoperabilität und Schlüsselfaktoren für die Bewertung vorhandener Ansätze oder Initiativen und Tools auf diesem Gebiet.

Un univers en extension: métadonnées et accessibilité de l'information numérique

Johannes Hofman

En développant un monde d'e-business et d'e-administration, nous créons, nous communiquons et nous utilisons un nombre sans cesse croissant d'informations numériques. La bureautique, le courrier électronique et l'internet sont des outils puissants à cet égard, mais, pour reprendre les mots de Tim Berners-Lee, «des outils puissants peuvent souvent être utilisés à des fins constructives et destructives»… Ces outils peuvent donc être utilisés pour le meilleur et pour le pire, et la question est de savoir comment en faire bénéficier le plus grand nombre.

Un autre aspect intéressant doit être pris en compte: les activités liées à l'information convergent dans un seul et unique monde virtuel, la toile mondiale, auquel on accède depuis son bureau (c'est-à-dire depuis son écran d'ordinateur). En d'autres termes, les technologies de l'information offrent des moyens radicalement différents d'enregistrer, de manipuler et d'utiliser l'information, mais elles exigent de nouvelles approches pour gérer et traiter cette information. Or, nous ne parvenons toujours pas à gérer l'information numérique de façon appropriée. La perte d'informations est l'un des risques des développements actuels, pas seulement en tant que telle, mais aussi parce que nous ne sommes pas conscients que ce risque

existe. En conséquence, nous ne savons pas le détecter là où il faudrait. Il nous faut donc gérer l'information numérique et son accessibilité.

Les moyens traditionnels de recherche et d'accès à l'information ne suffisent plus dans un monde numérique, compte tenu de la nature différente de cette information. Contrairement aux documents classiques sur papier, les documents ou l'information numériques ne sont plus des entités matérielles fixes, et sont donc en tant que tels intangibles et volatils. Il nous faut donc de nouveaux mécanismes adaptés et complexes pour les rendre visibles, récupérables, accessibles et compréhensibles.

Plusieurs projets et approches s'efforcent actuellement de satisfaire à ces conditions. Certaines disciplines professionnelles, notamment les bibliothécaires et les archivistes, se préoccupent essentiellement de l'utilisation de métadonnées de toutes sortes, tandis que les constructeurs de logiciels proposent des outils tels que l'indexation automatique et la recherche en texte intégral. Les thésaurus et les normes constituent d'autres instruments. La question est alors de savoir quelle approche est la plus utile, le cas échéant, dans quelles circonstances et dans quel but?

Pour pouvoir utiliser des outils, il faut comprendre la finalité de leur utilisation — en l'occurrence: la fourniture de services électroniques, l'accès à l'information, la communication, la conservation, la gestion, etc. Toutes ces activités correspondront à des cahiers des charges différents, avec éventuellement certains recoupements partiels. Deux perspectives au moins peuvent être distinguées: celle de la création et de la maintenance de l'information et celle de son utilisation. Si le créateur et l'utilisateur ne sont pas la même personne, leur point de vue sera différent et, partant, leurs besoins le seront également.

Le web implique également d'utiliser un langage commun de recherche et de communication. L'interopérabilité de l'information des divers domaines et compétences devient donc de plus en plus pertinente. Hormis l'interopérabilité, la possibilité de récupérer les données n'est pas suffisante. Les informations récupérées et trouvées doivent également être présentées et testées quant à leur fiabilité, le dernier aspect — et non des moindres — étant la nécessité de les comprendre afin de pouvoir les interpréter. Tous ces différents aspects influeront sur la gestion de l'information numérique et sur son accessibilité.

De nombreuses initiatives sont déployées, qui s'efforcent de trouver des solutions pour rechercher l'information sur le web de façon pertinente. Ces initiatives sont dirigées pour la plupart par les besoins d'une communauté, comme dans le cas de l'initiative OAI (Open Archives Initiative) ou du projet DCMI (Dublin Core Metadata Initiative). Le développement du «web sémantique» en constitue un autre exemple récent.

En ce qui concerne les développements intervenant dans la communauté archivistique, l'encodage de la description (EAD) et l'encodage du contexte (EAC) des archives sont d'autres moyens de présenter des informations archivées à certaines communautés d'utilisateurs. Ces solutions reposent sur des séries de métadonnées d'archivage existantes.

L'exposé évoquera les aspects entourant la récupération et l'accessibilité des informations numériques, notamment la création de métadonnées, l'interopérabilité et les principaux aspects de l'évaluation des approches, des initiatives et des outils actuellement disponibles.

Jean-Claude Van Lancker

Jean-Claude Van Lancker is customer services supervisor at Eurocontrol.

Eurocontrol case study

Jean-Claude Van Lancker

Introduction

Eurocontrol has as its primary objective the development of a seamless, pan-European air traffic management (ATM) system. The achievement of this objective is the only solution to the present and future challenges facing the entire aviation community, which are to cope with the constant growth in air traffic, while maintaining a high level of safety, to reduce costs and to respect the environment.

The agency develops, coordinates and plans the implementation of short and long-term pan-European ATM strategies and their associated action plans in a collective effort involving national regulatory authorities, air navigation service providers, the civil and military airspace users, industry, and other European institutions. It interacts with a wide variety of external institutions and organisations including product/service suppliers, national air traffic offices and other international bodies such as the International Civil Aviation Organisation (ICAO), the European Commission, the US Federal Aviation Authority (FAA), the European Space Agency (ESA) etc.

The organisation has 1 675 staff of which 1 160 are based in Brussels and 380 located in offices at Maastricht, Luxembourg and Brétigny. It is involved in the production of a wide range of technical, procedural, political and statutory documentation, much of which is subjected to detailed review and scrutiny at numerous levels and is translated into up to 20 different languages.

Current document management environment

Eurocontrol's activities are typically driven by tasks forces, committees, working groups, consultation groups and councils that are responsible for the advancement of specific Air Traffic Management and Control programmes or specific administrative, institutional or financial remits.

Documents are produced within the Agency predominantly using Microsoft Office products but often require more sophisticated tools such as desktop publishing (e.g. Adobe Publishing Collection) or drawing applications (e.g. AutoCAD and Autodesk MapGuide). The files produced using these applications are stored in a variety of locations: stand-alone servers, network drives and PC hard drives under locally managed file structures and document referencing systems. These repositories are distributed throughout the organisation with various levels of interface and interconnectivity.

Internal document production and life-cycle management, version control, progress tracking, translations etc. are all manually coordinated by authors (technicians, project leaders etc), librarians, administrative support staff, personal assistants and secretaries.

Distribution within the agency is achieved using a combination of posted hard copy, e-mailed attachments and increasing use of web-publishing techniques via the ORBITe Intranet and shared/public network drives. External distribution is largely reliant on couriered hard copy and fax, supplemented by limited, but increasing use of e-mailed attachments and web publishing.

Project drivers

Each of the business units has, in the absence of a corporate standard, tended to develop differing in-house practices and procedures for producing, storing and managing an ever-increasing volume of documentation. This fragmented approach to document management is increasingly incapable of meeting the organisation's operational and strategic needs and is adversely affecting efficiency and the timely delivery of service to Member States and stakeholders.

Document delivery deadlines have been missed, especially in relation to supply of information in sufficient time to prepare for meetings. The inadequacies of the current arrangements are resulting in the following problems:

- Excessively slow progression of documents through the production process due to:
  - — programme delays incurred as a result of reliance on transmittal of hard copy to and from contributors, especially when using international couriers;
  - — documents becoming 'lost within the system' due to an inability to efficiently track progression through the production process, especially during review stages, delays often being caused when contributors are absent without a mandated alternative contributor having been nominated;
  - — difficulty in controlling and coordinating document versions and associated contributors' comment;
  - — difficulty to control deadlines.

- poor utilisation of the Agency's knowledge base due to fragmented and uncoordinated information storage arrangements, both within and between directorates and sites;

- duplication of information both within and between directorates and sites.

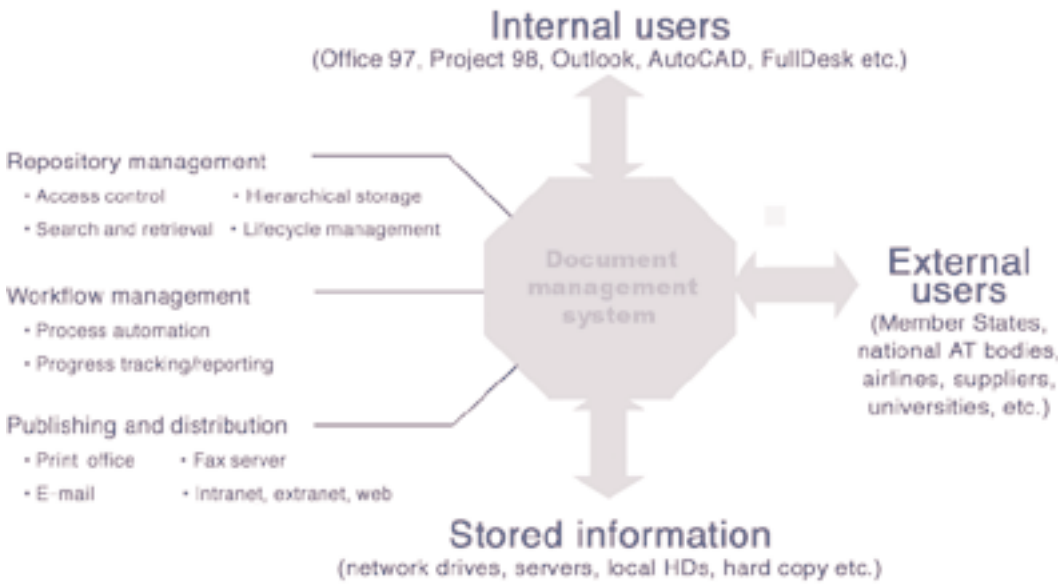
Identifying the need

The need for a document management solution was highlighted at a meeting of the Advisory Financial Group where it was requested that: 'priority be given to those projects which ensure timely, electronic and hard copy, delivery of Agency Working Papers to States and other stakeholders'. Moreover, the inflexible nature of the current arrangements will obstruct the Agency's ability to adapt to future changes in role or organisation. With the mandate to improve the methods of working, two projects were identified:

FAX automation project: this project had as specific objective to replace 'archaic' stationary FAX devices by a solution that integrates the management of the FAX communication within the overall Office environment. In particular that project had to investigate on the scanning facilities available and to adopt standard means and processes for the digitalisation of documents not readily available in electronic format.

Document management project: this project has a much broader business scope than the previous one, its objective being the implementation of an Agency-wide document management solution. In actual fact this project has not yet delivered its final solution but has selected a shortlist of the three potential suppliers.

Functional overview of requirement



Fax automation project

The prime goal of this project was to provide an integrated system that would automate the delivery of messages both in and out of the agency.

Eurocontrol already had a large base of HP printing solutions in place for its use of Office systems. On the other hand some specific solutions were using the Digital Sender device, a network appliance dedicated to converting paper documents into electronic format and sending the digitised document over the network.

The emerging trend of the Multi-function product (MFP) retained Eurocontrol's attention. Not only because it combines four current devices in one (printer, fax, copier, digital sender) reducing therefore substantially the cost of ownership, but also because it fits nicely into the office environment (proven technology that integrates seamlessly the network through LDAP connector).

After careful consideration of the various products available the Agency chose Topcall an integrated messaging suite combined with the digital sending functionality available on the different devices from Hewlett Packard. Aspects such as intelligent capture of the information, back-file conversion and document standard formats were central to the selection by Eurocontrol of the HP-MFP product line.

This combination provided a number of advantages:

— it allowed staff to scan in documents locally and have them routed electronically to the recipient(s) via e-mail or fax;

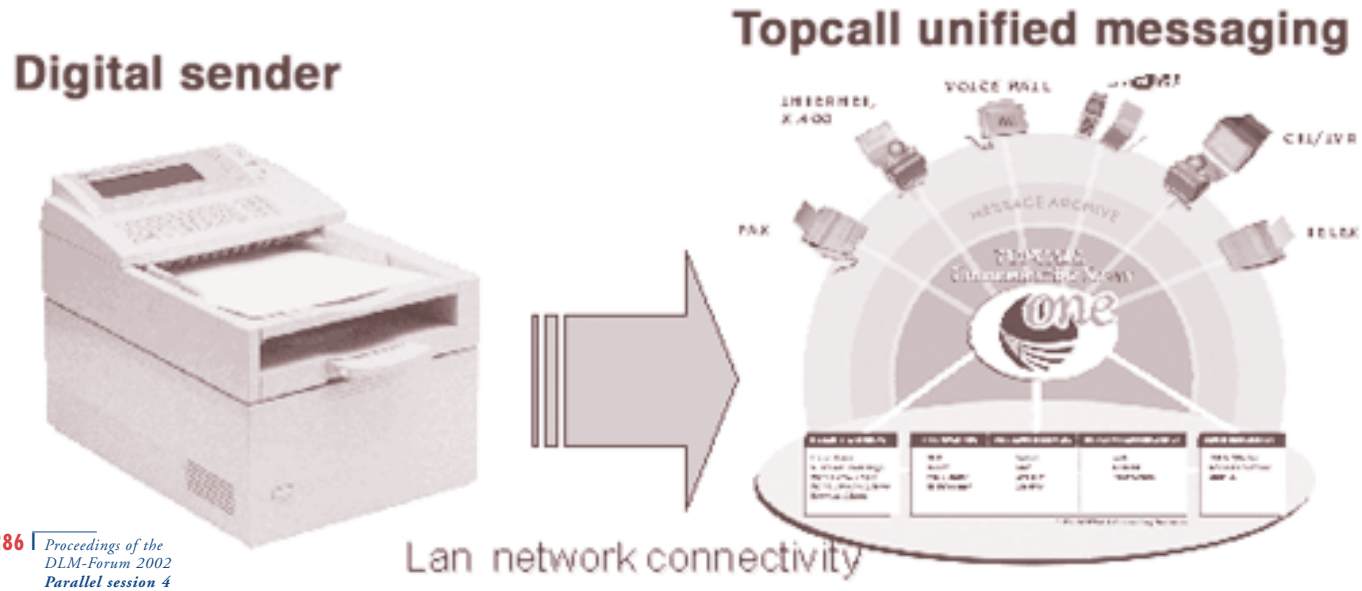
— documents could be stored as TIFF images or PDF files and then forwarded via e-mail or stored on the network stores;

— incoming telex or fax messages could be received as e-mail messages in a users Microsoft Outlook desktop environment;

— the MFP products opened 'copier functionality' to the users of the office systems, directly from their workplace. 'Networked' devices will progressively replace standalone copiers;

— the MFP products extended the simple use of scanner to different workflow applications;

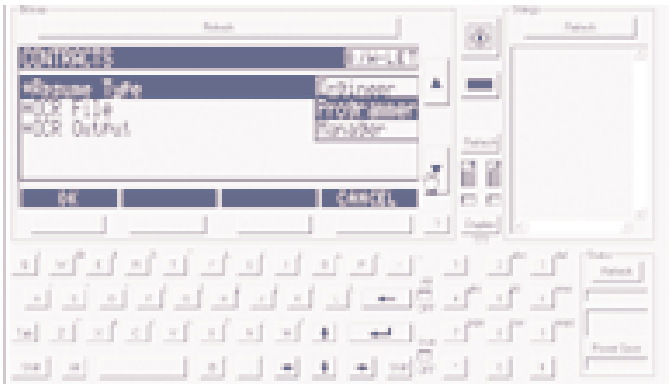
— The whole network of printer devices can be managed remotely by the system administrator though the HP Web JetAdmin Application Framework.



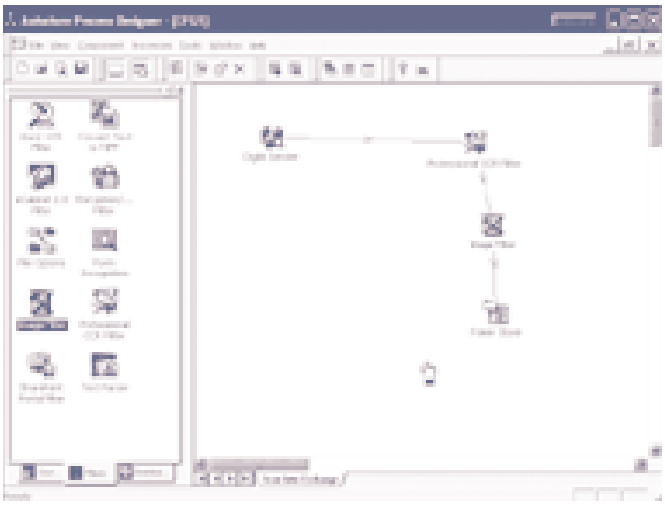
The HP digital sender and multi-function product were chosen for their versatility and ease of use. The digital sender is a colour-networked scanner with LED screen, keyboard and automatic document feeder. The LaserJet 4100 MFP and LaserJet 9000 MFP both have colour network scanning, LED screen, keyboard and LAN connections, as well as printing and copying functionality. The digital sending functionality on these devices, when combined with a software suite called AutoStore ([1]) 2002, becomes a highly configurable system that automates the process of document capture, processing and routing, into most popular applications (e.g. Microsoft SharePoint ([2]) Portal Server, Microsoft Exchange, Lotus Notes, Documentum and Filenet). AutoStore profiles can be defined, that allow the user of the digital sending device to determine how the electronic document will be processed. These document-processing features include searchable PDF, document encryption, optical character recognition (OCR) as well as forms processing.

LaserJet 4100

([1]) See www.nsius.com for more details.
([2]) See www.hp.com/go/sharepoint for more information.



The profiles can easily be altered and new ones added using a 'drag and drop' interface of the process designer. This allows the administrator to create complex but easy to use profiles without resorting to 'programming'.



Implementation issues

The storage of the scanned document as images although relatively small at 40 KB per page, soon started to occupy many gigabytes of space on the network. Users were then faced with the declaimer of being told to 'housekeep' their network drives, i.e. remove unwanted files, and the requirement to keep the document they had carefully scanned for future reference. As a matter of the fact, the availability of the document management solution will provide the expected longer-term solution to that issue, as amongst other objectives it will have to provide a mean to store once the only occurrence of a document version.

Lessons learnt

The availability of new scanning facilities was felt to be a critical success factor. For many different business units the availability of these new facilities had paved the way to the conversion of their local archive to digital files. As for example, the recruitment section uses the handy scanning and storing capabilities of a digital sender to speed-up the process of new staff recruitment. Electronic versions of application forms are now distributed as soon as they are available in human resources.

Next stage

Document management project: To complement the business requirements of information capture (covered in the FAX project), document management focuses more in detail on the aspects of document production (revision, access) and document retention (long term availability and preservation).

The project objectives are to achieve the following:

— improve the visibility, earlier delivery and re-usability of Eurocontrol's information and thereby increase the efficiency of document production;

— reduce the costs and time associated with distributing information by providing electronic alternatives to the current reliance on hard copy dissemination;

— improved information exchange with external stakeholders;

— increase process efficiency through the introduction of workflow management.

All the above objectives were documented together with the functional requirement and written into an invitation to tender (ITT) and sent to 100 potential suppliers, 10 proposals were received from which three were short-listed.

Business justification for the project

As stated before the Agency was experiencing problems managing its information assets and in some cases failing to deliver the information to the right people on time. After reviewing the potential areas where EDM and workflow technology could help it was decided to implement the technology and new methods of working in two specific areas. They were chosen as they had the most pressing requirements and between them represented similar ways of working to all the other business area of the Agency.

After carefully assessing the impact the implementation of a DMS and workflow automation would have on these two areas, it was calculated they would yield the following benefits:

— timely delivery of the Agency's service/product to Member States and stakeholders thus improving Eurocontrol 's image and customer satisfaction;

— enhanced visibility and availability of the Agency's information using electronic search and retrieval of documents across all of the Agency's directorates and sites;

— improved flexibility of electronically managed documents enabling the Agency to quickly respond to future;

— organisational changes with minimal disruption;

— improved productivity as staff is released from mundane and stressful administrative tasks to concentrate on core business objectives.

— early involvement of member sates and stakeholders in the document production lifecycle by electronically publishing draft versions for review and comment;

— reduction in document production times;

— decrease in document distribution costs as more use is made of electronic media, e.g. e-mail and web publishing.

As stated before the project has now been put on hold until January 2003 due to budgetary constraints but we are confident that when it is deployed it will deliver significant benefits, based on our experience with the fax-messaging project.

Estudio de casos de Eurocontrol

Jean-Claude Van Lancker

Cada una de las unidades de Eurocontrol, a falta de una norma corporativa, ha tendido a desarrollar diversas prácticas y procedimientos internos para presentar, almacenar y gestionar un volumen cada vez mayor de documentación. Este enfoque fragmentado de la gestión de documentos es cada vez menos capaz de cubrir las necesidades operativas y estratégicas de la organización y está afectando desfavorablemente a la eficacia y a la prestación puntual de servicio a los Estados miembros y a los interesados.

La necesidad de una solución para la gestión de documentos se subrayó en una reunión del grupo consultivo financiero, donde se solicitó *«que se dé prioridad a los proyectos que garantizan la entrega puntual, en papel y en versión electrónica, de los documentos de trabajo de la agencia a los Estados y a otras partes interesadas»*.

Se identificaron dos proyectos:

— **El proyecto de automatización del fax,** que tiene como objetivo específico sustituir aparatos fijos de fax «arcaicos» por una solución que integre la gestión de la comunicación por fax en el medio global de la oficina. Este proyecto, en especial, tuvo que investigar las posibilidades de escaneado disponibles y adoptar medios y procesos estándar para la digitalización de documentos no disponibles en formato electrónico.

El objetivo fundamental de este proyecto era proporcionar un sistema integrado que automatizase la entrega de mensajes, tanto dentro como fuera de la agencia. La disponibilidad de nuevos instrumentos de escaneado se consideró un factor crítico para el éxito. Para muchas unidades diversas, la disponibilidad de estos nuevos instrumentos preparó el camino para la conversión de sus archivos locales en ficheros digitales.

— **El proyecto de gestión de documentos,** de un alcance mucho mayor, tiene por objetivo aplicar una solución de gestión de documentos para toda la agencia. De hecho, este proyecto aún no ha logrado su solución final, pero ha realizado una preselección de tres posibles proveedores.

Como complemento de los requisitos de recogida de información (cubiertos en el proyecto FAX), la gestión de documentos se centra más detalladamente en los aspectos de la elaboración de los documentos (revisión y acceso) y en su conservación (disponibilidad y conservación a largo plazo).

Los objetivos del proyecto son, entre otros, mejorar la visibilidad, la entrega temprana y la reutilización de la información de Eurocontrol, reducir los costes y el tiempo de distribución de la información proporcionando alternativas electrónicas a la actual dependencia de la difusión en papel, mejorar el intercambio de información con terceros del exterior y aumentar la eficacia del proceso mediante la introducción de la gestión del flujo de trabajo.

Eurocontrol-Fallstudie

Jean-Claude Van Lancker

Jeder einzelne Geschäftsbereich von Eurocontrol hat bisher, da ein gemeinsamer Standard fehlte, versucht, verschiedene interne Praktiken und Verfahren zur Erstellung, Aufbewahrung und Verwaltung eines ständig an Umfang zunehmenden Archivguts zu entwickeln. Dieses uneinheitliche Konzept der Archivgutverwaltung wird dem operationellen und strategischen Bedarf der Organisation immer weniger gerecht und beeinträchtigt die Effizienz und die pünktliche Bereitstellung von Diensten für Mitgliedstaaten und andere Interessengruppen.

Die Notwendigkeit einer Lösung in der Archivgutverwaltung wurde deutlich, als auf einer Sitzung der Beratenden Finanzgruppe gefordert wurde, dass: *„solche Vorhaben vorrangig zu behandeln seien, die die fristgerechte Weiterleitung von Arbeitsunterlagen der Behörde in elektronischer Form sowie als Papiervorlage an Staaten und andere Interessensgruppen gewährleisten“*.

Zwei Vorhaben wurden festgelegt:

– Das **Fax-Automatisierungs-Projekt** verfolgt das Ziel, „herkömmliche“ stationäre Faxgeräte durch eine Lösung zu ersetzen, die die Verwaltung der Fax-Kommunikation in das gesamte Büroumfeld integriert. Insbesondere wurden im Rahmen dieses Projektes die verfügbaren Scan-Möglichkeiten geprüft sowie Standardmittel und -verfahren zur Digitalisierung nicht direkt in elektronischer Form verfügbarer Dokumente verabschiedet.

Das vorrangige Ziel dieses Projektes bestand darin, ein integriertes System bereitzustellen, das die Weiterleitung von Mitteilungen sowohl innerhalb als auch außerhalb der Behörde automatisiert. Die Verfügbarkeit neuer Scan-Einrichtungen wurde als kritischer Erfolgsfaktor angesehen. In vielen verschiedenen Geschäftsbereichen hat die Verfügbarkeit dieser neuen Einrichtungen den Weg zur Umstellung der Abteilungsarchive auf elektronische Dateien geebnet.

– Das **Dokumentenmanagement-Projekt** mit einem viel breiter angelegten Umfang, dessen Ziel darin besteht, eine behördenweite Lösung der Archivgutverwaltung umzusetzen. Bisher ist das Projekt noch zu keiner endgültigen Lösung gelangt, es wurde jedoch eine Vorschlagsliste mit den drei potenziellen Auftragnehmern erstellt.

In Ergänzung der Geschäftsanforderungen der Informationserfassung (im FAX-Projekt enthalten) konzentriert sich die Archivgutverwaltung im Einzelnen stärker auf die Erstellung von Dokumenten (Überarbeitung, Zugang) und die Aufbewahrung von Unterlagen (langfristige Verfügbarkeit und Erhaltung).

Ziele dieses Projektes sind unter anderem die Erhöhung der Transparenz, eine raschere Weitergabe und Wiederverwendung von Eurocontrol-Informationen, eine Reduzierung der Kosten und des mit der Informationsverbreitung verbundenen Zeitaufwandes durch die Bereitstellung elektronischer Alternativen zur gegenwärtigen Verwendung von Papiervorlagen, die Verbesserung des Informationsaustausches mit externen Interessengruppen und eine Erhöhung der Verfahrenseffizienz durch die Einführung eines Arbeitsflussmanagements.

Étude de cas: Eurocontrol

Jean-Claude Van Lancker

Chacun des départements d'Eurocontrol a eu tendance, en l'absence de règles communes, à développer des pratiques et des procédures internes différentes en ce qui concerne la production, le stockage et la gestion de documents toujours plus nombreux. Cette approche morcelée de la gestion documentaire répond de moins en moins aux besoins opérationnels et stratégiques de l'organisation, nuit à l'efficacité des services fournis aux États membres et aux parties intéressées et empêche la fourniture de ces services en temps voulu.

La nécessité de trouver une solution en ce qui concerne la gestion des documents a été soulignée lors d'une réunion du groupe consultatif dans le domaine financier, qui a demandé que «la priorité soit donnée aux projets permettant de fournir les documents de travail de l'agence en temps utile, sur support électronique et papier, aux États et autres parties intéressées».

Deux projets ont été identifiés:

- Le projet ayant trait à l'**automatisation des télécopies,** dont l'objectif spécifique est le remplacement des télécopieurs «archaïques» par une solution intégrant la gestion des télécopies dans l'environnement bureautique global. Ce projet devait permettre, en particulier, d'étudier les outils de scannage disponibles et d'adopter les instruments et procédés standard en vue de la numérisation de documents non disponibles immédiatement au format électronique.

  L'objectif principal de ce projet était de fournir un système intégré qui rendrait la transmission des messages automatique, tant à l'intérieur qu'à l'extérieur de l'agence. La disponibilité de nouveaux outils de scannage était considérée comme un facteur de réussite essentiel. Pour nombre de départements, la disponibilité de ces nouveaux outils a débouché sur la conversion des archives locales en documents numériques.

- Le projet de **gestion documentaire:** ce projet, qui a une portée beaucoup plus large, a pour objet la mise en œuvre d'un système de gestion documentaire à l'échelle de l'agence. En fait, ce projet n'a pas encore complètement abouti, mais a déjà permis d'établir une liste restreinte comptant trois fournisseurs potentiels.

  Pour répondre aux besoins opérationnels de la saisie des informations (qui fait l'objet du projet «Télécopies»), la gestion documentaire met davantage l'accent sur les différents aspects de la production (révision, accès) et de la conservation (disponibilité et conservation à long terme) des documents.

  Le deuxième projet vise notamment à améliorer la visibilité, la rapidité de transmission et les possibilités de réutilisation des informations d'Eurocontrol, à réduire les coûts et les délais de transmission des informations en fournissant des alternatives électroniques au papier, à améliorer les échanges d'informations avec l'extérieur, ainsi qu'à accroître l'efficacité des procédures grâce à l'introduction de la gestion des flux de travail.

Malcolm Todd

Malcolm Todd is a member of the Electronic Records Management Development Unit at the UK Public Record Office (PRO), led by Stephen Harries.
This involves him in administering the PRO's formal testing programme for electronic records management, which conducts formal compliance testing on software against the statement of functional requirements he will be speaking about at this conference.
He also writes electronic records management guidance for the UK central Government (the most recent example completed being on the subject of web sites and intranets) and is currently working on a metadata standard for UK central Government records management.
Prior to joining the PRO in 2001, Malcolm worked elsewhere in the UK Government as a records manager.

Stephen Harries

Stephen Harries is Head of the Electronic Records Management Development Unit in the Public Record Office of the United Kingdom.
Stephen has a background in information management and information science, and has worked in both public and private sectors, including higher education, software development and supply as a consultant.

Functional requirements for ERMS

Malcolm Todd, Stephen Harries

This paper describes the evolution of standard and generic functional requirements for electronic records management systems (ERMS) in the UK Government, the influence which these have had the development of the commercial software product market for ERMS in the United Kingdom, and how this fits in to the overall strategy for modern electronic records in the moves towards e-Government.

An attempt has been made to relate it to a number of other initiatives (some of which have been covered by other DLM contributors) to place it in the wider context.

ERM in the UK Government

As we move more and more into a knowledge-based economy, governments around the world have recognised the need to modernise the way in which the business of government is conducted, and the way in which services to the citizen are organised and delivered, in order to retain competitiveness on the world stage.

In the United Kingdom, the Modernising Government White Paper, published in 1999, set out a range of targets for the achievement of electronic government (then called 'information age government'). These included targets for making all government services to the public available through online channels by the year 2005, and a specific target for all government departments and agencies to have implemented the capability for managing electronic records electronically by the year 2004.

Over the last few years, departments and agencies have been tackling the difficult task of making government services available electronically by the target date of 2005. While most would agree that this has turned out to be a rather more difficult task than first anticipated in the initial flush of 'dot.com' enthusiasm, there are now some significant successes in making portal-based and e-enabled transactional services available across the Internet, and it seems clear that this is the direction in which governments must move for the future.

While much of the public profile rests on the customer-facing view of e-government, existence of a supporting infrastructure is also important to the sustainability of the venture. Electronic services will produce electronic records of transactions, business activities and decisions, and these records must be reliably captured and stored, and access to them retained over perhaps many years, to make the e-government service maintainable in the longer term.

For example, behind a portal service built to enable members of the public to submit an electronic tax return, there must be facilities to capture an authentic record of that return, and to ensure the information content remains accessible and complete for seven years through changes in technology, and then to destroy the record by a managed and audited process according to established business rules. During this time, the information content of the record may be needed to answer enquiries or to help with a business decision — perhaps even to be produced as evidence in court, where its continued reliability as an accurate and complete record may be subject to legal challenge. Legislation also requires that the information is prevented from use for inappropriate purposes during its lifetime — the obvious example is the processing of personal data.

The government is also facing the challenge of the implementation of the Freedom of Information Act 2000. For some of you here from countries where there has been such legislation in place for years, decades or even centuries this may seem a commonplace, but one of the features of the UK legislation is that it is fully retrospective. This means that it catches all records created in the past including approximately 25 years' worth from a period when there was no inkling when they were created and organised that they would be released within 30 years of their creation, if indeed they were kept that long. In some respects, the accumulation of uncontrolled or partly controlled electronic documents since circa 1990 is as difficult a problem as the longer legacy of paper.

At the same time, there is a growing emphasis on internal modernisation of government business processes. Just as the development of electronic service delivery is about changing the ways government does business, rather than doing the same thing but online — 'innovation not automation' — so electronic records management is about changing the role and relationship of records management to the organisation. Long the seen as the 'poor relation', the ERM is moving into the centre of information resource management, but in so doing it will have to change its nature.

Rather than simply automating the administration of paper, this move will unlock the content of records previously inaccessible in unwieldy paper folders, and make electronic records a central resource for corporate knowledge. ERM is beginning to be fully integrated with knowledge management and decision-making systems, and has a particular role to play in supporting evidence-based policy making — making policies that can be shown to work. In essence, it is a major platform for internal modernisation of government, where the changes predicated on the modernisation of external service delivery are rolled back through the rest of the organisation. Success or failure will be determined here.

For the records manager, this is a fundamental change — from management of the artefacts towards design of the environment in which the end user operates. These are the people who create records on a day-to-day basis, and who will take key decisions on whether a record is captured and classified correctly. If it is not, then the record is effectively lost — no one else will have time to tidy up afterwards. Records managers will need new skills to ensure that this activity takes place as a part of the normal course of business. They will no longer be able to establish control by physical custody of non-current material. This involves them in developing completely different skill sets from those required in the past: influencing, policy setting, interfacing with other information professionals, technology professionals etc., etc.

The challenge for electronic records management (ERM) is to enable the capture and flexible use of the information contained in electronic records, to provide support for modernisation and new ways of working, while retaining the reliable structure and stability of the corporate record knowledge base over the whole records life cycle. In the electronic environment, this poses particular challenges. We seek the same quality of outcome as paper records management systems aim to achieve (or better), but the means of reaching these must be quite different.

The move towards corporate electronic records management requires not just implementation of ERM software facilities to manage the 'data', but also implementation of a 'good fit' between policies, procedures and user practices designed to meet organisational needs; and these together imply an extensive cultural change in attitudes, behaviour and working practice, which can be encouraged but not mandated — a big challenge for any public-sector organisation.

Some of these policies and procedures will be embedded in software systems, and greater consistency and interoperability between organisations can be achieved by standardising on the requirements that define software packages. Organisations will then be able to implement with greater confidence that they are not creating 'islands' of electronic records, and the overall public record systems will be able to achieve a greater level of dynamism and momentum.

Role of the PRO

The Public Record Office (PRO), the National Archive of the United Kingdom ([1]), is the lead agency for records management in the UK Government, responsible for setting and maintaining standards for current records management in government departments and agencies, as well as overseeing selection of the historic record. The PRO has had a programme aimed at furthering issues arising from the introduction of electronic records in government for some years, encouraging infrastructure developments and producing standards and guidance. In this capacity, it has responsibility for oversight of achievement of the 2004 target for the ERM, which states that:

It is our aim that by 2004 all newly created public records will be electronically stored and retrieved.

As a small-sized organisation, the PRO does not usually become directly involved in specific ERM projects — there would be far too many of them! — but aims rather to shape and influence the environment in which these projects happen, leveraging opportunities and investments to build and strengthen infrastructure. Two of the major areas in which we work towards this are in influ-

([1]) Technically England, Wales and the United Kingdom (i.e. excluding Northern Ireland and Scotland).

encing the commercial software sector for electronic records applications, and in encouraging sharing and common structures within government departments.

Another issue is that we do not have the power to mandate our views on departments and agencies. The legislative framework for records management in central government in the United Kingdom was written in the 1950s for the paper environment and concentrates on access issues and the physical custody of archival records, with the management of current records happening within a framework of where the legislation is silent (e.g. disposal). This is in marked contrast to the wide-ranging legal powers exercised by, for example, the National Archives of Australia over the records of the Australian Commonwealth.

In addition, a unitary approach to central government IT projects would run counter to the traditions of the UK Government. In some ways (such as the power of central *vis à vis* local government) we are a highly centralised administration. According to other perspectives, the structural changes of the 1980s and 1990s broke up central government into a plethora of executive agencies responsible for aspects of political programmes headed by ministerial departments. These probably have their own information technology strategies and environments and records management business requirements. This will also go hand in hand with different legacy problems with paper and uncontrolled electronic documents from the rollout of client-server computing in more or less the same period. By the time of writing, they may also have their own private sector partnerships for the provision of information technology and other services.

1999 functional requirements

Even before the publication of the modernising government targets, the PRO had initiated and led a project to identify generic functional requirements for electronic records management in UK Government. Working in conjunction with a wide range of government organisations — from those primarily delivering transactional services, and casework-based organisations, to policy-making departments — over many months, the PRO was able to identify a set of requirements that were quite practical in addressing real problems on the ground in a feasible way, but which also provides a challenging but achievable benchmark for software suppliers. The requirements were published in the later part of 1999, and are still available from the PRO web site (http://www.pro.gov.uk/recordsmanagement/)

Aim and audience

The functional requirements project has two broad aims:

- to provide support for central government organisations in developing their own requirements for ERM, distilling and sharing those related to common practice across all of government, so that individual organisations did not have to 're-invent the wheel';
- to define a benchmark (as a fixed target) for software suppliers developing and upgrading products, which encouraged competitiveness in supplying quality products, and against which individual products could be assessed

We also had a mind to the wider demand for good practice guidance from the Public Record Office as with the National Archives. We have the function of providing leadership within the archival community in the United Kingdom. Whilst in the United Kingdom it cannot always be assumed that the word 'archives' also encompasses current records management, especially in the electronic environment as it can more safely be assumed in some delegates' home countries, we have been increasingly looked to for expertise in this area, notably by local government. We have to prioritise this against our specific responsibilities towards central government.

For the present purpose, though, it is worth reflecting that the requirements are accompanied by documentation of the underlying concepts in the requirements (this has been invaluable in assisting software suppliers from very different backgrounds from archives and records management to understand our needs in terms of ERM software). We also see the requirements as a part of a growing body of knowledge that the Public Record Office shares with government and more widely.

Scope and coverage

Why was it worth developing a specific set of requirements for this context? In the UK software market at that time, little choice existed, and this mainly between small companies with a restricted support base. Most of the products available in 1999 had been developed in North America or Australia.

In the United Kingdom, records management practice differs quite considerably from the first of these domains: in the United Kingdom there is a emphasis on management at the file level, rather than the document level, and different approaches to security, and privacy and access issues. A particular approach to appraisal is well established in paper records management — reviews at fixed periods — which, although likely to change in the new electronic environment, is likely to be with us for some time still. We wished to establish a set of requirements that could practically be implemented in today's environment, tackling the major issues of unmanaged office documents, e-mails, and so on — but without requiring extensive preparatory work that might not be achievable. Something, in other words, that took organisations forward in a positive way, without presenting too many unassailable barriers — something that would work.

In addition, we wished to encourage the market development for electronic records and document management systems in the United Kingdom: encourage those coming from elsewhere to build a strong and robust base in the United Kingdom; encourage UK-based companies to develop new products; encourage the larger corporate names that could support this application area at a whole-of-government level; and encourage the development of expertise and skills in implementation amongst the consultant and systems integrator community. Without a fixed target to aim at and agreed in advance within central government, how could we expect suppliers to undertake development work at their own cost to provide what was for them a different type of product from what many of them already produced?

The project, therefore, took a practical 'bottoms-up' approach, drawing on the input of operational records managers and those charged with planning for corporate implementation; rather than a theoretical or academic approach. The requirements sought to define statements that could be used in actual statements of user requirement, rather than a framework of principles through which such a statement might be derived. Very different, then, from projects such as that conducted at Pittsburgh to determine a theoretical framework of functional requirements.

The organisations which took part in the definition of the requirements were drawn from a wide range of types: policy-making departments with a tradition of semi-autonomous branches; large and geographically dispersed service delivery departments; casework-based operations; organisations from the criminal justice sector; and small single-focus executive agencies. There could be no single specification which encompassed the entirety of business requirement for all these diverse business operations and activities.

Rather than attempt to be all-encompassing, the specification of requirement aimed to identify the core areas of electronic records management that are generic to all these situations, while recognising that a core specification would be insufficient for business needs by itself. Any organisation which made use of the generic specification would need to consider adapting it to its own business needs, and would need to add its own particular 'line of business' requirements. Products which incorporated the core requirements would be likely to offer these in the context of a broader range of facilities, that might include document management, image management, content management, workflow, advanced search facilities, and so on. No attempt was made to incorporate all these aspects in one document — far too large and unwieldy — with the danger that no product could meet them all. All specifications and products that are compliant, however, have this core in common.

The specification falls into three areas:

- core requirements, the requirements are specific to the electronic records management application:
  - — capture and classification (record and fileplan organisation);

  - authentic storage and management;
  - export and disposal management;

- additional mandatory requirements, which are necessary for ERM, but are also found in other application areas:
  - audit and authentication;
  - access control;
  - hybrid (paper and electronic) file management;

- supplementary requirements, which may be needed in some but all implementations:
  - advanced search and retrieval;
  - document management;
  - bulk import.

No attempt is made to specify non-functional requirements.

Compliance testing scheme

Following publication of the 1999 requirements, the PRO initiated a formal compliance testing regime, to enable commercial software suppliers to submit their products to a formal testing programme. A series of 15 detailed testing scripts were devised to test rigorously the mandatory core and additional (audit and access control) requirements, together with standard sample test data.

Initially, we invited statements of interest in undergoing a formal test and compiled a shortlist to enable the first six approvals. Subsequently, testing has been an ongoing activity and we now have 16 listed approved products. It is the products that we approve in the version that we have tested against the functionality we specified originally. We do not approve suppliers, nor other versions of the same software (though we may retest a subsequent major version if substantially different from the previous one and submitted to us). We test neither individual implementations, customisations nor beta versions. The product has to be available and supported in the UK market already. The logic behind this is that we have to concentrate resources on products of immediate practical use to our client departments to justify the substantial resources involved in formal tests.

The tests are conducted at supplier premises, on standard configurations of packages under evaluation, by direct 'hands-on' application of test scripts by PRO personnel. We do not host the tests ourselves so there is no question that our technical environment is displaying their product at anything less than the maximum advantage. The full range of tests takes three days to complete, and includes exploration of issues uncovered by the standard test scripts if necessary.

Product sets which successfully pass the tests are added to a list of approved systems, published on the PRO web site. Approval is given for a period of two years, to a specified major version of a product set. Product sets may consist of independent ERM/EDRM packages, or integrations of two or more packages together: all integrations and independent products are treated as separate product sets. Approval does not in itself simplify subsequent government procurement procedures (as DOMEA approval does), but the Office of Government Commerce (part of the Treasury department in the United Kingdom) has the approved products list in mind when compiling framework procurement contracts to comply with EU regulations.

The compliance testing scheme has been very successful in encouraging the development of new products and systems, and in bringing new players to the UK marketplace. The requirements provide a benchmark for development, and the compliance testing provides a level of quality assurance to departments wishing to acquire and EDRMS. Listing of an approved product set does not guarantee that its use will resolve all of their electronic records management problems. Many of the products on the market can just as readily be configured out of compliance with our requirements as into it. Rather it gives some confidence that in the opinion of the Public Record Office the product has the capability of meeting the cross overnmenet generic requirement. Although

use of the list of approved systems has not been made mandatory, and the list does not in any way control the gateway of access to the market, appearance on the list is treated as a mark of high commendation by government organisations. Consequently, it is an accolade actively sought by suppliers interested in government business.

The requirements and compliance testing regime are primarily developed for central government departments; but the work has been widely taken up and used by UK local government authorities and other public-sector organisations: most recently, by the Ministries of Justice and Foreign Affairs in the Dutch Government. It has, in effect, become a *de facto* standard within the United Kingdom.

To return briefly to what was touched on earlier about this wider interest, without careful reference back to the core aims of the project and of the Public Record Office it would have been easy to have been drawn into an open-ended 'certification' or 'accreditation' scheme (these terms are not used by us). We do not encourage anyone to adopt the requirements as a single line in their requirements statements: 'must be UK Public Record Office approved' as this would probably be done in ignorance of the detailed and central government specific work that led to the writing of the requirements in the first place. This would make little sense. Unfortunately of the other existing standards (see below), very few have recognised formal evaluation schemes and the most prominent (DoD 5015.2) is rather different in nature.

Other requirements standards

At the time when these functional requirements were first developed, there were hardly any examples of practical standards in this area: the principal one being the United States Department of Defence standard 5015.2. Other relevant standards in the United Kingdom included the more general British Standards for Information Security (BS7799, now an ISO standard) and a Code of Practice on the evidential value of electronic documents for legal admissibility (PD0008). All these were taken into account in the production of the requirements.

Since that time, a small number of related standards have evolved, and the relationship with the PRO standard is discussed below.

ISO 15489

The international standard for information and documentation — Records Management sets out a general framework for the creation, capture and management of organisational records. This is of course a much higher level and abstract document than a detailed set of functional requirements, and covers a broader range of activities and contexts. PRO requirements generally fit within this framework, but tend to be slightly less prescriptive in assumption in some areas: for example, the current PRO requirements do not assume a formalised classification scheme and allow for a more generalised mechanism of organisation. There is also less emphasis on any particular approach to records-management design: a functional approach is not assumed implicitly or explicitly.

MoReq

The PRO requirements formed a major input into the evolution of MoReq, with some of the same people involved in both developments. The MoReq work drew on a wider field — both public and private sector — and a broader geographical area. It took a rather more 'academic' approach; and also developed some aspects of the PRO work towards more formalised models: for example, in adopting a formal classification scheme, rather than a loose 'fileplan' concept.

In general, though, there is substantial consistency between the two specifications, and the current revision of the PRO requirements will take into account most of the differences.

DOMEA

Delegates may recall an interesting paper was delivered at the DLM in Brussels in 1999 by Michael Wettengel and Andreas Engel about the DOMEA programme then associated very much with the moving of the heart of German administration from Bonn to Berlin and concerned with specifying concepts and scenarios for software suppliers to demonstrate the ability to support in a formal

evaluation situation. It seems to me that this operates at a slightly different and perhaps a higher conceptual level from our requirements. Certainly, that is the perception of the supplier who appears on both our approved listings with whom we have had the opportunity informally to compare the two schemes. We would not wish, however, to overemphasise the differences: we are both concerned with the outcomes of the systems in terms of their practical deployment in a government business environment.

Other EU Member States

Apologies to any attendees if their software-approval schemes have been omitted from a mention here — one task for this DLM conference is to find out as much as we can about what everyone else is doing in Europe.

US DoD 5015.2

The DoD 5015.2 standard has recently undergone a revision, expanding its scope into areas such as access control, and with a more structured approach to metadata. It does remain, though, heavily angled towards the US environment, with each requirement traced back to a specific piece of US legislation. There are some significant differences between US and UK practice: in the United Kingdom the emphasis is on management at the file/folder level while in the US the emphasis is more on management at document level.

These differences have proved to be a difficulty in dealing with North American based products in the PRO compliance testing programme. ERM products from this region are developed specifically towards the DoD specification, and often find it difficult to successfully and reliably adapt their design model to a UK-based approach. In practice, for the PRO requirements there is much more commonality with the European-based MoReq, and the hope is that a close integration here will encourage development of a strong and consolidated market which Europe-based software suppliers will be eager to supply and some of those suppliers from further afield might find it worthwhile to develop towards.

Revising and updating the requirements

There has been much progress in both the understanding and application of electronic records management, and in the broader context of e-government infrastructure and services, since the first publication of the PRO functional requirements in 1999. One of the more obvious is the new ways of collaborative working to support e-business just alluded to. Another (and linked) shift is the convergence of electronic document and records management technology.

It is now time to revise the functional requirements to bring them up-to-date and to incorporate some of the work which has emerged in the interim period. This revision is at present under way, and will include consideration of the major developments which impact on UK central Government.

e-government developments

There has been considerable development in the infrastructure and standards for e-Government services in the last two to three years. A strategic decision to locate future developments in an IP-based environment has evolved into, among other things, a wide-ranging standard for interoperability known as e-GIF (government interoperability framework). This establishes XML as the first choice for data-exchange purposes, and is being complemented by the development of a range of XML schemas. In addition, e-GIF also mandates a set of other standard formats for browser accessibility, which effectively sets out formats for medium term maintenance of electronic records. The latest version of e-GIF also includes a metadata standard for resource discovery, with the identification of a standard element set. Both of these are now mandatory on all public sector information systems in the UK Government.

As part of the framework for achieving the 2004 target for ERM capability, government departments and agencies are required to develop statements of user requirement, setting out their own business needs in relation to EDRM systems. While the PRO generic specification will form the basis of much of this work, we strongly discourage outright adoption of a standard set of requirements.

The onus is on departments to use the generic specification as a platform for thinking through their own business need, which should be fully justifiable by a business case. This is in any case a prerequisite to secure funding from the Treasury. Quite a number of these statements are now available, and they are being fed into the revision work, in order to identify additional requirements which are common rather than specific.

Information policy legislation

The last couple of years has seen passing of several significant pieces of information policy legislation which affect electronic records in the United Kingdom. A revised Data Protection Act strengthens requirements for managing personal information, and a new Freedom of Information Act is due for full implementation in 2005. This will require government departments to be very rigorous in managing their electronic documents and records within a formal environment and according to agreed business rules. Additionally, a Human Rights Act and various other EU directives have an impact on ERM needs.

MoReq

As mentioned previously, the MoReq work will largely be incorporated into both the revised functional requirements and the revised compliance testing regime.

Records management metadata standard

We are working on specifying a set of records management metadata that will support the revised requirements. The e-government metadata standard (e-GMS) specifies the resource discovery metadata standard for the UK Government. It is a DC-derived scheme.

This will be more closely aligned with specified system functionality than any other scheme than perhaps the DoDs (and our functional requirement and hence the metadata) is more detailed. This will mean that it will specify very precisely the metadata capture mechanisms which in turn should allow us to be more confident about obligation levels.

We aim to produce records-management metadata that is consistent with the e-GMS for the purposes of resource discovery whilst also supporting the detailed processes of robust electronic records management. As some of you may be finding yourselves, this is a highly challenging task and some of our American colleagues doubt whether DC is a suitable foundation for this.

Future interoperability

One of the main aims of the revision will be to develop further the potential for interoperability between ERM systems in government. At present, many of the companies operating in this area are in the small-to-medium category and none, on their own, seem likely to supply most or the entirety of the UK Government. To give greater confidence to departments in selecting a software system, and to prevent future 'lock-in' which would hinder technological migration, it is necessary to work towards more generalised standards of interoperability. In the broader view, this is also vital for supporting 'joined-up' services and more effective management of knowledge and information within and between government organisations.

This will work at two main levels:

- development of an XML schema for metadata exchange, incorporating both resource discovery elements and current records management elements; suppliers will be required to develop export and import facilities which support this schema; and
- development of more consistency in business rules for managing information at a whole-of-government level, with standardised disposal rules and data-exchange protocols.

In addition, the functional requirements for current records management will be complemented by requirements for the medium term sustainability of electronic records for business use while in departments. My colleague Richard Blake delivered a paper on this subject in a parallel session on 7 May.

The possibilities of a standard XML schema based on the metadata standard to support the transfer of records both between end-of-line businesses within the UK Government and between government departments and the Public Record Office for archival preservation is one that excites us greatly.

Development of the market

The take-up of electronic document management in the UK central Government was relatively limited, not least as most of the early solutions in that area had no functionality for the management of records. We have now witnessed a distinct convergence between the electronic document and the records management markets, with some very large international software houses realising that their market penetration will be limited if they do not either develop records management capability in house or interface with specialist products that have it.

There is currently noticeable interest in development partnerships with Microsoft Inc., not least for the collaborative working potentialities of its SharePoint Portal Server to be harnessed to robust records management capability. We have also seen some of the leading document management suppliers signalling that they wish to add ERM functionality to their product either as an optional module or as part of the core functionality. A number of UK Government departments are also large users of Lotus Notes and there are some solutions beginning to emerge to facilitate the management of records in that environment also. Whilst Microsoft itself is not expected to enter this market, there are signs that IBM is already doing so.

We think there are distinct signs that some of the North American suppliers are waking up to the fact that there is sufficient commonality in approaches to ERM across the EU area to make it worthwhile market in its own right. This is a highly encouraging sign.

Conclusions

We have come a long way since 1999 in encouraging commercial software vendors to produce products that are fit for deployment in the UK central Government. We have a number of products apparently written specifically to meet UK Government requirements and a far closer match with them than those available at the beginning of this process.

ERM is seen far more in the wider information management environment — indeed we tend nowadays to refer to 'EDRM'. This reflects its place as an enabling technology for new ways of providing services to the citizen: changing the organisation rather than simply automating existing processes.

There is a corresponding convergence between DM and ERM technologies in the range of available software also and more interest from far larger software providers. The last product we added to our list of approved systems comes from a private sector accountancy background — ironically, we tested it as the Enron scandal with its allegations of the company's auditors shredding important documents was breaking news in the United States.

Local government, other administrations and private sector organisations are taking ever more interest in what we are doing. In short, records management appears to be becoming more 'mainstream' rather than a specialist or 'niche' market.

Requisitos funcionales de los sistemas de gestión de documentos electrónicos

Malcolm Todd, Stephen Harries

En la administración británica, la Public Record Office (PRO) es el principal organismo de gestión de archivos, encargado de fijar y mantener normas aplicables a todos los organismos de la administración central. Esta función se ejerce para los archivos actuales y para los archivos históricos; además, habida cuenta de la evolución hacia los documentos electrónicos, es muy importante aplicar tales normas en el momento de la creación y la introducción de los documentos, así como en las fases de gestión permanente y selección histórica. En 1999, la PRO publicó unos requisitos genéricos para la gestión de documentos electrónicos en la administración del Reino Unido. Esta publicación se elaboró principalmente con el fin de aplicar sistemas, herramientas y procedimientos que permitan almacenar y recuperar los documentos digitales por vía electrónica en toda la administración central del país de aquí a 2004, como parte del programa de modernización de la administración. La elaboración de estos requisitos funcionales perseguía dos objetivos de igual importancia: proponer una plataforma común para los organismos y los servicios públicos con vistas a definir proyectos de gestión de documentos y archivos electrónicos y realizar procedimientos de contratación pública, y proporcionar un marco de referencia para los proveedores de *software* con vistas al desarrollo de los productos necesarios. En 1999, el mercado de la gestión de documentos y archivos electrónicos en el Reino Unido estaba aún poco desarrollado, con muy pocos productos disponibles.

La primera versión de estos requisitos funcionales fue un proyecto elaborado en colaboración con un amplio abanico de organismos administrativos —distintos ministerios, grandes direcciones descentralizadas de prestaciones de servicios, servicios de asistencia social y administración judicial, y pequeñas agencias de ejecución de vocación única—. Teniendo en cuenta que una única especificación no puede razonablemente cubrir un abanico tan amplio de actividades, los trabajos se centraron en la definición de un núcleo de procedimientos básico para la gestión de los archivos electrónicos en estos distintos contextos: introducción de datos, clasificación, seguridad del almacenamiento y de la gestión, eliminación y exportación. Los requisitos para la gestión de documentos electrónicos, la recuperación de información, el desarrollo de las operaciones y la gestión de las imágenes tienen requisitos más específicos, que cualquier administración puede añadir a este núcleo básico, junto con sus propios requisitos funcionales. No obstante, estos requisitos básicos de la gestión de los archivos electrónicos (añadidos a otros trabajos sobre herramientas de procedimiento y de aplicación) contribuyen a uniformizar las prácticas en todas estas organizaciones y a diversificar los contextos operativos en la medida de lo posible.

La PRO prosiguió estos trabajos desarrollando un sistema riguroso de prueba y evaluación basado en estos requisitos funcionales. Los proveedores de *software* pueden someter grupos de productos a este procedimiento de prueba oficial, y si superan la prueba, sus productos se añaden a la lista de sistemas aprobados por la PRO después de haber demostrado su capacidad para cumplir los requisitos genéricos. Este sistema ha suscitado una evolución en el mercado: algunos proveedores del mercado han elaborado varios productos especialmente en función de estos requisitos, y su influencia es manifiesta en las últimas versiones de los programas existentes. Aunque de momento la mayoría de las empresas capaces de proponer un producto compatible son PYME, las grandes empresas aprovecharán en adelante las grandes oportunidades que ofrece la administración electrónica.

A grandes rasgos, el sistema es comparable con el gestionado por el Ministerio americano de Defensa, basándose en la norma DoD 5015.2. Esta norma DoD es muy específica del contexto americano, y cada especificación está vinculada a un texto concreto de legislación americana. Además, en la práctica se observan grandes diferencias (la más importante de

entre ellas es la práctica británica de gestionar en el nivel de los expedientes más que en el de los documentos). Las especificaciones británicas son muy pragmáticas y tienen por objeto describir sistemas que podrían razonablemente ser utilizados por organismos públicos en su situación actual, aunque esta situación diste mucho de ser perfecta.

Las especificaciones funcionales de la PRO ejercen una influencia creciente fuera de la administración central. Las administraciones locales y un número mucho más elevado de organizaciones las están integrando ahora en las especificaciones de sus sistemas de gestión de documentos y archivos electrónicos. Las organizaciones internacionales han manifestado su interés, así como las administraciones públicas europeas y, en particular, la administración y las instituciones de enseñanza superior de los Países Bajos. Las especificaciones de la PRO constituyen el punto de partida para el desarrollo del proyecto de especificaciones tipo (MoReq) de la Unión Europea.

La modernización y la administración electrónica van a buen paso en el Reino Unido y en Europa desde 1999, y actualmente se está trabajando en la elaboración de una segunda versión revisada de las especificaciones funcionales. Esta versión deberá estar disponible en la primavera de 2002. Las especificaciones básicas no cambiarán esencialmente, sino que serán completadas con algunos aspectos que han aparecido más recientemente en el Reino Unido: metadatos, interoperatividad y normas de autenticación para la administración electrónica y, finalmente, legislación sobre la libertad de información y sobre protección de datos. Esta revisión tendrá también por objeto garantizar la compatibilidad con las especificaciones MoReq e incorporará muchas de ellas cuando sea necesario. Por último, garantizará la compatibilidad con la norma de gestión documental ISO.

Esta revisión identificará y precisará más formalmente los elementos tipo de los metadatos de gestión de archivos, con el fin de completar los trabajos actuales sobre la generación de metadatos para las administraciones públicas, ya en curso en el Reino Unido y Europa. También se está prestando atención a aspectos como la utilidad y la gradualidad. La revisión de las especificaciones se verá apoyada por un nuevo sistema revisado y mejorado de prueba y evaluación, con modalidades transitorias aplicables a los sistemas que ya figuren en la lista de sistemas autorizados. En el Foro DLM se dará información complementaria sobre la evolución de estos trabajos.

Funktionsbezogene Anforderungen an Systeme zur Verwaltung elektronischer Unterlagen

Malcolm Todd, Stephen Harries

Innerhalb der britischen Regierung ist das Public Record Office (PRO) die Leitstelle für die Aktenverwaltung und damit für die Aufstellung und Einhaltung von Normen in allen zentralen Ämtern und Behörden zuständig. Diese Funktion gilt für aktuelle Dokumente ebenso wie für das historische Archiv, und mit der Entwicklung hin zu elektronischen Unterlagen wird es natürlich sehr wichtig, diese Normen zum Zeitpunkt der Erstellung und Erfassung wie auch in den Phasen der weiteren Verwaltung und historischen Auswahl anzuwenden. Im Jahre 1999 veröffentlichte das PRO allgemeingültige Anforderungen für die Verwaltung elektronischer Unterlagen im britischen Staatsapparat. Der eigentliche Hintergrund dafür ist das Ziel, Systeme, Einrichtungen und Vorgänge für die Aufbewahrung und das Retrieval elektronischer Unterlagen in elektronischer Form im Rahmen des Modernisierungsplans für Ämter

und Behörden bis 2004 in der gesamten britischen Zentralregierung einzuführen. Die Vorgabe der funktionsbezogenen Anforderungen war mit zwei gleichermaßen bedeutsamen Zielsetzungen verbunden: die Bereitstellung einer gemeinsamen Plattform für Regierungsstellen bei der Festlegung von ERM- und EDRM-Projekten und die Durchführung von Beschaffungsaktivitäten; sowie die Festlegung einer Leistungsvorgabe für Software-Anbieter zur Entwicklung der erforderlichen Produkte. Im Jahre 1999 war der ERM/EDRM-Markt im Vereinigten Königreich mit einem begrenzten Angebot etablierter Produkte noch recht unterentwickelt.

Die erste Version der Anforderungen war ein Gemeinschaftsprojekt einer Vielzahl britischer staatlicher Stellen: mehrere Ministerien, große Versorgungsämter an ganz unterschiedlichen Standorten, Sozial- und Strafrechtsbehörden sowie kleine Einzelbehörden. In dem Bewusstsein, dass eine einzige Spezifikation den praktischen Zwecken einer solch großen Bandbreite von Tätigkeitsfeldern nicht genügen würde, konzentrierte man sich bei der Arbeit auf die Konzipierung des grundlegenden Kerns von ERM innerhalb dieses Gesamtrahmens: Erfassung, Klassifizierung, sichere Lagerung und Verwaltung sowie Aussonderung und Export. Anforderungen an EDM, Informationsretrieval, Workflow und Verwaltung von Grafiken sind häufiger spezifiziert worden; sie können diesem Kern von der jeweiligen Stelle zusammen mit ihren eigenen speziellen Arbeitsanforderungen hinzugefügt werden. Die ERM-Kernanforderungen (mit anderen Arbeiten zu Verfahrens- und Implementierungstools) helfen jedoch, die Verfahren übergreifend für alle diese Stellen und unterschiedlichen operationellen Rahmenbedingungen möglichst zu vereinheitlichen.

Das PRO führte diese Arbeit noch weiter, indem es ein detailliertes Erprobungs- und Evaluierungsprogramm auf der Basis der funktionsspezifischen Anforderungen entwickelte. Softwareanbieter können Produktpakete für dieses formale Testprogramm einreichen und werden im Erfolgsfall in eine Liste von PRO-zugelassenen Systemen aufgenommen, die nachweislich zur Erfüllung der allgemeinen Anforderungen in der Lage sind. Das Programm hat sich als sehr erfolgreich für den Ausbau neuer Angebote am Markt erwiesen. So haben gewerbliche Anbieter mehrere Produkte speziell für diese Anforderungen entwickelt, und bei bereits vorhandenen Produktpaketen sind Folgeversionen offensichtlich entsprechend angepasst worden. Obwohl es sich bei den meisten Firmen, die in der Lage sind, ein sachgerechtes Produkt anzubieten, derzeit um kleine und mittlere Unternehmen handelt, ziehen die umfangreichen Möglichkeiten, die sich mit E-Government eröffnen, inzwischen auch das Interesse großer Namen der Branche auf sich.

Das Programm ist im weitesten Sinne vergleichbar mit dem des US-Verteidigungsministeriums auf der Basis von DOD5015.2. Der DoD-Standard ist allerdings stark auf die USA bezogen, und jede Anforderung ist mit einem bestimmten Rechtsakt verknüpft. Auch in der praktischen Handhabung bestehen einige erhebliche Unterschiede (vor allem erfolgt im Vereinigten Königreich die Verwaltung nicht auf der Dokumenten-, sondern auf der Aktenebene). Die britischen Anforderungen sollten sehr praxisorientiert sein und Systeme beschreiben, die von Regierungsstellen sinnvoll unter den Umständen eingesetzt werden können, die sie bei sich vorfinden, auch wenn diese Umstände vielleicht nicht gerade vollkommen sind.

Die funktionsbezogenen Anforderungen des PRO gewinnen einen immer stärkeren Einfluss außerhalb der britischen Zentralregierung. Kommunalverwaltungen, also eine sehr viel größere Zahl staatlicher Stellen, beziehen jetzt EDRM-Systeme in ihre Spezifikationen ein. Ihr Interesse bekundet haben auch internationale Organisationen und der öffentliche Sektor Europas, insbesondere die Regierung und das Hochschulwesen der Niederlande. Die PRO-Anforderungen bildeten einen wichtigen Ausgangspunkt für die Entwicklung des MoReq-Projekts der EU.

Seit 1999 hat sich in Sachen Modernisierung und E-Government im Vereinigten Königreich und in Europa eine rasante Entwicklung vollzogen, und derzeit wird an einer zweiten, revidierten Ausgabe der ERM-Funktionsanforderungen gearbeitet, die im Frühjahr 2002 erscheinen soll. Die ERM-Kernanforderungen werden sich nicht grundlegend ändern, jedoch um

Aspekte ergänzt werden, die in letzter Zeit im Vereinigten Königreich stärker in den Vordergrund gerückt sind: Metadaten, Interoperabilität und Authentizitätsnormen für e-Government sowie Rechtsvorschriften zu Informationsfreiheit und Datenschutz. Außerdem geht es bei dieser Überarbeitung darum, dass die Neufassung mit den MoReq-Anforderungen im Einklang steht, von denen gegebenenfalls viele aufgenommen werden. Auch die Kompatibilität mit der ISO-Norm für Schriftgutverwaltung wird sichergestellt.

Bei der Revision werden Standardelemente für Metadaten der Schriftgutverwaltung benannt und deutlicher herausgestellt, um die im Vereinigten Königreich und in Europa bereits abgeschlossenen Arbeiten zu Fundstellen-Metadaten für staatliche Verwaltungen zu ergänzen und einzubeziehen. Beachtung finden werden auch Fragen der Nutzbarkeit und Skalierbarkeit. Die revidierten Anforderungen werden durch eine überarbeitete und verbesserte Prüf- und Bewertungsregelung gestützt, in der Übergangsmodalitäten für bestehende Einträge im Verzeichnis der zugelassenen Systeme vorgesehen sind. Weitere Informationen zum Fortgang dieser Arbeit werden auf dem DLM-Forum vorgestellt.

Spécifications fonctionnelles des systèmes de gestion de documents électroniques

Malcolm Todd, Stephen Harries

Dans l'administration britannique, le Public Record Office (PRO) est le principal organisme de gestion des archives, chargé de fixer et de maintenir des normes applicables à tous les organismes de l'administration centrale. Il exerce cette fonction pour les archives courantes et pour les archives historiques; en outre, compte tenu de l'évolution vers les documents électroniques, il devient très important d'appliquer de telles normes au moment de la création et de la saisie des documents aussi bien qu'aux stades de la gestion permanente et de la sélection historique. En 1999, le PRO publiait un cahier des charges générique pour la gestion des documents électroniques au sein de l'administration du Royaume-Uni. Cette publication intervenait principalement dans le but de mettre en œuvre des systèmes, des installations et des procédures permettant de stocker et de récupérer les documents numériques par voie électronique dans toute l'administration centrale du pays d'ici à 2004, en tant que partie intégrante de l'agenda de modernisation de l'administration. La production de ces spécifications fonctionnelles poursuivait deux objectifs de même importance: proposer une plate-forme commune aux organismes et aux services publics pour définir des projets de gestion de documents et d'archives électroniques et les conduire jusqu'aux procédures de passation des marchés, et fournir un cadre de référence aux fournisseurs de logiciels pour le développement des produits nécessaires. Le marché de la gestion des documents et des archives électroniques au Royaume-Uni était encore peu développé en 1999 et ne proposait que très peu de produits.

La première version de ce cahier des charges fonctionnel a été le fruit d'un projet élaboré en collaboration avec un large éventail d'organismes administratifs — ministères, grandes directions décentralisées de prestations de services, services d'assistance sociale et d'administration judiciaire et petites agences d'exécution à vocation unique. Prenant acte de ce qu'un seul et même cahier des charges ne pouvait raisonnablement pas couvrir un éventail d'activités aussi large, les travaux se sont concentrés sur la définition d'un noyau de procédures de base pour la gestion des archives électroniques dans ces différents contextes: saisie, classement, sécurisation du stockage et de la gestion, élimination et exportation. Les conditions requises pour la gestion des documents électroniques, la recherche documentaire, le dérou-

lement des opérations et la gestion des images font l'objet de spécifications plus fréquentes, que n'importe quelle administration peut ajouter à ce noyau de base, en même temps que son propre cahier des charges fonctionnel. Toutefois, ces spécifications de base de la gestion des archives électroniques (ajoutées à d'autres travaux sur les outils de procédure et de mise en œuvre) ont contribué à uniformiser les pratiques dans toutes ces organisations et à diversifier les cadres opérationnels chaque fois que cela était possible.

Le PRO a poursuivi ces travaux en développant un système rigoureux de test et d'évaluation à partir de ces spécifications fonctionnelles. Les fournisseurs de logiciels peuvent soumettre des groupes de produits à cette procédure de test officielle, et, s'ils réussissent l'épreuve, leurs produits sont ajoutés à la liste des systèmes bénéficiant de l'agrément du PRO après avoir fait la preuve de leur capacité à satisfaire aux conditions génériques. Ce système est parvenu à susciter des développements sur le marché (plusieurs produits ont ainsi été élaborés spécialement en fonction de ces spécifications par des fournisseurs du marché), et son influence est manifeste dans les dernières versions des progiciels existants. Même si la plupart des entreprises capables de proposer un produit compatible sont pour l'instant des PME, de grandes entreprises se saisissent désormais des opportunités significatives qu'offre l'objectif d'une administration électronique.

Dans ses grandes lignes, le système est comparable à celui géré par le ministère américain de la défense, reposant sur la norme DOD 5015.2. Cette norme DOD est très spécifique au contexte américain, chaque spécification étant corrélée à un texte particulier de la législation américaine. En outre, on observe des différences majeures dans les pratiques (la plus importante d'entre elles étant l'habitude britannique de gérer au niveau du dossier plutôt qu'à celui du document). Les spécifications britanniques se veulent très pragmatiques et visent à décrire des systèmes qui pourraient raisonnablement être mis en place par des organismes publics dans la situation dans laquelle ils se trouvent — même si cette situation est loin d'être parfaite.

Les spécifications fonctionnelles du PRO exercent une influence sans cesse croissante en dehors de l'administration centrale. Les administrations locales et un nombre nettement plus élevé encore d'organisations les intègrent désormais dans le cahier des charges de leur système de gestion de documents et d'archives électroniques. Des organisations internationales ont manifesté leur intérêt ainsi que des administrations publiques européennes, notamment l'administration et l'enseignement supérieur des Pays-Bas. Les spécifications du PRO constituent un point de départ majeur pour le développement du projet de spécifications types (MoReq) de l'UE.

La modernisation et l'administration électronique vont bon train au Royaume-Uni et en Europe depuis 1999, et des travaux sont désormais en cours pour élaborer une deuxième version révisée des spécifications fonctionnelles. Cette version devrait être livrée au printemps 2002. Les spécifications de base ne changeront pas fondamentalement, mais seront complétées par certains aspects apparus plus récemment au Royaume-Uni: métadonnées, interopérabilité et normes d'authentification pour l'administration électronique et, enfin, législation sur la liberté d'information et sur la protection des données. Cette révision aura également pour but d'assurer la compatibilité avec les spécifications MoReq et en incorporera un certain nombre chaque fois que nécessaire. Enfin, elle assurera la compatibilité avec la norme de gestion documentaire ISO.

Cette révision identifiera et précisera plus formellement les éléments types des métadonnées de gestion d'archives afin de compléter les travaux actuels sur la génération de métadonnées pour les administrations publiques, d'ores et déjà en cours au Royaume-Uni et en Europe. Des aspects comme l'utilisabilité et la gradualité font également l'objet d'un certain intérêt. Un nouveau système revu et amélioré de test et d'évaluation viendra en appui de la révision des spécifications, avec des modalités transitoires applicables aux systèmes figurant déjà sur la liste d'agrément. Des informations complémentaires seront données sur l'évolution de ces travaux lors du Forum DLM.

Parallel session 5

Wednesday, 8 May 2002

IMPROVING ACCESS TO KNOWLEDGE
Training and education for information experts and users

Chairperson:	Erik Ketelaar (The Netherlands)
Co-chair:	Elizabeth Shepherd (United Kingdom)
Rapporteur:	Catherine Dhérent (France)

Thijs Laeven

Mr Laeven is a freelance consultant in competence-driven human resource development and vocational education and training for knowledge and record-keeping management. He is senior partner of Innogration Management Consultants. On behalf of the Netherlands Cultural Heritage Association for Records Management and Archives (DIVA) and the Ministry of the Interior, he manages a virtual developers' network to produce a complete series of competence profiles for archivists and records managers in the Netherlands and he supervised the development of the www.divaprofielen.nl web site.

Professional competences for record-keeping staff

Thijs Laeven

Introduction

This paper reviews the development and the use of a complete series of professional competence profiles for record-keeping staff. The paper includes a demonstration of the web site that makes the profiles publicly accessible. The project was an initiative of the Ministry of the Interior and the Netherlands Cultural Heritage Association for Records Management and Archives.

In the period from summer 2000 until summer 2001 some 40 active professionals in the field of records management and archives have described the professional competences required in the coming years. They were engaged to develop a complete overview of all the professional competences that should be 'future-proof'. I will discuss the start of the project and the sources of inspiration and the design of the competences catalogue and the web site where the profiles can be generated. I will also discuss their usefulness in competence management and their potential transnational use.

Future-proof profiles

The competence profiles must be innovative and future-proof. The assignment that was commissioned to the working groups highlights the following:

1. the records continuum is a leading principle throughout the competence descriptions;
2. office automation and the emergence of ICT set the organisational context to respond to;
3. digitalisation of archival and record-keeping processes are to be omnipresent in the profiles;
4. a constant and dominant concern for service level and quality of performance;
5. demand driven by information needs in the primary business processes;
6. orientation towards document management strategies to create knowledge-based organisations;
7. change of perspective towards *ex ante* policy-making instead of mere *post hoc* operations;
8. awareness of lifelong learning as a necessity in an organisational learning setting;
9. compatibility of the profiles with those of neighbouring disciplines.

It should be stressed for the DLM-Forum audience that not only electronic document and records management is taken into account, but also the rather traditional tasks, in order to cover the whole domain including hybrid situations.

Need for new profiles

Taking into account the points 1 to 9 above, at least three other reasons motivated the contractors to have the profiles developed. First there was the expiration of a booklet with a series of so-called qualification profiles which needed an update. Second, these profiles limited themselves to administrative record-keeping in State government organisations only, without incorporating archival functions and other private or public organisations. Third, the sector itself was in urgent need for a comprehensive set of competence profiles to support further professionalisation and lifelong learning.

Besides, the ICT industry´s answer to the DLM message on electronic document and records management issued by the end of 2000 also stresses the need for a standard competence profile of the modern archivist and record-keeping professional in the information age to enhance professionalism and employability ([1]). These might be the starting point for the development of inno-

([1]) See *European archives news Insar newsletter* No 8 autumn 2000 or www.dlmforum.eu.org for the full text.

vative curricula for the initial and continuous education of information managers and record-keeping professionals.

(2) Michael Wettengel 'Core competencies for electronic record-keeping', Kew, Public Record Office, 3–4 June 1998.

Sources of inspiration

I will only mention the two principal sources from which we have got some inspiration. First, there is a paper by Michael Wettengel (2) in which he reports on the results of the working group on training matters of the DLM-Monitoring Committee. This report distinguishes six core competences, develops their content and associates them with learning objectives. The six competences are the following: archival, legal, organisational, methodological, information technology and systems design. These six competences are reduced to four competence domains in our approach by combining No 3 with No 4 and No 5 with No 6.

The second source of inspiration we found in Australia. In 1997 the 'National Records and Archives Competency Standards' were published. The content was thoroughly revised in 2001, not the basic principles. They give an exhaustive overview of 'what people do in the workplace at various levels and the standard to which they do it.' — i.e. professional tasks at seven different levels. They list 'all aspects of work performance (including) the particular skills and knowledge required to do the job and all the (…) attitudinal aspects of undertaking the job.' The standards cover 'the characteristics possessed by individuals that enable them to be either assessed or judged competent in a particular task (…).' — i.e. ways to assess the evidence of required knowledge, skills and attitudes.

The Australian standards contain a complete catalogue of learning 'units' that must be part of the curricula and programmes of initial and continuous education. The objective is to regulate and to standardise the quality of education in this field.

In our project we were less concerned with standardisation of learning objectives than with professional competences. Learning objectives are a great help to construct educational programmes in conformity with these standards and to judge the quality of such programmes. In our project we concentrated on the required professional knowledge, skills and attitudes that underlay professional behaviour or performance. These competence descriptions can be used by training organisations to develop learning objectives and programmes. In our view this is the responsibility of these organisations, not of the professional records and archives management sector itself. In other words: the sector may define its 'demands', but the training organisations are the supply side in the (education) market. Of course there is a link between the two sides, as there are links with the labour market (demand and supply side of employees) and with the Human Resource Management of record-keeping and archives organisations or departments (demand and supply side of development, career steering, mobility etc.). We consider our profiles as an instrument for the competence management (see below) within record-keeping and archives organisations or departments, of which education, recruitment of staff, human resource development, mobility and employability are very relevant aspects.

We were very inspired by the level distinction made by the Australians and by the level indicators or 'descriptors' they use. These indicators are very useful to describe the level of professional activities regardless of function name or professional content. They are generic and use elements like 'guidance' as opposed to 'autonomy', 'responsibility', 'complexity', 'routine' and 'constraints' versus 'judgment' and 'planning', or 'range of tasks and roles'. We adapted the system to our objectives and condensed the Australian seven levels to a more convenient three level distinction: strategic, tactical and operational.

Another very helpful element of the Australian standards was their distinction of what I would call 'moments' in the records continuum to classify the units. They use the following labels to identify four 'streams of activity':

- designing, creating and using systems that keep records;
- creating records and capturing them into record-keeping systems;
- maintaining and managing records over time;
- making records accessible.

Besides these specific labels, they use 'overall units' which sit across the four main streams and affect all areas of the record-keeping regime. All the units within the four main streams of activity and the overall units are worked out in full detail at seven levels. They are split up into 'elements' with corresponding 'performance criteria' and further details concerning 'required evidence' and assessment.

Development of the competences

A great number of field specialists from all the parties involved in the record-keeping and archives sector have been working together in a virtual developers' network to produce the set of competence profiles. All the individual contributions have been integrated into a first working document. After that the draft profiles were commented upon by an inner circle of developers and by an outer circle of future users and other specialists. After this feedback round they were revised. It was decided to 'publish' the profiles by means of an interactive web site. This obliged us to think over the user and demand driven interaction. Simultaneously we user-tested the draft version of the profiles. A series of experimental settings were selected to have the usability tested and the user guide developed. To give some examples of these user test situations: integration of the front-office function of archives and public libraries, matching of initial and post-initial school programmes, a trainee programme in a decentralised records management in public administration, development of function profiles for record-keeping managers and archivists in the non-public sector, actualisation of function profiles in a digital environment. These experiments were designed to test the transition of generic profiles to specific situations, hence to define the functional requirements of the user interface. At the same time they gave us concrete examples and demonstrations of 'how it works'. Eventually two half-products were delivered:

- the validated competence profiles; and
- the blueprint of the web site.

Competence catalogue

The profiles are divided into nine task-related profiles and one overall profile, as shown in Table 1. On the horizontal axis the nine domains cover the areas of the record-keeping and archival regime under recognisable labels. On the vertical axis three dominant orientations are shown: (strategic) policy-making, (tactical) instrumentation/implementation and operational respectively. Each cell in the graph is to be considered as a professional competence — to be understood as a bundle of knowledge, skills and attitudes required for a task, a job or a function to be executed in conformity with performance standards — at one of the levels. Within each cell you may find the descriptions of the knowledge, skills and attitudes that constitute together the corresponding competence. Number 10 profile is an umbrella layer of area-independent overall competences at the same three levels. These competences cover the full range of recurrent elements in the task-related profiles that apply to all areas. General competences (such as flexibility, integrity) are deliberately left out of the system; only competences related to record-keeping and archival core-business have been retained.

It is relevant to say that in the Netherlands for the record-keeping and archives sector there exist no national standards for function or task profiles. For instance, what is labelled records manager in one vacancy text may cover completely different tasks than in another. other indicator of the lack of unity is the absence of a specific union for record-keeping or archives workers, hence the absence of standard collective labour agreements. So we had to define generic competence profiles and made a web site to support users in making them appropriate for their own situation.

Table 1. The Netherlands competences catalogue

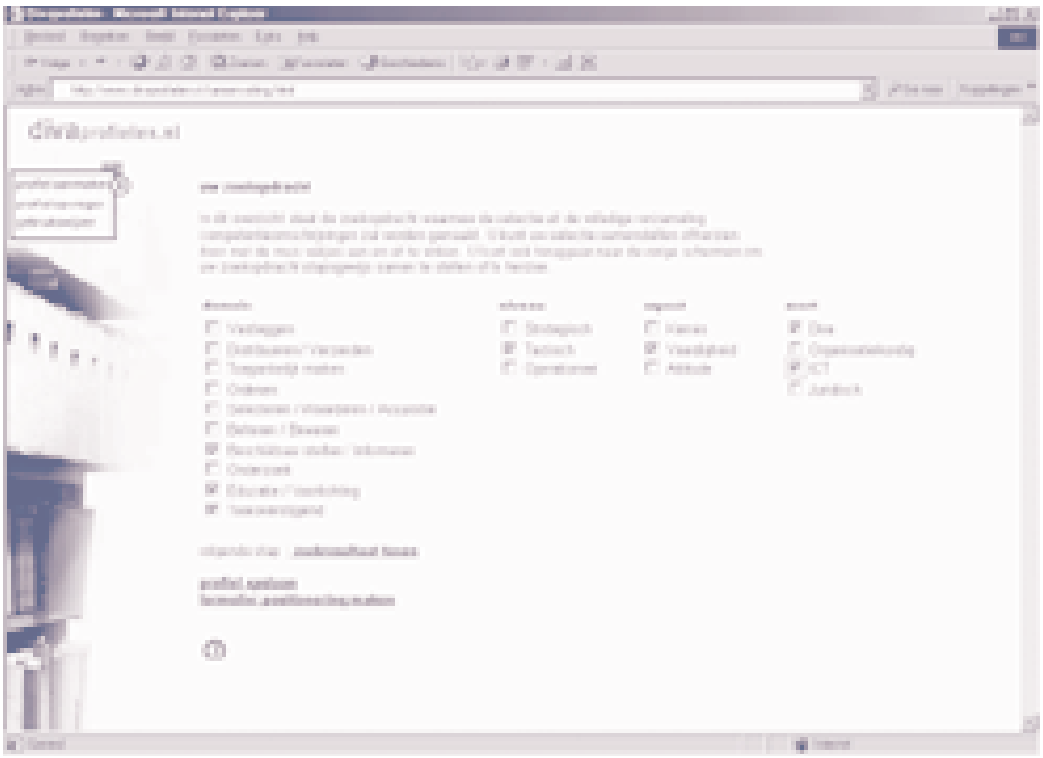
Record-keeping and archivals tasks in 9 domains plus 1 overall profile

capture	distribution & mail	accessibility and description	filing system	acquisition and appraisal and selection and disposal	preservation and manteinance	availability	research	communication and education	overall profile	*task's main orientation*
... task independent overall competences ...										
										strategic
										tactical
										operational

Competence web site

To enlarge usability and wide access to the profiles, it was decided to 'publish' them not in a traditional booklet but on a special web site. The website gives access to the database of all the (nearly 850) competence descriptions. The following URL www.divaprofielen.nl gives free and unlimited access to the web site. The user is guided through four screens that assist to formulate a query that makes a selection in the database. The selection is made on four parameters:

1. relevant domains (as shown in Table 1);
2. level (as shown in the table);
3. aspect (knowledge, skill, attitude);
4. content (record-keeping, organisational, ICT, legal).

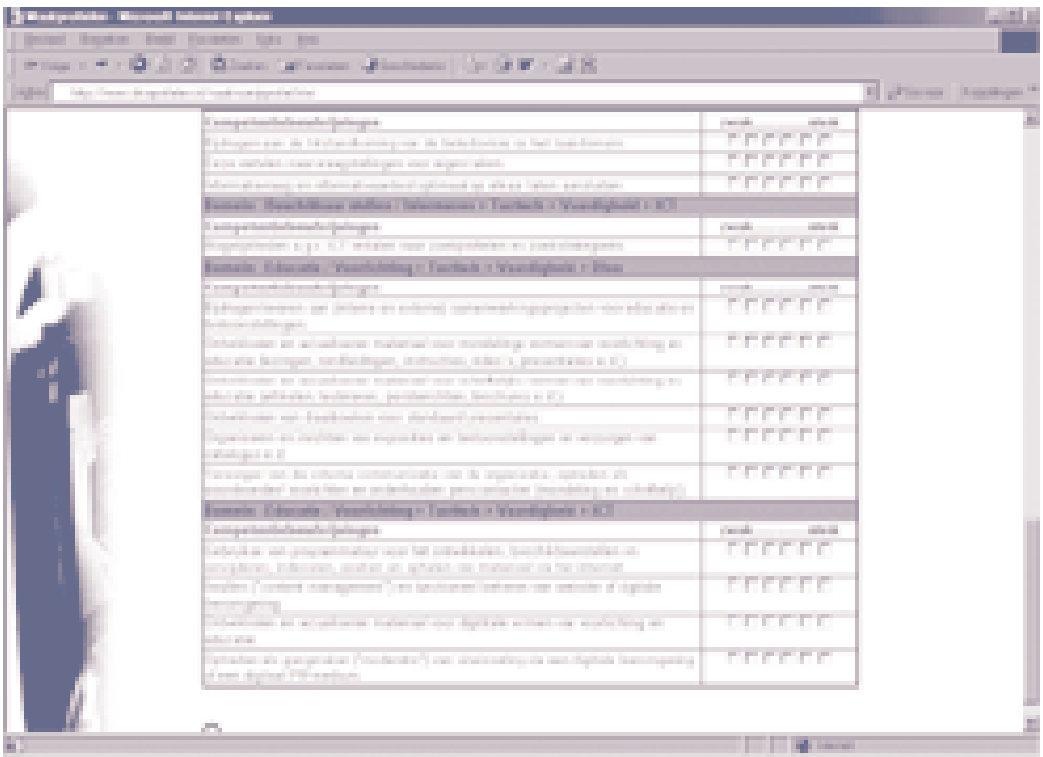


Screen dump No 1. The central page in the web site; simply ticking boxes to define a query that produces a competence profile

The screen dump No 1 shows the central web page in which the user can define a query simply by ticking boxes. The fewer boxes are marked, the more the user restricts the query and the more the results are exact, as you can imagine. Each of the four columns represents one of the parameters mentioned above. By ticking boxes the user makes the generic profiles specific for his or her own specific situation.

This tailormade selection is presented in a table which the user can either save on the web site (password-protected) or on the company's network/personal computer and print locally. The profile can also be presented in combination with some kind of positioning. The user (or the manager, the colleagues, the clients or the subordinates) can give scores that reflect the mastering of the competences by the subject. This may position the subject's mastering of the competences with reference to future competence needs. In this respect the profiles are useful for continuous post-vocational training needs identification and the management of learning programmes. The users targeted by the profiles are the professionals themselves, their (also non-specialist) managers and Human Resource Management staff. But also training institutions can use the profiles to develop their curriculum.

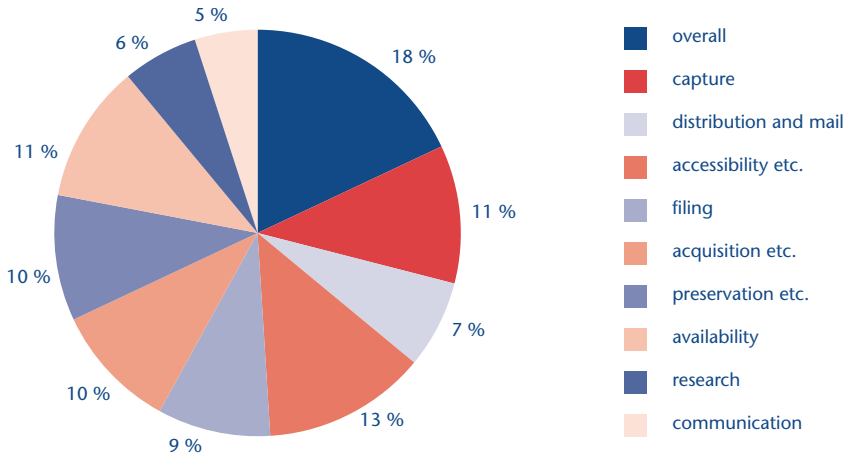
Screen dump No 2. Shows a result page in the web site; part of a competence profile for a front-office employee



The screen dump No 2 shows part of a profile for a front office employee. The right section of the page contains boxes that can be used for the positioning or the self positioning of the targeted employee. This can be done off line and the use can save the blank form and the filled out form locally. To make the results even more accurate, the user can wordprocess the profile and add or delete or reformulate competence descriptions.

User data of the web site

Estimations of the population of workers in the field of records management and archives in the Netherlands vary from some 5 000 people up to the double. In the first month after the web site becoming operational some 1 300 users visited the web site from which about 600 were more-than-once visitors. They had some 1 100 competence profiles generated. They were especially interested in the 'operational' level (45 %) and some less in the other two levels (each ±27 %). Refer to graph 1 for some more details. The knowledge and skills aspects both had a 36 % score and attitude 28 %. The record-keeping content was used in 45 % of the generated competence profiles; the organisational and the ICT content each about 22 % and the legal content about 11 %.



Graph No 1. User results after first month: domains used in generated competence profiles

Competence management

As stated earlier, we expect that the web site is especially suited for use in a competence-based human resource management context. The main functionality is: (1) to generate tailor made competence profiles; and (2) to support individual or collective (self-) positioning of employees. A manager can find an answer to the question 'What competences will be required in two years for the electronic document and records management staff in my local government administration?'. But he or she may also try to position the employees as far as their actual scores are concerned and estimate the 'distance' or discrepancy between the competences already present and those to be developed. You may think of applications like: recruitment and selection of staff, career management, placement of employees after a reorganisation, continuous learning and education or development, coaching and so on. You may also think of all kinds of activities related to 'bridging the gap' between the competences already present and those not yet present. Training institutions can use the competences catalogue (demand side in the competence market) to develop programmes and courses.

Future developments and extensions of the web site functionalities may include such things as a course selector that helps to find appropriate courses (initial or post vocational) to support employees in acquiring the relevant competences. Or a web-based vacancy matching device that can bring together personnel wanted and professionals looking for new situations.

The great advantage of the web site is to give to the different parties concerned a common 'language' to discuss competences.

Transnationality and standardisation

Inspiration for this approach was found in European and Australian examples (see above). As the competences in the Netherlands catalogue are generic, with only limited reference to specific Dutch law and by-law, they can be relevant for other European countries or easily be made so. As the nine points mentioned above are largely transnational and insist on electronic document and records management, the resulting profiles may be of some interest for the development of European competence standards. The database of competence elements and the interface of the web site can be brought into a European platform to develop common competences for the record-keeping and archives profession. Apart from the international relevance (think of transnational employability) you can also think of developing further the standard core competencies developed by the DLM working party (see footnote) into standard profiles.

Competencias profesionales del personal de los archivos

Thijs Laeven

El presente documento examina la evolución y la aplicación de una serie completa de perfiles de competencias profesionales, destinados al personal de los archivos. La presentación incluye una demostración del sitio web que permite difundir estos perfiles al público. El proyecto es el fruto de una iniciativa conjunta del Ministerio de Interior y de la Asociación neerlandesa para el patrimonio cultural.

Entre el verano de 2000 y el verano de 2001, un gran número de profesionales del ámbito de la gestión documental y los archivos procedieron a la descripción de las competencias profesionales requeridas en los próximos años. Estas personas se comprometieron a esbozar un cuadro completo de todas las competencias profesionales que deberían resistir el paso del tiempo.

Perfiles profesionales a prueba del tiempo

Los perfiles de competencias deben ser innovadores y resistir el paso del tiempo. La misión confiada a los grupos de trabajo destaca los siguientes principios:

1) La continuidad documental es un principio clave en todas las descripciones de competencias.
2) La ofimática y la aparición de las tecnologías de la información y de las comunicaciones (TIC) fijan el contexto organizativo que permite hacerle frente.
3) La digitalización de los procesos de archivo y gestión documental es una constante en los perfiles de cualificación.
4) El nivel de servicio y la calidad de la ejecución son una preocupación permanente y dominante.
5) La demanda está dirigida por las necesidades de información de los procesos de administración primarios.
6) Deben aplicarse estrategias de gestión documental para crear organizaciones del conocimiento.
7) Es necesario cambiar de perspectiva y adoptar decisiones ex ante en vez de actuar más tarde.
8) La formación permanente debe considerarse una necesidad en un contexto de aprendizaje organizativo.
9) Los perfiles deben ser compatibles con los de las disciplinas vecinas.

Catálogo de competencias

Los perfiles profesionales se organizan en nueve perfiles específicos y un perfil global, según figura en el gráfico 1. En el eje horizontal, los nueve ámbitos indicados cubren la gestión documental y el archivo con designaciones fácilmente reconocibles. En el eje vertical se definen tres orientaciones dominantes: elaboración de decisiones (nivel estratégico), instrumentación/aplicación (nivel táctico) y nivel operativo. Cada celda del gráfico contiene una serie de competencias —es decir, un conjunto de conocimientos, cualificaciones y actitudes requeridas para una tarea, un puesto o una función— en uno de estos niveles. El perfil número 10 corresponde a una capa de competencias globales independientes, en estos mismos tres niveles. Las competencias cubren la totalidad de los elementos recurrentes en los perfiles de tareas, que se aplican a todos los ámbitos. Las competencias generales (flexibilidad e

integridad) se dejan voluntariamente fuera del sistema: sólo se han mantenido las competencias relativas a las actividades básicas de la gestión documental y archivística.

Sitio web de las competencias

Con el fin de poder ampliar las posibilidades de acceso a estos perfiles, se decidió publicarlos en un sitio web especial en vez de en un folleto clásico. Este sitio web da acceso a la base de datos que contiene la totalidad de las descripciones de competencias (casi 850). El usuario se ve guiado por cuatro pantallas que le ayudan a formular una pregunta de selección en la base de datos. Esta selección se realiza sobre la base de cuatro parámetros:

1) ámbitos pertinentes (tal como aparecen en el gráfico);
2) nivel (tal como aparece en el gráfico);
3) aspecto (conocimientos, cualificaciones, actitudes);
4) contenido (gestión de archivos, aspectos organizativos, TIC, jurídico).

Esta selección personalizada se presenta en forma de un cuadro que el usuario puede guardar en el sitio web (con una contraseña de protección), o en el ordenador en red o el PC de la empresa, para luego imprimirlo localmente. El perfil de cualificación también puede presentarse en asociación con un determinado tipo de posicionamiento. El usuario (o el director, o incluso los compañeros, los clientes o los subordinados) pueden dar una puntuación que refleje el cumplimiento de estas competencias por parte del sujeto. Esta puntuación puede permitir clasificar al sujeto con respecto a las futuras necesidades. A este respecto, los perfiles son útiles para definir las necesidades en cuanto a formación continua y para gestionar los programas de formación. Los usuarios contemplados por los perfiles son los propios profesionales, sus directivos (incluso no especialistas) y el personal de gestión de los recursos humanos. No obstante, los institutos de formación también pueden utilizar estos perfiles para desarrollar sus programas de enseñanza.

El autor realizó una demostración en directo del funcionamiento del sitio web, que está operativo desde enero de 2002. De esta manera, también pudieron presentarse las primeras experiencias de los usuarios.

Transnacionalidad

Este enfoque se vio inspirado por ejemplos recogidos en Australia y en Norteamérica. Dado que las cualificaciones del catálogo de los Países Bajos son genéricas y apenas hacen referencia a la legislación y a la normativa neerlandesas, pueden ser aplicables a otros países europeos o llegar a serlo. Puesto que los nueve ámbitos mencionados anteriormente revisten un carácter ampliamente transnacional y hacen hincapié en la gestión de los documentos electrónicos, los perfiles resultantes pueden presentar un cierto interés para la elaboración de normas de cualificación europeas.

Fachkompetenzen von Schriftgutverwaltern

Thijs Laeven

Dieser Beitrag befasst sich mit der Entwicklung und Anwendung einer kompletten Reihe von Fachkompetenzprofilen für Schriftgutverwalter. Dazu gehört auch eine Demonstration der Website, auf der die Profile öffentlich einsehbar sind. Dieses Projekt geht auf eine Initiative zur Schriftgutverwaltung und Archivierung des Innenministeriums und des niederländischen Kulturgutverbands zurück.

Im Zeitraum Sommer 2000 bis Sommer 2001 hat eine große Zahl von Fachleuten auf dem Gebiet der Schriftgutverwaltung und Archivierung die in den kommenden Jahren benötigten Fachkompetenzen beschrieben. Sie beteiligten sich an der Erstellung einer vollständigen Übersicht über alle fachlichen Kompetenzen, die „zukunftssicher" sein sollten.

Zukunftssichere Profile

Die Kompetenzprofile müssen innovativ und zukunftssicher sein. In der Aufgabe, die den Arbeitsgruppen gestellt wurde, wird Folgendes hervorgehoben:

1. Das Schriftgut-Kontinuum ist Leitgrundsatz in allen Kompetenzbeschreibungen;
2. Büroautomatisierung und Einführung von IKT geben den organisatorischen Rahmen vor, auf den es sich einzustellen gilt;
3. die Digitalisierung von Verfahren der Archivierung und Aktenführung muss in den Profilen allgegenwärtig sein;
4. ständige und vorrangige Beachtung von Leistungsumfang und -qualität;
5. Nachfrage orientiert an Informationsbedarf in den primären Geschäftsprozessen;
6. Ausrichtung auf Strategien des Dokumentenmanagement zur Schaffung wissensbasierter Organisationen;
7. Veränderung der Sichtweise in Richtung einer vorausschauenden Entscheidungsfindung anstelle reiner Post-hoc-Abläufe;
8. Anerkennung des lebenslangen Lernens als Notwendigkeit für den Lernprozess in einer Organisation;
9. Kompatibilität der Profile mit denen benachbarter Disziplinen.

Kompetenzkatalog

Die Profile sind, wie in Graphik 1 dargestellt, in neun aufgabenbezogene Profile und ein übergreifendes Profil gegliedert. Auf der Horizontalachse sind die neun Aufgabenbereiche der Schriftgutverwaltung und Archivierung als Rubriken aufgeführt. Auf der Vertikalachse sind drei vorherrschende Ausrichtungen dargestellt: strategisch (Entscheidungsfindung), taktisch (Ausstattung/Umsetzung) und operativ. Jede Zelle in der Graphik enthält auf einer der Ebenen eine bestimmte Anzahl von Kompetenzen, die als Bündel von Kenntnissen, Fertigkeiten und Einstellungen zu verstehen sind, die für eine Aufgabe, eine Tätigkeit oder eine Funktion benötigt werden. Profil Nummer 10 ist eine übergreifende Ebene von bereichsunabhängigen Kompetenzen für diese drei Ebenen. Diese Kompetenzen decken die gesamte Bandbreite wiederkehrender Elemente in den aufgabenspezifischen Profilen ab, die für alle Bereiche gelten. Allgemeine Anforderungen wie Flexibilität und Integrität sind bewusst ausgelassen worden; geblieben sind nur Kompetenzen mit Bezug zum Kerngebiet Schriftgutverwaltung und Archivierung.

Kompetenzwebsite

Um die Nutzbarkeit der Profile zu verbessern und einen breiten Zugriff darauf zu ermöglichen, wurde entschieden, sie nicht in einer herkömmlichen Broschüre, sondern auf einer eigenen Website zu „publizieren". Über die Website hat man Zugriff auf alle (fast 850) Kompetenzbeschreibungen. Der Benutzer wird durch vier Fenster geführt, die dabei helfen, eine Anfrage zu formulieren, mit der dann in der Datenbank eine Auswahl getroffen wird. Die Auswahl erfolgt nach vier Parametern:

1. relevante Bereiche (laut Tabelle),
2. Ebene (laut Tabelle),
3. Aspekt (Kenntnisse, Fertigkeiten, Einstellung),
4. Inhalt (Schriftgutverwaltung, Organisation, IKT, Recht).

Die Ergebnisse der Anfrage werden in einer Tabelle dargestellt, die der Nutzer entweder auf der Website (passwortgeschützt) oder im Netz/PC des Unternehmens speichern und dann bei sich ausdrucken kann. Das Profil kann auch in Verbindung mit einer Art Positionierung dargestellt werden. Der Benutzer (bzw. der Vorgesetzte, die Kollegen, die Kunden oder Untergebenen) kann je nach Beherrschung der Kompetenzen Punkte vergeben. Damit lässt sich die Beherrschung der Kompetenzen im Hinblick auf künftige Anforderungen positionieren. In dieser Hinsicht sind die Profile nützlich für die Ermittlung des Weiterbildungsbedarfs und die Gestaltung von Lernprogrammen. Zielgruppe der Profile sind die Fachleute selbst, ihre (auch nicht im Fachgebiet tätigen) Vorgesetzten und die Mitarbeiter der Personalabteilung. Aber auch Bildungseinrichtungen können die Profile für ihre Lehrpläne nutzen.

Der Verfasser präsentiert eine Live-Demonstration der Website, die seit Januar 2002 online ist. Somit können auch erste Erfahrungen von Benutzern vorgestellt werden.

Transnationalität

Den Anstoß für diese Herangehensweise gaben australische und nordamerikanische Vorbilder. Da die Kompetenzen im niederländischen Katalog generisch sind und nur selten auf spezielle Rechtsvorschriften und Bestimmungen des Landes verwiesen wird, können sie für andere europäische Länder übernommen oder problemlos angepasst werden. Die oben genannten neun Punkte sind weitgehend transnational und konzentrieren sich auf die elektronische Schriftgutverwaltung, so dass die entsprechenden Profile von Interesse für die Entwicklung europäischer Kompetenzstandards sein könnten.

Compétences professionnelles des personnels des archives

Thijs Laeven

Le présent exposé passe en revue l'évolution et l'application d'une série complète de profils de compétences professionnelles destinés aux personnels des archives. La présentation inclut une démonstration du site web qui permet de diffuser ces profils auprès du public. Le projet est le fruit d'une initiative conjointe du ministère de l'intérieur et de l'Association néerlandaise pour le patrimoine culturel.

Entre l'été 2000 et l'été 2001, un grand nombre de professionnels de la gestion documentaire et des archives ont procédé à la description des compétences professionnelles requises dans les années à venir. Ils s'étaient engagés à brosser un tableau complet de toutes les compétences qui devraient «résister à l'épreuve du temps».

Profils professionnels à l'épreuve du temps

Les profils de compétences doivent être innovants et résister à l'épreuve du temps. La mission qui était confiée aux groupes de travail met en exergue les principes suivants:

1) la continuité documentaire est un principe clé dans toutes les descriptions de compétences;
2) la bureautique et l'émergence des TIC fixent le contexte organisationnel permettant d'y satisfaire;
3) la numérisation des processus d'archivage et de gestion documentaire est une constante des profils de qualification;
4) le niveau de service et la qualité d'exécution sont un souci permanent et prépondérant;
5) la demande est dirigée par les besoins informationnels des processus d'administration primaires;
6) des stratégies de gestion documentaire doivent être mises en œuvre pour créer les organisations de la connaissance;
7) il faut changer de perspective et adopter des décisions ex ante plutôt qu'agir après coup;
8) la formation tout au long de la vie doit être comprise comme une nécessité dans un contexte d'apprentissage organisationnel;
9) les profils de qualification doivent être compatibles avec ceux des disciplines voisines.

Catalogue de compétences

Les profils professionnels sont organisés en neuf profils spécifiques à une tâche plus un profil global, ainsi qu'il ressort du graphique 1. Sur l'axe horizontal, les neuf domaines indiqués couvrent la gestion documentaire et l'archivage sous des désignations aisément reconnaissables. Trois orientations dominantes sont définies sur l'axe vertical: élaboration des décisions (niveau stratégique), instrumentation/mise en œuvre (niveau tactique) et niveau opérationnel. Chaque cellule du graphique renferme un certain nombre de compétences — c'est-à-dire un ensemble de connaissances, de qualifications et d'attitudes requises pour une tâche, un poste ou une fonction — à l'un de ces niveaux. Le profil numéro 10 correspond à une couche de compétences globales hors domaines, recoupant ces trois mêmes niveaux. Les compétences couvrent la totalité des éléments récurrents dans les profils de tâches, qui s'appliquent à tous les domaines. Les compétences générales (flexibilité, intégrité) sont volontairement laissées en dehors du système: seules les compétences concernant les activités de base de la gestion documentaire et archivistique ont été retenues.

Site web des compétences

Afin de pouvoir élargir les possibilités d'accès à ces profils de qualification, il a été décidé de les publier non pas dans une brochure classique, mais sur un site web spécial. Ce site donne accès à la base de données renfermant la totalité des descriptions de compétences (soit presque 850). L'utilisateur est guidé par quatre écrans qui l'aident à formuler une requête de sélection dans la base de données. Cette sélection est opérée sur la base de quatre paramètres:

1) domaines pertinents (tels qu'ils apparaissent dans le graphique),
2) niveau (tel qu'il apparaît dans le graphique),

3) aspect (connaissances, qualifications, attitudes),

4) contenu (gestion d'archives, aspects organisationnels, TIC, juridique).

Cette sélection personnalisée se présente sous la forme d'un tableau que l'utilisateur peut soit sauvegarder sur le site web (mot de passe de protection), soit sur l'ordinateur en réseau ou le PC de l'entreprise, puis imprimer en local. Le profil de qualification peut également se présenter en association avec un certain type de positionnement. L'utilisateur (ou le dirigeant, voire les collègues, les clients ou les subordonnés) peut donner une note qui reflète la maîtrise de ces compétences par le sujet. Cette notation peut permettre de classer le sujet par référence aux futurs besoins de qualifications. À cet égard, les profils sont utiles pour identifier les besoins en matière de formation continue et pour gérer les programmes de formation. Les utilisateurs visés par les profils sont les professionnels eux-mêmes, leurs dirigeants (même non spécialistes) et les personnels de gestion des ressources humaines. Toutefois, les instituts de formation peuvent également utiliser ces profils pour mettre au point leurs programmes d'enseignement.

L'auteur a procédé à une démonstration en direct du fonctionnement du site, lequel est opérationnel depuis janvier 2002. Les premières expériences des utilisateurs ont donc pu être également présentées.

Transnationalité

Des exemples collectés en Australie et en Amérique du Nord ont inspiré cette approche. Étant donné que les qualifications du catalogue néerlandais sont génériques et ne font que peu référence à la législation et à la réglementation néerlandaises, elles peuvent être pertinentes pour d'autres pays européens ou le devenir aisément. Comme les neuf domaines mentionnés plus haut revêtent un caractère largement transnational et mettent l'accent sur la gestion des documents électroniques, les profils en résultant peuvent présenter un certain intérêt pour l'élaboration de normes de qualification européennes.

Peter Horsman

Peter Horsman (1947) works currently with the Archiefschool (Netherlands Institute for Archival Education and Research) in Amsterdam, where he coordinates the research programme. Before he held a position as Director of Information Policies at the National Archives in The Hague. He has over 25 years of experience in archives; the last 15 years he has been intensively involved in projects in the field of archival automation, including electronic records — among these the InterPARES project. He advised various organisations about record-keeping systems and archives management. From 1992–2000 he has been chair of the ICA Committee on information technology. At the moment he is a member of the ICA Committee on programme management. He gave papers and lectures on various subjects, in and outside the Netherlands, and published in a variety of journals. With Eric Ketelaar and Theo Thomassen he is editor-in-chief of the journal Archival science. *Peter Horsman was the project manager of the E-TERM project, developing a model curriculum with teaching materials for a traning course on electronic records management.*

E-TERM and beyond
Project overview

Peter Horsman

Introduction

E-TERM stands for European training programme on electronic records management. This paper reports briefly on the project, its objectives, its products, and its activities. Then, the second part sketches the outlines of a possible future implementation. This outline will demonstrate the flexibility of the E-TERM curriculum as it has been delivered by the project.

Need for continuous education

The time that archivists or records managers — I will use these terms as being interchangeable — could survive during their active working life, only based upon initial training and education has passed. As in most professions the working environment is currently continuously changing, and these changes follow each other rapidly. In particular information and communication technology requires updating of existing knowledge, acquiring new knowledge, understanding new phenomena, improving skills, and learning new ones. It is not even in the first place the technology that pushes the archivist to learn new things, and to keep learning, but rather the impact of the technology on organisational behaviour, documenting business processes, and on society as a whole. The role and form of information is changing, and consequently the form and structure of records and archives. Furthermore, user behaviour and expectations are changing as well. Record-keeping methods that might have been successfully in operation for decades, are being challenged now, and become subject to dramatic redesign. Authenticity and reliability of documents are no longer self-evident; the position and added value of the archivist and records manager are seriously questioned.

As is true for many professions, continuous education and training of record-keeping professionals is one possible answer, and one possible way to find solutions for problems, and answers to challenges. It is not for nothing that as of its very beginnings the DLM-Forum devotes one of its priorities to archival education.

Background

Indeed, the origin of the E-TERM project is firmly rooted in the DLM programme, and Hans Hofmann has been a driving force behind it. He convinced the Netherlands Archives School to invite various archival educators from all over Europe to discuss the possibilities of developing a model curriculum of a (vocational) training course on records management. This curriculum, the participants agreed upon. It could be based on an existing one, by then successfully deployed by the Archiefschool. But, such a project might benefit as well from a previous European Commission (Socrates) sponsored project: Recpro, which had developed a model curriculum for a university course on electronic records management.

Partners

As a result of a two-day meeting in spring 1999 in Amsterdam a consortium was formed, consisting of the former Recpro partners University of Northumbria at Newcastle (United Kingdom), the University of Tampere (Finland), the Fachhochschule Potsdam (Germany), and the Archiefschool, Amsterdam (Netherlands), to which joined University College London (United Kingdom), the Central Direction of the Italian State Archives, Rome, and the University of Porto (Portugal).

The consortium applied for a Leonardo funding, which finally — with some delay in decision-making — was granted by the end of 1999. I will not go into the details of the bureaucratic proce-

dures for application, contracting and reporting, however, but focus on the contents of the project.

(1) For the project blackboard software was used, but actually this software is interchangeable.
(2) As a matter of fact: one experience the project brought into light, is not only the convienience of electronic communication, but also the added value of real meetings. Social contact, knowing, and seeing each other increases mutual committment.

Objectives

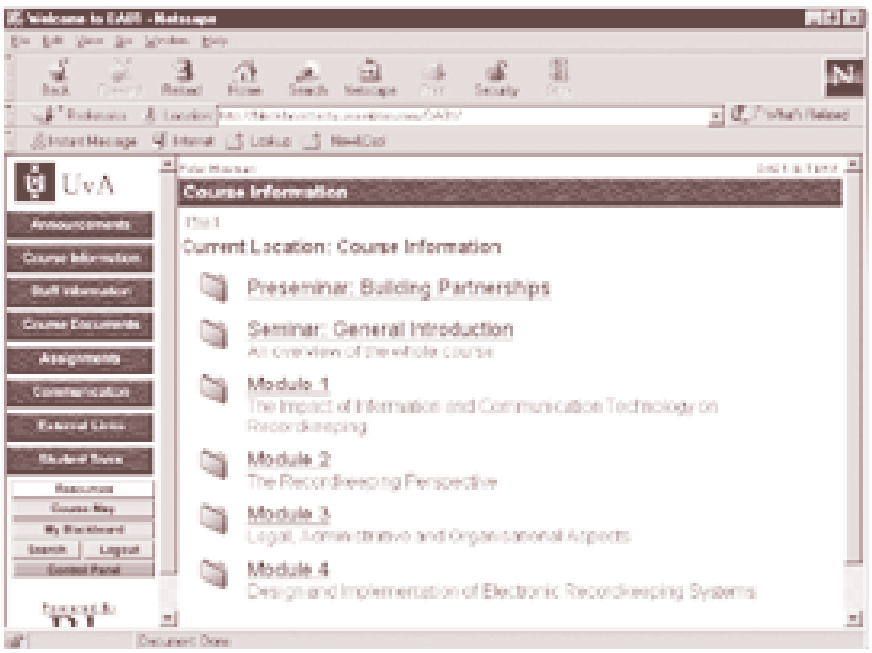
The main project goal was the development of a curriculum of a training course for practitioners, archivists, records managers, administrators, and eventually ICT professionals. The curriculum should be accompanied by teaching materials, such as texts and case studies. If there was one thing all partners fully agreed upon, it was certainly the need for materials suitable for training. Furthermore, the curriculum should be modelled in such a way — and the teaching materials selected accordingly — that the whole could be easily adapted for national or local needs, standards and practices. This implied, for instance, that individual training institutes, where ever in Europe, should be able to make modifications to add specific materials, and possibly to leave out materials selected in the project. These requirements called for a highly flexible course layout.

E-TERM deliverables

The project plan defined the deliverables: first, of course, the curriculum, that is the description of the modules (component parts); second the teaching materials that support the modules: texts, case studies, standards.

The third deliverable is the multilingual terminology, supporting professional communication, as well as the correct, transnational, interpretation of the texts and case studies. A software application was developed to organise the terms and their sources into a database.

Four: an open-learning environment in the form of an electronic learning system for structuring the project results, to experiment possibilities of distance learning, as well as for discussing results (1). Indeed, only a few meetings were planned during the project, most of the communication being done through the learning environment and regular e-mail (2). See Figure 1, below, for an example of a blackboard screen.



The fifth, and last deliverable is the E-TERM web site, for external communication and dissemination of the products.

During the project the plenimary products, including the curriculum were tested in pilots, which were carried out by the partners. The next sections discuss the products in some level of detail.

Course outline

The curriculum was initially designed along the themes, or rather: perspectives, of the original course of the Netherlands Archiefschool: (1) technology perspective, (2) legal, administrative and organisational perspective, (3) the record-keeping perspective, (4) standards and implementation, (5) developing records management strategies.

A preceding 'zero' module was then added, to inform archivists, records managers and ICT specialists about administration; to inform administrators and ICT specialists about record-keeping, and to update the ICT knowledge of both administrators and record keepers. This zero-module aims to facilitate working in the course with a mixed audience.

Teaching materials

As has been pointed out before teaching materials include texts, case studies, and standards. Each of these essential components should be linked to the curriculum through explanatory introduction, questions to guide reading and use, and specific assignments to make sure that the course participants really add to their knowledge, skills and understanding.

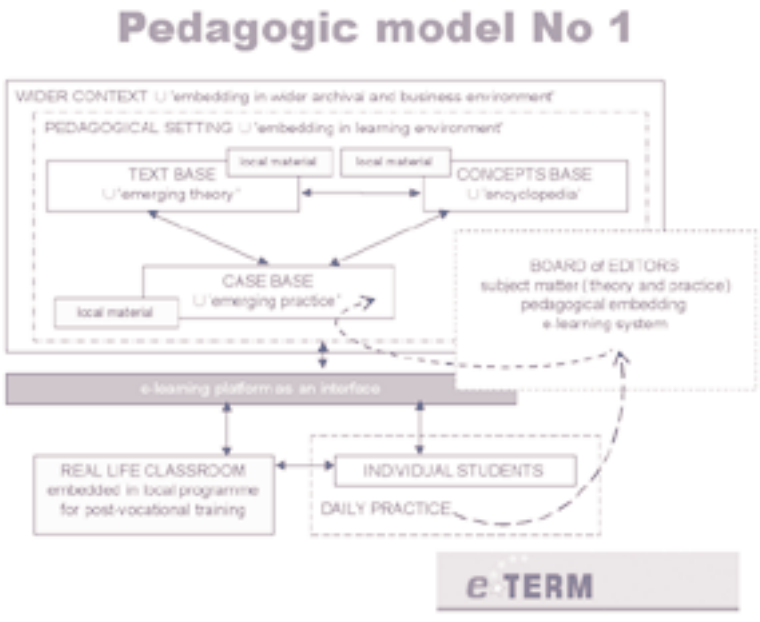
At the heart of the course are a 'master case-study', covering all identified aspects, and enabling the participants to compare a model case with their own daily working situation. The cases are meant to bridge the gap between theory and practice.

Pedagogic model

Underlying the course components, shortly described above, was a pedagogic model. Figure 2 depicts this model, showing how the various components relate to the learning process, and how this learning process is embedded in daily practice.

Project assessment

During the project, which formally ended by November 2001, the products were tested in several pilots, in the United Kingdom, in Germany, in Italy as well as in the Netherlands. First, the project itself benefited from these pilots. Second, the flexibility of the curriculum was clearly demonstrated: all pilots were different, with different course layouts, teaching styles and teaching materials, and target groups. Yet, looking beyond the pilots, one currently may discover new challenges in training. One of these emerging, pedagogic challenges is an orientation in curriculum devel-



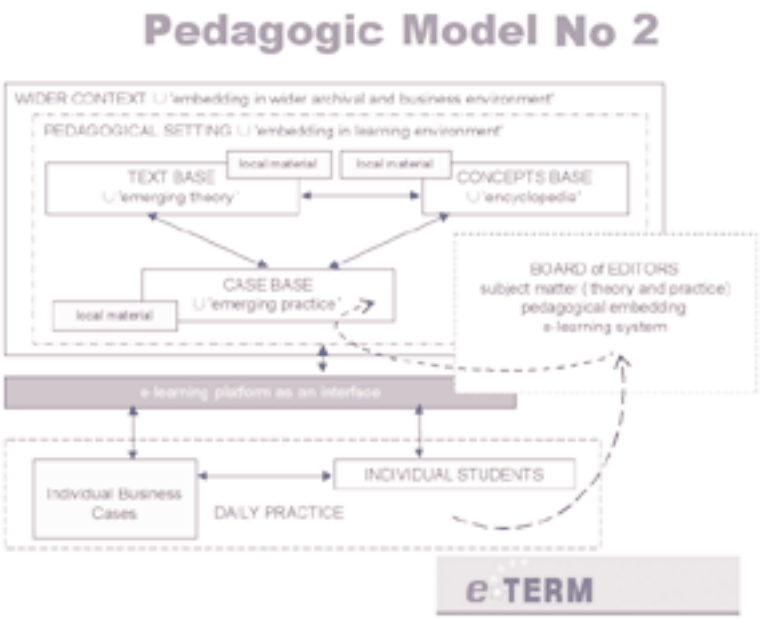
opment on competencies, rather than learning objectives ([3]). Or, to put it in another way: demand driven, or problem driven, instead of supply driven, or knowledge driven. In a way, E-TERM has been designed to teach, to transfer and explain knowledge, to elucidate concepts. Certainly, that is at least what initially is needed, but only to some extent. For a next step in development a necessary approach might be to start with questions the participants have to answer in their working practice, that is to say: to build the course on their — individual — business cases, to identify what knowledge and skills are needed to solve their problems, and finally to guide them in finding the answers as much as possible by themselves. The teacher becomes a coach, participants may learn from other cases, from theory, by seeking answers, by discussing in the (virtual and real) classroom.

([3]) See also the paper on this subject by Thijs Laeven, 'Professional competences for record-keeping staff'.

A model for going forward

The pedagogic model revisited

The original pedagogic model (Figure 1) had been constructed on a model master case. Even if this case study originated from a real life situation, it was modified to meet teaching purposes, and moreover, for the participants it remains something from another organisation, it is not their problem. An alternative model, presented in Figure 3, focuses on learning, rather than on teaching. It puts the real-life situation at the heart of the course, as an 'engine' for the learning process.



What is in a business case

Having for every participant in a training course an individual business case may have advantages for the particular participant; it has at least some difficulties for the teaching process. The most crucial question is how participants can share each other's experiences, how they can share learning. Therefore the business cases must be designed and structured in such a way that they can be mutually compared, that each participant is able to relate her/his own case to another case. Therefore the case descriptions should follow the same format.

A typical format may include a brief context description, a description and personal analysis of the problem, an identification of various stakeholders and other players in the process; finally, it should contain a personal vision on the desirable future situation. Of course, it can well be the case that in the course of learning these personal analyses and visions will be modified. At least, learning should bring the participant closer to the solution of the problem.

New course outline

Such a business case based approach requires likely a slightly different modularisation than has been foreseen in the initial E-TERM curriculum. A possible course outline might be built on a

(4) DIRKS underlies also the technical report which accompanies the ISO 15489 standard for records management.
(5) www.ucl.ac.uk/e-term

methodology for system development, that is to say: on a process of problem solving. It implies building the course curriculum on the process of implementation of record-keeping systems, rather than on emerging and existing record-keeping theories, or on the contextual aspects of records creation and use. However, the latter are certainly to be taken into account during development and implementation of record-keeping system, and consequently do appear in each of the modules. But different from the currently designed training course (E-TERM 1), they don't structure the curriculum.

Typically, modules of an E-TERM 2 course might include:

- writing the business case (an activity comparable with a preliminary survey in a development methodology);
- analysing business processes (comparable with a business systems analysis or definition study);
- identifying functional requirements for record-keeping;
- analysis of existing record-keeping systems, modifying existing systems or designing new record-keeping systems;
- implementation of new or modified systems; change management and related activities;
- systems and project evaluation.

A methodology, specifically developed for record-keeping systems is Australian DIRKS (4). This methodology, indeed, seems to offer good possibilities for structuring a course, but any other, national or organisational, methodology might work as well, as long as it can be adapted to the requirements of record-keeping systems development.

Conclusions

E-TERM, or rather E-TERM-1, was a first, important step in a process of developing training courses in the field of electronic records management for practitioners. Its results are available now from the E-TERM web site (5); they are ready to be rolled-out in European Union Member States. The course as it is, is complete in itself, and basically ready for national adaptation.

However, it is a first phase in an ongoing process. Expanding and in-depth development, as has been sketched in the previous sections, looks like a logical next step. It must provide the archivists and records managers with the required (and continuously changing) competencies to assess current record-keeping systems, and to improve or renew them according to user demands and organisational accountability requirements.

The E-TERM project demonstrated also the added value of cooperation. The project team operated in a networked environment, and would like to continue working as such. A future network of 'archives schools' may well be able to develop courses, to maintain them, to expand them, to share the best of national traditions, to continue bringing records and archives management on the required level — not only to preserve records, but above all, to satisfy users, and support both organisations and society.

Proyecto europeo de enseñanza de la gestión de los documentos electrónicos (e-term) ([1])

Peter Horsman

Con motivo de la reunión de expertos europeos en documentos electrónicos, celebrada en La Haya el 18 de junio de 1997 y que constituyó una continuación del Foro DLM, se abordó ampliamente el desarrollo de una formación europea en materia de gestión de los documentos electrónicos destinada a archiveros y a gestores de documentos. A resultas de este encuentro, la sección «Educación y formación» del Consejo Internacional de Archivos (CIA) presentó un balance muy positivo de los dos cursos presentados ese día: un seminario de cinco días organizado por la escuela de formación archivística de los Países Bajos y un curso organizado por la escuela de archivo de Marburgo (Alemania). Dado que el curso de Marburgo iba dirigido principalmente a los archiveros, y el curso de los Países Bajos a archiveros y a gestores de documentos, los participantes acordaron que el curso de los Países Bajos podía servir de modelo para el programa europeo que iba a prepararse.

A instancias de Hans Hofmann, archivero de la Comisión Europea e impulsor del Foro DLM, la escuela de archivo lanzó en la primavera de 1999 una invitación a instituciones homólogas de distintos países europeos. Todos los participantes habían admitido la necesidad de disponer de una formación, con material pedagógico adaptado. La solicitud se presentó en Bruselas y ganó una subvención en virtud del programa Leonardo da Vinci. El proyecto finalizó en diciembre de 2001.

Los participantes fueron los siguientes: University College London, University of Northumbria de Newcastle (ambas del Reino Unido), Ufficio centrale per i beni archivistici (Italia), Fachhochschule Potsdam (Alemania), Universidad de Tampere (Finlandia) y Universidad de Oporto (Portugal). El proyecto estuvo dirigido por la escuela de archivo de los Países Bajos.

El proyecto dio lugar a un programa de enseñanza y a una serie de materiales pedagógicos probados: textos (aspectos teóricos), estudios de casos (aplicaciones), terminología y ejercicios prácticos. El programa neerlandés inicial sirvió como punto de partida, pero fue modificado y ampliado.

El programa de enseñanza se concibe en cinco módulos, que confieren total flexibilidad según las necesidades locales. Estos cinco módulos pueden eventualmente ir precedidos de un módulo introductorio: introducción a las tecnologías de la información para archiveros y gestores (y administradores) de documentos, introducción a la teoría y a la práctica del archivo para profesionales de la informática y administradores y, por último, introducción a la gestión, destinada a los archiveros, a los gestores de documentos y a los profesionales de la informática.

Cada módulo aborda un tema concreto:

- impacto de las tecnologías de la información y de las comunicaciones (TIC) en las organizaciones;
- sistemas de archivo;
- aspectos jurídicos y organizativos;
- problemas de aplicación;
- estrategias y políticas.

Para cada uno de estos temas se buscaron y comentaron textos pertinentes, y se realizaron ejercicios prácticos. Un «caso modelo» puede utilizarse para concretar la teoría, de manera

([1]) European Training Programme on Electronic Records Management.

que los participantes puedan adaptar lo aprendido a su trabajo diario. El estudio de casos de menor alcance puede utilizarse para ilustrar temas concretos, como el correo electrónico, la transferencia de ficheros, la descripción, etc.

Como ya se había observado, la ambigüedad de la terminología constituye una de las dificultades concretas para la comprensión de la abundante literatura sobre los archivos electrónicos. Esta dificultad es aún más aguda en un medio multilingüe. Por ello, el proyecto ha tenido un resultado principal y esencial: la puesta a punto de una terminología multilingüe, que clarifica los conceptos de los distintos países.

Los distintos aspectos del proyecto se confiaron a los distintos socios del consorcio, en forma de lotes: Londres y Newcastle se dedicaron a los textos y a los estudios de casos; Potsdam e Italia, a la terminología, y Tampere y Oporto se dedicaron a las pruebas pero participaron también en otros trabajos. En realidad, todas las instituciones involucradas probaron el programa, y los resultados de los proyectos piloto resultaron muy útiles en la fase final del proyecto. Por último, Amsterdam se encargó de la gestión del proyecto y utilizó el programa como instrumento.

La formación se impartirá en su totalidad en un medio de aprendizaje electrónico. Los establecimientos de formación que deseen utilizar el aprendizaje a distancia pueden hacerlo. Por el contrario, los establecimientos que prefieran utilizar la enseñanza tradicional en salas pueden utilizar el mismo marco. Por supuesto, también es posible combinar los dos métodos de enseñanza.

En realidad, poco después del comienzo del proyecto, el consorcio decidió utilizar, con carácter experimental, un medio de aprendizaje electrónico (e-aprendizaje) particular para el desarrollo del proyecto. Por distintas razones, entre las que figura la difusión mundial y el seguimiento posventa, se escogió Blackboard. Aunque no se haya creado para desarrollar programas de formación, esta plataforma resulta muy útil para las actividades del consorcio. Los medios de aprendizaje electrónico se crean para aprovechar al máximo las ventajas de la TIC moderna: funcionan en un medio de Internet y poseen capacidad de correo electrónico, foros de debate, comunicación entre comunidades, e incluso enseñanza en línea. Todas estas posibilidades configuran un medio muy eficaz para diseñar cursos que impliquen una cooperación internacional.

La utilización de las nuevas tecnologías contribuye en efecto a la comunicación entre los socios. De esta manera, el presupuesto de las reuniones (gastos de viaje y alojamiento) fue bastante limitado. No obstante, esta visión tecnológica podría haber pecado de exceso de optimismo. Un encuentro real posee ventajas sobre un encuentro virtual, y los contactos sociales favorecen la comprensión mutua y el compromiso.

Ahora que el proyecto ha finalizado, los socios han decidido continuar su cooperación. Incluso con las dificultades tecnológicas, parece que la vía está abierta para verdaderos cursos europeos, organizados por un consorcio de establecimientos de formación para un público heterogéneo. ¿Y por qué no prever un objetivo aún más ambicioso: un centro europeo (virtual) de enseñanza superior en materia de archivos?

E-Term ([1]): ein europäisches Lern- und Lehrprojekt zu elektronischen Aufzeichnungen

P. Horsman

(1) E-TERM: European Training programme on Electronic Records Management

Auf der Tagung europäischer Sachverständiger zum Thema elektronische Aufzeichnungen am 18. Juni 1997 in Den Haag, dem Folgetreffen des DLM-Forums, wurde umfassend die Entwicklung eines europäischen Ausbildungsgangs zur elektronischen Schriftgutverwaltung diskutiert. Am Ende der Tagung legte die ICA-Sektion Aus- und Fortbildung im Archivwesen eine sehr positive Bewertung der beiden Ausbildungsgänge vor, die zuvor am selben Tag vorgestellt worden waren: einen fünftägigen Seminarzyklus der niederländischen Archiefschool und einen von der Archivschule Marburg entwickelten Lehrgang. Da sich der Marburger Lehrgang in erster Linie an Archivare wendet und der niederländische Kurs für Archivare wie auch für Schriftgutverwalter ausgelegt ist, einigten sich die Teilnehmer darauf, dass der niederländische Lehrgang als Modell für den zu konzipierenden europäischen Ausbildungsgang dienen sollte.

Auf Anregung von Hans Hofmann, Archivar der Europäischen Kommission und Initiator des DLM, lud die Archiefschool im Frühjahr 1999 Partnereinrichtungen aus mehreren europäischen Ländern ein. Alle Teilnehmer stimmten darin überein, dass ein Bedarf für einen Ausbildungsgang und entsprechende Lehrmaterialien besteht.

Daraufhin wurde in Brüssel ein Leonardo-Zuschuss beantragt und genehmigt. Das Projekt lief im Dezember 2001 aus.

Zum Konsortium der Teilnehmer gehörten: das University College London, die University of Northumbria in Newcastle, das Ufficio centrale per i beni archivistici (Italien), die Fachhochschule Potsdam (Deutschland), die Universität Tampere (Finnland) und die Universität Porto (Portugal); federführend war die Archivschule der Niederlande.

Aus dem Projekt gingen ein Lehrplan und ein geprüftes Paket von Unterrichtsmaterialien hervor: Texte (Theorie), Fallbeispiele (Praxis), Terminologie und Aufgabenstellungen. Der ursprüngliche Lehrplan der Niederlande diente als Vorlage, die dann abgewandelt und ausgebaut wurde.

Der Lehrplan hat einen modularen Aufbau, und seine fünf Module können an die jeweiligen Erfordernisse angepasst werden.

Wahlweise kann den fünf Modulen ein einführendes Modul vorangestellt werden: Einführung in die Informationstechnologie für Archivare und Schriftgutverwalter (und Verwaltungsmitarbeiter), Einführung in die archivische Theorie und Praxis für IKT-Personal und Verwaltungsmitarbeiter und Einführung in die Verwaltung für Archivare, Schriftgutverwalter und IKT-Personal.

Die Module des eigentlichen Lehrgangs behandeln folgende Themen:

- Auswirkungen der IKT auf Organisationen,
- Aktenführungssysteme,
- rechtliche und organisatorische Sichtweisen,
- Umsetzungsprobleme,
- Strategien und Maßnahmen.

Für jedes Thema wurden aussagekräftige Texte zusammengestellt, Kommentare verfasst und Aufgaben geschrieben. Die Theorie kann anhand eines „Musterfalls“ veranschaulicht wer-

den, so dass die Teilnehmer dann imstande sein müssten, das Gelernte in ihrer täglichen Arbeit anzuwenden. Kleinere Fallbeispiele dienen der Veranschaulichung bestimmter Einzelbereiche, z. B. E-Mail, Migration, Beschreibung usw.

Wie bereits ausgeführt, sind die Schwierigkeiten beim Verstehen der vielen neuen Veröffentlichungen vor allem auf terminologische Unklarheiten zurückzuführen. Noch deutlicher wird dies in einer mehrsprachigen Umgebung. Ein wichtiges und wesentliches Ergebnis des Projekts ist daher eine mehrsprachige Terminologie mit Begriffserklärungen.

Jeder Partner des Konsortiums wurde mit Teilen des Projekts betraut, die zu Arbeitspaketen zusammengefasst waren: London und Newcastle konzentrierten sich auf Texte und Fallbeispiele, Potsdam und Italien übernahmen die Hauptzuständigkeit für die Terminologie; Tampere und Porto waren für die Erprobung verantwortlich, trugen aber auch zu den Arbeitspaketen bei. Der Lehrplan wurde eigentlich von allen teilnehmenden Einrichtungen getestet, und die Ergebnisse der Pilotprojekte lieferten nützliche Erkenntnisse für die Schlussetappe des Projekts. Amsterdam war zuständig für das Projektmanagement und nutzte dafür den Lehrplan als zentrales Hilfsmittel.

Der gesamte Ausbildungsgang findet im Rahmen einer elektronischen Lernumgebung statt. Sie kann in dieser Form von Lehreinrichtungen, die Fernlehrgänge durchführen wollen, aber auch von Einrichtungen, die den traditionellen Unterrichtsraum bevorzugen, genutzt werden. Ebenso sind natürlich auch Kombinationen beider Lehrkonzepte möglich.

Kurz nach Anlauf des Projekts hatte das Konsortium beschlossen, versuchsweise eine bestimmte elektronische Lernumgebung für die Entwicklung des Projekts zu nutzen. Aus verschiedenen Gründen, so z. B. wegen seiner weltweiten Verbreitung und seines Leistungsumfangs, wurde Blackboard ausgewählt. Auch wenn diese Plattform nicht für die Entwicklung von Ausbildungsgängen konzipiert wurde, ist sie für die Aktivitäten des Konsortiums offenbar doch von großem Nutzen. Elektronische Lernumgebungen werden eingerichtet, um alle Vorteile der modernen IKT zu nutzen: Sie funktionieren innerhalb einer Internet-Umgebung und verfügen über Einrichtungen für E-Mail, Diskussionsforen, Gruppenkommunikation und sogar Online-Unterricht. Damit handelt es sich um eine sehr effektive Umgebung für die Gestaltung von Lehrgängen im Rahmen einer internationalen Kooperation.

In der Tat hat die Nutzung dieser neuen Technologien die Kommunikation zwischen den Partnern vereinfacht. Der finanzielle Aufwand für Zusammenkünfte (Reise und Unterkunft) hielt sich somit in Grenzen. Dennoch ist dieser Technologieansatz vielleicht zu optimistisch gewesen. Eine reale Zusammenkunft hat durchaus Vorteile gegenüber virtuellen Treffen, und soziale Kontakte tragen zu einem besseren gegenseitigen Verständnis und Engagement bei.

Nach Abschluss des Projekts haben die Partner beschlossen, ihre Zusammenarbeit fortzusetzen. Die Vorzüge der gemeinsamen Arbeit beim Unterricht auf dem Gebiet der elektronischen Aufzeichnungen liegen für alle Beteiligten auf der Hand. Trotz der technologischen Beschränkungen scheint der Weg für europäische Ausbildungsgänge offen zu stehen, die von einem Konsortium von Lehreinrichtungen für eine heterogene Zielgruppe durchgeführt werden. Und warum sollte man nicht ein noch ambitionierteres (virtuelles) europäisches Zentrum für Hochschulstudien und Weiterbildung im Fach Archivwesen in Angriff nehmen?

Projet européen d'enseignement de la gestion des documents électroniques [E-TERM ([1])]

Peter Horsman

Le développement d'une formation européenne à la gestion des documents électroniques, destinée à des archivistes et à des gestionnaires de documents, a été largement évoqué à l'occasion de la réunion des experts européens sur les documents électroniques qui s'est tenue à La Haye, le 18 juin 1997, réunion qui faisait suite au Forum DLM. À l'issue de cette rencontre, la section «Éducation et formation» du CIA avait dressé un bilan très positif des deux formations présentées plus tôt dans la journée, en l'occurrence un cycle de cinq jours de conférences organisé par l'école de formation archivistique des Pays-Bas et un cours élaboré par l'école des archives de Marbourg. Étant donné que la formation de Marbourg visait principalement les archivistes, mais que la formation néerlandaise concernait les archivistes et les gestionnaires de documents, les participants avaient approuvé l'idée selon laquelle la formation néerlandaise pouvait servir de modèle au programme européen à élaborer.

À l'instigation de Hans Hofmann, archiviste à la Commission européenne et impulseur du Forum DLM, l'école d'archivistique lançait au printemps 1999 une invitation à des institutions homologues dans divers pays européens. Tous les participants avaient admis la nécessité de disposer d'une formation, y compris en ce qui concerne le matériel pédagogique adapté. Une demande de subventions au titre du programme Leonardo avait donc été déposée à Bruxelles et approuvée. Le projet s'était achevé en décembre 2001.

Les participants étaient les suivants: University College London, University of Northumbria de Newcastle, Ufficio centrale per i beni archivistici (Italie), Fachhochschule Potsdam (Allemagne), université de Tampere (Finlande) et université de Porto (Portugal). Le projet était placé sous la direction de l'école d'archivistique des Pays-Bas.

Le projet a débouché sur un programme d'enseignement et sur une série de matériels pédagogiques validés: textes (aspects théoriques), études de cas (applications), terminologie et exercices pratiques. Le programme néerlandais initial a servi de point de départ, mais a été modifié et étendu.

Le programme d'enseignement est conçu en cinq modules, qui confèrent toute la flexibilité requise selon les nécessités locales. Ces cinq modules peuvent éventuellement être précédés d'un module introductif: introduction aux technologies de l'information à l'intention des archivistes et des gestionnaires de documents (et des administrateurs), introduction à la théorie et à la pratique de l'archivage à l'intention des professionnels de l'informatique et des administrateurs et, enfin, introduction à la gestion, destinée aux archivistes, aux gestionnaires de documents et aux professionnels de l'informatique.

Chaque module aborde une thématique particulière:

— impact des TIC sur les organisations,

— systèmes d'archivage,

— aspects juridiques et organisationnels,

— problèmes de mise en œuvre,

— stratégies et politiques.

Des textes pertinents ont été recherchés et commentés et des exercices pratiques rédigés pour chacun de ces thèmes. Un «cas type» peut être utilisé pour préciser la théorie, de telle façon que les participants puissent adapter les enseignements appris à leur travail quotidien.

([1]) E-TERM: European Training Programme on Electronic Records Management.

Des études de cas de portée plus modeste peuvent être utilisées pour illustrer des sujets particuliers, tels que l'e-mail, le transfert de fichiers, la description, etc.

Ainsi qu'il avait déjà été observé, l'ambiguïté de la terminologie constitue l'une des difficultés particulières rencontrées dans la compréhension de l'abondante littérature sur les archives électroniques. Cette difficulté devient encore plus aiguë dans un environnement multilingue. Aussi le projet a-t-il débouché sur un résultat majeur et substantiel: la mise au point d'une terminologie multilingue, clarifiant les concepts d'un pays à l'autre.

Chaque volet du projet a été confié aux différents partenaires du consortium sous forme de lots: Londres et Newcastle se sont consacrées aux textes et aux études de cas, Potsdam et l'Italie à la terminologie, tandis que Tampere et Porto s'engageaient dans les tests, tout en participant également à d'autres travaux. En fait, toutes les institutions partie prenante ont testé le programme et les résultats des projets pilotes ont été fort utiles au stade final du projet. Enfin, Amsterdam s'est chargée de la gestion du projet et a utilisé le programme en tant qu'instrument.

La formation sera dispensée dans sa totalité dans un environnement d'apprentissage électronique. Les établissements de formation souhaitant fonctionner à distance peuvent l'utiliser telle quelle. Par contre, les établissements préférant la salle de cours traditionnelle peuvent utiliser le même cadre. Bien sûr, il est également possible de combiner les deux méthodes d'enseignement.

En fait, très peu de temps après le démarrage du projet, le consortium a décidé d'utiliser, à titre expérimental, un environnement d'e-apprentissage particulier pour le développement du projet. Pour diverses raisons, parmi lesquelles sa diffusion mondiale et son suivi après-vente, c'est Blackboard qui a été retenue. Bien qu'elle n'ait pas été créée pour développer des programmes de formation, cette plate-forme semble très utile aux activités du consortium. Les environnements d'e-apprentissage sont créés pour exploiter au maximum les avantages des TIC modernes: ils fonctionnent dans un environnement internet et possèdent des fonctionnalités pour l'e-mail, les forums, les communautés et même pour l'enseignement en ligne. Toutes ces possibilités en font un environnement très efficace pour concevoir des cours impliquant une coopération internationale.

L'utilisation des nouvelles technologies contribue en effet à la communication entre les partenaires. Le budget des rencontres (frais de voyage et d'hébergement) a donc été tout à fait modeste. Toutefois, cette vision technologique pourrait avoir péché par excès d'optimisme. Une rencontre réelle possède des avantages sur une rencontre virtuelle, et les contacts sociaux favorisent la compréhension mutuelle et la motivation.

Maintenant que le projet est terminé, les partenaires ont décidé de poursuivre leur coopération. Même avec les contraintes technologiques, la voie semble ouverte pour de véritables cours européens, organisés par un consortium d'établissements de formation à l'intention d'un public hétérogène. Et pourquoi ne pas envisager encore plus ambitieux: un centre européen (virtuel) d'enseignement supérieur de l'archivistique?

Understanding the information-seeking behaviour of archival researchers in a digital age: paths, processes and preferences

Wendy Duff

Digital access to archival materials holds great promise. Electronic records, the world wide web, and archival networks present numerous opportunities to cut across geographic and cultural boundaries and 'democratise' access to archival material. Finding aids linked to digitally born or digitised records provide a central gateway to holdings that are geographically scattered. These systems enable users to retrieve and use the holdings of archives in Spain, England, France, China, Australia, or even Africa without leaving their office, their homes or perhaps an Internet café. For example, the National Archival Network in Great Britain envisions researchers 'sitting at their terminal with an Internet connection, linking to a central gateway to the archival network, where they would fill in the sort of search template that you might use to access an individual repository's catalogue. It would offer the ability to mix searches by place, person, subject, date, and keyword. The network would then perform this search across all electronic catalogues connected to the network simultaneously, and return a single integrated response to the user in the form of a pick-list. The user would then follow the hypertext links from the pick-list to the full finding aids, and have the freedom to move around the finding aid as well as see the specific entry found by the search.' ([1]) This network envisions a simple, efficient, and powerful tool to access finding aids across Great Britain in a single search. Great vision but will users be able to exploit this fabulous resource fully?

Canadian archives share Great Britain's National Archival Network's vision and are building a similar system called the Canadian Archival Information Network (CAIN). At the same time, archival institutions across North America are developing web sites, providing e-mail reference services and extending their programmes to remote locations around the globe. The web seems to present untold possibilities for increasing the number of people who use our resources. Researchers who have never physically visited an archives are accessing records remotely; and consequently, the demographic and geographic diversity of archival users is growing. This change is well under way in Canada. For example, The Ontario Archives, my provincial archives had only 19 062 visitors to their building in 2000/2001, but over 6 700 000 hits on their website. In the same year, 349 682 people visited the National Archives of Canada facility in person while the web site had over 72 million hits. Historians and administrators still use archives, but today they are joined by genealogists, school students, journalists, media specialists and a myriad of other researchers.

These changes provide a great opportunity but unfortunately these new users often lack the skills needed to navigate complex archival descriptive systems, the knowledge required to evaluate the authenticity or reliability of records, or an understanding of the context essential for interpreting material. Furthermore, remote researchers have little or no access to an archivist to help them traverse the complex web of finding aids or locate relevant records. Traditionally archivists gain knowledge of their researchers' needs and behaviours when they interact with users and help them locate information. They know their researchers' topics, the records that are being used, and any problems that occur. However, a digital world limits their ability to gather this information informally because users interact with information systems, not archivists. Ironically, as the need to understand users increases, the means of obtaining the information informally decreases. This results in new demands and a greater urgency to develop formal methods to gain an awareness of the information-seeking behaviour of archival patrons.

Recently a number of researchers have begun to study users in an archival context including how they seek and use information. This paper discusses the value of user studies and reviews methods for gaining a greater understanding of the information needs of archival users and for evaluating the usability of archival systems. It provides an overview of the findings of a number of archival user studies and makes recommendations for designing user-friendly, intuitive systems for accessing archival material. It concludes with a call for the development of a network for research on archival users studies and generic user-evaluation instruments.

Wendy Duff

Wendy Duff is an assistant professor at the University of Toronto, Faculty of Information Studies. She received her PhD. from the University of Pittsburgh. Her primary research interests are user studies, archival description, and electronic records. She has written numerous articles on the access to archival material, information-seeking behaviour of archival users, archival description, and electronic records. She has served as a member of the ICA Adhoc Commission on Descriptive Standards, the Encoded Archival Description Working Group, and chair of the Canadian Committee on Descriptive Standards. She is presently a member of the Canadian-US Task Force on Archival Description and the Encoded Archival Context Working Group. She was co-investigator for a digitisation and access project: Usage analysis funded by the Andrew W. Mellon Foundation. She has recently received grants from the Social Sciences Humanities and Research Council of Canada and the Delmas Foundation to investigate the use of archival material by scholars, how scholars evaluate the authenticity of documents, and how archival users find information, and the use of archival description in the scholarly process.

([1]) Kingsley, N., 'Making a new showcase for British archives: the creation of a national archival network.' In *Electronic access: Archives in the new millennium*, London, PRO, 1998, p. 19.

Conducting user studies

Archivists' expertise is grounded upon knowledge of records and record creating activities. At the heart of archival theory is the record, not its secondary use nor the various types of researchers who visit archives seeking information. Archivists often express little interest in how or why genealogists, lawyers, media specialists or scholars access or use material. Some archivists assume that they understand the needs of their users because they deal with them on a day-to- day basis. Others assume they know what researchers require; and consequently they do not need to gather feedback or study information seeking behaviour in an archival context. This represents a sort of 'archivist knows best' attitude. Closely aligned to this way of thinking is a belief that archives acquire and describe records for the future, not the present. This belief suggests that the wishes of today's researchers are not worthy of study and the needs of future researchers impossible to determine. However, designing usable web-based systems that effectively serve remote researchers requires a thorough understanding of how and why people seek information in an archival context. As Paul Conway points out, 'good reference service, like good business, means discovering patrons' needs, developing the means to meet these needs, and following up to measure the impact of services.' ([2]) Remote services also require concerted efforts to discover patrons' needs, ways to meet these needs and mechanisms for evaluating their impact. In a similar vein, Tim Ericson wisely suggests, 'If, after we brilliantly and meticulously appraise, arrange, describe and conserve our records, nobody comes to use them, then we have wasted our time.' ([3]) Unfortunately, archivists have little empirical knowledge about their patrons' needs; and consequently, they may fail to serve archival researchers effectively.

Librarians and information scientists on the other hand have conducted user studies since the 1940s. The research in this area grew slowly at first but its development was greatly aided by the establishment of the Centre for Research on User Studies at Sheffield University in the 1970s. This Centre served as a catalyst for various studies, trained numerous researchers and refined both qualitative and quantitative methods for conducting user studies for approximately fifteen years ([4]). Throughout the 1980s other researchers expended much effort developing conceptual frameworks, improving techniques, and contributing to the basic knowledge in this field. By the 1990s information science researchers began to concentrate their interest and research on issues related to the use of the Internet and digital libraries. The findings of a multitude of projects are documented in articles on information seeking behaviour, information use and the information needs of numerous different user groups. Archivists can benefit from this research, by building on the current development of conceptual models and employing the robust techniques for studying the needs of archival users. But to do so, we must first know the degree to which the existing information seeking behaviour research is applicable to an archival setting.

Archival user studies

Since the 1980s the call for studying the use and users of archives has grown and during the last five to ten years a few researchers have begun to answer this call. Students enrolled in archival programmes, professors of archivistics and professional archivists have begun to evaluate archival services and systems from the users' point of view. These studies have employed both qualitative and quantitative methods including paper-based and web-based surveys, interviews, focus groups, and experiments, etc. Unfortunately this overview only includes research conducted in Canada and the United States.

In North America students have examined remote reference correspondence, e-mail reference and on-site reference questions to understand the terms patrons used in their requests and to observe the effect of providing remote access to the archival system and reference service via e-mail. A number of master's students have carried out a few small studies but combined their results to provide interesting insights. For example, Karen Collins' study of reference questions indicated that users most frequently supplied information about the name, place, date and event in their query ([5]) and Kristin Martin found that 'remote users will likely not provide the archives with enough information when submitting their requests'. ([6])

Professional archivists have also studied their patrons' information seeking activities and gathered feedback to assist with the design of new systems. For example, the staff of the Claude Pepper

([2]) Conway, P., 'Research in presidential libraries: a user survey,' *Midwestern archivist* 11, No 1, 1986, pp. 35–55.
([3]) Ericson, T. L., 'Preoccupied with our own gardens: outreach and archivists,' *Archivaria* 31 (1990–91), p. 117.
([4]) Robert, N., and Wilson, T. D., 'The development of user studies at Sheffield University 1963–88,' *Journal of Librarianship* 20 (October 1988), pp. 270–290.
([5]) Collins, K., 'Providing subject access to images: a study of user queries,' *American archivist* 61 (Spring 1998) pp. 36–55.
([6]) Martin, K. E., 'Analysis of remote reference correspondence at a large academic manuscripts collection,' *American archivist* 64 (Spring/Summer 2001), p. 40.

Collection of the Florida State University Libraries conducted a user study of their online 'Pepper online archival retrieval and information system' (Polaris). Though their study had only a few participants the staff found the feedback 'invaluable.' Respondents to the Polaris study indicated that they sometimes become lost or confused using the online finding aids and they wanted to 'know were they are in the collection at all times.' ([7]) Not only did these respondents require systems with both advanced and basic search capabilities, they also wanted the 'help pages to be expanded.' ([8]) Interestingly, most studies of online systems indicate that users rarely consult a system's help pages, but Polaris users wanted these pages because finding aids would be utilised 'without the benefits of in-person assistance.' I am curious to see if users of an archival system actually consult help pages when available because sometime what users say they want is not necessarily what they will use. Employing observations, interviews, and direct measurement techniques, Paul Conway conducted an extensive study of researchers at the National Archives of the United States in 1990/1991. Not only did the research shed light on how users carried out their work across the organisation, but it also pointed to a number of concrete recommendations for the design of a user interface, user training and ways to improve services to NARA's patrons ([9]).Unfortunately, as far as I know, these recommendations were never implemented.

User studies also provide a rich research area for university professors of archival science. In the last few years, Helen Tibbo, Anne Gilliand-Swetland, Elizabeth Yakel and myself have carried out a number of different studies to gain a greater understanding of the use of archives. For example, in a survey of American historians Helen Tibbo found that 93 % of her respondents reported that they had consulted repositories' web sites ([10]). Anne Gilliland-Swetland, Yasmin Kafal and William Landis' study of students in grades four and fives use of the Online Archives of California indicated that to be most effective primary material 'required direct integration with curriculum activities and a range of other kinds of materials.' ([11])

For the last five years my colleagues at the University of Toronto and I have also been carrying out a variety of studies in this area. Time does not permit in-depth review of this research, but I hope a few examples will demonstrate their value. In a survey of all historians who teach in Canadian degree granting institutions, similar to a study done by Helen Tibbo, Barbara Craig, Joan Cherry and I found that historians use both formal and informal methods for locating archival resources and that they rely heavily on finding aids. For example, 93 % of the respondents rated finding aids and archival resources as either very or somewhat important sources for becoming aware and locating needed information. Footnotes and archivists were rated almost as high with 89 % of the respondents rating footnotes and 83 % of respondents rating archivists as very or somewhat important sources for locating resources. In comparison, only 45 % of the respondents rated the world wide web as either a very or somewhat important source. Although research results suggest that historians continue to use traditional methods for locating archival records, the web is gaining in importance and was rated as having higher importance than indexing or abstracting services but lower than bibliographies, footnotes, etc.

Fifty percent of the respondents of the survey noted that the lack of finding aids was a barrier to their current research, and 37 % had had to use access tools that were not detailed enough. A number of respondents suggested that the web and electronic finding aids would assist them in their research and many noted that not only did remote access to these tools save both time and money, they also assisted with the planning of a trip to a distant archives, directing students or research assistants, clearing freedom of information requirements, and ordering inter-library loans or photocopies ([12]).

Catherine Johnson and I analysed e-mail reference questions received by archives and found these requests usually included proper names, dates, places, subject, form and occasionally events. Furthermore, 13 % of all the questions in our study indicated that the user probably needed some form of education to use the archival system. Putting indexing and finding aids on the web did not seem to decrease e-mail questions, causing instead an increase in the number of requests for photocopies ([13]).

Catherine Johnson and I have also conducted in-depth interviews with a small number of historians and genealogists (10 each) to find out how they locate and use archival material. These interviews are part of a larger study which also includes a number of smaller projects that employ a variety of methods, e.g. diaries, participant observation and experiments.

([7]) Altman, B., and Nemmers, J. R., 'The usability of online archival resources: the Polaris project finding aid,' *American archivist* 64 (November 2001), p. 126.
([8]) Ibid, p. 127.
([9]) Conway, P., *Partners in research: improving access to the nation's archive user studies at the National Archives and Records Administration* (Pittsburgh: Archives and Museum Informatics, 1994).
([10]) Tibbo, H. R., 'Primarily history: historians and the search for primary source materials' Joint Conference on Digital Libraries, Portland Oregon, July 2002.
([11]) Gilliland-Swetland, A. J., Kafal, Y. B., and Landis, W. E., 'Integrating primary sources into elementary school classroom,' *Archivaria* 48 (1999), p. 110.
([12]) Duff, W. M., Craig, B., Cherry, J. M., 'Historians' use of archival sources: promises and pitfalls of the digital age', in progress.
([13]) Duff, W. M., and Johnson, C. A., 'A virtual expression of need: an analysis of archival reference questions,' *American archivist* 64 (Spring/Summer 2001), p. 43–60.

The interviews with the historians indicate that interviewees not only used finding aids, but also valued them highly. The results of the survey of Canadian historians supported this finding. Finding aids were consulted to orient historians to new collections, to provide context and background information for their research areas, and to facilitate the identification of relevant documents. We also found these historians were very adept at using the provenance method of retrieval. They collected names throughout their research to identify relevant material within both the collections and the finding aids. This study found that historians' questions often start broad as they gather all available background information and build up contextual knowledge about their area. Without contextual knowledge, the ability of researchers to find relevant information is difficult, if not impossible. Historians at the beginning of their research, therefore, appear vague about their information needs. However, as they gain more understanding their questions change and they do not seem to have a problem reframing their research or refining their questions ([14]).

I am still analysing the data from the interviews with the genealogists but it appears that genealogists rely heavily on maps and many suggested that after name access, geographic access is most important. Genealogists seem to love surfing the web in search of their ancestors. In a study of Early Canadiana Online (ECO), a digital library collection that consists predominantly of fiction, books on women and aboriginals, we were amazed to discover that 29 % of the respondents of the survey were genealogists. At first glance, the material seems of little relevance to genealogical research, but not only did genealogists use this digital library, they seemed to love it ([15]).

These various studies use different methods and focus on a variety of types of archives and users. Each study, however, provides some insight into how historians, genealogists, students etc., carry out their research and suggests methods for improving services and designing systems.

Implications for using digital materials and web-based finding aids

Studies of remote and onsite reference questions indicate that researchers' requests include names of people and places, as well as dates, forms of materials, types of events and sometimes subjects. Archivists should design systems with search functions that enable researchers to search for material using the terms they know including names, places, dates, form, and subjects. Archival systems should also facilitate browsing of digital surrogates and finding aids. Historians browse for information, often skimming through vast amounts of information to identify relevant material and to gather contextual and background information. Direct access to relevant information through sophisticated searching may meet the needs of some researchers, for example genealogists, but this type of access will meet the needs of all researchers.

E-mail reference requests received by archives are increasing exponentially. In their search for information researchers write to archives, often asking very general questions that need to be clarified ([16]). Martin suggests that archives create automatic forms that force researchers to supply certain types of information when making requests, such as a name, date, and place of birth etc, when requesting a birth certificate. Such a system would certainly help archivists to fill a request, but archives should do more. Archives should not only provide forms but they should also develop expert systems that would help users carry out their research more independently. An expert system could query researchers to get information and based on the data supplied the system could lead users to relevant material. Digitising paper guides, inventories and records is not enough. Archives also need to build systems that guide users with vague information needs and help them articulate and refine their requests.

Archival systems will require not only detailed help pages as the Polaris study found or the expert systems as I suggest, but also a kind of virtual archivist. Archivists play a key role in directing people to relevant information and remote users will continue to need this direction and advice. Research shows that archivists are often called on to provide expert advice on the relevancy of certain records to particular kinds of research. In a digital world their expertise is needed as much, if not more, than in a paper-based environment.

Archival web sites should be designed for users with various levels of expertise and archival knowledge. The new users that the web attracts will not know how to use a traditional finding aid or exploit archival resources. However, web pages specifically designed for different groups can be

([14]) Duff, W. M., and Johnson, C. A., 'Accidentally found on purpose: information-seeking behaviour of historians', *Library quarterly, in press.*

([15]) Duff, W. M., and Cherry,J. M., 'Use of historical documents in a digital world: comparisons with original materials and microfiche,' *Information research* 6 (October 2000): http://InformationR.net/ir/6-1/paper86.html

([16]) Martin, K. E., *American archivist*, op cit.

customised to meet particular needs. For example, genealogists would have pages with genealogical information, maps that provided access to sources related to a geographic area, templates for asking e-mail reference questions or the type of expert system that I mentioned. Pages designed for schoolchildren could contain interactive sources or be integrated with classroom assignments to help students learn about the value of archival material. Instead of providing only one type of information for everyone archivists should tailor their services to fulfill the requirements of various researchers with different information seeking behaviours.

Archives should not only meet the immediate needs of researchers but should also strive to educate the public about archival concerns. Archival websites have an important role to play in educating citizens about issues of authenticity and reliability of digital records. As content on the web proliferates identifying reliable authoritative records will become more challenging and forgeries of alteration more difficult to detect. The archival mission should not only focus on instructing governments and corporations in the proper care of records, but should also educate the general public about these matters. The world wide web provides an excellent mechanism for facilitating advocacy and outreach activities and disseminating information about record related issues. If archivists join forces they will identify new and creative ways to exploit the technology, to educate their users, and to facilitate access to archival records.

Recommendation

A full understanding of the information seeking activities in an archival context will require collaboration between archival professionals and archival academics. There are a few concrete steps that I believe we should take to foster this partnership. First, we should establish a network for research on user studies in an archival context. This network could provide training, carry out research, consult with archives interested in conducting studies and develop research instruments. Most archivists are not trained as social scientists; therefore they require assistance and access to tools to conduct their research. The archival community could benefit greatly from the development of generic evaluation tools or questionnaires that would gather user feedback on new digital collections or archival information systems. Designing robust questionnaires is challenging but generic evaluation tools could help alleviate some of the burden. A tool could provide standard questions that archivists could employ when developing a questionnaire for their institution. A generic evaluation instrument would not only benefit archives that wish to evaluate their system, but would also enable archivists to compare the results of different studies.

In 1991, Margaret Hedstrom presented a framework for electronic records research that highlighted the need to assume a broad perspective that set studies within their social, historical and cultural context. She suggested that 'a research framework can provide the structure for a series of smaller, practical projects that build on each other's results, contribute to an understanding of broader issues, and yield cumulative results from what might otherwise be disparate efforts.' ([17]) Electronic records research carried out over the last 10 years has contributed to a greater awareness of broader issues and yielded cumulative results as shown by the papers at this conference. The establishment of a research network for, and a coordinated effort on studying archival users could also lead to a better understanding of broader issues and yield cumulative results in this new but extremely important area.

([17]) Hedstrom, M., 'Understanding electronic incunabula: a framework for research on electronic records,' *American archivist* 54 (Summer 1991), p. 339.

Comprensión del comportamiento de los investigadores en la investigación documental en la era digital: vías, procesos y preferencias

Wendy Duff

Buscar y explotar información en un fondo de archivo nunca es tarea fácil. Localizar y explotar documentos digitales a través de la red puede ser aún más complicado. Numerosos investigadores cuentan con los archiveros para que éstos les ayuden a consultar la información archivada y a encontrar documentos, ya que los sistemas suelen ser muy vastos y desalentadores. Los archiveros también proporcionan a numerosos investigadores la información de referencia necesaria para comprender e interpretar las fuentes de archivo. Los usuarios que acceden a documentación a distancia no suelen tener archiveros próximos, incluso si su tarea de consulta y explotación de esta documentación puede resultar muy difícil. Por ello, los archivos deben concebir sistemas que respondan a las necesidades documentales de los usuarios. Estos últimos necesitan sistemas de archivo que anticipen el tipo de información buscada de manera fácil e intuitiva. Ahora bien, la concepción de sistemas explotables implica una profunda comprensión de la manera en que los distintos tipos de usuarios realizan la investigación documental.

Tradicionalmente, los archiveros han mantenido una relación bilateral con los usuarios, adquiriendo así conocimientos sobre el tipo de información buscada por éstos y sobre la manera en que éstos explotan los recursos archivísticos. A medida que los archivos amplían sus servicios a distancia, desarrollando sitios web y alcanzando a usuarios de todo el mundo, es necesario que los archiveros tengan una visión profunda del comportamiento de los lectores en materia de investigación documental. Desgraciadamente, muy pocos archiveros han estudiado a los investigadores, y disponen de pocos conocimientos prácticos sobre las personas que consultan sus archivos, sobre la manera en que buscan información y sobre lo que hacen con ella.

El presente documento examina la importancia de las investigaciones sobre los lectores y analiza las investigaciones en curso en este ámbito. También examina los distintos métodos de análisis de las necesidades y de evaluación de la forma de explotar los sistemas de información. El documento menciona distintas técnicas que pueden emplear los archiveros para comprender mejor la manera en que sus lectores realizan la investigación documental. Destaca la importancia de comprender mejor sus necesidades y asesora sobre la manera de realizar investigaciones de este tipo. Por último, el documento alude a una serie de análisis relativos a las prácticas de consulta de los investigadores y a la consulta de archivos en el proceso escolar.

Las investigaciones que se mencionan en el presente documento utilizan métodos cualitativos y cuantitativos que permiten analizar la manera en que los usuarios del ámbito escolar y los genealogistas llevan a cabo sus investigaciones en un fondo de archivos y utilizan distintas herramientas, entre las que figuran los instrumentos de investigación en papel y digitales y la consulta al archivero. También se muestra cómo los lectores determinan la autenticidad y la fiabilidad de las fuentes de archivo. La información en la que se basa la exposición procede de una investigación realizada a más de 600 profesores de historia de universidades canadienses, y está completada por entrevistas mantenidas con estudiantes de historia, una decena de historiadores, diez estudiantes de doctorado de historia y finalmente, diez genealogistas. La exposición se basa también en datos de los diarios de consulta de estudiantes de doctorado, cuando consultaban archivos.

Los resultados de estas investigaciones informan sobre la manera en que estudiantes y aficionados realizan sus investigaciones utilizando las herramientas archivísticas. Por otra parte, la exposición se basa en estas conclusiones para sugerir la posible utilización de herramien-

tas digitales e instrumentos de investigación en línea para los estudiantes y los investigadores principiantes. El documento muestra cómo los recursos en línea pueden mejorar y modificar la consulta de archivos y examina también los problemas que pueden plantearse cuando los investigadores acuden a la red y a otros instrumentos de investigación en línea. Se evalúa el posible impacto de la utilización de estos recursos en la manera tradicional de realizar investigaciones documentales, así como los problemas que pueden surgir cuando los investigadores acuden a la red y a otros instrumentos de ayuda en línea para buscar información. Así pues, ¿cuál es la relación entre el recurso material y la información de contexto necesaria para interpretar los archivos? ¿Cómo evalúan los lectores la autenticidad y la fiabilidad de los documentos en un mundo digital? ¿Qué se perderá (o se ganará) cuando la «eficacia» de la investigación en línea sustituya a los medios tradicionales mediante una navegación «ineficaz» en un gran maremágnum de archivos? El documento concluye proponiendo recomendaciones para concebir sistemas de archivos intuitivos y de fácil utilización y preconiza el desarrollo de un instrumento genérico para el estudio de los investigadores, instrumento que podrían utilizar los archivos para analizar el impacto de sus instrumentos de investigación en línea y sus recursos digitales.

Das Informationssuchverhalten von Archivbenutzern im digitalen Zeitalter verstehen: Pfade, Prozesse und Präferenzen

Wendy Duff

Informationen in einem Archiv zu finden und zu benutzen, ist keine einfache Aufgabe. Noch komplizierter kann es sein, digitale Unterlagen im Internet aufzuspüren und zu verwenden. Viele Benutzer wenden sich daher an Archivare, die ihnen beim Durchforsten von Archivinformationssystemen und Auffinden von Unterlagen behilflich sein sollen, sind doch diese Systeme bisweilen äußerst undurchsichtig. Die Archivare versorgen auch viele Rechercheure mit Hintergrundinformationen, die sie benötigen, um Archivquellen zu verstehen und zu interpretieren. In der Regel können Benutzer, die aus der Ferne auf Material zugreifen, nicht die Hilfe eines Archivars in Anspruch nehmen, obwohl ihre Aufgabe, Material zu suchen und zu nutzen, überaus schwierig sein kann. Archive müssen also Systeme konzipieren, die dem Informationsbedarf von Benutzern gerecht werden. Forscher wiederum brauchen Archivsysteme, die die von ihnen benötigten Informationen benutzerfreundlich und intuitiv antizipieren. Um jedoch benutzbare Systeme konzipieren zu können, muss man das Informationssuchverhalten der verschiedensten Archivbenutzer genau kennen.

Bisher waren die Archivare direkt für die Benutzer tätig und haben auf diese Weise Erkenntnisse über die Arten von Informationen erworben, die ihre Benutzer suchen, und erfahren, wie Rechercheure Archivquellen nutzen. Da die Archive nunmehr ihre Dienste auch für weit entfernt befindliche Benutzer anbieten, Websites entwickeln und virtuell für Benutzer in der ganzen Welt zugänglich werden, müssen die Archivare das Informationssuchverhalten ihrer Benutzer genau kennen. Leider haben sich bisher nur wenige Archivare eingehend mit ihren Nutzern befasst und verfügen deshalb kaum über empirische Kenntnisse darüber, wer ihr Archivgut nutzt, wie Benutzer Informationen suchen und was sie damit tun.

In diesem Beitrag wird der Wert von Benutzerstudien diskutiert, und es werden die derzeit dazu laufenden Arbeiten beleuchtet. Es geht um die verschiedenen Methoden der Ermittlung von Benutzeranforderungen und die Bewertung der Benutzbarkeit von Informationssystemen. Es werden mehrere Verfahren erörtert, auf die Archivare zurückgrei-

fen können, um das Informationssuchverhalten ihrer Benutzer besser zu verstehen. Auf die Bedeutung einer besseren Kenntnis der Bedürfnisse von Archivbenutzern wird hingewiesen, und es werden Hinweise für die Durchführung von Benutzerstudien gegeben. Darauf folgt ein Bericht über eine Reihe von Studien, die das Informationssuchverhalten von Forschern in Archiven und die Nutzung von Archivquellen bei der wissenschaftlichen Arbeit untersuchen.

Bei den in diesem Beitrag angeführten Studien kamen sowohl qualitative als auch quantitative Methoden zum Einsatz, um zu ermitteln, wie Wissenschaftler und Genealogen ihre Nachforschungen in Archiven durchführen und wie sie im Verlauf der Informationssuche eine Palette von Archivtools (z. B. gedruckte und digitale Findmittel) sowie den Archivar in Anspruch nehmen. Ferner wurde erforscht, wie sie die Authentizität und Zuverlässigkeit von Archivquellen bestimmen. Zu den Daten gehören Antworten einer Umfrage unter mehr als 600 Geschichtsdozenten kanadischer Universitäten sowie ausführliche Befragungen von Wissenschaftlern, die sich mit der Geschichte des Buches beschäftigen, und zwar zehn Historiker, zehn Geschichtsdoktoranden und zehn Genealogen. Weiterhin enthält der Bericht Daten aus Recherchetagebüchern, die von Doktoranden bei ihren Recherchen in einem Archiv geführt wurden.

Die Erkenntnisse aus diesen Studien vermitteln Einblicke, wie Wissenschaftler und Laien ihre Forschungsarbeit anhand von Archivmaterial durchführen. Aus diesen Erkenntnissen werden zudem mögliche Konsequenzen der Verwendung von digitalem Material und Online-Findhilfsmitteln für den Wissenschaftler und den unerfahrenen Forscher abgeleitet. Es werden Möglichkeiten beschrieben, wie Online-Ressourcen die Archivrecherche verbessern und verändern können, und Probleme angeführt, die auftreten können, wenn Forscher bei der Suche nach Informationen das Internet und andere Online-Findhilfsmittel nutzen wollen. Zudem werden die möglichen Auswirkungen dieser Ressourcen auf die herkömmliche Form der Recherche betrachtet. Welche Verbindung besteht beispielsweise zwischen der physischen Ressource und der textabhängigen Information, die zur Interpretation des Archivguts notwendig sind? Wie bewerten Forscher die Authentizität und Zuverlässigkeit von Archivgut in einer digitalen Welt? Was verliert (oder gewinnt) man, wenn die „Effizienz“ der Onlinerecherche den traditionelleren, aber „ineffizienten“ Prozess des Durchsuchens großer Mengen von Archivmaterial ersetzt? Der Beitrag schließt mit Empfehlungen für die Konzipierung benutzerfreundlicher, intuitiver Archivsysteme und einer Aufforderung, ein allgemeines Instrument für die Nutzerbewertung zu entwickeln, das von Archiven eingesetzt werden kann, um die Auswirkungen ihrer Online-Findhilfsmittel und digitalen Ressourcen zu untersuchen.

Comprendre le comportement des chercheurs dans leur recherche documentaire à l’ère du numérique: cheminements, processus et préférences

Wendy Duff

Rechercher et exploiter des informations dans un fonds d’archives n’est jamais une chose aisée. Localiser et exploiter des documents numériques via le web peut se révéler encore plus compliqué. De nombreux chercheurs comptent sur les archivistes pour que ceux-ci les aident à consulter des informations archivées et à trouver les documents, car les systèmes sont souvent vastes et décourageants. Les archivistes fournissent également à de nombreux cher-

cheurs les informations de référence nécessaires pour comprendre et interpréter les sources archivistiques. Les utilisateurs qui accèdent à distance à de la documentation n'ont généralement pas d'archivistes à proximité, alors même que la tâche de consultation et d'exploitation de cette documentation peut se révéler extrêmement délicate. Aussi les archives doivent-elles concevoir des systèmes répondant aux besoins documentaires des utilisateurs. Ces derniers ont besoin de systèmes qui anticipent le type d'informations recherchées de façon conviviale et intuitive. Or, la conception de systèmes exploitables implique une compréhension approfondie de la manière dont divers types de lecteurs se comportent dans leur recherche documentaire.

Les archivistes ont de tout temps servi les lecteurs dans un rapport bilatéral, acquérant ainsi des connaissances sur le type d'informations recherchées par les lecteurs et sur la manière dont ces derniers exploitent les ressources archivistiques. Étant donné que les archives étendent désormais leurs services à distance, développant des sites web et touchant virtuellement des lecteurs du monde entier, il est nécessaire pour les archivistes d'avoir une vision approfondie du comportement des lecteurs en matière de recherche documentaire. Malheureusement, très peu d'archivistes ont réalisé des enquêtes sur les chercheurs et ne disposent donc que de peu de connaissances pratiques sur les personnes qui consultent leurs archives, sur la manière dont celles-ci recherchent des informations et sur ce qu'elles en font.

Le présent exposé évoque l'importance des enquêtes sur les lecteurs et analyse les recherches en cours dans ce domaine. Il passe en revue les diverses méthodes d'analyse des besoins et d'évaluation de l'exploitabilité des systèmes d'information. L'exposé évoque diverses techniques que les archivistes peuvent employer pour mieux comprendre la manière dont leurs lecteurs pratiquent leur recherche documentaire. Il souligne l'importance de mieux comprendre leurs besoins et propose des conseils sur la façon de conduire des enquêtes de ce type. Enfin, l'exposé rend compte d'une série d'analyses portant sur les habitudes de consultation des chercheurs et sur la consultation d'archives durant la scolarité.

Les enquêtes dont il est question dans l'exposé font appel à des méthodes qualitatives et quantitatives, qui permettent d'analyser la manière dont les utilisateurs scolaires et les généalogistes mènent leurs recherches dans un fonds d'archives et utilisent divers outils, parmi lesquels les instruments de recherche sur papier et numériques et la consultation de l'archiviste. L'exposé montre également comment les lecteurs déterminent l'authenticité et la fiabilité des sources archivistiques. Les informations sur lesquelles se fonde l'exposé proviennent d'une enquête réalisée auprès de plus de 600 professeurs d'histoire enseignant dans des universités canadiennes, et sont complétées par des entretiens poussés avec des étudiants en histoire, une dizaine d'historiens, dix étudiants en doctorat d'histoire et, enfin, dix généalogistes. L'exposé s'appuie également sur les données des journaux de consultation d'étudiants en doctorat, lorsqu'ils consultaient des archives.

Les résultats des enquêtes renseignent sur la manière dont des étudiants et des amateurs mènent leurs recherches en utilisant les outils archivistiques. Par ailleurs, l'exposé s'appuie sur ces conclusions pour suggérer l'utilisation éventuelle d'outils numériques et d'instruments de recherche en ligne pour les étudiants et les lecteurs novices. L'exposé montre comment des ressources en ligne peuvent améliorer et modifier la consultation d'archives et examine aussi les problèmes pouvant se poser lorsque les chercheurs se tournent vers le web et vers d'autres instruments de recherche en ligne. L'impact possible de l'utilisation de ces ressources sur la manière traditionnelle de conduire des recherches documentaires est évalué. Ainsi, quel est le rapport entre la ressource matérielle et l'information contextuelle nécessaire pour interpréter les archives? Comment les lecteurs évaluent-ils l'authenticité et la fiabilité des documents dans un monde numérique? Qu'a-t-on à perdre (ou à gagner) lorsqu'une recherche en ligne «efficace» remplace les moyens traditionnels par une navigation «inefficace» dans un gros gisement d'archives? L'exposé propose en conclusion des recommandations pour concevoir des systèmes d'archives intuitifs et conviviaux et préconise de développer un instrument générique pour l'étude des chercheurs, instrument que les archives pourraient utiliser pour analyser l'impact de leurs instruments de recherche en ligne et de leurs ressources numériques.

Michele Boccaccio

Michele Boccaccio was born in Siracusa, Italy, on 12 July 1961. He obtained a degree in computer science at the University of Pisa in 1988. He began his career as a software engineer, working on software development projects in Italy and Belgium (Brussels) for companies belonging to the Olivetti group. In 1989 he started working for Telettra and participated in the system testing of the first GSM system developed by a consortium composed by Telettra, Matra Communications, Orbitel and Ericsson in France (Paris). In 1992 he was nominated System Testing Manager for the GSM Siemens SBS system, hosted by Siemens Telecomunicazioni in Cassina dè Pecchi. In 1994 he joined Nokia Networks. He played many roles inside the Customer Services organisation; firstly as System Acceptance Manager for Omnitel/Pronto Italia project, later as Technical Support Manager for Italy and afterwards for South Europe. In 1998 he was nominated Head of Network Services, Europe. During 1999 and 2000 as Director of the unit, he was in charge of the establishment and management of a new Research and Development Centre in Catania (Italy), the first Nokia centre in the field of advanced services for mobile telephony. In 2001 he became the Director of Institutional and Regulatory Affairs for Nokia Italy. He currently leads the Comunicando start-up as CEO of the company.

Education, training and operations

e-learning presentation: Comunicando's view

Michele Boccaccio

Trends and market segments

Public administrations/the public sector are widely facing a deep renewal process of their own internal operational procedures, technological infrastructures as well of the professional skills and managerial competencies of middle and top executive layers.

The world of universities, schools and educational institutions offers many business opportunities within the scope of special projects for advanced training, like, for example, cooperations with compulsory schools, universities, institutes for disabled persons, etc. In few researches carried out among professionals in the training and education field both in the public and private sector, e-learning competences of teachers and trainers have been estimated insufficient or hardly in line with the real needs. It is evident that there is a need for teachers, professors and trainers and pedagogical professionals to get trained in the world of e-learning (skills, competences, pedagogical approach, instructional design, technology, delivery methodology, facilitation, etc.).

Companies increasingly reveal a pressing need to train and develop own resources and employees to enter the net economy, to be able to maintain their own-success position in the market or to re-position or re-engineer with right competences. Training and learning must be of quality, with high technological content, strategically aligned with market directions and easily deliverable time- and cost-wise. The three different segments are important ones to focus and pay attention to. e-learning is definitely an answer to the above identified needs in terms of rapidity in response, quality in results, efficiency in implementing the solutions, etc.

Lifelong learners are those people who are pursuing learning opportunities either related to their own personal hobbies and interests. They will turn to more formal learning experience when structure, well-shaped and sophisticated learning solutions will be offered via the net. There are also other individuals in the market, such as lawyers, physicians, dentists, etc., who need to continue to increase job-related skills and their education. Many of them are mainly responsible to take care of searching, buying and paying for their own development.

Education and training requirements

Drivers to e-learning and the learning gap

The business scenario for training and learning is very attractive and is very much supported by the digital economy. It represents one of the most promising market segments in the world of the net economy. The Internet has revolutionised commerce, communication and information exchanges. This puts the whole discussion into perspective as the customer base of potential users and customers is increasing at high speed. Several factors are facilitating this growth. The larger and growing base of installed computers at homes and workplaces, more powerful equipment and modems, network security, infrastructure, and bandwidth improvements too. Access to the Internet is more reliable, cheaper and the consumer is a new consumer, educated by usage of the new powerful media, feeling at ease conducting online activities and commerce, aware of his/her purchasing power. The knowledge economy puts individuals, corporations and organisations in front of a huge paradigm shift. Globalisation and immediacy are the fundamental traits of most companies' operations and human daily activities: the entire world is in real-time by reach. Information and content are of utmost importance, they get digitised, multimedia approaches support the distribution of knowledge and this very precious content has to be managed to become a real asset. The convergence of communication, computing and content is transforming our lives: new richer media, more information in compressed chunks transferred everywhere via telecommunications highways. The knowledge economy is an innovation-based economy.

Creativity, imagination, competence and new ways of doing things are key to support any kind of organisation competitive advantages as well as assets for individuals to exploit opportunities made possible by the technological growth. So, education, training and learning are becoming critical processes to ensure a successful and sustainable growth for communities, social life as well as economy. In the knowledge economy, learning is a continuous process, part of daily work, anytime, anywhere, just when needed, encouraged, flexible, accessible and easy. Jobs requiring information management capabilities will grow; people will be more and more dealing continuously with information, content and knowledge.

In Europe, by 2010, half of the professional posts' demand will be made by companies, which produce products and/or technological services or made by those companies that will make intensive use of technology in their normal business operations. The occupation in the world of the information society is becoming more and more difficult and very much depending on the right competencies, the rapid capability to adapt to needs and to acquire new knowledge, competence and skills. The challenge for corporations, organisations, professionals and individuals is to continuously keep, maintain and update own skills in any competence area. The real shift in the market place is that telecommunications and information technology is crossing any kind of industry segment, so the ICT-related skills are becoming essential elements to perform any of the job-related activities in any business. Lack in competence and skills in the field of ICT is demonstrated already by several studies carried out at national and international level in Europe, especially in comparison with the United States. The skill gap is already obvious when looking at the labour market demand of professionals with high qualification and skills in software, Internet, telecommunications and related services. This demand is felt unanswered by the current offering. The lack of competencies in the field of IT could catch up in Europe in 2002 to become meaningful numbers, up to incapability, to fill up approximately 19 % of the jobs needed. (*Source*: Federcomin, Anasin, IDC 2000).

Technology and standards development

Many are drivers for the digital competitiveness: improved infrastructure (consolidation and widespreading of technologies that allow data transmission, sounds, ADSL, satellite broadcasting, GPRS, UMTS, etc.), increased access possibilities (lower-entry barriers for online services, wider range of services development opportunities), tougher competition, new business models, digital television and Internet TV, faster evolution in the communication media market, more and more increasing synergies between entertainment and the consumer world, where training and learning develops towards the concept of edutainment.

A lot of other joint factors are then influencing the development of new ways to deliver training and offering learning opportunities: the explosion and wide usage of Internet, extranet and intranet, the wide diffusion of browsing application and technology, powerful multimedia personal computers, desktop and laptop, wide offering of the 2.5 generation and 3 G (third generation) of mobiles phones and personal digital assistants, powerful and flexible programming languages as well as continually developed standards in the field of documentation and content management as well as in e-learning field.

Technology supports the entire learning and training process, from content creation and development to delivery and finally to reporting and evaluation and results assessment. Technology nowadays enables learning offering in the most effective, flexible and scalable manner. Technology and standards enables the creation of content, which becomes easy to reuse and update and refreshed. Technology is scalable and is based on open standards to leverage on existing computing and ERP (enterprise resource planning) systems and technological infrastructure as well as support integration to different and third-partied modules and tools and can be empowered to support increase in users into the system.

Standards are acquiring more and more value in the overall international discussion about knowledge as well as content management. There is a need for more open standards in the backbone technology and supporting tools on the learning process, standards are now gaining a great meaning in the field of content creation or content digitising. In fact developing and owning content, which is standard-compliant, is the only way to ensure that content can be shared across units of the same organisation, can be integrated everywhere across delivery systems and can be easily reused. The issue nowadays with standards is that they are still under development, not yet

widely distributed and there is a limit in the understanding on how to use them. Furthermore, products need to comply with more than one standard to easily fit almost everywhere, so confusion is somehow hindering the speed of diffusion. Yet, big development steps have been taken in recent times and the need to consolidate and clarify standards usage in the industry is urged by involving content developers, creators and publishers and let them widely understand what to comply with and what standards enable them to do.

Comunicando's view on e-learning

The wider arena of e-learning

e-learning can be defined as the use of technology and the Internet to manage, design, deliver, select, coach, support and share all kinds of learning. Its meaning is generally encompassing a big variety of training applications and processes including computer-based learning and CD-ROM, web-based learning, virtual classrooms, digital collaboration, satellite broadcasting, interactive TV, etc.

As mentioned already, the Internet is the real core issue; in fact, it has transformed the manner in which education occurs and creates new ways of learning. e-learning is a sort of result from the combination of the Internet and education and it should enable a more comprehensive approach to strengthen competence and skills development by providing resources and learning tools/applications to support individual different learning styles.

Learning goes mobile and wireless

Learning not only gets electronically empowered and enabled for wider reach, it also goes mobile and wireless — it is called m-learning. It develops from the combination of mobile devices and e-learning. This is the ultimate social integration of education/learning and technology. cell phones, PDAs, mobile computing and third-generation devices — these are equipment and software which enables m-learning. Today's advanced mobile computing with integrated phone capability is the best tool for e-learning and m-learning. It provides a full operating system, more memory, and much greater processing power. Some of the best applications for mobile computers are lecture notes, homework, interactive student chatting, mentoring, and remote access to live computer labs. Mobile devices become even more powerful when combined with wearable components that allow people to perform hand-free complex or manual tasks. The number of people worldwide using wireless devices is increasing as well as the number of people using wireless devices to access corporate intranets. The number of employees who spend the majority of their time outside the office is growing and a big part of Internet browsing will be carried out on wireless platforms and the mobile devices used will be more than landline PCs. Remote access possibility becomes an advantage and people can access knowledge bases from almost anywhere. There is an increase in productivity for employees as the workers can engage in e-learning at any time, while commuting, travelling, or waiting in line, or any other non-conventional available time for learning, etc. m-learning of course drives attention both from technical perspective as well as from content definition, development and from the learning experience that learners have when they are 'on the move'. So issues about instructional design and content creation and packaging have to be taken into account, such as rather small modular learning chunks than long storyboards, profiling and interaction much more personalised, etc. On the other side, from a technical view point the promise is to make sure that content can be seamlessly accessible from any mobile device and terminal, so that learning nuggets are properly recognised as they are standards compliant and can be managed and displayed on PDA screens, can be delivered by using graphics, videos and voice. The true challenge will be to combine three main factors: the learner, the content we need to deliver and the platform as backbone.

Delivery and content creation

Thanks to technology, highly powerful systems to manage and deliver training and learning are available, following the process up from early beginning to its final phase: they are the learning-management systems. The LMS is a system for both learners and instructors that assist in managing the entire learning process. An LMS takes over time-consuming preparatory and administrative tasks such as publishing content, generating assignments, marking tests, tracking learner progress, producing reports and even tracking attendance. An LMS also allows instructors to tailor programmes. An LMS can be used to incorporate any type of educational material into a mul-

timedia resource library that can be referenced by learners. Users interact with the system though their own customised portal, either to plan, organise and develop and manage learning solutions (learning supply view) or to get learning and plan their own development paths (learning view).

In a few words, the requirements for a good LMS are related to the following macro functional features: courses/programmes/learning-catalogue management, resources management, competency profiling and assessment, enrolment, reporting/tracking and data import/export features, workflow management, commence features, to system security and easy integration to third-parties technologies. From the technical viewpoint good characteristics are related to the type of architecture supporting the systems (n-tier architecture, single- or multi-server installation, etc.), infrastructure (server specs and operating system, databases, etc.), deployment (scalability, links to commercial ERP systems, etc.) customisation and profiling, hosting/housing versatility. Learning management systems support management of any type of learning delivery and channel, including the so-called traditional way, classroom training. Classroom training can again be easily planned, tracked and managed by the LMS, so all information relevant to training and learning are gathered and tracked in a consistent and digitised way as well as stored in common data bases. Managing content is already a fundamental issue, but it will acquire a greater meaning as a key element in managing the knowledge both in complex and simpler organisations.

Products available on the market to support electronic content design, development and creation are called authoring tools. They are multimedia creation tools and are used to develop content for all the content delivery systems to create media that can either stand alone or be added as a module into a management system. Standards are now gaining a great meaning in content creation/digitising. Developing/owning standard-compliant content is the only way to ensure that it can be shared across organisations and units of the same organisation, be integrated everywhere across delivery systems and be easily reused. Learning objects are the modular plug-and-play building blocks of e-learning content, and can include such media types as text, graphics, audio, video, animation, games, tests, and simulations. Content has to capture, stimulate the learner attention and concentration as well as to fit into the learner self-paced learning process. Content should stimulate thought, sight, sound and touch in many innovative ways, with videos, flash animations, etc. It must challenge the student attention and require his/her actions and contribution, with questions, pop-up windows, etc. It must be easily translated from different languages and moved across cultures. It must be easy to stop and start again, with bookmarks, control panel, etc. It must be movable everywhere. Content can be either purchased by external vendors (off the shelf courses), which are specialised in producing a course catalogue as their own product for the market or content can be developed based on own content, own knowledge (custom content).

Learning processes integration

As we described, tools and systems available to support e-learning have a functional scope, e.g. the perspective of the process they address and support. In fact, it is very important to understand how the tools and systems support the integration of the different learning processes. It is also relevant to keep the overall description of available systems and tools very simple, as in the market we are now starting to have quite a lot of commercial products and tools.

From strategy to implementation

It is important that the organisation has a learning and training vision and strategy that steers the direction towards targeted results. We need to be clear about the organisation business objectives and how the training strategy can support and enable them. It is of the utmost importance to understand what is the current training strategy (stage of execution, degree of awareness on the organisation and in the management layers, etc.). In a good strategic planning exercise, identifying the key stakeholders to support the strategy is very important.

Finally, mapping of the next steps to reach the strategic outcomes completes the planning phases. Once the strategy is ready and communicated and the business case is very well known in the organisation as well as perceived as supporting the overall business objectives, the time has come to move to implementation plans. The first actions are related to the project team creation and to the communication plan preparation. The second main step is to analyse the training-management process and set all requirements to be matched by the solutions and systems to be purchased. Vendors' selection, technology assessment and bid to the vendors come next. One of the

next suggested steps is related to the project pilot definition (geographical area, department, etc.) where to run the first trial before the roll out of the system to the entire organisation. If you also need to develop and/or buy content, it is important to carry out a need analysis to define what content goes 'e' and for which target group of employees. In case the content will be developed externally, vendors' selection, demo assessment and technology check if the packaging is standard-compliant and fits into the other infrastructure is vital. Content migration to the new systems and database will come as one of the immediate steps. Other subsequent activities are linked to the integration of the new e-learning project with other existing learning supply and to a proper marketing of the projects according to the communication plan already defined. Blueprinting and good documentation of each phase is very beneficial to manage the project and to refer back when the system will go live.

Change management is definitely one of the most challenging phases of an e-learning project as the latter is impacting on processes, on people and on learning culture habit. Communication in general is fundamental: you need to explain how learning changes, how daily life gets enhanced and improved by the support of the technology, to show how the system works, to do e-learning together, to support the organisation to create discipline and to facilitate the new way of learning (time for learning, respect for learning, rewarding for learning).

Best practices

Business to business

The case describes the B2B model which Comunicando is pursuing as part of its strategy. It is a prospect case still under non-disclosure, a university which, for a few years offers post-graduate courses to newly graduate students as well as to professionals who wish to further deepen their specific competence in the given subject matters. There are two main aims of the organisation. The first is to make learning content available on the Internet by utilising e-learning technology and methodology, in order to enrich the service offered to students who can already participate to classroom courses as to improve flexibility in content accessibility, further deepening of topics, collaborative opportunities. The second aim is to reach those students who, for geographical distance, full-time employment, other issues which may partially or totally hindering participation. Its overall target is to both increase richness and reach of its learning supply for the Italian market. The idea is to reproduce the campus environment by making all the main functionalities to manage and deliver the programmes available online and truly enabled by e-technology. The key issue is focusing in enabling all the players in the campus — directors of the programmes, professors and mentors, secretarial office and students — to access specified functionalities in the systems to perform their roles to the final goal of making programmes successfully delivering content and competence improvement. The business model is made up of two main components: developing, integrating and implementing the learning solutions management platform as well as creating and developing e-learning content modules. The management platform is a system which enables learning processes management, automation, content creation, virtual collaborative learning management, and creation of learning communities, testing, tracking and reporting.

Many of the management processes get streamlined and are made automated and therefore more efficient by use of technology.

The learning management platform is hosted by Comunicando.

This lets the customer be focused on their core competences, which are represented by know-how ownership and value in content design and creation, assessment and verification, mentoring and facilitation in learning.

The customer does not get the burden on big technological investments as well as does not need to worry about infrastructure management, maintenance, security, etc.

Two are the main access channels to the system and to functionalities: portal for the university (administration and management functions) and portal for the students (learning related functionalities).

The content creation approach is also very well applicable as it mainly aims to revive existing knowledge in the organisation and its transformation to a digitised content for further reusability and dissemination so as not to lose precious intellectual richness. A particular value may be offered by this approach to public administrations, which, by historical and organisational reasons, normally own a great deal of internal information and content distributed in the different organisational layers, which in many cases is not fully used and shared across the organisation.

Business to consumer

The case describes the B2C model. Comunicando also targets the consumer market by making e-learning available to individuals independently on the access terminal at their disposal: TV set, mobile devices, PC screens, third-generation technology, etc. The main backbone of the model is the learning solutions platform and all the functionality description provided earlier on. Main players are individuals who access the content seamlessly from home, on the move, etc; and content creators who offer their content to a bigger consumer audience. Comunicando hosts, administers and dispatches the contents to consumers.

Comunicando introduces the concept of edutainment and gaming approach applied to learning. Easy access via portals to facilitate navigation, collaborative environments, chat and discussion forums are enriching the proposed supply.

In this model, the technological backbone plays a key role as well as applications enabling access to any kind of terminal and device is important.

Many different types of content may be channelled and offered to the vast consumer target groups, clustered by interest areas, topics and themes and community (children, elders, housewives, musicians, lawyers, environment associations, etc.).

Thematic offering available online and accessible anytime, anywhere is fundamental as it may answer to many of the consumer needs. From a professional skills development viewpoint, it represents offering to those who wish to deepen competencies for career and vocational training. It also represents a huge offering opportunity to those who wish to pursue personal interests and hobbies development. An ethical improvement for the community development is given by the opportunity to offer easy access to content to the disabled people and to those who may have difficulties in participating in a 'traditional' knowledge-sharing session. Of course, learning for disabled people, especially with heavy problems, would deserve a longer and more dedicated context for discussion, especially from an instructional viewpoint. The benefits that e-learning brings to this important area of education are nevertheless clear.

Educación, formación y procesos

Michele Boccaccio

En este Libro Blanco analizamos las necesidades de educación y formación en la sociedad digital y los factores que dirigen el mercado, que empujan y mantienen los requisitos del aprendizaje electrónico en el contexto de rápido cambio de la economía digital. Tratamos brevemente las lagunas del aprendizaje, así como la demanda y la oferta actuales, a fin de subrayar la necesidad de proponer buenas soluciones de formación que permitan afrontar los retos de la economía y la adquisición de competencias en los próximos años. Examinamos las tendencias actuales del desarrollo tecnológico y las normas en materia de educación y formación.

A continuación partimos del análisis del concepto de formación por ordenador para explorar el campo más amplio del aprendizaje electrónico en el contexto socioeconómico, tal

como se perfila en lo sucesivo, y para examinar la manera en que avanza el aprendizaje móvil. También examinamos las herramientas y las tecnologías que tienen a su disposición las empresas, las organizaciones y los estudiantes para gestionar su formación en la economía digital. También se abordan y explican las principales ventajas que derivan del aprendizaje electrónico. Por último, examinamos los usuarios potenciales, los protagonistas y los segmentos de mercado (desde el punto de vista de las empresas), donde se observa una evolución significativa del desarrollo del aprendizaje electrónico (e-aprendizaje).

Asimismo, se analizan las buenas prácticas para proporcionar casos concretos para la evaluación comparativa y el posterior debate.

Por último, se celebrará un debate sobre la importancia de desarrollar una visión y una estrategia de aprendizaje y formación capaces de alcanzar los resultados previstos. Conviene clarificar los objetivos de actividad de la organización y la manera en que la estrategia de formación puede contribuir a su realización. Una vez que la estrategia está lista y difundida, que el modelo de actividad es perfectamente conocido en la organización y que se percibe como que puede apoyar a los objetivos generales de la empresa, ha llegado el momento de pasar a los planes de aplicación. Por último, mencionamos brevemente el problema de la gestión del cambio, que representa definitivamente una de las fases más delicadas de un proyecto de aprendizaje electrónico, dado que este cambio influye en los procesos, los individuos y la cultura en materia de formación.

Ausbildung, Weiterbildung und Abläufe

Michele Boccaccio

In dem Beitrag analysieren wir die Aus- und Weiterbildungsanforderungen in der digitalen Welt sowie die Markttriebkräfte, die die E-Learning-Anforderungen in dem sich rasant verändernden Umfeld der digitalen Wirtschaft vorantreiben und stützen. Wir gehen kurz auf das Qualifikationsdefizit und den aktuellen Stand bei Angebot und Nachfrage ein, um zu verdeutlichen, wie wichtig gute Lernlösungen sind, die dabei helfen, die Herausforderungen des wirtschaftlichen und wettbewerbsbasierten Szenariums in den kommenden Jahren zu bewältigen. Wir geben einen Überblick über die aktuellen Entwicklungstrends auf technologischem Gebiet und bei Branchenstandards im Segment Ausbildung und Lernen.

Im Anschluss an eine Analyse des Konzepts des computergestützten Unterrichts (Computer Based Training) befassen wir uns mit dem E-Learning im weiteren Sinne, dem sich aktuell herausbildenden sozioökonomischen Umfeld und dem derzeitigen Stand beim mobilen Lernen. Ein weiterer Schwerpunkt werden auch die Hilfsmittel und Technologien sowie entsprechende Organisationen sein, die den Lernenden für das Lernen in der digitalen Wirtschaft zur Verfügung stehen. Außerdem werden die wichtigsten Vorteile des E-Learning genannt und erklärt. Schließlich vermitteln wir einen Überblick über potentielle Nutzer, Akteure bzw. Marktsegmente (aus unternehmerischer Sicht) in Verbindung mit maßgeblichen Entwicklungstendenzen des E-Learning.

Des Weiteren werden nachahmenswerte Praktiken beleuchtet, um konkrete Fälle für das Benchmarking und die weitere Diskussion anzuführen.

Bei der abschließenden Analyse geht es um die Bedeutung einer Vision und Strategie für das Lernen und die Ausbildung, um die Richtung der anvisierten Resultate einzuschlagen. Wir müssen klare Vorstellungen von den Geschäftszielen einer Organisation haben und wie sie durch eine entsprechende Weiterbildungsstrategie gefördert und erreicht werden können. Sobald die Strategie erstellt und weitergeleitet worden ist und sobald die Geschäftsbewertung der Organisation genau bekannt ist und als den allgemeinen Unternehmenszielen dienlich wahrgenommen wird, ist der Zeitpunkt für die Umsetzung der Durchführungspläne gekommen; dieses Thema wird kurz erörtert. Abschließend gehen wir kurz auf die Thematik der Bewältigung von Veränderungen ein – wegen der Auswirkungen auf Abläufe, Menschen und Lernkultur fraglos eine der anspruchsvollsten Phasen eines E-Learning-Projekts.

Éducation, formation et processus

Michele Boccaccio

Nous analysons dans ce livre blanc les besoins de l'éducation et de la formation dans la société numérique et les facteurs de dynamisation du marché qui poussent et entretiennent les besoins en matière d'apprentissage électronique dans le contexte en mutation rapide de l'économie numérique. Nous discutons brièvement des lacunes de l'apprentissage et de la demande et de l'offre actuelles, de façon à mettre en lumière la nécessité de proposer de bonnes solutions de formation permettant de relever les défis de l'économie et de l'acquisition des compétences des toutes prochaines années. Nous passons en revue les tendances actuelles du développement technologique et des normes de l'éducation et de la formation.

Nous poursuivons ensuite à partir d'une analyse du concept de formation par ordinateur pour explorer le champ plus vaste de l'apprentissage électronique dans le contexte socio-économique, tel qu'il se dessine dès à présent, et pour examiner la manière dont l'apprentissage mobile décolle actuellement. Nous nous penchons en outre sur les outils et les technologies que les entreprises, les organisations et les apprenants ont à leur disposition pour gérer leur formation dans l'économie numérique. Les principaux avantages découlant de l'apprentissage électronique sont également discutés et explicités. Enfin, nous passons en revue les utilisateurs potentiels, les acteurs et segments de marché (du point de vue des entreprises), où l'on observe une évolution significative du développement de l'e-apprentissage.

Les bonnes pratiques sont également analysées pour fournir des cas concrets à l'étalonnage des performances et à la discussion ultérieure.

Enfin, une discussion sera menée sur l'importance de développer une vision et une stratégie d'apprentissage et de formation capables de mener vers les résultats visés. Il convient de clarifier les objectifs d'activité de l'organisation et la manière dont la stratégie de formation peut contribuer à les réaliser. Une fois que la stratégie est prête et diffusée, que le modèle d'activité est parfaitement connu dans l'organisation et qu'il est perçu comme pouvant appuyer les objectifs généraux de l'entreprise, le moment est alors venu de procéder aux plans de mise en œuvre. En conclusion, nous évoquons brièvement le problème de la gestion du changement, qui représente définitivement l'une des phases les plus délicates d'un projet d'e-apprentissage, ce changement influant sur les processus, les individus et la culture en matière de formation.

An intelligent tutoring e-learning module for training archivists

Paltin Sturdza, María Luisa Conde, Luis López Corral

Introduction

e-learning is a subset of technology-based training and encompasses all learning activities conducted on the Internet. We define e-learning as the use of a network for delivery, interaction, or facilitation of educational content. The educational content can be a two-minute tutorial, a college course, a corporate or certificate training, high school classes, and even online continuing education classes like those that you might take at night school or your community college.

Numerous studies have shown that employers can accelerate development, deployment, and the implementation time of business applications if they provide their workforce with high quality training. Moreover, companies can reduce administrative overhead and technical-support costs by implementing effective training programmes. e-learning is the fastest growing segment of the corporate training market as well as of the IT industry in general. It is only natural that the training of archivists will follow this trend.

Limitations of existing commercial e-learning products

e-learning packages have elaborate mechanisms of managing courseware, course distribution, and grading. They excel in giving the instructional designer tools to produce visually appealing and interactive screens, but behind the screens there is a shallow representation of content and pedagogy and therefore lack 'intelligence'. The knowledge they impart is not represented and stored in formal content models such as EPCs, semantic nets, production rules, frames or logical predicates. The knowledge resides in canned text, videos, drawings and does not constitute a true knowledge base in the sense of AI, cognitive psychology, or expert systems. In order to provide e-learning commercial systems with 'intelligence', companies are turning now to the technology of ITS (intelligent tutoring system).

Intelligent tutoring systems (ITS)

These systems emerged from artificial intelligence (AI). ITS attempts to emulate 'human-like' tutoring capabilities. The term 'intelligent' refers to a system with the ability to know what to teach, when to teach and how to teach. Such a tutoring system can be effective because it can respond to the specific needs of student, guide slow learners, challenge rapid learners and monitor the progress of each student as well as establishing a training plan. ITS have the ability to understand, learn, and solve problems just like the human counterpart.

The dream of the 'intelligent book' and 'intelligent communication of knowledge' imagined by Venger in 1987 is now closer to reality:

'Imagine active books that can interact with the reader to communicate knowledge at the appropriate level, selectively highlighting the interconnectedness and ramifications of items, recalling relevant information, probing understanding, explaining difficult areas in more depth, skipping over seemingly known material ... intelligent knowledge-communication systems are indeed an attractive dream.'

ITM: convergence of e-learning and ITS

ITS are difficult and expensive to build and lack the visual appeal and interactive screens of the commercial e-learning tools. ITS authoring tools need to be made robust enough to be used in production and commercial environments ITS authoring tools are still in the prototype stages which have had success in certain cases. A few prototypes have made it into the market.

Paltin Sturdza

Paltin Sturdza has graduated from the Florida State University in the US. He has 20 years of IT experience in both academia and the industrial world. Dr Sturdza has taught Computer Science at the Mountain Empire College, Pennsylvania State University (United States) and at the University of Alcala (Spain). He has collaborated in various research projects with the Polytechnic Institute and the Unisys Centre for Artificial Intelligence in Madrid (Spain). He has published and presented papers on expert systems development at various international conferences (United States, France, Germany, Spain, Italy, Brazil, Mexico) and large corporations such as IBM, MCI-Worldcom, BASF, Ciba-Geigy. He co-authored an expert system shell with Progilog (France) and the Spanish branch of Software AG (Germany). He has worked as an IT consultant for the IBM Global Service Division, BASF, and MCI-Worldcom.

María Luisa Conde

Director of the National Archives of the Spanish Adminstration. Ms Conde authored various manuals for administration archivists and will provide the subject matter expertise for the ITS.

Luis López Corral

Luis López Corral is currently the Head of the Computer Science department at the University of Alcala. With a Doctorate in Mathematics he is a full professor in computer science. He was formerly Director of the IT Division of the Community of Madrid and Regional Administration. His current research interests are in the use of mathematical modelling in the development of ITS.

We are proposing ITM (intelligent tutoring modules) as a novel solution for building the new generation of e-learning systems. Until now ITS and e-learning have evolved in parallel and separately. Our approach will allow these technologies to converge by combining the robustness of the current e-learning systems with the sophistication of the ITS systems.

The ITM project

The ITM will be implemented on the LUVIT e-learning commercial product as a platform. The ITM is a knowledge-based tutor, i.e. a computer-based educational system which behaves similarly to a human tutor. ITM uses artificial intelligence techniques that can determine what to teach, how to teach and learn certain teaching relevant information about the student being taught. ITM will represent the domain knowledge (domain knowledge model), the instructor's or teacher's knowledge (pedagogy model) and the student's knowledge level (student model).

The first novelty of our proposed system is that it combines the ITS (intelligent tutoring systems) technology with an e-learning commercial system (LUVIT) by the development of an ITM (intelligent tutoring modules). The ITM is more intelligent than other conventional e-learning systems that are lacking flexibility and learner-centred orientation. ITM adds a dynamic and adaptive dimension to self-paced instruction. For the user (student) it offers the advantage of highly individualised instruction. For the developer it offers another novelty in the form of the unified knowledge modelling. Through an advanced process model called EPC (event driven processing chains), and the implementation of rules using a commercial database, this model allows top-down development of the ITS software. It also provides a single method for writing specifications for both domain and pedagogical knowledge. It allows the developer to modify easily both the domain and the pedagogical knowledge. Our ITM solution provides (1) an ITM methodology for developing (authoring) intelligent tutors modules as a service, (2) the ITM executable application (the course content to teach) as a result of applying the methodology (3) the LUVIT system as an e-learning content management system.

Modelling archiving knowledge using the EPC (event-driven processing chains)

Manuals and books represent archiving knowledge as texts the student needs to read and assimilate in order to build a mental model which will allow him to solve concrete problems such as transferring documents from one archive to another. He must understand the textbook explanatory discourse and transform it into actions. As Confucius has stated, the ultimate criteria for understanding is the capacity to act correctly 'What I hear, I forget; What I see, I remember; What I do, I understand.' Confucius, 451 BC.

Components of the EPC model

Connectors	Conditions determine the execution of an action, they represent a state of affairs, attributes of an entity, the fact that an action has occurred, the result of an action which becomes the condition for the next action. Example: 'Documents are ordered'
Action	Actions are concrete happenings which produce some results and are determined or triggered by conditions. The are executable steps in some more complex process. Example: 'Eliminate paper clips from the documents'
Process	A process is composed of several actions. It is an abstraction and can not be executed per se. Example: 'Prepare documents' can be decomposed into several sub-processes ('order documents', 'eliminate paper clips')

Connectors

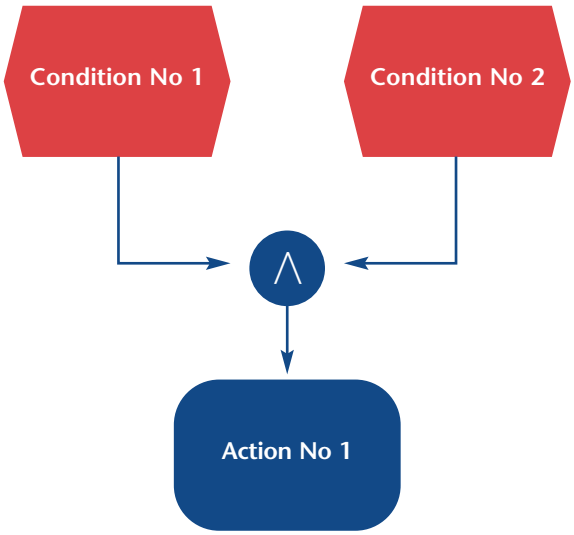
Connectors show the logical connections (decision paths) between the EPC components.

And

Non-exclusive or

Exclusive or (either ... or)

EPC frames



An EPC frame is composed of conditions, actions, processes and connectors. From the EPC frame executable rules will be derived: 'If Condition No 1 and Condition No 2 are true then perform Action No 1.'

An EPC frame can be made of at least one condition and one action or several conditions and several actions.

An interconnected set of EPC frames constitute an event-driven processing chain model.

The EPC model is both understandable by humans and executable by the computer. Its conversion into rules will ensure the interactive consultation with the learner. For implementing the EPC model we have chosen the ARIS easy design tool by IDS Professor Scheer (www.ids-scheer.com).

The linguistic knowledge engineering methodology for building ITMs

The purpose of linguistic knowledge engineering (LKE) is to build the EPC model which will generate the interactive knowledge bases to be used for teaching archiving knowledge. The interactive knowledge base will comprise a domain rule base (what to teach) and a pedagogical rule base (how to teach).

In a commercial e-learning system, the student has to read and assimilate a textbook. Then, he has to apply this knowledge to his specific environment. We propose to convert the archiving textbook into an 'interactive knowledge base', or 'interactive book' the user can consult and engage in a dialogue.

The LKE method provides the concepts for transforming the knowledge expressed in natural language of a text into an EPC. The end product will be an expert system which will be part of the intelligent tutoring module (ITM). This transformation requires building an EPC model which will act as bridge between the explanatory discourse of the text and the executable rules of the ITM expert system which will guide the student through a 'learning by doing' session.

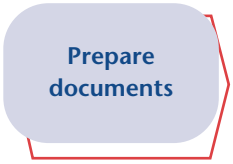
Without studying and learning the textbook and solely by using linguistic clues the knowledge engineer will identify conditions, actions, processes and connectors and construct a candidate EPC which will be submitted for approval and corrections to the subject-matter expert.

The following is a fragment from the chapter entitled 'Preparation of documents for transfer' from the 'Manual for handling administrative archives' (Spanish State Archives). Each sentence is numbered for reference in our analysis.

'(1) In the first place we need to review all the dossiers to be transferred in order to verify if none is missing. (2) In addition, we will verify that no document is missing from the dossier and all documents are in order … (3) After this operation each dossier will be ordered within its corresponding box. (4) You need to take off the suspended files or any other type of file, the classifiers, etc. (5) You also need to take off the staples, paper clips, rubber bands and any other extraneous body which could degrade the paper …(6) Finally, you need to number all the boxes to be transferred and this will allow the identification of these boxes in respect to the description of contents which will appear in the delivery confirmation (*Relacion de entrega)*.

LKE example analysis

We will analyse two sentences from the text and build a small EPC from it.



The title contains an abstract (decomposable) process 'prepare the documents' which will give the name to the 'process path' node. We have left out 'for transfer' from the model because these words do not indicate an action to take for the moment. They have only an explanatory value by showing the learner why he has to take the current action. The action 'transfer' will appear later in the model. We can see here a good example of the textbook mixing the explanatory with the executable discourse. This mixture of imperative actions with explanatory statements will be eliminated in the model where actions said by the ITM are actions done by the learner. While executing the actions suggested by the ITM system, the learner can pause and ask for explanations (causal or teleological). Thus the ITM separates clearly instructions to act from explanations and generates all possible explanations from the executable model on demand from the learner. This is clearly an advantage for the learner who, using textbooks, is limited only to those explanations the author deemed necessary to provide.

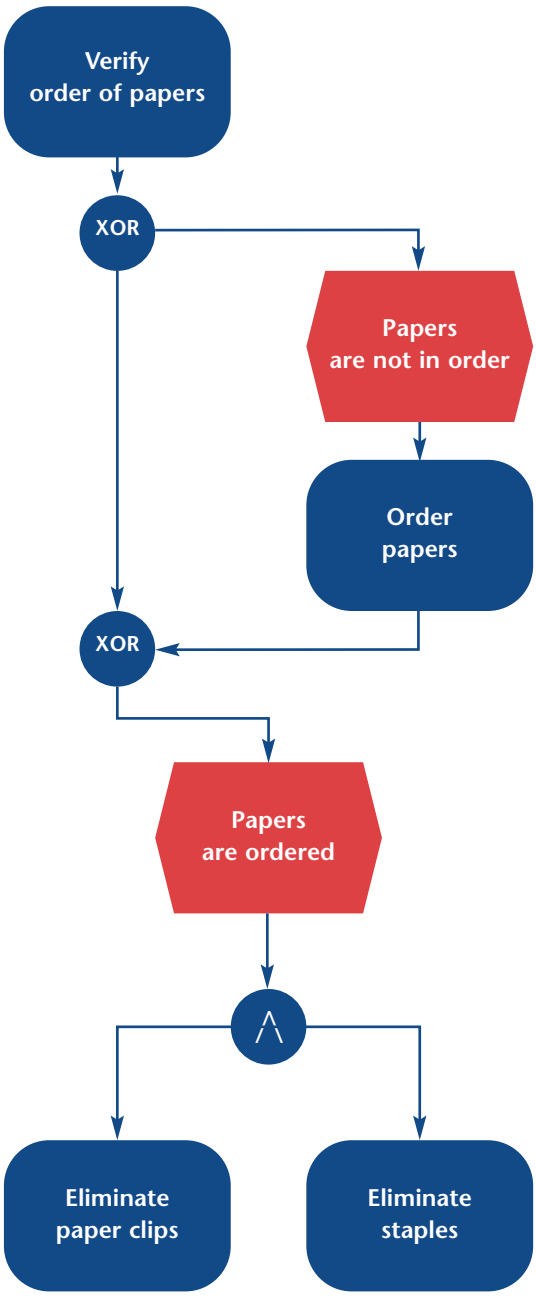
From EPC to executable rules

The EPC model shows graphically the behaviour of a system and constitutes thus an intuitive description and explanation of the domain knowledge. To make it executable we will have to convert it into rules. To do this we simply add an IF to the conditions and a THEN to the actions. For example the last EPC fragment can be written as the rule 'IF papers are ordered THEN eliminate paper clips and eliminate staples.'

The ITM system cannot test conditions directly as in a regular programme, it must ask the user. According to the user response it will activate a rule or another. The activation of a rule will result in an imperative action communicated to the user. The ITM guides the learner through the decisions and steps of a 'learning by doing' system. However, the ITM is not simply a another type of expert system. It provides also a sophisticated 'explain while doing' capability. After each rule is activated or 'fired' the learner has the ability to ask for explanations.

Humans have difficulties in blindly following instructions without knowing their purpose especially when they are attempting to learn a subject matter. For this reason the ITM system provides the 'why' type of explanation which displays on demand the purpose of an action.

Humans also might forget the causal chain of events which led to a present condition. For example, during a consultation with the ITM the learner might forget how he reached a certain conclusion and might want to know why he is prompted for a certain action. He can request a 'how' type of explanation and the system will provide the preceding step.



Sentence (2)

'In addition, we will verify that no document is missing from the dossier (expedient) and all documents are in order…'

This sentence contains two verification actions: verification of the completeness of the set of documents within a dossier and the verification of the order of these documents within the dossier.

We represent only the 'order verification' in our model.
The outcomes of the verifications are always exclusive (XOR connector).

If the verification result is that the 'papers are ordered', the actions to take are 'eliminate paper clips' and 'eliminate staples'.

If the verification result is 'papers not in order', the action to take is to order them. The result of this action is that 'papers are in order'.

Note that the XOR connector allows thew model to show that the same state ('papers are ordered') can be reached either directly as the result of the verification or indirectly through the action taken after it has been found initially as 'not in order'.

The action nodes in the model must be seen as imperative actions addressed by the system to the user. The user must confirm that the action has been carried out and that the condition 'papers are ordered' becomes true. Once the condition is confirmed, the user can proceed with the next action or actions, in this case 'eliminate paper clips' and 'eliminate staples'.

The relationship between imperative actions and results in the ITM system are different from the computer instructions in a programme where giving an instruction coincides with its execution (except for machine malfunctioning).

In addition to the causal and teleological explanations we consider helpful to have also the 'why not' type of explanation in which the learner can change his answers and see what are the consequences of not carrying out the recommendations of the ITM system. This explanation facility is also allowing the learner to explore various alternatives at any moment of the consultation.

Automatic generation of executable rules

The rules resulting from the EPC model are stored in a database as data where rows spanning the same rule number delimit a rule, a clause type 'P' indicates a premise (condition), a clause type 'C' indicates a conclusion (action), variables represent complete sentences, predicates 'y' indicate 'true' and 'n' indicate 'false'.

Rule Number	Clause type	Variable	Predicate	Text
01	P	Ver_ord	=”y”	Order of documents has been verified
01	P	Doc_ord	=”y”	Documents are in order
02	C	Ord_doc	=”n”	Do not order documents
02	P	Ver_ord	=”y”	Order of documents has been verified
02	P	Doc_ord	=”n”	Documents are not in order
02	C	Ord_doc	=”y”	Order documents

Storing the rules in this way offers the advantage that we can manipulate the rule base using a DBMS system which can help propagating rule modifications and by allowing queries to the rule base. Most important, it allows the static chaining of rules and the possibility to programme a rule generator which will create automatically the executable rules in a form similar to the one below:

Rule No 1

IF order of documents has been verified
documents are in order
THEN do not order documents

Rule No 2

IF order of documents has been verified
documents are not in order
THEN order documents

Rule No 3

IF documents are ordered
THEN eliminate paper clips and eliminate staples

Each rule clause is represented by a variable which will receive the values ‘yes’ or ‘no’.

As the arrow indicates, the rules are chained according to the principles that the result of one rule action becomes a condition in the next rule. Rule chaining is done by the inference engine of the expert system in the rule base before rule execution. In this way, during the consultation, the system achieves real time, executing directly the rules without spending time on searching and chaining them during the consultation. This feature has a significant role for the Internet-based applications where performance can be an issue.

From the rule database, executable rules are generated by a programme according to the following rule frame which we present as pseudo-code. Note that the rule generator can be adapted with small changes to generate rules in any programming language. Actual rules will be instantiations of the following rule frame:

```
IF  action-flag = true                  ← This variables ensures chaining with the previous rule.
    ASKUSER Condition No 1              ← The user is asked about existent conditions.
    ASKUSER Condition No 2

    IF  Condition No 1 = true AND       ← These are the executable rule clauses (tests and
        Condition No 2 = true             actions).
THEN DO ACTION No 1
ELSE DO ACTION No 2

    CALL How-explanations               ← Subroutines for causal, teleological, and ‘why not’
    CALL Why-explanations                 explanations
    CALL Change-answers
ENDIF
```

The application code which implements the ITM system described above is generated automatically from the rule base without programmer intervention. This reduces considerably the application development effort which can focus more on the high level knowledge engineering.

Multimedia interactive dialogue

The learner will interact with the rule base through screens, voice messages, pictures and videos.

In the ASK USER screen, the system will ask the user about conditions through lines of text and voice messages. Eventually, conditions asked about will be exemplified by pictures. In our example a picture of unordered documents will exemplify the second condition of the rule No 1. Example dialogue:

ITM: Have you verified the of documents?
USER: yes.
ITM: Are the documents in order?
USER: No

In the ACTION screen, the system processes the user answers it will display a text and voice message with recommendations for actions to be carried on by the learner. The ordering of papers can be exemplified by a video showing a real life action. The activation of rule No 2 will produce the following action prompt from the system:

ITM: I suggest that you order the documents.

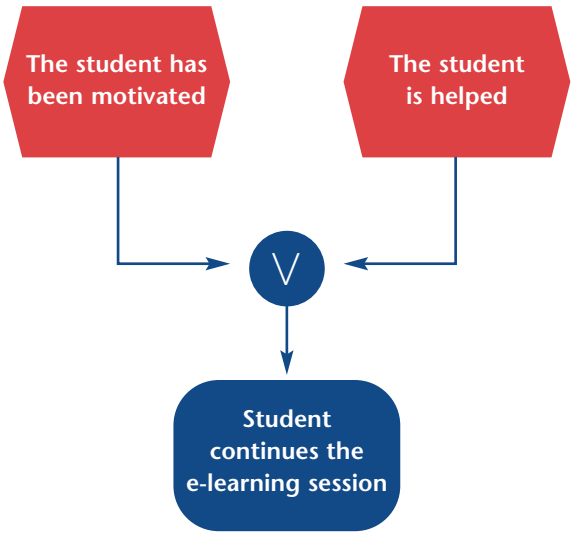
The EXPLANATION screen will allow the learner to ask for explanations regarding the activation of the current rule:

ITM: Do you want a 'why' type of explanation?
USER: Yes.
ITM: You need to order the documents because you are preparing them for transfer.
ITM: Do you want a 'how' type of explanation?
USER: Yes.
ITM: You need to order the documents because they were out of order.
ITM: Do you want to change your answers?
USER: Yes

The system will redisplay the ASK USER screen and the user will give a new answer that the documents are this time, ordered. The ITM will no longer prompt for the ordering action and will recommend instead the actions 'eliminate paper clips and eliminate staples'.

EPC: The unified model for knowledge representation

We have seen how the EPC can be used for representing domain knowledge. The same EPC can be used to represent pedagogical knowledge and the LKE itself.

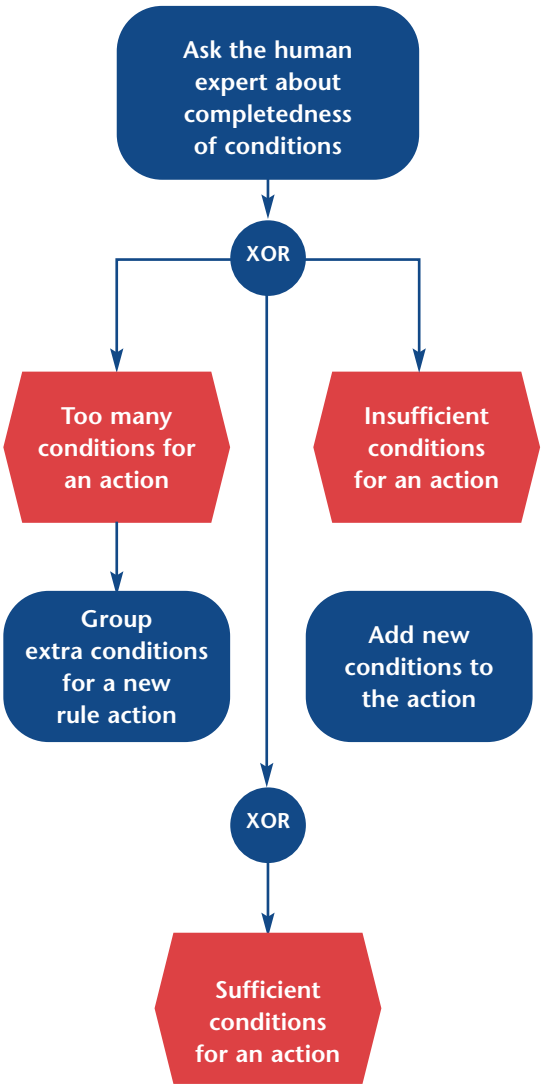


Representing pedagogical knowledge

Pedagogical knowledge can also be expressed with the same EPC model used for the representation of the domain knowledge.

The source of the pedagogical model to be developed can be found in various pedagogy textbooks. The model can also include contributions from teachers, specialists in the archiving practice.

The use of the EPC model for representing both domain and pedagogical knowledge is offering the content developers a unique integrated and unified knowledge tool.



Representing LKE itself

This diagram represents a fragment of a much larger model where the LKE itself is represented using the same EPC.

In this fragment, the knowledge engineer is building domain rules. A correct rule must have a sufficient number of conditions for an action.

Too few conditions will make the rule too general and cause the system not to be discriminating enough.

Too many conditions will make the rule cumbersome and difficult to understand. In this case it is recommended that conditions should be grouped and a new rule constructed.

Summary of the overall ITM architecture

Author	The domain human expert organises his knowledge and experience by writing manuals and textbooks.
LKE	The knowledge engineer applies the LKE methodology to the textbooks and asks clarifications from the human expert.
EPC	The result of applying the LKE methodology is the representation of the textbook knowledge in the form of an EPC model built by the knowledge engineer using the ARIS tool set.
Rule database	A post-processor programme reads the ARIS tool set reports from the EPC model and derives the corresponding rules which are stored in a database in tabular form.
Executable rules generator	A special programme, called a rule generator reads the rule tables and creates automatically the executable rules in a programming language of choice such as JAVA, C++, etc. in the form of IF … THEN … ELSE statements.
Consultation session with ITM	The ITM system uses the executable rules to produce an interactive dialog which guides the learner through a 'learning by doing' session and provide causal and teleological explanations. It also allows the user to explore various solutions by letting him change the answers at any time during the consultation.

Future projects

We believe that the intelligent tutoring modules (ITMs) could provide the necessary advanced technology for the current European initiatives in e-learning such as E-TERM, Socrates, Recpro. The best archiving practices could be rigorously modelled using the EPC models from which executable and interactive knowledge bases could be made available across Europe as teaching aids available for consultation on the Internet any time, any place. A common archiving knowledge-executable repository could be built as an addition to existing manuals. We hope that our pilot project at The University of Alcala de Henares will provide a glimpse of what the 'intelligent, interactive manual' could do on a European scale.

Un módulo de tutoría inteligente para la formación de los archiveros

Paltin Sturdza, María Luisa Conde, Luis López Corral

Los ***programas de aprendizaje electrónico (e-aprendizaje)*** son excelentes en cuanto a la oferta de herramientas que permiten elaborar pantallas interactivas y atractivas visualmente, pero detrás de las pantallas se oculta una escasa representación de contenidos y pedagogía, careciendo por tanto de «inteligencia». Los ***sistemas de tutoría inteligentes (STI)*** son el fruto de la inteligencia artificial y tratan de emular funciones de tutoría semejantes a funciones humanas. El término «inteligente» designa un sistema capaz de saber lo que es necesario enseñar, cuándo hacerlo y cómo.

La ***estrategia que proponemos*** consiste en combinar las técnicas de la tutoría inteligente con un sistema de aprendizaje electrónico del mercado, desarrollando un módulo de tutoría inteligente (MTI) destinado a un producto de aprendizaje electrónico preexistente. Este módulo evaluará la comprensión y los progresos realizados por el estudiante a través de las interacciones entre los distintos submódulos. Las instrucciones podrán ser adaptadas automáticamente por el *módulo pedagógico* en función de las necesidades del estudiante, sin intervención de un instructor humano. El módulo controlará las interacciones con el estudiante y creará un *módulo de estudiante* para cada individuo.

El ***módulo de conocimientos archivísticos*** contiene información sobre los archivos, en particular, los hechos y los conceptos, así como los procesos necesarios para solucionar los problemas que se plantean en el sistema de archivos. El ***módulo pedagógico*** contiene normas que permiten apreciar el nivel de comprensión de las cuestiones de archivo por el estudiante (tal como las representa el *módulo del estudiante*) y compararlo con la estructura real de los conocimientos (representada por el *módulo de los conocimientos*).

El ***módulo del estudiante de archivística*** almacena información específica de cada estudiante. También contiene información sobre la comprensión de éste, recurriendo a un modelo de estilo de aprendizaje. A tal efecto utiliza una herramienta de diagnóstico, contenida en el *módulo pedagógico*, para saber el estado de los conocimientos del estudiante sobre la disciplina. El *módulo del estudiante* es necesario para personalizar las instrucciones en función de las características personales del estudiante y sus necesidades de formación. Sin esta información, el *módulo pedagógico* del profesor no dispone de ninguna base para tomar decisiones, viéndose obligado a tratar a todos los estudiantes de la misma manera.

En la ***aplicación*** de nuestro sistema, los *conocimientos archivísticos* se representarán en forma de normas de producción, ampliamente utilizadas en los sistemas expertos. Los conocimientos temáticos se adquirirán utilizando metodología LKE (Linguistic Knowledge Engineering) y un sistema experto genérico, creado por el Dr. Sturdza, que deberá adaptarse al medio de aprendizaje electrónico en la red. El *módulo pedagógico* se aplicará utilizando el mismo sistema experto genérico. El sistema contendrá una base de normas pedagógicas así como un motor de inferencias, que aplicará estas normas en función del estado de los conocimientos del estudiante memorizado en el *modelo del estudiante*.

Nuestro ***nuevo enfoque respecto del desarrollo de un STI*** consiste en modelar los conocimientos específicos aplicando una versión modificada de las cadenas de procesos dirigidas por los acontecimientos (modelo EPC). El modelo EPC permitirá derivar directamente las normas de producción. Este modelo de las normas EPC se aplicará a la vez a los conocimientos temáticos y a los conocimientos pedagógicos y ofrecerá un marco de aplicación unificado, permitiendo así un considerable ahorro de esfuerzos en la aplicación del sistema de aprendizaje.

El modelo EPC ofrece una imagen gráfica de la base de normas, fácil de comprender y de modificar. Dado que las normas se obtienen directamente a partir del modelo, cualquier modificación de este último se reflejará en la norma ejecutable. De esta manera, el esfuerzo de mantenimiento del sistema se ve muy reducido. Esta característica será aplicable tanto a los conocimientos temáticos como a los conocimientos pedagógicos.

La elaboración del curso de archivística en el marco de las normas EPC se basará en la metodología LKE. Ésta reduce considerablemente la inversión necesaria para transformar textos de enseñanza archivística (manuales) en una base de normas que serán tratadas por el motor de inferencias del STI.

Esta tutoría inteligente proporcionará un módulo de «aprender haciendo», en el que se propondrán al estudiante situaciones realistas para que éste las resuelva.

Ein intelligentes E-Learning-Modul für die Ausbildung von Archivaren

Paltin Sturdza, María Luisa Conde, Luis López Corral

E-Learning-Pakete sind wunderbar, wenn es darum geht, dem Gestalter Hilfsmittel an die Hand zu geben, mit denen er visuell ansprechende und interaktive Darstellungen erstellen kann. Inhalte und pädagogische Elemente kommen jedoch dabei zu kurz, d. h. es fehlt an „Intelligenz". Mit dem ***Intelligent Tutoring System (ITS),*** das aus der Künstlichen Intelligenz (Artificial Intelligence – AI) hervorgegangen ist, wird versucht, „menschliche" Lehrfähigkeiten zu emulieren. Mit „intelligent" ist dabei ein System gemeint, das „weiß", wann und wie etwas zu unterrichten ist.

Unser ***Lösungsvorschlag*** besteht darin, durch die Entwicklung eines ITM (Intelligent Tutoring Module) für ein vorhandenes E-Learning-Produkt die ITS-Technologie mit einem im Handel erhältlichen E-Learning-System zu kombinieren. Dabei beurteilt das E-Learning-ITM den Kenntnisstand und die Fortschritte des Lernenden durch die Interaktion dieser Module. So kann dann der Unterricht mit dem *Didaktikmodul* ohne Eingreifen einer menschlichen Lehrkraft automatisch auf die Erfordernisse des Studenten zugeschnitten werden. Das E-Learning-ITM überwacht die Interaktionen jedes Lernenden und erstellt ein individuelles *Studentenmodul*.

Die Komponente ***Archivfachwissenmodul*** enthält Informationen über das Fachgebiet der Archivierung, z. B. Fakten und Begriffe, und die Verfahren, die zur Lösung von Problemen innerhalb des Systems notwendig sind. Das ***Didaktikmodul*** enthält Regeln, die es ermöglichen, den Kenntnisstand des Lernenden im Fach Archivierung (repräsentiert durch das *Studentenmodul*) zu beurteilen, und vergleicht den Kenntnisstand mit der eigentlichen Wissensstruktur (repräsentiert durch *Fachwissen*).

Das ***Archivarstudentenmodul*** speichert spezielle Informationen für jeden einzelnen Lernenden. Es enthält Informationen zum Kenntnisstand des Lernenden im Fach Archivierung auf der Basis eines Modells des Lernstils des Lernenden. Mit Hilfe eines Diagnose-Tools, das im *Didaktikmodul* enthalten ist, wird der Wissensstand des Lernenden im Lehrfach ermittelt. Das *Studentenmodul* ist notwendig, um die Anweisungen auf die Schwachstellen und den Lernbedarf des Lernenden zuzuschneiden. Ohne dieses Wissen fehlt dem *Didaktikmodul* des Tutors die Grundlage, um Entscheidungen zu treffen, und es ist gezwungen, alle Lernenden gleich zu behandeln.

Bei unserer *Implementierung* wird das *Archivfachwissen* über Erstellungsregeln dargestellt, die in Expertensystemen umfassend genutzt werden. Das Fachwissen wird mittels der LKE-Methodik (Linguistic Knowledge Engineering) einer von Dr. Sturdza eingerichteten Expertensystemoberfläche erworben, die dann an die E-Learning-Umgebung im Internet angepasst werden muss. Das *Didaktikmodul* wird mittels dieser Expertensystemoberfläche implementiert. Das System wird die Didaktikbasis sowie eine Inferenzmaschine enthalten, die diese Regeln je nach Wissensstand des Lernenden, der im *Studentenmodul* gespeichert ist, in Anwendung bringt.

Unsere ***neuartige Herangehensweise an die ITS-Entwicklung*** besteht in der Modellierung des Fachbereichswissens unter Zuhilfenahme eines modifizierten Formalismus für ereignisgesteuerte Prozessketten (Event Driven Process Chains, EPC). Das EPC-Modell wird eine direkte Ableitung von Erstellungsregeln ermöglichen. Der EPC-Regel-Formalismus wird sowohl für das fachliche als auch für das didaktische Wissen genutzt und bietet damit einen einheitlichen Implementierungsrahmen, der den Aufwand bei der Implementierung des Lernsystems deutlich verringert.

Das EPC-Modell bietet eine leicht verständliche und modifizierbare grafische Darstellung der Regelbasis. Da die Regeln direkt aus dem Modell abgeleitet werden, sind Abwandlungen des Modells über die ausführbare Regel problemlos zu übertragen. Auf diese Weise wird der Aufwand für die Laufendhaltung des Gesamtsystems erheblich gesenkt. Dieses Merkmal gilt sowohl für das fachliche als auch für das didaktische Wissen.

Die Entwicklung eines Lehrgangs Archivierung in einem EPC-regelgestützten Rahmen basiert auf der Methodik des LKE (Linguistic Knowledge Engineering). Durch LKE reduziert sich der Aufwand für die Umwandlung von Unterrichtstexten (Handbüchern) für Archivare in eine Regelbasis, die von der ITS-Inferenzmaschine verarbeitet wird, erheblich.

Das E-Learning-ITM wird eine „Learning-by-doing"-Komponente enthalten, in der realistische Situationen dargestellt werden, die der Lernende bewältigen muss.

Un module de tutorat intelligent pour la formation des archivistes

Paltin Sturdza, María Luisa Conde, Luis López Corral

Les ***progiciels d'e-apprentissage*** sont excellents dans l'offre d'outils de conception permettant d'élaborer des écrans interactifs et attractifs sur le plan visuel, mais derrière les écrans se cache une représentation superficielle du contenu et de la pédagogie qui manque d'«intelligence». Les ***systèmes de tutorat intelligents*** (STI) sont le fruit de l'intelligence artificielle et s'efforcent d'émuler des fonctionnalités de tutorat ressemblant aux fonctionnalités humaines. Le terme «intelligent» désigne un système ayant la capacité de savoir ce qu'il faut enseigner, quand le faire et comment.

La ***stratégie que nous proposons*** consiste à combiner les techniques du tutorat intelligent à un système d'e-apprentissage du marché, en développant un module de tutorat intelligent (MTI) destiné à un produit d'e-apprentissage préexistant. Ce module évaluera la compréhension et les progrès réalisés par l'étudiant au travers des interactions entre les différents sous-modules. Les instructions pourront être personnalisées automatiquement par le *module pédagogique* en fonction des besoins de l'étudiant, sans intervention d'un instructeur humain. Le module contrôlera les interactions avec l'étudiant et construira un *module de l'étudiant* pour chaque individu.

Le ***module de connaissances archivistiques*** renferme des informations sur les archives, notamment des faits et des concepts, ainsi que les processus nécessaires pour résoudre les problèmes se posant dans le système d'archives. Le ***module pédagogique*** renferme des règles qui lui permettent d'apprécier le niveau de compréhension des questions d'archivage par l'étudiant (telles que les représente le *module de l'étudiant*) et de les comparer à la structure réelle des connaissances (représentées dans le *module des connaissances*).

Le ***module de l'étudiant en archivistique*** mémorise des informations propres à chaque apprenant. Il renferme également des renseignements sur la compréhension de celui-ci, en ayant recours à un modèle de style d'apprentissage. Il utilise à cet effet un outil de diagnostic, contenu dans le modèle pédagogique, pour extraire l'état des connaissances de l'apprenant dans la discipline. Le *module de l'étudiant* est nécessaire pour personnaliser les instructions en fonction des dispositions personnelles de l'étudiant et de ses besoins de formation. Sans ces renseignements, le *module pédagogique* du tuteur ne dispose d'aucune base pour prendre des décisions et est alors contraint de traiter tous les étudiants de la même façon.

Dans l'***implémentation*** de notre système, les *connaissances archivistiques* seront représentées sous la forme de règles de production, largement utilisées dans les systèmes experts. Les connaissances thématiques seront acquises en utilisant la méthodologie LKE (Linguistic Knowledge Engineering) et un système expert générique, construit par l'auteur, qui devra être adapté à l'environnement d'e-apprentissage. Le *module pédagogique* sera implémenté en utilisant le même système expert générique. Le système renfermera une base de règles pédagogiques ainsi qu'un moteur d'inférences qui appliquera ces règles en fonction de l'état des connaissances de l'étudiant mémorisé dans le *module de l'étudiant*.

Notre ***nouvelle approche du développement d'un STI*** consiste à modéliser les connaissances spécifiques en appliquant une version modifiée des chaînes de processus dirigées par les événements (formalisme EPC). Le modèle EPC permettra de dériver directement les règles de production. Ce modèle des règles EPC sera appliqué à la fois aux connaissances thématiques et aux connaissances pédagogiques, et offrira un cadre d'implémentation unifié, permettant ainsi d'économiser considérablement sur l'implémentation du système d'apprentissage.

Le modèle EPC offre une image graphique de la base de règles, facile à comprendre et à modifier. Étant donné que les règles sont obtenues directement à partir du modèle, toute modification de ce dernier se répercutera aisément sur la règle exécutable. Il s'ensuit une réduction significative de l'investissement dans la maintenance. Cette caractéristique sera applicable à la fois aux connaissances thématiques et aux connaissances pédagogiques.

La rédaction du cours d'archivistique dans le cadre des règles EPC reposera sur la méthodologie LKE. Celle-ci réduit considérablement l'investissement nécessaire pour transformer des textes d'enseignement archivistique (manuels) en une base de règles qui seront traitées par le moteur d'inférences du STI.

Ce tutorat intelligent fournira un module de «formation sur le tas», dans lequel des situations réalistes seront proposées à l'étudiant et traitées par lui.

Parallel session 6

Wednesday 8 May 2002

CAPTURING AND TRANSFORMING INFORMATION
Automatic indexing technologies and implementing retrieval solutions

Chairperson:	Yannis Ioannidis (Greece)
Co-chair:	Tom Quinlan (Ireland)
Rapporteur:	Thekla Kluttig (Germany)

Bernhard Lindgens

Bernhard Lindgens has been responsible for applying information technology in fiscal investigations and in tax-fraud prevention at the national and international level in the German Federal Ministry of Finance since 1987. Since 2000 he has been responsible for the conception and development of a central Internet-inquiries service (Zentralen Internet-Ermittlung).

The intelligent recording, indexing and automatic categorisation of information

Report on practical experience in countering tax fraud in e-commerce

Bernhard Lindgens

According to a conservative estimate, the turnover in goods and services in the German Internet marketplace already amounted to DEM 7 000 million in 2000. The reasons for the expansion of electronic commerce are obvious. The use of the Internet as a trading and distribution platform involves only a fraction of the personnel, space and storage costs otherwise incurred by conventional shops and service operations. As many access and content providers are now offering premanufactured web shops at reasonable prices, e-commerce has become an option for businesses without any experience of Internet technology and the setting up of web shops.

The increase in e-commerce has led to constantly rising (turnover) tax losses for the revenue authorities of the EU Member States. Up to now, it has been impossible to give even an approximate estimate for the tax income lost. However, it is clear from surveys that in Germany at least, a not inconsiderable proportion of Internet goods and services providers are not registered with the responsible financial authorities, in contravention of clear legal requirements. It is not only illegal tax advantages due to the supposed anonymity in the world wide web that are forcing the revenue authorities to take rapid action. When Internet companies are not taxed, this necessarily leads to distortions in competition at a national level.

As early as 1998, the German revenue authorities were attempting to use technical solutions for central registration of Internet providers for tax purposes. In 2000, the company SER, acting for the Federal Treasury Department, carried out a study with the aid of the associated search programme 'SERglobalBrain' which produced some alarming results. Because of the great number of business fields, some disreputable, three separate searches were carried out: for apparent investment swindlers, porn providers and reputable trading businesses and service providers in the world wide web. Internet sites of the providers selected by the tax authorities served by way of examples for search patterns and learning quantities. The number of businesses registered for tax was, as expected, low for investment swindlers and porn providers. What was surprising was the high proportion of unknown, in terms of tax registration, apparently reputable Internet businesses. Even when compared manually with the revenue authorities' database of businesses, which is centrally managed in Germany (and contains all the businesses which are liable for tax in the individual *Länder*), and corrected using personal and business data from queries to national and international domain administrations, a considerable number of non-registered trading businesses remained.

These results were an important factor in commissioning the company SER early in 2001 to produce a feasibility study on the setting-up of a central Internet investigation with the aim of registering Internet providers for tax purposes. Moreover on 19 December 2001 the law against tax evasion introduced a legal basis for the observation of services offered on the Internet and the support of turnover tax on e-businesses. In this context it must be emphasised that this legally established permission to observe trading activities on the Internet is in no way intended to hinder the expansion of e-commerce. The task for the central Internet investigation, created specially at federal level, is only to ensure that Internet providers of goods and services which are liable to tax in Germany are registered for tax. In this way equality of taxation with other conventional businesses is assured. For this reason there is no monitoring of individual business transactions or even payment processes.

The working methods of the Internet investigation follow the concept of a learning organisation. Experiences from the tax law professional side are entered into a knowledge base using computers, in order to continually improve the quality and quantity of the Internet businesses to be qualified (knowledge management). The German Federal Treasury Department will use modern tech-

niques for recording, indexing and automatically categorising information when carrying out the component tasks listed below:

- periodic and automated recording of businesses supplying services on the Internet and which are liable for tax in Germany according to specialist categories (language, currency, business objective, geographical location of a company's office, contact and business addresses, telecommunications connections, provider data, advertising banner etc.);
- specified searches for problematic business areas for the purposes of taxation, e.g. 'promoters' of tax evasion and tax havens;
- automatic coordination of qualified URL and web pages with the database of businesses which are already registered for tax;
- recording web presences which have not been registered until financial-legal or criminal proceedings have been concluded.

To the best of our knowledge at this point, an EU-wide announcement regarding the necessary soft- and hardware will be made in the third quarter of 2001.

When attempting to describe verbally the criteria for an Internet business's liability to German taxation it soon becomes clear that the (cost-effective) use of conventional search engines and strategies, contrary to initial expectation, does not lead to the desired result. A central problem in Internet coordination for monitoring registration for tax is how to define 'German' businesses. The form of the Internet presence (language, currency, order forms, terms and conditions of business, etc.) is in no way conclusive when establishing a German tax liability. The solution is rather to convey an intelligent search routine providing a picture of how German businesses look. Internet entries for trading purposes may be divided basically into three main categories.

- **Category 1**

The correct type works only in its officially designated area of business, is registered for tax and uses its Internet presence as an extra and up to date advertisement or distribution channel. As a rule, this overwhelming majority of businesses may be found using the usual search engines.

- **Category 2**

The negligent type does not work in its supposed area of business or at least not primarily and may also have failed to register its trade. This includes for example many suppliers in the tourism branch, 'trend suppliers' like vendors of ringing tones for mobile phones and also purveyors of porn. These businesses may be found with the usual search engines but are usually registered with another commercial enterprise for tax purposes.

- **Category 3**

The wilfully unlawful type which consciously attempts to exploit the supposed anonymity and global nature of the world wide web in order (also) to obtain illegal tax advantages. An example would be capital investment swindlers. In this case one may assume that the business is not registered for tax. Since these suppliers are aware that the security authorities all over the world are now searching for offers of this kind, they are generally no longer using the catchwords that were usual only a short time ago, such as 'save tax', 'tax-free investments' and 'tax havens'. For this reason alone, it is futile to search for suppliers in Category 3 with the world wide web's established search engines.

As a result of the problem outlined above, the only possibility given the present state of technology is the use of an intelligent, knowledge-based software in order to search for and classify individual business activities. This software recognises the contents of web pages from texts it has learned (full-text search procedures) and a learning quantity of selected web presences which also allows searching with the assistance of metadata. Searching is associative, i.e. imprecision and departures in the words are permitted. Yet web pages with similar contents and a high degree of probability are established and then summarised through a classification. The procedure devised provides that individual classes are defined by tax experts and given a specific trade reference number.

After the training process has been completed a knowledge base is provided to be used in the automated mass registration of the businesses represented in the Internet and liable for tax in Germany. This mass registration is almost exclusively of businesses in Categories 1 and 2. These are selected in a preliminary process using the revenue authorities' business database in a mass comparison. The remainder of the addresses which are not identical are used with their links in the knowledge database as a negative exclusion. The businesses whose addresses are identical with those in the business database and their Internet sites are excluded from the first registration (positive exclusion). The positive exclusion is intended to ensure that non-critical businesses that are registered for (turnover) tax in an orderly manner will not be taken into account in further search descriptions. However, they are to be monitored for changes within the framework of periodic inquiries. Particularly in Category 2 the improper inclusion of another Internet business or one which does not belong to the registered business field may lead to income from e-business being completely overlooked in the taxation of the officially designated business activity.

An example of the typical progression of a search may elucidate how techniques are used for the intelligent recording, indexing and automatic categorisation of information in the coordination of registration for tax purposes.

- **Step 1**

Definition of the business activity to be investigated and laying down the relevant trading reference number for the class(es) concerned. The definition is provided by a tax expert, e.g. by giving the number 65.23.1 to the services branch in the area of equipment consultancy.

- **Step 2**

Defining the class(es). Commercial activities are equated with classes. Training texts are used to examine whether the classes have been correctly chosen. Any cases of overlap are discovered in the process.

- **Step 3**

The correct texts and documents for the business number and class are put together by the tax experts. Example: sites found on the Internet are considered by way of examples and the web pages marked where the nature of the business is clear: 'offshore trust', 'corporate shell', or 'tax-free plus capital'.

- **Step 4**

The preparation of texts and the implementation of the learning process are carried out by researchers with the support of a knowledge management component and a knowledge base at their workplaces. Example: the marked texts are input to the software in a class with the designation '65.23.1' as a learning quantity, learned by the software and marked as such.

- **Step 5**

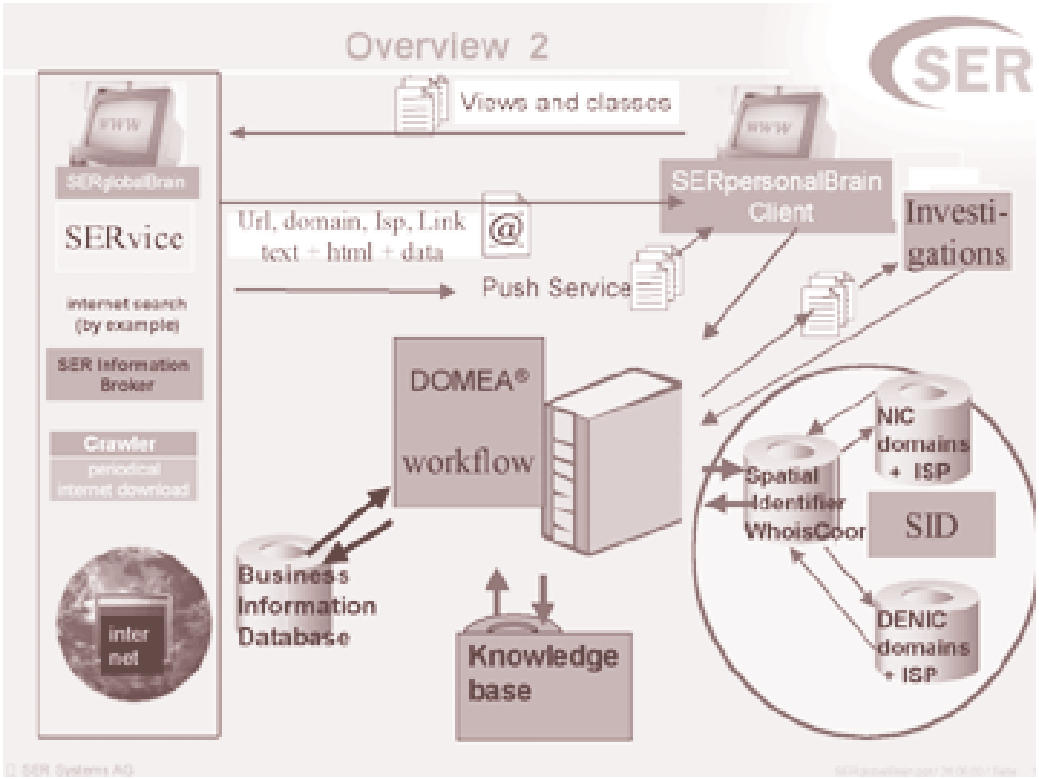
The search is started by the researchers at their workplaces after release by the system administrators. The confidence level — degree of probability with which the Internet presences found fit with the categories learned — are continually monitored by researchers.

- **Step 6**

After the necessary investigation and correction of personal and business data in national and international domain administrations coordination takes place with the revenue authorities' locally held business database.

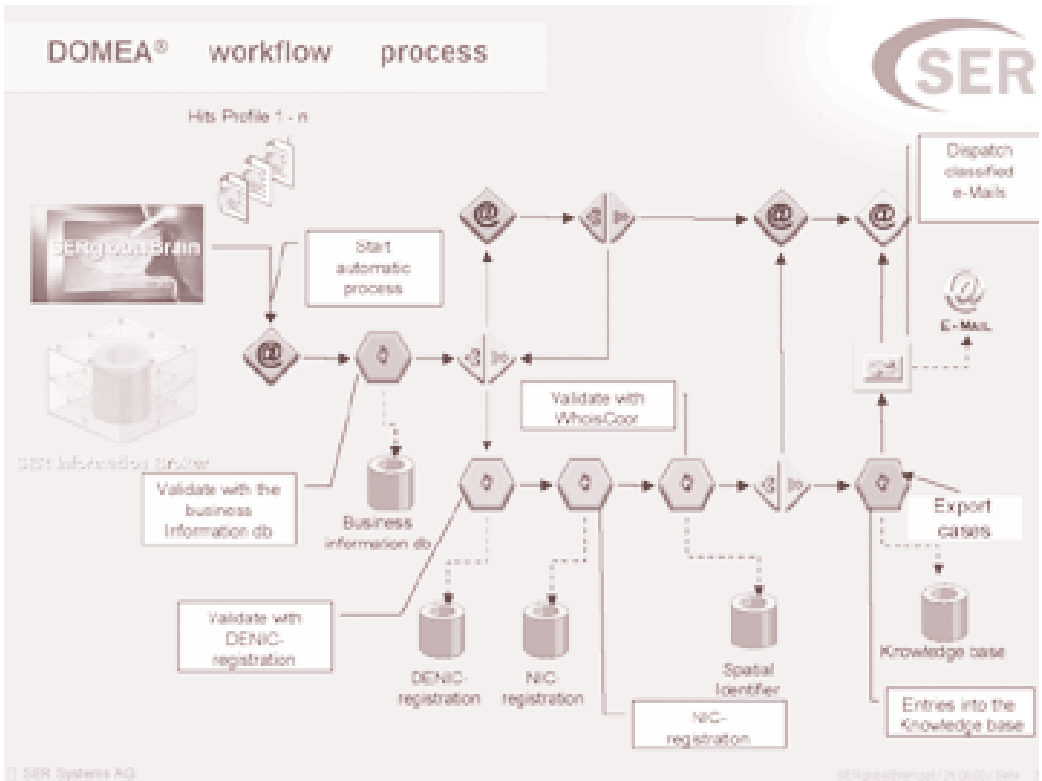
A representation of the file processing also provided in the overall system for the actual taxation of those Internet suppliers who are not listed in the business database, must be omitted at this point for obvious reasons.

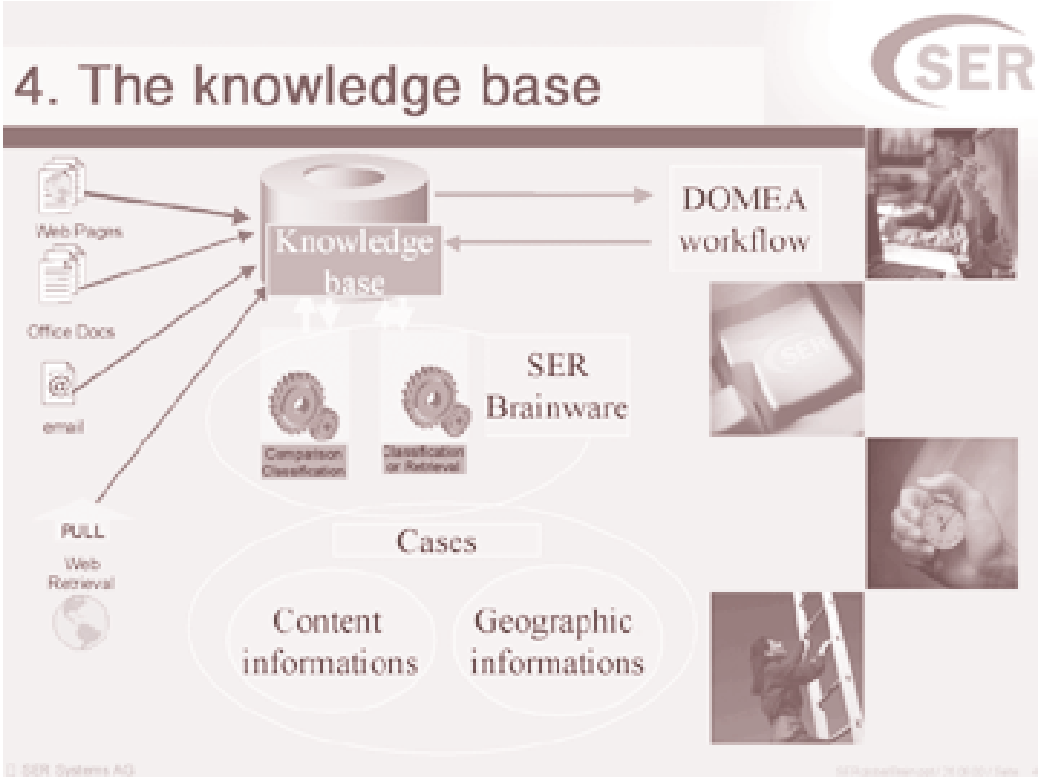
Illustrations



SER Internet Services with SERglobalBrain

System architecture





Introducción de datos inteligente, indización y clasificación automática de la información

Informe práctico sobre la lucha contra el fraude fiscal en las transacciones comerciales electrónicas

Bernhard Lindgens

El desarrollo del comercio electrónico supone para la administración fiscal de los Estados miembros de la Unión Europea unas depreciaciones fiscales en constante aumento. La imposición de las prestaciones de servicios proporcionadas por vía electrónica, en particular, ya no puede ser garantizada por los medios tradicionales. Ahora bien, estas depreciaciones fiscales en el comercio electrónico no sólo obligan a adoptar medidas urgentes, sino que la no imposición de las operaciones realizadas por Internet también da lugar necesariamente a distorsiones de la competencia a nivel nacional.

El Ministerio Federal de Hacienda alemán utiliza técnicas modernas de introducción inteligente, indización y clasificación automáticas de información para luchar contra el fraude fiscal en Internet. El objetivo que se persigue es garantizar el registro fiscal de los prestadores de bienes y servicios en Internet sujetos al impuesto en Alemania. La utilización de los motores y de las estrategias de investigación tradicionales no suele dar los resultados deseados. La calificación de empresario «alemán» constituye el problema clave a la hora de comprobar el registro fiscal de los operadores en Internet a efectos de la rectificación. La presentación de la presencia en Internet (lengua, divisa, formularios de pedido, condiciones comerciales, etc.) no permite por sí misma establecer sin errores una obligación fiscal en Alemania. La solución consiste en transmitir a un motor de investigación inteligente las características de la apariencia de los operadores alemanes en Internet. Las ofertas encontradas en la red se añaden

a los conocimientos adquiridos por el sistema y permiten mejorar la calidad de los resultados de la investigación. Si se trata de una empresa alemana, la rectificación se produce en una segunda fase, recurriendo a la base de datos central de empresas de la administración fiscal. Los prestadores de servicios que no aparezcan en esta base de datos serán comunicados automáticamente a la administración fiscal regional competente a efectos de una investigación posterior.

Además de la posibilidad de rectificación permanente en Internet, que acabamos de describir, para garantizar el registro fiscal de los operadores, las técnicas de introducción inteligente, indización y clasificación automáticas de los operadores profesionales en Internet permiten asimismo luchar con medios modernos y adecuados contra las maniobras de fraude fiscal.

Intelligente Erfassung, Indexierung und automatische Kategorisierung von Informationen

Praxisbericht zur Bekämpfung des Steuerbetrugs beim elektronischen Geschäftsverkehr

Bernhard Lindgens

Mit der Zunahme des elektronischen Geschäftsverkehrs sehen sich die Finanzverwaltungen der EU-Mitgliedstaaten mit stetig steigenden Steuerausfällen konfrontiert. Insbesondere die Besteuerung elektronisch erbrachter Dienstleistungen kann auf herkömmlichem Wege kaum sichergestellt werden. Doch nicht nur die zahlenmäßig schwer zu beziffernden Steuerausfälle beim elektronischen Geschäftsverkehr zwingen zu raschen Maßnahmen: Die Nichtbesteuerung von Internet-Unternehmungen führt zwangsläufig auch auf nationaler Ebene zu Wettbewerbsverzerrungen.

Zur Bekämpfung des Steuerbetrugs im Internet bedient sich das deutsche Bundesministerium der Finanzen moderner Techniken zur intelligenten Erfassung, Indexierung und automatischer Kategorisierung von Informationen. Ziel ist es, die steuerliche Registrierung in Deutschland steuerpflichtiger Anbieter von Waren und Dienstleistungen im Internet sicherzustellen. Der Einsatz herkömmlicher Suchmaschinen und -strategien führt häufig nicht zum gewünschten Ergebnis. Zentrales Problem beim Internet-Abgleich zur Überprüfung der steuerlichen Registrierung ist die Qualifizierung „deutscher" Unternehmer. Von der Gestaltung der Internet-Präsenz (Sprache, Währung, Bestellformulare, Geschäftsbedingungen usw.) kann keinesfalls zweifelsfrei auf eine deutsche Steuerpflicht geschlossen werden. Der Lösungsansatz besteht darin, einer intelligenten Suchroutine das Erscheinungsbild deutscher Internet-Unternehmungen zu vermitteln. Aufgefundene Web-Präsenzen werden der Lernmenge hinzugefügt und verbessern die Qualität der Suchergebnisse. Handelt es ich um deutsche Unternehmer, erfolgt im zweiten Schritt ein Abgleich mit der zentral geführten Unternehmerdatenbank der Finanzverwaltung. In der Unternehmerdatenbank nicht nachgewiesene Anbieter werden der zuständigen Landesfinanzbehörde zur weiteren Recherche automatisiert übermittelt.

Neben dem beschriebenen permanenten Internet-Abgleich zur Sicherstellung der steuerlichen Registrierung ermöglicht die intelligente Erfassung, Indexierung und automatische Kategorisierung gewerblicher Internet-Anbieter darüber hinaus eine zeitnahe und geschäftsfeldbezogene Bekämpfung von (Steuer)Betrügereien.

Saisie intelligente, indexation et classement automatique d'informations

Rapport pratique sur la lutte contre la fraude fiscale dans les transactions commerciales électroniques

Bernhard Lindgens

Le développement du commerce électronique confronte l'administration fiscale des États membres de l'UE à des moins-values fiscales en augmentation constante. L'imposition des prestations de services fournies par voie électronique, en particulier, ne peut plus être assurée par les moyens traditionnels. Or, non seulement ces moins-values fiscales dans le commerce électronique obligent à prendre des mesures d'urgence, mais la non-imposition des opérations sur l'internet aboutit aussi nécessairement à des distorsions de concurrence au niveau national.

Le ministère fédéral allemand des finances utilise les techniques modernes de la saisie intelligente, de l'indexation et du classement automatiques d'informations pour lutter contre la fraude fiscale sur l'internet. Le but recherché est de garantir l'immatriculation fiscale des prestataires de biens et de services sur l'internet assujettis à l'impôt en Allemagne. L'utilisation des moteurs et stratégies de recherche traditionnels ne conduit pas souvent au résultat souhaité. La qualification d'entrepreneur «allemand» constitue le problème clé lorsque l'on vérifie l'immatriculation fiscale des opérateurs sur l'internet aux fins du redressement. La présentation même de la présence sur l'internet (langue, devise, formulaires de commande, conditions commerciales, etc.) ne permet nullement de conclure sans se tromper à une obligation fiscale en Allemagne. La solution consiste à transmettre à un moteur de recherche intelligent les caractéristiques d'apparence des opérateurs allemands sur l'internet. Les offres trouvées sur le web sont ajoutées aux connaissances acquises par le système et permettent d'améliorer la qualité des résultats de recherche. S'il s'agit d'une entreprise allemande, le redressement intervient dans un deuxième temps en faisant appel à la base de données centrale d'entreprises de l'administration fiscale. Les prestataires n'apparaissant pas dans cette base de données sont communiqués automatiquement à l'administration fiscale régionale compétente aux fins de la recherche ultérieure.

Outre la possibilité de redressement permanent sur l'internet, que l'on vient de décrire, pour assurer l'immatriculation fiscale des opérateurs, les techniques de saisie intelligente, d'indexation et de classement automatiques des opérateurs professionnels sur l'internet permettent de lutter, par des moyens modernes et adaptés, contre les manœuvres de fraude (fiscale).

A survey of advancements in document recognition and document retrieval technologies

Hiromichi Fujisawa

1. Introduction

A survey of document recognition and document retrieval technologies, which can be seen in research and development in these 10 years, is presented. Actually, document archiving in terms of images started in the early 1980s when the optical storage technology emerged. In response to the new technology, commercial products such as 'optical (electronic) document filing systems' were developed, and large archiving systems such as those for patent automation were built in Japan and the United States. Digital library initiatives in 1990s in the United States, Europe and Japan have also spurred the technologies in consideration. Since then, the document recognition technology has been long expected to facilitate automatic indexing and full-text search of the image documents.

The researches on document recognition for this kind of application started in the mid-1980s, and have been very active especially over these 10 years. Although some of the developed techniques have been put into practice, document recognition can never be perfect in its accuracy, and new technologies are still under active research to achieve improved accuracy and higher functionality.

One of the new research trends is an attempt to make the existing information retrieval methods tolerant of recognition errors by admitting that 'recognition can never be perfect.' Another is to develop a totally new method of image document search. This paper reviews recent research work in these directions.

The reviewed work is from a series of symposia on document analysis and information retrieval, SDAIR, which were held annually from 1992 to 1996 at University of Nevada, Las Vegas, and from a series of IAPR ([1]) supported International Conferences on Document Analysis and Recognition, ICDAR, which have been held every other year since 1991. The SDAIR's were special in the sense that researchers both from document recognition and information retrieval were participated, and that a kind of merge of the two technologies were encouraged. The sources of this study also include other journals as well.

Information access functions should include not only document search by query terms, but also document categorisation, document clustering, filtering, relevance ranking, duplication identification, language identification, and so on. Some of the reviewed work targets these functions as well, trying to improve such operations even in the presence of recognition errors.

In this survey presentation, attempts will be made to show the research trends and prospects with a hope to find new concepts by merging the presented ideas and the practitioner's thoughts that could be brought from the audience.

2. Image document archiving

2.1. Ideal system concept

The early system concepts and the proposed techniques for image document archiving can be seen in the references (1-7). Figure 1 shows a conceptual block diagram of an idealistic system (5), which captures document images in full colour and decomposes the document contents into separate components of text, pictures, graphics, bibliographic information and structural information. In that system concept, further analysis is made to extract semantic information, associative links within documents, and associative links to outer worlds.

The keys of the concept are not only the full recognition of the document for full-text search, but also the smaller granularity of the accessed information. The conventional information retrieval

Hiromichi Fujisawa

Dr Hiromichi Fujisawa received the B.E., M.E. and Doctor of Engineering degrees in Electrical Engineering from Waseda University, Tokyo, in 1969, 1971, and 1975, respectively. He joined Central Research Laboratory, Hitachi, Ltd. in 1974. He has engaged in research and development work on character recognition, document understanding including mailpiece address recognition and forms processing, and document retrieval. Currently, he is a Senior Chief Researcher at Central Research Laboratory. From 1980 through 1981, he was a visiting scientist at the Computer Science Department of Carnegie Mellon University, United States. Besides working at Hitachi, he has been a visiting lecturer at Waseda University (1985 to 1997) and at Kogakuin University, Tokyo (1998–present). He is an IEEE Fellow, IAPR Fellow, and a member of ACM, etc. (*IAPR: International Association for Pattern Recognition).*

([1]) IAPR: International Association for Pattern Recognition.

usually returns whole documents as a result of search, but smaller elements of diagrams, graphs, pictures, and illustrations could be the information access targets.

In the 1980s, the research on document recognition, often referred to as 'document understanding', targeted a rather idealistic system concept and pursued better recognition performances. In the 1990s, however, a new direction emerged that encouraged approaches which are more realistic to overcome the problems, by admitting the limitation of the recognition technology. This line of work will be described in Section 4.

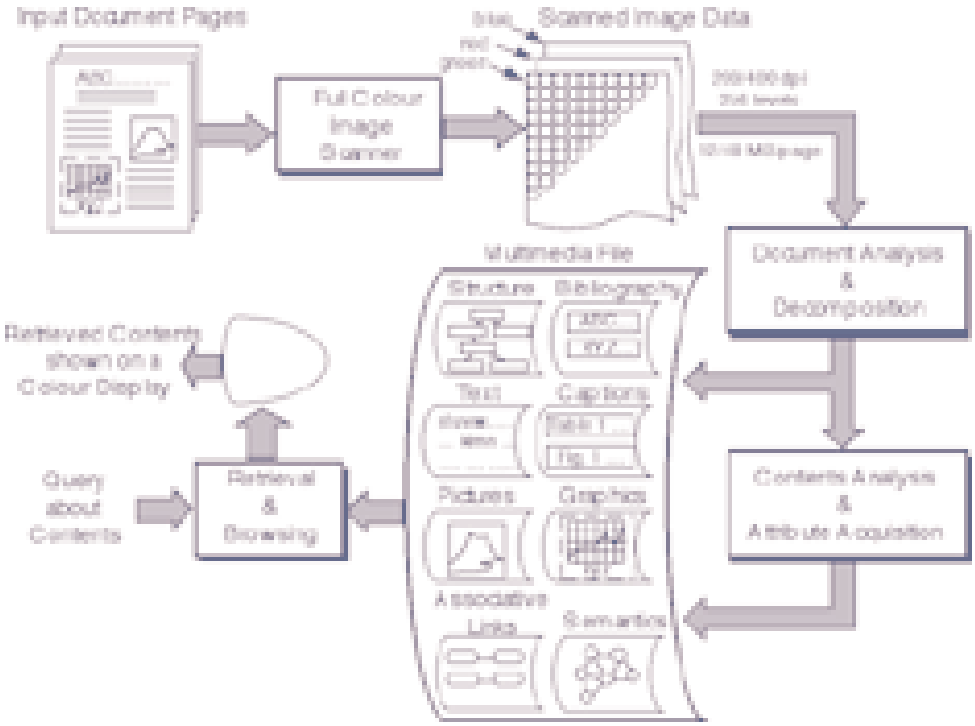


Figure 1: Multimedia contents retrieval system (5).

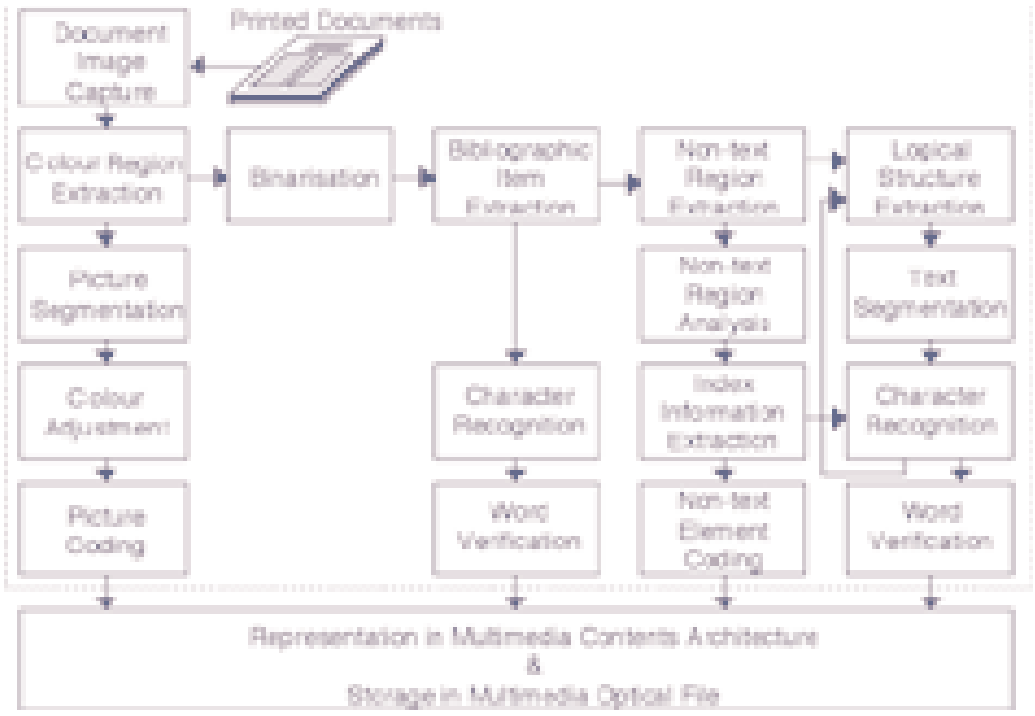


Figure 2: Document analysis and decomposition processes (5)

First, let us look at the document analysis and decomposition process shown in Figure 2. To extract information that should support and enable the access to the registered documents, the printed contents have to be analysed to produce indexes. Text information is the most reliable and effective medium for search, which can be reproduced from a binary (i.e. black and white) image by using OCR, or character recognition. In the ideal situation, pictures are first identified in the document image, and segmented as a separate medium of colour images. Other areas are binarised, from which such components as bibliographic text, non-text elements like graphics, and text bodies are extracted. This whole task is called document layout analysis, or document structure analysis.

By extracting the bibliographic items and text blocks in image forms, conventional OCR techniques can be applied to recognise characters. In Figure 2, word verification follows character recognition as a post-process. In recent techniques, integrated word recognition is applied, where the processes of character recognition and word verification are merged (or integrated) in a single process. In any case, it is hoped to detect character recognition errors and make corrections autonomously. If it works ideally, it is good for the information retrieval purpose. However, the dictionary or thesaurus used for this cannot be perfect, and this approach may have bad effects as well. Concretely speaking, new terms and unfamiliar proper names can be a source of a problem. It is ironical that such terms, which cannot be verified properly due to the dictionary imperfection, could be very good feature terms for information retrieval. For example, 'Chernobyl' must have been unknown to most of the dictionaries before 1986, but it must be a very good keyword (query term) for these kinds of documents.

Non-text elements such as diagrams, illustrations, graphs, etc. are also important and effective to answer the information demands. Some research work aims at indexing those components by extracting and recognising the text information embedded in those components (8).

2.2. Document recognition as applied to information retrieval

As mentioned, one of the biggest concerns is how to minimise recognition errors that may affect the retrieval performance. Those errors may occur at each step of document recognition described in the previous subsection. Namely, they may occur in:

(1) structural analysis

(2) segmentation

(3) character recognition.

The errors in structural analysis may discard some portions of text, confuse the reading order, and mix up bibliographic items. Segmentation errors may cause wrong character recognition. Character recognition errors may generate corrupted text. Each of these occurrences may deteriorate the retrieval performance. Although each of these error types needs to be considered, the research communities have focused mainly on the second and third types in conjunction with information retrieval.

So, in this paper, the effects of character recognition errors on the retrieval performance are first summarised. Then, three types of approaches taken in the reviewed work are presented. These are:

(1) document recognition methods adapted to information retrieval

(2) information retrieval methods adapted to document recognition

(3) image document retrieval methods without recognition.

The first approach includes improvement of recognition accuracy by applying linguistic knowledge processing, for example, or by recognising words instead of strings of characters. The second approach is to seek a robust search method for the corrupted text. The third approach is to develop new search schemes that do not rely on 'character recognition.'

2.3. Issues in capturing a huge amount of document images

While this paper will mainly deal with recognition issues, capturing a huge amount of document images reliably is also a tough task. Usually, image capturing consists of the following steps:

(1) splitting the sheets of paper from bound documents

(2) setting separated papers on the scanner

(3) capturing images by the scanner

(4) checking the quality of images: orientation, skew, displacement, binarisation, and picture quality.

(5) normalisation (if possible)

(6) rescanning in case of bad images

(7) checking the order of pages and missing pages

(8) extracting and entering index information

(9) writing the image data into a storage medium

(10) verification of the written data.

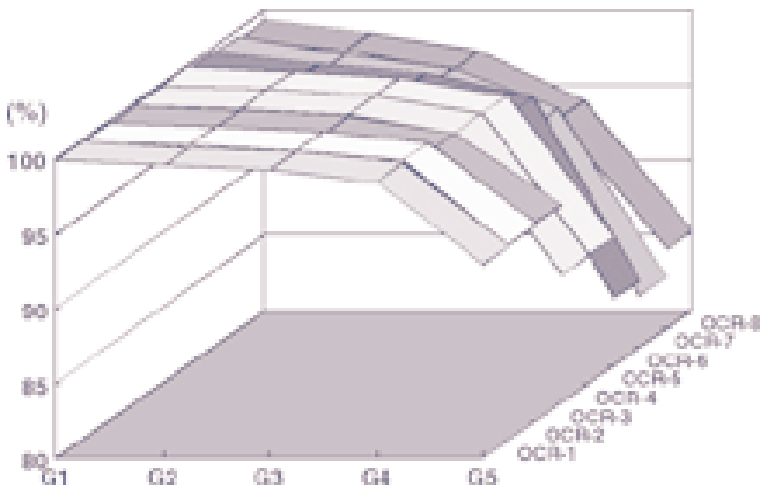
The steps (1) through (3) are mostly of mechanical processes. Issues include scanner resolution, scanning speed, mechanical throughput, paper feeder reliability, double feed protection, and limitation on paper size and thickness. The steps (4) through (7) are to guarantee the image quality. Often, the image quality check is done by human operators. Although there are algorithms for detecting and normalising such variations as orientation, skew, and displacement, it is difficult in general to detect the abnormality in the captured images automatically. Unexpected abnormality may include a folded page corner, smear, very low-quality print, loss of image due to scanner malfunction, and so on. These steps are practically very important for a huge archive construction and an archive where they want to discard physical paper documents after archiving.

The automatic operation of the step (8) is called 'automatic indexing.' The involved issues will be discussed in the following sections.

3. Effects of recognition errors

Before discussing automatic indexing, let us look at the OCR performances we can get. The study done by Information Science Research Institute (ISRI) of University of Nevada, Las Vegas (UNLV) showed very interesting characteristics of commercial OCRs (9–11, 17). When they tested OCRs from eight vendors against about 2 200 pages of printed English documents, they found that all eight OCRs attained the accuracy of better than 99.3 % for good quality documents (Figure 3). They divided the set of pages into five classes, G1 through G5, depending on their print quality (or image quality). Actually, the best three OCRs attained better than 99.9 % for this best quality

Figure 3: OCR performance characteristics (17)



group G1. The difference in recognition accuracies of different OCRs can be seen especially in low quality group G5. The best three OCRs attained the accuracy of about 93 % for G5, while others attained 83 % through 89 % accuracies as shown in Figure 3. Other statistics showed that more than a half of errors occurred in all cases were from only 20 % of the worst documents.

The tendency seen in Figure 3 is more prominent in the word recognition accuracy, which is not shown in the figure. The word recognition accuracy of the best group OCRs went down from 99.8 % (G1) to 86 % for G5. On the other hand, the accuracy of other OCRs went down from 98–99 % to 60–70 %. This shows that the problem lies in recognition of low quality documents.

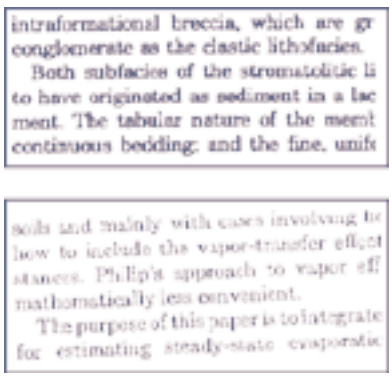
Another recognition characteristic is the effect of word length. Recognition of longer words is more difficult than shorter ones. While the best group OCR could keep the same accuracy level of word recognition for words of length 1 to 13, other OCRs gave 10-point drops for longer words. It is deemed that linguistic postprocessing contributes to this better recognition performance for the longer words.

Still another finding is regarding the relation between character recognition accuracy and word recognition accuracy. Our original expectation is that a drop in word recognition accuracy due to a drop in character recognition accuracy is much bigger when character recognition accuracy is low. Different from this expectation, the fact was that the relation was linear, and a 5-point drop in character recognition accuracy induced a 12-point drop in word recognition accuracy. This linearity means that error occurrences in character recognition are statistically dependent.

4. Review of the technical approaches to image document retrieval

4.1. Document recognition methods adapted to information retrieval

As described in the previous section, recognition errors are more prominent in low-quality document images. The errors are in terms of substitution, deletion, insertion, over-segmentation of a single character ('r n' for 'm', or 'l i' for 'h'), and merged multiple characters ('d' for 'c l', or 'm' for 'i n'). The types of errors other than substitution are caused by character-segmentation errors. In a worst-case scenario, errors also occur in word segmentation.


intraformational breccia, which are gr
conglomerate as the clastic lithofacies.
Both subfacies of the stromatolitic li
to have originated as sediment in a lac
ment. The tabular nature of the ment
continuous bedding; and the fine, unifo

soils and mainly with cases involving lo
how to include the vapor-transfer effect
stances. Philip's approach to vapor eff
mathematically less convenient.
The purpose of this paper is to integrate
for estimating steady-state evaporatio

Figure 4: Examples of low quality document images (17).

By considering these factors, studied are such approaches as word recognition by using HMM (hidden Markov models) algorithms (12: Hull, et al., pp. 174–185) (15: Elms, et al., pp. 203–216), and word recognition that does not require word segmentation (14: Myers, et al., pp. 177–188). The HMM approach was first invented for continuous speech recognition, where the tasks of word segmentation and recognition had to be done at the same time. It is now successfully used in character string recognition. It is effective not only in handwritten script writing of roman alphabets but also in low quality machine-printed characters.

Another approach is to take advantage of the multiple occurrences of the same words (12: Hull, et al., pp. 26–39) (13: Khoubyari, et al., pp. 217–232) (15: Hong, et al., pp. 177–190). Because important words appear multiple times in a document, the images of the same word can be collected in a scanned document by carrying out the clustering analysis of the segmented word images. By summing up the multiple images of the same word, we may get a word image of better quality. It is reported that the word recognition accuracy could be improved from 70 % to 92 %.

The research on document degradation model is another effort to give better recognition accuracy for low quality documents (13: Baird, et al., pp. 1–16) (15: Ho, et al., pp. 413–422). By creating a model of degraded document images, it is expected that a huge number of simulated, degraded images can be generated for the training of character recognition engines.

Character classifier combination is still another approach to get higher accuracy (27). By using multiple character recognition engines in parallel, possibly of different vendors, a big improvement can be obtained. There are several ways of combining the multiple recognition results

including probabilistic methods. One of the most popular ones is a majority voting mechanism; i.e., the character class that takes the majority of recognition results will be assigned to the unknown character class. These days, there are solution vendors who have contracts with several OCR vendors, and sell 'combination OCRs,' claiming that their solution reduces the error rate to one fifth of the conventional ones.

4.2. Information retrieval methods adapted to document recognition

An experimental study on the effect of recognition errors on retrieval performance showed the following (14: Croft, et al., pp. 115–126). In their experiment, a clean text corpus was artificially corrupted simulating recognition errors, and full-text indexing was applied without concerning the simulated recognition errors. Then, information retrieval performances were evaluated by using an Inquiry system. The findings are:

(a) when the recognition accuracy is rather high, the effect of such errors on search performance is small. More specifically, the difference in average retrieval precisions is less than 5 %;

(b) when the recognition accuracy is rather low, the effect is noticeable especially for short documents;

(c) in general, ranking is affected by recognition errors.

It is confirmed again that problems lie in low quality image documents for which recognition performance is rather low.

One of the modern information retrieval techniques is a vector space model. In the method based on VSM, words in each document are identified and the number of times the word occurs in that document is counted. By doing this, each document produces a vector whose element is a function of the number of times the corresponding term has appeared in that document. These vectors may have very high dimensionality of more than 10 000 for instance, but term selection (i.e. feature selection) is possible to reduce the dimensionality. Then, it is possible to measure similarity between two vectors mathematically. This method can be applied to document search, text categorisation, filtering, and topic identification.

A study of applying this vector space method to corrupted text showed that there is not a big difference (15: Ittner et al., pp. 301–316). A comparison was made between clean text and corrupted text that was generated by recognizing the fax-quality printed images of the same text. The recall rate and the precision for those were 51 % and 63 % for clean text, and 43 % and 66 % for corrupted text, respectively. The slightly better precision in corrupted text compared to that of clean data is considered to be due to the consistency in recognition errors. In the real life, documents are in different quality deteriorated by different reasons, and it is considered that this cannot be applied.

A research group has studied variations of the vector space method (16: Singhal, et al, pp. 149–162). The similarity measures for vectors are usually normalised to eliminate the vector length influence. Because there are several ways of doing this, they studied comparatively the normalization methods and found a better method for recognition-generated text.

Two other approaches, M-1 and M-2, are proposed as in (15: Fujisawa, et al., pp. 55–80). M-1 is to expand the query terms to include possibly recognition errors. For example, when the entered query term is 'recognition', the IR system expands it to 'recognition' and 'rccognition' because 'e' might be misrecognised as 'c' when the archived documents were indexed. The expansion process uses a so-called confusion matrix, which carries the probability of each character to be misrecognised as another. Figure 5a shows their experimental results for Japanese documents. Two curves represent the recall-precision relationships for good quality documents (Test set I) and for low quality documents (Test set II). For example, when one term is expanded to 8.2 terms on average, the recall has improved from 92 % to 96 % with a decrease of precision by six points.

The other method, M-2, is to modify both the recognition and retrieval systems. Namely, the OCR can be redesigned to output 'non-deterministic text' as a recognition result, which is a special type of text string that carries multiple recognition results for ambiguous decisions. For example, when

Figure 5a: Recall and precision for M-1 (15: Fujisawa)

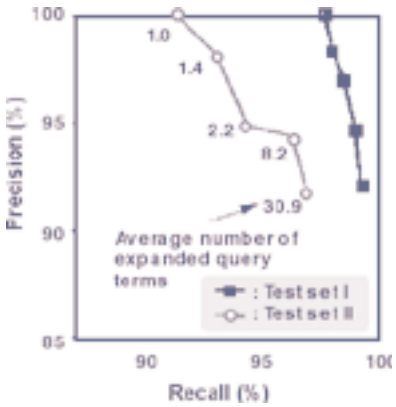
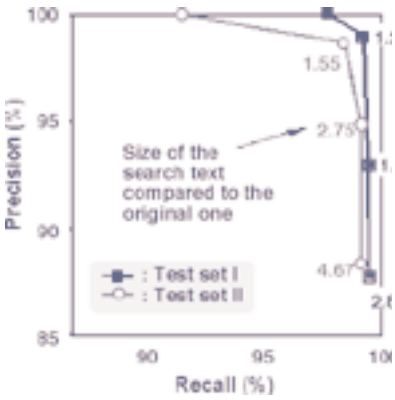


Figure 5b: Recall and precision for M-2 (15: Fujisawa)



a character pattern is ambiguous to decide whether it is 'c' or 'e,' a string '(ce)' is generated without determination. In this case, the recognition result of 'recognition' is 'r(ce)cognition'. Then, the retrieval engine may choose one of these alternatives, 'c' or 'e' when a query term is searched. Figure 5b shows again their experiments on Japanese documents. When OCR gives 55 % more characters, the recall has improved from 92 % to 98.5 %.

Work on the similar approach for English text can be seen in (21: Ohta, et al., pp. 950–956) which also uses a confusion matrix to expand query terms.

Methods to use fuzzy string matching (approximate string matching) are also in this kind of approaches (29) (21: Takasu, et al., pp. 957–961).

4.3. Methods for image document retrieval without document recognition

The last approach is that of not using OCRs. Instead of trying to recognise the characters, it is attempted to identify classes of characters. For example, characters {a, c, e, m, n, o, r, s, u, v, w, x, z} that do not have an ascender nor descender form a class 'x', and descender characters {g, p, q, y} a class 'g' and so on. This method is called character shape coding (CSC) (21: Smeaton, et al., pp. 974-978). This reduced output of a quasi-OCR is still text information, hopefully without misclassification. The similar approach for words instead of characters is called word shape coding (13: Trenkle, et al, pp. 105–122) (16: Chen, at al., pp. 163–178). The merits of this approach are a lighter processing load for transformation and small probability of erroneous coding. The latter is very important especially when low quality documents are concerned. The demerit is, however, a possibility of overloading of extracted tokens; i.e. different words may collide into the same token. It is hoped that multiple query terms may solve this problem.

Research work by Spitz on text categorisation by using this CSC shows that topic identification performance can be maintained or slightly improved using CSC derived from images (28).

In similar direction, Chinese version of CSC is studied (22: He, et al., pp. 685–688). They have devised single-byte Stroke Density Codes (SDC) representing two-dimensional features of Chinese characters. In their experiment, successful keyword spotting in Chinese image documents was shown.

5. Concluding remarks

Presented in this survey paper is the research work that can be seen in the conferences of these 10 years in the areas of document analysis and recognition. Without doubt, one of the major objectives of document analysis and recognition is the automation of document indexing. It has become quite economical to deal with digital images in computer systems. We can now capture, transfer, store, and view the image documents on computers in office and at home. It is only for

these two or three years that we can rely on images in office work. However, the final functions that need better technology are indexing and search of such documents.

Three approaches in R & D to these objectives have been presented. They are the document recognition technology appropriate for information retrieval, the information retrieval technology appropriate for document recognition, and a new technology for image document retrieval.

As we are now producing electronic documents increasingly, development of better technologies for image document archiving and retrieval needs to be hastened.

Acknowledgement

The author would like to express his gratefulness to Professor Tom Nartker of University of Nevada, Las Vegas, for his kind support for this study.

References

(1) Wong, K. W., Casey, R. G. and Wahl, F. H., 'Document analysis system,' *IBM Journal of Research and Development*, Vol. 26, No 6, 1982, pp. 647–656.

(2) Nagy, G., 'Hierarchical representation of optical scanned documents,' *7th Int. Conf. Pattern Recognition*, 1984, pp. 347–349.

(3) Nagy, G., Seth, S., and Stoddard, S. D., 'Document analysis with an expert system,' *Pattern recognition in practice II* (Gelsema, E. S. and Kanal, L. N., Eds.), 1986, pp.149–159.

(4) Higashino, J., Fujisawa, H., Nakano, Y., and Ejiri, M., 'A knowledge-based segmentation method for document understanding,' *8th International Conference Pattern Recognition*, 1986, pp. 745–748.

(5) Fujisawa, H., Yashiro, H. et al., 'Document analysis and decomposition method for multimedia contents retrieval,' *2nd Int. Symp. Interoperable Information Systems*, ISIIS-88, Tokyo, 1988, pp. 231–238.

(6) Dengel, A. and Barth, G., 'Anatasil: A hybrid knowledge-based system for document layout analysis,' *Int. Joint Conference on Artificial Intelligence*, 1989, pp. 1 249–1 254.

(7) Fujisawa, H., Shima, Y., Koga, M., and Murakami, T., 'Automatically organising document bases using document understanding techniques,' *Second Far-East workshop on the future database systems* (Ed. Q. Chen), Kyoto, 1992, pp. 244–253.

(8) Koga, M., Murakami, T., Shima, Y., and Fujisawa, H., 'A structured analysis method of graph image for document-image retrieval,' SPIE, Vol. 1906, Character recognition technologies, San Jose, CA, 1993, pp. 291–295.

(9) UNLV: Information Science Research Institute 1993 Annual Report, April 1993.

(10) ISRI 1994 Annual Report, April 1994.

(11) ISRI 1995 Annual Report, April 1995.

(12) *Proc. symposium on document analysis and information retrieval* (SDAIR), Las Vegas, 16–18 March, 1992.

(13) *Proc. Second SDAIR*, Las Vegas, 26–28 April, 1993.

(14) *Proc. Third SDAIR*, Las Vegas, 11–13 April, 1994.

(15) *Proc. Fourth SDAIR*, Las Vegas, 24–26 April, 1995.

(16) *Proc. Fifth SDAIR*, Las Vegas, 15–17 April, 1996.

(17) Rice, S. V., Jenkins, F. R., and Nartker, T., 'The fifth annual test of OCR accuracy,' *Information Science Research Institute Technical Report*, TR-96-01, 1996.

(18) *Proedural first international conference on document analysis and recognition* (ICDAR), Saint-Malo, France, 1991.

(19) *Proc. Second ICDAR*, Tsukuba, Japan, 1993.

(20) *Proc. Third ICDAR*, Montreal, Canada, 1995.

(21) *Proc. Fourth ICDAR*, Ulm, Germany, 1997.

(22) *Proc. Fifth ICDAR*, Bangalore, India, 1999.

(23) *Proc. Sixth ICDAR*, Seattle, United States, 2001.

(24) Doermann, D., 'The retrieval of document images: a brief survey,' *Prof. fourth international conference, document analysis and recognition*, Ulm, Germany, 1997, pp. 945–949.

(25) Nagy, G., 'Twenty years of document image analysis in PAMI,' *IEEE Trans. pattern analysis and machine intelligence*, Vol. 22, No 1, 2000, pp. 38–62.

(26) Taghva, K., Nartker, T., Borsack, J. and Condit, A., 'UNLV-ISRI Document Collection for Research in OCR and Information Retrieval,' *Proc. IS&T/SPIE conference document recognition and retrieval VII*, San Jose, 2000, pp. 157–164.

(27) Xu, L., Krzyzak, A., and Suen, C. Y., 'Methods of combining multiple classifiers and their applications to handwriting recognition' *IEEE trans. systems, man, and cybernetics*, Vol. 22, No 3, 1992, pp. 418–435.

(28) Spitz, A. L. and Maghbouleh, A., 'Text categorisation using character shape codes,' *Proc. IS&T/SPIE conf. document recognition and retrieval VII*, San Jose, 2000, pp. 174–181.

(29) Myka, A. and Guntzer, U., 'Fuzzy full-text searches in OCR database,' *Proc. forum on research and technology advances in digital libraries,* 1995, pp. 87–100.

Revisión de los progresos realizados en las tecnologías de reconocimiento y de investigación documental

Hiromichi Fujisawa

Nuestro propósito es presentar un análisis de las tecnologías de reconocimiento e investigación documental, que han sido objeto de trabajos de investigación y desarrollo en estos diez últimos años. Desde hace tiempo se supone que las técnicas de reconocimiento óptico de caracteres y de reconocimiento documental facilitan la indexación automática y la consulta en texto íntegro de documentos de imagen. Gracias a los recientes progresos realizados en el reconocimiento documental y en la informática, algunas de estas ideas ya se han puesto en práctica. No obstante, los resultados del reconocimiento documental nunca pueden ser perfectos, y siempre se están estudiando nuevas tecnologías. El presente documento examina los recientes trabajos en este ámbito.

Los trabajos analizados son el fruto de varios simposios sobre el análisis de documentos y la búsqueda de información, que se han organizado todos los años entre 1992 y 1996 en la Universidad de Nevada (Las Vegas), así como de una serie de conferencias internacionales sobre el mismo tema, que se celebran todos los años desde 1991. Los primeros presentaban una particularidad, en el sentido de que participaron investigadores de los ámbitos del reconocimiento de documentos y de la recuperación de información, fomentando así una especie de fusión entre las dos tecnologías.

Una de las cuestiones principales es minimizar los errores que se producen en el reconocimiento óptico de caracteres (OCR), cuando el texto generado por el OCR se utiliza a efectos de la investigación documental. En primer lugar, se presenta brevemente el impacto de los errores de reconocimiento en los resultados de investigación. A continuación se presentan tres enfoques del problema: 1) buscar una mayor precisión de reconocimiento, por ejemplo aplicando las técnicas de tratamiento lingüístico de los conocimientos; 2) desarrollar nuevas técnicas de recuperación que minimicen el impacto de los errores de reconocimiento óptico, por ejemplo aplicando métodos de correspondencia flexibles; 3) desarrollar un sistema enteramente nuevo que integre los dos procesos de reconocimiento y recuperación. En otros términos, este último enfoque consiste en modificar los procesos de reconocimiento y de recuperación. Así pues, el reconocimiento óptico de caracteres puede modificarse para producir un texto «no determinista», como resultado del reconocimiento, que es en realidad una cadena de texto que incluye varios resultados posibles para decisiones ambiguas. De esta manera, cada vez que sea ambiguo elegir entre una «e» o una «c», se genera una cadena de caracteres «[e c]» sin que se tome una decisión. Posteriormente, el motor de investigación podrá elegir una de las alternativas cuando se solicite un término de búsqueda. Otro enfoque consiste en no utilizar el reconocimiento óptico de los caracteres. Así pues, en vez de intentar reconocer los caracteres, se intenta definir una serie. Por ejemplo: los caracteres {a, m, o, r, s, ...}, que no son letras ascendentes ni descendentes, están representados por la clase «x», los caracteres {g, j, p, q, y} están representados por la clase «j», etc. Este resultado parcial de un cuasi reconocimiento óptico puede explotarse para la recuperación de información.

La investigación documental no sólo plantea problemas de términos de búsqueda, sino que también presenta otras dificultades: clasificación y agrupación de los documentos, clasificación por importancia, identificación de duplicados, identificación de la lengua, etc. Cuando se trata de procesar imágenes, las cosas se complican más aún. Algunas labores de investigación en materia de reconocimiento documental examinan también estas cuestiones, esforzándose en mejorar las operaciones, incluso cuando se trata de imágenes.

Aunque las técnicas de investigación pueden adaptarse a los documentos de imagen, sigue siendo importante tratar de obtener mejores resultados para el reconocimiento. También se examinan los recientes progresos en las técnicas de reconocimiento. Asimismo, se hablará también del taller de la IAPR (International Association for Pattern Recognition) sobre los sistemas de análisis documental. Entre los recientes progresos figuran: la segmentación avanzada de caracteres a partir de los modelos HMM (Hidden Markov Models) o de la información lingüística, una mejor aplicación de los conocimientos lingüísticos para solucionar las ambigüedades, la utilización de una combinación de clasificadores de caracteres múltiples para mejorar la precisión del reconocimiento, un análisis avanzado de la paginación, etc. La captura de imágenes y su pretratamiento, tal como la binarización y la extracción de modelos de caracteres de los documentos en color y/o a partir de un contexto complejo, son también cuestiones importantes.

En nuestra presentación nos esforzaremos por establecer una correspondencia entre los distintos enfoques y las demandas reales, que serán formuladas por el público, y por progresar en el plano tecnológico.

Übersicht über die Fortschritte bei Technologien der Dokumentenerkennung und des Dokumentenretrieval

Hiromichi Fujisawa

Es soll ein Überblick über Technologien der Dokumentenerkennung und des Dokumentenretrieval im Bereich der FuE in den letzten zehn Jahren vermittelt werden. Mit den OCR- und Dokumentenerkennungstechnologien verbindet sich seit langem die Erwartung, dass sie das automatische Indexieren und die Volltextsuche von Bilddokumenten erleichtern. Dank der jüngsten Fortschritte bei den Dokumentenerkennungs- und Computertechnologien werden jetzt einige Vorstellungen in die Praxis umgesetzt. Die Dokumentenerkennung kann jedoch niemals vollkommen sein, und neue Technologien befinden sich noch im Forschungsstadium. In diesem Beitrag werden die jüngsten Forschungsarbeiten auf diesem Fachgebiet beleuchtet.

Grundlage für die untersuchten Arbeiten waren eine Reihe von Symposien zur Dokumentenanalyse und zum Informationstretrieval, die von 1992 bis 1996 an der University of Nevada in Las Vegas veranstaltet wurden, sowie eine Reihe internationaler Konferenzen zur Dokumentenanalyse und -erkennung, die seit 1991 im Zweijahresturnus stattfinden. Das Besondere bei den Symposien bestand darin, dass Wissenschaftler aus den Bereichen Dokumentenerkennung und Informationsretrieval eingeladen waren und eine gewisse Verschmelzung dieser beiden Technologien angestrebt wurde.

Eines der größten Probleme ist die Minimierung der Fehler, die bei der OCR entstehen, wenn der mit der OCR erzeugte Text für das Wiederauffinden von Dokumenten benutzt wird. Zu Beginn werden die Folgen der OCR-Fehler für die Retrievalergebnisse zusammengefasst.

Anschließend werden drei Lösungsansätze für dieses Problem vorgestellt: 1. die Bemühungen um eine noch bessere Erkennungsgenauigkeit, z. B. Anwendung der Sprachkenntnisverarbeitung, 2. die Entwicklung neuer Retrievaltechniken, die die Folgen von OCR-Fehlern z. B. durch flexiblen Abgleich reduzieren, und 3. die Entwicklung eines völlig neuen Systems, das in der Verknüpfung der beiden Prozesse Erkennung und Retrieval besteht. Bei dem letztgenannten Ansatz geht es mit anderen Worten darum, sowohl die Erkennungs- als auch die Retrievalprozesse zu modifizieren. So lässt sich z. B. die OCR so modifizieren, dass im Ergebnis der Erkennung ein „nichtdeterministischer" Text ausgegeben wird, also eine besondere Art von Zeichenkette, bei der Zweifelsfälle mit einem Mehrfach-Erkennungsergebnis versehen sind. Ist beispielsweise unklar, ob es sich um ein „e" oder ein „c" handelt, wird eine Kette „[e c]" ohne Festlegung erzeugt. Dann kann die Retrievalmaschine bei der Suche nach einem Abfragebegriff eine der beiden Alternativen wählen. Ein anderes interessantes Konzept ist die Nichtverwendung von OCR. Statt zu versuchen, die Zeichen zu erkennen, wird versucht, eine Klasse von Zeichen zu identifizieren, z. B. die Zeichen {a, m, o, r, s,...}, die weder Ober- noch Unterlänge haben, werden als Klasse „x" dargestellt, die Zeichen {g, j, p, q, y} werden durch die Klasse „j" dargestellt usw. Diese reduzierte Ausgabe einer Quasi-OCR kann für das Retrieval benutzt werden.

Probleme des Dokumentenretrieval betreffen nicht nur die Suche mit Abfragebegriffen, sondern auch die Dokumenteneinordnung, die Dokumentenbündelung, die Auflistung nach Relevanz, die Identifizierung von Doppelungen, die Sprachenerkennung usw. Kompliziert wird es, wenn Bilddokumente verarbeitet werden sollen. In einigen der Forschungsarbeiten zur Dokumentenerkennung werden auch diese Fragen thematisiert, und es wird versucht, diese Vorgänge auch bei Bilddokumenten zu verbessern.

Obwohl die Retrievaltechniken an Bilddokumente angepasst werden können, ist es doch wichtig, bessere Erkennungsergebnisse anzustreben. Betrachtet werden auch jüngste Fortschritte bei den Dokumentenerkennungsverfahren. Eine weitere Quelle für diese Untersuchung wird der IAPR-Workshop zu Dokumentenanalysesystemen (DAS) sein. Zu den aktuellen Ergebnissen zählt die fortgeschrittene Zeichensegmentierung auf der Basis von HMM (Hidden-Markov-Modelle) oder linguistischen Informationen, eine bessere Nutzung von Sprachkenntnissen zur Lösung von Zweifelsfällen, die Anwendung einer Kombination von Mehrfachzeichenklassifikatoren zur Verbesserung der Erkennungsgenauigkeit, eine fortgeschrittene Layoutanalyse u. a. m. Weitere wichtige Fragen sind die Bilderfassung und -vorverarbeitung wie Binarisierung und Zeichenstrukturextraktion von Farbdokumenten und/oder einem komplexen Hintergrund.

Bei der Präsentation dieser Übersicht wird der Versuch unternommen, die unterschiedlichen vorgestellten Ansätze an den realen Anforderungen in den Anwendungsbereichen zu messen, die von den Zuhörern vertreten werden, und nach weiteren technologischen Lösungen zu suchen.

Tableau des progrès réalisés dans les technologies de reconnaissance et de recherche documentaires

Hiromichi Fujisawa

Notre objet est de présenter une analyse des technologies de reconnaissance et de recherche documentaires, qui font l'objet de travaux de recherche et de développement depuis ces dix dernières années. Les techniques de reconnaissance optique des caractères et de reconnaissance documentaire sont supposées depuis longtemps faciliter l'indexation automatique

et la consultation en texte intégral de documents vidéo. Grâce aux progrès récents réalisés dans la reconnaissance documentaire et dans l'informatique, certaines de ces idées sont désormais en pratique. Toutefois, les résultats de la reconnaissance documentaire ne peuvent jamais être parfaits et de nouvelles technologies sont toujours en cours d'étude. Le présent exposé passe en revue les travaux récents dans ce domaine.

Les travaux sous revue sont le fruit de plusieurs symposiums sur l'analyse et la recherche documentaires qui ont été organisés tous les ans entre 1992 et 1996 à l'université du Nevada, à Las Vegas, ainsi que d'une série de conférences internationales sur le même thème qui se tiennent tous les ans depuis 1991. Les premiers présentaient une particularité, en ce sens que des chercheurs travaillant aussi bien dans la reconnaissance documentaire que dans la recherche documentaire y participaient, encourageant ainsi une sorte de fusion entre les deux technologies.

L'un des principaux problèmes est de minimiser les erreurs intervenant dans la reconnaissance optique des caractères, lorsque le texte généré par reconnaissance est utilisé aux fins de la recherche documentaire. L'impact des erreurs de reconnaissance sur les résultats de recherche est tout d'abord présenté brièvement. Puis trois approches du problème sont présentées: 1) rechercher une précision de reconnaissance encore meilleure, par exemple en appliquant les techniques de traitement linguistique des connaissances; 2) développer de nouvelles techniques de recherche qui minimisent l'impact des erreurs de reconnaissance optique, par exemple en appliquant des méthodes de correspondance souples; 3) développer un système entièrement nouveau qui soit une intégration des deux processus de reconnaissance et de recherche. En d'autres termes, cette dernière approche doit modifier à la fois le processus de reconnaissance et celui de la recherche. Ainsi, la reconnaissance optique des caractères peut être modifiée pour produire un texte «non déterministe», en tant que résultat de la reconnaissance, qui est en fait une chaîne textuelle comportant plusieurs résultats possibles pour les points ambigus. Ainsi, chaque fois qu'il est ambigu de trancher entre un «e» ou un «c», une chaîne de caractères «[e c]» est générée sans qu'une décision soit prise. Le moteur de recherche peut ensuite opérer un choix lorsqu'un terme de recherche fait l'objet d'une requête. Une autre approche consiste à ne pas utiliser la reconnaissance optique des caractères. Ainsi, au lieu d'essayer de reconnaître les caractères, on essaie d'en identifier une série. Exemple: les caractères {a, m, o, r, s...}, qui ne sont ni des lettres montantes ni des lettres ascendantes, sont représentés par la classe «x», les caractères {g, j, p, q, y} sont représentés par la classe «j», etc. Ce résultat partiel d'une quasi-reconnaissance optique peut être exploité pour la recherche.

La recherche documentaire pose des problèmes de termes de requête, mais présente aussi d'autres difficultés: classement et groupement des documents, classement par pertinence, identification des doublons, identification du langage, etc. Lorsqu'il s'agit de traiter des images, les choses se compliquent quelque peu encore. Certains travaux de recherche en reconnaissance documentaire se penchent également sur ces questions, en s'efforçant d'améliorer les opérations même lorsqu'il s'agit d'images.

Même si les techniques de recherche peuvent être adaptées aux documents vidéo, il est tout de même important de viser l'objectif de meilleurs résultats pour la reconnaissance. Les progrès récents dans les techniques de reconnaissance sont également passés en revue. Il sera aussi question de l'atelier de l'IAPR sur les systèmes d'analyse documentaire. Les avancées récentes sont: la segmentation avancée des caractères à partir des modèles HMM (Hidden Markov Models) ou de l'information linguistique, une meilleure application des connaissances linguistiques pour résoudre les ambiguïtés, l'utilisation d'une combinaison de classificateurs à caractères multiples pour améliorer la précision de la reconnaissance, une analyse avancée de la mise en page, etc. La saisie des images et le prétraitement, comme la binarisation et l'extraction des modèles de caractères sur les documents en couleurs et/ou à partir d'un contexte complexe, sont également des questions importantes.

Nous nous efforcerons dans notre présentation de mettre en correspondance les diverses approches et les demandes réelles du terrain qui seront formulées par le public, et de progresser sur le plan technologique.

Implementing an electronic document and records management system at the UNHCR and the IOC

Maria Montserrat Canela Garayoa, Ineke Deserno, Marsha Gillis

Introduction

The United Nations High Commissioner for Refugees (UNHCR) is mandated by the United Nations General Assembly to lead international action to protect refugees worldwide. The UNHCR's Headquarters are in Geneva, Switzerland, and they have 250 field offices in 120 countries.

The International Olympic Committee (IOC) is an international non-profit organisation and creator of the Olympic Movement, for which it serves as an umbrella organisation. Its primary responsibility is to supervise the organisation of the summer and winter Olympic Games. The IOC's Headquarters are in Lausanne, Switzerland.

As recently as 10 years ago, the UNHCR and the IOC conducted most of their operations strictly with traditional office tools (typewriters, copiers, phones and faxes); by the end of the 1990s both had evolved from a predominantly paper environment to a predominantly electronic one. This evolution forced both organisations to grapple with new challenges for managing and preserving electronic information.

The need for an integrated system

While the new technology benefited document creators, it created new complexities for Records and Archives staff, who now faced the double challenge of managing both physical paper documents and their electronic counterparts. In the past five years both organisations realised that they required an integrated system to manage all documents that are either created electronically in-house and those received electronically from external sources.

E-mail: a key catalyst for EDRMS

E-mail and other office productivity tools such as word processors have clearly had a positive impact on the way we work. But e-mail — unlike other tools that essentially replaced existing mechanical technology (typewriters, calculators, etc) — has had a profound impact on organisational culture at the IOC and the UNHCR. Staff now depend on e-mail as their primary communication tool, using it to transmit and receive important knowledge.

The rise in e-mail use and the need to preserve knowledge within e-mail were two strong catalysts for records managers at the UNHCR and the IOC to choose EDRMS. Records managers also saw in EDRMS an opportunity to integrate existing e-mail procedures into a larger framework of document and record management practices.

EDRMS objectives and requirements

Both the UNHCR and the IOC initiated their electronic document and records management system (EDRMS) projects with similar objectives and requirements. They wanted a system that would:

- help the organisation manage records, framed within a comprehensive file plan associated with retention schedules;
- help users store and retrieve documents from their desktop, with minimal disruption to their daily activities;
- automatically capture metadata;
- integrate with existing e-mail applications to help users preserve e-mail records;
- integrate all of the above components into one unified system;

Montserrat Canela

Montserrat Canela has a degree in history and an MA in archives. She joined the UNHCR Archives and Records Centre in 2001, where she has a post of Associate Archivist, responsible for the records management part of the section. She has served for 10 years as director of the of Arxiu Central del Departament de Cultura de la Generalitat de Catalunya (Spain) and five years as director of Arxiu Historic Comarcal de Cervera (Spain). Before joining the UNHCR she was member of a team based in Barcelona that developed records management systems for several public and private institutions and that participated in various training sessions and published articles based on its professional experience.

Ineke Deserno

Ineke Deserno holds an MA in history and archives. She did an advanced study program at the School for Archives, Information and Library Studies at the University of British Columbia, Vancouver. She joined the UNHCR Archives and Records Centre in 1996. She is currently the Electronic Records Archivist and responsible for the coordination of the implementation of the integrated Electronic Document and Records Management System. From 1996 to the end of 1999 she was the Associate Archivist. Prior to the UNHCR she worked for a Dutch consultancy firm specialised in records management and archives. In March 2002 she started to work for the International Olympic Committee, where she coordinates the implementation of the electronic document and records management system.

Marsha Gillis

Marsha Gillis works at the UNHCR.

Coincidentally the two organisations selected the same tools to fulfil these requirements: Open Text's Livelink for collaborative document management and iRIMS for records management. Our personal involvement with the EDRMS projects at the UNHCR and the IOC provided a unique opportunity to review the implementation process and discuss the lessons we learned from two separate, yet related, perspectives. The UNHCR initiated their project in 1998 and is now (spring 2002) nearing the end of its headquarters implementation phase, whereas the IOC is now in the project initiation phase.

This paper discusses the implementation of EDRMS at both organisations, including:

- functional requirements of the system and the underlying principles that fed them;
- implementation: how the system was developed and introduced;
- challenges the project team faced during the implementation;
- conclusions: discussion of critical success factors, best practices and lessons learned.

Functional requirements

The most important requirement for both the UNHCR and the IOC was to integrate closely the document management and records management components of the new system. For both components, the requirements were very similar for the two organisations.

Document management

System usability was a critical factor underlying all document management requirements. The archives and records specialists knew that a new system would fail if the average user could not easily file and retrieve their documents. The need to develop a friendly, usable system translated into the following document management requirements:

- automation of procedures for capturing information, including metadata. The objective was to avoid burdening staff with additional filing work;
- parallel or complementary functionality for managing paper documents such as incoming faxes/mail and electronic documents such as e-mail or spreadsheets;
- collaboration and sharing tools;
- intuitive full-text search of content and attributes.

Records management

The records management component gives structure to the information stored in the central repository. In the initial project phase the following were identified as crucial record keeping requirements:

- standard file plan — An essential records management tool to ensure that records relating to one subject are grouped together in the same folder. The file plan expresses the relationship of the records and makes them available to users;
- application of disposition rules for records in the repository — by associating the organisation's retention schedules with the file plan, the retention process can be automated more easily. Established retention periods would determine when the records and associated metadata would be deleted from the central repository.

Both organisations had also identified remote (offsite) access to the document repository as a critical business requirement, due to the international scope of their mandates.

EDRMS implementation: a brief summary

To fulfil the functional requirements for EDRMS both the IOC and the UNHCR selected Livelink (document management) and iRIMS (records management). The implementation activities and

issues specific to each organisation will be presented at the conference as two case studies. However, it is possible to summarise the implementation of EDRMS by highlighting the similarities for both organisations.

Integration and usage of tools

The tool selection and degree of software integration was virtually identical for both the IOC and the UNHCR. Livelink was implemented as the organisation-wide document management tool. The implementation of Livelink Desktop ensured that document capture procedures were straightforward for users. Staff use Livelink Desktop to store and retrieve documents directly from their MS Office applications. The e-mail application was also integrated with Livelink so that users could store e-mails directly into the document repository.

Although the records management tool, iRIMS, was also integrated with Livelink, the use of iRIMS was limited to records and archives staff. They have primary responsibility for maintaining the file classification plan, creating folders, scanning paper documents, registering document profile information, and maintaining the index to paper files.

Conversion of paper documents into electronic files

Scanning of incoming paper mail is a key component of the EDRMS implementation. Faxes, however, are received directly into the repository via the fax server. Both types of incoming documents are immediately made available to the intended recipients.

Communication and training during implementation

During the early stages of the implementation the project team communicated information about the project to staff. The communication was established through an awareness campaign to promote the new tools and give staff an opportunity to ask questions about the system and its impact on their work.

Before the classroom training began, the training materials provided by the software vendor were customised to reflect the organisation-specific features of the applications. Then each staff member attended a classroom training session to learn the basic document management tasks. As a follow-up, staff were encouraged to seek individual training if necessary, and had access to a dedicated EDRMS support person.

Challenges

Considering that the UNHCR and the IOC are, in the first quarter of 2002, at different phases of the project life cycle, we must first discuss the real challenges faced by the UNHCR and then address the challenges we anticipate for the project at the IOC.

1. Organisational challenges

Working culture

The UNHCR is an emergency-based organisation with a high level of staff rotation, and these characteristics make knowledge transfer very difficult.

The IOC is a planning-based organisation that works on strict deadlines due to the nature of the Olympic Games. In repeating many of the same activities from one Games to the next, it is very important for staff to transfer past knowledge each time.

Staff at both organisations travel extensively and have a strong need to access information remotely.

Differing EDRMS requirements

In the requirements gathering phase of the UNHCR project, the business analysts identified considerable departmental variations in the functional requirements for EDRMS. For example, the

departments responsible for the regional refugee operations conduct considerable business via e-mail and experience high staff turnover due to the organisation's rotation policy.

Knowledge preservation, e-mail integration, version control and centralised subject files were therefore key requirements for EDRMS. In administrative departments (e.g. Human Resources) the document creation, storage and access requirements were significantly different as the primary work involves paper personnel files.

Communication between departments

The EDRMS project at the UNHCR included participants from two areas: the Records and Archives department and the Information Technology and Telecommunications department. Communication and collaboration between the two departments was an ongoing challenge throughout the project.

This was partly due to the division of responsibilities for the project: the IT department handled technical and contractual issues (hardware purchasing, installation and maintenance; plus vendor contract negotiations); and the Records and Archives department was responsible for all functional and practical issues (preparing functional specifications, user interface design, file plans, implementation scheduling, etc.). The former viewed EDRMS as lower priority than some of their other projects, whereas the Records and Archives department considered EDRMS to be a mission-critical system.

Some of these issues will be less pronounced at the IOC, which is a much smaller organisation and has dedicated resources from the Information Technology department working on the EDRMS project.

2. Records management challenges

Competing repositories of information

The existence of multiple information repositories at the UNHCR presented the EDRMS team with a challenging question: how to make the new system a complement to the others, rather than a competitor? At the UNHCR, the various repositories include: intranet; other electronic reference sources; local drives; e-mail inboxes; and paper repositories. Some of these sources contained some of the same information as the document management repository, which initially caused confusion for users.

The IOC may be able to avoid some of these issues because it has fewer repositories in general and plans to limit access to local drives from the beginning of the implementation.

Standard file plan

A standard, organisation-wide file plan was the foundation for the EDRMS implementations at both the IOC and the UNHCR. Prior to EDRMS the file plan existed solely for central paper files and was used primarily by archives. Staff did not typically follow the file plan, which was long and unwieldy. The challenge with EDRMS was to make the file plan work in an electronic environment while ensuring adequate flexibility to meet the needs of all departments.

Adjusting existing procedures to EDRMS

Users tend to rely on many paper-based procedures for their daily work, even if they have more efficient electronic means at their disposal. The EDRMS teams wanted to encourage staff to stop relying on paper by providing equivalents in electronic format.

It is virtually impossible to ask all users to follow identical procedures for using e-mail and desktop tools. Each person has their own communication style and each person integrates electronic tools into their work differently.

E-mail filing and retrieval

Because of its inherently informal style and tone, an e-mail message is difficult to file and retrieve. For example, a user may receive in one day 15 messages with very similar subject headings, but

must assess the content of each e-mail individually to determine which folder in the standard file plan is most appropriate. The subject heading used by the sender likely has little relationship to the organisation's file plan, making the task more difficult.

E-mail duplication

It is difficult to monitor the degree to which staff will file one e-mail in numerous locations. The duplication typically occurs when one external sender sends the same message to many people within the organisation. The receivers may then create individual replies to messages and soon it becomes very difficult to trace the thread and ensure that all 'discussions' that took place via e-mail messages and replies are captured in the knowledge repository.

E-mail volume

The sheer number of e-mail messages/attachments sent and received in one day is staggering. Of course not every message needs to be filed in a subject folder, but deciding what to file is daunting when users are already feeling overwhelmed by the volume of messages they receive and send each day.

3. Technical challenges

Lack of product integration

Soon after selecting the software tools, the UNHCR discovered that Livelink and iRIMS were, at that time, not fully integrated products and they did not meet the technical and functional requirements as expected. The UNHCR was one of the first organisations to implement both these tools together and as such, the team was not able to benefit from experience the vendor might have gained through extensive 'field testing' elsewhere.

The IOC will clearly be in a different position when they begin the full implementation of Livelink and iRIMS as integrated products. Both the vendor and the project team leaders have worked with the integration in a real environment and will therefore be able to tackle technical and functional problems much more quickly.

4. User acceptance challenges

User perceptions prior to EDRMS

The shift from paper to electronic document creation caused a shift in staff attitudes and perceptions. As information becomes electronic it becomes less 'visible'. One can easily visualise a room filled with 100 boxes, each containing 10 paper documents, but it is more difficult to imagine the 'virtual' space occupied by 100 directories on a network server, each containing 10 electronic documents. Users at both the UNHCR and the IOC tend to share two key perceptions that could hinder adoption/acceptance of EDRMS:

- electronic files are perceived to be more difficult to retrieve than paper files. Not only do users see retrieval as a personally challenging task (e.g. retrieving their own work from a potentially confusing set of shared directories), but they also doubt the ability of the organisation to preserve, manage and retrieve electronic documents (e.g. retrieving old files from last month's tape backup);
- filing or archival tasks are viewed as low-level tasks that should be the sole responsibility of secretarial staff.

Training

The level of computer literacy varies greatly at the UNHCR and so the training provided was not always sufficient to meet the needs of very basic users. For more skilled users — who easily grasped the new concepts and technology during classroom sessions — the challenge was to bridge the transitional gap from skills learned in class to skills applied on the job.

5. Future considerations

Now that the UNHCR has nearly completed the EDRMS implementation, the focus turns to maintaining the systems and identifying future challenges. We have identified several key challenges that the UNHCR will need to address as EDRMS evolves, and these issues will undoubtedly be of concern to the IOC as well:

- long-term preservation of information;
- eventual obsolescence of the current technology and possible difficulties for users to work with documents in the repository (e.g. file formats that do not remain compatible with future upgrades of software);
- system enhancements — the existing level of customisations in EDRMS will cause additional work when the organisation upgrades their operating system and e-mail software;
- integration with other records-generating applications such as financial, supply and human resources databases.

Success factors and lessons learned

Deploying an integrated document and records management system has been a valuable learning experience for the project team and the organisation. In this section we identify some of the factors that were critical to the success of the project. We also share the lessons learned and translate them into best practices for others embarking on similar projects.

1. Records management

Avoid competition among information repositories

All stakeholders must work together to delimit the scope and mandate of the different information repositories used within the organisation. The following questions must be clearly answered:

- What are the inconsistencies from one tool or repository to another?
- What are the policies for their use?
- What types of information are stored?
- How do the different repositories relate to each other?
- Where should links be made, if none already exist?

Classify documents appropriately

Each separate work unit at the UNHCR had their own workspace in the repository, a space that was set up initially by the records and archives staff according to the organisational file plan. The goal was to achieve a level of 'controlled flexibility' that would meet the needs of the individual departments while at the same time keeping overall management of the records in the hands of the specialists. Initially users tend to resist any imposition of standardised subject folders coming from a central 'authority'. Therefore to overcome this resistance it was necessary to allow each department to control their files at lower levels within the central filing framework.

Early in the EDRMS implementation phase the records managers determined that they had made the right decision to give specialised archives staff the sole responsibility for the classification component. As a result the users' existing filing procedures did not change significantly — their main responsibility is to choose subject folders when filing.

With a high volume of documents and records, and limited resources, it became very important for classification to be done on the folder level. This was another key factor in ensuring ease-of-use for the UNHCR and IOC staff. In both cases the records managers assign the classification file codes at the folder level, so that subfolders and documents added to a folder always inherit the

folder file code. Records schedules are applied according to file code, so that all folders, subfolders and documents that share the same classification code will have the same schedule.

Capture metadata carefully

To ensure the accuracy and usefulness of metadata it is crucial to capture or record the metadata information very carefully, keeping in mind any requirements for future retrieval. Specifically, users must be able to easily retrieve information without needing special training or knowledge of the original metadata capture process.

Encourage appropriate use of e-mail

E-mail creation and filing practices are very difficult to control in a large organisation without sophisticated software tools. As the EDRMS project at the UNHCR evolves, the organisation envisions better e-mail management through enhanced system functionality. For example, it may eventually be possible to prompt users to code each e-mail they send or receive as 'official' to ensure that it is captured in the document repository.

Without such controls the organisation must encourage users to make appropriate use of e-mail, by:

- providing clear guidelines for creating and filing e-mails;
- educating users about costs incurred from inappropriate use of e-mail;
- promoting the use of more cost-efficient collaboration or communication tools, such as the Livelink discussions function, which stores a message once so users don't have to send it to many colleagues.

2. Technical

The technical and hardware requirements for EDRMS must be included in the overall information technology strategy from the early stages of the project. The critical technology must be installed, configured, and rigorously tested prior to the rollout of EDRMS.

When the UNHCR first faced product integration challenges, it did not have many options. Few other organisations had, at that point, experience with integrating Livelink and iRIMS, which meant they had to bear the costs of customisation to ensure that their requirements would be met.

3. User acceptance

Test the system usability with users

An interface considered 'friendly' or 'usable' must nonetheless undergo rigorous testing by its target audience. The pilot implementation at the UNHCR allowed the project team to gather user feedback and then modify the user interface and data capture procedures accordingly. This pilot phase was critical to the success of the project—it helped to ensure that the end product was very easy-to-use, thereby increasing the overall level of user acceptance.

The most critical element for users was their ability to interact easily with the document repository from within their existing PC desktop tools (e.g. Microsoft Office and e-mail). The project team made this element a priority, ensuring that the procedures for storing and retrieving documents in EDRMS were as smooth and transparent as possible.

Ensure users have pre-requisite skills

To deliver classroom training effectively, the class participants must be grouped according to skill level wherever possible. At the UNHCR such grouping was not feasible because the training was delivered to one department at a time, and within one department the staff are at various skill levels.

The impact and effectiveness of training can be increased by establishing pre-requisites for participants. Specifically, each user would complete a brief skills assessment before attending EDRMS

training to determine if the pre-requisite skills were in place. Before learning a web-based document management tool, each class participant must be able to:

- demonstrate an understanding of basic concepts related to hierarchical file structures on a PC or network (i.e. relationships between drives, folders, sub-folders and files);
- use Windows file management tools to store and retrieve files or to move and delete files;
- use basic functions of a browser to navigate through a simple web site;
- use basic functions of their existing desktop tools (i.e. word processors or e-mail applications) if the EDRMS is closely integrated with these tools.

4. Future considerations

Ongoing training must be carefully planned to avoid a loss of momentum after the initial training rollout period. As well, adequate technical and procedural support must continue to be available, which will depend on available budget and staffing resources.

Most importantly, the Records and Archives department must continue to manage the document repository, as they are the key stakeholders within both the IOC and the UNHCR. They cannot work in isolation however, and must collaborate with other stakeholders and senior managers in the future.

Conclusion

In this paper we discussed the challenges of implementing an integrated document and records management system at two very different organisations, and in general we have concluded that these challenges are common to any such project. By focusing on the lessons learned we identified various factors that contribute to a successful EDRMS implementation:

- a solid technical infrastructure;
- a usable interface that favours user acceptance; and
- records management practices that meet the needs of records managers without burdening users in their daily work.

To conclude our discussion we look at several key issues affecting user acceptance.

Management support

From the project outset we knew that top management support would be essential for moving the project forward. But we did not anticipate the effect of middle managers' attitudes on their staff's perceptions of EDRMS. We observed a clear correlation between the positive attitude of a manager and a high level of system acceptance and use by the department staff. The reverse was also true — use of EDRMS was low or in some cases, non-existent, in departments where the manager had a very negative opinion of the system. To ensure a high degree of user acceptance it is important to obtain the support of both top executives and middle managers, so that communication or perceptions about the system will be predominantly positive.

Quantity of information

At the UNHCR we had expected that the first users of EDRMS in any given department would be the least likely to embrace the system immediately after deployment. This hypothesis was confirmed halfway through the implementation. As the number of documents in the system increased, the users perceived more value in EDRMS and consequently their level of system use increased.

Flexibility and control: finding the right balance

Throughout the project we observed an evolution from user resistance to user acceptance to user ownership of EDRMS. The degree of user acceptance and ownership was linked to the degree of

flexibility and control the users had over their 'workspace' within the document repository. Providing flexible options is an admirable and desirable objective, but some caution is necessary to ensure that records managers retain an appropriate level of control over the system and its core organisational structures.

Transforming theory into practice

As professional records managers we began these EDRMS projects armed with considerable information and knowledge of theoretical implementation issues and models. Through hands-on implementation, however, we learned how to bridge the distance between practical issues and theory. We also learned that to overcome each difficult challenge was to reach another project milestone, bringing us closer to our objective: to successfully integrate document and records management systems within a flexible, yet controlled framework.

Aplicación de sistemas de gestión de documentos y de archivos electrónicos en el ACNUR y el COI

Ineke Deserno, Montserrat Canela, Marsha Gillis

1. Introducción

El presente documento tiene por objeto informar sobre los sistemas integrados de gestión de documentos y archivos electrónicos del Alto Comisionado de las Naciones Unidas para los Refugiados (ACNUR) y del Comité Olímpico Internacional (COI). El documento describe las principales características de estos sistemas, la gestión de los archivos electrónicos y la aceptación y utilización de estos sistemas por los creadores de los documentos.

El ACNUR ha sido designado por la Asamblea General de las Naciones Unidas para realizar y coordinar la acción internacional en materia de protección de los refugiados en todo el mundo. La sede del ACNUR está en Ginebra, Suiza. La mayoría de las actividades relacionadas con los refugiados se realizan en las oficinas locales, que representan más de 250 repartidas en unos 120 países.

El COI es una organización internacional no gubernamental sin ánimo de lucro, creadora del movimiento olímpico. El COI aúna al conjunto de las organizaciones del movimiento olímpico. Su principal responsabilidad es supervisar la organización de los Juegos Olímpicos de verano y de invierno. Su sede está en Lausana, Suiza.

2. Funcionamiento de los sistemas

A finales de 1998, el ACNUR inició un proyecto de gestión integrada de documentos y archivos electrónicos. El COI inició su proyecto en 2001.

Los módulos de **gestión documental** incluyen las siguientes funciones:

- depósito de documentos electrónicos,
- colaboración y desarrollo de la operación,
- compartir información,
- búsqueda en texto íntegro de contenidos y atributos,
- integración con aplicaciones ofimáticas y correo electrónico,
- integración con escáner y fax.

Los módulos de gestión de archivos cumplen las siguientes funciones:

- plan de clasificación,
- normas de selección,
- información descriptiva,
- índice central de los archivos en papel y electrónicos,
- facilita la búsqueda de ficheros.

Es importante comprender que ninguno de estos módulos constituye una solución de conservación a largo plazo.

3. Introducción de los documentos electrónicos

El sistema de gestión de documentos y archivos electrónicos constituye el depósito de los documentos digitales. Estos documentos son los correos electrónicos, los documentos administrativos (en formato Word, Excel, etc.), los faxes y las imágenes de los mensajes recibidos, pasados por el escáner.

Los documentos que llegan se pasan por el escáner. La correspondencia de un día concreto puede consultarse y tratarse en una bandeja de entrada digital, durante todo el día. Al final del día, estos documentos se clasifican en los archivos del espacio de trabajo. Durante el día, el personal de los archivos registra, clasifica y codifica los documentos.

Los faxes que llegan se registran automáticamente en el espacio de trabajo adecuado, por medio de un servidor de fax. Estos faxes pueden consultarse y tratarse en la bandeja de entrada digital. De la misma manera que con el correo en papel, el personal de los archivos registra, clasifica y codifica el documento.

Los correos electrónicos y los documentos electrónicos son archivados por su creador. Los procedimientos automatizados y de fácil uso, integrados en el proceso de creación de documentos, contribuyen notablemente al éxito del proyecto. Así, en el ACNUR, cada vez que un usuario guarda un documento electrónico en una aplicación estándar (MS-Word, por ejemplo), aparece una pantalla que presenta el plan de archivo de la oficina. Basta con seleccionar la carpeta adecuada para guardar el documento en el archivo.

4. Requisitos de la gestión de archivos

El módulo de gestión de los archivos tiene por objeto organizar la información almacenada en los archivos electrónicos centrales. Las principales cuestiones archivísticas que deben abordarse en este nivel son las siguientes:

- creación de planes de clasificación
- integración de las normas de selección

Un plan de archivo es una herramienta de gestión archivística esencial, que garantiza que los documentos relativos a un mismo tema o una misma materia se agrupen en la misma carpeta. Este plan permite describir las relaciones entre los archivos y acceder a ellos para su consulta.

La información descriptiva sobre el documento archivado es capturada automáticamente por el módulo de gestión, que la asocia al documento archivado. Esta información comprende el nombre del creador, el destinatario, el objeto, la fecha de creación y el soporte.

El establecimiento de un vínculo entre las normas de selección y el plan de clasificación constituye una fase esencial de la gestión de los archivos y de la automatización del proceso de conservación. Los plazos de conservación determinarán la fecha en que los archivos y los metadatos asociados se eliminarán del depósito central.

Implementierung von Systemen für die elektronische Dokumenten- und Schriftgutverwaltung beim UNHCR und beim IOC

Ineke Deserno, Montserrat Canela, Marsha Gillis

1. Einführung

In diesem Beitrag sollen Informationen über integrierte Systeme für die elektronische Dokumenten- und Schriftgutverwaltung (Electronic Document and Records Management Systems, EDRMS) beim Amt des Hohen Flüchtlingskommissars der Vereinten Nationen (UNHCR) und beim Internationalen Olympischen Komitee (IOC) vermittelt werden. Themen sind die Hauptmerkmale der Systeme, die Verwaltung elektronischer Unterlagen und die Akzeptanz und Nutzung der Systeme durch die Urheber von Unterlagen.

Das UNHCR leitet und koordiniert im Auftrag der UN-Generalversammlung internationale Maßnahmen zum weltweiten Schutz von Flüchtlingen. Sein Hauptquartier befindet sich in Genf. Der größte Teil der UNHCR-Aktivitäten wird über seine derzeit mehr als 250 in etwa 120 Ländern befindlichen Außenstellen abgewickelt.

Das IOC ist eine internationale nichtstaatliche und nicht gewinnorientierte Organisation und Begründer der olympischen Bewegung. Seine Aufgabe besteht darin, als Dachorganisation für die olympische Bewegung zu fungieren. Vor allem ist es dafür zuständig, die Organisation der Olympischen Sommer- und Winterspiele zu überwachen. Das IOC-Hauptquartier befindet sich im schweizerischen Lausanne.

2. Funktionen des Systems

Ende 1998 startete das UNHCR ein Projekt für ein integriertes System für die elektronische Dokumenten- und Schriftgutverwaltung (EDRMS). Beim IOC lief dieses Projekt 2001 an.

Die Komponenten der **Dokumentenverwaltung** umfassen folgende Funktionen:

- Ablage für elektronische Dokumente,
- Zusammenarbeit und Workflow,
- gemeinsame Nutzung von Informationen,
- Volltextrecherche zu Inhalt und Attributen,
- Einbindung von Desktop-Anwendungen und E-Mail,
- Einbindung von Scann- und Faxmöglichkeiten.

Die Komponenten der Schriftgutverwaltung weisen folgende Funktionen auf:

- Klassifizierungsplan,
- Archivierungspläne,
- Profilangaben,
- zentraler Index für papiergebundene und elektronische Unterlagen,
- einfacheres Auffinden von Daten.

Zu beachten ist aber, dass keine der Komponenten eine Lösung für die Langzeitarchivierung darstellt.

3. Erfasste elektronische Aufzeichnungen

EDRMS ist die Ablage für elektronische Aufzeichnungen. Dazu gehören E-Mails, Desktopdokumente (z. B. Word und Excel), Faxmitteilungen und Bilddateien von gescannten eingegangenen Schreiben.

Eingehende Dokumente werden gescannt. Die Schreiben eines bestimmten Tages stehen einen Tag lang in einem digitalen Eingangsfach für eine Arbeitsabteilung zur Ansicht und Bearbeitung zur Verfügung. Am Ende des Arbeitstages werden die Dokumente in den Aktenplanordnern des Arbeitsraums der Arbeitsgruppe abgelegt. Im Tagesverlauf wird das Dokument von den Archivmitarbeitern registriert und codiert und erhält ein Profil.

Eingehende Faxmitteilungen werden durch einen Faxserver automatisch in den Arbeitsraum der jeweiligen Abteilung geleitet. Die Faxmitteilungen stehen im digitalen Eingangsfach zur Ansicht und Bearbeitung bereit. Wie bei papiergebundener Post wird das Dokument von den Archivmitarbeitern registriert und codiert und erhält ein Profil.

E-Mails und Desktopdokumente werden vom Urheber der Unterlagen abgelegt. Einen erheblichen Anteil am Erfolg des Projekts haben automatisierte und leicht anwendbare Verfahren im Prozess der Dokumentenerstellung. Beim UNHCR beispielsweise öffnet sich jedes Mal, wenn ein Benutzer ein elektronisches Dokument in einer Desktop-Standardanwendung (z. B. MS-Word) speichert, ein Bildschirmfenster mit dem Aktenplan für das Büro. Bei Auswahl des entsprechenden Ordners wird das Dokument dort abgelegt.

4. Anforderungen an die Schriftgutverwaltung

Die Komponente der Schriftgutverwaltung ist dafür verantwortlich, die in der zentralen elektronischen Ablage gespeicherten Informationen zu ordnen. Auf dieser Ebene müssen folgende zentrale Fragen gelöst werden:

- Aufstellung der Aktenpläne,
- Einbindung von Archivierungsplänen.

Ein Aktenplan ist ein wichtiges Hilfsmittel für die Schriftgutverwaltung, denn damit wird sichergestellt, dass Unterlagen zum selben Gegenstand oder zur selben Angelegenheit in ein und demselben Ordner zusammengefasst werden. Er bringt die Beziehung der Unterlagen zum Ausdruck und hält sie verfügbar.

Die Profilinformation über die Unterlage wird automatisch durch die Komponente Schriftgutverwaltung erfasst und mit der Unterlage in Verbindung gebracht. Diese Profilinformation beinhaltet folgende Angaben: Urheber/Absender, Adressat, Gegenstand, Erstellungsdatum und Medium.

Ein wesentlicher Schritt bei der Verwaltung des Schriftguts in der Ablage und der Automatisierung des Archivierungsprozesses ist die Zusammenfügung der Archivierungspläne der Organisation und der Aktenpläne. Mit den Aufbewahrungsfristen wird festgelegt, wann die Unterlagen und dazugehörigen Metadaten aus der zentralen Ablage gelöscht werden.

Die Komponente Schriftgutverwaltung bedeutet auch Einheitlichkeit und Übereinstimmung für alle Unterlagen, unabhängig vom Trägermedium oder dem Zeitpunkt ihrer Erstellung. Für die elektronischen, audiovisuellen und papiergebundenen Aufzeichnungen gelten die gleichen Aktenplancodes und Bewertungsverzeichnisse. Dadurch kann man einheitlich nach allen Datensätzen einer bestimmten Provenienz und/oder Thematik suchen.

Mit den EDRMS haben sich UNHCR und IOC für die elektronische Speicherung als dauerhafte Aufbewahrung ihrer Aufzeichnungen entschieden. Die Systeme bieten den

Archivabteilungen beider Organisationen automatisierte Hilfsmittel, mit denen sie elektronische Dokumente im Kontext, als Teil des gesamten Unterlagensystems einer Arbeitsgruppe bewerten können. Nicht dauerhafte elektronische Dokumente bzw. Papierunterlagen werden am Ende ihres Lebenszyklus kassiert. Dauerhafte elektronische Aufzeichnungen, die für das tägliche Geschäft nicht mehr benötigt werden, werden in das elektronische Archivsystem überführt, in die entsprechenden Bestände integriert und – genau wie die nichtelektronischen Unterlagen im selben Bestand – internen und externen Benutzern zur Einsicht bereitgestellt.

5. Benutzerakzeptanz von elektronischen Ablage- und Retrievalsystemen

In den letzten Jahren ist viel über die Konzipierung und Einrichtung von elektronischen Dokumenten- und Schriftgutverwaltungssystemen gesagt und geschrieben worden. Ein Aspekt, der nicht so umfassend diskutiert wurde, ist die Nutzung und Akzeptanz des Systems durch die Urheber von Unterlagen. Die Erfahrung hat uns gelehrt, dass der Erfolg eines EDRMS-Systems sehr stark von der Akzeptanz des Systems durch die Urheber abhängt. Bei der Präsentation in Barcelona werden auch die Schwierigkeiten vorgestellt, mit denen unsere Organisationen im Hinblick auf die Akzeptanz des Systems durch Urheber von Unterlagen zu kämpfen haben, und welche Erkenntnisse gewonnen wurden. Außerdem werden Vorschläge unterbreitet, wie dieser Aspekt bei der Planung, Entwicklung und Einführung eines solchen Systems zu verbessern ist.

Mise en œuvre de systèmes de gestion de documents et d'archives électroniques au HCR et au CIO

Ineke Deserno, Montserrat Canela, Marsha Gillis

1. Introduction

L'exposé a pour objet de renseigner sur les systèmes intégrés de gestion de documents et d'archives électroniques du Haut-Commissariat des Nations unies pour les réfugiés (HCR) et du Comité international olympique (CIO). L'exposé décrit les principales caractéristiques de ces systèmes, la gestion des archives électroniques, l'accueil réservé par les créateurs de documents à ces systèmes et leur utilisation.

Le HCR est mandaté par l'Assemblée générale des Nations unies pour mener et coordonner l'action internationale en matière de protection des réfugiés partout dans le monde. Le siège du HCR est situé à Genève, en Suisse. La plupart des opérations liées aux réfugiés sont réalisées par les bureaux locaux, soit plus de 250 bureaux répartis dans 120 pays environ.

Le CIO est une organisation non gouvernementale internationale sans but lucratif et est le créateur du mouvement olympique. Le CIO fédère l'ensemble des organisations du mouve-

ment olympique. Sa responsabilité première est de superviser l'organisation des jeux Olympiques d'été et d'hiver. Son siège est situé à Lausanne, en Suisse.

2. Fonctionnement des systèmes

À la fin de l'année 1998, le HCR inaugurait un projet de gestion intégrée de documents et d'archives électroniques. Le CIO a démarré son projet en 2001.

Les modules de **gestion documentaire** remplissent les fonctions suivantes:

— dépôt de documents électroniques,

— collaboration et déroulement des opérations,

— partage de l'information,

— recherche en texte intégral dans le contenu et les attributs,

— intégration avec des applications bureautiques et l'e-mail,

— intégration avec le scanner et la télécopie.

Les modules de **gestion d'archives** remplissent les fonctions suivantes:

— plan de classement,

— règlements de tri,

— informations descriptives,

— index central des archives «papier» et électroniques,

— traçabilité des emplacements de fichiers.

Il est important de comprendre qu'aucun de ces modules ne constitue une solution de conservation à long terme.

3. Saisie des documents électroniques

Le système de gestion de documents et d'archives électroniques constitue le dépôt des documents numériques. Ces documents sont les e-mails, les documents administratifs (au format Word, Excel, etc.), les télécopies et les images des courriers à l'arrivée passés au scanner.

Le courrier d'arrivée est scanné. La correspondance d'un jour donné peut être affichée et traitée dans la corbeille numérique d'une unité de travail, pour toute la durée de la journée. À la fin de cette journée, ces documents sont classés dans les dossiers de l'espace de travail. Au cours de la journée, le personnel des archives enregistre, décrit et codifie les documents.

Les télécopies à l'arrivée sont saisies automatiquement dans l'espace de travail approprié, par le biais d'un serveur fax. Ces télécopies peuvent être consultées et traitées dans la corbeille numérique. De la même manière qu'avec le courrier papier, le personnel des archives enregistre, décrit et codifie le document.

Les e-mails et les documents bureautiques sont classés par leur créateur. Des procédures d'utilisation automatisées et simples, intégrées dans le processus de création de documents, contribuent notablement à la réussite du projet. Ainsi, au HCR, chaque fois qu'un utilisateur fait une sauvegarde d'un document électronique dans une application bureautique standard (Ms-Word, par exemple), un écran présentant le plan de classement du bureau s'affiche automatiquement. Il suffit de sélectionner le dossier approprié pour sauvegarder le document dans les archives.

4. Cahier des charges de la gestion des archives

Le module de gestion des archives a pour fonction d'organiser l'information mémorisée dans des archives électroniques centrales. Les questions archivistiques primordiales qui doivent être résolues à ce niveau sont les suivantes:

— créer des plans de classement,

— intégrer les règlements de tri.

Un plan de classement est un outil de gestion archivistique essentiel, qui permet de veiller à ce que les archives concernant un même thème ou une même matière soient regroupées dans le même dossier. Ce plan permet de décrire les relations entre les archives et de communiquer celles-ci pour consultation.

Des informations descriptives sur le document archivé sont automatiquement saisies par le module de gestion et associées au document archivé. Ces informations sont: le nom du créateur, le destinataire, l'objet, la date de création et le support.

Établir le lien entre les règlements de tri et le plan de classement constitue une étape essentielle de la gestion des archives et de l'automatisation du processus de conservation. Les délais de conservation détermineront la date où les archives et les métadonnées qui leur sont associées seront éliminées du dépôt central.

Le module de gestion d'archives assure également la cohérence et l'unité de toutes les archives, quels que soient leur support ou leur date de création. Les mêmes codes de plan de classement et de règlements de tri sont appliqués à toutes les archives, électroniques, audiovisuelles et sur papier. Cette identité permet une consultation intégrée de tous les articles provenant d'un fonds et/ou sur un thème particulier.

En optant pour un système de gestion de documents et d'archives électroniques, le HCR et le CIO ont choisi le stockage électronique pour la conservation permanente de leurs archives. Ce système permet aux services des archives des deux organismes de disposer d'outils automatisés pour trier les archives électroniques dans leur contexte, en tant que partie intégrante de l'ensemble du système documentaire d'un groupe de travail. Les documents électroniques non permanents et/ou les documents «papier» seront éliminés à la fin de leur cycle de vie. Les archives électroniques permanentes devenues inutiles pour le fonctionnement au jour le jour seront versées dans le système d'archives électroniques, intégrées dans le fonds correspondant et mises à la disposition des chercheurs internes et externes, en même temps que les documents non électroniques du même fonds.

5. Accueil réservé par les utilisateurs au système de classement et de recherche

Beaucoup a été dit et écrit, ces dernières années, sur la conception et la mise en œuvre des systèmes de gestion de documents et d'archives électroniques. Un sujet a été moins évoqué que les autres: l'accueil et l'utilisation du système par les créateurs d'archives. Or, l'expérience nous a montré que la réussite d'un tel système est fortement tributaire de l'accueil que lui réservent les créateurs d'archives. L'exposé de Barcelone présentera les défis auxquels nos organisations doivent faire face eu égard à l'accueil de ce système par les services auteurs et aux leçons apprises. La présentation suggérera également quelques pistes pour améliorer cet aspect des choses au moment de la planification, du développement et de la mise en œuvre d'un tel système.

Matthias Meusch

Matthias Meusch, born 4 July 1967, studied history and English together with a high-school teaching qualification at the Justus-Liebig-University, Gießen, and the University of Illinois (United States). He worked on his doctorate from 1995 until 1998 and became a researcher at the Institute for the History of Medicine at Gießen University from 1997–99. From 1999 to 2001 he trained to become a senior archivist at the North-Rhine Westphalia State Archives in Münster and at the Marburg School for Archivists. From May 2001 to March 2002 he was a researcher at the North-Rhine Westphalia State Archives in Düsseldorf and contributed to a project by the Deutsche Forschungsgemeinschaft (DFG — German Research Association) entitled 'Development of retroconversion tools for archival finding aids'. Since March 2002 he has been responsible for the National Socialism Documentation Department and for computerisation in the Düsseldorf Principal State Archives; and is leader of the retroconversion project supported by the DFG.

Improving accessibility: DFG ([1]) project 'Development of retroconversion tools for archival finding aids'

Matthias Meusch

Introduction

Inventories as a problem? Surely the completion of a new inventory is in fact the solution to a problem. After all, the new finding aid describes a holding and gives access to the relevant records. Finding aids have traditionally been produced and disseminated in printed book form. It is only in the last few years that databases have made their appearance in archives, and then often only in larger ones. The vast majority of finding aids remain in analogue form — i.e. handwritten or typewritten — or are at best available as text files. And that does present a problem.

For many are now accustomed to searching online for books or articles. Archives too have in the last few years set about the task of satisfying changed user requirements for improved access to their material. However, so far all that most archives have provided on the Internet is an introduction to the archive concerned and a guide to its holdings. For instance, the archive portal 'NRW-Archive im Internet' ('Archives of North Rhine-Westphalia on the Internet') gives such details for over 400 archives covering every field in North Rhine-Westphalia. The logical next step, which is important for the archive sector in the age of information technology, is for finding aids to be made available not only in a given archive's own reading room but also on the Internet, so that searching both within a holding and at levels above that of a single holding or archive is made possible. There are also external demands on the profession to satisfy this objective: Sigrun Eckelmann, the Programme Director of the Deutsche Forschungsgemeinschaft, recently called for 'inventories to be presented on the Net as quickly as possible' ([2]). Good results have already been achieved in the English-speaking countries with the EAD (encoded archival description) presentation format: the standard coding of inventories and their presentation on the Internet makes searching possible at levels above that of an individual holding or archive ([3]).

The number of Internet-searchable German-archive finding aids is currently still very limited ([4]). The discrepancy between objective and reality is due to the fact that, as stated, the majority of existing inventories are available not in database format but as text files or in typewritten or even handwritten form.

This situation is addressed by the project 'Development of retroconversion tools for archival finding aids', which has been assisted by the Deutsche Forschungsgemeinschaft since 1 May 2001 ([5]). Since the problem arises in comparable form for all archives, the project aims for as wide a consensus as possible among archives in all the various fields. It therefore takes the form of a cooperative project, under the primary responsibility of the Nordrhein-westfälisches Hauptstaatsarchiv (Principal State Archive of North Rhine-Westphalia) in Düsseldorf, embracing the Bundesarchiv (Federal Archive), the North Rhine-Westphalian State and Local Authority Archives, represented by the two Archive Offices in the Rhineland and Westphalia, and the Westfälisches Wirtschaftsarchiv (Westphalian Trade and Industry Archive) in Dortmund. 'Retroconversion' means retrospective digitisation, a notion borrowed from the library sector. Libraries have been digitising their card catalogues since the 1980s; this was at first done mainly by manual copying, but imaging techniques are now also occasionally used ([6]). Both methods require a great deal of manual work. However, manual conversion of the German archives' finding aids awaiting digitisation is precluded by their enormous volume. We are therefore aiming for a substantially automated software-based process. Inventories are structured as free forms, which are read out by software tools in other sectors too: an efficient OCR system, a layout analysis and a keyword-based approach permit structuring of inventory pages and of individual description units with their individual elements. One reservation, however, still applies for the time being: a convincing technical solution is not yet in prospect for handwritten inventories. For this reason, we are for the present confining ourselves to the conversion of typewritten finding aids and those in the form of text files. A detailed plan for tackling

([1]) Deutsche Forschungsgemeinschaft (the central public funding organisation for academic research in Germany).

([2]) Eckelmann, S., 'Die Archivförderung der DFG — Schwerpunkte und Perspektiven', *Online-Findbücher, Suchmaschinen und Portale.* Beiträge des 6. Archivwissenschaftlichen Kolloquiums der Archivschule Marburg, ed. Angelika Menne-Haritz, Marburg 2001 (Veröffentlichungen der Archivschule Marburg, No 35), pp. 173–181; this quotation is taken from p. 177.

([3]) On the potential for and limits of the use of EAD in German archives, see Camp, A., Kiesling, K., Dhérent, C. and Grau, B., *Online-Findbücher, Suchmaschinen und Portale* (as note 2); the contributions of Daniel Pitti, Seamus Ross, Anne van Camp and Black-Veldtrup, M.,'Archive vor der Globalisierung? Beiträge zum Symposion des Nordrhein-Westfälischen Hauptstaatsarchivs in *Verbindung mit den Allgemeinen Reichsarchiven in Brüssel (Belgien) und Den Haag (Niederlande) 11–13. September 2000* Düsseldorf, ed. Black-Veldtrup, M., Dascher, O. and Koppetsch, A., Düsseldorf 2001; and the volume: Eckelmann, S.,

handwritten originals is to be compiled as a basis for an application for a follow-up project. For this purpose, over the next year analyses of the cost and labour requirements for conventional solutions such as manual copying and imaging techniques are to be undertaken, as a pointer at least to the beginnings of a suitable approach. The latest developments in the field of handwriting pattern recognition (OCR/ICR) will also be considered in this connection.

I should now like to give you a brief description of the current status of our project and its results so far. This will be divided into two parts. The first outlines the fundamental capabilities and functionalities of the software to be written in the project, while the second introduces an exchange and target format that we have developed for the retroconverted data.

The retroconversion tool

The project centres on the development and testing of a software tool for the substantially automated conversion of typewritten archival finding aids and those in the form of text files to a database environment. The function of this software is to recognise the individual items of information in a single-item description (e.g. reference number, inclusive dates or title) as such and to read them out separately into the target and exchange format described in more detail below. The particular difficulty arising concerns the variable formats of the finding aids: column, table and run-on text layouts, or even index cards. In the case of text files, the range extends from the consistent application of standard formats to completely 'wild' formatting using blanks and tabs. The structure of the single-item descriptions, too, is highly variable from one finding aid to another. In the absence of standardisation of archival finding aids, the software to be developed must therefore be flexible and adaptable. It must also be capable of recognising both relational and hierarchical structures within a description.

The software is currently being developed by the Hamburg firm of CCS compact computer systems ([7]) in close collaboration with the project staff. It comprises not so much an entirely new program as the adaptation of a successful, high-performance layout analysis tool called docWorks that has been on the market for some years. The company is drawing for this purpose on its experience of software adaptation for the EU-wide Meta-E project ([8]). Let us now consider the workflow of a retroconversion process using the tool now being developed, as far as this can be determined at this stage (the process of technical implementation is after all only just beginning) (see Figure 1).

Typewritten finding aids are first scanned in. The inventory pages, now in the form of TIFF files, are imported into the tool and undergo layout analysis on the basis of individually defined rules. The results are in turn compared with a master set of rules, categorised and undergo text recognition by a commercially available OCR program. Next comes a correction stage, in which the results of the layout analysis and of the OCR are checked for recognition errors and if necessary corrected manually. The individual items of description data in a single-item description are then read in structured form into an XML instance based on the DTD (document type definition) of the exchange format. Finding aids already in text-file form are transferred to an internal format by an analysing Windows printer driver before the individual items of description data are again extracted by means of the rule-based layout analysis described above.

To ensure that the tool has the required flexibility, it must be possible for the operator to parameterise this layout analysis for the relevant finding aid. Before a retroconversion begins, the tool must be adapted to the specific inventory by defining a pattern template. This is done in two stages. First, the individual recognition fields on an inventory page (or index card) must be specified and, second, these must be assigned to the appropriate headings of the exchange format, which has a tree structure. Hence, after this parameterisation, the software should, for example, automatically be able to identify that a particular field on an inventory page is a reference number or the 'contents' part of a description and assign the contents of these fields to the corresponding headings of the exchange format. In the case of the reference number the basis for identification may be its position — often offset to the left of or above the title — while the 'contents' part can be recognised by predetermined keywords such as 'Enthält' ('contents'), 'Enthält u.a.' ('contents include') or 'Intus'. Inclusive date ranges can sometimes be identified by their position (usually at top or bottom right) or by their specific structure, e.g. by the definition of a regular expression such as 'Laufzeit (date range) = ((four digits) – (four digits))'.

Kreikamp, H. D., Menne-Haritz, A., and Reininghaus, W., *Neue Medien im Archiv: Online-Zugang und elektronische Aufzeichnungen. Bericht über eine Studienreise nach Nordamerika, 10–21. Mai 1999*, Marburg 2000 (Veröffentlichungen der Archivschule Marburg, No 32). As a part of the DFG project 'Internet presentation of online inventories using the EAD system', the Bundesarchiv (Federal Archive) is preparing the Internet presentation of EAD inventories and assessing the potential for joint European action.

([4]) An overview of the German inventories available on the Internet can be found on the web site of the Archivschule Marburg: http://www.uni-marburg.de/archivschule/fv61.html. Hartmut Weber estimates that in the medium term 10 % of all unrestricted-access inventories will be available on the Internet; see Hartmut Weber, Digitale Repertorien, virtueller Lesesaal und Praktikum im www — neue Dienstleistungsangebote der Archive an die Forschung, on: http://webdoc.sub.gwdg.de/edoc/p/fundus/2-00/weber.pdf, p. 16.

([5]) See http://www.archive.nrw.de/dok/retrokonversion01

([6]) A large number of contributions on this subject can be found on http://bibliotheksdienst.zlb.de/; see, for example, Angelus, A., Eichhorn-Berndt, C., Schnelling, H., 'Digitalisierung des Realkataloges (Hartwig-Katalog) der Universitäts- und Landesbibliothek Sachsen-Anhalt und seine Visualisierung im Internet', *Bibliotheksdienst, Heft 3*, 2000. See also the DFG project 'Codices Electronici Ecclesiae Coloniensis (CEEC)' on: http://www.ceec.uni-koeln.de/

([7]) See http://www.ccs-gmbh.de/

([8]) See http://meta-e.uibk.ac.at/

The layout analysis consists of a set of rules contained in an SQL-server-based database. The rules are based on a widely used script language called TCL (version 8.3). However, extensive IT skills or programming capabilities will not be needed for parameterisation — i.e. adaptation of the layout analysis to an inventory. Since docWorks has a user interface based on Windows standards, adaptation will be menu-driven or use drag-and-drop. The commands entered in this way are then automatically converted into TCL rules without the need for the operator to work with the program code. The tool will thus be usable by people without an IT background, an essential requirement for its eventual widespread adoption.

Once a template has been defined, it can be stored and used for further retroconversion processes. The more patterns are available, the fewer new definitions will be necessary. The tool is intended to present the user with the template that is most likely to be appropriate, which s/he can then adapt by simple means. These pattern templates are of course determined not purely by the layout, but also by the type of archived material described in the finding aid. Hence there are separate patterns for, say, action files, acts and charters, and maps. At present case files ('parallel files'), posters, photographs and films are also covered.

Once the tool has been parameterised from the first few pages of an inventory, ideally the retroconversion process can then continue automatically for the rest of the finding aid. This ideal case is represented by an inventory with consistently uniform structure and page layout, without irregularities of any kind — for instance, later additions. However, this will be the exception in most archives. As a rule, a semiautomatic mode will have to be used, in which the software stops when it encounters an irregularity — e.g. a handwritten addition — and requests manual intervention by the user.

A manual mode, in which each description unit is converted singly, is provided for especially difficult originals. This may seem extravagant, but we hope it will ultimately still be more economic than having the finding aid copied manually. To verify this, the exact labour input in terms of time and staffing is to be determined and documented in the real-service phase of the project by the conversion of some 300 finding aids.

The XML instance that is the end-result of the retroconversion process can be transferred to an archival database by means of a suitable import interface. This interface must be provided by the database operators, and to this end contact has already been made with the relevant firms. Although this is not a part of the project itself, one of our main aspirations is, as stated at the beginning, the presentation of these retroconverted finding aids on the Internet. There is no difficulty in making XML instances available online using style sheets based on the XSL (extensible stylesheet language) XML co-standard. Alternatively, they can be put on the Internet from a database.

So much for the actual workflow in the main part of a finding aid, the single-item description. However, the entire finding aid, with all its components, must be converted. These include the title page, the introduction, possibly with footnotes and bibliographic references or graphics — e.g. organisational charts — the contents, classification, indexes and concordance lists. All these elements must also be reproducible in the exchange format we have developed, which I shall now briefly introduce.

The exchange format

Our exchange format comprises a document type definition (DTD) for reading the retroconverted data into an XML instance, by means of which it can then be imported into a database or direct on to the Internet. XML allows (data) elements to be hierarchically interleaved and specified by 'attributes' so as to satisfy archival requirements.

A number of special-purpose archive database systems from different suppliers are used by German archives. Since our tool is intended not merely for use in isolation but to be adopted on a widespread basis — perhaps even as a kind of standard — in the archive sector, as universally valid as possible a data structure for the exchange format must be aimed for. This will allow the individual components of archival description information to be described and their layout, order and possible frequency to be specified.

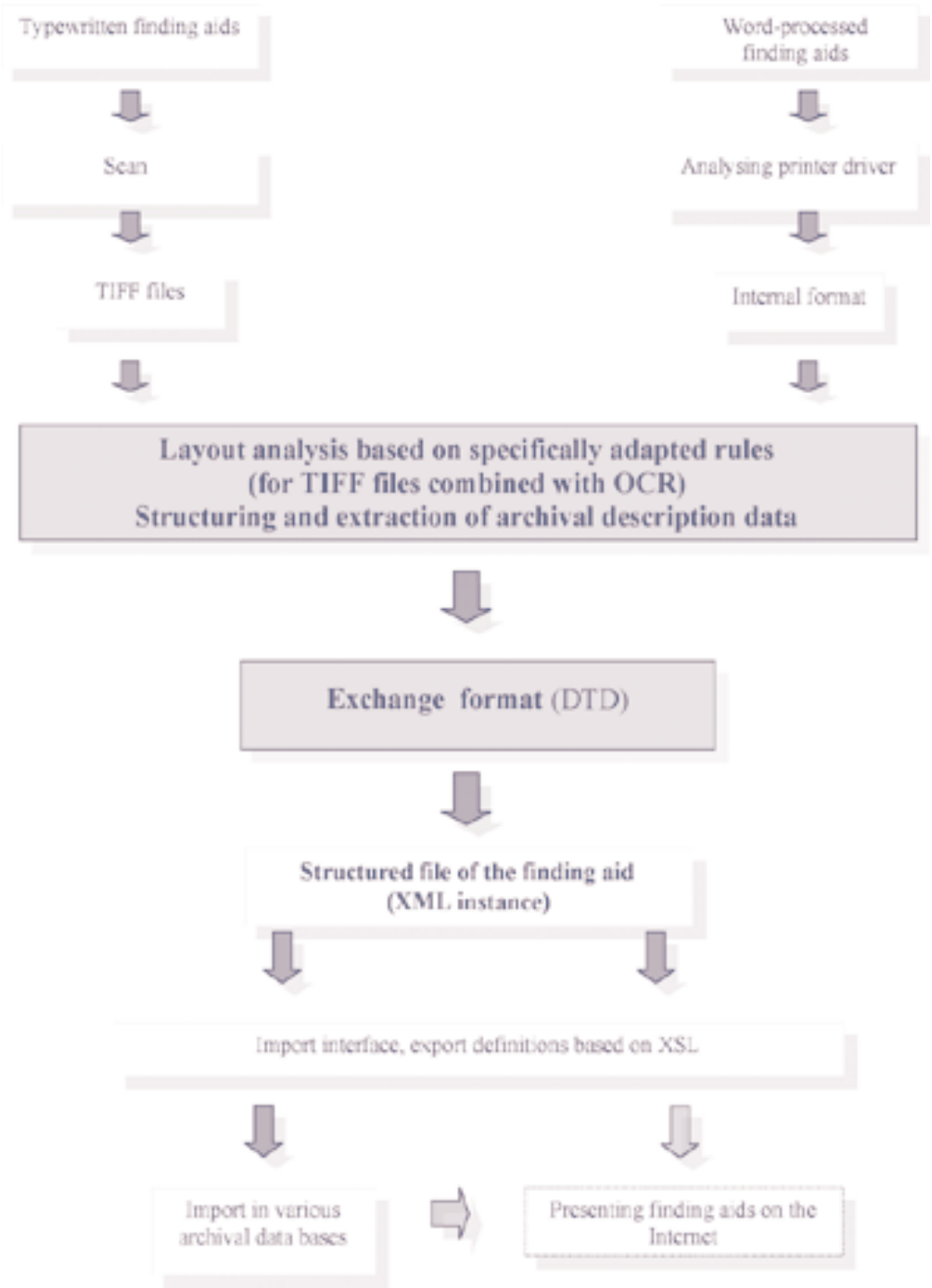


Figure 1. Workflow of a retroconversion process

The following fundamental requirements of an exchange format as defined above were specified on commencement of the project ([9]):

- it must be possible to reproduce the hierarchical and relational links in archival finding aids without information loss;
- as 'flat' as possible a file format must be chosen; it should be standardised and non-proprietary (ASCII/ANSI);
- account must be taken of existing archival standards (e.g. EAD) while at the same time maximising openness to future technical developments and those in the field of archival science.

It was therefore virtually inevitable that XML, as a platform-independent and now widely used standard, would be chosen for defining the exchange format; after all, it has the great advantage of consistent separation of structure, content and layout ([10]).

([9]) See the project descriptions given by Black-Veldtrup, M., Das DFG-Projekt 'Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel', in: *Online-Findbücher* (as note 2); Bischoff, F. M., 'Das Projekt Retrokonversion von Findbüchern', *Archive vor der Globalisierung?* (as note 3), pp. 148ff.

([10]) On this point, see also Sahle, P., 'Sinnsuche in der Badewanne', *Tagungsbericht: Standards und Methoden der Volltextdigitalisierung, Trier, 8. und 9. Oktober 2001*, on: http://computerphilologie.uni-muenchen.de/jg02/sahle.html. See also Menne-Haritz, A., 'Die Herstellung internationaler Kompatibilität archivischer Erschließungsangaben mit XML-Austauschformaten, in: Der Zugang zu Verwaltungsinformationen — Transparenz als archivische Dienstleistung', *Beiträge des 5. Archivwissenschaftlichen Kolloquiums der Archivschule Marburg*, ed. Brübach, N., Marburg 2000 (Veröffentlichungen der Archivschule Marburg, No 33), pp. 129–146.

(11) The format was finalised after requesting information from the administrations of all the German archives on the archival administration and database systems they used.

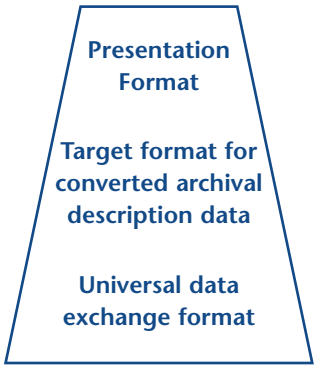


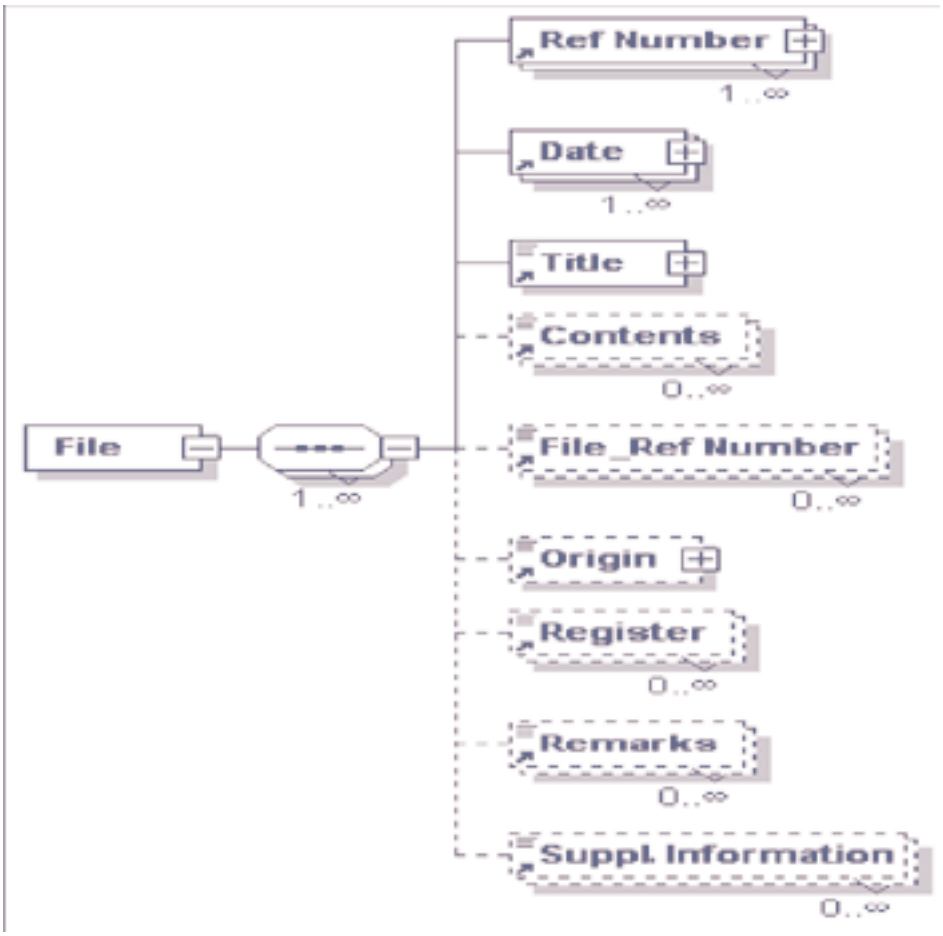
Figure 2. Archival exchange formats: a variety of differentiation and scope

Three different format types are initially distinguished for the purposes of the project: a general exchange format, a target format for the retroconverted data and an Internet presentation format for finding aids (see Figure 2). A comparatively low degree of differentiation and hence a relatively limited number of data fields would suffice for online presentation. Conversely, a target format for importing retroconverted inventory data calls for greater differentiation and therefore has more elements and greater complexity (it is based on the description data contained in the analogue inventories). This applies even more to a format intended to permit cross-archive data exchange between different database systems. What kind of format have we opted for? As stated, the project does not at this stage cover the technical implementation of the presentation of inventory data on the Internet. There seemed little point in separating the target format from the general exchange format. For this reason, we finally chose a fairly highly differentiated format (a DTD with some 150 elements at present) that can be used as both target and exchange format (11).

The particular form this assumes will be illustrated by the example of the description of a seal. The description data in an archival database may sometimes be highly differentiated. For instance, the individual categories of the description (shape, colour, material or type of seal) may be placed in separate fields. In an analogue inventory, by contrast, these individual items of information may be presented as a coherent block of text, and although the retroconversion tool can recognise this as such, it will usually have difficulty in subdividing it. With our DTD, either the entire description can be read in as a single information block, or more detailed structuring by individual headings can be carried out.

A fundamental design principle of this DTD is that associated descriptive elements are put together in 'containers'. For instance, all the individual elements needed for the archival description of a file, such as the reference number, title and date, are subordinated to the container element 'File' (see Figure 3).

Figure 3



To maximise flexibility, only certain plainly essential core structures and elements are stipulated as compulsory in the chosen conception, which, however, provides for a high degree of differentiation and specification. To this end, the number of compulsory elements required for the individual components of an inventory or unit of description is small compared with the number of additional options and possible attributes.

To avoid redundancy and excess in the exchange format, the general aim has been to define and combine elements in such a way that they can be used on different levels and in different contexts. Attributes are used for specification and unequivocal assignment. Hence a specific element for the dates of acts or charters will not be found. Instead, the container 'Laufzeit' ('inclusive dates') contains as its functional equivalent the element 'LZ_Text' ('date text') — which is used for alphanumeric inclusive date range information rather than 'LZ_Anfang' ('start date') and 'LZ_Ende' ('end date'), including a number of attributes to allow more specific dating. In accordance with almost universal practice, a spare field and a remarks field are provided for further information that does not fit into this framework.

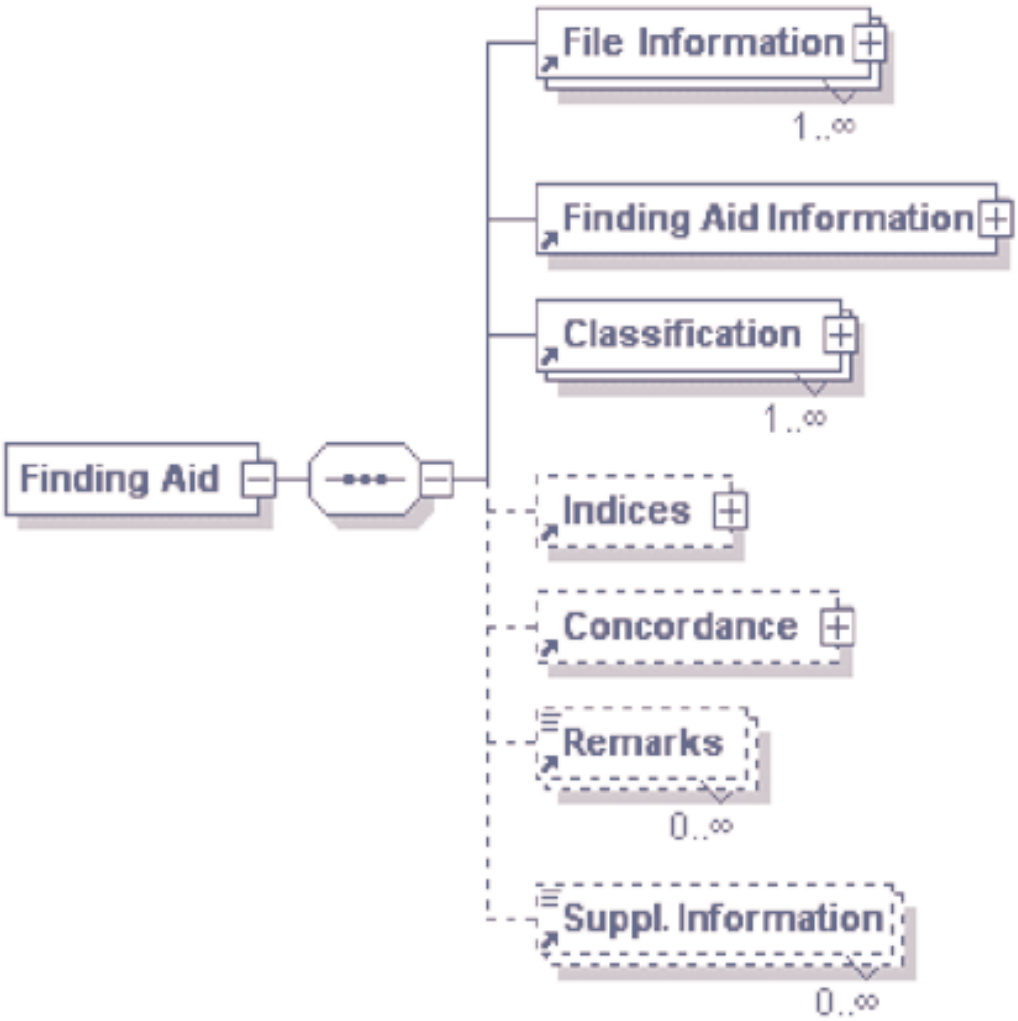
```
<Laufzeit Dat_Art=»Erschlossen»>

        <LZ_Text> 1320 Januar 02</LZ_Text>

</Laufzeit>
```

Figure 4. Sample reproduction of the dating of an act or charter in an XML instance

What does the structure of our DTD actually look like? How does it reflect the information contained in archival finding aids with its hierarchies and relations? A comprehensive, detailed account is beyond the scope of this presentation, and I shall therefore confine myself to a brief description of some fundamental structural features.

The starting point — or 'root element' in XML terminology — is the relevant finding aid. The first level includes a number of containers for detailed information on the individual components of an inventory and on the appropriate associated XML instance (the element 'Dateiinfo' (file information)). Information on the finding aid and instance, as well as the element 'Klassifikation' (classification), is compulsory. Indices and concordances, as well as remarks, can of course be appended optionally. There is always a spare or auxiliary field specifiable by attributes (see Figure 5).

Figure 5. First-level DTD container elements

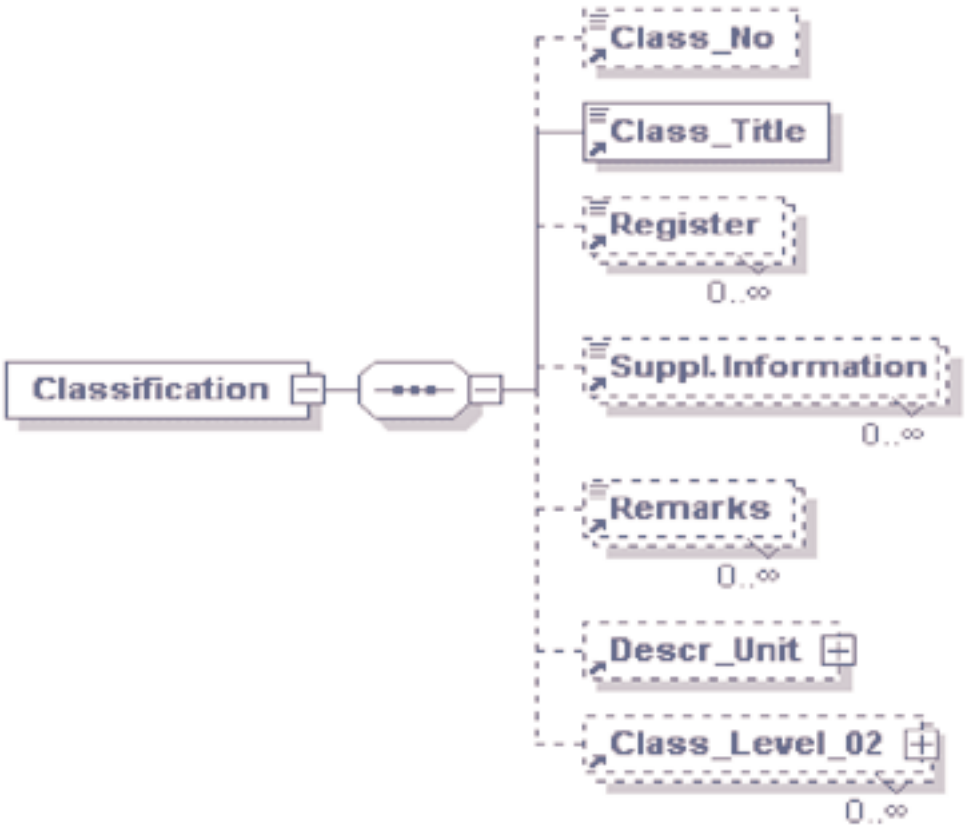


The actual description data is reproduced hierarchically in the container 'Klassifikation' ('classification'). Where necessary, the next lower hierarchical level can be reached with the element 'Klass-Stufe_02' (class level 02), which is structured in exactly the same way as 'Klassifikation' (classification). Up to 10 classification levels can be distinguished in this way. Hierarchical volume series, individual processes within a file and multilevel indices can be reproduced along the same lines (see Figure 6).

A particularly important element is 'VE', an abbreviation of the German term for description unit, which contains the element sets of the different kinds of archival material (e.g. files, diplomas, maps or films) for describing the individual units of description (see Figure 7).

The DTD also includes provision for both vertical and horizontal extensions of various kinds (see Figure 8). Since the project concerns the digitisation of individual finding aids, the exchange format is of course primarily directed to this level. However, the present finding-aid DTD can be integrated without difficulty into a holding-related DTD to be defined at a later date, thus allowing the reproduction of guides to archival holdings. Another possibility is a reference to reproductions of individual records. Other points worth mentioning are compatibility with existing standards and openness to future developments, which, as stated above, are vital requirements that the exchange format must satisfy. To this end, for instance, a mapping system is being developed for the transfer of data from our exchange format to EAD by the XSLT XML co-standard.

Figure 6. Hierarchical structure of the container element 'classification'



Outlook

The software for the retroconversion of archival finding aids is currently in the course of development. A prototype will be completed in the next few weeks and will then undergo six months of further development in a 'prototyping phase' involving close collaboration between our own project staff, project staff at the Münster State Archive and engineers of the software house. The end-product is scheduled for completion in the autumn. The tool will then be used for the conversion of some 300 finding aids in a real-service phase lasting until the end of the project in summer 2003, and the costs and labour requirements involved will be determined. The results will be

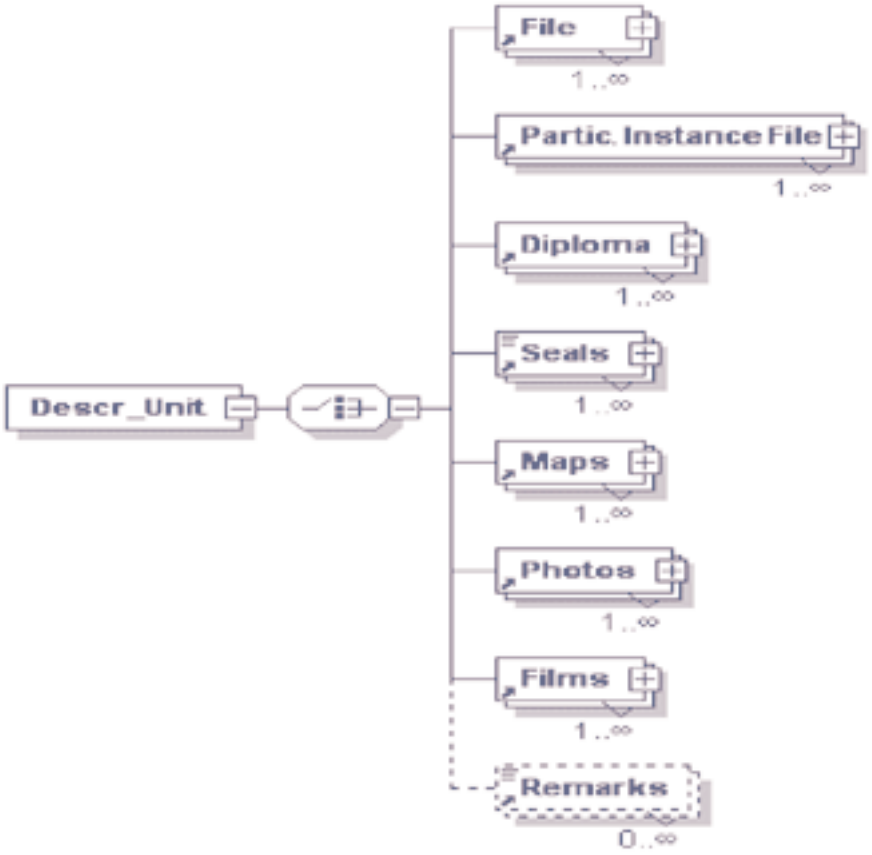


Figure 7. Container element 'Description Unit' (Descr_Unit)

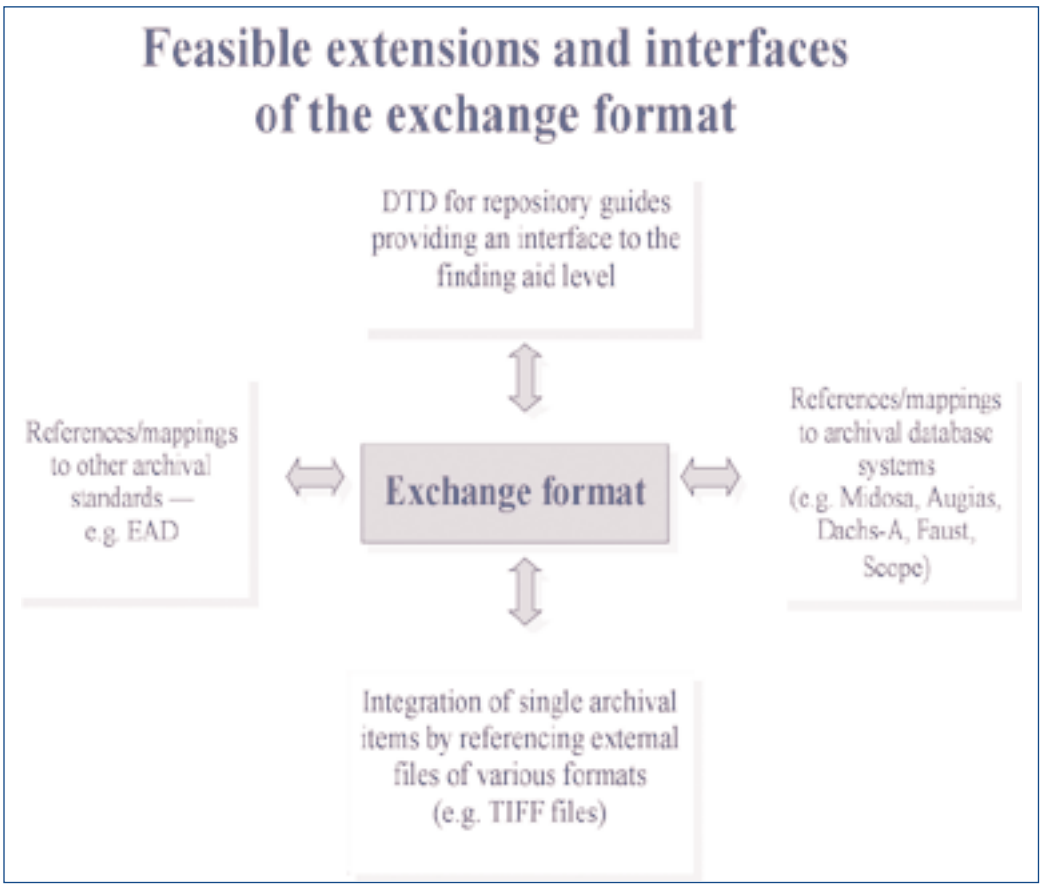


Figure 8. Feasible extensions and interfaces of the exchange format

recorded in a final report and also presented at a workshop at the Hauptstaatsarchiv (Principal State Archives) in Düsseldorf.

Following internal project consultations, the prototype of the finding-aid DTD was recently made available to a wider circle of specialists in an alpha version together with an associated tag library, with a view to stimulating what is hoped will be an intensive debate involving a number of different archives. This seems to us to be the prerequisite for maximising acceptance of the system in the archive sector. The importance of standards was repeatedly stressed at the first plenary session of this year's DLM-Forum. We should very much like our DTD to develop into a standard for exchanging archival data.

It is hoped that the exchange format will be widely adopted; the software tool will also be freely usable — that is to say, apart from licence fees no costs will be incurred by public archives in the Federal Republic of Germany. When the project comes to an end — i.e. after the spring or summer of 2003 — it will be possible to make the software available to other interested archival bodies. In addition, as far as North Rhine-Westphalia is concerned, for the period after the end of the project it is planned to establish retroconversion centres at the State Archives and the two archive offices, where retroconversion work for other, smaller archives in their areas will also be undertaken. This could become a pilot with effects extending beyond North Rhine-Westphalia and perhaps even beyond the Federal Republic of Germany.

Proyecto del DFG para el desarrollo de herramientas modernas para la retroconversión de instrumentos de búsqueda de archivos

Matthias Meusch

A partir de finales de los años noventa, los archivos de Renania del Norte-Westfalia comenzaron a proponer a sus investigadores la posibilidad de investigar en sus fondos y colecciones en línea, a través de Internet. Uno de los primeros proyectos financiados por el consorcio de investigación Deutsche Forschungsgemeinschaft (DFG) dio lugar a la creación del portal de Internet «NRW Archive im Internet» (http://www.archive.nrw.de/), que permite acceder a los fondos de más de 400 servicios de archivos de distintas fuentes (servicios de archivos estatales y comunales, archivos de la Iglesia, archivos económicos, etc.).

La segunda fase consiste ahora en poder buscar también en Internet instrumentos de investigación en gran cantidad. Los instrumentos de investigación en línea utilizados hasta ahora se han elaborado esencialmente a partir de bases de datos, o bien se han introducido a mano con grandes dificultades, o incluso se presentan en forma de imágenes que es necesario indizar manualmente. La gran mayoría de los instrumentos de investigación, principalmente los más utilizados, están escritos a mano o mecanografiados, o, en el mejor de los casos, figuran en forma de ficheros de textos, que no se prestan a su utilización en base de datos ni en línea. Ahora bien, habida cuenta de este gran volumen, la conversión puramente manual es imposible. Por eso, el DFG apoya, desde el 1 de mayo de 2001, un proyecto piloto dirigido a desarrollar una herramienta informática que permitirá automatizar todo lo posible esta retroconversión. Los cuatro servicios de archivos públicos de Renania del Norte-Westfalia, a saber, el Rheinisches Archiv- und Museumsamt, el Westfälisches Archivamt (los archivos municipales de Renania del Norte-Westfalia también están representados por medio de estos dos servicios), el Westfälisches Wirtschaftsarchiv y los archivos federales, se han asociado al proyecto, que está bajo la tutela de los archivos públicos centrales del Estado federado, en Düsseldorf.

El presente proyecto no tiene por objeto digitalizar los archivos, sino solamente la información que permite su explotación, es decir, los metadatos. A diferencia de las bibliotecas, donde un arsenal reglamentario, como por ejemplo las «normas de clasificación alfabética» (RAK) o las «instrucciones prusianas» (PI), velan por una determinada uniformidad y normalización, por lo que respecta a la descripción de archivos existen muy pocas directrices de este tipo. No sólo el volumen de la información contenida (profundidad de descripción) en cada título del índice, sino también la disposición de las páginas del directorio digital, por ejemplo, son muy diferentes entre los distintos instrumentos de investigación. Por esta razón, la herramienta que se desarrolle deberá ser flexible.

A partir de ficheros de texto o de páginas de directorios digitales mecanografiadas escaneadas, la herramienta de retroconversión deberá ser capaz, mediante la utilización de palabras clave, códigos, formatos o características de paginación, de reconocer la información de explotación (título, firma, plazo, etc.) como tal y convertirla a un formato de intercambio de manera estructurada y sin ninguna pérdida. Por lo que se refiere a este formato de intercambio, se tratará de una definición de tipo de documentos (DTD) programada en XML, la estructura de cuyos elementos se acerca a la información contenida en los directorios digitales. Esta DTD deberá ser capaz de representar una jerarquía de clasificación con exactitud, como series de volúmenes, por ejemplo. El formato deberá poder integrar los vínculos jerárquicos y los vínculos de relación dentro de un fondo, sin pérdida de información. Por otra parte, esta DTD no deberá estar protegida por una exclusividad, sino que deberá concebirse en interés general de forma que permita intercambiar datos con el mayor número posible de sistemas de bases de datos y con las estructuras de datos más diversas; si es posible, también deberá servir como norma para intercambiar datos entre los sistemas de información de archivos. En su versión actual, el formato soporta expedientes y actos jurídicos, sellos, mapas, planos, fotografías, carteles y películas.

Antes de iniciar el procedimiento de retroconversión, debe parametrarse la herramienta según la estructura del instrumento de investigación que vaya a convertirse. La introducción de los parámetros debe efectuarse en principio en el nivel del usuario, es decir, por medio de menús o mediante el desplazamiento del ratón, y las intervenciones en el código sólo pueden admitirse para volúmenes de parámetros más importantes. Los parámetros, una vez introducidos, se memorizan y se utilizan como modelos para los trabajos de conversión posteriores. A medida que aumenta el número de procedimientos de retroconversión, disminuye el número de nuevas definiciones, mientras que el número de adaptaciones de los parámetros existentes sigue progresando. La conversión propiamente dicha, es decir, el reconocimiento y la estructuración de los datos de explotación, se hace automáticamente en la mayoría de los casos, y al menos de forma semiautomática por regla general. Puede preverse un método totalmente manual para los casos más difíciles. Debe calcularse y tenerse en cuenta el tiempo necesario y el coste financiero de la retroconversión de un instrumento de investigación, para que sirva de referencia a proyectos de retroconversión posteriores, y en particular a proyectos financiados por terceros.

En la primavera de 2002 estará disponible un prototipo de esta herramienta. En los seis meses siguientes, el *software* se perfeccionará y concluirá en cooperación con los colaboradores del proyecto y los programadores. En el período de explotación real que seguirá hasta el final del proyecto, previsto para abril de 2003, se convertirán unos 150 instrumentos de investigación, mecanografiados o en forma de ficheros de texto, mientras que se calculará la inversión necesaria para esta operación. Al final, los instrumentos de investigación se presentarán en forma de instancias XML, que podrán convertirse sin problemas en bases de datos o prepararse con vistas a su presentación en línea en Internet.

El proyecto también prevé experimentar la posibilidad de una retroconversión de instrumentos de investigación manuscritos. Por instrumentos de investigación manuscritos se entienden los instrumentos escritos en caracteres latinos o alemanes *(Sütterlin),* que generalmente están escritos por varias manos, así como los instrumentos de investigación mecanografiados que incluyen un gran número de adiciones manuscritas. En principio, en el marco de un estudio de mercado, tres vías son posibles y aplicables: 1) transformación de

los directorios escaneados en ficheros de texto con ayuda de un *software* de reconocimiento inteligente de caracteres (RIC); 2) conversión del instrumento de investigación en forma de imágenes, seguida de una indización manual a efectos de la investigación; 3) copia del instrumento de investigación, realizada internamente o por prestadores de servicios externos. La viabilidad y la rentabilidad de las tres soluciones se analizarán a partir de una muestra de una treintena de instrumentos de investigación manuscritos.

Al final del proyecto, en la primavera de 2003, debería poderse proponer una herramienta que permita convertir rápidamente los instrumentos de investigación más importantes, o incluso la totalidad, de un servicio de archivos, sin ningún problema y sabiendo con exactitud lo que costará en términos financieros y humanos. La herramienta será flexible, de tal modo que permita su adaptación a los instrumentos de búsqueda de todos los tipos de archivos. Uno de los capítulos del proyecto, titulado «Retroconversión de instrumentos de investigación manuscritos», permitirá aportar conocimientos sobre las posibilidades y los medios disponibles en este ámbito, indudablemente el más problemático.

Por ello, el proyecto del DFG para el desarrollo de herramientas para la retroconversión de instrumentos de búsqueda de archivos no es más que una primera etapa hacia la digitalización generalizada de los instrumentos de investigación y su comunicación por Internet. El objetivo supremo consiste en continuar por la vía de la apertura de los archivos y su anclaje en la sociedad moderna de la información digital, dirección por la que Alemania ya está bien encaminada estos últimos años.

Das DFG-Projekt zur Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel

Matthias Meusch

Schon Ende der 90er Jahre haben die nordrhein-westfälischen Archive damit begonnen, ihren Benutzern auch online über das Internet Möglichkeiten zur Recherche in ihren Beständen zu bieten. Ergebnis eines ersten von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekts war das Internet-Portal „NRW Archive im Internet“ (http://www.archive.nrw.de/), das Zugang zu den Beständen von über 400 Archiven unterschiedlicher Sparten (z. B. Staats-, Kommunal-, Kirchen- und Wirtschaftsarchive) bietet.

In einem nächsten Schritt sollen nun auch Findmittel in großem Umfang über das Internet recherchierbar werden. Die bisher online vorliegenden Findmittel wurden größtenteils aus Datenbanken konvertiert oder mühsam von Hand eingegeben, oder man präsentiert Images der Findbuchseiten, die nachträglich manuell indiziert wurden. Die weitaus größte Zahl der Findmittel, vor allem auch diejenigen mit der höchsten Benutzungsfrequenz, liegen hand- oder maschinenschriftlich oder bestenfalls als Textdateien vor und sind nicht datenbank- beziehungsweise onlinefähig. Angesichts der großen Masse verbietet sich jedoch eine rein manuelle Konversion. Die DFG fördert deshalb seit dem 1. Mai 2001 ein Pilotprojekt, in dessen Rahmen ein Software-Tool entwickelt wird, das eine weit gehende Automatisierung dieser Retrokonversion ermöglichen soll. Unter Federführung des NRW Hauptstaatsarchivs Düsseldorf sind an diesem Projekt die vier nordrhein-westfälischen Staatsarchive, das Rheinische Archiv- und Museumsamt, das Westfälische Archivamt – über die beiden Archivämter sind auch die nordrhein-westfälischen Kommunalarchive mit eingebunden –, das Westfälische Wirtschaftsarchiv und das Bundesarchiv beteiligt.

Im Rahmen dieses Projekts sollen keine Archivalien digitalisiert werden, sondern nur Erschließungsinformationen, also sog. Metadaten. Im Unterschied zum Bibliotheksbereich, wo bei der schon vor 20 Jahren begonnenen retrospektiven Konversion der Zettelkataloge Regelwerke wie die „Regeln der alphabetischen Katalogisierung“ (RAK) oder die „Preußischen Instruktionen“ (PI) für eine gewisse Einheitlichkeit und Normierung sorgen, existieren solche Richtlinien für die archivische Verzeichnung nur in sehr begrenztem Maße. Sowohl der Umfang der enthaltenen Informationen pro Titelaufnahme (Verzeichnungstiefe), als auch das Layout etwa einer Findbuchseite sind jeweils, mitunter von Findmittel zu Findmittel, unterschiedlich. Dementsprechend flexibel muss das zu entwickelnde Werkzeug sein.

Das Retrokonversions-Tool soll aus Textdateien oder eingescannten maschinenschriftlichen Findbuchseiten anhand von Schlüsselwörtern, Schriftauszeichnungen, den verwendeten Formatvorlagen oder Layout-Merkmalen die einzelnen Erschließungsinformationen (z. B. den Titel, die Signatur oder die Laufzeit) als solche erkennen und strukturiert und verlustfrei in ein Austauschformat einlesen. Bei diesem Austauschformat wird es sich um eine in XML programmierte DTD handeln, deren Elementstruktur sich zum einen an den in den Findbüchern vorzufindenden Informationen orientiert. So muss sie in der Lage sein, Klassifikationshierarchien genauso abzubilden wie zum Beispiel Bandserien. Sowohl die hierarchischen als auch die relationalen Zusammenhänge innerhalb eines Bestandes müssen verlustfrei in das Format eingebunden werden können. Zum anderen aber soll sie nicht proprietär, sondern so allgemeingültig aufgebaut sein, dass sie den Datenaustausch mit möglichst vielen Datenbanksystemen mit möglichst unterschiedlichen Datenstrukturen ermöglicht und insofern als Standard auch für den Austausch zwischen Archivdatenbanksystemen dienen kann. In der derzeitig aktuellen Version unterstützt das Format Sach- sowie die sog. Massen- oder Fallakten, Urkunden, Siegel, Karten und Pläne, Fotos, Plakate und Filme.

Zu Beginn einer Retrokonversionsaufgabe muss das Tool auf die Struktur des zu konvertierenden Findmittels parametrisiert werden. Dies muss grundsätzlich auf Anwenderebene, also etwa menügesteuert oder durch „drag & drop“ geschehen können, lediglich für Adaptierungen größeren Umfangs müssen Eingriffe in den Programmcode möglich sein. Die einmal eingerichteten Vorlagen sollen abgespeichert werden und für künftige Konversionsvorhaben als Mustervorlagen zur Verfügung stehen. Mit fortschreitender Zahl von Retrokonversionsvorgängen werden so immer weniger Neudefinitionen, sondern in wachsendem Ausmaß nur noch Adaptionen bestehender Vorlagen nötig. Die eigentliche Konversion, das heißt die Erkennung und Strukturierung der Erschließungsdaten, sollte im günstigsten Fall automatisch, in der Regel aber zumindest halbautomatisch vor sich gehen. Für besonders schwierige Fälle soll ein vollständig manueller Modus vorgesehen werden.

Aufwand und Kosten der Retrokonversion eines Findmittels sollen dabei ermittelt und festgehalten werden, um als Datenbasis für künftige, etwa drittmittelfinanzierte Retrokonversions-Vorhaben Anwendung finden zu können.

Im Frühjahr 2002 wird ein erster Prototyp des Retrokonversionstools zur Verfügung stehen. Während den sich daran anschließenden sechs Monaten wird diese Software dann in Zusammenarbeit zwischen den Projektmitarbeitern und den Programmierern weiterentwikkelt und fertiggestellt. Während des darauf folgenden Echtbetriebs sollen bis Projektende im April 2003 etwa jeweils 150 maschinenschriftliche und als Textdateien vorliegende Findmittel konvertiert und der dazu nötige Kostenaufwand ermittelt werden. Das Ergebnis stellen Findmittel in Form von XML-Instanzen dar, die problemlos in Datenbanken konvertiert oder für eine online-Präsentation im Internet aufbereitet werden können.

Auch die Möglichkeiten einer Retrokonversion handschriftlicher Findmittel sollen erprobt werden. Unter handschriftlichen Findmitteln verstehen wir solche in lateinischer oder deutscher (Sütterlin) Schrift, die in der Regel in mehreren verschiedenen Händen abgefasst sind sowie auch maschinenschriftliche Findmittel mit einer großen Anzahl handschriftlicher Ergänzungen. Grundsätzlich stehen hier drei Wege zur Verfügung, die im Rahmen einer Machbarkeitsstudie Anwendung finden sollen: 1. Umwandlung eingescannter Findbücher in

Textdateien mittels einer ICR-Software; 2. Ablegen der Findmittel als Images mit anschließender manueller Indizierung für die Recherche; 3. Abschreiben der Findmittel in Eigenleistung oder durch kommerzielle Dienstleister. Alle drei Wege sollen anhand eines Samples von 30 handgeschriebenen Findmitteln auf Machbarkeit und Rentabilität hin geprüft werden.

Nach Abschluss des Projekts im Frühjahr 2003 soll ein Tool zur Verfügung stehen, das eine schnelle Konvertierung der wichtigsten – oder vielleicht auch aller – Findmittel eines Archivs problemlos und unter genauer Kenntnis des zu erwartenden Kosten- und Personalaufwands ermöglicht. Dabei wird das Tool so flexibel sein, dass es auf die Findmittel verschiedener Archivgutarten adaptiert werden kann. Das Teilprojekt „Retrokonversion handschriftlicher Findmittel" wird Erkenntnisse über die Möglichkeiten und Wege in diesem sicherlich problematischsten Bereich bereitstellen.

Das DFG-Projekt zur „Entwicklung von Werkzeugen zur Retrokonversion archivischer Findmittel" stellt somit nur einen ersten Schritt dar auf dem Weg zu einer flächendeckenden Digitalisierung von Findmitteln und deren Bereitstellung im Internet. Übergeordnetes Ziel ist es, den Weg der Öffnung der Archive, ihre Positionierung in der modernen, digitalen Informationsgesellschaft, bei der man in den letzten Jahren in Deutschland schon ein gutes Stück vorangekommen ist, fortzusetzen.

Projet DFG de développement d'outils modernes pour la rétroconversion d'instruments de recherche d'archives

Matthias Meusch

Dès la fin des années 90, les archives de Rhénanie-du-Nord - Westphalie ont commencé à proposer aussi à leurs chercheurs des possibilités de recherche dans leurs fonds et collections en ligne par l'internet. L'un des premiers projets financés par le consortium de recherche Deutsche Forschungsgemeinschaft (DFG) a débouché sur la création du portail internet «NRW Archive im Internet» (http://www.archive.nrw.de/), qui permet d'accéder aux fonds de plus de 400 services d'archives de différents ressorts (services d'archives de l'État et communaux, archives de l'Église, archives économiques, etc.).

La deuxième étape consiste maintenant à pouvoir rechercher aussi sur l'internet des instruments de recherche en grand nombre. Les instruments de recherche en ligne utilisés jusqu'à présent ont été pour l'essentiel convertis à partir de bases de données, ou bien entrés péniblement à la main, ou encore présentés sous la forme d'images qu'il fallait ensuite indexer manuellement. L'énorme majorité des instruments de recherche, principalement ceux qui sont le plus souvent utilisés, sont écrits à la main ou dactylographiés ou, dans le meilleur des cas, sous la forme de fichiers-textes, qui ne se prêtent pas à une utilisation en base de données ni en ligne. Or, compte tenu de cette masse volumineuse, une conversion purement manuelle est impossible. Aussi, le DFG soutient-il, depuis le 1er mai 2001, un projet pilote qui vise à développer un outil logiciel qui devrait permettre d'automatiser le plus possible cette rétroconversion. Les quatre services d'archives publiques de Rhénanie-du-Nord - Westphalie, en l'occurrence le Rheinisches Archiv- und Museumsamt, le Westfälisches Archivamt — les archives municipales de Rhénanie-du-Nord - Westphalie sont également représentées par le biais de ces deux services —, les Westfälisches Wirtschaftsarchiv et les Archives fédérales, sont associés au projet, placé sous la tutelle des archives publiques centrales du *Land,* à Düsseldorf.

Le présent projet n'a pas pour but de numériser des archives, mais seulement les informations en permettant l'exploitation, c'est-à-dire les métadonnées. À la différence des bibliothèques, où un arsenal réglementaire, dont notamment les «Règles de la classification alphabétique» (RAK) ou les «Instructions prussiennes» (PI), veille à une certaine uniformité et normalisation, de telles directives n'existent encore qu'en quantité très limitée pour la description des archives. Non seulement le volume d'informations contenues (profondeur de description) dans chaque titre répertorié, mais aussi la disposition d'une page de répertoire numérique, par exemple, sont très différents d'un instrument de recherche à l'autre. Aussi l'outil à développer doit-il être souple.

À partir de fichiers-textes ou de pages de répertoires numériques dactylographiées scannées, l'outil de rétroconversion doit pouvoir, en utilisant des mots clés, des codes, les formats utilisés ou les caractéristiques de mise en page, reconnaître les informations d'exploitation (titre, signature, délai, etc.) en tant que telles et les convertir dans un format d'échange de façon structurée et sans aucune perte. En ce qui concerne ce format d'échange, il s'agira d'une définition de type de documents (DTD) programmée en XML, dont la structure des éléments se rapproche des informations contenues dans les répertoires numériques. Cette DTD doit être en mesure de pouvoir représenter une hiérarchie de classement exactement comme des séries de volumes, par exemple. Le format doit pouvoir intégrer les liens hiérarchiques et les liens relationnels à l'intérieur d'un fonds, sans perte d'informations. Par ailleurs, cette DTD ne doit pas être protégée par une exclusivité, mais doit être conçue dans l'intérêt général pour permettre d'échanger des données avec le plus grand nombre possible de systèmes de bases de données et avec des structures de données les plus diverses qui soient; si possible, elle doit aussi servir de normes pour échanger des données entre des systèmes d'information d'archives. Dans sa version actuelle, le format supporte des dossiers et des actes juridiques, des cachets, des cartes, des plans, des photographies, des affiches et des films.

Avant d'entamer la procédure de rétroconversion, l'outil doit être paramétré selon la structure de l'instrument de recherche à convertir. L'entrée des paramètres doit en principe s'effectuer au niveau de l'utilisateur, c'est-à-dire grâce à des menus ou par déplacement de la souris, les interventions dans le code ne pouvant être admises que pour des volumes de paramètres plus importants. Les paramètres une fois entrés sont mémorisés et utilisés à titre de modèles pour les travaux de conversion ultérieurs. Plus le nombre de procédures de rétroconversion augmente, plus le nombre de nouvelles définitions diminue, cependant que le nombre d'adaptations des paramètres existants continue de progresser. La conversion proprement dite, c'est-à-dire la reconnaissance et la structuration des données d'exploitation, se fait automatiquement dans le meilleur des cas, et au moins semi-automatiquement en règle générale. Un mode entièrement manuel doit être prévu pour les cas les plus difficiles. Le temps nécessaire et le coût financier de la rétroconversion d'un instrument de recherche doivent être calculés et notés, pour servir de référence à des projets de rétroconversion ultérieurs, notamment des projets financés par des tiers.

Un prototype de cet outil sera disponible au printemps 2002. Dans les six mois qui suivront, le logiciel sera perfectionné et finalisé en coopération entre les collaborateurs du projet et les programmeurs. Entre la période d'exploitation grandeur réelle qui s'ensuivra jusqu'à la fin du projet, prévue en avril 2003, 150 instruments de recherche environ, dactylographiés ou sous forme de fichiers-textes, seront convertis, cependant que l'investissement nécessaire à cette opération sera calculé. Au final, les instruments de recherche se présenteront sous la forme d'instances XML qui pourront être converties sans problème dans des bases de données ou préparées en vue d'une présentation en ligne sur l'internet.

Le projet prévoit également d'expérimenter la possibilité d'une rétroconversion d'instruments de recherche manuscrits. Par instruments de recherche manuscrits, on entend des instruments écrits en caractères latins ou allemands (Sütterlin) qui sont généralement rédigés à plusieurs mains, mais aussi des instruments de recherche dactylographiés comportant un grand nombre d'ajouts manuscrits. En principe, trois voies sont possibles et applicables dans le cadre d'une étude de marché: 1) transformation des répertoires scannés en fichiers-textes à l'aide d'un logiciel de reconnaissance intelligente de caractères (RIC); 2) conversion

de l'instrument de recherche sous forme d'images, suivie d'une indexation manuelle pour la recherche; 3) copie de l'instrument de recherche en interne ou par des prestataires extérieurs. La faisabilité et la rentabilité des trois solutions seront analysées à partir d'un échantillon d'une trentaine d'instruments de recherche manuscrits.

À l'issue du projet au printemps 2003, un outil devrait pouvoir être proposé, lequel permettra de convertir rapidement les instruments de recherche les plus importants — et pourquoi pas la totalité — d'un service d'archives, sans aucun problème et en sachant précisément ce qu'il en coûtera en moyens financiers et humains. L'outil sera souple de façon à permettre son adaptation aux instruments de recherche de tous les types d'archives. L'un des volets du projet, intitulé «Rétroconversion d'instruments de recherche manuscrits», permettra d'apporter des connaissances sur les possibilités et les moyens à disposition dans ce domaine, assurément le plus problématique.

Aussi, le projet DFG de développement d'outils de rétroconversion d'instruments de recherche d'archives n'est-il qu'une première étape en direction de la numérisation généralisée des instruments de recherche et de leur communication sur l'internet. L'objectif suprême est de continuer sur la voie de l'ouverture des archives et de leur ancrage dans une société de l'information numérique moderne, une voie sur laquelle l'Allemagne s'est déjà bien engagée ces dernières années.

The project on digitisation of archival photo documents and accommodating them on the Internet

Gennadij Zalaev ([1])

Gennadij Zalaev

Dr Zalaev, born in 1947, is a Doctor of computer science, correspondent member of the International Academy of Information, associate professor at the Russian State University of Humanitarian (RGGU), where he lectures a course in electronic archives.
1982 — Head of the Department of Computerised Archive Technologies, Russian State Archives Scientific and Technical Documentation.
1979–1982 — senior investigator at the International Centre for Scientific and Technical Information.
1995–96 — Expert of the project 'Study for the development of the information services market in Russia', Tacis N Telrus 9402.
1999 — Speaker at the 1999 DLM-Forum
2000–01 — Technical director of the web project 'Archives of Russia'
Scientific professional interests — information systems, digitisation of archive documents, Internet, electronic libraries and archives.

([1]) Due to unforeseen circumstances, Gennadij Zalaev could not take part in the DLM-Forum 2002.

I. Premise of the project

The sector of the archive information always represented a rather independent component of the Soviet and Russian information infrastructure. The main reason of it is limitation of access to the archive documents for broad circles of the contributors and citizens. Limitation of access increased with years and its radicals go from the period when the archive service was included into the structure of the Ministry of Inner Affairs of the USSR.

The model of an information potential of Russia can be represented as an iceberg whose large, invisible part is a retrospective information storing in archives and forming the Archive Fund.

The Archive Fund of the Russian Federation represents the set of documents which belong to the state and have political, economic, scientific, welfare or historical significance and numbers more than one billion documents of different kinds: the documents on paper carriers, photos, cinema and video, audio documents etc. These documents cover all the aspects of human activity and all branches of knowledge. The Russian State Archive for Scientific and Technical Documentation is one of the leading Russian archives.

The Russian State Archive for Scientific and Technical Documentation (until 1995, The Russian Research Centre for Space Documentation) is a state archive and research centre. It was created on April 1974 on a decision of the government of the Soviet Union. It is the only Russian archive, which has different kinds of documents on a history of the Russian space-rocket engineering and manned astronautics. The archive formed its funds by documents of the organisations — participants of space activity, documents on space subjects from organisations of mass media, personal archives of the experts on space-rocket engineering and Russian cosmonauts, memoirs of the figures of science and engineering.

The documents of various kinds are stored in archive funds: the reports on the completed research and experimental design projects, cinema, photo, audio-, video and telemetry documents. These archived documents allow visually observe the path of the Soviet and Russian astronautics for a period since 1930 up until the present time, and also development of international cooperation in the field of space researches. For example, the Soviet Union-USA Souz-Apollo Space Programme, the Inter-Cosmos International Space Programme etc. More than 300 000 units of storage make the documentary basis of the astronautics history: its formation, development, basic directions of the various space programs; stages of practical mastering of space and international cooperation in this area.

It is necessary to mark that the certain category of archive documents in Russian archives is practically unavailable because of its physical condition. First of all it applies to unique and especially important documents, and also shabby ancient documents to which usage can damage them.

Thus, we have two aspects hindering usage of the archive information: political aspect and technological aspect.

The processes of democratisation in the public life of Russia and modern information technologies allows the country to proceed from the concept of limitation of access to archival information to the concept of broad access to archive documents.

The modern information technologies allow the realisation of potentials of usage of documents stored in archives. The most interesting and perspective directions for the extension of access to archival documents are the technologies of digitisation, modern CD disks and the Internet.

The Russian State Archive for Scientific and Technical Documentation has conducted a number of practical activities on digitisation of the archive documents and their recording on optical disks in 1993–96. The first of such activities was digitisation of fund of the documents connected to life and activity of the German Chancellor, Joseph Wirt. The complex of these documents is stored at the Russian Center for Storage of Historical Documentary Collections. The Fund contains both paper documents (manuscript, typewriting, printed production, newspaper) and photos. The purpose of activity was to grant to German experts facsimile copies of the documents of this fund in a digital kind.

The second project is digitisation of photo documents from funds of the Russian State Archive for Scientific and Technical Documentation connected with life and activity of the first cosmonaut Uri Gagarin. Within the framework of the project:

— 850 photos were digitised and the array of digital copies on CD disk was formed.

— The database of the descriptions of digital copies of the photo documents was created.

The database contained the descriptions of the photo documents, keywords, place of shooting, date of shooting, colour — black-and-white or colour, type of the carrier (positive, negative) and archive number.

The software of the database is CDS ISIS soft system.

Thus, it was conducted, not only simple digitisation of the documents and recording of their digital copies on CD disk but also the information retrieval system with digital copies was developed. We call this system the electronic archive. The system was created for the purposes of usage. The user of a system received relevant inquiry of the description of documents at realisation of information search in the database and could view digital copies appropriated to the relevant descriptions of the documents on the display screen. It improved the quality of the information service, reduced time of obtaining of the originals of archive documents. However, it did not decide a problem of the extension of access to the archived documents.

In 2000 we developed the conception of the project of presentation of the information on the archived documents on the Internet by the way of electronic Internet — catalogue. In 2001 the conception was upgraded and is oriented on usage of digital copies of photo documents.

II. The purpose of the project

The main purpose of the given project is the creation and submission in the Internet and on CD disks of digital copies of the archive documents on a history of a development of the outer space from a collection of the Russian state archive of the technological documents, and also system development of search. The project is directed on providing of access to archive materials to numerous users, as in Russia, and all over the world. Though the storehouses of many Russian archives are well organized on the basis of traditional technologies, the documents retrieval demands durable efforts directly in archives.

The important area of activity realised within the framework of the project is the creation of information resources in the Internet about development of outer space, based on collections of The Russian State Archive for Scientific and Technical Documentation.

More than 80 000 photos and negatives are stored in this archive. The archive also possesses the largest collection of films and the largest collection of audio documents, including negotiations between side of spacecraft and Mission Control Center.

III. The concept of selection of the documents for the electronic Internet — catalogue and digitisation

The concept determined sequence of scanning of the photo documents. The collection has more than 85 000 photo documents on the history of development of outer space. Therefore, it was necessary to determine what photo documents to scan first of all, what in the second queue, in a

third etc. We should determine the order of scanning of the photo documents (photos, negatives, slides etc.). The following queue of scanning containing three stages was developed. At the first stage, the most important documents are scanned. The most important documents are those of large historical significance.

In connection with that the descriptions of photo documents are entered in the electronic Internet — catalogue for maintenance of broad access, it was decided at the second stage to scan the documents having international significance. Funds of the Russian State Archive for Scientific and Technical Documentation contain the documents connected with activity of various international programs on research of space: flight of docked spacecraft and docking in space of USSR and US spacecraft, flights of international space crews under the Inter Cosmos, program flights including astronauts of various countries. The photo documents on the Souz–Apollo Programme were selected for scanning from these complexes of the archive documents. The documents cover a period from 1971 to 1975.

At the third stage we have decided to scan photo documents, which we took on the basis of their descriptions from the database of the History of Cosmos Research Information Retrieval System by way of their input in the database since first. Naturally, we excluded those documents, which were scanned on first and at the second stages.

IV. The characteristic of stages of the project

- The first stage is of activity of the project

At the first stage 464 black-and-white and colour photo documents were scanned. There are photos, negatives and slides.

Examples of the descriptions of the especially valuable photo documents:

Example 1.

Cosmonaut Uri Gagarin with his daughter Elena

Example 2.

Korolev Serge (portrait). 1935.

It is an early snapshot of the chief designer of space-rocket systems Korolev.

Example 3.

Cosmonaut Uri Gagarin (portrait). 1961.

- The second stage is of activity of the project

At the second stage were scanned about 450 photo documents of the Souz-Apollo mission. There are photos, negatives and slides.

Examples of the descriptions of the especially valuable photo documents:

Example 1.

Flight and docking of the spacecrafts ‘Souz’ and ‘Apollo’.

Example 2.

Joint meeting of the Soviet and American experts in the field of space research to the ‘Souz-Apollo’ programme at Johnson Space Centre.

- The third stage of activity — selection from the database

The photo documents were selected as follows at the third stage. We printed out archive numbers of the photo documents, which descriptions are contained in the database of the History of Cosmos

Research Information Retrieval System. Then archive numbers of the photo documents digitised on the first and second stages were eliminated from this list. Thus, the formed list of archive numbers actuates archive numbers of the non-digitised documents by way of input of their descriptions in the database of the History of Cosmos Research Information Retrieval System. The portion of photo documents for scanning was formed on the basis of this list.

The concept of a building of a system

The system consists of two blocks. The first block is array of the digitised copies of photo documents (photos, negatives, slides). One digital copy is one file. A filename, which is assigned to a digital copy, is equivalent to archive number of the photo document, i.e. number under which the file is stored in archive.

The second block is a retrieval system. The retrieval system contains the descriptions of photo documents, search indexes, control instruction and Help. The descriptions of the photo documents contain following elements:

— archival number
— keywords
— headings
— date of shooting.

The search is possible on any element.

The series review of all digital copies is possible on collection option. The navigating of review allows viewing of digital copies, being transferred 'forwards and back'.

As a result of search we receive description of the documents relevant to inquiry and their digital maps.

The archival number is a connection element (key) between the array of digital copies and retrieval system.

The system exists in two kinds:

1. electronic internet — catalogue placed on a site of the Russian State Archive of the technological documentation.
2. electronic fund of use on the CD disk.

Their difference consists in following.

Digital copy of the electronic Internet — catalogue has the much smaller sanctions and accordingly smaller sizes of files. It was made so to prevent unauthorised copying of digital copies of photo documents in commercial purposes for example, to use at preparation of issuing.

The users working in the Russian State Archive for Scientific and Technical Documentation have charge-free access to a system of electronic fund.

Realisation of the project

Within the framework of the project the electronic catalogue including now 3 000 of the most interesting photos and their description from a collection of archive is created.

The activities were conducted under the following scheme deciding three problems:

1. The first problem is the formation of the fund for safety maintenance.

The photos from negatives were scanned with the sanction 600 dpi in the format TIF. The limitation of the size up to 1.5 Mbytes for each digital copy was specially made. The TIF format was

selected to exclude losses of quality inherent to the JPEG format. It allows maximum approach a digital copy to a first copy. On the basis of these digital copies the fund for safety maintenance was formed.

It is necessary to mark, that in Russia now there is no standard on organisation of fund of insurance copying on the basis of digital copies of the archival documents, therefore we use the term 'fund for safety maintenance'.

For the accounting and managements of digital copies special software were developed.

2. The second problem is formation of the electronic fund for use.

The digital copies from the TIF format with usage of the PhotoShop software were converted to the JPEG format with the sanction of 300 dpi.

3. The third problem is the formation of the electronic internet catalogue.

For the Internet version the sanction of each digital copy with the help of PhotoShop software was decreased up to 40 dpi. The digital copies were placed on a web site in the system of the electronic Internet catalogue.

For scanning of the photo documents the following equipment was used:

— the Cano Scan FS2710 scanner for scanning of negatives;

— the Umax Astra 4000U scanner for scanning of negatives and photos.

Seven experts from various divisions of the Russian State Archive for Scientific and Technical Documentation participated in the project.

The activities on creation of a photo of the catalogue were financed by Unesco.

The electronic Internet catalogue is placed in the Internet:
http://rgantd.ru/elcatalog/photocat.htm

V. Perspective

1. In 2002 activities on the creation of the English version of the electronic catalogue with usage of the computer translation system with subsequent editing of texts should begin. The analysis has shown that usage of such methodology for translation of the volumetric database on the English language gave acceptable results, from the point of view of quality, at a rather low cost of activities.
2. Continuation of activities on a loading operation in a system of the electronic Internet catalogue. It is supposed to digitise 1 000 photo documents.
3. The system of the electronic Internet catalogue of the descriptions of the film documents is developed. This catalogue contains only descriptions of the film documents on a history of a development of the outer space, which are stored in the Russian State Archive for Scientific and Technical Documentation. The subjects of the film documents are very diverse and actuate shootings on board spacecraft, shooting from spacecrafts etc.

The realisation of full (in Russian and English languages) version of the electronic catalogue with digital copies of the photo documents will allow to organise broad access of the European citizens, scientists and representatives of mass media to the archive photo documents on a history of a development of the outer space in the USSR and Russia.

Proyecto de digitalización de archivos de documentos fotográficos y su instalación en Internet

Gennadi Zalaev ([1])

(1) Debido a circunstancias imprevistas, el Sr. Zalaev no ha podido participar personalmente en el Foro DLM 2002.

El potencial de información en Rusia es un verdadero iceberg, con una gran parte sumergida constituida por información retrospectiva, almacenada en archivos.

Las modernas tecnologías de información permiten explotar el potencial de consulta de estos archivos. Las técnicas de digitalización, los CD e Internet son las perspectivas más interesantes y prometedoras para ampliar el acceso a estos documentos archivados.

El proyecto presentado tiene como objetivo principal crear e introducir en Internet y en CD copias digitales de los archivos de los distintos fondos de los archivos nacionales rusos de documentación científica y técnica, así como desarrollar un sistema de búsqueda documental. El proyecto tiene por objeto proporcionar acceso al material de archivo a numerosos usuarios, en Rusia y en el resto del mundo. Si bien los almacenes de los archivos rusos están bien organizados según las tecnologías tradicionales, la recuperación de documentos exige esfuerzos continuados directamente en los archivos.

La creación de recursos documentales en Internet relativos al desarrollo de la investigación espacial, a partir de las colecciones de archivos nacionales rusos de documentación científica y técnica (Centro de investigación y documentación espacial rusa hasta 1996) constituye un ámbito principal de actividad en el marco del proyecto. Este archivo constituye el mayor fondo de archivos documentales, vídeos, fotografías y archivos sonoros que describen la historia de la evolución de la conquista espacial y la construcción de naves espaciales en Rusia.

Las actividades comenzaron prácticamente en 1996, cuando se aprobó la solución de digitalizar los documentos fotográficos que describen la vida y la actividad del primer cosmonauta del planeta, Yuri Gagarin. En el marco del proyecto se digitalizaron y grabaron en CD 850 fotografías. Además del motor de búsqueda, se elaboró una base de datos que contiene las descripciones de las fotografías, las palabras clave, los lugares de toma, la fecha de toma de las fotografías, el método (color o blanco y negro), el tipo de soporte y el número de archivo.

El programa informático utilizado fue el sistema ISIS. En los archivos hay almacenadas más de 80 000 fotografías y negativos y unas 3 500 películas. El archivo también posee una amplia colección de documentos sonoros, incluidas las conversaciones entre la tripulación y el centro de control.

El proyecto abarca también la creación de un catálogo electrónico, que incluye en una primera fase 3 000 fotografías de las más interesantes con su descripción, seleccionadas en una colección.

Las actividades se llevaron a cabo aportando respuestas a los tres siguientes problemas:

1. Primer problema: constitución de un «fondo de seguridad»

Las fotografías de los negativos se escanearon con una resolución de 600 dpi en formato TIFF (copias digitales «originales»). Cada copia se limitó voluntariamente a un tamaño de fichero de un máximo de 1,5 megabites. Este formato se escogió para evitar las pérdidas de calidad inherentes al formato JPEG y para acercarse al máximo al ejemplar original. El «fondo de seguridad» se constituyó a partir de estas copias digitales «originales». Es importante destacar que en Rusia no existen normas que regulen la organización de un fondo de seguridad,

constituido a partir de la copia digital de documentos archivados; por eso utilizamos este concepto de «fondo de seguridad».

Se elaboró un *software* especial para el archivo de estas copias digitales y su gestión.

2. Segundo problema: constitución de un fondo de explotación

Las copias digitales en formato TIFF, destinadas a su utilización con el programa PhotoShop, se transfirieron (convertidas) al formato JPEG, con una resolución de 300 dpi.

Se elaboró un *software* de utilización basado en el sistema ISIS para la búsqueda documental. El sistema de búsqueda y la organización de las copias digitales se registran en un CD.

3. Tercer problema: constitución de la versión de Internet

La resolución elegida para cada copia digital se redujo a 40 dpi para la versión de Internet. Los ejemplares digitales se colocaron en un sitio web en un catálogo electrónico.

El equipamiento utilizado fue el siguiente: escáner CanoScan FS2710 para los negativos y escáner Umax Astra 4000U para negativos y fotografías.

La Unesco financió las actividades relativas a la creación de un catálogo de fotografías. Este catálogo está disponible en Internet (http://rgantd.ru/elcatalog/photocat.htm) y su versión en CD se está preparando.

Este año comenzarán, en principio, las actividades de creación de la versión inglesa del catálogo electrónico, con ayuda de un sistema de traducción automática y de un módulo de edición de textos. El análisis puso de manifiesto que la aplicación del método de transferencia (traducción) de esta base de datos voluminosa en lengua inglesa daba resultados aceptables desde el punto de vista de la calidad, a un coste bastante reducido.

La realización de la versión íntegra del catálogo electrónico y de las copias digitales de las fotografías permitirá un amplio acceso de los ciudadanos europeos, científicos y representantes de los medios de comunicación a estos archivos visuales, que informan acerca de la historia de la conquista espacial en Rusia.

Das Projekt der Digitalisierung von archivierten Fotodokumenten und ihrer Bereitstellung im Internet

Gennadi Zalaev [1]

Das Informationspotential in Russland kann man sich als Eisberg vorstellen, dessen großer, nicht sichtbarer Teil die in den Archiven lagernden historischen Informationen sind.

Mit Hilfe der modernen Informationstechnologien können die Potentiale erschlossen und die in den Archiven aufbewahrten Dokumente genutzt werden. Die interessantesten und zukunftsträchtigsten Wege für einen breiteren Zugang zu diesem Archivgut bieten sich mit der Digitalisierung, modernen CDs und dem Internet.

(1) Herr Zalaev konnte wegen unvorhersehbarer Umstände nicht am DLM-Forum teilnehmen.

Hauptanliegen dieses Projekts ist die Erstellung digitaler Kopien von Archivgut, das sich in der Sammlung des russischen Staatsarchivs für wissenschaftlich-technische Dokumentation befindet, und deren Speicherung im Internet oder auf CD, sowie die Entwicklung eines Systems für die Informationsrecherche. Dabei geht es darum, Archivmaterialien einer Vielzahl von Benutzern in Russland und weltweit zugänglich zu machen. Obgleich die Bestände vieler russischer Archive auf der Grundlage herkömmlicher Technologien gut organisiert sind, ist das Wiederauffinden von Dokumenten dort überaus aufwändig.

Im Mittelpunkt des Projekts steht die Einrichtung von Informationsressourcen im Internet zur Entwicklung der Weltraumforschung auf der Basis von Beständen des russischen Staatsarchivs für wissenschaftlich-technische Dokumentation (bis 1996 Russisches Forschungszentrum für Weltraumdokumentation). Dieses Archiv verfügt über den größten Bestand an Dokumentar- und Videofilmen, Fotos und Tonaufzeichnungen zur Geschichte der Weltraumforschung und der Raumfahrttechnik in Russland.

Die praktischen Arbeiten begannen 1996, als das Konzept zur Digitalisierung der Fotodokumente über Leben und Wirken des ersten Menschen im Weltall, Juri Gagarin, genehmigt wurde. Im Rahmen dieses Projekts wurden 850 Fotos digitalisiert und auf CD gespeichert. Als Suchhilfsmittel wurde eine Datenbank entwickelt, die die Beschreibungen der Fotos, Stichwörter, Ort und Datum der Aufnahme, Art des Fotos (farbig, schwarz-weiß), Art des Datenträgers und Archivnummer enthielt.

Als Software wird das ISIS-System verwendet.

Im Archiv lagern mehr als 80 000 Fotos und Negative sowie etwa 3500 Filme. Außerdem besitzt das Archiv eine umfangreiche Sammlung von Tondokumenten, u. a. zum Funkverkehr zwischen Raumschiff und Flugleitzentrum.

Im Zuge des Projekts wird ein elektronischer Katalog erstellt, in den in der ersten Phase die 3000 interessantesten Fotos und ihre Beschreibung aus einem Teil der Archivbestände aufgenommen werden.

Die Aktivitäten wurden nach folgendem Schema zur Lösung von drei Problemen durchgeführt:

1. Erstes Problem: Bildung eines Sicherheitsbestands.

Die Fotos von Negativen wurden im TIFF-Format mit einer Auflösung von 600 dpi (mit je einer digitalen Kopie) gescannt. Dabei wurde eigens eine Begrenzung der Größe auf 1,5 MB für jede digitale Kopie festgelegt. Das Format wurde ausgewählt, um Qualitätsverluste auszuschließen, wie sie beim JPEG-Format auftreten, und um bei der digitalen Kopie eine optimale erste Kopie zu erhalten. Mit den digitalen Kopien desselben Typs wurde ein Sicherheitsbewahrungsbestand gebildet. Hierzu ist anzumerken, dass es in Russland gegenwärtig keine Norm für die Organisation von Sicherungsbeständen zum Kopieren auf der Basis von digitalen Kopien von Archivdokumenten gibt; deshalb verwenden wir den Begriff „Sicherheitsbewahrungsbestand“.

Für die Ablage (Aufbewahrung) und Verwaltung der digitalen Kopien wurde eine spezielle Software entwickelt.

2. Zweites Problem: Bildung eines Nutzungsbestands.

Die digitalen Kopien wurden mit Hilfe des Programms PhotoShop aus dem TIFF-Format in das JPEG-Format mit einer Auflösung von 300 dpi umgewandelt.

Für das Informationsretrieval wurde ein System auf der Grundlage des ISIS-Systems entwikkelt. Retrievalsystem und digitale Kopien wurden auf CD gespeichert.

3. Drittes Problem: Bildung der Internetversion.

Für die Internetversion wurde die Auflösung der digitalen Kopien auf 40 dpi verringert. Die digitalen Kopien wurden in Form eines elektronischen Katalogs auf eine Website gestellt.

Folgende Geräte wurden verwendet: ein Scanner vom Typ CanoScan FS2710 für das Einscannen der Negative und ein Scanner vom Typ Umax Astra 4000U für das Einscannen von Negativen und Fotos.

Die Erstellung des Fotokatalogs wurde von der Unesco finanziert. Dieser Katalog befindet sich im Internet (http://rgantd.ru/elcatalog/photocat.htm) und wird auch auf CD vorbereitet.

In diesem Jahr sollen die Arbeiten zur Erstellung einer englischsprachigen Version des elektronischen Katalogs mittels automatischer Maschinenübersetzung mit nachfolgender redaktioneller Bearbeitung der Texte beginnen. Untersuchungen haben gezeigt, dass bei Einsatz dieser Methode für die Übertragung (Übersetzung) der riesigen Datenbank in die englische Sprache relativ kostengünstig eine annehmbare Qualität erzielt werden kann.

Mit der Erstellung der vollständigen Version des elektronischen Katalogs mit digitalen Kopien der Fotodokumente wird es möglich sein, den europäischen Bürgern, Wissenschaftlern und Vertretern der Massenmedien einen breiten Zugang zu den Archivfotobeständen zur Geschichte der Weltraumforschung in Russland zu gewähren.

Projet de numérisation d'archives de documents photographiques et installation sur l'internet

Gennadi Zalaev ([1])

Le potentiel d'informations en Russie est un véritable iceberg, dont la vaste partie immergée est constituée d'informations rétrospectives stockées dans les archives.

Les technologies modernes de l'information permettent d'exploiter le potentiel de consultation de ces archives. Les techniques de numérisation, les CD et l'internet sont les perspectives les plus intéressantes et les plus prometteuses pour élargir l'accès à ces documents archivés.

Le projet présenté a pour principal objectif de créer et de verser sur l'internet et sur CD les copies numériques des archives des divers fonds des Archives nationales russes de la documentation scientifique et technique, ainsi que de développer un système de recherche documentaire. Le projet est axé sur la possibilité, pour les chercheurs de Russie et du monde entier, d'avoir accès à ces archives. Même si les magasins des différents services d'archives russes sont parfaitement organisés selon les technologies traditionnelles, la recherche documentaire nécessite des efforts soutenus directement dans ces services.

La création de ressources documentaires sur l'internet concernant le développement de la recherche spatiale, à partir des collections des archives nationales russes de la documentation scientifique et technique (centre pour la recherche et la documentation spatiales russes jusqu'en 1996), constitue un domaine d'activité majeur dans le cadre du projet. Ce service est le plus gros fonds d'archives de documentaires, de films vidéo, de photographies et d'ar-

([1]) À cause de circonstances imprévues, M. Zalaev n'a pas pu participer personnellement au Forum DLM 2002.

chives sonores retraçant l'histoire de l'évolution de la conquête spatiale et de la construction de vaisseaux spatiaux en Russie.

Les activités pratiques ont démarré en 1996, lorsque le choix de numériser les documents photographiques retraçant la vie et l'activité du premier cosmonaute de la planète, Youri Gagarine, a été approuvé. Au total, 850 photographies ont été numérisées et enregistrées sur des disques CD dans le cadre du projet. Outre le moteur de recherche, une base de données a été élaborée, qui renferme les descriptions des photographies, les mots clés, les lieux de prise de vue, leur date, le mode de tirage (couleurs ou noir et blanc), le type de vaisseau et le numéro d'archives.

Le logiciel utilisé est le système ISIS. Plus de 80 000 photographies et négatifs et 3 500 films environ sont entreposés dans les archives. Celles-ci possèdent également une vaste collection de documents sonores, incluant les conversations entre l'équipage et le centre de contrôle.

Le projet porte également sur la création d'un catalogue électronique, regroupant dans un premier stade 3 000 photographies parmi les plus intéressantes et leur description, sélectionnées à partir d'une collection.

Les opérations ont été menées en apportant des réponses aux trois problèmes décrits ci-après:

1) Premier problème: constitution d'un «fonds de sécurité»

Les photographies des négatifs ont été scannées avec une résolution de 600 dpi au format TIF (copies numériques «originales»). Chaque copie a été volontairement limitée à une taille de fichier d'au maximum 1,5 Mo. Ce format a été retenu pour exclure toute perte de qualité inhérente au format JPEG et pour approcher au maximum l'exemplaire original. Le «fonds de sécurité» a été constitué à partir de ces copies numériques «originales». Il est important de souligner qu'en Russie il n'existe pas de normes régissant l'organisation d'un fonds de sécurité constitué à partir de la copie numérique de documents archivés; c'est pourquoi nous employons cette notion de «fonds de sécurité».

Des logiciels spéciaux ont été élaborés pour l'archivage de ces copies numériques et leur gestion.

2) Deuxième problème: constitution d'un fonds d'exploitation

Les copies numériques au format TIF, destinées à une utilisation par le logiciel PhotoShop, ont été transférées (converties) au format JPEG, avec une résolution de 300 dpi.

Un logiciel d'utilisation reposant sur le système ISIS a été élaboré pour la recherche documentaire. Le système de recherche et l'organisation des copies numériques sont enregistrés sur un CD.

3) Troisième problème: constitution de la version internet

La résolution retenue pour chaque copie numérique a été réduite à 40 dpi pour la version internet. Les exemplaires ont été placés sur un site web dans un catalogue électronique.

L'équipement utilisé a été le suivant: scanner CanoScan FS2710 pour les négatifs et scanner Umax Astra 4000U pour les négatifs et les photographies.

Les activités relatives à la création d'un catalogue de photographies ont été financées par l'Unesco. Ce catalogue est disponible sur l'internet (http://rgantd.ru/elcatalog/photocat.htm), et sa version sur CD est également en cours de préparation.

Cette année devrait voir le lancement des activités de création de la version anglophone du catalogue électronique, à l'aide d'un système de traduction automatique et d'un module d'édition de texte. L'analyse a montré que l'application de la méthode de transfert (traduction) de cette base de données volumineuse en langue anglaise donnait des résultats acceptables du point de vue de la qualité, à un coût assez faible.

La réalisation de la version intégrale du catalogue électronique et des exemplaires numériques des photographies permettra un large accès des citoyens européens, des scientifiques et des représentants des mass media à ces archives visuelles qui relatent l'histoire de la conquête spatiale en Russie.

Markus Junker

Dr Markus Junker is a leading academic at the German Research Centre for Artificial Intelligence (DFKI). He is responsible, inter alia, for the joint project 'Adaptive READ', supported by the German Federal Ministry for Education and Research, which focuses on adaptive systems for the retrieval of knowledge from documents. Dr Junker studied computing from 1989 to 1995 at Kaiserslautern University, where he was awarded a doctorate in 2001 on 'Learning heuristic rules for text categorisation'. Since 1996 he has worked at the DFKI in Professor Dengel's field of research. His main interests are the use of machine-learning processes to support information searches in electronic document collections.

Andreas Dengel

Professor Dr Andreas Dengel is the Academic Director of the German Research Institute for Artificial Intelligence (Deutschen Forschungszentrum für Künstliche Intelligenz — DFKI) and, since 1993, professor for information technology at Kaiserslautern University. Dr Dengel founded Insiders Information Management GmbH. From 1980 to 1986 he studied information technology and economics. He then worked in Siemens' research laboratory in Munich, at Xerox Parc in Palo Alto, California, and at Stuttgart University, where he was awarded his PhD in 1989. Dr Dengel has on many occasions won recognition for his work; for example he has received the prestigious 'Young Investigator Award' from the International Association for Pattern Recognition and the Alcatel. SEL Research Prize for Technical Communication. His academic interests focus on the areas of knowledge management and document analysis.

Innovative solutions for the automatic categorisation and indexing of electronic documents

Markus Junker, Andreas Dengel

1. Introduction

Category systems for structuring document collections are widely used and are generally accepted as a useful instrument for finding documents. Automatic systems for categorisation try to facilitate the assignment of a document through pre-defined and automatically deducible assignment rules. In the past few years, work has been done on so-called learning categorisation processes (also referred to as machine-learning processes) in the areas of information retrieval (IR) and artificial intelligence (AI), which do not require a lengthy manual formulation of assignment rules. The processes are trained with a large number of correctly categorised documents. Basing themselves on the training documents, they are able to assign a document to the correct category automatically. A particularly interesting possible application for this process is automatic indexing. Up to now, automatic indexing processes could only extract descriptors from documents. By contrast, self-learning processes can assign descriptors automatically from any document languages.

In Section 2, we begin by briefly outlining the history of the development of self-learning automatic categorisation processes. Section 3 describes the main ways in which the processes work. The possible applications of the processes in document description in libraries and archives will be discussed in Section 4. Section 5 will present findings about the learning processes which are important for practical applications as well as their possible limitations in archives and libraries. Section 6 sets out our conclusions.

2. The history of machine-learning processes in document categorisation

The use of machine learning processes in document categorisation originated in Information Retrieval, an area within computer science at the beginning of the 1990s (Lewis 1992). During the first five years almost all experiments were conducted using a collection of about 20 000 Reuters newspaper articles in the area of economics. Typical categories corresponded, for example, to commodities (corn, rice etc.). The first experiments used so-called linear classifiers. In 1994, for the first time, a rule learning process, or rather a related method, the so-called decision tree process was applied to the problem (Lewis et al. 1994). At that time there were no significant differences in the degree of precision which could be obtained using the different methods.

From 1995 onwards, there was a strong increase in scientific interest in machine-learning processes for document categorisation. Interest was also aroused in the area of artificial intelligence. Other collections were put together for evaluating the processes. Today, dozens of different processes have been evaluated, mainly statistical learning processes familiar from other applications. There was also a lot of experimentation with various document representations. Unfortunately, it was not possible to compare a large number of works directly, as the experimental conditions often varied incomprehensibly. Nevertheless, in the last three years processes have been compared with one another in an increasingly professional manner, so that today a number of statements may be regarded as accepted (Mladeni_ 99, Yang et al. 99)

It is only recently that machine-learning processes for document categorisation have been used in commercial applications. In particular, these processes have been increasingly integrated into document management systems, where they have been used for sorting out business correspondence and internal publications.

3. Functions

Diagram 1 shows the main scenario in the application of self-learning document categorisation processes. For example, in order to recognise invoices, a certain number of invoices and a certain

number of documents which are not invoices must be made available to the machine-learning processes. Providing sample documents which do not belong to the category to be learned is essential and in practice doing so is not normally a problem. The result of the learning process is a so-called classifier for the invoice category. This is — with a certain degree of precision — able to determine whether a new document is or is not an invoice.

- **Document representation**

The representation of the objects to be categorised in the so-called 'attribute value representation' is essential for the application of self-learning categorisation processes. This requires, firstly, the intellectual determination of attributes through which an object should be represented. For example, in order to categorise vehicles, one could give the number of wheels for each vehicle and the number of seats. By describing sample vehicles one can learn from machine-learning processes, for example, whether a vehicle is or is not a motorbike or whether or not it is a bus.

In order to apply self-learning processes to document categorisation, the documents must be in the form of electronic ASCII texts. In the case of documents in hard copy this means that they must be scanned in first and prepared using optical character recognition (OCR). Then, the documents

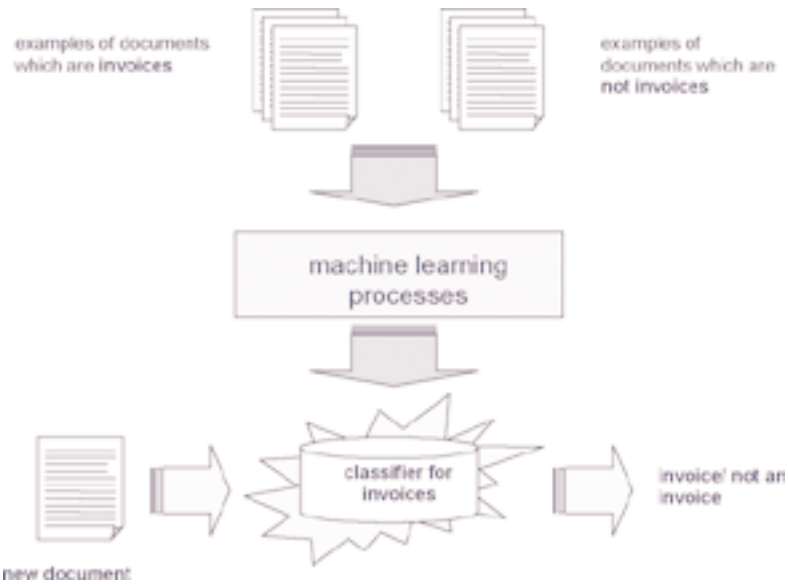


Diagram 1: Scenario where self-learning processes are applied to document categorisation

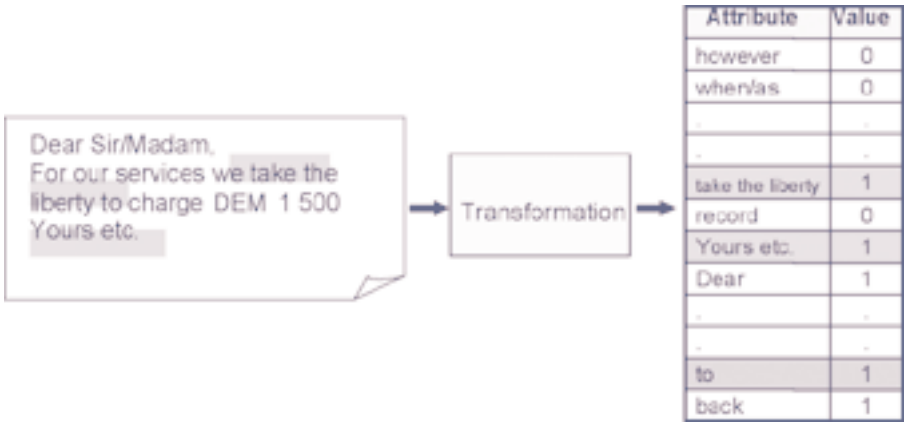
in ASCII format must be transformed into attribute value representation. This requires, first of all, the determination of attributes. Typically a collection of documents is used. For every individual word appearing in these documents an attribute is introduced which may be compared with a full text index such as that used for example by an Internet search engine. Diagram 2 demonstrates what the result of such a transformation can look like using a sample

document (an invoice). The individual words in the original document can be seen to correspond to attributes. The attribute value of a document indicates in each case how often a particular word appears in the document to be represented.

Learning processes

Building on the sample documents in attribute-value representation, a number of machine-learning processes may be used. Two well-known types of learning process are so-called rule learning processes and statistical learning processes (which may be regarded as including the well known neuronal networks).

Diagram 2: Transformation of documents into the attribute value representation



- **Rule-learning processes**

Rule-learning processes generate categorisation rules as classifiers. On the basis of sample documents, they automatically try to generate conditions for attribute values that decide to which category a document is to be allocated. For example, a rule learner may learn a rule in the form:

(calculate>0) or ((take the liberty>0) and (DM>0)) → Invoice

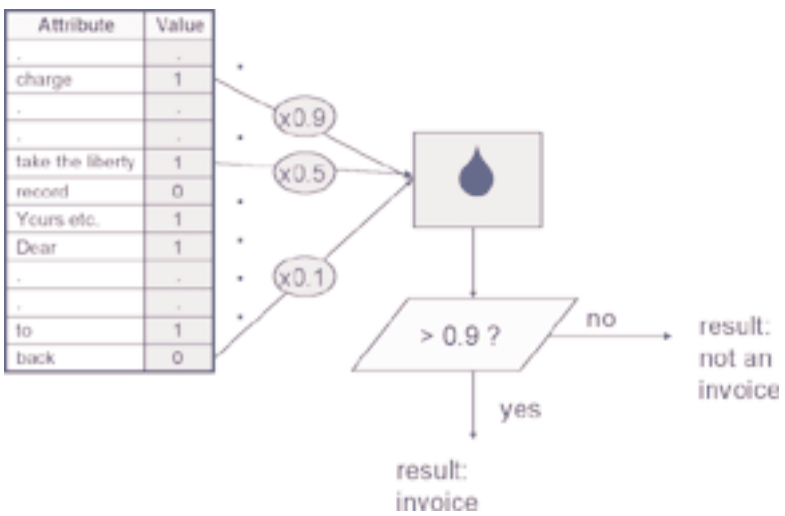
When the concept 'charge' appears at least once in the document to be categorised or the words 'take the liberty' and 'DM' occur at least once together, the document in question is assigned to the 'Invoice' category.

In principle, the classifier for a category may also be indicated as a conventional search expression, e.g. 'charge or (take the liberty and DM)'. More complex tests such as the search for particular word parts, phrases or word roots are generally only possible where these are already provided in the documents' attribute value representation.

- **Statistical learning processes/neuronal networks**

The most simple statistical processes/neuronal networks produce so-called linear classifiers from the examples. Diagram 2 shows the mode of functioning of a learned linear classifier by way of an example. In principle, a linear classifier consists of a weighting factor for every possible attribute (in the diagram only the factors 0.9, 0.5 and 0.1 for the words 'charge', 'take the liberty' and 'back' are represented) and a so-called threshold value (in the diagram the value 0.9 in the rhombus). The concrete attribute values for a document are multiplied by their respective weightings, then added up and the sum is compared with the threshold value. If it is greater than the threshold value, the document is assigned to the corresponding category.

Diagram 3: Composition of a linear classifier



Linear classifiers are the best known statistically derived classifiers. They present a limitation in that the appearance of a word in a document is always given the same weighting in a decision for or against a particular category. It is easy to imagine a scenario in which the weighting given to a word when taking the decision would depend on the appearance or non-appearance of a certain other word. So, for example, the appearance of either the word 'DM' or 'take the liberty' on its own is a weak indication of the category 'Invoice' ('DM often appears on delivery notes, while 'take the liberty' may appear in the most varied documents). However, where 'DM' has already appeared, far greater significance may be attached to the additional appearance of 'take the liberty'. Such connections may be represented by non-linear classifiers, which theoretically allow any offset schemes for attribute values.

4. Possible applications in libraries and archives

The original area of application for self-learning document categorisation systems was the categorisation of newspaper articles using examples. More practically oriented were approaches using document-management systems which often sort a business's incoming post according to contents criteria. As far as we know, processes of this kind have not yet been used for document description in archives and libraries.

An obvious potential use for machine-document categorisation processes is arranging documents into existing category systems. Documents which have already been ordered into systems may serve as examples in the learning processes.

Another, particularly interesting possible application is the extension of conventional automatic document indexing processes. Automatic document indexing has aroused great interest in archives and libraries in recent years. In research projects — in Germany for example there are THEAS, MILOS and Kascade — new automatic indexing processes have been developed which use linguistic approaches for producing canonic word forms and elaborated weighting processes (Lepsky et al. 00). What previous approaches to automatic indexing have in common is however that the descriptors assigned correspond more or less directly to the appearance of a particular term in the original document (e.g. in the title and/or contents). Descriptors with a higher value, e.g. occurring in intelligent manual indexing using controlled vocabularies and thesauri (e.g. the keyword file, SWD in German speaking countries) may not be generated using these processes.

Machine-learning document categorisation processes could be aided by a consideration of the problem of assignment/non-assignment of a particular descriptor — as a categorisation task. Using a number of sample documents, to which the descriptor was either assigned or not assigned in the past, the processes can independently decide on the assignment of a descriptor to a new document. Whether or not a descriptor is assigned is decided by the classifier learned for the descriptor. The number of individual learned properties in the document is the overall deciding factor on the assignment of a descriptor. This leads to a reliable approach in respect of ambiguous expressions in a text. It is not the appearance of an ambiguous concept alone that decides in cases of doubt whether or not a descriptor is to be assigned but rather its appearance in combination with other concepts which make its meaning clear. The same approach may also be used to deal with the problem of synonyms. In assigning a descriptor, a syntactically perfect different, synonymous concept may be used by the learning process in the classifier.

5. Findings to date and open questions

Experimenting with machine-learning document categorisation processes in IR and AI has led to numerous practically relevant findings. The most important of these for potential use of the processes in archives and libraries are set out below.

Experiments have shown that the precision of learning processes depends very much on the category to be learned and the quantity of training examples. Also, results are improved when individual documents have a large quantity of text, i.e. where in addition to the title they contain summaries, contents or even the full document. Where there are enough examples (100 to several thousand positive examples and the same number of negative examples), it is possible to achieve more than 98 % recall and precision.

In particular with statistical classifiers such as the linear classifier, the choice of the threshold value gives rise to a simple possibility of adjusting recall and precision exactly to the application time. Moreover, using so-called sampling processes such as cross-validation, a very exact estimation of recall and precision of the learned classifiers may be achieved. In applications this means that for each category/descriptor suggested by the system the probability of its correctness/ incorrectness may be fairly exactly indicated. The categories/descriptors which are often incorrectly omitted by the system (these have a low recall) may also be found relatively easily.

In IR and AI several experiments have been conducted using various attributes for representing documents. There has been particular interest in the transmission of simple, single-word based attributes with word sequences or particular phrases as well as in the effect of a root form modification. In all cases attributes are determined automatically from document collections or using linguistically derived processes. In particular, there was no manual selection or, still less, assignment of attributes. The experiments conducted showed that the precision of the processes can only be slightly increased by enriching document representation (an exception is the breaking down of compound words in German). It must be taken into account here that most of the experiments were conducted with documents which were at least as long as summaries. The findings can probably only be partially applied to very short texts for representing documents (for example, where only the titles are used in attribute value representation). In these cases one may expect due to a lack of repetition of contents in the texts that more meaningful attributes may be needed in order to achieve greater precision with regard to a category/descriptor.

To achieve a degree of precision, the choice of learning process is more important than the attributes used. Considerable improvements have occurred in recent years in this area. The Support Vector Machine (SVM) has proved to be one of the best learning processes. This is a process derived from learning theories, which learns linear and non-linear statistical classifiers (Joachims 1998). Furthermore, it has demonstrated that linear classifiers are not significantly worse than non-linear. Rule learning processes are still less advanced than statistical processes. However, they have the advantage that it is relatively easy to edit their classifiers manually at a later stage (Junker 2001). The very best results are achieved using hybrid processes which combine the ideas of rule learning with statistical processes (Junker et al. 2002).

In measuring the speed of individual processes one must distinguish between the training and the application stages. As a general observation, rapid training in a learning process leads to longer times required for categorisation/assignment of a descriptor. Rule learning processes require — depending on the number of examples and the category to be learned — from a few seconds to a few minutes per category/descriptor on today's hardware while SVM as a linear learning process is normally considerably faster. It is also important to remember that neither SVM nor rule-learning processes are incremental in the following sense. In order to bring new examples into line with the training material, the learning process must be started from scratch with all the examples. This entails relatively long learning periods. In both rule learning processes as well as linear classifiers a decision about classification/descriptor assignment can be taken in seconds.

As represented earlier, document categorisation learning processes have not yet been used in archives and libraries. Application in this area involves some challenges for the processes in question. Firstly, problems may arise due to what may in practice be an insufficient degree of precision in processes for assigning categories/descriptors. Such problems may be expected particularly where there are not enough examples for a category/descriptor. The problem can be solved by integrating the process into a procedure model in which people will ultimately decide in cases of doubt. A second problem may be the time required for learning processes, particularly in the training period. The processes for learning a category/descriptor are usually relatively quick. However, dealing with very large numbers of categories/descriptors could lead to practical limitations. A solution could be splitting processes so that various institutions, each with their own hardware, take responsibility for an area of the categories/descriptors to be learned.

6. Conclusions

Machine-learning processes for document categorisation represent a fascinating possibility for overcoming the limits of conventional automatic categorisation/indexing processes. Up to now

they have not, as far as we know, been used in archives and libraries. However, extensive preliminary work has been carried out in information retrieval and artificial intelligence which may be drawn upon. The application of the process in archives and libraries however involves some challenges. Given their great potential, the application of the process should be further developed.

Bibliography

(Joachims 1998) Joachims, T., 'Text categorisation with support vector machines: learning with many relevant features', *Proceedings of the 10th European Conference on Machine Learning (ECML 98)*, p. 137–148; Chemnitz, Germany, 21–24. April, 1998.

(Junker 2001) Junker, M., 'Lernen heuristischer Regeln für die Textkategorisierung', Dissertation, Department of Computer Science, University of Kaiserslautern, 2001.

(Junker et al. 2002) Junker, M., Dengel, A., and Abecker, A., 'Text categorisation using learned document features', *Hybrid methods in pattern recognition*, Word Scientific Publishing Co., planned publication, summer 2002.

(Lewis 1992) Lewis, D. D., 'Representation and learning in information retrieval', dissertation, Department of Computer Science, University of Massachusetts, 1992.

(Lewis et al. 1994) 'A comparison of two learning algorithms for text categorisation', *Proceedings of the 3rd Symposium on Document Analysis and Information Retrieval (SDAIR 94)*, pp. 81–93, Las Vegas, NV, United States, 11–13 April 1994.

(Lepsky et al. 2000) Lepsky, K. and Zimmermann, H. H., 'Katalogerweiterung durch Scanning und automatische Dokumenterschließung — Results of DFG-Projekts Kascade', *Zeitschrift für Bibliothekswesen und Bibliographie 47*, Heft 4, 2000.

(Mladeni_ 1999) Mladeni_, D., 'Feature selection for unbalanced class distribution and naive Bayes', *Machine-learning. proceedings of the 16th International Conference (ICML 99)*, pp. 258–267, Bled, Slovenia, 27–30. June 1999.

(Yang et al. 1999) Yang, Y. and Lui, X., 'A re-examination of text categorisation methods', *Proceedings of the 22nd Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval (SIGIR 99)*, pp. 92–49, Berkeley, CA, United States, 15–19. August 1999.

Soluciones innovadoras para la clasificación y la indización automáticas de documentos electrónicos

Markus Junker, Andreas Dengel

Hasta ahora, los métodos de creación de índices automáticos sólo podían extraer descriptores a partir de los documentos. Nuevos procedimientos de autoaprendizaje, desarrollados recientemente en la investigación documental, permiten ya proporcionar automáticamente descriptores en cualquier lengua de documentación. La condición previa es aplicar estos procedimientos a archivos ya dotados de descriptores. Estos procedimientos se desarrollaron en principio para ayudar a la clasificación automática y presentan en este ámbito también posibilidades de aplicación interesantes.

Los sistemas de clasificación que permiten organizar fondos de documentos están muy extendidos y generalmente se consideran un instrumento útil para la investigación documental. Los sistemas automáticos de clasificación intentan facilitar la clasificación de un documento utilizando normas de asignación definidas de antemano y ejecutables automáticamente. La investigación documental y la inteligencia artificial de estos últimos años se han interesado por los procedimientos de clasificación de autoaprendizaje, que funcionan sin necesidad de definir manualmente normas de asignación largas y molestas. Estos procedimientos se han probado en un gran número de documentos correctamente clasificados. A partir de documentos de entrenamiento, son capaces de clasificar automáticamente nuevos

documentos en la categoría adecuada. Internamente, estos procedimientos tratan, calculando la distribución estadística de las palabras, de aislar aquellas propiedades que permitirán asignar el documento a una categoría concreta.

Los sistemas de autoaprendizaje también pueden aplicarse a soluciones innovadoras en la creación de índices automáticos. Los métodos automáticos tradicionales son limitados en el sentido de que sólo pueden utilizar como descriptores términos extraídos de un documento. Estos sistemas no pueden generar descriptores de mayor valor, tal como los que resultan de la indización manual inteligente a partir de glosarios y diccionarios comprobados. Gracias a los sistemas de autoaprendizaje, es posible asignar a un documento cualquier tipo de descriptor. Tal como ocurre con los sistemas de clasificación de autoaprendizaje, también es necesario disponer de un número suficiente de documentos previamente indizados. Los métodos de autoaprendizaje serán entonces capaces de aislar aquellas propiedades del documento que permitirán que se le asigne tal o cual descriptor.

Distintos métodos de aprendizaje, como las redes neuronales, los clasificadores estadísticos y los procedimientos de aprendizaje de normas, pueden utilizarse para la clasificación automática, y también para la indización automática. Estos sistemas son diferentes en su rapidez de entrenamiento y en su velocidad de clasificación o de indización, y también en su precisión y en la posibilidad que ofrecen o no de poder corregir la clasificación y la indización automáticas.

En nuestra exposición presentaremos los resultados más recientes de la investigación en el ámbito de los sistemas de aprendizaje que permiten la clasificación automática, y discutiremos sus potencialidades y sus límites en cuanto a soluciones innovadoras de indización y clasificación automáticas.

Innovative Lösungen zur automatischen Kategorisierung und Indexierung von elektronischen Dokumenten

Markus Junker, Andreas Dengel

Bisher konnten Verfahren zur automatischen Indexierung lediglich aus den Dokumenten Deskriptoren extrahieren. Neue, in den letzen Jahren im Information Retrieval entwickelte selbstlernende Verfahren können automatisch Deskriptoren aus beliebigen Dokumentationssprachen vergeben. Voraussetzung für die Anwendung ist ein Archiv mit bereits intellektuell vergebenen Deskriptoren. Die Verfahren wurden ursprünglich zur Unterstützung der automatischen Kategorisierung entwickelt und haben auch in diesem Bereich interessante Anwendungen.

Kategoriensysteme zur Strukturierung von Dokumentbeständen haben eine weite Verbreitung und sind allgemein als nützliches Instrument zum Auffinden von Dokumenten akzeptiert. Automatische Systeme zur Kategorisierung versuchen die Zuordnung eines Dokumentes durch vorher definierte und automatisch abarbeitbare Zuweisungsregeln zu erleichtern. In den letzen Jahren wurde im Information Retrieval und in der Künstlichen Intelligenz an sogenannten lernenden Kategorisierungsverfahren gearbeitet, die ohne die aufwändige manuelle Festlegung von Zuweisungsregeln auskommen. Die Verfahren werden mit einer Menge korrekt kategorisierter Dokumente trainiert. Basierend auf den Trainingsdokumenten sind sie in der Lage, ein neues Dokument automatisch den entsprechenden Kategorien zuzuordnen. Intern versuchen die Verfahren über die Berechnung von

statistischen Wortverteilungen genau die Eigenschaften von Dokumenten zu isolieren, die für ihre Zuordnung zu einer bestimmten Kategorie sprechen.

Die genannten Lernverfahren lassen sich auch für innovative Lösungen zur automatischen Indexierung verwenden. Herkömmliche automatische Indexierungsverfahren sind beschränkt in der Art, dass sie lediglich die aus einem Dokument extrahierten Terme als Deskriptoren verwenden können. Höherwertige Deskriptoren, wie sie sich bei der intelligenten manuellen Indexierung über die Verwendung kontrollierter Vokabulare und Thesauri ergeben, können von diesen Verfahren nicht generiert werden. Mit den genannten Lernverfahren ist es ohne weiteres möglich, einem Dokument beliebige Deskriptoren automatisch zuzuweisen. Analog zu den lernenden Kategorisierungsverfahren muß hierzu eine ausreichend große Anzahl bereits indexierter Dokumente vorgegeben werden. Die Lernverfahren sind dann in der Lage, genau die Dokumenteigenschaften zu isolieren, die für die Vergabe bestimmter Deskriptoren spricht.

Sowohl zur Unterstützung der automatischen Kategorisierung als auch für die automatische Indexierung stehen eine Reihe verschiedener Lernverfahren wie Neuronale Netze, Statistische Klassifikatoren und Regellernverfahren zur Verfügung. Sie unterscheiden sich bezüglich ihrer Geschwindigkeit beim Training, bezüglich ihrer Geschwindigkeit bei der Kategorisierung/Indexierung, bezüglich ihrer Genauigkeit und bezüglich der Möglichkeit, noch korrigierend in die automatische Kategorisierung und Indexierung einzugreifen.

In unserem Beitrag wollen wir jüngste Forschungsergebnisse im Bereich von Lernverfahren zur automatischen Kategorisierung vorstellen und ihre Potentiale und Grenzen in innovativen Lösungen zur automatischen Indexierung und Kategorisierung diskutieren.

Des solutions innovantes pour le classement et l'indexation automatiques de documents électroniques

Markus Junker, Andreas Dengel

Jusqu'à présent, les procédés de création d'index automatique ne pouvaient extraire de descripteurs qu'à partir des documents. De nouvelles procédures autoapprenantes, développées récemment dans la recherche documentaire, permettent désormais de fournir automatiquement des descripteurs dans n'importe quel langage de documentation. La condition préalable est d'appliquer ces procédures à des archives déjà dotées de descripteurs. Ces procédures ont été développées à l'origine pour aider au classement automatique et présentent dans ce domaine aussi des possibilités d'application intéressantes.

Les systèmes de classement permettant d'organiser des fonds de documents sont largement répandus et sont généralement considérés comme un instrument utile pour la recherche documentaire. Les systèmes automatiques de classement essaient de faciliter le classement d'un document en utilisant des règles d'affectation définies au préalable et exécutables automatiquement. La recherche documentaire et l'intelligence artificielle de ces dernières années se sont intéressées aux procédures de classement autoapprenantes, qui opèrent sans qu'il soit besoin de définir manuellement des règles d'affectation longues et fastidieuses. Ces procédures ont été rodées sur un grand nombre de documents correctement classés. À partir de documents d'entraînement, elles sont en mesure de classer automatiquement un nouveau document dans la bonne catégorie. En interne, ces procédures essaient, en calculant la

distribution statistique des mots, d'isoler celles des propriétés qui permettront d'affecter le document à une catégorie précise.

Les systèmes autoapprenants peuvent également être appliqués à des solutions innovantes dans la création d'index automatique. Les procédés automatiques traditionnels sont limités, en ce sens qu'ils ne peuvent utiliser comme descripteurs que les termes extraits d'un document. Ces systèmes ne peuvent pas générer des descripteurs de plus grande valeur, tels qu'ils résultent de l'indexation manuelle intelligente à partir de glossaires et de thésaurus vérifiés. Grâce aux systèmes autoapprenants, il est tout simplement possible d'affecter à un document n'importe quels descripteurs. À l'instar des systèmes de classement autoapprenants, il faut là aussi disposer d'un nombre suffisamment important de documents préalablement indexés. Les procédés autoapprenants sont alors en mesure d'isoler les propriétés du document qui permettront de lui attribuer tel ou tel descripteur.

Différents procédés d'apprentissage, tels que les réseaux neuronaux, les classificateurs statistiques et les procédures d'apprentissage de règles, peuvent être utilisés pour le classement automatique, mais aussi pour l'indexation automatique. Ces systèmes sont différents dans leur rapidité d'entraînement et dans leur vitesse de classement/d'indexation, mais aussi dans leur précision et dans la possibilité qu'ils offrent ou non de pouvoir corriger le classement et l'indexation automatiques.

Nous présenterons dans notre exposé les résultats les plus récents de la recherche dans le domaine des systèmes d'apprentissage permettant le classement automatique, et nous discuterons de leurs potentialités et de leurs limites dans des solutions innovantes d'indexation et de classement automatiques.

Workshop 1

Tuesday 7 May 2002

THE USE AND IMPLEMENTATION OF METADATA STANDARDS

Chairperson: Walter Koch (Austria)

Co-chair: Ian Campbell-Grant (United Kingdom)

Maria Guercio

Maria Guercio entered the National Archives in 1978 and worked for almost 10 years in the Soprintendenza archivistica of Rome, particularly in the area of business archives, municipal archives and the record-keeping system.
In 1987 she joined the Central Office of Archives and worked in the Research Department for descriptive standards. In 1992 she was the Director of the International Service for the Ministry for Cultural Heritage. In 1994 she worked in the Office for the Archival Automation and Electronic Records, then — in 1996 — as the person responsible for the international activities of the National Archives, with special reference to the European Union and the Council of Europe. Since 1996 she has been part of the Commission, working on the new Italian legislation for electronic records management.
Since 1988 she has been the Italian representative and since 1990, also the Secretary (and Chairman in 1995–1996) of the ICA Committee on Current Records. In 1996, she was a member of the ICA Commission for Programme Management. She taught courses on electronic record-keeping system at the University of Macerata (1993–99) and at the University of Rome (1995–98)
In 1998 she joined the University of Urbino as Associate Professor in archival science and electronic records and since 2001 she has been full professor at the same university. She is Director of the Urbino Institute for Archival and Librarian Science and the Chair of

Records classification and content management: old functions and new requirements in the legislations and standards for electronic record-keeping systems

Maria Guercio

Records management and content management: overlapping and convergence

A long introduction would be necessary to discuss the potential that technological innovation offers to those designing record-keeping systems, with reference to the tools improving and qualifying information (classification, metadata and content management ([1])). Because of the time available for my presentation, I will simply comment on the duplicity and ambiguity of the topic I'm going to cover, on the potential supports to the decision-making process and also the risks of improper use of concepts and tools that was defined by experts as 'another casualty of consulting faddism, much as did business process reengineering (BPR) or total quality management (TQM), which, in many cases, did not deliver sustainable value to customers' ([2]). The main purpose of this contribution is participating in the debate on a possible prospective interaction and cooperation between information-science practitioners and record-keeping system practitioners with particular attention paid to the drawing up of national and international standards and guidelines.

The thrust towards a qualified content management by the various stakeholders in the computerisation process, software houses in particular, has a twofold meaning: on the one hand it is a significant opportunity for innovation and improvement of traditional functions not supported by old technologies, in this way the record systems (as deposits rich in information content fully and immediately re-available) are placed at the centre with a strategic role in the ongoing process of change; on the other hand it implies (typically present in a transition stage where information specialists are dominant and preservation specialists are marginalised) the risk of:

— losing the notions of record archives, record structures and relations (structure, context and legal function of documentary objects and their aggregations) in favour of an indistinct and disqualified 'information' dimension, leading to numerous attempts of codifying and identifying structured content by using 'automated, algorithmic and data driven techniques' ([3]);

— renouncing the complexity of stable and significant relations which are created, maintained and communicated between the decision-making process and its evidence in a traditional record-keeping system;

— perpetuating the marginality of a specialised professional role among the useful or even necessary investments to be carried out in the current management activities: traditionally this professional role was active at the end of the record producing chain, while in the digital environment it should be present at the phase of conception and the design of the records system itself.

In this context, with these potentials and limitations, new attention has been drawn on new tools and ways to handle records and on the active roles that archivists and record managers have during the production process of documentary sources. Caution and innovative intelligence must guide the record-keeping professions over this period of change, which is both exciting and contradictory, characterised by inter-professional contaminations and continuity of traditions. Paul Browning and Mike Lowndes state 'We are all hybrid now', in an interesting essay ([4]) on the integration of library management systems (LMS) and content management systems (CMS). A similar expression can be used in the archival community to describe the state of research on record management system (RMS). In order to obtain real progress however, with the aim of integration, it is necessary that borders are clearly set, complementarities granted and strengthened and confusion avoided. It is especially important that records systems are designed, developed and subsequently acknowledged as structured systems of contents, relations and metadata, functional to specific purposes targeted towards record production and management and not for a generic need for content retrieval.

([1]) The tem used in specialised texts is knowledge management, defined as a 'system and managerial approach to collecting, processing, and organising enterprise-specific knowledge assets for business functions and decisions'. Data mining and text mining are considered two sub-sectors, strictly linked to classification tools for record information and advanced research functions.
([2]) Hsinchun Chen, *Knowledge management systems — a text mining perspective*, Tucson, The University of Arizona, 2001, p. 3.
([3]) *Ibidem*, p. 4.
([4]) 'Content management systems: Who needs them?', *in Ariadne issue, 30 (20 December 2001), http://www.ariadne.ac.uk/issue30/techwatch/intro.html*

The document and record-management facilities are included in CMS, but they are understood and described as low-level and limited-use facilities (imaging, filing, archiving). On the contrary in information science literature, the activities of record attribute definition based on specific typologies, preparation and use of classification plans, indexing and retrieval are erroneously considered only in terms of CMS facilities to be developed from scratch. Actually, even though not so frequently in the private sector, document management systems in digital environments often include — and they should do more in the future — the advanced use of classification and indexing systems, they are recently enhanced thanks to sophisticated IR tools (controlled dictionaries, thesauri, etc.).

The first aspects to clarify are the unique features of content-management systems and particularly the characteristics and specifications of their information-retrieval function which enables a real enhancement of the purposes of the — traditional or revised — classification systems envisaged in record management systems.

CMSs, more than being a technology or a specific product, identify a set of activities or, more recently a set of software products, aimed at organising the contents of an information system (not necessarily a documental system) in a digital environment to grant interoperability (in terms of retrieval and reuse of contents) between systems or heterogeneous parts of the same system. The favoured environment is the web and the specific objectives are:

— the development of conditions guaranteeing the reuse and integration of information from different origins;

— efficient information retrieval;

— maintenance of contents and structures of the information system;

— user-friendliness: content production, organisation and reuse must take place through automatic capture and preservation of all metadata indispensable for a correct management of the system over time.

CMSs were originally focused on the development of web interfaces. Today they are looked upon with growing interest as tools able to manage and qualify the whole information system of an organisation or a network of organisations. Their overlapping with electronic record management systems (ERMS) is an often unavoidable but not necessarily negative event, especially when it gives rise to (not to be taken for granted) guided and aware processes of mutual recognition, integration and convergence.

The first step in this direction implies the analysis and assessment of ERMS tools, in particular those tools constituting the functional requirements of a record system and which promoted the transition towards electronic management systems: univocal identification and systematic organisation of records according to functional classification and filing principles.

Some preliminary remarks are necessary here.

— Information systems and record systems are integrated resources. They are difficult to differentiate during the creation stage, but more and more differentiated during the following maintenance and preservation processes when they constitute the archival memory of the creator.

— The growing interaction (real coincidence from the time point of view) between record production and record communication (i.e. publication on the web, development of company's intranets and cooperative work) entails (without any further thrust within the professional community) the risk of a process resulting into undifferentiated and flattened products, loss of quality vis-à-vis the speed of communication bordering on immediateness.

— Content versioning, content integration and process workflow tools are indispensable but not sufficient requirements both for content management programmes and record management systems, especially in digital environments.

— Electronic record systems imply the adoption of all basic elements of a CMS ([5]), i.e. preparation of pre-defined approaches to record typologies ('template-based self-service authoring for

the graduate courses for cultural heritage preservation. She has been the coordinator of the Italian Group within the InterPARES project for the preservation of electronic records and Research Director of the Italian partnership within the NARA-NPACI-Supercomputer Centre project for the preservation of archival persistent objects. She is also consultant for the archives for the Italian Authority for the Information Technologies in the public sector and the author of many articles and essays. These two projects have been funded by the National Archives, the National Council on Research and the Ministry for Universities (ITL 300 000 000 and ITL 200 000 000). She is Technical Director of the research project for the definition of a standard for electronic record-keeping systems at the High School of the Public Administration. She published many articles and books, particularly the Manuale di archivistica informatica (Roma, Carocci, 2001).

([5]) *Ibidem*, p. 9.

non-technical content providers'), 'roles-based security, workflow management (submit, review, approve, archive), integration with existing data/databases and user authentication systems, metadata management, flexible output — write once, publish many times'.

— The use of XML as the communication and sharing meta-language is now generally accepted, although remarkable investments are necessary to reconvert legacy contents and because of the lack of market tools to produce XML contents.

— This impulse towards convergence of record resources leads to the need to carry out a twofold process. On the one hand the specialisation and improvement of traditional record management tools and on the other the development of software systems or programmes focused on content integration: 'e-portal, document-management, groupware, workflow, data warehousing, search engine, web-based training, and messaging e-mail' ([6]).

— In the context of record management functions and activities, archival classification (if adequately developed) is an advanced tool closer to knowledge management than information management, because of the complex syntactic, semantic and functional analysis it requires.

([6]) For quite some time now companies, which are centres of excellence in the sector of document management, have improved and developed traditional search facilities by using tools adopted by KM systems too.

The challenges of ICT and the roles of standards and guidelines: from machine-readable records to machine-understandable records

The most demanding challenges are still unanswered. This is due to the delay of archivists, the neglect of administrators and the carelessness of ICT experts for the traditions and tools of record-keeping. The development of standards for the electronic record-keeping is certainly able to support the evolution of information systems in the direction set by e-government programmes. However technical and technological solutions must be based on the identification of qualified and defined responsibilities and dedicated centres able to face the complexity of an ever more demanding and rapid communication.

The paperless office, based on the immediate application of information technologies proved to be an unattainable dream. It did not take into account the intrinsic properties — now and in the future still insurmountable — of paper production, the lack of computer skills of employees and inadequacy of technologies that only recently appeared to provide mature solutions in terms of imaging, workflow, computerised circulation, etc. Today an office-with-less-paper is possible and desirable (and perhaps indispensable). Especially an office with a higher quality and rationality in the organisation and use of record information made available by the use of state-of-the-art information technologies to supplement the work of educated and skilled human resources.

The development of the Internet and the increasing creation of web-oriented information and services have drawn attention, aroused energies, attracted investments on the need to have significant, non-redundant, easy to identify and use digital contents (documents and reference structures) in an automatic mode too. This is the case for example of the maintenance and preservation of record materials generated in newsgroups and related e-mail systems and of the advanced management of information-rich reports and memoranda that can be a valuable resource if organised by text-mining methods. Solutions are not always consistent with the archival nature of the materials, nor are the archivists or record managers adequately involved in computerisation processes, neither are they able to provide consultancy vis-à-vis the complexity of the question. Yet this is one of the most significant problems to be solved.

In many contexts and sectors the non-sufficient or absent quality of available contents is one of the most important reasons why the market for electronic document management systems (EDMS) and electronic record management systems (ERMS) is backward compared to other sectors of ICT, and this is also why record-keeping functions find it difficult to draw the attention of public and private top managers. This extant culture gap is linked on the one hand to a misunderstanding of the contribution that new technologies can offer to the full development of record information systems to the creator, on the other hand to the insufficient supply of integrated technological and organisational solutions by different technical stakeholders (archivists and information-scientists). The hope to skip some stages thanks to the use of more mature technologies able to replace the daily work of rationally and systematically ordering the information produced or received by organisations is just a delusion. This is proved by the limited results obtained by projects to develop self-organised, self-updated, self-explanatory information systems, specifically in

the web environment. They can be useful as initial filters for the maintenance of large quantities of information but are inadequate to meet the original and still essential needs of collection and typical use of record systems.

The improved performance and availability of telematic networks can be translated into a real wealth of content or rather knowledge, if new application products and their implementation consider the typical problems of recorded information (identification of documents and their structures, i.e., semantically and functionally relevant essential elements, recognition/reconstruction of stable and rational relations, administrative record flow management, primary and secondary responsibility management, indexing and functional retrieval). There is no new thing under the sun. However some technological solutions are new and promising and so is also the awareness — found in some national laws and in the guidelines of the European Union — of the need to define minimal functional requirements on the basis of acquired knowledge and experiences.

The development of XML and DTD formats for the maintenance of document typologies, maintenance of metadata and their automatic use, appears to be an innovation that is going to play a strategic role in interoperability over time and space. The obligation of exchanging record information by means of XML formats and the use of shared profiles is for example envisaged by the Italian legislation setting the minimal requirements for electronic management of documents (dpr 445/2000 and dpcm, 31 October 2000). The Australian legislators together with the professional communities have worked with determination to develop a real theory for archival metadata and to draw up record management standards. In February 2001 the European Union funded a study that is going to be very useful in guiding the computerisation projects of record systems.

Some operational questions have therefore found answers in terms of management and technological solutions, in fact it is generally recognised that:

— record systems must always be organised in compliance with minimal functional requirements;

— record-keeping systems must be consistent with the nature of the objects they keep and must enhance the potential for their integrated use, for example by organising metadata ([7]) in a system of structured information and using the potentials of the web to guarantee both the production and keeping of indispensable hierarchical administrative or documental relations and the new hypertextual links ([8]);

— when there are no controls, rules and procedures to produce record resources, and this is frequently the case in public organisations too, methods and tools must be developed for subsequent automatic retrieval of the metadata necessary for the keeping, maintenance and retrieval of the necessary information. This can be done through a theory on archival metadata to be developed on the basis of the work already carried out by some international projects ([9]). The theory should promote the identification of the nature and structure of record systems and their attributes, describe their functional maintenance, define methods and procedures for the preservation both of record contents and metadata for their critical interpretation over time and space.

These promising yet preliminary developments imply an advanced analysis of the archival functional requirements necessary to manage archival digital resources and their continual revision in the light of technological innovations, which is much needed for standardisation purposes too.

Content management in an electronic record subsystem: the innovative role of archival classification

Such a complex analysis on wide-ranging issues and objectives risks producing more or less exhaustive lists of minimal conditions and requirements, which are of no real use. At this transition stage of confusion and uncertainty, attention should be focused on the inclusion of any type of facilities and functions.

The proper creation and maintenance of record resources in digital environments minimising the risks of loss and manipulation and guaranteeing their enjoyment by the users, require a real organisational strategy based on some essential functions translated into a structured set of metadata.

([7]) The archival community is discussing with attention the issue of metadata and recognises (always) the necessary and unquestionable role of structured and semi-structured information for the existence of archival systems. However it is necessary to remember — as recently stressed by Margaret Hedstrom — that 'Metadata are expensive to create, capture and manage and organizations are unlikely to adopt metadata frameworks unless direct benefits ca be demonstrated', 'Record-keeping metadata. Presenting the results of a working meeting', *Archival science*, 2001,1, p. 247.

([8]) From this point of view, indications are provided by the MoReq study on database technologies adequate to support the management of highly structured information, which however is only a small portion of the record production of an organisation. Today organisations tend to produce mainly e-mail messages and their attachments, spreadsheets, digitised images of paper documents, electronic microfilms, sound records, etc. Sector studies have now proved that if the software is able to recognise the contents of a document, file or archives and if these contents can be identified and maintained with a consistent and systematic approach, many activities can be carried out automatically with a quality exceeding that of the manual work of skilled professional operators.

([9]) Compare the results of the project conducted by the Australian Government, Australian Government Locator Service (http://www.naa.gov.au/record-keeping/gov_online/agls/summary.html) and the indications provided by the European Moreq project, 'Model requirements for the management of electronic records' (http://www.ispo.cec.be/ida). Also remarkable is the proposal of ISO standards reference model for an open archival information system (OAIS) for the protection of airspace archives and data systems with the aim of providing methodological guidelines to the permanent or long-term preservation of digital information starting from the functional maintenance of necessary metadata.

In particular the following elements are of paramount importance and require investments during the design stage or in case of retrospective retrieval:

— univocal and certain identification of records (registry system);

— formal maintenance of the relating information content (identification and management of document typologies and profiles through the management of metadata frameworks);

— organisation/sorting of record materials by managing context and relation information (classification and file system metadata).

Under these circumstances the classification system widely, although not always properly, developed for paper systems is a very important tool. Often information scientists, and archivists too, have underestimated its role either believing that the basic search functions of electronic systems would suffice or, as many archivists did, ignoring the need for innovation of traditional tools in a stage of enormous technological progress. Only in the last few years people have become aware of the need to develop advanced classification and filing systems in an electronic environment too. These systems must organise records according to the nature and function of each document and according to the institutional mandate of the creating organisations. However this awareness is present only in some archival traditions ([10]). The solutions proposed by ICT, even though appealing, cannot deceive expert archivists ([11]) as to the risks of fragmentation and dispersion they introduce in the mass management of record products as all sector studies, case studies and surveys at international and national level have pointed out ([12]).

It is because of this substantial failure that new expressions have emerged: knowledge or content management or, more recently, data mining, with a more specific reference to the development of record structures and profiles. Classification, indexing, archival description are all integral and advanced parts according to my opinion. Through them the old and contradictory needs of the record management process can be met in a new and enhanced way. For example the need to limit the number of copies of documents and at the same time, make documents immediately available to a growing number of users able to search for them and reuse them in different contexts.

As already mentioned the advocates of information processes targeted to content management find difficulties with the present systems because of the growing need for:

— flexibility;

— easy, rapid and efficient retrieval for increasing quantities of information;

— high-level of control on access.

They believe that the development of high-quality software programmes to manage content is going to be a no longer deferable requirement for a correct storage of information and records, for search, access and publication functions, integrated use of complex and ever changing materials with reference both to contents and also context relations, new media and formats. Controlling different versions, managing modifications, creating new documents in hybrid and interactive systems are further needs requiring more organisational efforts on the part of record creators.

These are well known issues to the archivists, although up to now they have been managing records substantially stable for shape and content (the reclassification of documents and files is a limited phenomenon in the paper world also because of the organisational effort it implies) and with limited movements and reuse. After all, it is this non-easy access, use and reuse of the traditional archival system that transformed it, over the decades, in a poorly appealing and insignificant obligation for companies and public administrations; a very expensive commitment, which is no longer matched by a clear and immediate advantage for creators, who even in the public sector are characterised by short-term approaches. On the other hand, technological developments have placed information needs and, consequently, the archival function, at the centre of the information and organisational systems of organisations once again, but only if adequate tools and skills are available. To check if document recording and their classification/filing are requirements able to meet these new needs, a compare and contrast analysis should be carried out between what a renewed archival tradition can offer and what the market of technological innovation can supply.

([10]) In Canada, Australia, Italy, Germany, countries of northern Europe, for example, innovative experiences have been launched in this sector. However it is difficult or at least complex to verify and discuss the outcome of these experiences to lay the foundations for the definition of international standards. Even the most important international conferences neglect these issues as shown by the recent European Conference in Florence or the presentations at the Congress of the International Archives Council held at Seville.

([11]) Since documentary products are tools for the preservation of memory, from ancient times on the problem of rationally organising, with method, these products has been dealt with. However it is it is mainly in modern times, with the growth of complex administrative and documentary systems, that renewed attention and more sophisticated tools have been devoted to the issue. 'The arrangement of things and notions is needed ... both to find, and preserve and convey' wrote F. Bacon at a time of great development of human knowledge when logical systems to preserve memory were needed in all sectors. F. Bacon, *Scritti* (a cura di Paolo Rossi), Turin, Utet, 1975, p. 262. It is certain that memory preservation is possible only through an organising process, although only recently information scientists appear to have acquired this awareness.

([12]) Regarding Italy, see for example, the results of the census carried out by the Central Office for Archival Heritage by Anai on the production, management and preservation systems of electronic documents. The census, conducted on a sample of 16 organisations, showed the lack of control on electronic production of documents and the high level of dispersion. cf. M. Guercio, *La conservazione a lungo termine di documenti elettronici: la partecipazione italiana al progetto InterPARES*, in 'Archivi per la storia', 2001, 1–2, pp. 283–306. Similar remarks can be found — regarding the Australian community — in the essay by Adrian Cunningham, 'Six degrees of separation: Australian metadata initiatives and their relationships with international standards', *Archival science*, 2001, 1, pp. 271–283.

The facilities and services of content management systems used in record systems require an advanced processing of the attributes/metadata needed to:

— manage workflow systems targeted to content enhancement;

— post compound documents on the net;

— develop search facilities and improve outcome in terms of quality and speed;

— share activities;

— protect records from losses and manipulations.

In particular the objectives of content management programmes regard:

— univocal identification and appropriate metadata management;

— control of versions, intended as a tool to maintain and control record creation processes over time as well as the analysis, revision, approval and signing stages;

— search and retrieval, both in terms of content and identifying attributes and structure, by using sophisticated information retrieval modes (by specific terms; synonyms, contextual use of terms in order to exclude useless searches or indexes);

— control accesses by parts of documents or record typologies, if correctly identified;

— manage storing stages (online, near-line, offline).

In summary, these are the traditional functions of record management and, in particular, those functions that are carried out thanks to record profiling and record classification/filing activities appropriately redesigned for an advanced digital environment.

Classification/filing systems developed and improved in digital environments can therefore be the basic framework for converging and advanced CMS and RMS, naturally focused on record systems ([13]). 'Classification is, in fact, an instrument for organising records developed in the modern era to support the record-keeping function of administrative apparatuses that are ever more complex and articulated'. Classification means the recognition, identification and functional arrangement of records according to logical and consistent criteria within distinct functional archival units (files and series).

More specifically, a classification system strictly interrelated — this is a basic requirement — with the filing system enables:

— the determination, on the basis of predefined and qualified criteria, of those records which are part of the archival fonds and their connections to their specific administrative/record-keeping contexts;

— the identification and maintenance of the archival relations which are established among records in the exercise of his/her functions by the creator;

— the retrieval of records according to functional criteria as expressed from the archival units of arrangement;

— simple enhancement and preservation functions;

— the definition of dynamic methods of record management enabling periodic updating;

— the reconstruction of the historical evolution of the records system over time as a snapshot of the different articulations that the system underwent.

Since classification is systematic (all records are classified), logical and functional, the digital environment offers new potentials to increase the quality of both organisational and retrieving systems:

— more rapid retrieval of information from records (especially if supported by indexes or other retrieval tools);

([13]) For an analysis of archival functional requirements and, in particular, for classification and filing see Guercio, M., 'Principles, methods and instruments for the creation, preservation and use of archival records in the digital environment', *The American archivist* 64 (Autumn–Winter 2002), 2, pp. 238–269.

(14) The results of the first stage of the project (guidelines for the electronic management of records, classification models for administrative functions, draft manual of procedures for record management) are published on the site of Scuola Superiore (www.sspa.it)

— diversification and specialisation for record types, which are richer in information (for example, through the creation of rule-based records repositories of decisions, minutes, memoranda, general reports, etc.);

— pre-establishment of filing procedures, insofar as it remains possible and appropriate to provide a predefined, but still somewhat flexible, framework through which specific types of appropriate record arrangements can be organised for the structure of a highly articulated institution enjoying considerable autonomy;

— efficient integration with various administrative procedures (e.g. control of workflow, management of the assignment of tasks, statistical functions, etc.).

Consistent principles for advanced classification plans guarantee centralised system control, therefore homogeneity, quality of input information and search results, but also flexibility of solutions. A classification system based on standard principles, themselves founded on the analysis of business procedure activities and supported by advanced content management facilities, is a quality instrument. This tool can be shared both for the structure of classes (principle of matching, type of relations and number of the classes) and the concrete identification of the approach to common functions concerning administrative functions.

With the aim of developing integrated models for archival classification a research project was launched at the Scuola Superiore per la Pubblica Amministrazione (www.sspa.it). This project involved several branches of the Italian Public Administration (Ministry of Finance, Ministry of Treasure, Ministry of Foreign Affairs, Ministry for Cultural Heritage, Information Science Authority of the Public Administration, Customs Services) (14).

The novelty of the project is to be found both in the nature of the classification headings and in the structure of the approach: classes are targeted to the identification of functions and activities more than subjects; functions are in turn grouped in two large categories relating to administrative functions and institutional functions respectively. Items relating to the first category are those concerning functions common to all public administration services, and partially private sector organisations (human resource management, financial resource management, instrumental resource management, direction and coordination, etc). The study is not limited to the identification of heading denominations, but their description is included so that the classification plan is transformed into a guide for operators and internal and external users of the record system (see attached model).

Today more than in the past, the development of shared models requires an updated control system of the record system. These shared models imply flexibility and uniformity simultaneously and an ongoing updating process. Therefore in addition to advanced tools, the system must include a set of procedures and rules both internal and external to the public administration 'that regulate and describe, in a uniform and controlled manner, the modalities of external access and of internal use of the plan, assembling them in a single instrument, a manual for record-keeping procedures established within each creator'. The existence of a defined and qualified set of rules is a prerequisite that was only recently recognised as essential to govern the system since it can guarantee transparency and control. The Italian legislator, for example, made it compulsory for those public administrations introducing electronic record-keeping systems (dpcm 31 October 2000) providing a list of the essential components too.

The quality of procedures, of the classification plan, the quality of available metadata and their processing depend on the quantity and quality of investment on archives and ITC that each creator is able to make. Undoubtedly, in order to make progress, in compliance with minimal requirements established by the Moreq study, classification systems should provide more functions. They should not be limited to a simple hierarchy of functional headings but should indicate the typology and the nature of files created under each heading, describe briefly the meaning and the content of each last level class, define the principles of retention and preservation and the levels of access, etc. The more the control of the records creation is done in a flexible and controlled way, the more qualified, useful, developed are the results, both for internal and external users. Moreover, during their recording stage documents are to be managed in a univocal, systematic and consistent manner and document typologies should be identified already at the design stage of the record management system.

It is necessary to remember that there are very few classification systems which satisfy these requirements and have the appropriate characteristics, either because of the negligence of agencies in dealing with the problems of organising their archival fonds, or because of the distractions, noted several times, of archivists who are busy in other, more rewarding areas. The primary reason, however, is the difficulty of the work of data gathering, analysis and definition that a good system requires and the lack of methodology in this area. These are exactly the negligence and oversights that CMSs aim at solving. They can only be solved in continuity with the methods and tools used for centuries by archivists and recently introduced in national regulations, international standards, in guidelines approved to enhance the quality of electronic record-keeping systems.

Clasificación de documentos y gestión de contenidos: funciones antiguas y nuevos requisitos en las legislaciones y normas para el sistema electrónico de mantenimiento de archivos

Maria Guercio

La clasificación de los documentos es una actividad tradicional, comúnmente utilizada en el trabajo de archivo para garantizar la organización funcional de los archivos a efectos de su control legal, de una recuperación eficaz y de una evaluación sistemática. Las nuevas tecnologías constituyen un reto en este ámbito, porque ofrecen herramientas potentes para mejorar la calidad de la clasificación y para añadir nuevas funcionalidades, pero también porque parecen proporcionar distintas soluciones a la creación y al mantenimiento del contexto de los archivos mediante técnicas de indización y herramientas de recuperación de datos. La gestión de conocimientos o la gestión de contenidos son los nuevos términos que suelen proponerse en vez de la terminología común de archivos, y también como alternativa a las herramientas tradicionales. La funcionalidad y el significado de estas actividades (clasificación y gestión de contenidos) requieren un análisis tanto a nivel teórico como práctico y un esfuerzo de comparación para verificar cómo pueden combinarse para definir una estrategia eficaz para organizar los archivos y mantener su contenido, su estructura y su contexto.

Una cuestión básica, por ejemplo, se refiere a la validez de la clasificación de los archivos para garantizar una acumulación ordenada y coherente de documentos con fines de investigación y conservación, pero sobre todo para soportar las funciones ejercidas por los organismos de la administración pública y las empresas. A lo largo del desarrollo de los sistemas electrónicos de mantenimiento de archivos y normas multinacionales, nacionales u organizativas, había y hay una larga discusión, en especial sobre los medios de mantener el sistema de archivos continuamente actualizado y de vincularlo efectivamente al proceso administrativo del cual constituye residuo y testimonio. El presente documento se centrará en los principales aspectos de este debate e intentará definir otras estrategias mediante el análisis de:

- los principios tradicionales de la función de clasificación para crear, tener acceso y mantener los archivos;
- los recientes progresos en el enfoque teórico y la función de los elementos de clasificación como metadatos de archivo en el medio digital (para asegurar la fiabilidad y la autenticidad);
- la función de la clasificación en las normas y directrices existentes (ISO, MoReq) y en las legislaciones nacionales (específicamente en las normas italianas sobre sistemas electrónicos de mantenimiento de archivos).

Se prestará una atención específica a los proyectos y a los experimentos realizados en estos últimos años para desarrollar sistemas de clasificación comunes, basados en el análisis de las

funciones y desarrollados con procesos avanzados de digitalización: por ejemplo, el acertado proyecto (cuyos resultados se resumirán) promovido por la Scuola superiore della pubblica amministrazione y desarrollado por un grupo de ministerios italianos y administraciones centrales (Ministerio de Hacienda, Ministerio del Patrimonio Cultural, Ministerio del Tesoro, Ministerio de Justicia, Ministerio de Asuntos Exteriores y Autoridad para la Tecnología de la Información) con el objetivo de definir un sistema de clasificación común para las funciones administrativas y transformar el sistema de clasificación, al mismo tiempo, en una herramienta de interoperabilidad y una herramienta para facilitar el acceso de los ciudadanos y mejorar la transparencia.

Por último, a través de ejemplos y experimentos concretos, el documento se centrará en nuevas aplicaciones y desarrollos de las funciones de clasificación con relación a:

- tesauros y herramientas de indización electrónica;
- organización lógica y estructural de términos de clasificación relacionados con el control de los procedimientos empresariales.

Klassifikation von Archivgut und Content Management: herkömmliche Verfahren und neue Anforderungen in den Rechtsvorschriften und Normen für die Verwaltung von elektronischen Archiven

Maria Guercio

Die Klassifikation von Archivgut ist eine traditionelle Tätigkeit, die gewöhnlich in der Archivkunde ausgeübt wird, um die funktionale Organisation der Archive im Hinblick auf eine rechtliche Kontrolle der Archive, eine effiziente Auffindung und eine systematische Beurteilung zu gewährleisten. Die neuen Techniken sind eine Herausforderung in diesem Bereich, da sie wertvolle Instrumente sind, um die Klassifikationsqualität zu erhöhen und neue Funktionen bereitzustellen, jedoch auch, weil sie verschiedene Lösungsansätze für die Schaffung und Beibehaltung des Archivkontexts in Form von Indexierungstechniken und Hilfsmitteln zur Auffindung von Informationen bieten. Wissensmanagement oder Content Management sind neue Begriffe, die die geläufige Archivterminologie und auch die herkömmlichen Hilfsmittel ergänzen. Die Funktionalität und die Bedeutung dieser Tätigkeiten (Klassifikation und Content Management) erfordern neben einer theoretischen und praktischen Analyse eine vergleichende Prüfung, wie sie für die Festlegung einer wirksamen Strategie zur Organisation von Archiven und zur Beibehaltung ihrer Inhalte, Strukturen und Kontexte kombiniert werden können.

Ein zentrales Thema ist beispielsweise die Tauglichkeit der Archivklassifikation für die Gewährleistung einer geordneten und kohärenten Sammlung von Archivgut zum Zwecke der Forschung und der Erhaltung, jedoch vor allem zur Unterstützung der Arbeit staatlicher Stellen und von Unternehmen. Die Entwicklung von Systemen zur Verwaltung von elektronischen Archiven sowie von multinationalen, nationalen oder organisatorischen Vorschriften wurde und wird begleitet von einer eingehenden Diskussion, die sich vor allem darauf konzentriert, mit welchen Mitteln man das Archivsystem kontinuierlich aktualisieren kann, und wie man es wirksam in den Verwaltungsprozess einbeziehen kann, dessen Überbleibsel und Nachweis es gleichermaßen ist. Der Vortrag konzentriert sich auf die wich-

tigsten Aspekte dieser Debatte und versucht, weitere Strategien festzulegen, indem er folgende Punkte analysiert:

- die traditionellen Grundsätze der Klassifikationsfunktion im Hinblick auf den Prozess der Einrichtung, des Zugangs und der Konservierung des Archivs,
- die jüngsten Entwicklungen im theoretischen Ansatz und die Rolle von Klassifikationselementen wie Archiv-Metadaten im digitalen Umfeld (zur Gewährleistung der Verlässlichkeit und der Authentizität),
- die Funktionen der Klassifikation innerhalb der bestehenden Normen und Leitlinien (ISO, Moreq) sowie innerhalb der nationalen Gesetzgebungen (insbesondere in den italienischen Vorschriften für Systeme zur Verwaltung von elektronischen Archiven).

Besondere Aufmerksamkeit gilt Projekten und Versuchen der vergangenen Jahre zur Entwicklung gemeinsamer Aktenpläne auf der Grundlage von Funktionsanalysen und im Rahmen moderner Digitalisierungsprozesse: beispielsweise dem erfolgreichen Projekt (dessen Ergebnisse zusammengefasst werden), das von der Scuola superiore della pubblica amministrazione gefördert und von einer Gruppe italienischer Ministerien und zentraler Verwaltungen (Finanzministerium, Ministerium für das kulturelle Erbe, Schatzamt, Justizministerium, Außenministerium, Amt für Informationstechnologien) entwickelt wurde, um einen gemeinsamen Aktenplan für administrative Aufgaben festzulegen und den Aktenplan gleichzeitig zu einem Instrument der Interoperabilität sowie zur Erleichterung des Zugangs der Bürger und zur Erhöhung der Transparenz zu entwickeln.

Schließlich konzentriert sich der Vortrag – anhand von Beispielen und konkreten Versuchen – auf neue Nutzungsmöglichkeiten und Entwicklungen der Klassifikationsfunktionen im Zusammenhang mit

- elektronischen Indexierungsinstrumenten und Thesauren,
- der logischen und strukturellen Gliederung von Klassifikationsbegriffen in Bezug auf die Überwachung von Geschäftsvorgängen.

Classement des documents et gestion des contenus: l'intégration de fonctions traditionnelles et d'exigences nouvelles dans les dispositions législatives et les normes applicables aux systèmes d'archivage électronique

Maria Guercio

Le classement de documents est une activité que l'archivistique exerce de longue date pour garantir une organisation fonctionnelle à des fins de contrôle juridique, de recherche efficace et d'évaluation systématique. Les nouvelles technologies constituent un véritable défi dans ce domaine, dans la mesure où elles offrent non seulement des outils extrêmement performants pour accroître la qualité du classement et le doter de fonctionnalités supplémentaires, mais également des solutions alternatives à la création et à la maintenance du contexte archivistique par les méthodes d'indexation et les outils de recherche d'information classiques. Des concepts tels que la gestion des connaissances ou la gestion du contenu viennent désormais enrichir, outre la terminologie, l'éventail des instruments traditionnels. La fonctionnalité et la portée de

ces activités (classement et gestion du contenu) exigent une analyse à la fois théorique et pratique, ainsi qu'un travail de comparaison destiné à déterminer la manière de les conjuguer pour parvenir à une stratégie optimale en termes d'organisation des documents et de sauvegarde de leur contenu, de leur structure et de leur contexte.

Ainsi, par exemple, l'un des aspects essentiels concerne la capacité d'un classement d'assurer un rassemblement méthodique et cohérent des documents à des fins de recherche et de conservation, mais surtout d'étayer les fonctions exercées par les entreprises et les agences publiques. La mise au point de systèmes d'archivage électronique, de même que l'élaboration de règles multinationales, nationales ou organisationnelles ont suscité, et continuent de susciter, un vaste débat portant, notamment, sur les moyens permettant d'assurer une mise à jour permanente du système et son lien efficace avec le processus administratif, dont il constitue à la fois le reliquat et la preuve. L'exposé se concentrera sur les principaux aspects de ce débat, et tentera de dégager des stratégies ultérieures en analysant:

— les principes traditionnels de la fonction de classement dans le processus de création, d'accès et de maintenance des documents;

— les récentes évolutions au niveau de l'approche théorique et le rôle d'éléments de classement tels que les métadonnées pour la gestion des archives dans l'environnement numérique (garantie de fiabilité et d'authenticité);

— la fonction du classement dans les normes et les lignes directrices (ISO, MoReq) et dans les législations nationales (et plus précisément dans les règles italiennes régissant les systèmes d'archivage électroniques).

Une attention toute particulière sera accordée aux projets et aux expériences menés ces dernières années pour développer des systèmes de classement communs, basés sur une analyse des fonctions et mis au point grâce à des procédés avancés de numérisation: on peut citer, à cet égard, la réussite d'un projet (dont une synthèse des résultats sera proposée) dont la promotion revient à la Scuola superiore della pubblica amministrazione et le développement à une série de ministères et de services centraux italiens (ministère des finances, ministère du patrimoine culturel, ministère du Trésor, ministère de la justice, ministère des affaires étrangères, Agence pour les technologies de l'information). Il avait pour but de définir un système commun de classement qui, tout en assumant les fonctions administratives requises, puisse servir d'instrument d'interopérabilité et d'outil destiné à faciliter l'accès des citoyens et à améliorer la transparence.

L'exposé illustrera pour terminer, à l'aide d'exemples et d'expériences concrètes, les nouvelles applications et évolutions des fonctions de classement en liaison avec:

— les thésaurus et outils d'indexation électroniques,

— l'organisation logique et structurelle des éléments de classement liés au contrôle des procédures d'entreprise.

International standard for archives and records management ISO 15489

Nils Brübach

In the following, the progress and results to date of an international standardisation project which has been running since 1997 in the area of archives and records management will be presented. The United States, Canada, Australia, France, Great Britain, Sweden, Ireland and the Federal Republic of Germany are actively involved in ISO 15489 Archives- and Records-Management. Whilst this standardisation project is being carried out under the auspices of ISO TC (Technical Committee) 46 its results are primarily relevant to records management in both the private and public sector and to archives. ISO 15489 is the first standardisation project to attempt systematically to analyse and compare approaches, methods and workflows in the area of records management and to define common requirements, at an international level. In so doing, quite different traditions and concepts of records management have come face to face.

The standard's aim is the creation of a framework for the management and preservation of documents which are created in private or public organisations for internal or external usage, regardless of their physical composition and logical structure. The standard will not apply to the management and preservation of documents in archives, even though in some countries there is an overlap between records and archives management. The standard focuses on records management, that is on the preliminary stages of the archivist's core tasks, but this is a good reason for archivists to follow its development and organisation. The concept of 'archives' in the title of ISO 15489 follows the usual definition of the concept in English speaking countries. This must be qualified by the observation that, from a German point of view, it must be emphasised that the functional division between records and archives which has existed in Germany since the 16th century is often no longer applied to the preservation of digital records ([1]).

1. The progress of standardisation project ISO 15489

The international standardisation process in the area of archives and records management originated in an Australian initiative in 1996. From 1993 to 1996, Standards-Australia drew up a national standard 'AS 4390 Records Management', which represented the first comprehensive regulation in the area of records and archives management in Australia. ([2]) This standardisation project was also remarkable at an international level as it constituted the first successful attempt to produce and implement a unified regulation which would not be limited to any one sector. Up to then it had been usual in other countries, as in Germany, to formulate and implement regulations for individual administrative branches or internal recommendations and rules for particular organisations.

At the end of 1996 the Australian standard was presented to ISO (International Organisation for Standardisation) with the aim of having it elevated into an international standard ([3]).The Secretariat-General of the ISO submitted the Australian standard to be voted on by the relevant Technical Committee (TC) 46, 'Information and documentation'. In May 1997, the Australian draft standard was rejected by a majority of the member countries in the TC46. However, a standardisation project was approved by them in the area of archives and records management. Attention was focused particularly on internationally active organisations and globally operating commercial enterprises. In November 1997, a working committee, SC (sub-committee) 11 was set up within TC 46 for the standardisation process. Active participants of the committee are Australia, which holds the post of Secretary, Canada, France, Great Britain, Sweden, the United States and Germany as well as a number of international organisations active in the area of archives and records management, such as ICA and the International Records Management Trust. Some other countries have observer status.

In order to coordinate national work, shadow committees were set up in the countries actively involved. The German shadow committee, Working Group 15 in the NABD (Normenausschuss für Bibliotheks- und Dokumentationswesen, Working Committee on Librarianship and Documentation) was set up in August 1998 with members from public and private archives, research institutions in the area of management studies, software companies and the business sector.

Nils Brübach

Dr Nils Brübach was born in Hannover in 1962. He was educated as an archivist, economic historian and librarian at the Universities of Erlangen, Bamberg, Duke University, and Marburg Archives School. In 1993 he completed a PhD thesis on the structure, function and information processes of trade fairs.
From 1993 to 1995 he was junior archivist at the State Archives at Magdeburg,. His responsibilities included research into record systems and the appraisal and transfer of records from closed-down former GDR-Heavy-Industry-Combinates to the State Archive.
Since 1995, he has been a lecturer of archival science at Marburg Archives School (in 1998 he became senior lecturer). He has done research into the intrinsic value of archival material and developing migration strategies for it and, since 1996, he has led consultancies in archives and records management for several organisations in the public sector and private companies. Since 1998, he has been a member of the working group on ISO 15489 at German Standards. In 2000 he became a full Member of the Committee on Descriptive Standards of the ICA.

([1]) See Menne-Haritz, A., *Schlüsselbegriffe*, p.42, and Schärli, T., *Das archivische Lebensphasen-Modell im Kontext elektronischer Informationssysteme*, in *Digitale Herausforderungen für Archive. 3.Tagung des Arbeitskreises 'Archivierung von Unterlagen aus digitalen Systemen'*, ed. by Wettengel, M., Koblenz 1999, p.53., see especially Bearman, D. and Hedstrom, M., 'Reinventing archives for electronic records, *Electronic records management programme strategies*, Arbor, A. 1995. The application of the concept of archive to the area of electronic records management is clearest in: Gulbins, J., Seyfried, M. and Strack-Zimmermann, H., *Elektronische Archivierungssysteme*, Berlin *inter alia* 1993, pp. 252–259 and pp. 290–294.

([2]) *Australian Standard: Records management,*. Standards Australia, 1996, AS 4390, Homebush NSW.

([3]) The following account reflects an evaluation of the official documents of ISO/TC46/SC 11 'Archives/Records Management' and AA15 'Archives and records management' given by the

NADB of the DINe.V.(German Standardisation Institute). According to the ISO guidelines, 'in-committee confidence' must be observed regarding ongoing standardisation projects. Thus, specific references and quotations must be omitted. The progress of the standardisation project is, however, reflected in a general form in: ISO/IEC Directives Part 3: Rules for the Structure and Drafting of International Standards, third edition, New York 1997. See also Gründung eines DIN-Arbeitsgremiums zum Thema 'Archiv- und Schriftgutverwaltung' in: *Der Archivar 51/1998*, pp.761–762; Kappel, *R.*, Normierungsfragen im Archivwesen. Bericht über den Arbeitsausschuss 'Archiv- und Schriftgutverwaltung' im DIN, in*: Archiv und Wirtschaft* 32/1999 No 3, p. 141.

(4) Structure of AS 4390: Part 1: General; Part 2: Responsibilities, Part 3: Strategies, Part 4: Control; Part 5: Appraisal and disposal; Part 6: Storage.

(5) See, *inter alia*,. Sutton, M. J. D., Document management for the enterprise: principles, techniques and applications, New York 1996, pp. 75–135, and Brogan, M., Pederson, A. and Huma, A., 'Documenting society. a multimedia, self-contained training course in the development and management of record-keeping regimes in the office and archival domains', CD-ROM, Perth/Australia 1998, Module 2: Developing of record-keeping principles.

The Australian standard was accepted as a working draft for the new international standardisation project at both national and international levels.

The original Australian standard AS 4390 consisted of six modular parts (4). By way of an introduction, Part 1 described the standard's aim and area of application. It contained a terminology list and provided links to other standards. Part 2 dealt with responsibilities. Who benefits from records management? Within an organisation, who should be responsible for records management? The third part described the strategies used in structuring records management systems and defined the functional requirements. Part 4 was central. Under the title 'control', the sequence of events in business' records management systems, e.g. recording, classification functions (file plans) and the requirements of comprehensibility and of the documents' role were defined. Part 5 of the Australian standardisation project described selection and appraisal, on the basis of the analysis of business processes, and determining how long records should be stored. Part 6 dealt with preservation itself. Which measures are necessary for ensuring permanent access and secure information retrieval and what should storage conditions be like for analogue and digital records?

The Australian standard was primarily focused on record management systems in private commercial enterprise. In this it reflected Australian practice exclusively and followed the concept of 'records continuum'.

This concept is characterised by an archivist's perspective and unlike European approaches does not distinguish between records management for use in the institutions in which the documents originated and a later, permanent use and access for all. Its aim is a consistent and permanent management of records from the time of their creation or production until their permanent preservation and their function as retrievable items in archives. It thereby presents the irrefutable advantage of making it possible to intervene in the records management system at an early stage and so ensure it functions competently as well as ensuring then the possibilities of permanent archiving and use. However, this should not be allowed to lead to an archivist's approach being taken to documents which constitute records rather than archives. The assisting function of digital records management systems for internal records management in organisations must be the first focus of attention (5).

From the perspective of the European participants, there were six points of criticism of the Australian standardisation project. First the approach itself was criticised. It was a mixture of model-based idealistic construction and procedures (management). Due to this mixture, it was not fully comprehensible and from a German point of view was not suited to direct implementation. Secondly, from a German point of view the position of the archive and its special functions, in particular its co-decision rights in selection and appraisal, were not sufficiently taken into account. The approaches set out were very concrete but they did not allow for a division between records and archives. Thirdly, it failed to allow for the incorporation of national laws and administrative rules. Fourthly, there were gaps in the links to pre-existing standards and standardisation projects and some approaches directly contradicted other standards. In particular, the Australian terminology departed significantly from proposals made by TC46/SC3 (terminology) and the ICA Committee for terminology. The proposed requirements for the storage and permanent preservation of documents in the Australian standard also deviated sharply from the regulations of ISO 11799 (storage requirements for library and archive materials) in some points. Fifthly the Australian initiative concerned almost exclusively analogue paper documents. Digital records as a component of future document management systems were insufficiently considered. Finally, the concept of 'records' in the Australian standard consisted solely of the requirements for legality and accountability. This was not compatible with the functions and structures of electronic files which are typical in records management in the public sector, as described for example in the German DOMEA concept as early as 1998.

SC 11 therefore analysed the national positions from the first voting procedure and in November 1998 first totally revised the structure of the Australian standard and then defined the aim of the new international standard. The use of the concept 'archives' was now restricted to the preliminary stages of an archivist's work. An editorial committee with members from Australia, Germany, France and Sweden was set up. Its task was to incorporate national comments into the draft text on an ongoing basis and to keep these up to date.

A new draft text has existed since the beginning of 1999. It had little in common with the original Australian initiative and largely corresponded to the text of the published ISO standard. In particular the SC 11 decided, in May 1999, to divide the standardisation project. The division into a quality standard (ISO 15489) and a technical report was due to the fact that there were different traditions of records management at the international level, and a different understanding of 'records' and 'archives' and their functions had made it impossible to find a unified position. A division of the standardisation concept into a quality standard which defines the general conditions for records management and a technical report which also offers alternative solutions as 'best practice' in records management seems suitable for reconciling different traditions and approaches at an international level. The European countries were involved in elaborating both parts of the standardisation project and Germany oversaw the technical report.

The standard is, to a certain extent, a lowest common denominator defining the general conditions and functions of records management. Thus, it concentrates on the areas where international agreement is possible and meaningful. The technical report is understood as a 'code of best practice', i.e. it will offer a whole range of different solutions and procedures, according to which document management systems should be organised in order to achieve the quality requirements in the standardisation. An important point is to offer solutions which are compatible with the various national traditions and approaches in records management. The technical report is an instrument provided by ISO whose function is to set out standards briefly and in a comprehensible and useable manner and which should not consist of a pure description of processes ([6]). ISO distinguishes between three types of technical reports. Type 3 was chosen for ISO 15489 since it could set out 'data on the state of the art in relation to standards' ([7]). The division of the standardisation project was largely the result of a German initiative.

In November 1999, SC 11 in Melbourne adopted a new draft text of the quality standard, which was submitted as a Committee Draft for further examination and public, international discussion until 15 March 2000. Parallel with this a first draft text of the technical report was discussed. The international reaction to the submitted text was lively and overwhelmingly positive in the second half of 2000. After the incorporation of suggestions, especially from North America, a further and final revision of the standard and the international vote at which, interestingly, no objections were raised, ISO 15489 was published at the beginning of October 2001 at a session of the TC 46 SC 11. The Swedish and German shadow committees decided to publish the standard in national translations. The DIN-ISO 15489 is complete and currently being printed.

The technical report was completed six months later. In 2000 a comprehensive draft text was drawn up under German leadership and was voted on by all members of the SC 11. The technical report was submitted to the international voting procedure in parallel with the standard at the end of 2000 and was published together with the standard at the end of 2001.

2. Aspects of the standard's contents.

ISO 15489 contains 10 chapters ([8]). After a description of its aim and reference to other ISO standards which are relevant in the area of records management there follows a short list of definitions of terminology. Here it was possible to refer to ISO 5127 'Information and Documentation-Terminology', in which most of the definitions of terminology for archives drawn up by the KA Committee for Terminology, have been sourced. There were only a few terms used for the first time in ISO 15489 which had their own definitions. In the course of the final revision a chapter was included which set out the benefits of efficient records management. The tenor of ISO 15489 is particularly clear here. Those to whom the standard is addressed — not only records managers and archivists but also management and office personnel — are clearly described in this section as are their respective interests in efficient records management. The service function of records management within an organisation and its functions are the focus here. Chapter 5 of the standard describes the requirements of records management systems and also defines the main characteristics of records. It was stressed that these could be applied independently of the physical and logical structure of the documents and so could also be used for digital records. The requirements of records management systems are such that the conditions and processes of conventional and digital systems are considered in the standard. In the sixth chapter the legal conditions are dealt with. Only those areas of the law which are relevant to records management are paraphrased and it is emphasised that national

([6]) ISO directives, Part 3, pp. 16–20.
([7]) Quoted from ISO directives, Part 3, p. 16.
([8]) See ISO-15489 'Records management', ISO Geneva 2001.

(9) Further information at: http://www.iso.org/iso/fr/commcentre/isobulletin/articles/2002/pdf/recordsmanage02-02.pdf

legal and administrative rules must be prioritised. Sections 7 and 8 deal with responsibilities, design, strategies and the implementation of records management systems. It is in this section that archival considerations have been given special consideration. However this is done without confusing the two functionally independent areas of records and archives. Chapters 9 and 10 define the operational side, i.e. they give general directions on the workflow in a records management system. Various operation levels are also included in it. The standard is not a directive on records but indicates that an adequate structure for records management systems, the implementation of workflows and their control are tasks for management. The eleventh chapter deals with training. The structure of the technical report follows the same division.

3. Moving away from paper: ISO 15489 and future perspectives

ISO 15489 is a framework directive which provides guidelines which do not contradict existing European and national regulations and records management practices. The text is sufficiently concrete to define the meaning, function and elements of records management clearly. On the other hand it is conceived loosely enough to allow room for alternative procedures. These procedures are described in detail in the technical report. Indices and an identical division of the main chapters of both the standard and the technical report ensure that they are easy to use and of practical benefit in the implementation and organisation of records management systems.

The new ISO standard's most important aspect is that it follows a functional approach and does not attempt to replace functions by technology, where this is not possible. On the other hand one should also avoid simply 'electrifying' existing records management systems — but not overlook the potential of new functions for IT applications. In this way the standard is also open to future initiatives. As it was being drafted it was taken into account that standards become outdated. They do so even more quickly when they are tied to particular technical developments. Conversely, a rigid insistence on particular procedures for carrying out a function may conceal how a function might be better achieved using new instruments with other procedures. ISO 15489 takes these considerations on board, e.g. in the use of metadata or the application of various systems for the classification of digital documents. In the area of metadata in particular it is guided not by theoretical consideration or concepts alien to the subject but by the concrete needs of digital records management in practice.

The ISO standard and accompanying technical report were also written with an awareness of similar initiatives. The close connection with the results of other projects ensures two-way compatibility. On the one hand the results of the standardisation project should not contradict those of other national or international projects. On the other, as in every supranational initiative, the results must be in harmony with various national initiatives and traditions if they are to be capable of being implemented. The strategy should be that of a 'merger of equals' in which various concepts are related and not in competition with each other. At this point we must refer in particular to the MoReq project carried out within the EU's IDA programme. This aspect is to be considered again when the standard is revised after five years at the latest.

ISO 15489 offers coherence at four levels (9). It is a central document in a network of national and international initiatives in the area of records management and may serve as a standard to judge how far and in what manner the central functions have been realised. Secondly, it offers the possibility of internationalising national, traditional initiatives, on the basis of which existing systems function, and building bridges between different traditions. This aspect will be of particular importance to internationally operating organisations and enterprises as it will only be possible to develop, build and implement information programmes, internal strategies and records management systems which are globally unified or compatible at least and applicable in practice, by following the standard, since it offers functional reference criteria. Thirdly coherence between records management systems and archives may be established. The standard identifies functional interfaces which allow a problem-free incorporation of digital documents with compatible systems for the purpose of permanent access, while safeguarding the necessary function of digital documents. And fourthly, it allows coherence between functions and technology since it offers the possibility of designing IT solutions in the area of 'document related technologies'. These will, from the outset, meet the basic requirements for a functionally correct and high-quality implementation of records management operations applied to digital documents.

This last aspect in particular should be supported by and built on in the future by the ISO TC 46 SC 11's other work projects in the areas of 'records management and metadata' and 'records management in document-related technologies'.

La Norma Internacional de Gestión de Archivos ISO 15489

Nils Brübach

Este documento revela una historia de éxito: la creación de la primera norma internacional en el ámbito de la gestión de archivos, ISO 15489, lanzada en octubre de 2001 por el Comité técnico «Información y Documentación» (TC 46/SC 11: «Gestión de archivos»). Se tratarán tres temas. En primer lugar, se mostrará cómo una red de viejas tradiciones, tanto en gestión de archivos como en concepción en materia de archivos, se ha tornado fructífera para crear una base internacional para la cooperación y el desarrollo de dos documentos: una norma basada en la función y orientada en razón de la calidad, y un informe técnico que ofrece las mejores prácticas para la gestión de documentos electrónicos, preparado para su aplicación a escala mundial. La nueva norma ISO está relacionada con el grupo de normas ISO 9000, pero ofrece soluciones por sí misma, que se presentarán en la segunda parte del documento. Se presentarán los elementos de contenido de base, así como el enfoque de la norma en general y respecto de los metadatos, tema del taller.

Una norma tiene valor sólo en la medida en que la comunidad de usuarios designada la utilice realmente. El lugar de la norma ISO 15489 en una red de mejores prácticas, incluidas en otro proyecto, tal como los « modelos de requisitos para la gestión de archivos electrónicos» en concreto, se examinará en la tercera sección. Se incluirá una breve reseña de las perspectivas futuras de nuevos ámbitos de trabajo y posibilidades de cooperación en la normalización internacional en el ámbito de la gestión de archivos electrónicos.

Internationale Normung für die Schriftgutverwaltung. Die ISO 15489 „Archives and Records Management"

Nils Brübach

In diesem Vortrag wird eine Erfolgsgeschichte vorgestellt: die Schaffung der ersten internationalen Norm im Bereich der Archiv- und Schriftgutverwaltung: die ISO 15489, die im Oktober 2001 vom Technical Committee „Information and Documentation" (TC 46) SC 11 „Records Management" eingeführt wurde. Drei Themen werden behandelt. Zunächst wird gezeigt, wie ein herkömmliches Netz sowohl in der Archiv- und Schriftgutverwaltung als auch im Archivdenken nutzbar gemacht wurde, um eine internationale Basis für die Zusammenarbeit und die Entwicklung von zwei Dokumenten zu schaffen: eine funktionsgestützte qualitätsorientierte Norm und einen technischen Bericht mit bewährten Verfahren für die Verwaltung elektronischer Unterlagen, die demnächst weltweit umgesetzt werden können. Die neue ISO-Norm bezieht sich auf die Normungsgruppe ISO 9000, bietet jedoch eigenständige Lösungen, die im zweiten Abschnitt dieses Vortrags vorgestellt werden. Dargelegt werden ferner die inhaltlichen Kernelemente sowie das Konzept der Norm im allgemeinen und im Hinblick auf das Seminarthema „Metadaten".

Jede Norm ist nur so gut, wie die künftige Nutzergemeinschaft sie anwendet. Der Platz der ISO 15489 in einem Netz bewährter Verfahren, das auch andere Projekte wie insbesondere die „Musteranforderungen für die Verwaltung elektronischer Aufzeichnungen“ umfasst, wird im dritten Abschnitt untersucht. Hierzu gehört auch ein kurzer Abriss der zukünftigen Perspektiven im Hinblick auf neue Arbeitsfelder und Möglichkeiten der Zusammenarbeit bei der internationalen Normung im Bereich der e-Archivgutverwaltung.

La norme internationale ISO 15489 «Records Management»

Nils Brübach

Cet exposé décrit une réussite, à savoir la création de la première norme internationale dans le domaine de la gestion de documents: la norme ISO 15489, lancée en octobre 2001 par le comité technique TC 46 «Information et documentation», sous-comité 11 «Gestion des archives courantes et intermédiaires». Trois aspects seront abordés. On montrera, premièrement, comment un réseau de traditions anciennes, tant dans le domaine de la gestion de documents que dans celui de la réflexion archivistique, s'est transformé en une base internationale de coopération et de développement de deux documents: une norme basée sur les fonctions et orientée sur la qualité, et un rapport technique décrivant des bonnes pratiques de gestion de documents électroniques, prêtes à être appliquées partout dans le monde. Tout en étant rattachée au groupe de standards ISO 9000, la nouvelle norme ISO propose, à son propre compte, des solutions qui seront présentées dans la deuxième partie de l'exposé. Les éléments fondamentaux du contenu, ainsi que l'approche de la normalisation en général et du thème «métadonnées» de la session en particulier, sont également décrits.

Toute norme ne vaut que par l'utilisation concrète qu'en fait la communauté des usagers à laquelle elle s'adresse. Le troisième volet de l'exposé sera donc consacré à la place de la norme ISO 15489 dans un réseau de bonnes pratiques comprenant d'autres projets (MoReq pour la gestion de documents électroniques, en particulier). Les perspectives quant aux nouveaux sujets d'études et aux possibilités de coopération en matière de normalisation internationale de la gestion des documents électroniques seront esquissées dans ce contexte.

Guide to the conservation of data and electronic documents for teleprocedures, intranets and Internet sites — formats, media

French government activity in the conservation of data and electronic records

Pierre Fuzeau

Pierre Fuzeau

Pierre Fuzeau is executive partner in serda group, one of the first consulting and training companies dedicated to knowledge management, content management and records management in France. He is also publishing director of the French magazine Archimag, specialised in information technology.
Pierre Fuzeau is specialised in the management of projects in the field of: knowledge management, EDMS, workflow, Internet/intranet, records management, archiving/documentation since 1985.
He contributes to design specialised methodologies to conduct those of this projects like FlowGEID (designing workflow and EDMS processes) or Classtoo (designing competitive intelligence in companies). He also represent France in the Marchive consortium (designing records keeping systems in large administrations in Europe) conceived within the framework of the Esprit IV programme of the European Union.
He teaches in the DESS history and archives of the University of Angers (France) since 1994 and in that of ISD of Tunisia since 1998. The contents of this teaching relates to the control of data-processing projects, organisation analysis, anddesign of organisational and technological solutions. He assembled with the university of Angers the first graduate training on the records management in France.
He is graduate from the EHESS (Ecole des hautes études en sciences sociales).

This guide is intended chiefly for managers in charge of teleprocedures and intranet and Internet sites. In the modern environment, these managers have to take account of the technical and organisational aspects of the conservation of data and digital documents.

The guide describes the formats, metadata, media and organisation which will allow conditions to be determined for the conservation of digital documents produced and received, and offers recommendations for the practical implementation of such conservation.

The expected results after implementing these measures are as follows:

- permanent access to and use of the data contained in medium to long-term documents, independently of any technological changes (software, hardware, specifications and standards);
- lower service operating costs due to a reduction in circulation, redundancy of copies and a reduction in the storage of documents on paper media;
- where applicable, the ability to produce legal proof of a deed by means of a digital document which can be used within the scope of legal proceedings.

The guide offers the following elements:

- a model for the processing of digital documents to be conserved;
- the formats, metadata and types of media to use;
- reference action plans capable of implementation;
- recommendations for the tools to be applied;
- orders of magnitude for budgets;
- a framework for the formulation of a set of technical conditions of tender.

Matters relating to the digitisation of paper documents or microfilms are not covered in this guide, since digitisation is carried out upstream of the conservation process.

1. Conservation

1.1. Definition of the conservation of data and electronic documents

Conservation refers to an organisational, functional and material system which permits data and documents to be accessed and restored over time. The process of conservation differs from computer backups in that the transferred files are selected, and the duration of conservation of the data may range from the useful working life within a department to statutory durations such as 10 years for contractual documents, 30 or 40 years for credit dossiers, or unlimited duration for documents of historical value.

It will be noted that conservation can be handled either internally or externally, and may be contracted out.

(1) May delegate responsibility to an ad hoc project leader.
(2) May also be the person responsible for the management of archives.

Some documents to be conserved are probative in nature: they serve as evidence of a transaction. This quality demands the adoption of an appropriate system for guaranteeing data and documents vis-à-vis third parties.

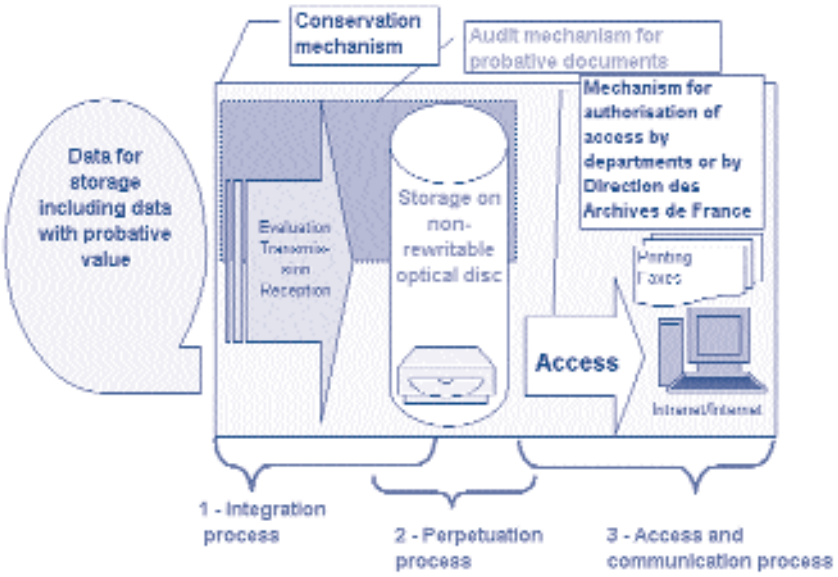
1.2. Parties involved

Within the scope of any project pertaining to information systems, the conservation process involves numerous parties, in particular those listed below:

Party	Responsibility
Teleprocedures or intranet/ Internet manager	Supervises work concerned with the definition and implementation of the conservation system and the integration of his/her recommendations into specifications for teleprocedures, intranets or Internet sites (1).
IT department	Ensures coherence and supports the system defined and adopted for conservation.
Assistant manager	Contributes expertise regarding the conduct of projects, and his/her knowledge of the domain.
Personnel of all ranks	Support implementation of the conservation system.
Training department	Organises training activities for agents directly involved in the conservation process and for users within the departments.
Users within departments	Affected by questions of access to data during or after processing.
Agent responsible for conservation (2)	Implements conservation activities. Applies the process to the various document types: teleprocedures, web, etc.
Auditor (where applicable)	May be internal or external. Produces audit reports which strengthen confidence in the authenticity of the original digital documents.
Direction des Archives de France	Prepares regulations regarding conservation and processing. Oversees compliance with regulations. Validates management charts. Issues communication authorisations in derogation of archive law.
Mission des Archives Nationales or Direction des Archives Départementales	Represent Direction des Archives de France in its supervisory capacity. Contribute their archiving skills for the evaluation of documents and preparation of conservation charts, as well as for the preparation and gathering of metadata. Issue authorisations for the removal of documents. Manage the conservation of definitive archives.

1.3. The three processes: integration, perpetuation and access

In order to satisfy requirements for the conservation of data and digital documents, information systems must incorporate the following processes and elements:



Data conservation comprises three distinct processes: the integration process, the perpetuation process and the access and communication process.

— The integration process chiefly concerns identification of the digital documents to be conserved, the recommended formats and the metadata to be associated with the documents.

— The perpetuation process concerns storage and the recommended media.

— The access and communication process concerns the management of access rights.

2. The integration process

2.1. Identification and evaluation of digital documents to be conserved

The identification and evaluation of documents to be conserved are two indispensable stages whose proper implementation undoubtedly has a significant influence on the success of the project in terms of the conservation component.

2.2. Recommended formats

Certain formats are to be preferred according to the type of document. The table below gives a summary of the various formats and will help you to make a choice on the basis of four criteria concerning the recommendability, durability, openness and frequency of use of a given format. These criteria will be represented by the following characters or symbols:

1. Recommendation

   The format is:

   a. recommended

   b. possible

All other unmentioned formats should be avoided except in special circumstances.

2. Durability

The format is:

a.  durable

b.  not very durable

This criterion takes account of the technical and usage context.

3. Open format

a.  open

b.  not open

This criterion takes account of the openness or the 'proprietary' nature of the format. If the rights to the format are owned by a particular organisation or individual, the format is considered to be 'not open'.

4. Frequency of use

a.  widely used

b.  little used

Depending on the extent of their use in the world at large, formats are classified as being either widely used or little used.

Document	Format and comments
Text	
TXT 	Recommended: text files are the most secure over time because they are the simplest. When a text is capable of being conserved in text format without any adverse effect on its comprehensibility (content and context), files in this format are to be preferred.
RTF 	Possible: RTF (rich-text format) is a commonly used format in office computing (Microsoft© and Apple©). It is used chiefly on the Microsoft Office platform. However, it offers no guarantee of durability or stability. Corruption can occur in complex documents (links, images, etc.).
Published	
PDF 	Possible: PDF (portable document format) is a proprietary standard disseminated and promoted by the company Adobe, thanks to the free distribution of its Acrobat Reader software. It is derived from the postscript format. The advantage of this format is its portability in all environments and its stability for most documents produced by specific software packages. It can be used as the central format in electronic information systems.
Structured text	
XML 	Recommended: XML (extensible mark-up language) is a derivative of SGML. It was designed and promoted by the group W3C (World Wide Web Consortium), which is a reference body in the Internet domain. Although not an ISO standard, it today constitutes the recommended standard for French Governmental departments. Documents are structured for greater re-usability thanks to the separate management of content (XML sheet), structures (DTD and schemata), metadata (RDF) and styles (CSS and XSL). Durability will be ensured by monitoring developments in this recommendation.

SGML 	Possible: SGML (structured general mark-up language) is not, strictly speaking, a format but rather a language. It consists of an international standard supplemented by additional specifications such as style sheets, printing pages, exchange formats, document structures, etc. Its disadvantage is that few software packages use this standard, which is complicated to implement.
HTML 	Possible: HTML (hyper-text mark-up language) is a simplified implementation of SGML, and is widely used on Internet sites. It is not very stable, and few improvements can be expected in the future. It is not suitable for complex or long texts.

Fixed image (dominant) — monochrome

UIT T4 ([3]) **recommended** 	Recommended: this format is a derivative of TIFF (see below). It is today the most widely used format in the world, thanks to its use by fax machines. In technical terms, the compression algorithm avoids losses but does not give the benefit of a higher rate of compression.
PNG 	Recommended: PNG (portable network graphics). This is intended as a replacement for the GIF format (see below), and incorporates improvements with regard to error detection and interlacing speed, as well as a higher rate of compression. This format is promoted by the group W3C. It is an emerging format, but as yet little used and incorporated into few tools.
TIFF 	Possible: TIFF (tagged-image file format). This format uses types of compression which make it possible to guarantee quality and losslessness. It is very widely used. TIFF exists in various versions, the most recent of which incorporates a function for the management of multi-page documents.
GIF	Possible: GIF (graphic interchange format) is a very widely used format thanks to Compuserve, which adopted and disseminated it. There are two sub-formats, 87a and 89a, corresponding to fixed and animated formats respectively. This is a proprietary format (owned by Unisys). GIF files should be migrated as soon as possible to PNG files.

Fixed image (dominant) — colour

PNG 	Recommended: (portable network graphics) — see above.
JPEG 	Possible: (joint photographic experts group). This format is widely used in the Internet world. It permits a high rate of compression (in the order of 1/40). The disadvantage is a loss of information. TIFF See above. GIF See above.

Animated image and sound

MPEG 2 	Recommended: MPEG 2 (Moving Picture Expert Group). This compression standard for video images and sound works on all platforms. Durability will be ensured by monitoring developments in this standard.

Sound

MPEG 3 ([4]) 	Recommended: MPEG 3. This standard was developed from MPEG 2 and uses only the portion relating to sound. The sound is stereophonic. Durability will be ensured by monitoring developments in this standard.

([3]) Replaces the earlier name 'CCITT Group 3'.
([4]) Or MPEG 2 layer 3.

WAV	Possible: WAV (RIFF WAVE). Originally a Microsoft Windows sound file format, this has since been extended to cover other platforms. Stereophonic sound. A proprietary format.
Databases	
Plain text	Recommended: The content of databases should be conserved in the form of plain text provided with separators (delimiting commas). Information about the structure of the data (fields) should also be conserved. The character sets to use are ISO 646 or ISO 8859-1
Plans	
CGM	Recommended: CGM (computer graphics metafile). Standard for the storage and exchange of two-dimensional graphical data. Originally a vector format, but recently widened to include raster storage abilities. Four international standards have been developed, specifying how CGM should be used in e-mail messages with MIME extensions and on the Internet. Little used in software packages.
STEP	STEP (standard for exchange of product data). Standards for the exchange of data on industrial products. Widely used.
DXF	Possible: DXF, a proprietary format used by the software package Autocad, which is very widespread.
Encryption	
RSA	Recommended: RSA (Rivest, Shamir et Adleman) dual-key encryption system (asymmetric, i.e. one public key and one private). RSA is in the process of standardisation.

2.3. Methods and preparation for the transmission of documents to be integrated

The transmission of documents to be conserved may be effected by means of teleprocedure or office automation applications. Such applications may allow users to verify the transmission (review and approval function). In any event, it is for the manager to make a choice regarding the incorporation of this function.

In addition to gathering the documents to be transferred, there is also the task of compiling the associated metadata file in one of the formats recommended above, containing at least the metadata listed below:

(a) Document date and identifier

(b) Generating organisation

(c) Creator

(d) Signatory

(e) Recipient organisation

(f) Addressee

(g) Document language

(h) Access and communicability control

(i) Conservation rule

(j) Format and original software

(k) Keyword

The following metadata currently constitute the basic elements for utilisation:

Type of metadata	Content
(a) Document date and identifier (unique)	To be given in the form yyyy/mm/dd, followed by a unique identifier consisting of a batch number followed by an accession number.
(b) Generating organisation	Use the AdER ([5]) meta-directory format (country name/organisation name/organisational unit name)
(c) Creator	Use the AdER9 meta-directory format (uniqueIdentifier)
(d) Signatory	Use the meta-directory format and specify the signatory's responsibility (uniqueIdentifier).
(e) Recipient organisation	Use the AdER meta-directory format (see metadata 'B')
(f) Addressee	Use the format of the AdER meta-directory (see metadata 'C')
(g) Document language	Use the ISO 639 standard for classification
(h) Access	In conformity with regulations and communicability control
(i) Conservation rule	Determines the duration of administrative conservation and the final destiny: e.g. 10/C for administrative conservation of 10 years from the document date, where C means 'to be conserved as definitive archive'. Note that if the dossier of which the document forms a part has a later closure date, the duration is extended.
(j) Format and original software	TXT, HTML, TIFFG3, PDF, etc. and designer/software name/software version e.g. PDF-ADOBE/Acrobat/version 4.05a
(k) Keyword	Terms which will make it possible to relate the document to a dossier or/and submit a request. For example, one might insert a company name, dossier number, title, insurance policyholder number, etc. It is possible to provide several keywords, which must be separated by commas.

([5]) For a description of the structure of the inter-governmental meta-directory service under the LDAP standard, see the following web site: http://www.mtic.pm.gouv.fr/standards/referentiels/meta_annuaire_v1.shtml)

3. The perpetuation process and the storage of digital documents

3.1. Type of medium to use

As mentioned in paragraph 3.7 of the electronic archiving standard NF Z 42-013, the storage medium must be an 'optical medium in which the writing of the bits encoding the data is effected by irreversible transformation of one or more constituents of that medium'.

It is obligatory for the medium to be of the optical type, i.e. it must use laser technology (not magnetic or magneto-optical) and must be non-rewritable (or of the WORM type — write once read many). The deformation of the active surface of the medium must be definitive and irreversible.

The medium to use at present is the recordable compact disc or CD-Rom, which offers the following advantages:

— compatibility of the medium with the vast majority of CD-ROM readers;

— reliability of the medium (a conservation life of approximately 10 years is obtained under normal conditions of relative humidity and temperature);

— low cost: a simple CD-writer currently costs less than EUR 300;

— standardisation;

— a capacity of between 550 and 600 MB of useful data.

In the medium term, the recordable digital versatile disc or DVD-R will make it possible to achieve capacities of 5–20 GB. Indeed, it is possible to use this technology today if a sufficient budget is provided.

Note that for the purpose of accessibility by internal users in their routine activities, document bases may be replicated on departmental machines.

3.2. Operations to be carried out for copying on to the storage medium

The operations to be carried out are as follows:

Operation	In the short term, i.e. 2001–02
Initialisation of medium	An initialisation file is compiled in the form of a TXT file.
Recording/writing	Each XML envelope (document and metadata files) is written in sequential order.
Closure of medium	A closure file is compiled in the form of a TXT file.
Checking of 'Master' medium	Verification of the medium using tools such as ScanDisk. Re-reading by means of sampling, with a periodicity of five years (complete verification over 5 years).
Copying of medium	No comments
Control of copy (copies)	Verification of the medium using tools such as ScanDisk.

Periodically, indexes and databases will be recorded on to CD (or DVD). The database structure will be described in an initial file and the data recorded in 'plain' format using the ISO 8859-1 character set.

3.3. Physical conservation of media

Conservation relies on the following considerations:

— the physical device: the conservation life of these media is currently quoted by manufacturers at approximately 10 years. This forecast assumes conditions of stable humidity (40–60 %) and constant temperature (15–20°C) in a non-acidic and non-aggressive environment (air filtration). These are standard provisions for the conservation of magnetic media;

— one of the sets of recorded media will be regarded as the 'master'. This is set aside for the exclusive purposes of conservation and will not be subjected to any use. It is conserved in a secure location;

— the other sets will be subjected to use (see the 'Access process' section);

— a media management database will be set up.

The operations are as follows:

Operation	Description
Secure storage of 'Master'	Use of existing system for computer backups
Operational storage of copies	Use in jukebox systems, CD towers or simple readers (see 'Access and communication process' section)

Destruction procedures will need to be defined, but these will not be implemented in the short term. However, incomplete or defective media must be physically destroyed, taking account of the security level applicable to the content of the documents contained on the media (crushing of media).

3.4. Safeguarding the conservation of data on media

This involves regular checks on the condition of the recorded data.

Verification should be carried out annually by applying test software to 20 % of the media.

Every verification operation is systematically tracked in the form of structured data comprising the following information:

— site of the verification, with the computer IP number where applicable;

— time stamp of the verification;

— name of the organisation;

— last name and first name /login of the person responsible for the operation;

— test method;

— description of the sample;

— record of results;

— summary and decision.

3.5. Changes of media and their management

Technological changes in terms of document formats, media and user access interfaces occur at increasingly short intervals. This being the case, the best course is to migrate the XML envelopes from one technology to the other and keep track of these migrations.

Consequently, old reader units will not be conserved and the data will be transferred on to the new media accepted by the new units.

If the new readers accept the existing media, checks must be carried out to ensure that reading is exhaustive.

The same problem applies to the formats used and the document management and access tools.

The principles of migration are as follows:

— migration tools are applied only after the tests yield positive results; in other words, a preparation and testing phase is observed before carrying out the migration operation;

— the full description of the migration process must be conserved, i.e. information concerning the original format, target format, methodology, tools used, list of migrated documents and test results;

— the reference for each migration must be recorded in the metadata of the XML envelopes;

— each migration must unconditionally entail a transfer on to a new medium. The reference to the previous medium must be preserved.

The operations are described below:

Operation	In the short term	In the long term
Planning of migration	No migration to be envisaged in the short term	Specifications to be defined
Application of migration procedure/preparation phase	No migration to be envisaged in the short term	The procedure will differ according to whether it concerns a migration of format, medium or operational software. Specifications to be defined, including the series of tests to be carried out.
Application of migration procedure/application phase	No migration to be envisaged in the short term	The procedure will differ according to whether it concerns a migration of format, medium or operational software. Specifications to be defined.
Recording and conservation of each migration dossier	No migration to be envisaged in the short term	An information database is compiled on the migrations carried out.

3.6. Management of perpetuation

Storage and maintenance operations on the conservation system must be recorded. For each medium, the operations carried out are conserved, particularly the histories of operations such as migrations, physical localisations of media at the time of their storage, etc.

Every operation is systematically recorded in the form of structured information comprising:

— identification of the medium (master or copy, unique numbering, etc.);

— characterisation of the medium (type, content, etc.);

— tracking of operations carried out on the medium (verification, test, migration, etc.);

— planning of operations on the medium (verification, test, migration, etc.);

— localisation of media. Site of verification, with the computer IP number where applicable.

In the short term, the parameterisation of a database may involve the recording of the following information:

— medium identifier;

— type of medium;

— creation date;

— person responsible for the recording;

— description of the content and dating of the content;

— recorder used;

— recording software used;

— physical location of the medium;

— movements made, e.g. tests carried out, withdrawals for document consultation, etc.

For the backing up of system files and index bases, standard backup methods will be applied.

3.7. Useful information for storage

Certain elements of information regarding the environment are connected with the documents to be conserved, and should be incorporated into initialisation or closure files.

Initialisation file:

— site of initialisation, with the computer IP number where applicable;

— time stamp of the initialisation;

— name of the organisation;

— last name and first name/login of the person responsible for the operation;

— name and version of the software used to write the disc;

— serial number of the reader/writer used;

— accession number of the medium (be sure to use a numbering system based on the uniqueness of the identification of each medium).

Metadata to be added to the digital document at the time of recording:

— accession number of the medium + sequential accession number in the medium;

— time stamp of the recording.

Closure file:

— site of closure, with the computer IP number where applicable;

— time stamp of the closure;

— name of the organisation;

— last name and first name /login of the person responsible for the operation;

— name and version of the software used to write the disc;

— serial number of the reader/writer used;

— accession number of the medium;

— total number of XML envelopes;

— total number of bytes representing XML envelopes.

If the files amount to more than 600 MB, it will be necessary to activate the multi-volume function of the transfer software.

Time stamp: if the capturing of information regarding the date/hour/second/millisecond is important, the accuracy and precision of such information will lend increased credibility to the content of the document concerned. In order to be capable of guaranteeing and proving the accuracy of this information, radio card systems will be incorporated into the servers to which the production computers are connected.

3.8. Storage costs (January 2001)

CD-R writers have a unit cost in the order of EUR 300–800 including taxes. For writers installed in series (several CD-R discs written in parallel, acceleration of writing processes, etc.), the cost may be as high as EUR 4 500–10 000 including taxes.

DVD-R writers have a unit cost in the order of EUR 6 000–8 000 including taxes.

4. The access and communication process

4.1. Management of access rights and access to the requested document

Two types of access must be provided for:

— access to the digital document by the generating departments, which is effected in the form of a 'replay'.

This access limit for government departments means that the documentary database must incorporate access rules. These rules must be adapted to any organisational changes:

— access to the digital document by departments other than the generating department, and by users and researchers.

Documents deriving from (issued or received by) centralised or decentralised government departments constitute public archives and are covered by the provisions of Article 3 of the Law of 3 January 1979. Consequently, Direction des Archives de France has the authority to oversee application of the law, and in particular to authorise access to public documents by private individuals and researchers. Such authorisation may be granted by means of derogation where the documents concerned are not immediately communicable. Access to documents is becoming an increasingly sensitive issue as 'search and display' operations become technically easier thanks to Internet standards. As an example, data concerning the salaries or performance evaluations of personnel is naturally confidential, and access is restricted to authorised parties.

The responsible departments may intervene to authorise a user to access a document. At this stage, the authorisation will be in the form of a message or fax, or the user may receive the document itself via a messaging service or fax machine. If directly concerned, he or she may receive a paper copy.

In order to satisfy document access needs, use may be made of the functions offered by an electronic document management (EDM) application. This type of application offers the following functionalities:

— searching: multi-criteria searches, cross-referencing, Boolean operators, fuzzy logic searching;

— security: access control;

— display: use of viewer libraries permitting on-screen display, manipulation of the obtained image, printing abilities (which may be deactivated depending on the status of the user);

— publication: ability to publish via Acrobat for forwarding;

— recording and monitoring of access requests and document transmissions;

— keeping of access history, with statistics;

— integration of these applications on an Intranet or the Internet.

The budget to be projected for this type of application is as follows:

Component of the EDM system	Budget estimate
EDM software package (cost of licences)	Based on several stations (5–10): less than EUR 50 000 including taxes (includes commissioning). For an EDM application incorporating other functionalities: EUR 800–1 500 including taxes per agent workstation.
Document server	Based on a Pentium III-1000 server with 512 MB RAM, 30 GB storage, Raid 5, backup system, NT: 10,000 Euros inc. taxes

Tower/jukebox = dedicated server	From EUR 15 000 including taxes for a CD tower, and EUR 45 000–150 000 including taxes for a jukebox + server
Commissioning and configuration	EUR 7 500–45 000 including taxes

4.2. Conservation management

The management of conservation activities demands a high degree of technicality and a global approach on the part of the organisation concerned. It must therefore be handled very rigorously and as an ongoing concern.

The management operations are as follows:

Operation	Description
Supervision of activity	— compilation of a manual of standard procedures for third parties engaged in conservation, and application of this manual — definition of the documents to be generated by this activity in order to prove the reliability of the process, and application of these documents — definition of management charts for the activity, and utilisation of these charts — keeping of statistics — keeping of a history of operations
Ongoing training programme	Training in new documentary technologies Training in the activity of a conservation third party (management level and agent level) Training in project management and project supervision
Definition and implementation of organisational and technological projects	Creation of columns relating to digital documents in management charts Compilation/adaptation of specifications for invitations to tender Supervision of implementation of the new mechanism
Application of security mechanisms	Creation of system backups Installation of anti-intrusion system and filtering of access to equipment

4.3. Auditability of the conservation system

Conservation activities must be auditable and periodically audited so that, where applicable, it can be demonstrated to any requesting authority that documents are being conserved in conformity with the proper conditions.

In order to prove the reliability of operation of the conservation system, audit and self-audit measures must be applied regularly and systematically.

This activity may be effected by means of expert intervention, in which case the estimated budget would be EUR 3 500 including taxes for two or three auditor days.

Self-auditing can also be practised, in which case the checklist below should be used to identify/analyse those points which are satisfactory and those which require improvement.

Is there a procedures manual?

— Are there specialised personnel dedicated to this activity?

— Have the personnel been trained?

— Have the management charts been completed?

— Is the EDM software in place?

— Is the EDM software operational?

— Is the hardware (writer, server, tower/jukebox) in place?

— Are the security measures for the 'master' medium and copy (or copies) operational?

— Is the hardware (writer, server, tower/jukebox) operational?

— Are the physical measures for controlling equipment access in place?

— Are the index backup measures operational?

— Are the measures for tracking operations carried out on digital documents operational? For integration, for storage, for accesses and for publications?

The periodicity of self-audit may, for example, be every six months, and should also include the compilation of documentation on the self-audit histories.

Conservación de datos y documentos electrónicos para teleprocedimientos, intranets y sitios de Internet. Formatos y medios

Actividad del Gobierno francés en materia de conservación de datos y documentos electrónicos

Pierre Fuzeau

En enero de 2001, el Gobierno francés publicó la *Guía para la conservación a medio y largo plazo de datos y documentos electrónicos para teleprocedimientos, intranets y sitios de Internet. Formatos y medios.* La circular sobre la elección del formato para los documentos electrónicos se publicó el 23 mayo de 2001. La guía y la circular son aplicables a todos los documentos electrónicos que deban conservarse. Ambos textos fueron encargados por ATICA/Primer Ministro (Agence pour les Technologies de l'Information et de la Communication dans l'Administration).

Ésta fue una de las primeras contribuciones destinadas a conservar documentos en la administración pública francesa. Este trabajo tenía por objeto introducir una reflexión sobre este tema, definir directrices, aumentar el acceso a la información y la transparencia y reducir costes. Los objetos de protección eran el formato, los metadatos, el soporte y la organización.

En un principio el proceso de diseño consistió en una petición de comentarios, un grupo de trabajo y ayuda exterior. Los socios de este proyecto eran ATICA/Consejo Científico, ministerios, administraciones locales, universidades e industrias. El trabajo se basó en la norma relativa a la guía de información digital (Foro DLM), en la norma francesa NFZ 42 013 sobre almacenamiento de archivos electrónicos, en la norma ISO sobre gestión de documentos (15489) y en la norma XML.

Las recomendaciones se basan en el modelo que cubre los siguientes procesos de integración (evaluación/transmisión/recepción), conservación (almacenamiento), y acceso y comunicación (dentro del mecanismo de autorización de acceso por departamentos o por la Direction des Archives de France).

Aufbewahrung von Daten und elektronischen Unterlagen für Televerfahren, Intranets und Internetseiten. Formate, Medien

Maßnahmen der französischen Regierung im Hinblick auf die Aufbewahrung von Daten und elektronischen Unterlagen

Pierrre Fuzeau

Im Januar 2001 veröffentlichte die französische Regierung den „Leitfaden für die mittel- und langfristige Aufbewahrung von Daten und elektronischen Unterlagen für Televerfahren, Intranets und Internetseiten. Formate, Medien". Das Rundschreiben bezüglich der Formatwahl für EDV-Archive wurde am 23. Mai 2001 veröffentlicht. Der Leitfaden und das Rundschreiben gelten für alle EDV-Archive, die aufbewahrt werden sollen. Beide Texte wurden von der dem Premierminister unterstehenden Atica (Agentur für Informations- und Kommunikationstechnologien in der Verwaltung) in Auftrag gegeben.

Es handelte sich hierbei um einen der ersten Beiträge, um ein Gedächtnisarchiv in der öffentlichen französischen Verwaltung anzulegen. Diese Arbeit sollte dazu dienen, über dieses Thema nachzudenken, Leitlinien festzulegen, den Zugang zu Informationen zu verbessern, die Transparenz zu erhöhen und die Kosten zu senken. Schwerpunkte bei der Aufbewahrung waren Formate, Metadaten, Datenträger sowie Organisation.

Das Entwicklungsverfahren wurde ursprünglich durch eine Aufforderung zur Abgabe von Stellungnahmen, eine Arbeitsgruppe und eine externe Beratung eingeleitet. Beteiligt an diesem Projekt waren der wissenschaftliche Beirat von Atica, Ministerien, Kommunalverwaltungen, Hochschulen und Privatunternehmen. Die Arbeiten stützten sich auf Referenzwerke wie den Leitfaden für digitale Informationen (DLM-Forum), die französische Norm NFZ 42 013 für die Archivierung elektronischer Unterlagen sowie die ISO-Norm für die Archivgutverwaltung (15489) und die XML-Standards.

Die Empfehlungen beziehen sich auf das Muster, das die folgenden Prozesse der Integration (Evaluierung/Übermittlung/Empfang), der Aufbewahrung (Lagerung), des Zugangs und der Kommunikation (im Rahmen der Zugangsregelung der Dienststellen oder der „Direction des Archives de France") umfasst.

Conservation des informations et des documents numériques pour les téléprocédures, les intranets et les sites internet: formats, supports

Action du gouvernement français en matière de conservation de données et de documents électroniques

Pierre Fuzeau

Le gouvernement français a publié en janvier 2001 le *Guide pour la conservation des informations et des documents numériques à moyen et long terme pour les téléprocédures, les intranets et les sites internet: formats, supports*. La circulaire relative au choix des formats pour la réalisation de documents constitués de données numériques a été publiée le 23 mai 2001.

Le guide et la circulaire s'appliquent à tous les documents électroniques à conserver. Le commanditaire de ces deux textes a été l'ATICA (Premier ministre).

Il s'agit de l'une des premières contributions au maintien de la mémoire au sein de l'administration publique française qui vise à susciter la réflexion sur ce thème, à définir des lignes directrices, à améliorer l'accès et la transparence et à réduire les coûts. Les principaux éléments de la conservation abordés dans le guide sont les formats, les métadonnées, les supports et l'organisation.

Le processus de conception a été initialisé par un appel à commentaire, un groupe de travail et une consultation extérieure. Les partenaires du projet étaient le comité scientifique de l'ATICA, les ministères, les collectivités territoriales, les universités et les entreprises privées. Les travaux se sont appuyés sur des référentiels tels que le Guide de l'information numérique (Forum DLM), la norme française NFZ 42 013 sur l'archivage électronique, la norme ISO 15489 (Records Management) et les standards XML.

Les recommandations se basent sur un modèle couvrant le processus d'intégration (évaluation/transmission/réception), le processus de pérennisation (stockage) et le processus d'accès et de communication (dispositif d'autorisation d'accès par les services ou par la direction des Archives de France).

MoReq metadata — beyond Europe?

Marc Fresko

Background

MoReq stands for Model Requirements for the Management of Electronic Records. It is a formal specification of requirements for electronic records management systems (ERMS).

The need for a comprehensive specification of requirements for electronic records management was first articulated by the DLM-Forum in 1996. Subsequently, the European Commission Enterprise DG's Interchange of Data between Administrations (IDA) programme ([1]) commissioned the development of this model specification.

IDA ran an open competition in 1999, following which a development contract was awarded to Cornwell Affiliates plc ([2]). Work on the specification started in 2000, and the completed specification was published in early 2001. It is now available for download in various formats from several websites, including www.cornwell.co.uk/MoReq.html.

The MoReq specification

The MoReq specification sets out functional and non-functional requirements. It is written to be equally applicable to public and private sector organisations which wish to introduce ERMS, or which wish to assess the ERMS capability they currently have in place, in any European country. The target audience is:

— potential ERMS users: as a basis for preparing an invitation to tender;

— ERMS users: as a basis for auditing or checking an existing ERMS;

— training organisations: as a reference document for preparing records management training, and as course material;

— academic institutions: as a teaching resource;

— ERMS suppliers and developers: to guide product development by highlighting functionality required;

— record management service providers: to guide the nature of the services to be provided;

— potential users of outsourced record management services: as an aid in specifying the services to be procured.

The specification is written with an emphasis on usability. Throughout, the intention has been to develop a specification which is useful in practice. Feedback from users suggests that this intention has been met successfully.

The project team

Development was carried out by a small team of specialist consultants from Cornwell Affiliates plc, supported by a guiding team of experts drawn from several countries, and validation organisations from both the private and public sectors. The EC Project Manager was Mr Paul Murphy; the participants were:

- Project Director: Keith Cornwell, Managing Director, Cornwell Affiliates plc.
- Project Manager: Marc Fresko, Director EDMS and ERMS Consulting, Cornwell Affiliates plc.

Marc Fresko

Marc Fresko is the EDMS and ERMS Consulting Services Director for Cornwell Affiliates plc, an independent consulting firm based in the United Kingdom. He has provided consultancy advice on the use of office systems since 1979. The majority of his practice is in the public sector, where he has advised several government departments and agencies in the United Kingdom and other countries, the English and Scottish Parliaments, the Public Record Office, the British Library and the European Commission. Private-sector clients include financial institutions and professional practices.
Having worked with electronic document management systems since 1985, he has more recently been heavily involved with electronic records management (ERM). He has advised numerous clients on implementation approaches for ERM, and has assisted them in specifying and procuring systems. He was the project leader on the European Commission 'MoReq' project — originally conceived by the DLM Forum — which developed a unique generic specification for ERM systems. He was also a member of the team which produced the UK Public Record Office ERMS Specification, which is a de facto standard in the United Kingdom.
He has particular interests in 'legal admissibility' issues, document management terminology, and digital preservation. In this connection, he is an active member of the UKAIIM Standards Committee, and related BSI and ISO committees which work in these areas. He is widely published in a variety of publications, and is frequently invited to speak at conferences and similar events in the United Kingdom and elsewhere.

([1]) http://www.ISPO.cec.be/ida
([2]) http://cornwell.co.uk. The company has since then changed its name to Cornwell Management Consultants plc.

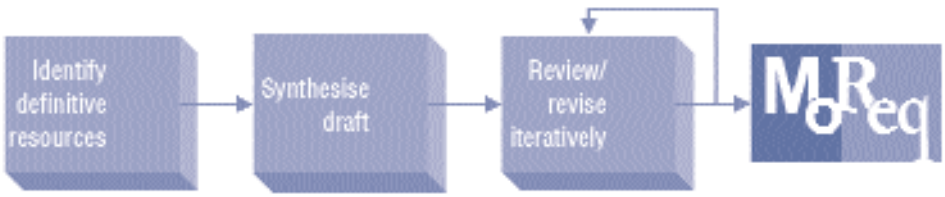
The MoReq team at its kick-off meeting, Heathrow, February 2000

- Consultant: Martin Waldron, Managing Director, In-Form Systems Ltd.
- Semi-independent reviewer: Keith Batchelor, Records Management Consultant.
- External reviewer: Francisco Barbedo, Archivist Expert, Porto State Archive (Portugal)
- External reviewer: Dr Nils Brübach, Marburg Archives School (Germany).
- External reviewer: Miguel Camacho, European Projects Manager, Sadiel S.A. (Spain)
- External reviewer: Prof. Luciana Duranti, School of Library, Archival and Information Studies, University of British Columbia (Canada).
- External reviewer: Prof. Mariella Guercio, Institute for Archival and Librarian Studies, University of Urbino (Italy).
- External reviewer: Peter Horsman, Senior Consultant, Netherlands Institute for Archival Education and Research.
- External reviewer: Jean-Pierre Teil, Project leader, National Archives (France).
- External reviewer: Hans Hofman, International Council for Archives.

The author considers himself fortunate to have had the privilege of working which this distinguished and able team, and is very grateful to all of them for their contributions to MoReq.

Development: methodology

The methodology is reasonably traditional. In summary, definitive records management resources were analysed and synthesised, into a draft which then underwent iterative review and development. This is shown schematically below.



Identify definitive resources

Project team members' knowledge, supplemented by desk research, discovered a number of significant resources on the subject of records management and electronic records management.

Interestingly, virtually all the resources discovered were in the English language. The only exceptions were in Italian and Norwegian. The Italian language resource contained only legislative requirements, and was not used as a core resource. The Norwegian NOARK 4 specification appeared to contain highly relevant material, but its translation would have been beyond the resources allocated to the project ([3]).

The resources were analysed for relevance to a requirements specification; we found that some documents contained work which, though important, could not be applied directly to a requirements specification. Relevant resources, and resources required by the project contract, formed the core resources for MoReq, namely:

([3]) NOARK 4 is hundreds of pages long, and contains dozens of technical diagrams; translation could only realistically be performed by specialists involved in its production. A translation activity was started in Norway during the course of MoReq, and a partial translation is now available.

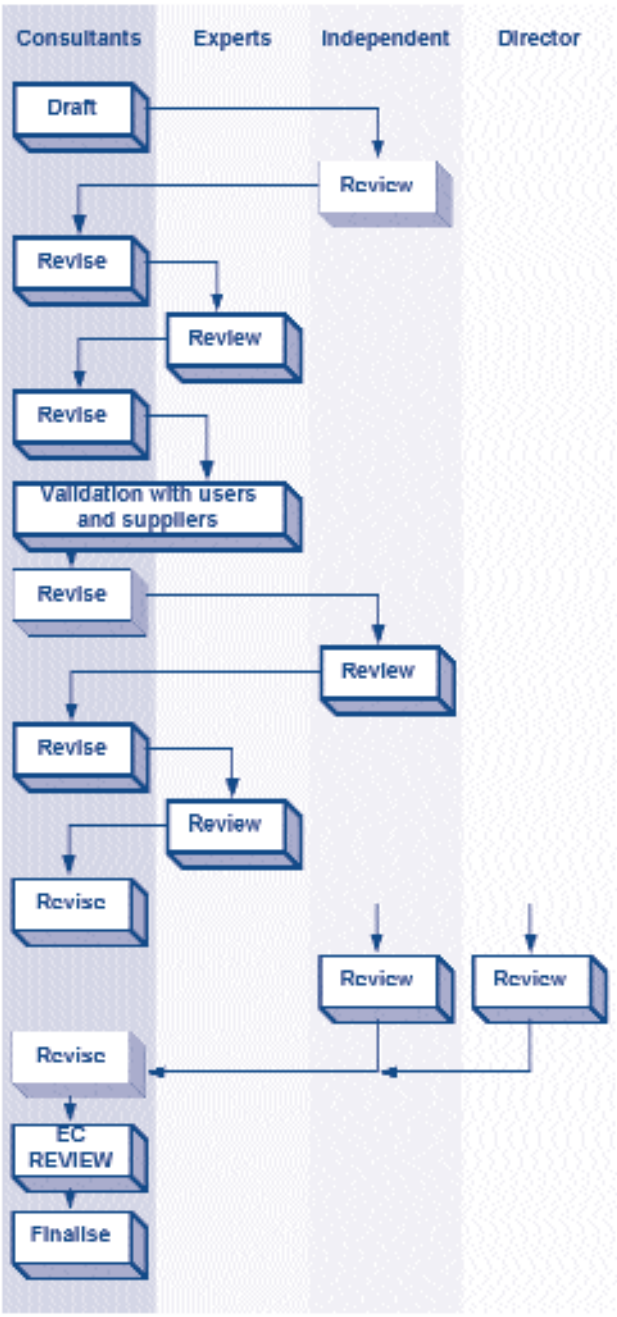
Name and ownership or source	URL or publication details
Dublin Core Metadata Element Set, Version 1.1: Reference Description	http://purl.oclc.org/dc/documents/rec-dces-19990702.htm or http://mirrored.ukoln.ac.uk/dc/
Functional requirements for electronic records management systems (GB Public Record Office)	http://www.pro.gov.uk/recordsmanagement/eros/invest/default.htm
Functional requirements for evidence in record-keeping (US University of Pittsburgh)	http://www.lis.pitt.edu/~nhprc/
Guide for managing electronic records from an archival perspective (Committee on Electronic Records, International Committee On Archives, ICA Study 8)	http://data1.archives.ca/ica/cer/guide_0.html
Code of Practice for legal admissibility and evidential weight of information stored electronically (British Standards Institution)	Published by British Standards Institution (www.bsi-global.com) as BSI DISC PD 0008
Guidelines on best practices for using electronic information (DLM Forum)	http://europa.eu.int/ISPO/dlm/documents/guidelines.html
ISAD(G): General International Standard Archival Description, Second Edition (Committee on Descriptive Standards, International Council on Archives)	http://www.ica.org/cgi-bin/ica.pl?04_e
The preservation of the integrity of electronic records (UBC-MAS Project) (University of British Columbia)	http://www.slais.ubc.ca/users/duranti/
Records Management, ISO 15489 (International Organisation for Standardisation)	To be published by the International Organisation for Standardisation; the standard was at the stage of a draft international standard at the time of writing.
Records/Document/Information Management: Integrated Document Management System for the Government of Canada — Request for Proposal — Requirements (RDIM) (National Archives of Canada)	originally published in 1996 ar http://www.archives.ca/06/4rdims.pdf; may now be unavailable, see also http://www.rdims.gc.ca/
Standard 5015.2 'Design Criteria Standard For Electronic Records Management Software Applications' (US Department of Defence)	http://jitc.fhu.disa.mil/recmgt/

(4) The MoReq project manager was fortunate to have been involved in the earlier development of the PRO specification.

This list does not claim to represent a complete resource list; it simply shows the resources available in early 2000 and selected as core for this project.

Synthesise draft

The core resources were analysed for content and structure. Specifically, the main sub-divisions (e.g. chapter headings) of relevant core resources were listed, compared and correlated. This led to the first structure — the table of contents — of the MoReq specification. This structure evolved as the project progressed, but remained recognisable throughout.



The first draft was then prepared. In this process, the UK Public Record Office (PRO) specification — being the closest in its objectives to the eventual objective of MoReq — was particularly influential (4). Requirements were taken from the core resources, adjusted as necessary to suit the concepts and terminology adopted in MoReq, and modified for the European and international context where necessary. Changes were introduced to allow for the electronic paradigm; these changes, intended to sweep away some of the limitations on record-keeping practice which has evolved with centuries of reliance on paper and other non-electronic records, proved to be contentious. The consultants also added requirements which did not appear in the core resources, e.g. for usability, scalability, workflow.

In parallel, a reference model and glossary were developed as a necessary medium of communication.

Review processes

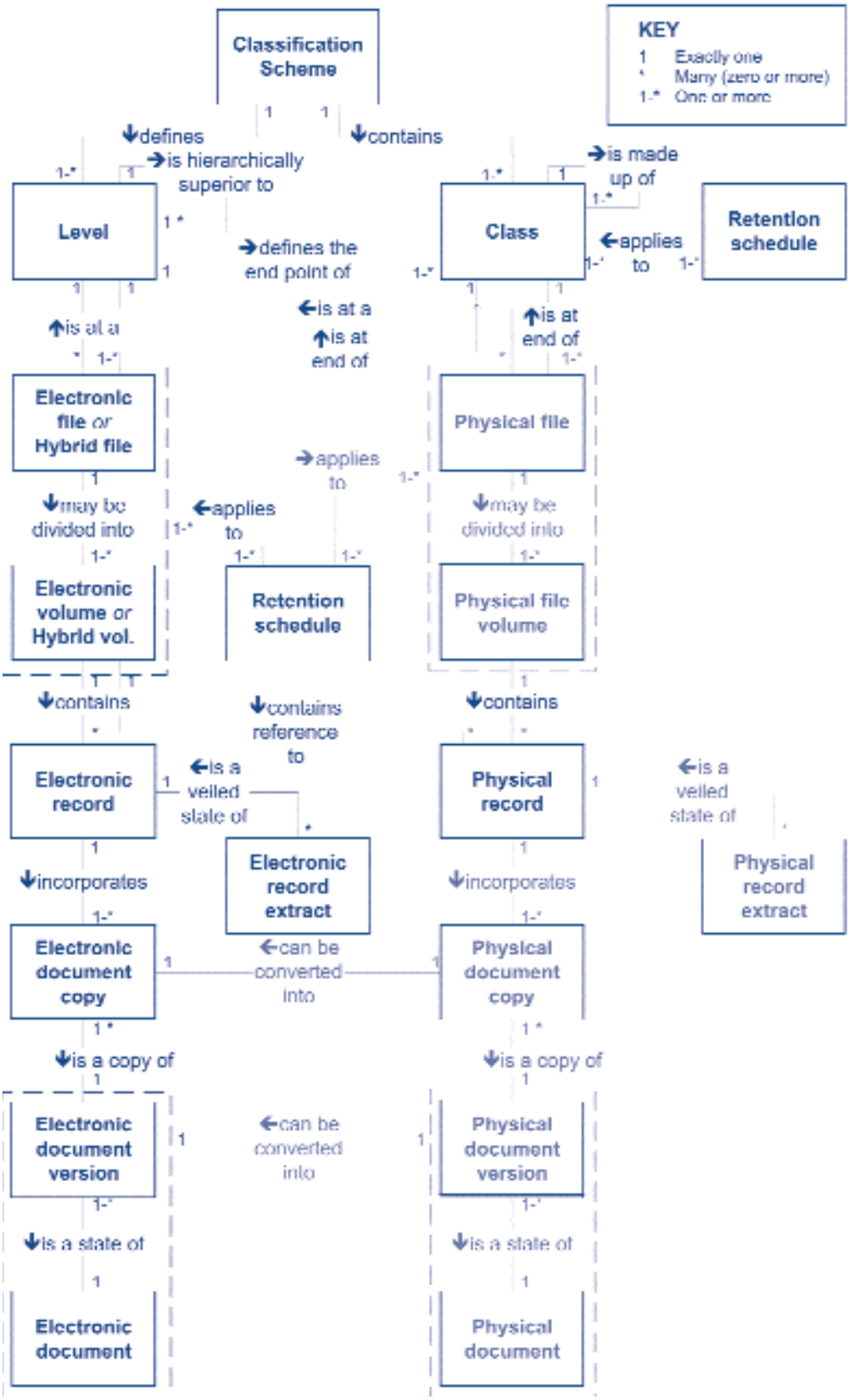
The review process was iterative. A key feature of the reviews was the input from, and debate with, the international team of reviewers. After the kick-off meeting at the start of the work, all the reviewing and debate took place through the medium of e-mail exchanges. The process also included reviews by a 'semi-independent reviewer', who had no other contact with the development process and so was able to perform reviews independently, without being influenced by design decisions. The first semi-independent review compared the MoReq first draft in detail against all the core resources, to ensure its completeness. The final review consisted of a quality assurance check by the project director.

The review process is illustrated diagrammatically below.

In the event, the level of reviews was even higher than described here; the Commission Project Officer contributed many invaluable insights, and the final review was supplemented by an additional consistency check by another consultant.

Entity-relationship model

The specification includes an entity-relationship model which is intended to clarify the meaning of some of the glossary definitions by demonstrating how some of the key terms are related. This is reproduced below.



Important features of this model include:

— the recursive relationship between levels;

— the recursive relationship between classes;

— the inclusion of non-electronic records and hybrid files;

— the concepts of ‘copy’ and ‘volume’.

Metadata model

The specification includes a list of the 127 metadata elements required to implement the requirements. This stops short of being a complete metadata schema, as it does not include, for example, domain-specific metadata needs; its scope is limited to those metadata elements explicitly called for by the specification. The 127 elements were identified by a detailed serial examination of the functional requirements.

The elements describe several entities and combinations of entities, namely:

- classification scheme;
- class and file;
- file or file volume
- volume;
- record;
- record extract;
- user;
- role.

Each element is described briefly, and its cardinality is shown. The requirements giving rise to each element are referenced ([5]).

One feature of the model which we believe to be unique is its recognition that some metadata elements may need to be associated with either files or file volumes, depending on the nature of the file.

Context of the MoReq metadata model

US DoD 5015.2

Arguably, the most widely-recognised specification for electronic records management in the world is the US Department of Defence's Military Standard 5015.2. The original 1997 version did not include a metadata model. However, the 2001 revision, which in early 2002 is still going through the final stages of approval, includes a number of separate tables which, together, constitute a simple metadata model. It lists only 50 metadata elements, none of which address aggregations; and fully one third of which concern security classification. This and several other features reflect the special needs of the defence sector.

UK Public Record Office specification

A more comprehensive model is found in the UK's Public Record Office 'Functional Requirements for ERMS', which is widely used as a de facto standard in the UK public sector. First published in 1999, it sets out minimum metadata for files, volumes, records, and users. In all, 73 elements are identified. At time of writing, a substantial revision is under way, due for publication in 2002. This is expected to feature a comprehensive metadata model for electronic records management, albeit specific to the needs of the British Civil Service.

Dublin Core and Pittsburgh models

The MoReq specification includes a brief appendix which relates the MoReq metadata model to two core references, the Dublin Core model and the Pittsburgh metadata model. The correlations are imperfect, reflecting the differences in the concepts underlying the models, for example the library orientation of the Dublin Core metadata set and the archival/records keeping orientation of MoReq.

Differing approaches to metadata

The differences between the metadata models briefly surveyed above is striking; the models include fundamental differences of scope and concept. It is consistent with the considerable differences between the specifications themselves. To be sure, they need not be identical, as they address different purposes and target audiences. Nonetheless these specifications contain a level of inconsistency which is surprising at times.

([5]) A limitation of the specification is that the cross-referencing is only unidirectional, i.e. it is not possible to see which metadata elements are associated with a requirement simply by looking at that requirement. This leads to the possibility of inconsistency when the specification is customised. The specification includes instructions intended to mitigate the effects of this limitation.

Beyond Europe

The MoReq specification aimed to address a wide range of record-keeping needs in Europe. It has been well received, suggesting that it met a real need. The same need exists elsewhere, from North America to the Far East. And some developing countries also have a growing need for a usable model.

Multinationals will face a problem from the increasing proliferation of electronic records specifications as they seek to implement ERMS, a key infrastructural application. They obviously are not keen to implement different packages, or different instantiations of one package, in different countries. They have a need for a common approach.

ERM software suppliers too are ill-served by the international diversity. They are faced with unwelcome choices if developing for an international market which has varying and sometimes incompatible requirements.

None of these needs are met at present.

It therefore appears that the time may be ripe to initiate a broader action which seeks to meet these needs in a broad, consistent and usable manner. Our knowledge of ERMS is now mature enough for us to conceive of a model specification which can be adapted for use beyond the borders of Europe. Development of this specification would require commitment from a variety of countries; it would represent a considerable challenge, but one we should not shirk. Is now the time to launch this initiative?

MoReq sobre metadatos: ¿más allá de Europa?

Mark Fresko

Antecedentes

La abreviatura MoReq procede de *Model Requirements for the Management of Electronic Records* (modelo de requisitos para la gestión de archivos electrónicos). Se trata de una especificación formal de requisitos para sistemas de gestión de archivos electrónicos (ERMS), que establece requisitos funcionales y no funcionales. La especificación está concebida para ser aplicable por igual a organizaciones de los sectores público y privado que deseen aplicar, desarrollar o evaluar capacidades de gestión de archivos electrónicos en cualquier país europeo.

La necesidad de una especificación amplia de requisitos para la gestión de archivos electrónicos se formuló por primera vez en el Foro DLM de 1996. Posteriormente, el programa sobre intercambio de datos entre las administraciones (IDA) de la Dirección General de Empresa de la Comisión Europea encargó el desarrollo de esta especificación modelo.

Tras la celebración de un concurso abierto en 1999, se adjudicó un contrato de desarrollo a Cornwell Affiliates Plc. Los trabajos sobre la especificación comenzaron en 2000 y concluyeron a principios de 2001. La especificación completa está ahora disponible para su descarga en distintos formatos desde varios sitios de Internet, incluido http://www.cornwell.co.uk/moreq.html.

Modelo de metadatos de MoReq

La especificación incluye una lista de los 127 elementos de los metadatos necesarios para aplicar los requisitos. No llega a ser un esquema completo de metadatos para aplicaciones específicas, pues no incluye, por ejemplo, las necesidades de metadatos por ámbitos específicos; su alcance se limita a los elementos de los metadatos explícitamente requeridos por la especificación.

Los elementos describen metadatos para varias entidades y combinaciones de entidades, a saber:

- Sistema de clasificación
- Clase y fichero
- Fichero o volumen de fichero
- Volumen
- Archivo
- Extracto de archivo
- Usuario
- Función

Cada elemento se describe brevemente y los requisitos que dan lugar a cada elemento se señalan con una referencia. Una característica del modelo que consideramos única es su reconocimiento de que algunos elementos de los metadatos pueden tener que asociarse con ficheros o con volúmenes de ficheros, dependiendo de la naturaleza del fichero.

Contexto del modelo de metadatos MoReq

DoD 5015.2 de Estados Unidos

Aparentemente, la especificación más ampliamente reconocida a escala mundial para la gestión de archivos electrónicos es la norma militar 5015.2 del Departamento de Defensa estadounidense. La versión original de 1997, que sigue vigente, no incluye un modelo de metadatos. Sin embargo, la revisión de 2001, que a principios de 2002 sigue sujeta a las últimas fases de la aprobación, incluye varios cuadros separados que, juntos, constituyen un sencillo modelo de metadatos. Se enumeran unos 50 elementos de metadatos, un tercio de los cuales relativos a la clasificación de seguridad. Esta y otras varias características reflejan las necesidades especiales del sector de la defensa.

Especificación de la Public Record Office del Reino Unido

En la Public Record Office (PRO) del Reino Unido existe un modelo más completo, los «requisitos funcionales para la gestión de archivos electrónicos», que se utiliza ampliamente como norma *de facto* en el sector público británico. Esta norma, publicada por primera vez en 1999, fija unos metadatos mínimos para ficheros, volúmenes, archivos y usuarios. En total se establecen 73 elementos. En el momento de elaborar el presente documento, esta norma está siendo objeto de una revisión sustancial.

Modelos de Pittsburgh y Dublin Core

La especificación MoReq incluye un breve apéndice que relaciona el modelo de metadatos MoReq con dos referencias básicas, el modelo Dublin Core y el modelo de metadatos de Pittsburgh. Las correlaciones son imperfectas y reflejan las diferencias de los conceptos que están en la base de los modelos, por ejemplo la orientación de biblioteca del conjunto de metadatos de Dublin Core y la orientación hacia el mantenimiento de archivos de los MoReq.

Distintos enfoques respecto de los metadatos

Las diferencias entre los modelos de metadatos anteriormente examinados son sorprendentes. Estas diferencias son coherentes con las grandes diferencias existentes entre las propias especificaciones. Es evidente que no necesitan ser idénticas, pues tienen distintos propósitos y audiencias. No obstante, estas especificaciones contienen un nivel de incoherencia que a veces sorprende.

Más allá de Europa

La especificación MoReq tiene por objetivo abordar una amplia gama de necesidades sobre mantenimiento de archivos en Europa. Ha recibido una buena acogida, lo que sugiere que cubre una necesidad real. La misma necesidad existe en todas partes, desde Norteamérica hasta Extremo Oriente. Y algunos países en vías de desarrollo también tienen una necesidad creciente de un modelo utilizable.

Las multinacionales se enfrentarán a un problema debido a la creciente proliferación de especificaciones sobre archivos electrónicos, a medida que intenten aplicar el ERMS (sistema de gestión de archivos electrónicos), que es una aplicación clave de infraestructura. Las multinacionales no son, obviamente, partidarias de aplicar distintos paquetes ni distintas concreciones de un paquete en distintos países. Necesitan un enfoque común.

Los proveedores de *software* de gestión de archivos electrónicos también se ven perjudicados por la diversidad internacional. Se enfrentan a opciones incómodas a la hora de desarrollar un producto para un mercado internacional con requisitos diversos y a veces incompatibles.

Actualmente no está cubierta ninguna de estas necesidades.

Parece, por tanto, que puede haber llegado el momento de emprender una acción más amplia que intente cubrir estas necesidades de manera amplia, coherente y utilizable. Nuestros conocimientos sobre ERMS están lo bastante maduros como para que podamos concebir una especificación modelo que pueda adaptarse a su uso más allá de las fronteras de Europa. El desarrollo de esta especificación requeriría el compromiso de varios países; representaría un considerable reto, pero no debemos acobardarnos. ¿Ha llegado el momento de lanzar esta iniciativa?

MoReq für Metadaten – über Europa hinaus?

Marc Fresko

Hintergrund

MoReq bedeutet **Mo**del **Req**uirements for the Management of Electronic Records (Musteranforderungen für die Verwaltung elektronischer Aufzeichnungen). Es handelt sich hierbei um eine formelle Spezifikation für funktionale und nichtfunktionale Anforderungen an Systeme zur Verwaltung von elektronischen Archiven (ERMS). Die Spezifikation gilt gleichermaßen für Organisationen des öffentlichen und privaten Sektors, die ERMS-Fähigkeiten in einem europäischen Land anwenden, entwickeln oder beurteilen wollen.

Die Notwendigkeit einer umfassenden Spezifikation für Anforderungen an Systeme zur Verwaltung von elektronischen Archiven wurde zum ersten Mal auf dem DLM-Forum im Jahr 1996 betont. In der Folge gab das Programm für den Datenaustausch zwischen Verwaltungen (IDA) der Generaldirektion Unternehmen der Europäischen Kommission die Entwicklung dieser Musterspezifikation in Auftrag.

Nach einer offenen Ausschreibung im Jahr 1999 ging der Entwicklungsauftrag an Cornwell Affiliates Plc. Die Arbeiten an dieser Spezifikation begannen in 2000 und wurden Anfang 2001

abgeschlossen. Die vollständige Spezifikation kann in verschiedenen Formaten von verschiedenen Webseiten, beispielsweise www.cornwell.co.uk/MoReq.html, heruntergeladen werden.

MoReq für Metadatenmodelle

Die Spezifikation enthält ein Verzeichnis der 127 Metadatenelemente, die notwendig sind, um die Anforderungen umzusetzen. Es handelt sich bei speziellen Anwendungen nicht um ein vollständiges Metadatenschema, da es beispielsweise keine domainspezifischen Metadatenanforderungen enthält; ihr Umfang beschränkt sich auf die ausdrücklich in der Spezifikation geforderten Metadatenelemente.

Die Elemente beschreiben Metadaten für mehrere Einheiten und Einheitenkombinationen, insbesondere:

— Klassifikationsschema,

— Kategorie und Datei,

— Datei oder Dateiband,

— Band,

— Archivstück,

— Archivauszug,

— Nutzer,

— Schriftrolle.

Jedes Element wird mit den entsprechenden Anforderungen kurz beschrieben. Ein Merkmal des Modells, das unserer Meinung nach einzigartig ist, ist die Anerkenntnis, dass einige Metadatenelemente je nach Art der Datei mit anderen Dateien oder Dateibänden verknüpft werden müssten.

Kontext der MoReq-Metadatenmodelle

US DoD 5015.2

Die wahrscheinlich am meisten anerkannte Spezifikation für die Verwaltung elektronischer Aufzeichnungen in der Welt ist die Military Standard 5015.2 des US-Verteidigungsministeriums. Die Originalversion von 1997 – die nach wie vor in Kraft ist – beinhaltet kein Metadatenmodell. Allerdings enthält die Fassung von 2001, die Anfang 2002 noch endgültig genehmigt werden musste, mehrere Einzeltabellen, die gemeinsam ein einfaches Metadatenmodell darstellen. Aufgeführt werden rund 50 Metadatenelemente, von denen rund ein Drittel Geheimhaltungsstufen betreffen. Dies und mehrere andere Merkmale verdeutlichen die besonderen Anforderungen des Verteidigungssektors.

Spezifikation des britischen Public Record Office

Ein umfassenderes Muster findet sich in den „Functional Requirements for ERMS“ des britischen Public Record Office, die als De-facto-Norm im öffentlichen Sektor im Vereinigten Königreich weit verbreitet sind. Die Erstauflage stammt aus dem Jahr 1999. Genannt werden Mindestmetadaten für Dateien, Dateibände, Archivstücke und Nutzer. Insgesamt werden 73 Elemente benannt. Zur Zeit der Drucklegung erfolgt eine grundlegende Überarbeitung.

Dublin-Core- und Pittsburgh-Modelle

Die MoReq-Spezifikation beinhaltet einen kurzen Anhang, in dem das MoReq-Metadatenmodell mit zwei Kernreferenzen, dem Dublin-Core-Modell und dem Pittsburgh-

Metadatenmodell, in Verbindung gebracht wird. Auch wenn einige Unterschiede in den Konzepten, die den Modellen zugrunde liegen, herausgearbeitet werden, beispielsweise der bibliotheksorientierte Ansatz des Dublin-Core-Metadatensets und der archiv- bzw. archivverwaltungsorientierte Ansatz der MoReq, sind die Bezüge dennoch wenig aufschlussreich.

Unterschiedliche Metadatenkonzepte

Zwischen den weiter oben kurz beschriebenen Metadatenmodellen bestehen erhebliche Unterschiede, die auf grundlegende Divergenzen auf Ebene der Spezifikationen selbst hindeuten. Um verlässlich zu sein, müssen sie keineswegs identisch sein, da sie unterschiedliche Ziele und Zielgruppen ansprechen. Gleichwohl sind diese Spezifikationen bisweilen erstaunlich inkohärent.

Über Europa hinaus

Mit den MoReq-Spezifikationen sollte europaweit eine Reihe von Fragen im Bereich der Archivkunde gelöst werden. Ihre freundliche Aufnahme bestätigt diesen Bedarf. Ein ähnlicher Bedarf besteht auch anderswo, von Nordamerika bis hin zum Fernen Osten. Auch in einigen Entwicklungsländern besteht wachsender Bedarf an tauglichen Modellen.

Multinationale Unternehmen, die ein System zur Verwaltung von elektronischen Archiven, d. h. eine Infrastruktur-Anwendung, an der quasi kein Weg vorbeiführt, einführen möchten, werden mit dem Problem der zunehmenden Verbreitung von Spezifikationen für EDV-Archive konfrontiert und vermutlich nicht daran interessiert sein, je nach Land unterschiedliche Software-Pakete oder verschiedene Instantiierungen desselben Software-Pakets zu verwenden. Sie benötigen vielmehr ein einheitliches Konzept.

Hersteller von ERM-Software beklagen ebenfalls die internationale Uneinheitlichkeit. Sie stehen vor schwierigen Entscheidungen bei der Entwicklung von Produkten für einen internationalen Markt mit unterschiedlichen und bisweilen inkompatiblen Anforderungen.

Auf keine dieser Fragen ist bisher eingegangen worden.

Daher scheint der Zeitpunkt gekommen, breiter angelegte Maßnahmen zu ergreifen, um diesem Bedarf umfassender, kohärenter und pragmatischer zu entsprechen. Unser Wissensstand über Systeme zur Verwaltung von elektronischen Archiven ist mittlerweile ausreichend, um ein Modell entwickeln zu können, das auch für Anwender jenseits der Grenzen Europas geeignet ist. Die Entwicklung dieser Spezifikation wäre ohne die Beteiligung zahlreicher Länder unmöglich; die Herausforderung wäre enorm, aber wir dürfen uns der Sache nicht entziehen. Ist jetzt der Zeitpunkt gekommen, um diese Initiative zu starten?

MoReq pour métadonnées: au-delà de l'Europe?

Mark Fresko

Contexte

MoReq (Model Requirements) désigne les spécifications types, fonctionnelles et non fonctionnelles, des systèmes de gestion des documents électroniques. Elles ont été rédigées de manière à être applicables aux organisations publiques comme aux organisations privées qui, établies dans n'importe quel pays européen, souhaitent mettre en œuvre, développer ou évaluer des capacités dans ce domaine.

La nécessité d'une définition exhaustive de spécifications en matière de gestion de documents électroniques a été exprimée pour la première fois lors du Forum DLM de 1996. La DG Entreprises de la Commission européenne y a donné suite dans le cadre de son programme IDA (Interchange of data between administrations).

À l'issue d'un appel d'offres ouvert organisé en 1999, un contrat de développement a été attribué à Cornwell Affiliates plc. L'étude des spécifications a commencé en 2000 pour s'achever au début de l'année 2001. Les spécifications complètes peuvent désormais être téléchargées en différents formats à partir de plusieurs sites web (notamment http://www.cornwell.co.uk/MoReq.html).

Modèle MoReq pour métadonnées

Les spécifications comprennent la liste des 127 éléments que doivent présenter les métadonnées pour satisfaire aux exigences. Il ne s'agit pas pour autant d'un schéma complet pour tout type d'application, dans la mesure où il n'intègre pas, par exemple, les besoins de métadonnées spécifiques dans certains domaines; il se limite aux éléments de métadonnées explicitement requis par les spécifications.

Les éléments décrivent les métadonnées pour diverses entités et combinaisons d'entités, à savoir:

- système de classification;
- catégorie et fichier;
- fichier ou dossier;
- dossier;
- document;
- extrait de document;
- utilisateur;
- rôle.

Chacun de ces éléments est brièvement décrit avec référence aux exigences sous-jacentes. Le modèle présente, à notre avis, une caractéristique unique, à savoir la reconnaissance du fait que certains éléments de métadonnées doivent, selon la nature même du fichier considéré, être associés à des fichiers ou à des dossiers.

Contexte du modèle MoReq pour métadonnées

Norme américaine DoD 5015.2

La spécification la plus largement reconnue à l'échelle mondiale en matière de gestion de documents électroniques est incontestablement la norme 5015.2 du ministère de la défense des États-Unis. Sa version initiale, qui date de 1997 et qui est toujours en vigueur, ne contient pas de modèle pour métadonnées. La révision 2001, qui était toujours aux derniers stades du processus d'agréation au début de 2002, inclut, en revanche, une série de tables distinctes qui constituent, ensemble, un modèle simple pour métadonnées. Cette nouvelle version énumère une cinquantaine d'éléments de métadonnées, dont le tiers environ concerne la classification de sécurité — un aspect qui, parmi d'autres, reflète les besoins spécifiques du secteur de la défense.

Spécification du PRO (Public Record Office) britannique

Les exigences fonctionnelles applicables aux systèmes de gestion de documents électroniques *(Functional requirements for ERMS)* du PRO proposent un modèle plus complet, qui

sert, de facto, de norme générale dans le secteur public au Royaume-Uni. Publié pour la première fois en 1999, ce modèle décrit les métadonnées minimales pour les fichiers, dossiers, documents et utilisateurs. Il identifie, au total, 73 éléments. Une révision fondamentale est en cours à l'heure où nous écrivons ces lignes.

Dublin Core Metadata Initiative et modèle Pittsburgh

Les spécifications types MoReq comportent une courte annexe consacrée à deux modèles de référence en matière de métadonnées, à savoir le Dublin Core et le Pittsburgh. Reflétant certaines divergences au niveau de la conception des modèles sous-jacents — orientation «bibliothèque» de la série de métadonnées Dublin Core et orientation «archivage» du MoReq, par exemple —, les corrélations sont néanmoins imparfaites.

Approches différentes des métadonnées

Il existe, entre les modèles pour métadonnées brièvement décrits ci-dessus, des différences frappantes qui traduisent de profondes divergences au niveau des spécifications proprement dites. La fiabilité de ces dernières n'exige nullement qu'elles soient identiques, puisqu'elles visent des objectifs et des publics cibles différents, mais le manque de cohérence entre elles atteint parfois un degré qui ne manque pas de surprendre.

Au-delà de l'Europe

Les spécifications types MoReq avaient pour but de répondre, au niveau de l'Europe, à toute une série de besoins dans le domaine de l'archivage, et l'accueil qu'elles ont reçu témoigne de la réalité de cette attente. Ces besoins existent également ailleurs, depuis l'Amérique du Nord jusqu'à l'Extrême-Orient, et tendent à se manifester aujourd'hui dans certains pays en développement en quête d'un modèle utilisable.

Les multinationales désireuses d'instaurer un système de gestion de documents électroniques — application infrastructurelle pratiquement incontournable — vont être confrontées au problème de la prolifération des spécifications en la matière et ne seront guère tentées d'utiliser des progiciels différents, ou des instanciations différentes du même progiciel, selon le pays. Elles ont besoin d'une approche commune.

Les fournisseurs de logiciels de gestion de documents électroniques souffrent, eux aussi, de ce manque de cohésion internationale. Comment prendre la bonne option lors du développement de produits destinés à un marché international caractérisé par des exigences diverses, voire incompatibles?

Aucune réponse n'est actuellement apportée à ces attentes et interrogations.

Le moment semble donc venu d'initier une action de plus grande envergure en vue de satisfaire ces besoins d'une manière plus large, plus cohérente et plus pragmatique. Notre connaissance des systèmes de gestion de documents électroniques est désormais suffisamment avancée pour que nous puissions développer un modèle capable de s'adapter à des utilisateurs situés au-delà des frontières de l'Europe. Un tel développement serait impossible sans la participation de toute une série de pays: le défi est donc majeur, mais nous ne devons pas nous y soustraire. L'heure d'une telle initiative n'a-t-elle pas sonné?

Workshop 2

Tuesday 7 May 2002

THE USE AND IMPLICATIONS OF DIGITAL SIGNATURES

Chairperson: Seamus Ross (United Kingdom)
Co-chair: Peter Berninger (European Commission)

Maike Bielfeldt

Ms Maike Bielfeldt was born in Hamburg (Germany) on 4 July 1970 where she went to school from 1976 to 1989. After graduating from high school she went through an apprenticeship as banking clerk in Hamburg. From 1991 to 1996 she studied Economics at the University of Hamburg (Germany) and Valencia (Spain).
Ms Bielfeldt entered the Association of the German Chambers of Commerce and Industry DIHT in November 1996 as trainee. In May 1997 she became the managing director of the subsidiary DE-CODA GmbH, also serving as the assistant to the CEO of the DIHT from January 1998–December 2001. Since January 2001 she has been serving as director of the IHK24 Ltd, where she could expand her activities in the e-business world. Working in the sector of security in the Internet and in electronic commerce Ms Bielfeldt has gained extensive knowledge and experience in this field, has established excellent contacts throughout Europe and is in constant communication with organisations and cooperations in her respective area of expertise.

([1]) Due to unforeseen circumstances, Ms Bielfeldt could not attend the DLM-Forum 2002.

The significance of the digital signature for European business

Maike Bielfeldt ([1])

Digital signatures have become an essential tool in today´s new marketplace, the Internet. Legislation has followed suit. The digital signature is now as powerful as the handwritten signature was in the past. This underlines the importance and impact of the digital signature for any kind of on-line activity, ensuring authenticity and integrity of any set of data: The acceptance of the digital signature will allow for new and better applications which will facilitate business procedures, knowledge management and administrative processes and open the realm of previously unheard-of possibilities. This in mind, DE-CODA was created in 1997 to proactively promote the digital signature. DE-CODA (http://signatur.ihk.de) together with her partners has come up with various applications first and foremost for the German Chamber of Commerce and Industry Organisation — and is currently reaching out to other clients as well.

Applications

DE-CODA, by now one of the major German players in the field of digital signatures, has successfully developed and implemented various applications supported by the digital signature.

Supplying the German Chambers of Industry and Commerce with software solutions based on the digital signature, DE-CODA currently runs two major projects one of which has become a successful real-time operation this very month. Both applications have found wide acceptance within the German Chambers of Industry and Commerce and will eventually replace all related paper-based procedures.

The advantages of the electronic processing of administrative data are obvious:

- access to forms is facilitated through the Internet;
- incomplete and inconsistent data is automatically rejected by the application;
- the applicant can check on the status of his application at any time; a history will show data regarding time, date, changes to the document, etc;
- the online submission saves time and money (transportation, manpower);
- once entered, the data can be used in different contexts and for different procedures (in observance of the data protection laws);
- an integrated protocol of authenticity accompanies each communication and thus allows for an identity check by the applicant or the CCI employee.
- the applications can be used by any company independently of its size and trade;
- about 300 000 vocational training contracts and about 800 000 certificates of origin underline the potential of our applications.

Certificates of origin

The ‘certificate of origin’ is one of the most important export documents in international trade. Our award-winning application ‘certificates of origin’ (ISSE Award 2000) promotes online submission, processing, approval and online modification of certificates of origin across borders. To complete this task successfully the digital signature is an essential component, ensuring the identity and authenticity of the data transmitted. The application ‘certificate of origin’ allows for the electronic submission of those certificates. They are printed out in the Chambers of Commerce and Industry, and — if authorised — will be sent to the company. The advantages are obvious compared to the current standard procedures: The application saves time and means less work. But

not only the application process but also the processing of certificates becomes easier. With this innovative application it is possible to take part in any step of the processing. Immediate responses and comments during the process save additional time. It is not necessary anymore to send the document back and forth. Additionally, a list is generated which shows the authorised or declined application for the certificate of origin. Thanks to the digital signature the online processing here has become possible. An integrated authentification protocol allows for easy verification of each party participating in the application process. You can have a look at this application at http://signatur.ihk.de/index.html

Vocational training contracts

Every year more than 300 000 new vocational training contracts are collectively registered at the German Chambers of Commerce nationwide — as required by German law. Today, our second pilot project promotes the electronic transmission of the contracts to the Chambers of Commerce and Industry as well as their digital recording. The application of digital signatures has become a vital factor in achieving this goal, ensuring authenticity of the document as well as its integrity.

'IHK24'

The German Chamber of Commerce and Industry Association features yet another project, which enjoys the support of DE-CODA's digital signature. The 'IHK 24' started out in Hamburg (http://www.ihk24.de) and was created with the intention of becoming an online platform for all of the Chambers of Commerce and Industry and their respective member companies, being accessible 24 hours a day. The online-platform offers data and facts important to the business world. It features a newsletter, a multitude of databases, business statistics, an Internet shop, job listings, a calendar of events, an intranet and links to other helpful tools within the Chamber of Commerce and Industry world. The ChamberCard serving both as a membership card in the real world and as an ID in the virtual world applies the digital signature. It allows members to create their own protected networking systems and thus ensures protected and safe business communication.

The key terms for our future endeavors will be 'interoperability' of both the different digital signature providers and their applications and a closer linkage of the existing projects to a knowledge management system which is currently being developed.

'Octane'

Another key feature is the international perspective of our approach. 'Octane', www.octane24.com, a project initiated by different European players addresses the needs of European and international SMEs (small- and medium-sized companies) that face prohibitive costs when trying to establish trustworthy relationships with international business partners. 'Octane' will deploy an open contracting service enabling SMEs to negotiate and establish contracts electronically within a secure environment. DE-CODA will help those partners to achieve that goal.

Octane defines 'electronic contracting' as the complete process that is required to achieve a legally supported business relationship that is accompanied by an electronic contract as the common and neutral representation for all those obligations upon which the involved parties have agreed.

Electronic contracting needs a secure technical and legal framework. Without the respective security technologies underlying the electronic transactions and contracts, the risk of manipulation is obvious. It is well known that each online transaction without any trusted and secure framework is a risk to the acceptance of online trading. The need for supporting mechanisms and service concepts that provide security to electronic transactions for buyers and sellers is clearly evident. Internet dealers have long awaited legislation that now assigns the same legal power to both the hand-written and the qualified digital or electronic signature. The option to finally agree and sign the electronic contract on the basis of a binding legal framework is a major necessity within the negotiation management process.

Octane closes the gap between supporting electronic commerce processes with applicable innovative products and services for the complete electronic contracting process, including a secure

(2) Such as ebXML, Java, X.509.

infrastructure and smart card-based signatures. Some aspects referring to the underlying legal background for electronic contracting, however, cannot fully be answered by this project. While the ad-hoc establishment of virtual organisations is often addressed by organisational and technical support, Octane utilises modern public key and certificate management services and takes the exploitation of this technology one step further: towards the SMEs via trusted organisations such as the Chambers of Commerce and Industry (CCI).

How does 'Octane' work? A contract editor supports and facilitates the composition and negotiation of bilateral or multilateral contracts in electronic form, based on a repository of model contracts and templates for different business scenarios. The contracting service also comprises a coordinating function, the electronic notary, to organise and monitor the signing process secured by the electronic signature. This component acts as a trusted 'third' party not only for contracts that require a physical notary public under current legislation, but rather for any commercial transaction that demands supporting services for non-repudiation, fair exchange of signatures, and for the electronic archiving of contracts.

In order to validate, transfer and promote the technology for the end-user, the project is currently establishing major trials in the business-to-business (virtual organisations, trade companies) as well as in the business-to-administration environment (CCIs).

From a business perspective, Octane accurately addresses the fundamental needs of commerce and trade, which are to 'find reliable information', to 'establish trust', and to 'negotiate and document an agreement'. The ability to execute these processes efficiently will allow for a measurable financial advantage in terms of expenses saved. Moreover, for the majority of cases it means the difference between success and failure. With Octane, the cost barrier for performing all these basic actions becomes rather irrelevant, and expanding regional trade and cooperation becomes feasible even for the smallest of companies.

The Octane system architecture supports all these different functions in an integrated, unified way, based on common electronic commerce standards (2), allowing interoperability between the different service components. Octane also contributes to the acceptance of digital signatures by deploying software and smart card-based digital signatures, depending on the requested level of security.

The main components of the Octane system architecture thus are:

- a business information service (BIS), needed to support companies in evaluating potential contractors. Standardised business information, provided online and in real-time, helps evaluate business register information and power-of-attorney credentials in order to verify roles and rights of signatories. The BIS component is required at the beginning of the electronic contracting process, i. e. the information phase, when companies search for trustworthy transaction partners;
- the open contracting service (OCS), which is a server-based software component designed to support contract composition and negotiation. This application allows organisations to set up and negotiate bilateral or multilateral contracts as electronic documents in a dynamic virtual setting. The contract editor is supported by model contracts and templates for different business scenarios. This technology provides a radical reduction of 'time-to-contract' by avoiding any media break during the negotiation, signing, and performance phases;
- the electronic notary, used to verify electronic signatures and certificates of contracting parties. This component acts as a trusted 'third' party not only for contracts that require a 'physical' notary public under common legislation, but rather for any commercial transaction that demands supporting services for non-repudiation and fair exchange of signatures. The electronic notary also supports electronic archiving of contracts and coordinates the signing process.

Thus, the digital signature could become an essential and successful tool for promoting and fostering business and administrative contacts in Germany, within the European Union, and beyond.

The European Chamber network 'ChamberSign'

It is the intention of The European Chambers of Commerce and Industry to set up an international network of registration authorities which, together with the trust centre D-TRUST promote and distribute the digital signature application. The network of Chambers of Commerce and Industries at home and abroad guarantees availability and easy access for its member companies.

With this purpose in mind, 10 European Chamber organisations founded the association 'ChamberSign' in July 1999. Chambers of commerce everywhere play an important role in achieving confidence in electronic commerce by acting as registration authorities. They are able to achieve such a high level of trust thanks to their recognition as leading organisations in international trade: chambers of commerce produce and handle international documents, have set up standardised rules and regulations for international trade, are members of international chamber networks and act solely in the general interest of their member businesses.

The primary goal of ChamberSign is to achieve interoperability of digital signatures used by European companies (certified by the Chambers of Commerce) in cross-border business-to-business electronic commerce. European legislation giving legal value to digital signatures will be in place in all EU Member States in the near future, making the ChamberSign network an excellent infrastructure for secure cross-border electronic transactions between European companies.

The chamber organisations involved represent 579 local chambers of commerce and 13.4 million enterprises, most of which are SMEs. Further expansion of the network to countries in Europe and beyond is planned for the near future.

ChamberSign will ensure that enterprises that receive electronically signed data over the Internet from international companies are able to unmistakably check the identity of the sender, authenticate the message, verify its integrity and make it impossible for the sender to deny having signed it.

Because chambers of commerce have a wide experience in certification processes, are in the best position to effectively verify effectively the credentials of their member companies and have an international network of regionally based Chambers at hand, ChamberSign certificates can be expected to become an important confidence building tool for cross-border business-to-business electronic commerce. ChamberSign certificates will be employed for a wide variety of purposes, including the sending of electronic purchase orders and invoices to companies abroad, the conclusion of cross-border contracts over the Internet and electronic delivery of international trade documents, such as certificates of origin and ATA carnets.

Future goals

- to establish and market a chamber certificate for electronic signatures which would be issued by the brand ChamberSign and the local chambers of commerce throughout Europe;
- to promote a series of criteria by which Chamber registration authorities commission services from certification authorities;
- to provide a means by which chamber associations benefit mutually from their best practice policies;
- to cooperate and promote greater awareness of the benefits of e-commerce and to develop and market products and services of interest to small- and medium-sized enterprises (SMEs).

Promotional activities do have to join in with political activities in today´s world. Only then interoperability and standardisation not only of technical but also of procedural practices can become a reality that will foster business throughout the EU and beyond.

Importancia de las firmas digitales para la economía europea

Maike Bielfeldt (1)

(1) En el último momento, la Sra. Bielfeldt no pudo estar presente en el Forum DLM 2002.

Nuestra empresa, DE-CODA GmbH, se creó en 1997 para promover de manera dinámica la firma digital. DE-CODA (http://www.de-coda.de) y sus socios ya han realizado distintas aplicaciones, fundamentalmente para la Federación de Cámaras de Comercio e Industria (CCI) alemana, y proponen sus servicios como consultores a las CCI alemanas. Además, DE-CODA ha apoyado la creación de autoridades de registro en más de cuarenta CCI en Alemania. En colaboración con la autoridad de certificación D-TRUST, DE-CODA ofrece equipos de iniciación con tarjetas inteligentes. DE-CODA propone asimismo cursos de informática, desarrolla herramientas de relaciones públicas genéricas y personalizadas, y organiza seminarios, debates y presentaciones. La empresa también ha elaborado presentaciones completas ambulantes para Cámaras de Comercio e Industria y PYME. Por último, ayuda a las PYME a incorporar la firma digital en sus procesos de trabajo individuales, con vistas a una mayor eficacia.

La actividad empresarial electrónica requiere unas normas de seguridad muy elevadas. La utilización de mecanismos que permitan identificar claramente a los socios comerciales en un medio anónimo es primordial para que evolucionen las transacciones por Internet. La firma digital propone un nivel de calidad muy superior a los métodos biométricos. Gracias a una firma electrónica cualificada, la autenticidad (origen y creador de los datos) y la integridad (exactitud) de la información pueden definirse de manera segura, mientras que la codificación de los correos electrónicos constituye otra opción.

Al ser las firmas digitales una herramienta esencial del nuevo mercado, es decir, de Internet, tanto el legislador europeo como el legislador alemán han seguido esta tendencia. La Directiva de 1997, que fijó una norma europea para las firmas electrónicas, supuso un paso importante que contribuyó a preparar el terreno para el desarrollo de un comercio electrónico seguro.

La firma digital es ahora en Alemania (casi) tan importante como la firma manuscrita en el pasado. Esta evolución destaca la importancia y el impacto de la firma digital en todo tipo de actividades en línea, garantizando la autenticidad y la integridad de cualquier clase de información. Se trata de una cuestión especialmente importante para las PYME, que no solamente podrán participar en las transacciones de comercio electrónico, sino que también podrán comunicarse con sus socios nacionales e internacionales en un medio seguro y protegido, desde la búsqueda de socios fiables, pasando por la negociación, hasta la celebración de contratos. La firma digital también puede utilizarse en plataformas de Internet, permitiendo a las PYME crear su propio sitio web, así como redes intranet y extranet.

La empresa DE-CODA ha desarrollado y aplicado distintas aplicaciones que incorporan la firma digital, tanto para el mercado nacional como para el internacional.

La aplicación del «certificado de origen», que le valió una recompensa, facilita la promoción del tratamiento en línea, la autorización y la modificación de los certificados de origen más allá de las fronteras. La firma digital es un componente principal, que garantiza la identidad y la autenticidad de los datos transferidos. Otra aplicación, la de los «contratos de formación profesional», constituye otro ejemplo de éxito basado en la firma digital. Cada año, más de 300 000 nuevos contratos de formación profesional se registran colectivamente en las CCI alemanas, tal como dispone la legislación alemana. Actualmente, nuestro segundo proyecto piloto tiene por objeto fomentar la transmisión electrónica de los contratos a las CCI y su registro electrónico. La firma digital es vital para lograr este objetivo en un medio jurídico seguro.

La Federación de Cámaras de Comercio e Industria alemana está aplicando otros proyectos, que se benefician del soporte de la firma digital de DE-CODA. De esta manera, la platafor-

ma IHK 24 se concibió como una plataforma en línea para el conjunto de las CCI y sus empresas afiliadas, y es accesible las 24 horas del día. La *ChamberCard* confiere una autorización de acceso personalizada al mundo virtual de la IHK 24.

A nivel internacional, DE-CODA se ha dedicado activamente a desarrollar la compatibilidad entre las distintas firmas digitales y su reconocimiento mutuo. En 1999, las CCI europeas (Eurochambres) fundaron ChamberSign. Desde entonces, DE-CODA participa de manera activa y constructiva en esta institución, que apoya las iniciativas de normalización internacional, así como el reconocimiento mutuo y la aceptación de las firmas digitales existentes. El centro de certificación D-TRUST de DE-CODA proporciona ahora la firma digital a ChamberSign.

La red de CCI, a nivel nacional e internacional, garantiza la disponibilidad y la accesibilidad a las empresas afiliadas. Teniendo en cuenta este objetivo, una decena de federaciones europeas de Cámaras de Comercio constituyeron la asociación «ChamberSign» en julio de 1999. En todas partes, las Cámaras de Comercio desempeñan un papel esencial para instaurar la confianza en el comercio electrónico, actuando como autoridades de registro. Si son capaces de suscitar un nivel de confianza tan elevado, es porque son reconocidas como organizaciones vanguardistas en el comercio internacional: las Cámaras de Comercio elaboran y tratan documentos internacionales, han fijado normas y normativas armonizadas para el comercio internacional, están afiliadas a redes internacionales y actúan exclusivamente en interés general de las empresas afiliadas.

Las federaciones de Cámaras de Comercio involucradas representan 579 Cámaras locales y más de 13,4 millones de empresas, la mayoría PYME. La red debería seguir extendiéndose a otros países en Europa y más allá en un futuro próximo.

Las Cámaras de Comercio tienen una gran experiencia en procesos de certificación, están bien situadas para comprobar con eficacia los certificados de las empresas afiliadas, y disponen de una red internacional establecida en las regiones, por lo que los certificados ChamberSign podrían convertirse en una herramienta principal de confianza en el comercio electrónico internacional entre empresas. Los certificados ChamberSign se utilizarán con distintos fines, entre los que figuran la expedición de órdenes de compra electrónicas y facturas a empresas en el extranjero, la celebración de contratos internacionales por Internet y el envío electrónico de documentos comerciales internacionales, como certificados de origen y carnets ATA.

DE-CODA también colabora con ComNetMedia y sus socios internacionales. ComNetMedia desarrolla aplicaciones basadas en la firma digital, destinadas al mercado europeo. Así Octane, uno de nuestros proyectos internacionales, funciona como un servicio abierto de contratación, que permite a las PYME negociar y celebrar contratos por vía electrónica en un medio seguro. De esta manera, DE-CODA crea siempre nuevas sinergias con socios interesados en favorecer la extensión de la firma digital.

La experiencia de DE-CODA y sus distintas aplicaciones ponen de manifiesto que la firma digital permite una mayor eficacia tanto para el sector privado como para el sector público, a nivel nacional e internacional. De esta manera, la exposición se centrará en las condiciones jurídicas necesarias para favorecer la aceptación de la firma digital y su utilización en Alemania y en Europa, así como en su gran impacto en la actividad económica europea e internacional.

(1) Frau Bielfeldt konnte wegen kurzfristiger Verhinderung nicht am DLM-Forum teilnehmen.

Die Bedeutung der digitalen Signatur für die europäische Wirtschaft

Maike Bielfeldt (1)

Unser Unternehmen, die DE-CODA GmbH, wurde 1997 gegründet, um die digitale Signatur aktiv zu fördern. Gemeinsam mit ihren Partnern hat die DE-CODA GmbH (http://www.de-coda.de) bereits verschiedene Anwendungen vor allem für den Deutschen Industrie- und Handelstag entwickelt und fungiert als Beraterin für die IHK in ganz Deutschland. Darüber hinaus unterstützt die DE-CODA die Einrichtung von Registrierungsstellen in mehr als 40 verschiedenen IHK. In Zusammenarbeit mit der Zertifizierungsstelle (Trustcenter) D-TRUST bieten sie nunmehr sogenannte Starter-Kits auf Chipkartenbasis an. Zum Angebot der DE-CODA gehören Computerkurse, die Entwicklung allgemeiner und kundenspezifischer PR-Tools, Workshops sowie Gespräche und Vorträge. Zudem hat das Unternehmen komplette Roadshows für IHK und KMU entwickelt. Die DE-CODA hilft mittelständischen Unternehmen dabei, die digitale Signatur mit dem Ziel höherer Effizienz in die individuellen Arbeitsabläufe einzubinden.

Für E-Business bedarf es hoher Sicherheitsstandards. Die Anwendung von Mechanismen, die Geschäftspartner in einem anonymen Umfeld eindeutig identifizieren können, ist heutzutage für Fortschritte im Internet-Geschäftsverkehr zu einem entscheidenden Faktor geworden. Dabei bietet die digitale Unterschrift ein weitaus höheres Sicherheitsniveau als biometrische Verfahren. Mit der qualifizierten elektronischen Signatur lassen sich Authentizität (Datenursprung/Urheberschaft) und Integrität (Korrektheit) von Daten mit Sicherheit feststellen. Eine weitere Option ist die Verschlüsselung von E-Mails.

Digitale Signaturen sind zu einem wesentlichen Werkzeug auf dem neuen Marktplatz der Gegenwart, dem Internet, geworden, was auch eine entsprechende Änderung der europäischen und deutschen Rechtsvorschriften nach sich zog. Die EU-Richtlinie von 1997 stellte einen bedeutenden Schritt dar, der durch die Festlegung eines europäischen Standards für digitale Signaturen mit den Weg für einen sicheren e-Commerce ebnete.

In Deutschland besitzt die digitale Signatur heute (fast) die gleiche Rechtskraft wie die eigenhändige Unterschrift in der Vergangenheit. Daran wird die Bedeutung der digitalen Signatur und ihr Einfluss auf Online-Aktivitäten aller Art sichtbar, so dass Authentizität und Integrität aller Daten gewährleistet sind. Besonders wichtig ist diese Problematik für KMU, die damit in einer sicheren und geschützten Umgebung nicht nur an E-Commerce und E-Trade teilnehmen, sondern auch mit nationalen und internationalen Geschäftspartnern kommunizieren können, von der Suche nach zuverlässigen Partnern über die Verhandlungsführung bis hin zum Vertragsabschluss. Mit der digitalen Signatur lassen sich zudem Internet-Plattformen unterstützen, so dass es den KMU möglich wird, eigene virtuelle Websites, Intra- und Extranets zu erstellen.

Von der DE-CODA wurden verschiedene durch die digitale Signatur unterstützte Anwendungen für den nationalen und internationalen Markt erfolgreich entwickelt.

Die mit einem Innovationspreis ausgezeichnete Anwendung „Ursprungszeugnisse“ hilft bei der grenzüberschreitenden Online-Bearbeitung, -Genehmigung und -Änderung von Ursprungszeugnissen. Dabei bildet die digitale Signatur eine äußerst wichtige Komponente und sichert die Identität und Authentizität der übermittelten Daten. Ein weiteres Beispiel für den erfolgreichen Einsatz der digitalen Signatur ist die Anwendung „Berufsausbildungsverträge“. Jedes Jahr werden entsprechend den gesetzlichen Anforderungen bundesweit von den IHK mehr als 300 000 neue Ausbildungsverträge in Verzeichnisse eingetragen. Unser zweites Pilotprojekt fördert nunmehr die elektronische Übermittlung der Verträge an die IHK und deren digitale Eintragung. Dabei wurde die digitale Signatur zu einem entscheidenden Faktor bei der Erreichung dieses Ziels in einem sicheren Rechtsumfeld.

Beim Deutschen Industrie- und Handelstag laufen noch weitere Projekte, denen die digitale Signatur der DE-CODA zugrunde liegt. „IHK24“ wurde als rund um die Uhr erreichbare Internet-Plattform für alle Industrie- und Handelskammern und ihre jeweiligen Mitgliedsunternehmen konzipiert. Hier dient die so genannte „ChamberCard“ als persönliche Zugangsberechtigung zur virtuellen IHK24-Welt.

International fördert die DE-CODA die Kompatibilität der verschiedenen digitalen Signaturen und ihre gegenseitige Akzeptanz. Im Jahre 1999 gründeten die europäischen Kammerorganisationen den Verein ChamberSign. Seither ist die DE-CODA ein aktiver und konstruktiver Partner für diesen Verein, der die internationale Normung ebenso unterstützt wie die gegenseitige Anerkennung und Akzeptanz vorhandener digitaler Signaturen. Mittlerweile stattet D-TRUST, das Trustcenter der DE-CODA, ChamberSign mit der digitalen Signatur aus.

Das Netz der Industrie- und Handelskammern im In- und Ausland garantiert die Verfügbarkeit und den problemlosen Zugang für die Mitgliedsunternehmen. Ausgehend davon gründeten zehn europäische Kammerorganisationen im Juli 1999 den Verein „ChamberSign“. Indem IHK als Registrierungsstellen fungieren, spielen sie eine wichtige Rolle bei der Schaffung von Vertrauen in den elektronischen Geschäftsverkehr. Ein solch hohes Vertrauensniveau können sie erreichen, da sie als führende Organisationen im internationalen Handel anerkannt sind. Handelskammern erstellen und bearbeiten internationale Dokumente, haben standardisierte Vorschriften und Regelungen für den internationalen Handel aufgestellt, sind Mitglied internationaler IHK-Netzwerke und handeln einzig im allgemeinen Interesse ihrer Mitgliedsunternehmen.

Die beteiligten Kammerorganisationen vertreten 579 lokale Handelskammern und 13,4 Millionen Unternehmen, größtenteils KMU. Für die nächste Zeit ist eine Erweiterung des Netzes auf weitere Länder Europas und darüber hinaus geplant.

Da Handelskammern über breit gefächerte Erfahrungen bei Zertifizierungen verfügen, besitzen sie die besten Voraussetzungen für eine wirksame Prüfung der Zeugnisse ihrer Mitgliedsunternehmen und können ein internationales Netzwerk regionaler Kammern nutzen. Es ist davon auszugehen, das Zertifikate von ChamberSign zu einem wichtigen vertrauensbildenden Werkzeug beim grenzüberschreitenden elektronischen Geschäftsverkehr zwischen Unternehmen werden. ChamberSign-Zertifikate werden für die verschiedensten Zwecke zum Einsatz kommen, so bei der Versendung von elektronischen Bestellungen und Rechnungen an Unternehmen im Ausland, beim Abschluss grenzüberschreitender Verträge über das Internet und bei der elektronischen Zustellung internationaler Handelsdokumente wie z. B. Ursprungszeugnissen und Carnets ATA.

Die DE-CODA arbeitet jedoch auch mit ComNetMedia zusammen und wirkt an deren internationalen Projekten mit. ComNetMedia entwickelt auf digitalen Signaturen basierende Anwendungen für den europäischen Markt. So dient beispielsweise OCTANE, eines unserer internationalen Projekte, als offener Vertragsservice, mit dessen Hilfe mittelständische Firmen Verträge elektronisch in einer sicheren Umgebung aushandeln und abschließen können. Auf diese Weise stellt die DE-CODA ständig neue Synergien mit interessierten Partnern her, um ihre digitale Signaturanwendung zu fördern und weiterzuverbreiten.

Die Fachkompetenz der DE-CODA und ihre diversen erfolgreichen Anwendungen zeigen, dass die digitale Signatur einen wesentlichen Effizienzzuwachs für den privaten ebenso wie für den öffentlichen Sektor im nationalen und internationalen Maßstab bietet. Im Mittelpunkt des Beitrags werden somit die rechtlichen Voraussetzungen zur Förderung der Akzeptanz und des Einsatzes der digitalen Signatur in Deutschland und in Europa sowie die bedeutenden Auswirkungen der digitalen Signatur auf die europäische und die internationale Wirtschaft stehen.

Importance des signatures numériques pour l'économie européenne

Maike Bielfeldt [1]

(1) Étant empêchée au dernier moment, Mme Bielfeldt n'a pas pu être présente au Forum DLM 2002.

Notre société DE-CODA GmbH a été créée en 1997 pour promouvoir de façon dynamique la signature numérique. DE-CODA (http://www.de-coda.de) et ses partenaires ont d'ores et déjà réalisé diverses applications, d'abord et avant tout pour le compte de la fédération allemande des chambres de commerce et d'industrie, et proposent leurs services de consultants aux CCI allemandes. En outre, DE-CODA est venue en appui de la mise en place des autorités d'enregistrement dans plus d'une quarantaine de chambres de commerce et d'industrie en Allemagne. En collaboration avec l'autorité de certification D-TRUST, DE-CODA offre désormais des kits de démarrage faisant appel à des cartes à puce. DE-CODA propose également des cours d'informatique, développe des outils de relations publiques génériques et personnalisés et organise des séminaires, des discussions et des présentations. La société a également élaboré des présentations complètes itinérantes pour les chambres de commerce et d'industrie et les PME. Enfin, elle aide les PME à incorporer la signature numérique dans leurs processus de travail individuels, en vue d'une plus grande efficacité.

L'e-business nécessite des normes de sécurité exigeantes. L'utilisation de mécanismes permettant d'identifier clairement les partenaires commerciaux dans un environnement anonyme est devenue primordiale pour faire évoluer les transactions sur l'internet. La signature numérique propose un niveau de qualité nettement supérieur aux méthodes biométriques. Grâce à une signature électronique qualifiée, l'authenticité (origine des données/créateur) et l'intégrité (exactitude) des informations peuvent être définies de façon sûre, cependant que le chiffrement des e-mails constitue une autre option.

Les signatures numériques étant devenues un outil essentiel du nouveau marché d'aujourd'hui, c'est-à-dire l'internet, le législateur européen comme le législateur allemand ont suivi. En fixant une norme européenne pour les signatures électroniques, la directive de 1997 a posé un jalon important, qui a contribué à préparer le terrain au développement d'un commerce électronique en toute sécurité.

La signature numérique est devenue en Allemagne aussi importante ou presque que l'était la signature manuscrite par le passé. Cette évolution souligne l'importance et l'impact de la signature numérique sur les activités en ligne de toute nature, garantissant l'authenticité et l'intégrité de n'importe quelle série d'informations. C'est un enjeu particulièrement important pour les PME, qui non seulement pourront participer aux transactions d'e-commerce, mais pourront aussi communiquer avec des partenaires nationaux et internationaux dans un environnement sûr et protégé — depuis la recherche de partenaires fiables en passant par la négociation jusqu'à la conclusion des marchés. La signature numérique peut aussi être utilisée sur des plates-formes internet, permettant aux PME de créer leur propre site web et des réseaux intranet et extranet.

La société DE-CODA a développé et mis en œuvre diverses applications supportées par la signature numérique, aussi bien pour le marché national que pour l'international.

L'application du «certificat d'origine», qui lui a valu une récompense, facilite la promotion du traitement en ligne, de l'agrément et de la modification des certificats d'origine par-delà les frontières. La signature numérique est une composante majeure, qui garantit l'identité et l'authenticité des données transférées. Une autre application, celle des «contrats de formation professionnelle», constitue un autre exemple de réussite reposant sur la signature numérique. Chaque année, plus de 300 000 nouveaux contrats de formation professionnelle sont enregistrés collectivement auprès des chambres de commerce et d'industrie allemandes, ainsi qu'en dispose la législation allemande. Actuellement, notre deuxième projet pilote vise à encourager la transmission électronique des contrats aux CCI et leur enregistrement

électronique. La signature numérique est devenue vitale pour atteindre cet objectif dans un environnement juridique sûr.

La fédération allemande des chambres de commerce et d'industrie met en œuvre d'autres projets, qui bénéficient du support de la signature numérique DE-CODA. Ainsi, la plate-forme IHK 24 a été conçue comme une plate-forme en ligne pour l'ensemble des CCI et de leurs sociétés affiliées, et est accessible 24 heures sur 24. La ChamberCard confère une autorisation d'accès personnalisée au monde virtuel de l'IHK 24.

Au niveau international, DE-CODA s'est employée activement à développer la compatibilité entre les différentes signatures numériques et leur reconnaissance mutuelle. En 1999, les chambres de commerce et d'industrie européennes (eurochambres) ont fondé ChamberSign. Depuis lors, DE-CODA participe de façon active et constructive à cette institution, qui soutient les initiatives de normalisation internationale, ainsi que la reconnaissance mutuelle et l'acceptation des signatures numériques existantes. Le centre de certification D-TRUST de DE-CODA fournit désormais la signature numérique à ChamberSign.

Le réseau des chambres de commerce et d'industrie, aux niveaux national et international, garantit la disponibilité et l'accessibilité aux sociétés membres. Avec cet objectif à l'esprit, une dizaine de fédérations européennes de chambres de commerce ont constitué l'association ChamberSign en juillet 1999. Partout, les chambres de commerce jouent un rôle essentiel pour instaurer la confiance dans le commerce électronique en intervenant comme autorités d'enregistrement. Si elles peuvent susciter un niveau de confiance aussi élevé, c'est parce qu'elles sont reconnues comme des organisations de pointe dans le commerce international: les chambres de commerce produisent et traitent des documents internationaux, elles ont fixé des règles et des réglementations uniformisées pour le commerce international, elles sont affiliées à des réseaux internationaux et agissent exclusivement dans l'intérêt général des sociétés affiliées.

Les fédérations de chambres de commerce partie prenante représentent 579 CCI locales et plus de 13,4 millions d'entreprises, dont la plupart sont des PME. Le réseau devrait continuer à s'étendre à d'autres pays en Europe et au-delà dans un avenir proche.

Les chambres de commerce étant très expérimentées dans les processus de certification, étant les mieux placées pour vérifier efficacement les certificats des sociétés affiliées et disposant d'un réseau international de CCI implantées dans les régions, les certificats ChamberSign pourraient devenir un outil majeur de mise en confiance dans le commerce électronique international interentreprises. Les certificats ChamberSign seront utilisés à des fins diverses, parmi lesquelles l'expédition d'ordres d'achat électroniques et de factures à des sociétés à l'étranger, la conclusion de contrats internationaux par l'internet et la fourniture électronique de documents commerciaux internationaux, tels que les certificats d'origine et les carnets ATA.

Toutefois, DE-CODA collabore également avec ComNetMedia et ses partenaires internationaux. ComNetMedia développe des applications reposant sur la signature numérique destinées au marché européen. Ainsi, Octane, l'un de nos projets internationaux, intervient comme un service ouvert de passation de marchés qui permet aux PME de négocier et de conclure des contrats par voie électronique dans un environnement sécurisé. DE-CODA crée donc ainsi toujours de nouvelles synergies avec des partenaires intéressés pour favoriser l'expansion de la signature numérique.

Le savoir-faire de DE-CODA et ses diverses applications démontrent que la signature numérique permet des gains d'efficacité considérables tant pour le secteur privé que pour le secteur public, sur le plan national et international. Aussi l'exposé sera-t-il centré sur les conditions juridiques requises pour favoriser l'accueil de la signature numérique et son utilisation, en Allemagne comme en Europe, ainsi que sur son impact majeur sur l'activité économique européenne et internationale.

Raimo Pohjola

Born in 1942. After studying history, political science and economics at the University of Turku graduated in history 1967. After graduation, he became a lecturer at the University of Turku and studies at the Universities of Stockholm and Göttingen. In 1972 he moved to the National Archives of Finland. From 1974 to 1985 he was Director of the Provincial Archives of Joensuu. In 1985, he was again working in the National Archives of Finland as Head of Unit and in 1997 as Deputy Director-General. His current responsibilities include records management, IT matters and personal and economic administration.

Implications of electronic signatures — the situation in Finland

The act on electronic service in the administration and the act on electronic signature

Raimo Pohjola

The EU directive on electronic signatures and standardisation

The EU directive on electronic signatures (1999/93EC) was issued on 13 December, 1999, and it entered into force on January 19, 2000. Before that the Commission had already issued a communication in 1997 which required common legislation regarding the recognition of the electronic signature as well as the minimum requirements for certifiers.

In order to meet the requirements of the directive on electronic signatures, the European industries and standardisation organisations have established a joint project called EESSI. In the framework of the project several fields connected with the products, systems and services of the certifiers will be standardised. Standardisation work is based on the public key infrastructure which meets the requirements of safe electronic transactions.

Electronic signature in some countries

There have been projects on electronic government and electronic signature in certain European countries since the end of the 1990s. In the Nordic countries these issues were studied actively. The first pieces of legislation in this field were enacted in Italy and Germany. In Italy an act that aims at the simplification of administrative procedures came into force in April 1997. Section 15 of the act entitles the use of electronic service in the administration in ways that are specified in a decree. On the basis of the act mentioned above was enacted a decree on electronic signatures which came into force in March 1998. The act and the decree gave electronic records, signatures, contracts and payments a legal validity.

In Germany the Act on Electronic Signatures (Signaturgesetz) came into force already on 1 August, 1997. It was part of a larger act called the Act on Information and Communication Services (Informations- und Kommunikationsdienste Gesetz). The act contained stipulations, for example, on how to become an actor in the information and communication branch, on the duty to provide guidance, on how to cease activities in the field as well as on time stamps. The Act on Electronic Signatures has been replaced by a new act (Gesetz über Rahmenbedingungen für elektronische Signaturen und zur Änderungen weiterer Vorschriften). There is also a decree under preparation which will define the requirements set for providers of certification services.

In the United States the Federal Act on Electronic Signatures came into force on 6 January 2000. The act gives the electronic signature the same validity as the traditional handwritten signature has. The act does not require any certification services to be connected with the use of electronic signatures. What is required, however, is that there must be the consumer's explicit consent before electronic signatures can be used. The aim of the act is to make the use of electronic signatures and electronic records possible by stipulating that legal transactions cannot be denied legal effect on the grounds that transactions have been done electronically.

Also Sweden, Norway and Denmark have acts on electronic signatures. In Sweden the act came into force on 1 January 2001, in Norway on 1 July 2001 and in Denmark 1 October, 2000. These acts put into effect the EU directive on electronic signature. According to these national acts, an advanced electronic signature has the same legal effects as the traditional handwritten signature. Only in the Norwegian act is there a stipulation that also other electronic signatures than advanced ones can have the same effect as the handwritten signature. A Danish speciality in the field of legislation com-

pared with the other Nordic countries is the office of the so called revisor. The certifiers must name a third party called a revisor, authorised by the State, to inspect their systems.

Electronic service in administration and electronic signature in Finland

In Finland the Act on Electronic Service in the Administration entered into force on 1 January 2000, that is, soon after the directive was passed. The act contained provisions on the organisation of electronic services, on electronic signature and on certifiers. The European Commission has, however, pointed out that Finland did not take into account, while preparing the act, Directive 98/34/EC regulating the provision of the information concerning technical standards and stipulations. In order to amend this defect there are two legislative projects under way. The one will result in a new act on electronic service in the administration and the other in an act on electronic signatures and certification. The bills are in the Parliament at the moment. The acts will be passed at the same time and also at the same time will be repealed the Act on Electronic Service in the Administration which came into force on 1 January 2000.

Electronic signature and the method of public key

In Finland the electronic signature is based on the method of public key which is at the moment the most widespread technique to realise electronic signatures. It is possible to verify the identity of a signatory by using the certificate that has been given to him or her by the certifier. The signature, the certification and the certifier constitute together the infrastructure of the public key which enhances the reliability of electronic services. In this system the user has two keys, the one is private, the other public.

The keys function together so that a message which is encrypted with the public key can be opened by the private key and vice versa. The public key is, as its name indicate, public — available to all — for example in a register maintained by the certifier. The private key is in the possession of the signatory. The private and public keys are connected by a complex mathematical equation so that it is not possible to derive them from one another. The longer the Akey lengths@, the safer the method. The encryption based on the public and the private key is called asymmetric encryption and it enables, besides encryption, also the creation of an electronic signature.

Besides the method of public key, electronic signature is based on the so-called hash function. By using the hash function the message is compressed into a digest that has a certain length. Electronic signature is realised so that the signatory uses the hash function and creates thus a digest and encrypts it with his or her private key. The signatory sends the message with the encrypted digest to the receiver. The receiver opens the encrypted digest with the signatory's public key and creates also a digest concerning the same message with a special signature verification programme. By comparing the two digests it is possible to verify the integrity of the message, that is, that the message is not changed during the transmission.

The opening of the encrypted digest by the public key of the signatory proves that the signatory has the private key that matches the public key. A third party (certifier) has given to the signatory a certification and guarantees that this particular private key matching the public key belongs only to the signatory in question. When the public key is verified in this way the receiver of a message can be sure that the signatory of the message is the person mentioned in the certificate. The certificate and the public key can be sent, for example, along with the message or the receiver can get it from the certifier's reference service.

Central concepts

The central concepts are defined in the forthcoming two acts (electronic signatures, electronic service):

Electronic message

— information sent by way of electronic data interchange and easily stored as a written document.

Electronic document
— an electronic message which has an effect on the lodging or handling of a matter or on the service of a decision.

Electronic signature
— data in electronic form which are attached to or logically associated with other electronic data and which serve as a method of authentication.

Advanced electronic signature
— an electronic signature which meets following requirements:
(a) it is uniquely linked to the signatory; (b) it is capable of identifying the signatory; (c) it is created using means that the signatory can maintain under his sole control; and (d) it is linked to the data to which it relates in such a manner that any subsequent change of the data is detectable.

Signature-creation data
— unique data, such as codes or private cryptographic keys, which are used by the signatory to create an electronic signature.

Signature-verification data
— unique data, such as codes or public cryptographic keys, which are used for the purpose of verifying an electronic signature.

Certificate
— an electronic attestation which links signature-verification data to a person and conforms the identity of that person.

Qualified certificate
— a certificate which meets the requirement laid down in Annex I of the Directive 1999/93EC and is provided by certification-service provider who fulfils the requirement laid down in the Annex II of the Directive 1999/93/EC.

Certification-service provider
— an entity or legal or natural person who issues certificates.

Electronic signature and the National Archives Service

The prerequisites for using electronic signatures and certificates are good in Finland because the principles of free consideration of evidence and freedom of contract are part of our legal system. A court must decide what is the truth in a particular case after the careful consideration of all the evidence and matters connected with it. The free consideration of evidence applies also to electronic signatures and to the legal transactions based on them. The traditional signature does not necessarily have a decisive role when a document's weight as evidence is considered.

From the point of view of the National Archives Service the critical question is the preservation of the data in the register of certificates as well as the preservation of the electronic signatures. The present Act on Electronic Service in the Administration still in force stipulates that the data in the register of certificates must be preserved permanently. The Parliament wanted to guarantee in this way good personal data protection. The government's proposal for the act refers also to legal protection. It was referred especially to the possible need to establish later the validity of an electronic signature that has been used in a transaction belonging to the sphere of civil law. The point is, that in these cases it is not possible to rely on administrative archives.

The National Archives Service opposed this stipulation. First of all it was in conflict with the right of the National Archives Service to decide on the permanent retention of records. The permanent preservation of the data in the register of certificates was regarded as unnecessary and, especially in electronic form, technically difficult. The new act will not contain this stipulation. The new proposed retention period is 10 years which is not problematic from the point of view of the National

Archives Service. According to some opinions even this retention period is too long. It is regarded as especially problematic that the new act will give the certifiers the right to collect information from the list of revocations about the queries made in order to find out if a particular certificate is in force or not. This information is, however, not open.

There is both in the old Act on Electronic Service in the Administration and in the new one a regulation about the filing of an electronic document. An electronic document shall be filed in a manner allowing for the later verification of its integrity and authenticity.

In Finnish records management the tradition of registration and registry offices is very strong. According to the present Act on Electronic Service in the Administration, all incoming electronic records must be registered. The same regulation is included in the forthcoming act, too. An entry in a register has to include information on the time of the arrival of the record as well as on the confirmation of its integrity and authenticity. In connection with registration the sender of the document must be verified, too. Thanks to registration the long-term preservation of electronic signatures becomes unnecessary.

Although an electronic signature guarantees the integrity of an electronic record, the signature is always dependent on the source file. When the records creator needs to migrate the preserved files from the old format to a new one, the signature becomes obsolete.

The electronic signature has its greatest significance in the safeguarding of the integrity and authenticity of the document during the transmission between the sender and the receiver. According to the present knowledge of the National Archives Service, the electronic signature loses, to a great extent, its significance when:

— the verification of the authenticity and integrity of the document has taken place;

— the verification has been registered;

— the document or its unaltered, semantically intact content has been transferred to a reliable IT system of an authority.

Trustworthiness is based on the unbroken chain of transactions described above as well as on the reliable IT systems of authorities. The authenticity and integrity of records are safeguarded by the inner logic of the system, by the measures taken to ensure data protection, by the chain of transactions and by the inspections made on the system.

In case the authority of an electronic record as evidence is questioned later, the agency in question can justify it by referring to the relevant register (diary) entry as well as to the reliability of the archiving system required by the Act on Electronic Service in the Administration.

According to the both acts on electronic service in the administration, the National Archives Service is competent to provide instructions and guidance on diary entries, registration and filing of electronic documents.

Electronic signature and the long-term preservation of electronic records

On the basis of what have been said above, it seems unnecessary to preserve the electronic signature after the transmission between the sender and the receiver has taken place. However, the short- and long-term preservation of electronic records is problematic in many ways. An IT system must be able to preserve an electronic record semantically unaltered. It must also be able to cement the document, so to speak, and to prevent its manipulation.

There are two options here:

— control of use based on different access rights;

— encrypted point of comparison which guarantees the indisputability of a record.

The use of encryption would mean the application of a technique resembling electronic signature. The means of encryption could not, however, be the certificate used in electronic transactions because its time of validity is limited. In Finland an issued certificate is valid for three years.

In Finland there has also been some discussion about the possibility to sign electronically by the National Archives Service the whole corpus of the electronic records that are to be transferred to the Service in one go. The operation would be repeated when the certificate's period of validity expires. In case the data be migrated, the new signature would be needed, too.

In the Finnish National Health Service there is a new pilot project, jointly operated by several hospitals, to create a regional electronic archives. In the hospitals taking part in the project the doctors sign electronic patient records electronically. To the information that is signed in this way are added also codes needed to audit the integrity of the records as well as the time stamp of the signature. The time stamp is there to guarantee that the signatures made by the certificate stay in force although the certificate itself is one day obsolete or otherwise out of use. The signed information is locked both in the patient application and in the so called notary database. This procedure documents the responsibilities regarding the treatment of patients.

Some patient records (e.g. x-ray pictures) differ from those normally created in electronic transactions. These records are stored in the systems as images. In these cases it is not enough to guarantee semantic integrity but also the external constancy of the records must be safeguarded.

Implicaciones de las firmas digitales: la situación en Finlandia

Ley sobre el servicio electrónico en la administración y ley sobre la firma electrónica

Raimo Pohjola

La directiva europea sobre las firmas electrónicas se promulgó el 13 de diciembre de 1999 y entró en vigor el 19 de enero de 2000. Desde entonces, la legislación de transposición de la directiva se ha promulgado en Suecia, Dinamarca, Noruega y Alemania.

Por lo que respecta a Finlandia, la Ley sobre el servicio electrónico en la administración entró en vigor el 1 de julio de 2000 (puede consultarse una versión en inglés de la ley en la siguiente dirección Internet: http://www.om.fi/2838.htm). Esta ley será sustituida próximamente por otra nueva. También se ha previsto una ley especial sobre las firmas digitales. Los dos textos se adoptarán previsiblemente en la primavera de 2002.

Siguen abiertas numerosas cuestiones relativas a las firmas digitales. Una de estas cuestiones se refiere a la conservación permanente de las firmas digitales. No obstante, la legislación finlandesa puede aportar una solución a este problema. Esta solución se basa en la tradición del registro, muy arraigada en la administración finlandesa. En virtud del título 25 de la Ley sobre el servicio electrónico en la administración, todos los documentos electrónicos deben registrarse a su llegada. Esta inscripción en el registro debe hacer constar información diversa: fecha de llegada del documento y confirmación de su integridad y su autenticidad. El propio remitente del documento es objeto de una comprobación en el marco de este registro. Gracias a esta operación, la conservación a largo plazo de las firmas se vuelve innecesaria.

Si bien la firma digital garantiza la integridad de un documento electrónico, esta firma depende siempre del fichero de origen. Cuando el autor del documento convierta los ficheros conservados de un formato antiguo a otro nuevo, la firma quedará obsoleta.

En el proceso de conversión, una firma electrónica pierde su integridad y su autenticidad. Por eso es fundamental comprobar la autenticidad y la integridad de los documentos electrónicos desde el principio de su ciclo de vida (es decir, inmediatamente tras su recepción). En la práctica, esto supone comprobar que existe una firma digital adjunta al documento. Si tal es el caso, se registra la existencia de la firma. De esta manera, la autenticidad y la integridad de un documento electrónico se basan en el archivo y en las inscripciones en un registro. Durante la conservación a largo plazo, es importante conservar intactos el contenido semántico y la integridad de los documentos electrónicos. A partir de un procedimiento planeado y controlado, es decir, el registro y el archivo, es posible conservar la autenticidad y la integridad de los documentos. A largo plazo, las firmas electrónicas ya no serán necesarias, y por consiguiente, los problemas técnicos vinculados a su conservación permanente desaparecerán.

El registro y el archivo adquieren mayor importancia cuando se trata de archivos que deben conservarse a largo plazo. Es probable que un plazo de conservación de una decena de años como máximo no plantee ningún problema desde el punto de vista de las firmas digitales.

Se ha propuesto una solución que recurre al ASCII como un posible medio para resolver los problemas que plantea la conservación permanente de las firmas digitales. Se trata de elaborar un resumen a partir del fichero ASCII del documento. Estos resúmenes son posteriormente menos vulnerables a los efectos secundarios de las conversiones.

Los archivos correspondientes al Derecho privado plantean, por su naturaleza, un problema particular. La idea básica en Finlandia es que la identificación (o certificado) electrónica que se utiliza en los servicios electrónicos de la administración también debe poder utilizarse en las transacciones privadas. El problema es que los documentos creados en el curso de estas transacciones no figuran en archivos administrativos, lo que significa que será imposible comprobar su autenticidad e integridad a largo plazo.

No obstante, es probable que los documentos de Derecho privado sólo se conserven en forma electrónica por un tiempo limitado. Estos documentos son por ejemplo órdenes de pedido, recibos y contratos. No es útil conservarlos más de 10 años. El plazo de conservación máximo para documentos de este tipo podría ser, por ejemplo, 30 años. Por lo que se refiere al sector privado, es evidente que los documentos más cruciales, que requieren una conservación más larga, se elaborarán en papel y se firmarán a mano, al menos en un futuro próximo.

Auswirkungen der digitalen Signatur – der Stand in Finnland

Gesetz über elektronische Dienste in der Verwaltung und Gesetz über die elektronische Signatur

Raimo Pohjola

Die EU-Richtlinie zur digitalen Signatur wurde am 13. Dezember 1999 verabschiedet und trat am 19. Januar 2000 in Kraft. Seither wurden ähnliche Rechtsvorschriften in Schweden, Dänemark, Norwegen und Deutschland erlassen.

In Finnland trat das Gesetz über elektronische Dienste in der Verwaltung am 1. Januar 2000 in Kraft (eine englische Übersetzung findet sich unter http://www.om.fi/2838.htm). Dieses

Gesetz wird demnächst durch ein neues abgelöst; außerdem wird es ein spezielles Gesetz zur digitalen Signatur geben. Beide Rechtsakte werden wahrscheinlich im Frühjahr 2002 erlassen.

Im Zusammenhang mit digitalen Signaturen sind noch viele Fragen offen. Eine davon ist die Langzeitsicherung dieser Signaturen. Die finnischen Rechtsvorschriften können jedoch eine Lösung für dieses Problem anbieten. Die Grundlage dafür bildet die Tradition der Registrierung, die in der finnischen Verwaltung sehr stark verankert ist. Gemäß § 25 des Gesetzes über elektronische Dienste in der Verwaltung müssen alle eingehenden elektronischen Unterlagen registriert werden. Der Eintrag im Register muss Angaben zum Zeitpunkt des Eingangs der Unterlage sowie zur Bestätigung ihrer Integrität und Authentizität enthalten. In Verbindung mit der Registrierung muss auch der Absender des Dokuments verifiziert werden. Aufgrund der Registrierung ist hier die Langzeitsicherung digitaler Signaturen überflüssig.

Obgleich eine digitale Signatur die Integrität einer elektronischen Unterlage garantiert, ist sie stets von der Ursprungsdatei abhängig. Muss der Urheber die aufbewahrten Dateien vom alten in ein neues Format migrieren, wird die Signatur obsolet.

Während des Migrationsprozesses verlieren elektronische Signaturen ihre Integrität und Authentizität. Daher ist es wichtig, Authentizität und Integrität elektronischer Unterlagen gleich zu Beginn ihres Lebenszyklus' (also unmittelbar nach Eingang) zu verifizieren. In der Praxis ist also zu kontrollieren, ob dem Dokument eine digitale Signatur beigefügt ist. Ist diese vorhanden, wird dieser Sachverhalt registriert. Somit basiert die Authentizität und Integrität einer elektronischen Unterlage auf Registereinträgen und Archivierung. Während der Langzeitaufbewahrung müssen semantischer Inhalt und Integrität von elektronischen Unterlagen aufrechterhalten werden. Die Erhaltung der Authentizität und Integrität elektronischer Unterlagen ist auf der Grundlage eines planmäßigen und kontrollierten Verfahrens, d. h. Registrierung und Archivierung, möglich. Auf lange Sicht würden keine digitalen Signaturen benötigt, und es gäbe keine technischen Probleme mehr im Zusammenhang mit der Langzeitaufbewahrung digitaler Signaturen.

Hinsichtlich der Unterlagen, die langfristig aufbewahrt werden müssen, nimmt die Bedeutung der Registrierung und Archivierung zu. Man kann davon ausgehen, dass eine Aufbewahrungszeit von bis zu zehn Jahren bei digitalen Signaturen kein Problem darstellt.

Als mögliches Verfahren zur Lösung des Problems der Langzeitsicherung digitaler Signaturen wurde eine Lösung auf der Basis von ASCII vorgeschlagen. Aus dem ASCII-Format der Unterlagen sollte ein Digest angefertigt werden. Digests wären gegenüber den Nebenwirkungen von Migrationen weniger anfällig.

Eine spezielle Angelegenheit sind Unterlagen privatrechtlicher Art. In Finnland gilt der Grundgedanke, dass die gleiche elektronische Kennung (Zertifikat), die bei elektronischen Diensten in der Verwaltung zum Einsatz kommt, auch bei privaten Transaktionen verwendet werden kann. Hierbei besteht das Problem, dass sich die bei diesen Transaktionen entstehenden Unterlagen nicht in Verwaltungsarchiven befinden, was bedeutet, dass die Verifizierung ihrer Authentizität und Integrität langfristig nicht erfolgen kann.

Es ist jedoch anzunehmen, dass Unterlagen privatrechtlicher Art nur begrenzte Zeit in elektronischer Form aufbewahrt werden. Dazu zählen beispielsweise Bestellungen, Quittungen und Verträge. Hier ist eine Aufbewahrung von mehr als zehn Jahren nicht notwendig. Die längste Aufbewahrungsfrist für diese Art von Schriftgut dürfte bei etwa 30 Jahren liegen. Natürlich werden die wichtigsten Schriftstücke im privaten Bereich, die eine längere Aufbewahrungsfrist erfordern, zumindest auf absehbare Zeit auf Papier ausgefertigt und auf herkömmliche Weise unterschrieben werden.

Implications des signatures numériques: la situation en Finlande

Loi sur le service électronique dans l'administration et loi sur la signature électronique

Raimo Pohjola

La directive européenne sur les signatures électroniques a été promulguée le 13 décembre 1999 et est entrée en vigueur le 19 janvier 2000. Depuis, la législation de transposition de la directive a été promulguée en Allemagne, au Danemark, en Norvège et en Suède.

Pour la Finlande, la loi sur le service électronique dans l'administration est entrée en application le 1er juillet 2000 (une version en anglais de la loi peut être consultée à l'adresse suivante: http://www.om.fi/2838.htm). Cette loi sera prochainement remplacée par une nouvelle. En outre, une loi spéciale sur les signatures numériques est également prévue. Les deux textes devraient être adoptés au printemps de 2002.

De nombreuses questions entourant les signatures numériques sont toujours en suspens. L'une de ces questions porte sur la conservation permanente des signatures numériques. Cependant, la législation finlandaise peut fournir une solution à ce problème. Cette solution repose sur la tradition de l'enregistrement, très solidement ancrée dans l'administration finlandaise. En vertu du titre 25 de la loi sur le service électronique dans l'administration, tous les documents électroniques à l'arrivée doivent être enregistrés. Cette inscription au registre doit faire figurer divers renseignements: date d'arrivée du document, confirmation de son intégrité et de son authenticité. L'expéditeur du document fait lui-même l'objet d'une vérification dans le cadre de cet enregistrement. Du fait de cette opération, la conservation permanente des signatures à long terme devient superflue.

Même si une signature numérique garantit l'intégrité d'un document électronique, cette signature dépend toujours du fichier source. Lorsque l'auteur du document doit convertir des fichiers conservés d'un ancien format à un autre, la signature devient obsolète.

Lors du processus de conversion, une signature électronique perd son intégrité et son authenticité. C'est la raison pour laquelle il est fondamental de vérifier l'authenticité et l'intégrité des documents électroniques dès le début de leur cycle de vie (soit immédiatement à leur réception). Cette vérification implique dans la pratique de vérifier qu'une signature numérique est bien jointe au document. Si tel est le cas, la présence de la signature est enregistrée. L'authenticité et l'intégrité d'un document électronique repose ainsi sur des inscriptions dans un registre et sur l'archivage. Durant la conservation permanente des pièces, il importe de conserver intacts le contenu sémantique et l'intégrité des documents électroniques. À partir d'une procédure planifiée et contrôlée, c'est-à-dire l'enregistrement et l'archivage, il est possible de préserver l'authenticité et l'intégrité des documents. À long terme, les signatures électroniques ne seront plus nécessaires et, partant, les problèmes techniques liés à leur conservation permanente disparaîtront.

L'enregistrement et l'archivage gagnent en importance lorsqu'il s'agit d'archives à conserver sur une longue période. Il est probable qu'un délai de conservation d'une dizaine d'années au maximum ne pose aucun problème du point de vue des signatures numériques.

Une solution faisant appel à l'ASCII a été suggérée comme un éventuel moyen de résoudre les problèmes que pose la conservation permanente des signatures numériques. Il s'agit d'en produire un condensé à partir du fichier ASCII du document. Ces condensés sont ensuite moins vulnérables aux effets indésirables des conversions.

Les archives relevant du droit privé posent un problème particulier. L'idée de base retenue en Finlande est que l'identification (certificat) électronique utilisée dans les services électroniques de l'administration doit pouvoir être utilisée aussi dans les transactions privées. Le problème en l'espèce est que les documents créés au cours de ces transactions ne figurent pas dans des archives administratives, ce qui signifie qu'il est impossible d'en vérifier l'authenticité et l'intégrité à long terme.

Toutefois, il est probable que les documents relevant du droit privé ne seront conservés sous forme électronique que pendant une durée limitée. Ces documents sont par exemple des bons de commande, des reçus et des contrats. Il n'est pas utile de les conserver au-delà de 10 ans. Le délai de conservation maximal pour des documents de ce type pourrait être de 30 ans. Pour ce qui est du secteur privé, il est évident que les documents les plus cruciaux, qui nécessitent des durées de conservation plus longues, seront produits sur papier et signés à la main, au moins dans un avenir proche.

Digital signatures: not a solution, but simply a link in the process chain

James Currall

1. Introduction

Digital signature technology is maturing fast. There are many organisations that will act as certification authorities for public keys and the technology to use public and private keys is available from within a wide variety of packages.

Computers and digital technology can be harnessed to make things happen, to implement processes which must be carried out. It is the processes not the technology that is important. Technology should be used, as appropriate, to 'oil the wheels' of process rather than forcing processes to take on a particular form.

In many cases the 'digitisation' of processes has simply resulted from doing exactly the same thing using a computer as was done in the paper world. Sometimes this is the appropriate way to proceed, but in all cases the purpose of the process needs to be examined in order to establish an appropriate way of digitising it. Many things are done in a particular way in the paper world because it is the paper world and that imposes severe constraints. The digital world opens up new possibilities and often processes can be re-engineered in the digital world to be much less time-consuming or inefficient and with considerably less duplication of effort.

This situation allows attention to turn to the processes in which these technologies can effectively be used and to the issues that their use raises in terms of: public understanding, legal admissibility, practicality, public key certificate availability and longevity of the Digital Signature itself.

2. What does a digital signature represent?

A traditional written signature seeks to serve one or more of three purposes:

- authentication (establishing the identity of the signatory);
- integrity (that the document signed is unchanged since signing);
- non-repudiation (that the author can't deny having signed it).

These three functions work well with top copy documents, but in a world full of photocopies all sort of changes could have been made or signatures added from other documents, etc., unless each copy is individually verified by the parties involved.

2.1. What they represent

A digital signature is designed to serve the three purposes listed above, but not look like a traditional written signature. It relies on an action being performed using something that only the signatory has access to (a particular private key). It combines this with a unique attribute of the 'document' to be signed. Digital signatures vouch for the fact that the individual who 'signed' the 'document' has access to the private key and knows any passphrase that protects it. This is similar to someone having access to a bank card and the PIN which protects it.

The guarantee that the 'document' was signed by the person whose signature it is, relies on the security surrounding storage and use of the private key. In much the same way, the guarantee that money has been taken out of a bank account by the person whose account it is relies on the how and where the card and PIN are stored. At a higher level, someone accepting digital signatures needs to trust that they really belong to the person that they purport to belong to. This is achieved by the keys used for signing and checking signatures being countersigned by a person or body

James Currall

James Currall, BSc PhD, is Information Strategy Coordinator, Information Services, and Senior Research Fellow, at the Humanities Advanced Technology and Information Institute, University of Glasgow
James joined the University of Glasgow as a statistical computing consultant in 1989 and has contributed to the research work of over 30 departments in the university. He comes from a background of biology and statistics. Since 1994 he has been the manager of the User Service Team in the University Computing Service, with overall responsibility for training, documentation, web development, software support and advisory services. He was also deputy director of a UK-wide support centre for computers in the teaching of statistics.
In 1997 he broadened his interests into areas of information beyond numbers, when he jointly put together the university's information strategy framework with Michael Moss and subsequently he was made information strategy coordinator. From 1997–2000 he co-led (with Michael Moss) the effective records management project funded by the Joint Information Systems Committee (JISC) and Technology Application Programme (JTAP), which carried out fundamental research on management of organisational records in digital form.
The information strategy developments have lead on to intranet and security work and James has just finished a project looking at the use of MiddleWare technologies (particularly with respect to authentication and other security) to permit people to publish information in a secure environment without the need for them to have detailed technical knowledge. During 2000 and 2001 he was also involved in a project, with South Bank University, investigating information security culture.
For the last year, James has been a Senior Research Fellow of the Humanities Advanced Technology and Information Institute, where he is developing research into information access drawing on his Computing Service work. In

the last few years he has worked on and given talks on a wide variety of topics concerned with the use and potential of technology in teaching, statistics, information, security and various combinations of these.

that either knows the signatory or takes steps to check up on the details. The confidence that one can have in a digital signature depends crucially on this being carried out properly.

A digital signature does not, of course, say anything about the trustworthiness of that individual (but then neither do traditional written signatures).

2.2. How they work

Digital signatures are not direct analogues of their written cousins. The main differences are:

- they can be applied by anyone who gets access to the file required for encryption (called the private key) who also finds out the passphrase required to unlock them;
- they are not directly viewable, but require special software to produce and subsequently to check the signatures.

The first difference makes them more similar to the seals that were used in the past than traditional hand-written signatures, in demonstrating the authenticity of paper documents. The second means that there is considerable investment required in training, software installation, etc. Signatures are by their nature personal, the same one being used for all purposes, unique and cannot be revoked except in jurisdictions where hands can be judicially removed. A more helpful concept is perhaps the seal.

Written signature	Digital signature	Seal
Uses property intrinsic to signatory	Uses an external object	Uses an external object
Means to make it cannot be stolen	Means to make it can be stolen	Means to make it can be stolen
Means to make it cannot be copied (but an instance of a signature can be forged)	Means to make it can be copied	Means to make it can be copied
Difficult to establish order of signing	Order of multiple signatures can be established	Order of multiple seals can be established
Mechanism stays the same and can't be revoked	Mechanism can be destroyed (revoked) and substituted by an entirely different one	Mechanism can be destroyed (revoked) and substituted by an entirely different one
Same signature for all purposes	Different signatures for different purposes	Different seals for different purposes

Seals were used widely until the late nineteenth century to authenticate documents and still are for certain 'important' documents, such as a degree certificate. There were personal seals and seals associated with offices (or roles) held by individuals and with institutions (as there still are). The legend around the circumference of the seal indicated its nature (policy of use). The digital signature fits this model rather better than it does the written signature.

Once we develop the policy, and process behind the use of the digital signature, as we are beginning to do, we will see that there can be digital signatures for an organisation as a whole, units within the organisation and roles held by office holders, as well as individuals. These different entities will require their own digital signatures which will be used in accordance with practices built on policies which endow credibility. These will eventually become acceptable to the various authorities and the courts. They may be revoked and a record retained of their 'impression' within archives to validate subsequent authenticity. For all this to happen procedures need to be put in place that establish policies for the registration, use and distribution of digital seals (signatures) within an organisations.

2.3. Conclusion

Do digital signatures have similar characteristics to traditional written signatures? I suggest that digital signatures have considerably more attributes in common with the seals used in the past, than with written signatures. The consequence of this is that we need to see digital signatures in a rather different light from their written cousins and the way that we embed these technologies into processes must also be reviewed. In particular the fact that different signature may be used under different circumstances, for different purposes and when acting in different roles. A 'one signature fits all purposes' philosophy does not seem to be appropriate or practical.

3. Fitting digital signatures into process

Digital signatures do not in themselves solve problems in the exchange and handling of digital documents (and other materials). What they do is to allow processes to be devised where authenticity and digital documents are crucial elements. In many cases, much of the process may not be digital or automated. Examples from the University of Glasgow, such as our remote site examination results approval discussed in the first sub-section below, demonstrate this.

In a number of instances, digital authenticity can be established in ways that do not require digital signatures (or any form of public key infrastructure). This issue is discussed in the second subsection below.

There is a growing literature on how easy or otherwise a public key infrastructure is to use for encryption and the creation and verification of digital signatures. What is becoming clear is that a universal signature which serves all purposes in the way that the actual scribble of a hand-written signature does is unlikely to work with digital signatures. This issue is discussed in the third subsection below.

3.1. The use of PGP for examination results transfer at the University of Glasgow

3.1.1. The problem

The University of Glasgow, in common with many institutions has its activities spread over a number of geographically separated campuses. In 1999, a new campus on the site of the Crichton Hospital in Dumfries opened to students. After the first diet of examinations in January 2000, the results had to be transmitted to the central Registry in Glasgow by courier and the certified results, for display at Crichton, had to be transferred from the Registry to Crichton by courier. The round trip of exam results from the examiners meeting to the Registry and back took several days, whilst the same process takes only a few hours on the main University campus at Glasgow. The problem with doing the whole thing digitally using e-mail revolves around the following requirements:

- that the registry has the sheet of results signed by the examiners before the results are processed;
- that the exam results should be transferred securely to the registry;
- that the registry and the university Senate are confident that the results have not been altered in transit between the examiners and the registry.

The problem revolves around getting examination results and appropriate assertions that they represented what the examiners had agreed to from remote locations to the Registry in a timely manner.

The university registry approached the computing service to ask if a solution to this problem could be found. The computing service decided that a process involving public key encryption using the software package PGP (pretty good privacy) could provide the answer and in June 2000 PGP was used successfully for examination results transfer from Crichton to the registry.

3.1.2. PGP

PGP is a software package that has been freely available for anyone who wants to use it for non-commercial purposes. It provides for privacy, authentication, integrity and non-repudiation through the use of encryption and digital signatures.

3.1.3. Procedures

3.1.3.1. Setting up PGP at Crichton

A process was devised by which exam results could be transferred securely from Crichton to the registry and whereby the registry could have confidence that the results were identical to the ones on the sheets signed by the examiners. Documents were written, which outlined the set-up processes, the exam mark encryption and signing and the exam mark signature verification and decryption. PGP was installed on the computers of the staff who were to take part in the results transfer process and a presentation was given to the majority of the academic and administrative staff at Crichton on the examination-mark transfer process and the use of PGP in it.

3.1.3.2. Using PGP for examination mark transfer

The most difficult tasks were to work out the procedures to follow when installing the technology, exchanging keys and actually using PGP for examination mark transfer. There is a need to ensure that no-one involved is in a position to falsify any of the keys or, once the procedures are in use, the exam results themselves. Confidence is required that the results:

- were not tampered with in transit;
- were not viewed in transit;
- came from the right person;
- were genuine and accorded exactly with the paper copy which the examiners signed at the end of their meeting.

In this example, a responsible individual not associated with the exam itself undertakes to forward a spreadsheet containing a copy of the marks approved by the examiners to the registry. In attaching their digital signature, they are explicitly asserting that the exam results are as on the examiners sheet, that the results have been signed by the examiners and that they have the paper copy signed by the examiners. The individual forwarding the results is acting as a surrogate for the examiners, which simplifies the technology required to complete the process. Each examiner could have been set up with a key pair and they could all produce digital signatures of the file containing the results, but there are a number of difficulties with this and a process of this nature would be complex, time consuming and error prone and would likely fail.

In summary the procedure followed at Crichton and which will be implemented elsewhere is as follows:

One person (from a small pool of approved signatories) takes responsibility for a particular set of exam results. This person is not an examiner for the exam in question. This signatory produces a digital signature for the file of examination results using their private key. Signatories are made aware that their signature attests to the fact that they have in their possession a printed copy of the examination results signed by the examiners at the close of their meeting and that the file is complete, accurate and genuine. This digital signature along with the file of results, encrypted with the Registry Examinations public key, is sent to the registry who use their private key to decrypt it. The registry checks the digital signature using the corresponding public key to ensure that the file is unaltered and that it was signed by the right person. If everything verifies, they proceed to process the exam results and print them on a printer in a secure room at Crichton a short time later.

3.1.4. Evaluation

The technology is fairly easy to set up and use. PGP guides the user through the installation process and key pair creation. The importance of fully understand the implications of having a private key and keeping it secure should not be underestimated, nor should a thorough understanding of what it means to make a signature using it. The greatest challenge in this project was working out the procedures to follow and then fitting the technology into the process. Devising and documenting the process and getting the approval of the Clerk of Senate for it took about three days of effort, whilst installation and training took a half day.

3.1.5. Conclusions

Technology such as this can support digital processing when there is a particular application in mind and the procedures are well designed and understood. It is easier to manage encryption and digital signatures when the people using it are known to each other and can trust each other's public keys and where there are a small number of people who need to have PGP and manage their own private keys.

3.2. Why alternatives?

At first sight it may seem perverse to tackle the issue of introducing a PKI into an organisation by discuss alternatives. However, PK techniques provide a technology, not a solution and whilst it could be applied to a wide variety of problems, it might not be the most cost-effective, error-free or easily managed solution to those problems. This section examines some of the issues.

3.2.1. Cost of PKI

Taken at face value, the cost of using public key encryption technologies is very low. There is good, free software available for use on desktop and server machines to generate key pairs, sign/verify and encrypt/decrypt. If one individual wishes to use the technology in their communications with another individual, the cost is minimal (assuming the motivation and a basic level of understanding). The two people concerned simply install the software, exchange public keys and get on with it. If the situation involves a group of a dozen people who know each other, then they all have to exchange public keys (66 exchanges). Even with people who know each other the problem of key exchange quickly gets out of hand.

Two developments are then needed to take matters forward for large groups of people and to extend the scope to people who don't know each other:

Key countersigning

Each public key needs to be countersigned by a party that is trusted by all the people involved (a certification authority), so that each individual does not have to check that each signature or encryption key does belong to the person it purports to belong to.

Key servers

Each person's public key need to be stored on a publicly accessible key server from which the encryption and signing software can retrieve it when required, rather than relying on all parties exchanging keys with each other. This development relies on the previous one being in place.

The two developments described above are what is meant by a public key infrastructure (PKI). Running a PKI has costs associated with it:

- the cost of servers on which to run the key server and certification authority activities (the former may coexist with other services, but the latter needs very tight security and probably needs a dedicated machine);
- the cost of designing the process of issuing and countersigning keys for individuals. The credibility of the public key set-up rests on getting the process right, so that relying parties know that the 'document' was actually signed by the person that they think it was;
- the cost of actually issuing, checking identities and countersigning keys for individuals. There are elements of these tasks which are difficult to automate in a credible way and therefore require human intervention.

If there are significant costs involved, there must be sufficient payback in benefits, efficiencies, etc.

3.2.2. PK understanding

Public key encryption techniques are easy once you understand fully what is going on. There are elements of the techniques which to many people appear counter-intuitive. Some rapidly get over this problem, others do not. For PKI to work on an institutional scale, there must be a 'critical mass' of understanding.

3.2.3. PK software

The software available for public key techniques on desktop machines is relatively easy to use, in that it has all the usual pull-down menus, drag and drop, dialogue boxes, etc. There have been a number of studies which have concluded that significant improvements in usability are needed before the 'average' person can cope adequately (Whitten and Tygar 1998, Chadwick, Tassabehji and Young 2000). The cost of providing training to overcome this problem may be considerable.

3.2.4. Risks of 'failure'

If people do not fully understand what they are doing with PK technology and do things incorrectly, the result may be much worse than not having used the technology at all. Material which people would not have considered sending by e-mail without the technology may accidentally be sent unencrypted whilst the sender believes that it is encrypted. People might think that they have applied digital signatures to documents and not have in fact done so, making verification of actions or auditing almost impossible. Failures in this way could have considerable costs attached to them.

3.2.5. An example of an alternative — expense claims

Traditionally in many institutions an expense claim involves:

(1) the claimant filling in a form detailing the claim and signing it;

(2) the claim plus receipts being passed to a head of department or similar counter-signatory for checking and countersigning and the identification of the budget to be debited;

(2) the claim plus receipts being passed to the finance office for checking, debiting against the account and authorisation of payment.

One way of digitising this process would be to have the forms filled in digitally and digitally signed by the claimant before passing on to the Head of Department (HoD) by e-mail for checking of signature and details, countersigning and passing on to the finance office by e-mail. This scenario raises a number of problems:

- What about the receipts?
- Everyone involved needs PK technologies (and needs to know how to use them)

There does not seem to be a cost saving in this simple digitisation of existing process.

What is really going on here?

(1) This is reasonably self explanatory except the purpose of the signature should be clarified. The signature indicates that the claimant asserts to have incurred the costs detailed and has followed the rules. The receipts are attached to provide weight to that assertion and in some institutions claims without the receipts are not accepted (the assertion is therefore not trusted anyway!).

(2) The HoD checks a number of things (implicitly or explicitly) and adds a signature to assert that they have been carried out.

(3) The claim comes from the person it purports to come from (in theory the claimants signature helps in this but in practice it probably does not):
   — the claim should have been incurred in the first place;
   — the claim looks reasonable;
   — the receipts back up the claim.

(4) The finance office repeats much of what the HoD has done (perhaps with more care, but with less personal knowledge of the circumstances) and probably checks the HoDs signature to guard against attempts to circumvent stages 1 and 2.

Checking should be carried out in one place and the person best placed to do this is the HoD. In that case the receipts need go no further than the HoD. The form could be filled in digitally, print-

ed out and signed and passed to the HoD for checking. The HoD having checked, simply needs to authorise payment from the appropriate budget and file the claim form and receipts for the period necessary to comply with audit and Inland Revenue requirements. The compliance function currently carried out by the finance office in this scenario could be replaced by sample audit.

Whilst some details of the above 'solution' might need to be improved, it illustrates the approach of re-examination of the process to facilitate digital working rather than attempting to 'bolt on' a cumbersome PK technology 'solution'.

The UK Inland Revenue have already demonstrated, in their online tax submission procedures, that a signature on a document is not necessarily required for an assertion that a form has been filled in correctly (according to a prescribed set of rules) to be made. The fact that an individual has used a specific userid issued by the Inland Revenue is considered sufficient.

4. Public understanding

An important paradigm shift occurs when moving from the paper world to the digital world, information (and much process) moves from being directly experienced by the person interacting with it, to being experienced only through the inter-mediation of computers and computer software. This raises issues of understanding and responsibility for actions supposedly taken by an individual (my computer does a lot of things that I don't explicitly authorise it to do, including sending information about me and my machine to web sites!). Several studies have shown that 'intelligent' people can experience very real difficulties using encryption and digital signature technologies, which seriously limit their usefulness and legislators who ignore this do so at the peril of society in general. To what extent can I be held personlly responsible for actions taken by software, which I don't understand, on my machine? I raise this issue as an important point for discussion.

4.1. Legal acceptance

Work has been underway in the European Union and in the United Kingdom to clarify the legal status of digital signatures. Both the European Union and the UK legislation take a technologically neutral approach and refer to 'electronic' rather than digital signatures. Their provisions, however, seem to be based on the assumption that the signatures are generated by public key cryptography.

In December 1999, the EU Council adopted the EU directive on electronic signatures (EU 1999/93/EC) which establishes a harmonised legal framework for electronic signatures throughout the EU, in particular addressing criteria for their legal recognition. The directive should have been adopted into the national law of each EU Member States by July 2001.

In the United Kingdom the Electronic Communications Act (2000) came into force on 25 July 2000 making electronic signatures legally admissible as evidence in court as to the authentication of parties to and integrity of an electronic communication. It did not, however, confer equivalent status to a traditional hand-written signature on any particular types of electronic signature. Exactly how an electronic signature will be treated in court will be for the court to decide as the explanatory notes that accompany the act make clear:

It will be for the court to decide in a particular case whether an electronic signature has been correctly used and what weight it should be given (e.g. in relation to the authentication or integrity of a message) against other evidence.

4.2. Usability

People have a fairly good grasp of the mechanism and implications of written signatures and are able therefore to use them accurately and reliably. Can we be sure that the same applies to the mechanism and implications of digital signatures? Two usability studies that have recently been reported are of relevance to our discussion:-

- 'Usability of security: a case study' (Whitten, A. and Tygar — 1998) — reports on an experiment in which a group of people were given a task to do which required them to use PGP. The

success of participants in this study was not very high because they had difficulty understanding exactly what they were doing with the software and the relationship between public and private keys.

- 'Experiences of using a public key infrastructure for the preparation of examination papers' (Chadwick, D. W., Tassabehji, R. and Young, A. — 2000) — reports on an attempt to convert an existing examination paper preparation process involving: question setters, administrators, examination boards and external examiners into a digital process using e-mail and PKI, without changing the process. This study was beset by many problems, including participant motivation, hardware, software and the fact that the process really needed changing to get it to work in a digital form.

4.3. Non-repudiation

Non-repudiation is the ability to provide evidence to a third party that an action has occurred and evidence which refutes denial. It can be specifically applied to non-repudiation of approval, sending, origin, submission, transport, receipt, knowledge or delivery services. Digital signatures can be used to provide non-repudiation of origin. Since only the originator should be able to access the private key, the signature is unforgeable evidence that the originator generated the message.

McCullagh and Caelli (2000) have argued that:

> *the technical meaning of the term non-repudiation either shifts the onus of proof from the recipient to the alleged signatory or entirely denies the signatory the right to repudiate a digital signature.*

This would be the reverse of the position with traditional hand-written signatures where the recipient or 'relying party' must prove that the signature belongs to the person denying it.

In practice the signatory would be able to repudiate a digital signature and the court would need to decide as Adams, Carlisle and Lloyd conclude:

It is incorrect to say that a public key infrastructure (PKI) can provide non-repudiation; this can never be done if humans must be involved in the dispute resolution phase. Rather, a PKI can be said to provide or implement a service that supports non-repudiation.

This broadly applies to the use of digital signatures and not just to non-repudiation.

4.4. Conclusion

Without a clear understanding of what is going on by all parties to a transaction involving digital signatures, it would be unsafe to rely on digital signatures as evidence in disputes, particularly those involving individuals rather than organisations. The Latin maxim *caveat emptor* would seem to be appropriate here, in the sense that relying parties should take sufficient precautions to ensure that the there are adequate guarantees for a transaction as they have the most to lose. It may be that signing should be carried out under the strict control of a public notary, with witnesses to the process, as is the case with important written signatures. The technology could support this way of working if the process required it.

5. Finally

Now that the technology is easily available we need to discuss the much more difficult problems of the technology in use. This workshop represents a good opportunity to do that.

References

Adams, C. and Lloyd, S., 'Understanding public-key infrastructures: non-repudiation': http://www.knowcisco.com/contents/15870166X/ch5s04.shtml

Chadwick, D. W., Tassabehji, R. and Young, A., (2000), 'Experiences of using a public key infrastructure for the preparation of examination papers', Computers and Education 35, 1. Available via: http://www.sciencedirect.com/

Electronic Communications Act 2000: http://www.hmso.gov.uk/acts/acts2000/20000007.htm

EU Directive 1999/93/EC, (13 December 1999) on a Community framework for electronic signatures: http://europa.eu.int/eur-lex/en/oj/index.html

Introduction to the Internet service for self-assessment (SA): http://www.inlandrevenue.gov.uk/efiling/sa_efiling/sa_intro.htm

McCullagh, A. and Caelli, W., (2000) 'Non-repudiation in the digital environment', *First Monday* Volume 5, issue 8: http://firstmonday.org/issues/issue5_8/mccullagh/index.html

Whitten, A. and Tygar, J. D., (1998) 'Usability of security: a case study': http://reports-archive.adm.cs.cmu.edu/anon/1998/abstracts/98-155.html

Las firmas digitales no son una solución, sino un simple eslabón de la cadena

James Currall

La tecnología de las firmas digitales está evolucionando rápidamente. Existen numerosas organizaciones que actúan como autoridades de certificación de claves públicas, y un amplio abanico de *software* propone tecnología que permite utilizar claves públicas y privadas. Esta situación permite centrarse en los procesos que recurren a estas tecnologías y en los interrogantes que suscita esta utilización: acogida por parte del público, admisibilidad jurídica, viabilidad, certificación de las claves públicas y longevidad de la propia firma digital.

¿Poseen realmente las firmas digitales características similares a las firmas manuscritas tradicionales? Yo sugiero que las firmas digitales tienen muchos más puntos comunes con los sellos utilizados en el pasado que con las firmas manuscritas. Por tanto, debemos considerarlas desde un punto de vista distinto a las firmas manuscritas, y también debemos revisar las modalidades de incorporación de estas tecnologías en los procesos.

Las firmas digitales en sí no solucionan los problemas planteados por el intercambio y la manipulación de documentos digitales (y de otros materiales). Su función es permitir concebir procesos en los que la autenticidad y los documentos digitales sean elementos primordiales. Muy a menudo, la parte esencial del proceso no puede digitalizarse ni automatizarse. Ejemplos procedentes de la Universidad de Glasgow, como nuestro sistema de aprobación de los resultados de exámenes a distancia, lo demostrarán. En una serie de casos, la autenticidad digital puede establecerse según modalidades que no implican necesariamente la firma digital (ni ninguna otra forma de infraestructura de clave pública). Según mi experiencia, un proyecto que utilice firmas digitales presenta riesgos de fallos, no tanto porque la tecnología de la firma digital no esté a la altura del reto, sino porque el conjunto del proceso no ha sido concebido ni aplicado correctamente en sus distintos aspectos (manual, automático, digital y en papel).

Un gran cambio de modelo que se produce cuando se pasa del papel al formato digital consiste en que la información (y la parte fundamental del proceso) pasan de ser experimentados directamente por la persona que interactúa con ella, a ser experimentados sólo por medio de los ordenadores y el *software*. Este fenómeno plantea la cuestión de la comprensión y la responsabilidad de las medidas que el individuo debe adoptar («¡mi ordenador hace muchas cosas que yo no he autorizado explícitamente!»). Varios estudios ponen de manifiesto que personas «inteligentes» pueden tener grandes dificultades al utilizar técnicas de codificación y firma digital, lo que limita gravemente su utilidad, y que los legisladores que ignoran este fenómeno lo hacen con riesgo para la sociedad en general. Plantearé esta cuestión como un punto importante de debate.

Ahora que la tecnología es fácilmente accesible, conviene abordar problemas mucho más delicados de su uso. Este taller constituye una buena oportunidad para hacerlo.

Digitale Signaturen: keine Lösung, sondern lediglich ein Glied in der Prozesskette

James Currall

Die Technologie der digitalen Signatur entwickelt sich in raschem Tempo. Viele Organisationen fungieren als Zertifizierungsstellen für öffentliche Schlüssel, und die Technologie für die Verwendung öffentlicher und privater Schlüssel ist im Rahmen vielfältiger Pakete verfügbar. Somit können wir uns nunmehr den Prozessen, bei denen sich diese Technologien wirkungsvoll einsetzen lassen, sowie den Fragen zuwenden, die ihr Einsatz aufwirft, und zwar im Zusammenhang mit Öffentlichkeitswirksamkeit, rechtlicher Zulässigkeit, Praktikabilität, Verfügbarkeit öffentlicher Schlüssel und Langlebigkeit der digitalen Signatur selbst.

Besitzen digitale Signaturen ähnliche Eigenschaften wie herkömmliche eigenhändige Unterschriften? Meiner Ansicht nach haben digitale Signaturen weitaus mehr mit Siegeln gemeinsam, wie sie in der Vergangenheit verwandt wurden, als mit handschriftlichen Unterschriften. Daraus folgt, dass wir digitale Signaturen mit anderen Augen betrachten müssen als ihre handschriftlichen Verwandten. Ebenso gilt es, die Einbettung dieser Technologien in andere Prozesse noch einmal zu überprüfen.

Digitale Signaturen für sich genommen lösen keine Probleme beim Austausch und beim Umgang mit digitalen Dokumenten (und anderen Materialien). Sie ermöglichen lediglich die Konzipierung von Prozessen, wenn Authentizität und digitale Dokumente entscheidende Elemente darstellen. In vielen Fällen ist ein Großteil des Prozesses unter Umständen gar nicht digital oder automatisch. Beispiele der Universität Glasgow wie unsere Fernbestätigung von Untersuchungsergebnissen sind Beleg dafür. Oft lässt sich die digitale Authentizität mit Mitteln feststellen, die keiner digitalen Signaturen (oder einer Public-Key-Infrastruktur) bedürfen. Meiner Erfahrung nach misslingt ein Projekt, bei dem digitale Signaturen verwendet werden, weitaus eher deshalb, weil der gesamte Prozess (manuelle, automatische, digitale und papiergebundene Aspekte) nicht richtig konzipiert und umgesetzt worden ist, als aufgrund der Tatsache, dass die Technologie der digitalen Signatur den Ansprüchen nicht gerecht wurde.

Ein wichtiger Paradigmenwechsel beim Übergang von der Papierwelt zur digitalen Welt besteht darin, dass sich Informationen (und ein Großteil der Prozesse) verlagern, und zwar vom direkten Erleben durch denjenigen, der mit den Informationen interagiert, hin zum Erleben ausschließlich durch die Vermittlung von Computern und Computersoftware. Daraus ergibt sich die Frage, inwieweit der Einzelne eigentlich von ihm auszuführende Handlungen versteht und dafür verantwortlich ist („Mein Computer macht vieles, was ich ihm gar nicht ausdrücklich erlaubt habe!“). Wie verschiedene Studien zeigen, kann der Umgang mit Technologien der Verschlüsselung und der digitalen Signatur für „intelligente“ Menschen durchaus sehr schwierig sein, was den Nutzen dieser Technologien erheblich einschränkt. Fehlt dem Gesetzgeber diese Einsicht, schadet dies der Gesellschaft insgesamt. Dieses Thema werde ich als wichtigen Punkt zur Diskussion stellen.

Jetzt, da die entsprechende Technologie überall verfügbar ist, müssen wir über die weitaus schwierigeren Probleme der in Gebrauch befindlichen Technologie sprechen. Dieser Workshop bietet dafür eine gute Gelegenheit.

Les signatures numériques ne sont pas une solution, mais un simple maillon de la chaîne

James Currall

La technologie des signatures numériques est en évolution rapide. Il existe de nombreuses organisations qui interviennent comme autorités de certification de clés publiques, et un large éventail de logiciels proposent la technologie permettant d'utiliser des clés publiques et privées. Cette situation permet d'attirer l'attention sur les processus qui font appel à ces technologies et sur les questions que cette utilisation suscite: accueil réservé par le public, admissibilité juridique, praticabilité, certification des clés publiques et longévité de la signature numérique en soi.

Les signatures numériques possèdent-elles vraiment des caractéristiques semblables aux signatures manuscrites traditionnelles? J'avancerais l'idée que les signatures numériques possèdent davantage de points communs avec les sceaux et cachets utilisés par le passé qu'avec les signatures manuscrites. Il s'ensuit qu'il nous faut les considérer d'un point de vue autre que les signatures manuscrites et qu'il nous faut également revoir les modalités d'incorporation de ces technologies dans les processus.

Les signatures numériques en elles-mêmes ne règlent pas les problèmes posés par l'échange et la manipulation de documents numériques (et d'autres matériels). Leur rôle est de permettre de concevoir des processus dans lesquels l'authenticité et les documents numériques sont des éléments primordiaux. Très souvent, l'essentiel du processus peut ne pas être numérisé ni automatisé. Des exemples empruntés à l'université de Glasgow, comme notre système d'agrément des résultats d'examen à distance, en feront la démonstration. Dans un certain nombre de cas, l'authenticité numérique peut être établie selon des modalités qui n'impliquent pas nécessairement de signature numérique (ni une quelconque forme d'infrastructure de clé publique). D'après mon expérience, un projet utilisant des signatures numériques présente des risques de défaillance, moins parce que la technologie de la signature numérique n'est pas à la hauteur du défi que parce que l'ensemble du processus n'a pas été conçu et mis en œuvre correctement dans ses divers aspects (manuel, automatique, numérique et «papier»).

Un changement majeur de modèle se produit lorsque l'on passe du papier au numérique, à savoir que l'information et l'essentiel du processus ne sont plus vécus directement par la personne en interaction, mais plutôt par l'intermédiaire des ordinateurs et des logiciels. Ce glissement pose la question de la compréhension et de la responsabilité des mesures que l'individu est supposé prendre (mon ordinateur accomplit un grand nombre de choses que je ne l'autorise pas explicitement à faire!). Plusieurs études montrent que des personnes «intelligentes» peuvent connaître des difficultés réelles à utiliser les techniques de chiffrement et de signature numérique, ce qui limite sérieusement leur utilité, et que les législateurs qui ignorent ce phénomène le font au péril de la société en général. Je poserai ce problème en tant qu'argument essentiel de discussion.

Maintenant que la technologie est facilement accessible, il convient de discuter des problèmes nettement plus délicats de cette technologie dans son application. Cet atelier constitue une bonne opportunité pour le faire.

Jean-François Blanchette

Jean-François Blanchette received a BSc in computer science in 1995, a MSc in computer science with a specialisation in cryptography in 1997, both from the Université de Montréal, and a Ph.D. in social studies of science and technology from Rensselaer Polytechnic Institute in 2002. In 2001, he served on a task force mandated by the French Ministry of Justice to make proposals regarding the adaptation of authentic acts to electronic signatures. He is currently an associate researcher at the Research Centre on Public Law, at the University of Montreal, and serves on the Quebec Chamber of Notaries' advisory committee on electronic authentic acts.

Civil law authenticity meets digital signatures — a report from the trenches

Jean-François Blanchette

1. Introduction

The French Civil Code lays out precise rules for admissibility and weighing of evidence in litigation. The base rule is that French law requires that proof be pre-constituted, that is, parties must think in advance about proving their rights and/or obligations, by manufacturing a written document at the same time that the contract is concluded. Parties are free to conclude contracts by whatever means they choose — handshake, verbal agreement, signed documents, or telepathy for that matter — but they are required to produce signed documents if they want to prove those obligations in the context of litigation. Article 1341 of the Code articulates this fundamental requirement:

A written document must be established, either by a notary or as a private act, for any dealing over a sum or a value established by decree (before 1980, 50 FRF, currently 5 000 FRF), and no proof by testimony is admissible against or beyond the content of this document, nor with regard to what could have allegedly been said before, during, or since the writing of the document.

In addition, French evidence law distinguishes between two kinds of written evidence: private acts comprise the documents which parties have themselves drafted to provide evidence of contractual obligations, with few form requirements beyond their signature in the hand of the parties to the act. While such act always prove against testimony (as stated in Article 1341 above), they can only serve as evidence if they are recognised by the parties to which they are opposed. In contrast, authentic acts are those documents drafted and received by public officers, officers of the State with the professional privilege of acting as trusted witnesses of private contractual agreements. Because they are produced and cared for by an officer of the State, such documents occupy the highest rung in the French evidence hierarchy, and their date, origin, and truthfulness is deemed valid in and of themselves. The written documents at the base of the French *Etat de droit* are authentic: real estate contracts; records of civil status; and court decisions.

Calls for reform of such a system, founded on the primacy of paper-based writing, have been heard repeatedly since the end of the 1970s, when machine-based processing of information became the norm within the insurance and banking industry. An initial reform in 1980 granted admissibility to non paper-based writing (especially microfilms), but did not attempt to define them. Throughout the 1980s and 1990s, the French legal community resounded with repeated calls for a more thorough re-examination of evidence law than had been afforded by the 1980 reform. Questions over the admissibility as written evidence of faxes, photocopies, and electronic messages led to conflicting jurisprudence which threatened the rationality and coherence of the French evidence regime. Both doctrinal analysis and jurisprudence were unanimous in arguing that the evidence system outlined in the Civil Code, founded on the primacy of paper-based writing, could not be interpreted so as to grant admissibility to writing which had been transmitted or stored in electronic form. In the late 1990s, the idea that the law could provide a useful mechanism for increasing consumer trust in electronic transactions and thus, e-commerce and e-government, has provided the impetus for the enactment of reform, both at the European and national level.

2. The electronic signature bill

The adaptation of the French evidence law system to electronic writing and electronic signatures was enacted on 13 March 2000, with the bill 'regarding the adaptation of French evidence law to information technologies and electronic signatures.' ([1]) The bill itself incorporates elements from three distinct sources: (1) a working group mandated in 1996 by the Ministry of Justice to reflect on the problem of 'electronic proof.' ([2]); (2) a commission of the Conseil d'état mandated in 1997 by the Prime Minister to reflect globally on the adaptation of the French legal and regulatory envi-

([1]) 'Loi no. 2000-230 du 13 mars 2000 portant adaptation du droit de la preuve aux technologies de l'information et relative à la signature électronique', *Jounnal Officiel,* 14 March 2000, 3968.

([2]) For a summary of this working group's conclusions, see Catala, Gautier, Huet et al. (1999) 'L'introduction de la preuve électronique dans le Code civil.' *La Semaine Juridique Édition Générale* 47(24 November), pp. 2069–2076.

ronment to the rise of information networks ([3]); (3) the parliamentary process (1999–2000) which introduced important amendments to the proposed draft bill.

The first working group's most important conceptual contribution was to suggest that a system founded on the primacy of written proof could be smoothly adapted to the new context of electronic networks, by simply recognising that writing is independent of media:

> *Article 1316 — Documentary, or written, evidence, results from a series of letters, characters, numbers, or any other signs or symbols endowed with an intelligible signification, whatever their media or the means of their transmission.*

Given this, electronic writing is not fundamentally different from paper-based writing and should enjoy the same evidential weight, when its *origin* and *integrity* can be properly ascertained:

> *Article 1316-1 — Electronic writing is admitted as evidence on the same terms as writing on paper media, provided that the person from which it emanates is duly identified and that it is drawn up and preserved under conditions sufficient to guarantee its integrity.*

When it came to signatures however, the bill adopted a definition provided by the *Conseil d'état,* a definition meant to serve as a place holder for the European directive's notion of 'advanced electronic signatures':

> *Article 1322-2 — The signature which is necessary to perfect a private act identifies the one who affixes it and manifests his consent to the obligations which flow from this act.*
>
> *When it is electronic, it consists in the use of a trustworthy identification mechanism guaranteeing the link with the act to which the signature is attached. The trustworthiness of this mechanism is presumed, until proof of the contrary, when the signature is created, the identity of the signatory ensured, and the integrity of the act guaranteed, under conditions established by a decree from the Conseil d'état.*

This application decree, published on March 31 2001, states that the trustworthiness of the electronic signature mechanism can be presumed when it ensures that the electronic signature '(a) is unique to the signatory; (b) is created using means which the signatory can keep under its sole control; (c) guarantee, with regard to the act to which it is attached, a link such that any subsequent modification to the act is detectable.' That is, such an electronic signature provides for identification, non-repudiation (whatever this means), and integrity.

These definitions are meant to define a proper framework for electronic private acts, but leave electronic authentic acts out of their scope ([4]). The French notarial profession pushed for an amendment to be included to the draft bill, the addition of a second paragraph to Article 1317 of the Civil Code, which would now read:

> *Article 1317 — The authentic act is that one which has been received by public officers with the right to operate in the place where the act has been drafted, and with the required solemnities. It can be established on electronic media if it is drafted and preserved under conditions established by decree from the Conseil d'état.*

The Senate Law Commission justified the move on the basis of expediency: while electronic authentic acts would clearly not become a reality in the close future, to seize the current occasion to define them in principle would save the Government another iteration of a somehow tedious legislative process. As the *rapporteur* explained, 'we must not close the door to what will one day become electronic authenticity. The recourse to an administrative decree will make it possible to use electronic authentic acts as soon as the material conditions are right.' That is, the proposed amendment made the principle of dematerialised authentic acts possible, while leaving the precise provisions for their proper production and care to forthcoming administrative decrees, to be updated as technological evolution required. Still, Elizabeth Guigou, Minister of Justice, warned that the 'dematerialisation' of authentic acts posed much greater challenges than that of private acts, in particular, with regard to their preservation:

> *Contrary to private acts, which must only be preserved for a maximal duration of 30 years, authentic acts must be preserved for an unlimited duration. Technological solutions must thus guarantee*

([3]) See Conseil d'état (1998) *Internet et les réseaux numériques*, La Documentation Française, Paris.

([4]) In fact, art. 1 of the directive on electronic signatures specifically excludes such acts from its scope: 'It does not cover aspects related to the conclusion and validity of contracts or other legal obligations where there are requirements as regards form prescribed by national or Community law nor does it affect rules and limits,contained in national or Community law,governing the use of documents.'

the long-term preservation of the acts. But current electronic technologies can only guarantee preservation of information for a limited period, due to their rapid obsolescence. Of course, it is possible to move information from one media to another following technical evolution, but the recuperation of information will have to be secure. I say it very clearly, the technical conditions for the dematerialisation of authentic acts are not there yet. In this regard, I entirely subscribe to the opinion of your rapporteur, who insisted, during the debates of the Commission, that this amendment not modify the base rules which govern the drawing up of authentic acts, notably, the personal presence of the signatory in the study of the public officer entrusted with the task of obtaining her consent and to confer authenticity to the act. Nevertheless, I subscribe to the idea of inscribing the principle of dematerialisation, so as not to later modify the Civil Code, while pushing back to a later time the material conditions for its realisation. I thus note that this solution is fully in accord with our common objectives. ([5])

([5]) Guigou, Elisabeth (2000): Adaptation du droit de la preuve aux technologies de l'information et relatif à la signature électronique — Intervention of Elizabeth Guigou, Minister of Justice and Guardian of the Seals, Senate, 8 February 2000.

The resulting legal framework for electronic evidence is thus now characterised by three main parameters:

- electronic writing (i.e., private acts) is admissible provided the conditions under which it has been created and preserved guarantee its integrity and its imputability. Once admitted as evidence, it carries the same weight as writing on paper;
- any electronic signature mechanism which identifies the signatory, manifests her consent, and guarantees a link with the document is acceptable to sign a private act. However, only 'advanced electronic signatures' will be automatically granted a presumption of trustworthiness;
- authentic acts (whose definition remains unchanged) are admissible in electronic form, when they have been drafted and archived under conditions yet to be defined.

3. The working group on electronic authentic acts

A mere two weeks after the law was enacted, a working group was formally convened by the Ministry of Justice on 31 March 2000, vested with a double mission: critical reflection on the concept of electronic authenticity and concrete propositions as a basis for drafting of the application decree. Its fundamental objective would be to 'research the conditions for a new electronic formalism which could be substituted to the actual requirements associated with the paper media.' In addition to the particular investigations which would have to be conducted for each category of authentic acts (decisions, records of civil status, notarised acts, etc.), the group was asked to focus particularly on the four following questions:

- How can the deep guarantees offered by authenticity (counsel of the parties, control of the expression of their consent, etc.) be preserved in the case of dematerialised acts?
- Under which conditions and following which forms can the electronic signature of the public officer and of the parties be affixed to authentic acts?
- How can the unlimited archiving and preservation of dematerialised authentic acts be insured?
- Under which conditions can 'copies' of dematerialised authentic acts be produced? What will be the evidential value of such copies?

The Ministry of Justice recognised that such questions were exceedingly complex and that their careful consideration would necessarily prove a lengthy process. Nevertheless, given that several of the legal professions (i.e., the notaries) had already begun to explore public-key infrastructures and certification services, and given the high symbolic charge which suffused the question of electronic authenticity, the Ministry suggested/requested that a preliminary report be submitted within six months — an incredibly short delay given the complexity and the elaborate character of the institutions in question. By and large, the working group was expected to simply adopt the definitions already provided by the first application decree, perhaps contributing to further defining the conditions for the deployment of such secure electronic signatures within the legal professions. The Civil Code defined 'signatures' and 'electronic signatures'; the first application decree, 'secure (i.e., advanced) electronic signatures.' The second application decree would contextualise these definitions for the case of electronic authentic acts.

The legal professions entrusted with the production and care of authentic acts (e.g., notaries) also brought to the table infrastructural commitments. The *Conseil supérieur du notariat* (CSN) has invested heavily in a public-key infrastructure which, it claimed, fully met the framework provided by the Civil Code and the first application decree. As the President of the CSN ensured in a professional newsletter, the March 2000 law:

> *… introduces in the virtual world created by new technologies comparable dispositions to those which we have known for a long time in the traditional universe of contract law. Moreover, the explicit reference to authentic acts, desired by a unanimous parliament, strengthen the role of the notary, promoter of legal security, by widening this eminent mission to computerised transactions. One must see here a kind of recognition of the efforts of the profession, notably through its REAL network, to rapidly adapt to the information technology revolution. The door of the future is open, let's begin walking on the new paths with confidence.* ([6])

While the representatives of the profession repeatedly pointed to the REAL network as an impressive demonstration of the readiness of the profession for the Information Age, it is still unclear which practical needs does the network answer, beyond showcasing the profession's technological prowess? What would a rank-and-file notary gain by undergoing the laborious and costly process of subscribing to REAL, acquiring the signature card, and installing all necessary hardware and software? So far, the board of the *Conseil supérieur du notariat* has only offered vague promises: 'Electronic messaging and related services will give us gains in productivity. These exchanges will be fully secure, thanks to our partner and system developer, 'La Compagnie des signaux.' ([7]) The REAL network had not been developed as a response to a practical problem, as no market for electronic transactions involving civil law authenticity has yet emerged.

In addition, the most fundamental question which faced the notarial profession with regard to the 'dematerialisation' of authentic acts hinged on the requirement for the parties to the act to physically appear in front of the notary, so that he can verify their free consent and their full understanding of the obligations incurred by them within the contract. As Pierre Catala explains:

> *The real difficulties with the application of new technologies to authentic acts are... not located either before the contract or after the contract. Even though solemnities are prescribed in both of these phases, they are formalities extrinsic to the production of the act. They are satellites of authenticity ... The heart of the problem lies in the writing of the authentic title proper. The title draws its entire strength from the simultaneous presence of the parties and of the public officer (or of a habilitated clerk), so that the solemnities required by texts can be simultaneously accomplished. Authenticity cannot be accomplished without the physical presence of the contracting parties before the special witness habilitated to receive the act. Holding this as essential, it is around it which we must invent the possible adaptations of current formalism to future technologies.* ([8])

The physical presence of the notary and of the parties seems an incontrovertible element of the notarial ritual. Yet, the appeal of a 'dematerialised' authentic act hinges precisely on the possibility of concluding contracts at a distance, without the parties being physically present — e.g., concluding a real estate transaction without requiring all parties to physically meet. How can civil law authenticity be compatible with such a setting, given that its essence lies precisely on the physical appearance of parties in front of the notary? Early in the debates, representatives of the profession formulated what would become the official position of the profession: however electronic authentic acts would come to be defined, they would involve the physical presence of the notary, as this presence, this trusted witnessing of consent, lay at the very heart of the profession, in fact, was the very substance of authenticity. As one eminent notary stated:

> *the physical presence of the notary is essential to the authentic act; thus, the consent of the parties cannot be witnessed at a distance. If this principle must be modified to adapt the authentic act to the electronic medium, such a modification must then also apply to the paper media, since the concept of authenticity cannot be different according to the underlying media of the instrumentum.*

The biggest problem which this argument faced was that the notarial profession had already sought to, in fact, succeeded in, breaching the principle of physical presence. In the early 1970s, it had successfully lobbied the Government so that the notarial contractual scene could be divided into two discreet moments, performed by two distinct authorised parties. The witnessing of the consent of the parties to the act could be entrusted to a specially authorised employee of the

([6]) Decorps, Jean-Paul (2000), 'Porte ouverte sur l'avenir', *La lettre aux Notaires* (62), 1.
([7]) Gard, Olivier (1999), 'Le notaire, la carte et le réseau, *Notaires — Vie Professionelle* (216), pp. 75–76.
([8]) Catala, Pierre (2000), 'Le formalisme et les nouvelles technologies,' *Desfrenois* 15/16, 897-910.

notary *(clerc habilité)* while the final signature on the act would remain the notary's privilege. The rationale behind such a move was simple: efficiency. Because the actual performance of the signing ritual as required by law is necessarily a lengthy one, the profession argued that a specially authorised clerk could just as well read the act to the parties, witness their consent and their signature, with the act to be formally signed by the notary at a time of her convenience. The move profoundly shocked one of France's most eminent scholars of notarial law, Jacques Flour, who vehemently critiqued the 1971 decree in a famous article, 'On a new notion of authenticity':

> *… until 26 November 1971, the requirement of the presence of the notary has never been considered as the vain manifestation of some mystical kind of formalism. This presence did not simply belong to the 'required solemnities' mentioned by Article 1317 without detailing them because it was sufficient to refer to the detailed prescriptions of the* loi de ventôse. *These solemnities are the external and contingent rites: the required presence of supplementary witnesses for certain types of acts and the qualities required of them; the purely material requirements for the drafting of the act... The presence of the notary is of an entirely different order.... (The authentic act)'s evidential force is explained by the fact that one can — and one must — trust the affirmations of a public officer. The notary, as is classically defined, is a 'privileged witness'. If he does not witness the affixing of the signatures, he is not a witness; the act can thus not be an authentic one.* ([9])

Writing *(l'écrit)* makes possible the virtualisation of the contractual scene, and the trustworthiness of authentic writing flows directly from that of its manufacturer, the public officer. It is this central role of writing which Flour points to when he notes that 'when it comes to solemn acts, there is no distinction between *instrumentum* and *negotium*. That is, the manufacturing of the *instrumentum* in accordance with the rules dictating the appropriate forms, is itself the *negotium*, that is, a performance with the legal consequence of binding the parties to the obligations contained in the contract. And, as Flour sternly remarked, the party responsible for the proper manufacturing of the *instrumentum* is — must be — the notary, under penalty of authentic acts losing their special evidential value ([10]).

For all of Flour's criticism, the principle of *clercs habilités* remained. In retrospect, the breach it effected would come to seem innocuous in comparison with the implication of the March 2000 electronic signature bill, which proposed to tamper with the very core of notarial authenticity, the physical presence of the notary. For some, such tampering is inevitable, and the profession will just have to bite the bullet if it is to survive globalisation — especially in its concrete manifestation as European integration. The European Affairs Commission of the Union of Latin Law Notaries has stated that 'it will fall to the notarial profession to decide ... if it must use (information technologies) as a simple means of communication or if it can consider using them for reception — and not only for production — of authentic acts.' That is, either the profession will confine the role of information technologies to that of a sophisticated tool for the production and management of documents, or it will use them as a telecommunications tool whereby some configuration of distributed parties and notaries can perform a new contractual scene, a scene bearing little resemblance to what the profession has known before. The problem is how to design this new scene without losing the very soul of the profession — the trusted witness, authenticity and its special evidential value.

The final draft of the working group's report was submitted to the Ministry of Justice in June 2001 ([11]). By then, the blaze that had been the new economy was already smoldering, and the impetus for rapid and radical adaptation of French evidence law to information technologies had lost much of its symbolic charge. The report could thus suggest that each profession (notaries, officers of civil status, court clerks, etc.) be left to organise dematerialisation of authentic acts at the pace which best fitted its needs and capabilities. The report has then made a series of recommendations with regard to how the application decree could deal with the 'dematerialisation' of bodies, documents, and signatures.

Physical presence — The report avoids any metaphysical confrontation over the meaning of physical presence, simply stating that 'whatever the new possibilities which the electronic media offers with regard to witnessing consent at a distance, to modify the actual principle of the physical presence of the public officer would imply a questioning of the principle of authenticity which the legislator has not sought.' It reiterates that the issue cannot be sidestepped as 'the physical presence of the public officer cannot be reduced to a simple requirement for concluding contracts. It is of

([9]) Flour, Jacques (1972), 'Sur une notion nouvelle d'authenticité (Commentaire des articles 11 et 12 du décret no. pp. 71-041 du 26 novembre 1971) (a),' *Defrenois* 92, 977-1017.

([10]) Flour's emphasis on the physical presence of the notary is understandable if one considers that the notarial system arose as a technological response to the need for a more efficient and portable means of proof than the one founded on the availability of the original witnesses of the contractual scene. That is, it is a technology of virtualisation, as Béatrice Fraenkel explains: 'To be able to forego witnesses to confirm a commitment means that the document which states such a commitment is no longer dependent on the physical existence of the contracting parties. It is sufficient that they have been present once, at the moment of its creation ... the legal effect of the document is no longer tied to the length of their lives... the legal instrument is even more independent of the formal circumstances surrounding a ceremony where statements, gestures, and manipulation of objects were all necessary to establish the legal effects of the contract. The scene of the modern contract is entirely different: it is writing which holds the central role, surrounded by the notary and the parties. And it is the affixing of autograph signatures which subsume the gestuality of bodies.' Fraenkel, Béatrice, *La signature, genèse d'un signe.* Paris: Gallimard, 1992, p. 24.

([11]) The report was published in April 2002: de Lamberterie, Isabelle (ed.) *Les actes authentiques électroniques: réflexion juridique prospective.* Paris: La Documentation Française.

the essence of the authentic act.' If the French State wants to make the physical presence of the notary optional, it will have to get rid of authenticity as currently defined. To close any loophole, the report underlines that if the legislator eventually sees fit to modify the principle of physical presence for electronic authentic acts, it will have to do the same with regard to paper acts, since the 13 March law unequivocally affirms the equivalence of paper and electronic written proof.

Document — The report recognises that preserving the legibility and proof value of dematerialised authentic acts will likely lie beyond the ability of individual notaries, town halls or courthouses, and that more centralised forms of archiving will need to emerge. Any such centralisation will however significantly collide with the territorial and functional monopolies of public officers. The working group underlines that it has not been mandated to suggest any modifications to such privileges, and suggests that any such modifications will be more appropriate within the context of a general reorganisation of archival institutions. The report suggests that the preservation of electronic documents can and should be addressed from the beginning of the document production chain and that, 'without wanting to sound too conservative,' the development of hybrid archiving systems, mixing paper and electronic technologies, can provide a productive avenue.

Signature — Even after the Ministry of Justice's request that the working group clarify its position with respect to 'secure electronic signatures,' the report has remained resolutely non-committal. With regard to the signature of the public officer, it has refused to define any specific requirement other than it be affixed by a physically present public officer. It underlines that the problem of ensuring the integrity of authentic acts should be dissociated from the question of signing them, and that further consideration relative to the problems of archiving `secure electronic signatures' will be necessary before they can be deemed applicable to authentic acts.

Conclusion

At the doctrinal level, legal scholars have provided a definition of electronic writing which cleanly separates writing from media and means of transmission, thus enabling the passage from paper-based to electronic-based written proof with the least modifications to the logic of the system. At the level of practices however, media can hardly be considered in such abstract terms: however much writing may be independent from media, it is obviously always manifested through some particular one! In fact, rather than 'dematerialised' (as the expression goes in France), electronic writing has never been so dependent on its material environment — be it physical media, encoding format, software environment, hardware platforms, or network protocols — for both its legibility and its proof value.

Furthermore, the civil law culture of writing is also a culture of the media on which such writing is produced, distributed, read, archived. The procedures which detail the proper manufacturing of official written acts — for notaries, the 1971 decree 'relative to the acts drawn up by notaries,' for officers of civil status, the *Instruction générale relative à l'état civil* — depend not only on the materiality of such documents, but on the institutions entrusted with their production and care, and on the beliefs attached to their cultural, legal, and technical performance. These elements suggest that the design of new forms of written evidence suitable for the context of electronic networks will not emerge solely through technological solutions (such as the 'advanced electronic signatures' promoted through the European directive). Rather, as the case presented here clearly underlines, defining criteria for electronic authenticity is at once a legal, technical, institutional, and cultural problem. By foregrounding the role of digital signature technology at the expense of a more inclusive approach to document lifecycle management, the European directive on electronic signatures (and its transposition into various national laws) has created a highly restrictive framework for the conceptualisation and realisation of electronic archiving systems able to efficiently preserve both the legibility and proof value of documents. Drawing from its rich store of knowledge and experience, the archival community can help evolve such first-generation legal instruments into forms more appropriate to the practical realities of electronic information preservation.

La autenticidad de la ley civil resuelve las firmas digitales: un informe de las trincheras

Jean-François Blanchette

En el sistema francés de presentación de pruebas en los procedimientos civiles, se da prioridad a los documentos públicos, es decir, a los documentos redactados y firmados por un funcionario ministerial o público. Son documentos públicos, entre otros, los contratos firmados ante notario, los documentos de estado civil y las decisiones judiciales. La ley de marzo de 2000 sobre la firma electrónica dispone que los documentos públicos electrónicos son admisibles como prueba, a condición de que se hayan redactado y archivado en las condiciones que se precisarán posteriormente en un decreto de aplicación. En abril de 2000, el Ministerio de Justicia confió a un Grupo de Trabajo la misión de estudiar y presentar propuestas relativas a las condiciones prácticas aplicables a la redacción y el archivo de documentos públicos electrónicos. En esta exposición, presentaré las principales características de la adaptación de la legislación francesa en materia de presentación de pruebas a los documentos y a las firmas electrónicas, las distintas iniciativas de las profesiones jurídicas para informatizar los documentos públicos electrónicos y, finalmente, las conclusiones del Grupo de Trabajo encargado de estudiar las especificaciones de la redacción y el archivo de este tipo de actos, en particular el papel de las firmas digitales y la autenticidad en Derecho civil para asegurar las transacciones «desmaterializadas».

Die Zivilrechts-Authentizität trifft die digitale Signatur – ein „Frontreport“

Jean-François Blanchette

Im französischen Zivilrecht liegt bei Beweismitteln das größte Gewicht auf „authentischen“ Schriftstücken, d. h. auf Dokumenten, die von staatlich zugelassenen Beamten erstellt und unterzeichnet worden sind. Zu authentischen Schriftstücken zählen u. a. notarielle Verträge, Personenstandsurkunden und Gerichtsbeschlüsse. Laut Gesetzentwurf vom März 2000 über elektronische Signaturen sind elektronische authentische Schriftstücke als Beweismittel unter der Voraussetzung zulässig, dass sie unter in einer späteren Durchführungsverordnung klarzustellenden Bedingungen erstellt und archiviert wurden. Im April 2000 erteilte das Justizministerium einer Arbeitsgruppe den Auftrag, die praktischen Bedingungen für die Erstellung und Archivierung elektronischer authentischer Schriftstücke zu erforschen und entsprechende Vorschläge zu unterbreiten. In diesem Beitrag werde ich auf die wichtigsten Bestandteile der Anpassung des französischen Beweisrechts an elektronische Schriftstücke und elektronische Signaturen eingehen, auf verschiedene Bemühungen seitens der Juristen, elektronische authentische Schriftstücke mittels EDV zu erstellen, sowie auf die Schlussfolgerungen der Arbeitsgruppe im Hinblick auf die Anforderungen für die Erstellung und Archivierung elektronischer authentischer Schriftstücke und hierbei vor allem auf die Rolle digitaler Signaturen und die versprochene zivilrechtliche Authentizität für die Sicherung „virtueller“ Transaktionen.

L'authenticité de la loi civile rencontre les signatures digitales: un rapport des tranchées

Jean-François Blanchette

Dans le système français d'administration de la preuve dans les procédures civiles, la priorité est donnée aux actes «authentiques», c'est-à-dire à des documents rédigés et signés par un officier ministériel ou public. Les actes authentiques sont, entre autres, les contrats reçus par notaire, les pièces d'état civil et les décisions de justice. La loi de mars 2000 sur les signatures électroniques dispose que les actes authentiques électroniques sont admissibles à titre de preuve, à condition qu'ils aient été rédigés et archivés dans les conditions qui seront précisées ultérieurement dans un décret d'application. En avril 2000, le ministère de la justice donnait mission à un groupe de travail d'étudier et de faire des propositions concernant les conditions pratiques applicables à la rédaction et à l'archivage d'actes authentiques électroniques. Dans cet exposé, je présenterai les principales caractéristiques de l'adaptation de la législation française sur l'administration de la preuve aux documents et aux signatures électroniques, les diverses initiatives des professions juridiques pour informatiser les actes authentiques électroniques et, enfin, les conclusions du groupe de travail chargé d'étudier les spécifications de la rédaction et de l'archivage de ce type d'actes — en particulier, le rôle des signatures numériques et l'authenticité en droit civil pour sécuriser les transactions «dématérialisées».

Workshop 3

Wednesday 8 May 2002

THE LEGAL ADMISSABILITY OF DIGITAL STORAGE

Chairperson: Claes Gränström (Sweden)
Co-chair: Maria Pia Rinaldi Mariani (Italy)

Jos Dumortier

Dr J. Dumortier is professor in Law and IT, Director of the Interdisciplinary Centre for Law and Information Technology (ICRI).
Jos Dumortier graduated in law at K.U. Leuven (1973). After postgraduate studies in Nancy (Centre Européen Universitaire, 1974) and Mannheim (DAAD, 1975), he became research fellow at the Institute for Labour Law at K.U. Leuven (Prof. R. Blanpain). In 1981 he finished his PhD in Law on a thesis about private international conflicts of law in the area of labour relations. After 1981 and until 1992 he continued to work part-time as a labour lawyer with Simont, Gutt and Simont (presently Stibbe Simont Monahan Duhot) in Brussels. At K.U.Leuven he worked during the same period as the Law Librarian and IT Director of the Faculty of Law. From 1981 until 1983 he studied Information

Electronic signatures and trusted archival services

Jos Dumortier, Sofie Van den Eynde

I. Introduction

Recent legislation on a European level has abolished some of the legal barriers regarding electronic document creation and preservation. The European Directive 1999/93/EC ([1]) deals with electronic signatures. Where the use of electronic documents is legally permitted, so-called 'qualified electronic signatures' must receive a status that is equivalent to the legal status that handwritten signatures normally have in relation to paper documents. Furthermore, the electronic commerce directive ([2]) obliges the Member States to ensure that the legal requirements applicable to the contractual process (the archival phase included!) neither create obstacles for the use of electronic contracts, nor result in such contracts being deprived of legal effectiveness and validity on account of their having been made by electronic means. Legal rules enforcing the use of paper documents in the contractual process must be progressively abrogated. Consequently, a lot of important documents, that are now still mostly paper-based, such as contracts and other legal instruments, will in the future solely exist in their original electronic form. It is needless to say that this evolution could lead to an increasing need for electronic archival services.

The PKI-based technique of the digital signature plays an important role in this new legal framework. The current state of the law in Europe results in the fact that only PKI-based digital signatures can bring forth so-called 'qualified' electronic signatures. The digital signature technique allows authenticating electronic information in a way that the origin of the information, as well as its integrity can be verified. A digital signature is an encrypted hash code that is deduced from and attached to the electronic information that has to be authenticated. If only one bit changes over time, the verifier of the digital signature will notice that the integrity has been affected. The verifier can also be sure that the electronic information originates form the signatory, since he is the only one who knows his own private key.

As a result of this new legal framework, archivists are increasingly challenged to deal with digital signatures as an organic part of electronic documents. However, there exists a lot of resistance in the archival community against the preservation of digital signatures. This is well-illustrated by the report of the InterPARES Authenticity Task Force, entrusted with the task of identifying 'conceptual requirements for assessing and maintaining the authenticity of electronic records ([3]). The Task Force adopted an unequivocal position with regard of the role of digital signature technologies and PKI as a means of ensuring the authenticity of records:

'Digital signature and public key infrastructure (PKI) are examples of technologies that have been developed and implemented as a means of authentication for electronic records that are transmitted across space. Although record-keepers and information technology personnel place their trust in authentication technologies to ensure the authenticity of records, these technologies were never intended to be, and are not currently viable as a means of ensuring the authenticity of electronic records over time.' ([4])

In this paper we will first investigate the need for electronic signatures to be verifiable years later from a legal point of view. We will also focus on several standardisation initiatives that were launched by the European Commission within the framework of the European ICT Standardisation Board regarding long-term validation and long-term archival of digital signatures. This analysis will draw the attention to the fact that digital signatures can in fact be archived with the expertise and the stability epitomised in a trusted archival service. To conclude, this paper will instigate the establishment of a legal framework for this kind of trust service.

II. Electronic signatures, digital signatures and public key encryption

Although digital signatures are known best as a substitute for handwritten signatures with legal value (= electronic signatures), the technique of the digital signature has many other applications.

([1]) Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures, OJ, 19 January 2000, L13/12.
([2]) Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, *OJ*, 17 July 2000, L 178/1.
([3]) Blanchette J. P., 'Dematerialising' written proof: French evidence law, cryptography and the global politics of authenticity, doctoral dissertation submitted to the Department of Science and Technology of the Rensselaer Polytechnic Institute, 2001, p.308, writes: 'The fundamental premise of the InterPARES project is that authenticity is not primarily a function of technology, but rather, of institutions. Archivists have historically been entrusted with the task of providing this function, within either private or public institutions, and they remain the most appropriate, professionally organised, socially recognised, historically legitimate profession to accomplish similar functions in the electronic environment.' The interesting dissertation of J. P. Blanchette is available at http://www.rpi.edu/~blanc/thesis.pdf
([4]) See the draft final report of the InterPARES Authenticity Task Force, http://www.interpares.org/documents/atf_draft_final_report.pdf, p. 8

It can be used in all cases where the origin and the integrity of electronic data have to be guaranteed ([5]). These qualities are very important for documents that are stored in archives. A digital signature added to the (signed) record by the archivist, allows the verifier of the signature to check the identity and the authority of the archivist. That is how the authenticity of a record 'as a record' can be checked in a network environment, the future work area of archivists. The presence of the digital signature of the archivist in the metadata of a record indicates that that record has the status of an archived record. The use of the digital signature technique also creates the opportunity for checking the integrity of electronic records. When used in this manner, the digital signature functions as a 'seal' ([6]). By creating and archiving an encrypted, and thus inaccessible hash code, it can be noticed at all times when the plain text has been tampered with ([7]).

It is true that PKI-based digital signatures have been considered for a while as an essential tool for archival purposes. It has been argued that all electronic data should be secured with digital signatures before entering the archive. The integrity of the electronic data stored in the archive could then, in this view, be controlled by the verification of the archivists' signature.

Scepticism about this theory appeared as soon as it became clear that, when using digital signatures, control of the integrity is only possible if the electronic data remain completely unchanged at the bit-level. This raises a problem when archivists want to migrate electronic data to new formats or software platforms in order to keep them accessible and legible. Some people have immediately concluded that digital signatures are therefore not useful and hence not relevant for archival purposes. If a digital signature secures a document before it enters the archive, the signature should be stripped and translated into metadata.

From a legal point of view however, in order for signed documents to keep their value over time, it is often important that the original electronic signature is still present. Signatures can be needed for non-repudiation purposes in an evidential context, for example. Many European countries require proof that non-commercial transactions are embodied in a signed document ([8]). Recent developments in the context of e-government have also made clear that signed electronic communication with the government must be archived. The government who picks the lowest bidding firm in the context of a public contract conducted by electronic means, will want to be able to prove before court that this firm is bound by its price offer.

The content of the document and the signature are one indivisible artefact in the paper world. A traditional signature has all the characteristics of a classical one-way function: it is easy to process in one direction but very difficult to reverse the process, i.e. the signature is easy to affix but difficult to remove. This is not the case with electronic signatures: an electronically signed document is not different from an electronic document that has not been signed except that it has appended to it another series of bits that can be used to identify the signatory and verify the integrity of the document. Thus, an electronic signature can very easily be stripped from a document for fraudulent purposes without leaving a trace.

Although they have the same functions from a legal viewpoint, traditional signatures and electronic signatures are two very different concepts that need to be treated differently. Never before in the history of written communication has a signatory had to worry about how the signature will be linked to the content of the document that he is signing. When using electronic signatures, this becomes now a very relevant issue. The content of the document and the digital signature should be concatenated and the hash-code of the concatenation should be lodged with an independent entity that will time stamp the hash-code ([9]). The hash code establishes the bond between signature and content. The time stamp must be included in the metadata of the document. The intervening independent entity is designated in this paper by the term 'trusted archival service'.

III. Long-term validation of electronic signatures

The European Commission took the view that the requirements identified by the e-signature directive needed to be supported by detailed standards and open specifications so that products and services supporting electronic signatures can be known to provide legally valid signatures. A mandate was issued to European standardisation bodies, CEN/ISSS and ETSI, to analyse the future needs for standardisation activities. Under the auspices of the European ICT Standardisation Board

Sciences (Infodoc) at the Université Libre de Bruxelles. Between 1984 and 1992 he was part-time lecturer in Information Science at the University of Antwerp and lecturer in Computer Law at Adecho in Brussels. In 1985 he became part-time lecturer in Legal Informatics and in 1993 full-time Professor in Law and IT at K.U. Leuven. In 1990 he was the co-founder of the Interdisciplinary Centre for Law and Information Technology of which he became the first Director. Since 1991 he is active in lecturing, research and consultancy in the area of Law and IT and published several books and articles on this subject. He is a board member of several scientific and professional organisations such as the European Academy (Germany), Beltug (Belgian Telecommunications Users Association) and SGOA (Mediation and Arbitration in IT Conflicts) and he participates in the editorial board of several periodical publications. He is the editor of the International Encyclopaedia of Cyberlaw. From 1995–99 he was part-time staff

([5]) The possible use of digital signatures for the preservation and authentication of records through time has been analysed in the framework of the DAVID-project (which stands for Digital Archiving in Flemish Administrations and Institutions, http://www.antwerpen.be/david). See also: Van den Eynde, S., *The OAIS reference model as starting point in search of the role of public key infrastructure for electronic archives*, Leuven, Interdisciplinary Centre for Law and Information Technology, August 2001, p. 63 (in Dutch only).

([6]) As opposed to the digital signature used as an electronic signature with legal value in the sense of the European e-signature directive.

([7]) The possibilities of digital signature technology must not be overestimated though. To guarantee integrity, we probably must combine this technology with carriers of the 'Write once read many' type.

([8]) Kötz, H., *European contract law: formation, validity and content of contracts, contract and third parties,* Oxford, Clarendon, 1998, p. 78.

([9]) McCullagh, A. et al., 'Signature stripping: a digital dilemma', *Journal of Information, Law and Technology*, 2001/1, http://elj.warwick.ac.uk/jilt/01-1/mccullagh.html

member in the office of the Minister of Justice. ICRI is currently involved in the DAVID project (digital archiving in Flemish institutions and administrations). The DAVID project is executed in association with the Antwerp City Archives and is financed by the Scientific Research Foundation. ICRI was involved in this project to study PKI in an archival context.

Sofie Van den Eynde

Sofie Van den Eynde, born in 1976, obtained her degree in law at the K.U. Leuven University in 1999. She graduated magna cum laude. During her studies she has specialised in tax law and in criminal law. She is currently employed at the Interdisciplinary Centre for Law and Information Technology of the University of Leuven (Belgium), where she is working on a four-year extending project called 'DAVID', which is the abbreviation for 'Digitale Archivering in Vlaamse Instellingen en Diensten' (Digital Archiving in Flemish Institutions and Administrations, http://www.antwerpen.be/david). This project is being executed in association with the Antwerp City Archives. Sofie is responsible for investigating the legal aspects of digital archiving in the scope of this project. She is the author of several publications in this field.

the European Electronic Signature Standardisation Initiative (EESSI) was launched. The first result of this initiative was an expert report about future standardisation requirements. This report affirms that trusted archival services could play an important role in supporting electronic signatures that may need to be used in evidence long after they were created and identifies it as a topic requiring further study since no standards exist yet for the use of such services in support of electronic signatures ([10]). In the mean time, ETSI has published a standard 'electronic signature formats' defining all the elements necessary to prove the validity of a signature long after the normal lifetime of the critical elements of an electronic signature ([11]). This so-called validation chain has to be archived.

Thus, it is not enough that just the electronic signature and the content of the document are present in the archives when a signed document is needed years later. In order to perform validation, the certificate used by the signatory must be obtained, and its validity at the time of signature creation must be proofed. It is possible that the certificate was valid at the time of signature creation, but had expired or had been revoked or suspended some time later. By consequence, the certificate status information must be archived as well ([12]). Signature validation must be performed immediately after, or at least as soon as possible after signature creation time, and not only at archival time, in order to obtain certificate status information that was issued by the CA as closely as possible to the moment of signature creation.

Only the moment of signature creation has an archival value. A signature that has been found to be valid at signature creation time shall continue to be so for the same document months or years later. Evidence must be provided that the document was signed before the certificate became invalid. Thus, the time of signature creation must also be determined and archived. A time stamp can provide for such evidence. A time stamp is a set of computer data, consisting of the hash code of the digital signature and the time of stamping, signed by a time stamping authority (TSA). It proves that the digital signature was formed before the certificate became invalid. Anyone who wants to make sure that he can rely on a signed electronic document for proof years later, must obtain a time stamp before the certificate becomes invalid. The sooner after signature creation time the time stamp is obtained, the better it is for legal certainty.

IV. Long-term archiving of electronic signatures

The availability in the archives of a complete validation chain is a necessary, but not the only condition to make long-term signature validation possible. In the end, all we can preserve in an electronic context are bits. However, it has been clear for a long time that it is very difficult to keep a set of bits indefinitely. With the lapse of time, the set of bits becomes illegible (to the computer and thus to humans) as a result of the technological obsolescence of the application programme and/or of the hardware (e.g. the reader). This paradox causes difficulties for the long-term validation of electronic signatures. Every alteration to the record after signature creation time will make signature validation fail. Reformatting operations that must keep a record legible for the future (such as migration, conversion) are thus detrimental to the durability of the signature.

The problem of the durability of PKI-based digital signatures has been poorly studied so far because of its complexity. Still, some solutions are found in literature, mostly coming from the industrial community. It should be mentioned that we are confronted here with a unique problem from a historical perspective. Although the authentication tools that were used in the past, such as handwritten signatures, seals, stamps, fingerprints etc. are also subject to reformatting (e.g. microfilming) because of the obsolescence of the paper carrier, they never become completely useless after reformatting. There is always at least a copy available that can be compared with other original authentication tools.

The original signatory may resign the reformatted record with his private key. This solution is of course unimaginable from a legal viewpoint. A person cannot resign a document, as it would be a totally new document with the same content however, but signed at a later date. Furthermore, a signatory could refuse to sign, or he could have deceased.

Another option would be to strip the digital signature, or in other words not to archive the digital signature. A trusted third party could certify that the original signatory signed a previous version of

([10]) Nilsson, H., van Eecke, P., Medina, M., Pinkas, D. and Pope, N., *European electronic signature standardisation initiative*, Final Report of the EESSI Expert Team, 20 July 2000, 69, available at: http://www.ict.etsi.fr/eessi/Documents/Final-Report.pdf

([11]) *Electronic signature formats*, ETSI TS 101 733 v.1.3.1 (2002).

([12]) It is the responsibility of each certification authority (CA) to make available in repositories on the Internet all the information needed to validate any signature that was created by means of a certificate issued by that CA. This includes making public at a regular basis information about the time a certificate expired, or was revoked or suspended.

the reformatted record, but which is identical with respect to the content. In this scenario the digital signature is not archived along with the reformatted record as it is cut off as soon as the record is subjected to reformatting. However, this solution puts up legal and practical barriers. A signature is an utmost personal authentication tool, which can never be replaced by a declaration of a trusted third person. Such a record has at best the value of a certified copy. Current legislation in many European countries still requires the submission of an original record though. Furthermore, in a network environment one can never prove the identity of the trusted third party who certified the existence of the signature, without using digital signatures. This is the same problem of a certification authority that certifies the identity of the signatory in a certificate. In order for the public to trust that certification authority, the certificate must be signed by the certification authority with its own private key. A trusted third party that certifies the existence of a signature e.g. in the metadata (the metadata have the same function here as the certificate in a PKI), must create the opportunity for the public (e.g. the participants in the archival network) to believe the metadata. Therefore, one will want to know the origin of the metadata. In the current state of technology, digital signatures are indispensable for this purpose. The problem of long-term authentication of the signatory is thus replaced by the problem of the authentication of the metadata.

The World Wide Web Consortium ([13]) has also tried to formulate answers for the problem of signature-volatility. The W3C is responsible for the development of XML (extensible markup language), which is nowadays the most popular standard for structured information exchange. The XML 1.0 Recommendation defines multiple syntactic methods for expressing the same information. That is why XML applications tend to represent the same content in different ways. Therefore, XML 'canonicalisation' was designed ([14]). The canonicalisation method uses an algorithm to generate the canonical form of a given XML document. The canonical form is the common denominator so to speak for all possible syntactic representations of a given content. A digital signature over the canonical form of an XML document allows the hash calculations to be oblivious to changes in the original document's physical representation.

It would be naïve however to believe that XML signatures (namely a digital signature generated from a hash code over the canonical form of the XML document) will solve the problem of digital signatures becoming obsolete. To begin with, the canonicalisation method developed for XML 1.0 may not be applicable to future versions of XML without some modifications. The transfer of an XML document to this newer version will invalidate the signature, since the canonical form cannot be carried indefinitely into the future. At the moment, software companies are implementing XML in their products. The multiple use of XLM and its vendor independent character give XML the status of de facto standard. But it is not very likely that XML will be maintained as a common format forever. IT will keep evolving and it is unthinkable that there will never be a better alternative for XML. A canonical form that takes all current and future formats into account is unfortunately still IT science fiction.

The only effective solution in our view for the problem of signature durability is the archival of the original binary representation of the document. This solution was proposed by the European Electronic Signature Standardisation Initiative (EESSI) in the study report 'Trusted Archival Services' ([15]). A TAS must guarantee that it will still be possible to validate an archived document years after the initial archival date, even if the applications that have been used at signature creation time are no longer in use. In other words, the TAS should maintain a set of applications (viewers as well as signature validation applications) together with the corresponding platforms (hardware, operating systems) or at least an emulator of such applications and/or environment in order to guarantee that the signature of the document can still be validated years later.

To achieve this goal in the best possible way, the TAS must only accept documents in a format that can still be understood when the format will no longer be in use. Only open file formats that are vendor-independent qualify for long-term archival. Indeed, open file formats, such as XML are being so clearly delineated that the content of such documents should be legible even 50 years from now. For today's undocumented, proprietary formats on the contrary (such as Microsoft's Word), it is likely that it won't even be possible to build emulators since no one besides the vendor has an insight in the syntax of these formats. Every TAS must therefore publish a list of supported document formats. Such a list may be exhaustive (for public services dedicated to a wide audience) or very restricted (for private services dedicated to very specific types of archival). Every time a document is submitted, the TAS must check the format before accepting it for archival.

([13]) http://www.w3c.org
([14]) Canonical XML, Version 1.0, W3C Recommendation, 15 March 2001 http://www.w3c.org/TR/2001/REC-xml-c14n-20010315
([15]) European Commission, August 2000, 37.

(16) Rothenburg, J., *An experiment in using emulation to preserve digital publications*, Amsterdam, National Library of the Netherlands, 2000, p. 74, see http://www.kb.nl/coop/nedlib/results/NEDLIBemulation.pdf

The costs and expertise required for this solution, justifies that the task of archiving signed documents will be appointed to an independent third party such as a TAS. Although contractual freedom also applies for the manner in which contracts are archived, private persons will not always be able to securely keep signed documents in their own possession. Governments must ensure TASs to be introduced in their public key infrastructure for archiving important governmental documents. Every community, from universities to notaries, from public authorities to businesses, must start thinking about how they will make use of a TAS in the future.

It is true that the emulation strategy (16) fails to gain a lot of approval in the archival community. Therefore, we propose to organise TASs in accordance with a decentralised archival model. In this model, the TAS is defined as a component of a distributed service and is only archiving signatures (and the original bitstream) while other generic archival services can remain responsible for archiving documents and keeping them legible. In this way, it is possible to reconcile signature durability and content readability.

V. A common legal framework for trusted archival services

A TAS should be able to present and validate signed documents years after their initial date of archival. As it was already indicated in the final report of the EESSI expert team, standards must be developed for the use of trusted archival services in support of electronic signatures. A clear Community framework regarding the conditions applying to TASs will strengthen confidence in and general acceptance of electronic signatures. This new legal framework should determine:

- that Member States must ensure that by accepting a signed document for archival, a TAS is liable for damage caused to an entity or a legal or natural person who relies on the TAS for presenting the signature and the validation chain as regards the immediate and complete character of the presentation. Breach of this 'obligation de résultat' should mean that liability is indisputable. A TAS should not be admitted to proof that he has not acted negligently since the loss of evidence is irreversible. Therefore, a TAS must obtain appropriate insurance to bear the risk of liability for damages;
- that there is a need for urgent standardisation to determine which conditions must be fulfilled for a validation chain to be complete. The completeness of the validation chain determines the outcome of the validation process, and thus the legal effects of the signature;
- that the archives of a TAS can never be destroyed. For the case where a TAS ceases its activities, procedures must be drafted to steer the transfer of the archives to another TAS. In order to prohibit that a TAS goes into failure, a very strict investigation regarding the financial situation and prospects should be carried out prior to the start of his activities;
- that a TAS must employ personnel who possess the expert knowledge, experience and qualifications necessary for the archival services provided;
- that a TAS must use trustworthy systems to store the documents, the signatures and the validation chains so that only authorised persons can make entries and changes;
- that a TAS, before entering in a contractual relation with a person who wants to archive a document to support the signature, must inform that person of the precise terms and conditions of the storage, such as the term of storage and the accepted file formats. Such information, which may be transmitted electronically, must be in writing and in understandable language. Relevant parts of this information must also be made available on request to third parties relying on the archived document for proof.

VI. Conclusion

It is clear that the archival of 'qualified electronic signatures' was not taken into account when writing the e-signature directive. Article 12 of the e-signature directive states that the European Commission shall review the operation of the directive and report thereon to the European Parliament and to the Council by 19 July 2003 at the latest. The archival of digital signatures used as a substitute for handwritten signatures with legal value so that they are verifiable years later for non-repudiation purposes, is one of the issues that needs to be considered in this review. We believe that the establishment of a legal framework for trusted archival services and the development of standards can take away the uncertainty that is prevalent in the archival community regarding the issue of digital signatures.

Firmas electrónicas y servicios de archivos terceros

Jos Dumortier, Sofie Van der Eynde

La legislación europea reciente ha suprimido algunas de las barreras jurídicas relativas a la creación y la conservación de los documentos electrónicos. La Directiva comunitaria 1999/93/CE regula las firmas electrónicas. En los casos en que se admite jurídicamente la utilización de documentos electrónicos, las «firmas electrónicas cualificadas» deben gozar de un estatuto equivalente al estatuto jurídico que las firmas manuscritas poseen generalmente con relación a los documentos en papel. Por otra parte, la directiva relativa al comercio exterior obliga a los Estados miembros a velar por que las condiciones jurídicas aplicables al proceso contractual (fase de archivo incluida) no generen obstáculos a la utilización de los contratos electrónicos ni impliquen una pérdida de eficacia jurídica y validez de tales contratos, por la razón de que hayan sido celebrados por medios electrónicos. Las normas jurídicas que regulan la utilización de los documentos en papel en el proceso contractual deben derogarse progresivamente. En consecuencia, numerosos documentos firmados, de gran importancia, que actualmente siguen estando en su mayoría en papel, tales como contratos y otros instrumentos jurídicos, sólo existirán en el futuro en forma electrónica.

El legislador europeo ha fomentado la utilización de las firmas digitales basadas en una infraestructura de clave privada (ICP). En consecuencia, los archiveros encontrarán cada vez más documentos electrónicos con firmas digitales. Ahora bien, parte de los archiveros son muy escépticos respecto de la aparición de las firmas digitales en sus archivos, y algunos incluso se niegan a admitir en sus fondos documentos electrónicos firmados. No obstante, una firma digital debe considerarse como parte constitutiva del documento electrónico firmado, y debe por tanto archivarse junto con el texto íntegro del documento. Los autores de la presente exposición analizan la necesidad de poder comprobar las firmas electrónicas años más tarde.

No obstante, este escepticismo no carece completamente de fundamento. Cuando se utiliza una firma digital, el control de la integridad sólo es posible si no se ha introducido ninguna modificación en los datos electrónicos en el nivel de la información binaria. Esta condición plantea un problema cuando los archiveros desean convertir datos electrónicos a nuevos formatos o transferirlos a plataformas informáticas para que puedan ser accesibles y legibles. En su exposición, los autores mantienen que las firmas digitales pueden en realidad archivarse con todos los conocimientos técnicos y la estabilidad de un servicio de archivos tercero de confianza (Trusted Archival Service), sin abdicar en nada en cuanto a la legibilidad del documento. Como conclusión, la presente exposición analizará la instauración de un marco jurídico para este nuevo tipo de servicios de confianza.

Elektronische Signaturen und Trusted Archival Services

Jos Dumortier, Sofie Van der Eynde

Mit den neuesten Rechtsvorschriften auf europäischer Ebene wurden einige rechtliche Hindernisse in Bezug auf die Erstellung und Aufbewahrung elektronischer Dokumente beseitigt. Die Richtlinie 1999/93/EG befasst sich mit elektronischen Signaturen. Wenn die Verwendung elektronischer Dokumente gesetzlich zulässig ist, müssen so genannte „qualifizierte elektronische Signaturen" einen Status erhalten, der handschriftlichen Unterschriften im Zusammenhang mit papiergebundenen Dokumenten gleichgestellt ist. Darüber hinaus sind die Mitgliedstaaten laut Richtlinie über den elektronischen Geschäftsverkehr verpflichtet, dafür zu sorgen, dass die für den Vertragsabschluss (einschließlich der Archivierung des Vertrags) geltenden gesetzlichen Bestimmungen weder die Anwendung elektronisch geschlossener Verträge behindern noch dazu führen, dass derartige Verträge aufgrund der Tatsache, dass sie elektronisch geschlossen wurden, ihrer Rechtsgültigkeit oder sonstigen Gültigkeit beraubt werden. Rechtsvorschriften, die die Verwendung von papiergebundenen Schriftstücken beim Vertragsabschluss vorschreiben, sind schrittweise außer Kraft zu setzen. Folglich werden viele wichtige signierte, bisher vorwiegend papiergebundene Dokumente wie Verträge und andere Rechtsinstrumente künftig ausschließlich in ihrer ursprünglichen elektronischen Form bestehen.

Vom europäischen Gesetzgeber wird für das Signieren der Einsatz von digitalen Signaturen auf PKI-Basis gefördert. Somit werden sich Archivare in zunehmendem Maße mit digital signierten elektronischen Dokumenten konfrontiert sehen. Die Vertreter des Archivwesens stehen jedoch dem Eindringen digitaler Signaturen in ihre Archive zum Teil recht skeptisch gegenüber, und einige weigern sich sogar, signierte elektronische Dokumente als Archivgut zu akzeptieren. Digitale Signaturen sind jedoch als organischer Bestandteil signierter elektronischer Dokumente zu betrachten und sollten zusammen mit dem Reintext des Dokuments archiviert werden. In diesem Beitrag untersuchen die Autoren die Notwendigkeit einer Verifizierung elektronischer Signaturen nach Ablauf mehrerer Jahre.

Die Skepsis ist allerdings nicht ganz unbegründet. Beim Einsatz digitaler Signaturen ist eine Kontrolle der Integrität nur möglich, wenn die elektronischen Daten auf Bit-Ebene vollkommen unverändert bleiben. Das führt dann zu Problemen, wenn der Archivar elektronische Daten auf neue Formate oder Softwareplattformen migrieren will, um ihre Zugänglichkeit und Lesbarkeit zu erhalten. In dem Beitrag führen die Autoren an, dass sich digitale Signaturen jedoch mit der Kompetenz und Stabilität, wie sie ein so genannter *Trusted Archival Service* verkörpert, ohne Abstriche an die Lesbarkeit des Dokuments archivieren lassen. Als Fazit wird die Schaffung eines Rechtsrahmens für diese neue Art von Vertrauensdienst angeregt.

Signatures électroniques et services d'archives tiers

Jos Dumortier, Sofie Van der Eynde

La législation européenne récente a aboli certaines des barrières juridiques concernant la création et la conservation des documents électroniques. La directive communautaire 1999/93/CE régit les signatures électroniques. Chaque fois que l'utilisation de documents électroniques est juridiquement admise, les «signatures électroniques qualifiées» doivent bénéficier d'un statut équivalent au statut juridique que les signatures manuscrites possèdent généralement par rapport aux documents sur papier. Par ailleurs, la directive relative au commerce extérieur fait obligation aux États membres de veiller à ce que les conditions juridiques applicables au processus contractuel (phase d'archivage comprise!) n'engendrent pas d'obstacles à l'utilisation des contrats électroniques ni n'entraînent une perte d'efficacité juridique et de validité de tels contrats, au motif qu'ils ont été conclus par des moyens électroniques. Les règles juridiques régissant l'utilisation des documents sur papier dans le processus contractuel doivent être abrogées progressivement. En conséquence, de nombreux documents signés d'une importance majeure, qui reposent actuellement toujours pour l'essentiel sur le papier, à l'instar des contrats et d'autres instruments juridiques, n'existeront plus à terme que sous forme électronique.

Le législateur européen a encouragé l'utilisation des signatures numériques reposant sur une infrastructure de clé privée (ICP). En conséquence, les archivistes vont rencontrer de plus en plus souvent des documents électroniques renfermant des signatures numériques. Or, un certain nombre d'entre eux sont très sceptiques sur l'apparition des signatures numériques dans leurs archives, et certains refusent même d'admettre dans leurs fonds des documents électroniques signés. Toutefois, une signature numérique doit être considérée comme partie constitutive du document électronique signé, et doit donc être archivée en même temps que le texte intégral du document. Les auteurs du présent exposé analysent la nécessité de pouvoir vérifier les signatures électroniques des années plus tard.

Pour autant, le scepticisme n'est pas totalement sans fondement. Lorsque l'on utilise une signature numérique, le contrôle d'intégrité n'est possible que si aucune modification n'est apportée aux données électroniques au niveau de l'information binaire. Cette condition pose un problème lorsque les archivistes souhaitent convertir des données électroniques dans de nouveaux formats ou les transférer sur des plates-formes logicielles pour pouvoir les rendre accessibles et lisibles. Dans leur exposé, les auteurs soutiennent que les signatures numériques peuvent en fait être archivées avec tout le savoir-faire et la stabilité d'un service d'archives tiers de confiance *(trusted archival service),* sans rien concéder à la lisibilité du document. En conclusion, l'exposé analyse la mise en place d'un cadre juridique pour ces services de confiance d'un nouveau genre.

Torbjörn Hörnfeldt

Torbjörn Hörnfeldt, Principal Administrative Officer at the Division for Inspection and Consulting, has worked for the Swedish National Archives (SNA) in different capacities since 1988. Originally employed to do systems development and programming he became involved in the revision of all regulations promulgated by the SNA in 1991. The revised regulations concerns everything from storage facilities to regulations dealing with information and communications technology. He has since the early 1990s worked primarily with issues in the regulatory domain. Hörnfeldt has been a project leader for several national projects, chaired the national standardisation group working with electronic document imaging and participated in international projects, the most important being InterPARES. Torbjörn Hörnfeldt studied computer and systems science at the Computing Science Department at Uppsala University between 1981 and 1985 and subsequently lectured at the same department until he left for the position at SNA.

Electronic signatures: some legal and practical issues

Torbjörn Hörnfeldt

Introduction

The theme and title of this workshop is 'the legal admissibility of digital storage'. Immediate reflections are whether the title should be completed with a question mark or an exclamation mark? Being a workshop I guess it has to be the former, even though one could expect it to be an exclamation mark given the current enthusisam for e-government, e-archives and e-everything.

Looking at the title of my presentation you may be forgiven for thinking that I really should have given it at yesterday's workshop on 'the use and legal implications of digital signatures'. And in a way you would be right, because I will talk about electronic signatures, but on the other hand you would also be wrong since I will actually not limit this presentation to electronic signatures. Instead I will try to discuss some of the complications and problems that seem to be the necessary complements when trying to regulate the creation, preservation and dissemination of electronic records at State authorities and the Swedish National Archives. In doing this I think that while some of the things I mention may be specific for Sweden, the general issues are not.

A short historical background

When I first came into contact with 'archival problems' concerning electronic records there was still a widely held opinion among public officials and archivists that the proper way to handle safekeeping and preservation of their evidential value was printing them out on paper or computer output microfilm (COM). So, the best practices for preserving electronic records was actually to move from the electronic domain.

At this time, circa 1990, the Swedish National Archives was revising the archival regulations for State public authorities and one of the principal difficulties was how to handle the fact that an increasing amount of records was created in the electronic domain only it was also becoming evident that electronic records had values, attributes, uses etc. that were very different from conventional paper records and that these could not be preserved if the electronic records were printed out on paper or transferred to computer-output microfilm. Furthermore, it had also become evident that the world's collected resources of paper and microfilm would not be sufficient if this strategy was adopted on a large enough scale. Thus, the new regulations issued by the Swedish National Archives stipulated that State public authorities could use any technical means available to create electronic or other records as long as the record could be preserved in its original form. This was quite the reverse from what had been stipulated earlier and the step taken by the SNA was quite unprecedented at that time. Regulations concerning the creation of records had been in force since the early 20th century and special dispensation was necessary for State public authorities if they wanted to create records using other materials and methods than those stipulated.

To protect the 'recordness' of records, electronic or not, the regulations issued by the Swedish National Archives were in principle implemented as functional requirements rather than as technical specifications with the exception of the technical requirements concerning paper and microfilm. One could say that the message to the State public authorities was that they could do as they wanted with regards to records creation as long as the records could be preserved without loss of information, authenticity, combinatorial possibilities and the possibility to retrieve information.

Questions posed at the paradigm shift

Taking this step, at the time a veritable paradigm shift, there were a number of questions left to be answered. The following are a few of them. Since there is in reality little possibility to preserve electronic records in exactly the form they are created the first thing that is necessary to determine is at what instant the record is created or finalised. Is it when you press the button ordering the computer to save to disk? Or is it when you sign the record with an electronic signature.

Second, is it necessary to determine if the record is what you see on the screen or what is actually being stored in one form of memory or another?

Third, what happens when you transfer the electronic record from one electronic environment to another, i.e. does anything happen to the electronic record when it is not retained in the original context?

Fourth, what are the acceptable changes one can make to an electronic record, if any? Does a change in the bitstream, that is a change in the way an electronic record is stored, change the electronic record in any way.

At the time we are discussing, around 1990, some of these questions were quite simple to answer. The determination of when a record is created or finalised is mostly governed in legislation not dependent on the technology used and there is even given some latitude to, for example, a public agency to choose to accept a record as valid or not. And as far as an electronic record being accepted as evidence in court, it is for the (Swedish) court to decide whether it accepts the record, formally finalised or not. Does the introduction of electronic signatures change this in any way?

Is the record what you see on the screen or is it what is stored in some form of memory? In Swedish legislation it is not self-evident since if you preserve a bitstring it may or may not be an electronic record depending on you having the hard- and software necessary to view it on screen or print it out on paper. The consequence of this is that the status of a bitstring is close to the status of the famous Schrödingers cat used to illustrate quantum theory. You simply do not know if a record is a record until you have accessed it. The legislation I am mentioning here is concerned primarily with freedom of information.

Is the record still valid as a record when taken out of its original context? Or to phrase it differently, is an electronic record necessarily the same as a contiguous bitstream? Given the existence and acceptance of compound records it is necessary to take into account that a record can be more than one bitstream and that it is necessary to have some metadata associated with the record to ensure that it can be preserved and presented as was originally intended.

Is it still the same record if the bitstream is changed in any way? Earlier I mentioned that the proof of a record being just that depends on if you can access it, not whether it is the same bitstream or not. One could even argue that the transformation of the stored bitstream is necessary to create and present an electronic record. Transformation of one bitstream to another, for instance when changing from one character set to another, should from this viewpoint not mean that record has been changed in any way if the transformation is reversible. That is, if you can transform the bitstream back to what it originally was. However, if the change is not reversible the record has obviously changed.

Questions after the paradigm shift (too little freedom is bad, so is too much freedom)

After the archival regulations had been issued in 1991 the reaction, if not immediate, from the State public authorities was one of consternation. If there had been too little freedom in the choice of materials and methods used to create records earlier, it now seemed to be as large a problem that there was too much freedom. To remedy this the Swedish National Archives issued complementary regulations concerning transfer of electronic records to State archival institutions in 1994. While the regulations were focused on the transfer process only they have been (mis-)used by many State public authorities as the blueprint for exemplary creation and preservation of elelctronic records. At the same time new technology, moving on, has resulted in the fact that many electronic records are created that do not fit the blueprint. This might be addressed in some ways by the pending revision of the regulations, but the technical standards that are specified in the regulations will never be quite up to date to the cutting edge of technology. I will return to the question of how the National Archives may handle this later.

The conclusion of this must be that too little freedom is bad, but so too is to much freedom.

Electronic signatures and archives

During the 1990s there was a lot of activity, especially among larger Swedish State authorities, concerning electronic signatures. While electronic seals had been in use by the Swedish Customs since 1990 this used technology that was expensive and based on the ISO X.-standards. With the increasing popularity and use of the Internet by authorities, corporations and the public the electronic signature was viewed by many as the final part in the puzzle that would allow for the paperless office and what we call e-government today. A lot of work was done, especially by the Swedish Agency for Administrative Development, the National Taxation Board and the National Insurance Board. This work was temporarily suspended as it became evident that the European Union would present a European directive that would make it possible to replace conventional, signed records with elelctronically signed records.

The European Directive 1999/93/EC was implemented in Swedish legislation on 2 November 2000 and does contain one section that has some bearing upon the use of electronic signatures and archives. In effect it is said that a qualified electronic signature can replace a conventional signature if there is nothing expressly prohibiting it. Also mentioned is the fact that communication with or between authorities can be regulated elsewhere. This could, and probably will be done in the regulations issued by the National Archives.

While the directive on elelctronic signatures has opened up a new realm of commercial possibilities it has also opened up a can of worms for the public authoritites and the National Archives. This is especially true for the Swedish National Archives since there is now not only too much freedom in the choice of technologies used to create electronic records, there is now also the fact that these records might have electronic signatures affixed to them.

A possible scenario can be summarised as follows:

— a member of the public can create an electronic record, sign it electronically and send it to a public authority (this applies equally to authorities and corporations of course);

— if the public authority has the required hard- and software, i.e. an Office package, that allows it to read the record it is now an official record that will become part of the elelctronic archives of the public authority;

— as part of the public authorities electronic archives the electronically signed record must be preserved and kept accessible to the public during the time it is meant to be kept in the archives, this includes the electronic signature;

— if the archives of the public authority shall be transferred to the National Archives, the National Arcives must in effect have the proper hard- and software to read the record and make it accessible to the public, again, this includes the electronic signature;

— if the electronically signed record is needed as proof at court or is necessary to verify a economic transaction, i.e. to retrieve VAT from another country, it is necessary that it can be read and made accessible, again including the electronic signature;

— I could go on ...

The problems associated with the scenario above are not simple to solve. Of course the National Archives can issue regulations concerning the formats to be used when creating elelctronic records and elelctronic signatures. Whether they would be possible to enforce within Sweden is one thing that is definitely uncertain, that it would not be possible to do so on a European level is definitely certain. This is also one of the reasons that the DLM-Forum and the cooperation beteween archivists, legal advisers, researchers and business is a necessity for everyone working with issues concerning the preservation of electronic records, especially those that carry elelctronic signatures.

Firmas electrónicas: algunas cuestiones jurídicas y prácticas

Torbjörn Hörnfeldt

Este documento trata algunas de las complicaciones y problemas que parecen surgir necesariamente al intentar regular la creación, conservación y difusión de documentos electrónicos en la administración pública y en los archivos nacionales suecos.

Cuando los archivos nacionales suecos revisaron la normativa sobre archivo para las administraciones públicas alrededor de 1990, se puso de manifiesto que los documentos electrónicos tenían valores, cualidades, aplicaciones, etc. muy diferentes de los documentos convencionales en papel, que no podrían conservarse si los documentos electrónicos se imprimieran en papel o se transfirieran a microfilms generados por ordenador (salida de ordenador a microfilm). Así pues, la nueva normativa publicada por los archivos nacionales suecos en 1991 establece que la administración pública puede utilizar cualquier medio técnico disponible para crear documentos electrónicos u otros, siempre que el documento pueda conservarse en su forma original.

La reacción de la administración pública a la nueva normativa en materia de archivo fue de consternación: cuando antes había habido muy escasa libertad para elegir los materiales y métodos utilizados para crear documentos, ahora parecía que había demasiada libertad. Para solucionar esto, los archivos nacionales suecos publicaron normas complementarias sobre la transferencia de documentos electrónicos a los archivos estatales en 1994. Si bien esta normativa se centraba únicamente en el proceso de transferencia, fue utilizada erróneamente por muchas administraciones públicas como modelo para la creación y conservación de documentos electrónicos.

Durante los años noventa, con la creciente popularidad y utilización de Internet, la firma electrónica pasó a ser considerada por muchos como la pieza final del rompecabezas que permitiría la gestión sin papel y la administración electrónica. La Directiva europea 1999/93/CE se aplicó en la legislación sueca el 2 de noviembre de 2000 y contiene una sección que tiene un cierto impacto en el uso de las firmas y los archivos electrónicos. En efecto, establece que una firma electrónica cualificada puede sustituir a una firma convencional si no hay nada que lo prohíba expresamente.

Mientras que la Directiva sobre la firma electrónica ha abierto un nuevo abanico de posibilidades comerciales, también ha abierto una caja de Pandora para los archivos nacionales, puesto que ahora no sólo hay demasiada libertad en cuanto a la elección de las tecnologías utilizadas para crear documentos electrónicos, sino que, además, estos documentos pueden contener firmas electrónicas. Los problemas asociados a esto no tienen fácil solución. Los archivos nacionales podrían publicar normativas sobre los formatos que deben utilizarse para crear documentos electrónicos y firmas electrónicas. No está claro que pudieran aplicarse a Suecia, pero sí es cierto que no sería posible su aplicación a escala europea.

Elektronische Signaturen: einige rechtliche und praktische Fragen

Torbjörn Hörnfeldt

In diesem Papier werden einige der Komplikationen und Probleme erörtert, die anscheinend immer wieder beim Versuch, elektronische Archive anzulegen, aufzubewahren und freizugeben, in den staatlichen Behörden und im Schwedischen Staatsarchiv auftreten.

Als das Schwedische Staatsarchiv die Archivvorschriften für staatliche öffentliche Behörden um das Jahr 1990 neufasste, wurde deutlich, dass sich elektronische Archive in Bezug auf Werte, Attribute, Benutzung usw. erheblich von konventionellen Papierarchiven unterscheiden, und dass diese nicht aufbewahrt werden können, wenn die elektronischen Archive auf Papier ausgedruckt oder auf Computer-Mikrofiche übertragen würden. Deshalb legten die 1991 vom Schwedischen Staatsarchiv erlassenen neuen Vorschriften fest, dass staatliche öffentliche Behörden alle zur Verfügung stehenden technischen Hilfsmittel einsetzen können, um elektronische oder andere Archive anzulegen, solange das Archiv in seiner ursprünglichen Form erhalten werden kann.

Die staatlichen öffentlichen Behörden reagierten konsterniert auf die neuen Archivvorschriften: Wo bisher zu wenig Spielraum bei der Wahl von Material und Methoden zur Einrichtung von Archiven bestand, schien es nun zuviel Spielraum zu geben. Das Schwedische Staatsarchiv erließ daraufhin im Jahr 1994 ergänzende Vorschriften betreffend die Übergabe von elektronischen Archiven an staatliche Archiveinrichtungen. Während sich die Vorschriften nur auf das Übergabeverfahren bezogen, wurden sie von vielen staatlichen öffentlichen Behörden als Vorlage für die Schaffung und Aufbewahrung von elektronischen Archiven benutzt bzw. missbraucht.

Mit zunehmender Popularität und Nutzung des Internet wurde die elektronische Signatur in den 90er Jahren von vielen als das letzte Teil des Puzzle angesehen, das einen papierlosen Büro- und *e*-Government-Betrieb ermöglichen würde. Die europäische Richtlinie 1999/93/EG wurde am 2. November 2000 in schwedisches Recht umgesetzt und enthält einen Abschnitt, der Bezug nimmt auf die Nutzung elektronischer Signaturen und Archive. Dort heißt es, dass eine qualifizierte elektronische Signatur eine herkömmliche Unterschrift ersetzen kann, wenn dies nicht ausdrücklich verboten ist.

Die Richtlinie über elektronische Signaturen hat nicht nur neue kommerzielle Möglichkeiten eröffnet, sondern auch Probleme für das Staatsarchiv geschaffen, da es nun nicht nur zu viele Freiheiten bei der Auswahl der zur Einrichtung von elektronischen Archiven verwendeten Technologien gibt, sondern diese Archive darüber hinaus auch elektronische Signaturen enthalten können. Die hiermit verbundenen Probleme sind nicht leicht zu lösen. Das Staatsarchiv könnte Vorschriften im Hinblick auf die bei der Einrichtung von elektronischen Archiven und elektronischen Signaturen zu verwendenden Formate erlassen. Ob dies in Schweden durchzusetzen ist, ist ungewiss, dass es auf europäischer Ebene unmöglich ist, ist sicher.

Signatures électroniques: aspects juridiques et pratiques

Torbjörn Hörnfeldt

Le présent document traite de quelques difficultés et problèmes apparemment inévitables dès que l'on tente de réglementer la création, la conservation et la diffusion de documents électroniques au niveau des autorités publiques et des Archives nationales suédoises.

Lorsque ces dernières ont procédé, vers 1990, à la révision des règles en matière d'archivage applicables aux autorités publiques, il est apparu que les documents électroniques avaient une valeur, des attributs, des utilisations, etc., très différents des documents papier traditionnels, caractéristiques qui ne pouvaient être conservées si lesdits documents électroniques étaient imprimés sur papier ou transférés sur microfilms COM (computer-output microfilm). De nouvelles règles ont été arrêtées par les Archives nationales suédoises en 1991, autorisant les pouvoirs publics à utiliser tout moyen technique disponible pour créer des documents, électroniques ou non, à condition que ces documents puissent être conservés dans leur format original.

Les pouvoirs publics ont été consternés par ces nouvelles règles: alors que la liberté qui leur avait été laissée jusque-là pour ce qui est du choix des supports et des méthodes de création des documents était trop restreinte, la latitude dont ils disposaient désormais leur semblait beaucoup trop importante. Afin de remédier à cette situation, les Archives nationales suédoises ont arrêté en 1994 des règles complémentaires concernant le transfert des documents électroniques aux institutions nationales chargées de la gestion des archives. Ces règles, qui ne portaient que sur la procédure de transfert, ont été (erronément) interprétées par nombre d'autorités publiques comme constituant un document régissant la création et la conservation des documents électroniques.

Durant les années 90, nombreux sont ceux qui ont commencé, au vu de la popularité et de l'usage croissants de l'internet, à considérer la signature électronique comme la dernière pièce du puzzle qui permettrait l'élimination du papier dans les bureaux et l'avènement de l'administration en ligne. Le 2 novembre 2000, la Suède a transposé la directive 1999/93/CE du Parlement européen et du Conseil, dont une partie peut être mise en relation avec l'utilisation des signatures et des archives électroniques. Il en ressort en effet qu'une signature électronique qualifiée peut remplacer une signature conventionnelle si rien ne l'interdit expressément.

Alors que la directive sur les signatures électroniques a créé tout un éventail de nouvelles possibilités sur le plan commercial, elle a également compliqué la situation pour les Archives nationales: à une trop grande liberté de choix dans les techniques servant à créer des documents électroniques vient en effet s'ajouter la présence possible d'une signature électronique. Les problèmes qui en découlent ne sont pas simples à résoudre. Les Archives nationales pourraient arrêter des règles relatives aux formats à utiliser pour la création des documents électroniques et pour les signatures électroniques. S'il est permis de douter de la possibilité d'appliquer de telles règles en Suède, leur inapplicabilité à l'échelon européen ne fait aucun doute.

Christine Pétillat

Christine Pétillat entered the French Archives nationales in 1974, as an assistant in the Centre of Contemporary Archives; in 1978 she took over the position of Archivist of the Ministry of social affairs and eight years later was appointed Archivist of the Prime Minister services. From 1989 to 1995 she was responsible for coordinating the word of archivists in the different ministerial departments. She is currently the Director of the Centre of Contemporary Archives, a position she has held for six years.

Electronic documents and the law — the current situation in France

Christine Pétillat

The spread of new technologies, which has been actively encouraged by the French Government for a number of years, is resulting in a dematerialisation of both business transactions and administrative procedures.

Yet strong growth of e-commerce and electronic exchanges in open networks such as the internet cannot happen without increased user trust and therefore the introduction of safeguards with respect to security, confidentiality and evidence.

France, whose legal system relies heavily on the written word, has undertaken to transpose the Community directive of 13 December 1999 into national law in successive stages, in an effort to meet the need for legal stability while remaining open to the innovations that will inevitably emerge in the sphere of information dissemination media.

The approach followed progresses from the general to the specific, namely it begins with general principles before tackling more technical aspects:

— Act No 2000-230 of 13 March 2000 adjusted the law of evidence to take account of the new information technologies, and provided for the recognition of electronic signatures;

— Decree No 2001-272 of 30 March 2001 then detailed the implementing provisions for electronic signatures.

However, the process undertaken two years ago is not yet complete. In particular, the work carried out on authentic instruments has not yet resulted in any regulation.

Some difficulties have nonetheless been highlighted, including most notably the long-term conservation of electronic documents and the signature processes attached to them.

An irreversible trend

In legal terms, the deployment of electronic resources immediately raised the issue of the validity of intangible data as regards the legal obligations pertaining to evidence. This has given rise to much debate. A few initial responses were put forward before the scale of the new developments made it necessary to carry out an overall review of the legal framework.

The spread of the phenomenon

Paperless and near-paperless technologies, initially restricted to business-to-business communication, are now being used for interrelations between the private and public sectors. And given the government's ambitions with respect to reforming the State, they are set to be extended very shortly to most of the relations between the public services and their users.

Electronic data interchange (EDI), a technology for transferring data between computers which emerged at the beginning of the 1970s as part of the effort to simplify international trading procedures, initially allowed businesses to exchange a number of details, in particular of a financial, accounting and commercial nature, with all of their partners: suppliers and customers, subcontractors, banks, insurance companies, service providers, etc.

The public authorities subsequently began accepting these electronic management methods, in various areas such as taxation, customs, social affairs and statistics, with a view to automating the issuing and processing of their compulsory documentation.

A first step was taken with Article 47 of the amending finance act for 1990, which allowed businesses, for tax reimbursement purposes, to send their invoices to their clients electronically by means of a system approved by the authorities. Decree No. 91-1403 of 27 December 1991, amended by Decree No. 95-309 of 20 March 1995, established a system for the transfer of tax and accounting data. This system enables the transmission on digital media, through relay organisations, of the documents that self-employed persons and private-sector companies must submit to the tax authorities each year. The decree of 26 February 1993 defined the principle of the dematerialisation of statistical customs declarations and authorised the transmission of goods exchange declarations by electronic means. The decree of 7 February 1994 subsequently extended this authorisation to include the declarations and payment of contributions required under the social security code. Then the act of 11 February 1994, known as the Madelin Act, gave a legal base to these innovative experiments and authorised enterprises to use distance declaration systems in their dealings with all administrative authorities.

At the same time, the concern that France, in spite of having developed the unique Minitel venture, might lose ground if it remained dependent on that low-capacity network prompted the French Government to step up the introduction of new information technologies. PAGSI, an action programme to speed up the country's move into the information society, was launched on 1 January 1998.

The programme was announced in a landmark speech by Lionel Jospin in August 1997, and was bolstered with financing from the State Reform Fund and the establishment of an interministerial framework. Its purpose is to use new technologies to modernise the administration, make public services at both central and local level more effective and improve relations between local authorities and the general public.

In the space of a few years, very significant progress has been achieved thanks to the investment made. Since October 2000, France has had a revamped and efficient public portal: http://www.service-public.fr/. Four thousand public-sector web sites have been created, along with 1 000 forms covering the administrative steps most frequently carried out by citizens. Electronic healthcare sheets are transferred through the health/social affairs network, judicial record certificates are issued remotely, and, following a reform of public contract procedures, public procurement will in certain cases be carried out using electronic means.

Despite these undeniable advances, a report submitted to the Prime Minister in the spring of 2001 called for further measures to be taken in order to make genuine services available to users. Accordingly, e-government has recently been given a new boost. The aim is to establish a consistent system designed as a central cog, rather than an added extra, in the workings of the State. A priority has been given to making more extensive use of dematerialised procedures, both for administrative tasks and for relations with enterprises and associations. The aim is now to take the protection of personal data forward and, in the near future, to make an 'electronic safe' available to all citizens for the purposes of managing their dealings with the public administration.

Despite a broad-based commitment to paperless methods — reflected *inter alia* in the forthcoming introduction of a new generation of distance procedures accessible to the general public — many people remain concerned about the securing of networks and the internet from both a legal and a technical point of view. It is therefore essential to achieve advances that can remove these obstacles and gain the support of the widest possible audience for these new practices.

An old issue in a new guise

Where evidence is concerned, the precedence of the written word and superiority of the preconstituted document have been established in French law since the 16th century (edict of Moulins). The current communication practices, involving a 'depersonalisation' of relations and 'dematerialisation' of transactions, do therefore not sit easily with the administrative formalism and binding rules applicable to legal instruments. Most importantly, digital media lack certain properties of paper on which the logic of evidence has hitherto relied.

Evidence of a legal transaction is information which, in an adversarial procedure, allows the judge to be certain of what was agreed and established at an earlier date. The French civil code sets out

five means of establishing this certainty (Article 1315–4): Documentary evidence (the written word), oral evidence (testimonies), presumptions, admissions of parties and oaths. Written evidence ranks first in this hierarchy. Furthermore, it is compulsory under Article 1341 of the code to draw up an instrument in writing and sign it, before a notary or privately, in the case of any transaction involving a sum of money above a certain amount (set at FRF 5 000 since Decree 80-533 of 15 July 1980). Furthermore, that same article stipulates that testimonial evidence is prohibited and must be disregarded where a written document exists, even if a lesser sum is involved. Documentary evidence therefore takes absolute precedence, since it can be invalidated only by other proof in writing.

However, the law of evidence does not rest exclusively on written evidence. The principle of the freedom of evidence is laid down in administrative law, criminal law and — with respect to commercial transactions — in Article 109 of the commercial code (with the reservation, however, that while a non-trader can invoke the principle against a trader, a trader cannot impose it on a non-trader).

The civil code itself sets out exceptions to the requirement for written evidence. Article 1347 provides for an exception where 'there exists a commencement of proof in writing'. It is even accepted in case law that the instrument establishing the likelihood of the alleged fact does not have to be either written or signed by the person it is being invoked against provided it is his or her intellectual work. Furthermore, Article 1348 allows greater flexibility where one of the parties 'did not have the material or moral possibility to procure a written proof of a legal transaction', making it possible to present, 'if the original instrument [has not been kept], a copy which is a reproduction that is not only faithful but also enduring. Is deemed enduring an indelible reproduction of the original which involves a non-reversible alteration of the medium'.

As the civil code did not include, until the revised version of 13 March 2000, any definition of 'written document', or 'signature', or 'original' the question of the admissibility of electronic writing as evidence has given rise to various and divergent opinions. Questions such as 'are digital documents written documents?', 'can they be regarded as preconstituted documents within the meaning of Article 1341 of the civil code?', and 'do they have probatory force?' tend to elicit contradictory responses from legal experts owing to the diversity of approaches that can be taken.

In everyday language, the written word is distinct from the spoken word in that it is enduring. The enduring marks can be alphabetical signs, codes, etc., which are understood to be writing. There is therefore nothing to say that digital inscriptions are not writing.

However, even though legal texts make no reference to paper as a specific medium, the probatory mechanism of documentary evidence quite obviously derives from the intrinsic properties and functionalities of paper. A priority is given to written evidence owing to the enduring quality it confers on information. In the case of paper, three aspects are taken into account: conservation over time, but also the physical and logical non-reversibility of the information. From the invariability of form is inferred an invariability of substance, thus establishing the original nature of the instrument's semantic content.

This line of reasoning fails in the case of a digital medium or a medium created by means of a digital process since these do not, by far, offer the same advantages. In addition to the difficulty of identifying their author, digital records present a fundamental problem in terms of how long they can last. The media they are stored on are evanescent (magnetic media) or unpredictable (CD-ROMs), since the apparent good state of the container is not necessarily indicative of the state of the content. They are not free-standing but are closely dependent on processing systems that are themselves prone to short-term obsolescence and must systematically be brought in line with the new hardware and software emerging on the market if they are to remain accessible. And every transfer, whether spatial (transmission of a message) or temporal (migration of a system), carries the risk of altering or losing information.

In short, digital media are in complete contradiction with the fixed, static qualities required of information that must have probatory force. In order to remedy this state of affairs, it is necessary not only to preserve the media carefully and ensure that no data is lost, but also to accept the

responsibility of demonstrating that the data has remained unchanged since the beginning. Only a system that can ensure the soundness of information throughout the links in the chain and the traceability of any change made to an electronic document is likely to meet this challenge.

The solutions explored

Given that digital data did not offer the relevant properties of traditional media and that the available exceptions to the rule of written evidence were not satisfactory either, practitioners worked around the difficulty by freeing themselves, through a contractual agreement pursuant to Article 1134 of the civil code, from the constraints of the law in order to develop an appropriate mechanism to validate their transactions.

Exploiting the potential of digital media, the users of electronic interchange methods attempted, through contractual relations with other enterprises and in their dealings with the public authorities, to fill the apparent legal vacuum and secure their transactions. An interchange agreement provides a solution for dematerialising documents as far as legal obligations and the law of evidence are concerned, since it identifies the messages exchanged and determines how any disputes as to the content of these messages will be settled. Technical means of ensuring the security of transfers and transactions compensate for the depersonalisation of relations and the obstacles this could cause (authentication, identification, etc.).

This approach was pioneered in the field of taxation with the procedures established for transferring tax and accounting data and drawing up dematerialised invoices. Tax and accounting data is transferred via a relay organisation which provides electronic certification. The certification plays exactly the same role as a hand signature (authentication of the signatory, non-repudiation of the instrument by its author, commitment as regards the content of the document). The relay organisation also ensures that the message received is identical to the message sent. Likewise, enterprises may send their invoices electronically if the system used can ensure that the messages sent and received are identical (the invoice message must bear at the very least the indications provided for under the legislation in force), produce a summary list that must be preserved and supply, on request from the authorities, plain text and hard copy versions.

Article 4.1 of the Madelin Act authorising distance declarations by enterprises specifies that 'each contract must *inter alia* set out the rules relating to the identification of the author of the instrument, the integrity, readability and reliability of the transfer, its date and time, the acknowledgement of receipt and its conservation'. The two above examples taken from the sphere of taxation are indicative of the ways and means used to this end: authentication by a third party, development of electronic formalities and an electronic signature, and provisions relating to checks and archiving.

The involvement of a certifying third party that is neutral and independent, for the purpose of ensuring that the transactions are completed and a trace of them is kept, is perfectly consistent with the rationale of securing messages on the networks without interfering with the applicative content. This third-party witness, whose legitimacy stems from the parties involved rather than the law, appears to be the best option for ensuring reliable identification of the sender and addressee, the integrity of the data transferred between them, the non-repudiation of the sender, and proper time and date stamping, and for producing irrefutable evidence that can be accepted by both parties in the case of a dispute.

As for other aspects, the systems used must meet specifications that have been set by decree. They must ensure constant security throughout the processing chain and whatever the operation, and must notably provide for detection of intrusion, protection against external or internal manipulation and the non-reversibility of records. Verification is carried out by means of automatically generated logbooks, precise documenting procedures and even the involvement of external auditors. In short, the dematerialisation systems must act like 'black boxes' that cannot change without leaving a trace of their successive states.

As the need for evidence is not immediate but is likely to arise in the more or less distant future, conservation over time is a significant consideration. The archiving media and methods must meet the same criteria as copies before these can be allowed to replace originals: they must be a reproduction

of the original that is not only faithful but also enduring. 'Enduring' means that the medium must be non-reversible, thus implying that the copy cannot be modified or falsified. The use of non-rewritable optical media such as write-once/read-many (WORM) discs, whose burning results in a definitive alteration of the medium, seems like an interesting option in this context. Nonetheless, its reliability will depend on a number of precautions being taken when designing and running the system.

With the aim of identifying the state-of-the-art of electronic archiving, the standard NF Z 42-013 of July 1999, drawn up by professionals under the auspices of the French standardisation association (*Association française de normalisation*), sets out a number of recommendations relating to the technical and organisational measures required when recording, storing and retrieving digital documents in order to ensure their conservation and integrity.

The agreement-based method described above, combined with measures to ensure the security of transactions and the preservation of data, has been upheld in case law. However, contractual relations must be established prior to using the system, thus making it considerably more difficult to implement in open networks.

The review of the legal framework

As the long-standing debate took on a new intensity with the spread of the internet, not everyone agreed that legislative action should be taken. Some observers even went so far as to claim such action would be premature, for the rapid advances in technology would very quickly make new legislation obsolete. Given the pressing need to ensure mutual recognition and interoperability in Europe, a major step was nonetheless taken with the decision to undertake a far-reaching transformation of the legal framework through legislation and subsequently fine-tune its implementation in practice by means of provisions laid down by regulation.

The 'digital revolution' brought about by the act of 13 March 2000

The legislative reform of the law of evidence was unanimously voted by both chambers. Its rapporteur, addressing the French national assembly, called it a 'digital revolution' on account of its clear commitment to breaking the historical link between the concept of written document and its emblematic medium: paper. However, it does respect the spirit of the previous system while incorporating into the civil code a new clause that abolishes the virtual monopoly hitherto enjoyed by paper with respect to evidence.

Clarified concepts

Although writers and case law had already recognised that digital documents qualified as written documents, it was considered useful and informative to clarify matters once and for all by incorporating definitions of the concepts of 'written document' and 'signature', constantly used in the production of evidence, into the civil code. To ensure their long-term validity, a deliberately broad phrasing was chosen to define these concepts, avoiding any reference to current technologies that are bound to be replaced in future.

The new Article 1316 provides that 'documentary evidence, or evidence in writing, results from a sequence of letters, characters, digits or any other signs or symbols having an intelligible meaning, whatever their medium and the ways and means of their transmission may be'.Since the nature of written document depends neither on the physical medium nor on how it is transferred in the case of distance communication, a digital written document has exactly the same status as a written document on any other medium.

In the case of signatures, a functional approach was taken. 'The signature necessary to the execution of a legal transaction identifies the person who apposes it. It makes clear the consent of the parties to the obligations which flow from that transaction. When it is apposed by a public officer, it confers authenticity to the document. Where it is electronic, it consists in a reliable process of identifying which safeguards its link with the instrument to which it relates.' As it merely lists the ends without mentioning the means used to achieve them, the reform of the act of March 2000 opens up the concept of signature, which is no longer merely the inscription of a name on paper in handwriting but can result from any modern process.

Explicit recognition of the probatory value of electronic writing and signatures

Data written on digital media is granted the status of a written document with probatory force as a logical consequence of the definition given of documentary evidence. Digital evidence is therefore not a distinct concept but one that is part and parcel of the concept of documentary evidence.

Admittedly, a digital document is required to fulfil a number of legal and technical conditions: it is admissible 'provided that the person from whom it proceeds can be duly identified and that it can be established and stored in conditions calculated to secure its integrity'. However, the fact that it must be possible to ascertain the integrity and author of an instrument are minimal conditions common to all forms of evidence. One consequence of this analysis is that the reservations about the admissibility of digital documents are not specific to that media and do not make it in any way exceptional.

Documentary evidence on paper and on digital media are strictly equivalent. No precedence is provided for in the event of a contradiction between a traditional document and a digital document. It is up to the judge to settle the difference by 'determining by every means the most credible instrument, whatever its medium may be'.

Likewise, 'the reliability [of the electronic signature] shall be presumed, until proof to the contrary, where an electronic signature is created, the identity of the signatory secured and the integrity of the instrument safeguarded, subject to the conditions laid down by decree in Conseil d'État'. The description given in this article is similar to that of the 'advanced electronic signature' set out in the Community directive.

The inclusion of authentic instruments

The review implemented by the act of 13 March 2000 includes all instruments; its scope therefore also covers authentic instruments. In this respect, the French legislators have gone one step further than their counterparts in the other Member States of the Union and have taken an original initiative:

'An authentic instrument is one that has been received by public officers empowered to draw up such instruments at the place where the instrument was written and with the requisite formalities.' Thanks to the role of the public officer acting in the capacity of public authority and to his or her signature authenticating the instrument, this instrument has three main effects: probatory force subject to a plea of forgery, legal date and enforceability.

Among these instruments, a distinction is made between instruments of an administrative nature drawn up by an official within the framework of his or her competence (civil status records), judicial and extrajudicial instruments (court judgements, writs, documents issued by court registrars, expert reports established further to a delegation of powers, etc.) and voluntary jurisdiction instruments drawn up to take note of an instrument or legal fact (notarial deeds, contracts, etc.).

The principle of the possibility of dematerialising authentic instruments and affixing electronic signatures to them is provided for, on condition that they are 'established and stored in conditions fixed by decree in Conseil d'État'.

The mechanism recommended by Decree 2001-271 of 30 March 2001

One year after the act was published, a decree was adopted (as provided for to implement the new Article 1316–4 of the civil code) in order to determine the conditions in which the electronic signature can be deemed reliable. It considers in turn the three distinct operations (creation, verification and certification) involved in producing electronic signatures, and outlines the provisions governing the secure electronic signature that is now a part of French law.

The concept of secure electronic signature

The set of definitions given in Article 1 include a definition of a secure electronic signature expressed in terms of the rules it must imperatively obey: it must be 'uniquely linked to the sig-

natory', 'created using means that the signatory can maintain under his sole control' and 'linked to the data to which it relates in such a manner that any subsequent change of the data is undetectable'.

Compared with an ordinary electronic signature, signatures regarded as 'secure' are subject to more specific criteria and meet the three objectives relating to law of evidence: identification of the author of the message, willingness to engage in a commitment (authentication and non-repudiation), and guarantee as to the integrity of the message emitted.

Article 2 stipulates that, in order to be presumed reliable, the signature must be 'created by a secure electronic signature creation device' and must be 'verified by the use of a qualified electronic certificate'.

Although no explicit mention is made, the decree refers to a public key infrastructure (PKI) based on an asymmetric key cryptography process — the most widely used process today. The concept of certificate used for the purposes of verification makes sense only within the context of this technology.

Certification, the cornerstone of the mechanism

Annexes I and II of the Community directive are transposed in Chapter III of the decree, which governs on the one hand the provision of certification services and on the other the approval of these services by qualified bodies to ensure that they comply with the decree's requirements.

The compulsory elements of the qualified certificate apply both to the service provider and the signatory for whom the certificate is issued. The certificate must contain the identification and secure electronic signature of the service provider, the conditions for use (period of validity) and the limits on the value of transactions for which it can be used. The signatory must be identified by his or her name or by a pseudonym, as well as by his or her title or profession where applicable. The certificate must also bear the relevant signature verification data, an identity code and an indication that it was issued as a qualified certificate.

The certificate must be issued by a certification services provider in accordance with certain rules. When issuing the certificate, the service provider must verify the accuracy of the information contained in the certificate, including the identity and validity of the public key. The provider must also provide the client with specific information in writing.

Service providers must operate a certificate directory service, use trustworthy systems to store certificates and enable them to be used as legal evidence. They must run a secure and speedy revocation service, employ suitably qualified staff and implement appropriate security measures and systems.

As provided for under the directive, no prior authorisation is required to undertake activities as a certification services provider. However, a complex 'provider qualification' system has been established. An entity belonging to the Ministry for Industry accredits the bodies responsible for the evaluation and qualification of certification service providers.

Furthermore, public bodies appointed by decree of the Prime Minister and the Minister for Industry carry out checks on all service providers, whether qualified or non-qualified, established in France.

Securing the creation of signatures

Electronic signature creation software is deemed to be secure if it meets a number of requirements and has been certified to be compliant with these requirements.

There are four requirements. The first three relate to the technical means for protecting the data used for electronic signature creation (the cryptographic keys). These must ensure that the data is unique, is protected against forgery and cannot be derived, and must also safeguard the signatory against use of the signature by others. Under the fourth provision, the software must allow the

user to establish the content of the instrument before signing it and must not alter the data to be signed. This is ensured by the 'hash' function which produces a 'fingerprint' of the document.

The certificate of compliance is issued either by approved bodies belonging to the Prime Minister's services responsible for the security of information systems or by another body created for this purpose by another Member State of the European Community.

A steering committee on certification reporting to the Prime Minister will be set up to monitor the implementation of the evaluation and certification procedures. A decree must still be adopted to establish that committee, identify its tasks and determine the software evaluation and certification procedures, as well as the procedures for approving the bodies authorised to evaluate the electronic signature creation software.

Will every new version and/or development of the software have to be certified? Will the bodies have to keep source versions as references? Pending the adoption of the decree, a number of questions remain unanswered. The general impression is that the evaluation and certification system could well turn out to be cumbersome.

Two further decrees (to create the bodies responsible for checks on the service providers and to establish the certification steering committee) must still be adopted before the decree of 30 March 2001, which transposes the annexes of the Community directive of 13 December 1999 quite faithfully, can enter into force.

Issues pending

Given the particular characteristics and exceptional importance of the probatory force of authentic instruments, the Ministry of Justice deemed preferable to undertake a substantive analysis prior to drawing up the decree provided for under Article 2 of the act of 13 March 2000. While noting that the move towards new media is already well in progress, the findings suggest the need for great caution owing to the fact that the processes currently available to dematerialise the instruments — documents whose powerful effects call for unlimited conservation — are not themselves immune from time.

In-depth analysis

The Directorate for Civil Affairs and the Seal entrusted a working group with the task of looking into the issues involved in issuing and preserving dematerialised authentic instruments. This working group operated within the framework of the public interest grouping *Droit et Justice*, headed by Isabelle de Lamberterie, and brought together persons responsible for producing authentic instruments (judges, registrars, notaries, municipal officers, bailiffs and lawyers), conservation and archiving specialists, academics, researchers and experts in the technologies involved. On the basis of a theoretical approach to the concept of authenticity and a review of the practical attempts to manage these instruments using electronic means, the summary report submitted in 2002 suggested a number of avenues for the drafting of the future instrument.

The specificity of authentic instruments

In essence, an authentic instrument is an instrument that has been drawn up by a public officer. However, it is also subject to other considerations and formalities. While these can vary depending on the type of instrument, in some cases they are required for the instrument to be valid.

In an effort to identify the underlying characteristics of the concept of authenticity amid the wide diversity of authentic instruments mentioned above, the working group established by the Ministry of Justice highlighted the common properties that might be said to form its substance: the presence of a public officer, without whom there can be no authenticity; the signature of the instrument by the public officer and, depending on the case, by the other parties to the instrument; and the formalities specific to each type of instrument.

The presence of the public officer, acting as primary witness to the affixing of signatures — notwithstanding certain exceptions to the general rule (competence of clerks for notarial deeds,

delegation of powers to municipal officers for civil status certificates, etc.) — is central to the process and one of its major components.

As for the signatures, the — essential — signature of the public officer confers the authenticity while those of the parties play a different role. By affixing his or her signature, the public officer certifies having fulfilled all the necessary verifications as to the identity of the parties, their legal capacity, their informed understanding of the commitment they are entering into and their freedom of consent. By signing, the appearing and declaring parties confirm their identity and acceptance of the instrument, thus reiterating what has already been established by the signature of the public officer.

Certain additional formalities, that can take various forms depending on the instrument, are intended to ensure that a number of conditions are fulfilled. They often provide for a particular spatial presentation; some derive from formal conditions relating to texts (notarial deeds must be written in a single context without spacing, the notary's signature on the first authentic copy of a deed bearing the executory formula and on copies must imperatively be placed on the final page, etc.)

Certain strict rules also govern the conditions under which the originals and copies can be established and invoked. These conditions vary depending on the type of instrument.

The life of authentic instruments — namely, all that has to do with communication, affixing declarations and conservation — is subject to procedures and responsibilities that are laid down in regulations. Lastly, safeguard measures are provided for (duplicate registers for civil status certificates, for instance) in order to ensure preservation.

It may seem surprising, given the complexity of the three stages in the management of authentic instruments (creation, use and conservation), that the act of 13 March 2000 opted to make explicit mention of them and envisage the possibility that they might be dematerialised. In fact, this was not provided for in the initial bill but results from an amendment taken further to pressure from notaries.

There can be no doubt, however, that it does reflect a very real need, for experimentation with the digitisation of authentic instruments is already under way.

A tide that cannot be stemmed

Most of the legal professionals involved have already, to various extents, used digital means to issue, transmit or preserve authentic instruments. This experimentation is often restricted to a specific stage in the life of the document because of the shortcomings of the law, the obstacles encountered and the lack of trust in the new technologies. Certain attempts have nevertheless yielded successful results and reflect genuine expectations in this field.

Digital media are now very frequently used as supporting tools for drawing up instruments. Notarial deeds, judgments and writs are created on computer systems with the help of templates and clause lists.

The need for a hand signature generally results in the digital process being interrupted and the document being printed for validation.

After the drafting stage, more technology is used in the form of libraries of digitised instruments facilitating searches for and retrieval of copies or extracts. Storage on CD-ROMs is opted for on account of their small size and convenient access for the purposes of issuing copies, but conservation continues to be based on traditional methods.

It is interesting to note that, alongside these sporadic enhancements, some professions are actively thinking about setting up exchange networks. The notaries have developed an infrastructure project (REAL plan) providing for the distribution of microchip cards to all members. The aim is to share joint collections and enhance communication among professionals and between them and their clients.

Civil status records no doubt offer the most scope for applications in municipalities, and have above all been the subject of a very comprehensive approach on the part of the central civil status department of the Ministry for Foreign Affairs. The department, whose activity has increased more than twofold over the last 10 years, has dematerialised the instruments, their updating and the processing of requests (including the signature of copies and extracts by officers). A digital signature — a kind of stamp consisting of an original digitised signature — has been designed and can be accessed and affixed by public officers only after a series of security checks (protocols and passwords) has been completed. Nonetheless, although the paper register, once it has been digitised, is no longer updated, the instruments, after they have been drawn up, are still printed and signed by hand, since this is the only admissible form of signature under the current regulations.

The process is under way and it would be futile to resist it. Instead, we should support it before we are confronted with irreversible damage.

Towards a new legal framework

Contending that digital media must not jeopardise the substantive safeguards offered by authentic instruments, the report submitted in February 2002 put forward a number of constructive proposals and a broad outline of the rules the forthcoming regulations should provide for.

It recommended to adopt both a general decree pursuant to Article 1317 of the civil code and specific decrees tailored to each category of instrument.

The general decree would set out the fundamental principles of authenticity and would be restricted to matters common to all instruments. It would set out the criteria required to ensure standardisation but would not go so far as to recommend or impose a particular process or technique. Furthermore, it would leave the various parties involved entirely free to organise their move to paperless methods at their own pace.

The physical presence of the public officer and the parties would continue to be imperative when the instrument is issued, whatever the options offered by electronic media with respect to distance registration of consent. The same would apply to the requirement for the public officer to verify the identities of the parties to the instrument. This verification could not merely consist of the submission of a certificate used for electronic signatures.

Electronic signature by the public officer would be possible provided that his or her authority to sign and authenticate the instrument can be checked before and after the signature takes place. Furthermore, since it would not be admissible for a private-sector service provider to take part in a process involving public officers, the certificate, in the case of a secure electronic signature, would have to be delivered either by a public body, a professional organisation or the administrative authorities.

Isabelle de Lamberterie's report also suggests that the authentic instrument should contain a number of additional details about the circumstances in which it was established.

The great unknown: sustainable conservation

In so far as the use of digital media and the use of paper have substantially different implications regarding conservation, the distribution of responsibilities, and even the institutional arrangements currently in force, should perhaps be reviewed.

Technical shortcomings

The readability of the instrument must be preserved, even after the technical tools used to create and display it on screen have disappeared. In the current state of our knowledge, the only feasible means of keeping up with the technological advances that are bound to occur over the years is to migrate the hardware, the software and the encoding formats systematically as the current ones become obsolete.

If the instrument has been signed with a cryptographic electronic signature, it will be necessary to ensure the readability not only of the document itself but also of the signature verification device.

However, any migration process will necessarily break the attached cryptographic signature, since there will be no way of distinguishing between a malicious amendment and one resulting from a migration.

How can this catch-22 situation be resolved? The solution most commonly put forward is the development of a 'resigning' or 'oversigning' system. But this will not allow the original signature to be preserved. The idea of adopting the XML format and drafting instruments using appropriate DTDs, while attractive with respect to preserving the document presentation (a significant consideration for authentic instruments), is but a short-term option that cannot have an effect on the sustainability of cryptographic signatures.

Even keeping the instruments in a central location, while ensuring greater reliability of migration operations, still leads to the same result.

Organisational uncertainty

Economic and technical reasons plead in favour of pooling the means and systems used for the storage of authentic instruments at regional or even national level. Yet these documents are still kept where they were issued for a number of years (100 years for notary public records and civil status certificates, 50 years for judgments), and are then transferred to the archive services of the relevant *département*.

Rather than scattering authentic instruments, grouping them in the care of either professional organisations or a specific body under the supervision of the French national archives would facilitate the deployment of the important resources required and would ensure the reliability of migrations.

This option, however, fails to take into account the changes that may occur in the life of many instruments (additional indications, etc.) and the duties specific to public officers.

Notaries believe they should be given direct responsibility, for the first 100 years, for all instruments, brought together in centralised digital records such as the ones being implemented by their Austrian counterparts.

The decree cannot decide on the possible changes required to the way archiving is organised. However, it could provide that the function of ensuring the integrity of the instrument should be dissociated from the traditional signature function. Solutions combining paper and digital media could also be envisaged.

Conferring probatory force on a document requires rigorous archiving and a perfect vision of the longer term. As long as we lack the means to ensure genuine sustainability and the structural integrity of the information, we must proceed with caution to prevent dematerialisation from having harmful consequences.

El documento electrónico y la justicia: situación en Francia

Christine Pétillat

La propagación de las nuevas tecnologías, apoyada activamente desde hace algunos años por el Gobierno, conduce a una desmaterialización de las transacciones en el sector privado y en los procedimientos administrativos.

Sin embargo, el desarrollo del comercio y los intercambios electrónicos en las redes abiertas de tipo Internet supone un refuerzo de la confianza de los usuarios y en consecuencia la introducción de garantías de seguridad, confidencialidad y prueba.

Francia, cuyo sistema jurídico se caracteriza fundamentalmente por la superioridad de los documentos escritos, emprendió por etapas sucesivas la transposición de la directiva europea de 13 de diciembre de 1999, esforzándose por reconciliar la necesidad de estabilidad del Derecho con las inevitables innovaciones que conocerán en el futuro los vectores de circulación de la información.

El planteamiento seguido va de lo general a lo particular, es decir, parte de los principios generales para centrarse luego en los aspectos más técnicos:

— la Ley n° 2000-230, de 13 de marzo del 2000, adaptó el derecho de prueba a las nuevas tecnologías de la información y garantiza el reconocimiento de la firma electrónica;

— el Decreto n° 2001-272, de 30 de marzo del 2001, precisó el conjunto del dispositivo de aplicación de la firma electrónica.

No obstante, el proceso iniciado desde hace dos años no alcanzó aún su resultado; la reflexión abierta sobre las escrituras públicas aún no fue consagrada por la normativa.

Es necesario decir que ya se han detectado y sacado a la luz algunos obstáculos entre los que destaca el de la conservación a largo plazo de los documentos electrónicos y los métodos de firma que les afectarán.

Das elektronische Dokument im Rechtsverkehr: die Lage in Frankreich

Christine Pétillat

Die Verbreitung der neuen Technologien, die auf Regierungsebene seit einigen Jahren aktiv unterstützt wird, führt im Privatsektor wie auch in den Verwaltungsverfahren zu einer Entmaterialisierung der Bearbeitungsvorgänge.

Der Aufschwung des elektronischen Handels und Austauschs über offene Netze wie dem Internet setzt jedoch eine Stärkung des Vertrauens der Nutzer und damit die Einführung von Garantien in Bezug auf Sicherheit, Vertraulichkeit und Beweisführung voraus.

Frankreich, dessen Rechtssystem stark durch die vorherrschende Schriftform geprägt wird, ist schrittweise an die Umsetzung der europäischen Richtlinie vom 13. Dezember 1999 gegangen, indem es sich bemühte, das Bedürfnis nach Rechtsstabilität mit den unvermeidlichen Innovationen in Einklang zu bringen, die im Bereich der Informationsträger zu erwarten sind.

Die durchgeführten Maßnahmen gehen vom Allgemeinen zum Besonderen, d. h., sie gehen von den allgemeinen Prinzipien aus, um dann die eher technischen Aspekte zu behandeln:

— Das Gesetz Nr. 2000-230 vom 13. März 2000 passte das Beweisrecht an die neuen Informationstechnologien an und gewährleistete die Anerkennung der elektronischen Signatur.

— Das Dekret Nr. 2001-272 vom 30. März 2001 präzisierte dann die gesamte Regelung zur Verwendung der elektronischen Signatur.

Der vor zwei Jahren eingeleitete Prozess ist jedoch bislang noch nicht abgeschlossen; insbesondere fanden die begonnenen Überlegungen zu öffentlichen Urkunden noch keine Verankerung in Rechtsvorschriften.

Gesagt werden muss, dass bereits jetzt einige Hindernisse aufgezeigt wurden, so in erster Linie die Langzeitaufbewahrung der elektronischen Dokumente und der mit ihnen verbundenen Signaturverfahren.

Le document électronique face à la justice: la situation en France

Christine Pétillat

La propagation des nouvelles technologies, qui se trouve depuis quelques années activement soutenue au niveau gouvernemental, conduit à une dématérialisation des transactions dans le secteur privé comme dans les procédures administratives.

L'essor du commerce et des échanges électroniques sur les réseaux ouverts de type internet suppose cependant un renforcement de la confiance des utilisateurs et donc l'introduction de garanties en matière de sécurité, de confidentialité et de preuve.

La France, dont le système juridique est fortement marqué par la prééminence de l'écrit, a entrepris par étapes successives la transposition de la directive européenne du 13 décembre 1999 en s'efforçant de concilier le besoin d'une stabilité du droit avec les inévitables innovations que sont appelés à connaître dans le futur les vecteurs de circulation de l'information.

La démarche suivie va du général au particulier, c'est-à-dire qu'elle part des principes généraux pour envisager ensuite les aspects plus techniques:

— la loi n° 2000-230 du 13 mars 2000 a adapté le droit de la preuve aux nouvelles technologies de l'information et assuré la reconnaissance de la signature électronique;

— le décret n° 2001-272 du 30 mars 2001 a ensuite précisé l'ensemble du dispositif de mise en œuvre de la signature électronique.

Le processus engagé depuis deux ans n'a toutefois pas atteint à ce jour son aboutissement; la réflexion commencée sur les actes authentiques n'a notamment pas encore été consacrée par la réglementation.

Il faut dire qu'ont été d'ores et déjà mis en lumière certains obstacles, au premier rang desquels figure celui de la conservation à long terme des documents électroniques et des procédés de signature qui leur seront attachés.

Digital storage — the need for an overall solution observing different legal fields

Per Furberg

Per Furberg

Having worked within different governmental committees from 1988 to 1997 with the task to consider IT-related legal issues in different areas, e.g., civil law, procedural law, criminal law and administrative law, Per Furberg was appointed District Court Judge in Gothenburg in November 1997. Today, he works at Setterwalls, with IT related issues. Per Furberg has been member of various international working groups. e.g., the OECD Group of Experts on Guidelines for the Security of Information Systems, the Council of Europe Committee of Experts on Problems of Criminal Procedural Law Connected with Information Technology, and the Council Of Europe Committee of Experts on Crime in Cyperspace. He has further been Consulting Expert regarding IT-related issues in different projects sponsored by the European Commission.

The need for balanced solutions regarding long-time storage and to secure future demonstration of authenticity has to be observed. Requirements regarding electronic substitutes for signed paper documents and other instruments equipped with a high value as evidence cannot be satisfied by technical standards alone. The legal and administrative demands have to be adapted, from case to case, on the one hand to the level of protection needed and on the other to the fact that complicated technical security measures may be hard to describe and evaluate in legal proceedings. These obstacles should be compensated by clear and distinct routines and understandable interfaces adopted to users, supervisors and court or arbitration proceedings.

Further, the discussion regarding trusted archival services has to be combined with the estimation of costs. In the public sector archival services are often apprehended as 'low-price services', lacking the foundation for costly development work and administration by groups of experts. Also in the private sector the introduction of electronic signatures and PKI may be held back if far-reaching demands regarding archiving will counteract the effects of IT-based rationalisations.

Digital storage — new dimensions

Legal issues brought to the fore by new technical and administrative routines are commonly gathered under headings like 'legal validity', 'liability' or 'legal admissibility'. However, the legal issues are complex and tied to different legal fields and the need for an overall solution is often neglected in preference to the special question of the hour. This is especially the case when time-honoured paper based procedures within the public administration or the private sector are replaced by digitally stored data. The quality — from a security perspective — of an original example gives a widely recognised protection from forgery and the like, normally combined with physical protection in closed rooms etc. By safely preserving the carrier, the sheet of paper, durability is achieved and the need for protection of the authenticity is satisfied by handwritten signatures.

However, legal demands are built on qualities that are partly missing in a digital document. The self-evident physical connections among carrier, text and signature in a paper document and the possibility to read a document without a transformation to human readable form do not exist in the IT environment. This trend was started already by the introduction of microfilm where the information and the carrier were given separate life cycles.

The next step towards limiting the importance of the data carrier was taken by the computerisation. Copying was not only a possibility but a necessity — binary data has to be copied to the primary memory to enable automatic data processing. Further, it is necessary to transform binary data to human readable form while, on the other hand, it is enough to enlarge micro film. As a result of the frequent updating of computer software it also becomes necessary to convert binary data to secure the readability. These measures will decompose also the original binary patterns of data representing the record.

Consequently, the relations between text, 'signature' and carrier in an electronic document is not univocal and 'locked' in the same way as in the paper environment and separate life cycles are given not only to the information and the data carrier but also to the information and the binary pattern of the data. This confers new dimensions to the preservation and protection of such records.

Partial modifications of a record — compensating measures

These new dimensions have been reflected by the regulations introduced by the Swedish National Archives in accordance with the Swedish Archives Act. Public records may be 'sorted out' only in

accordance with a decision by the National Archives. The problem is how to apply these regulations when the information, the binary data representing the information and the data carrier are given separate life cycles. Which kind of partial modifications will be regarded as 'sorting out' the record?

To satisfy the demands for safe long-time storage the National Archives have taken their starting point from the maintenance of the original contents, contrary to the traditional storage of a physical original example. This protection is aiming at the information (in a broad sense), not necessarily at the data carrier or the binary pattern. This approach has enabled a flexible transition from paper based to electronic records, by scanning incoming paper documents and accepting the delivery of documents in electronic form.

To avoid defective documents and files as a result of partial modifications, the National Archives have considered measures, when an authority transfers a record to another data carrier, to be a type of 'sorting out', if the transfer has involved a loss of:

(1) information;

(2) possible information combinations;

(3) the possibility to establish the authenticity of the information; or

(4) search ability.

With the support of so-called compensating measures — e.g. routines to control the transfer to new data carriers or the conversion of data from one format to another and to certify the correctness of the procedures — electronic records have been widely accepted by the public authorities as substitutes for paper based originals. These measures aim at preserving the integrity as well as the authenticity on a long-term basis.

Digital signatures and legal concepts

However, nowadays the developments in the field of information security and electronic records have been founded on digital signatures and related cryptographic routines. Such routines may support different life cycles of the information and its carrier but not a decomposition of the original binary patterns of data representing the record. No other unique characters to secure by technical measures exist than those bound to the bit pattern. By securing this pattern, the digital signature provides a 'lock' for the authenticity, corresponding to the lock for text and signature given by paper.

However, the legal analysis has mainly focused on certification authorities and administrative issues tied to supplying customers with certification services. The needs to adjust the regulations applicable on the electronic documents are often neglected. There is a need to find uniform solutions for different legal fields. However, laws and regulations regarding document management are normally built on legal concepts established long ago, influencing everyday life almost like laws of nature and enacted to satisfy the needs in a paper-based environment where a manipulation of a document has to be a material attack, traceable upon the physical object. An individual — often unconsciously — makes a visual authenticity control when he is reading an important paper document. A manipulation of a digital record consists of untraceable alterations of a bit pattern which may be manipulated during the transfer among different states and medias.

Consequently, it has often been questioned whether wordings like 'written' and 'signed' are meant to include electronic equivalents to paper media. Another area, brought up in different legal fields, is the interpretation of provisions founding its legal effects on the physical location of a document, its delivery etc. The problem is that the debate often has been limited to statements indicating that electronic records are or are not a 'record', 'written' etc. according to the wordings of laws and regulations. This discussion only on rare occasions arrives at the stage where the laws and regulations are analysed and conferred to IT.

A starting point from objects or functionality?

The legal issues may be analysed from a variety of starting points. A study, based on IT and data's character together with descriptions of the digital object, may aim at considering whether the electronic record is a physical object or something immaterial and if it is possible to possess and be the owner of a digital record in the same way as a traditional physical object. A similar angle of approach, taking the legal system and judicial classifications into different kinds of property as the starting point, is to ask whether the electronic record shall be considered to be chattels or another legal title, tied to the digital data. The problem is that any interpretation from the mentioned starting points – concealed in an analysis of terms and definitions – will probably reflect the interpreter's own intent, not the legislator's, regarding these legislative questions. The laws and regulations have usually not been adopted during the time of computers. Details in the wordings or the legal technical solutions can therefore hardly give any answers.

The functions which the electronic document is intended to fulfil will probably make a more fruitful starting point. The determining factor will be whether the document can offer the same functionality in electronic form. There is nothing new with such a 'functional' starting point. Traditional physical objects have been suited to function within certain administrative or commercial patterns of transactions and, when these patterns have changed the legislature has adapted the objects' functions to these new patterns. How the electronic document is constructed and represented, if it is considered to be a physical object etc. will consequently have to come second. These differences are of limited interest from a legal point of view if the digital instrument, independently of how it is constructed, fulfils the same functions.

However, this does not mean that the character of the electronic document and the kind of property it represents should be completely ignored. Wordings and existing law have a considerable power over our thoughts. Therefore an analysis is hardly possible without, on the one hand, metaphors such as 'documents', 'signatures' kept in 'electronic archives', 'handed over' and 'received', and, on the other hand, a comparison with rules and regulations regarding different kinds of property and instruments to establish the legal validity.

No once-for-all-solution

The rapid developments in the IT field have made it possible to recreate almost any function fulfilled by paper documents. One reason for the legal uncertainty may be that paper documents are normally shaped to include all the qualities needed to fulfil any of these functions, while IT-based material as a rule only comprises the functions which are needed according to the design of the information system. Consequently, the homogeneous paper document is replaced by a wide range of electronic substitutes, the quality of which may be hard to evaluate without expert knowledge.

This development and the focus in general on the information as such — not on the construction of the electronic record and the security mechanisms — will make it challenging to apply traditional legal concepts. It will become apparent that the desirable once-for-all-solution, pointing out the electronic document concept, fit for each application, does not exist. The development work for international standardisation seems, however, to try to recreate an electronic version of a universal concept document including all its qualities, independent of the needs realised within the current application.

Virtual places

Further, it is often stated that time and place have lost their meaning in the electronic environment. It is true that data in transit representing e.g. an application for a summons or an injunction will reach their destination much faster with the support of IT and may be communicated globally, but the same need to divide risks and responsibilities between sender and addressee will exist, if the item of mail is delayed, mutilated or does not arrive. When electronic filing and electronic mailboxes are put into practice the authority creates an electronic equivalent to a post-office box assigned as the electronic place of delivery, either housed on the authority's premises or outsourced or otherwise located somewhere else, but functionally equivalent to a mail receiving function within the authority's office.

Thus, it is necessary to create clear borderlines for electronic places, e-mail boxes and the like and to elaborate distinct views on how to judge when these borderlines are reached and crossed. The question is how these different aspects should be reflected when archiving such materials.

Digital bearers

The same need exists to clarify when digital bit strings, designed to carry legal rights, will be apprehended as documents, coins or bills, the possession and transfer of which will be decisive for the parties' legal rights. A paper document may, as a physical example, be in a person's possession, thereby establishing a legal right. This function is not possible to recreate in the IT environment without special security measures. As already pointed out, the unique aspect is related to a unique pattern of data rather than to a unique physical example. Consequently, the function of a paper document to transmit authority as an original example cannot in a simple way be recreated digitally if the mobility should be preserved. For that reason certain applications, such as doctor's prescriptions and shipping documents, demand some form of registration in the IT environment; confer the functionality within the EU project 'Bolero'. Another possible solution to preserve this function is in principle to substitute the functions related to the possession of a unique and tangible piece of paper by the exclusive and personal knowledge of a non-tangible secret code which is given only to the person who has the authority.

In this context, the functionality of the information systems and the existence of services to create the electronic equivalent to possession and *traditio* of physical examples should be decisive for the interpretation of legal acts enabled by IT. According to the travaux préparatoires of the Swedish Law on electronic money (e-money), a digital bit string secured from double spending is considered to be the carrier of a certain legal right, controlled by the 'possessor' and possible to 'hand over' (*traditio*). The legal principles applicable to other traditional bearer instruments, such as bearer bonds, bills and coins or chattels should be applicable to e-money. On the other hand, the provisions regarding non-negotiable claims may not be applied to e-money, as the stored monetary value is handed over when a payment with e-money is accomplished (prop. 2001/02:85 s. 60). This is a good example of the possibility to solve the legal issues within current law. However, it is unclear how this functionality should be reflected when the instruments and transactions are archived.

The need for flexibility

Within international fora, experts on IT and the law are often pleading that a wide opportunity should exist to apply current legislation directly or by analogy with regard to electronic documents. The only missing link is said to be detailed standardisation on different levels. However, these possibilities are questioned domestically and protracted reviews to clarify the legal situation are ongoing. Further, adjustments of laws and regulations regarding archiving are only the starting-point to adopt other sets of rules to IT. The approach chosen by the Swedish Government regarding e-money makes it, simplistically described, possible to treat electronic 'documents' as documents, electronic 'cash' as cash, account-based systems as account-based disposals, within current law. On the other hand, an approach based mainly on the historical meaning of the wordings, disregarding the practical usage, will call for extensive amendments of laws and regulations.

Hopefully flexible views on partial modifications and compensating measures to secure the integrity and authenticity of electronic documents will be accepted to enable balanced solutions for long time storage. There is a risk on the one hand to focus only on technological improvements and the creation of complex and costly standards, putting a burden on the authorities and the business community to use measures not necessary for having reasonable assurance — conferred with the use of traditional paper-based routines — that the users and the judicial system will accept the material as authentic. On the other hand the lack of generally accepted methods and necessary facilities to maintain the signed objects may hamper full-scale systems for electronic records management. The question is how to create flexible solutions on a long term basis without being hampered by laws and regulations or standards incompatible with the speedy development driven by IT.

An example from practical life

The need for a practical turn may be illustrated by the following example. If an incoming document to a public authority is created with a record format, hard to retain and transform on a long term basis and furnished with an advanced electronic signature supported by a certificate by an issuer unfamiliar to the receiving authority and given a format incompatible with that authority's information system, the practical solution would be to verify the signature and convert the mate-

rial in combination with secure registration and attestation of these measures and their reliability. But, words and traditional concepts have a considerable power over our thoughts and the immediate reaction may be to argue for the preservation of the 'original' example saying that a 'copy' is not secure as the 'original' signature cannot be verified anymore. However, a further study of data's character will show that there is no original example in the IT environment, only original contents normally represented by a multiplicity of 'copies'.

The next objection against verifying the signature and converting the record to one format suitable for the authority's system and for long time storage — followed by signing the record and sorting out the record with the received format — will probably be that these measures will leave no possibility to verify the 'original' signature. Consequently, a certified copy should not be able to give a reasonable protection for the authenticity of the 'copied' text and the information regarding who is the drawer of the record. However, the situation is different in the electronic environment when advanced or qualified electronic signatures and PKI are applied. A signature on a paper-based original is not verified by an expert until its authenticity is questioned and the material to carry out such verification is not obtained until a verification procedure by an expert is initiated. But, the use of PKI means that the material needed for verification is obtained beforehand and that a 'real' verification is normally completed as soon as the record has arrived at the addressee's mail box. Thus, an attestation of a converted electronic record, after verification, that the electronic signature was authentic and that the contents of the converted record are unaltered may be comparable with a certificate of authenticity by an expert after having obtained the material for comparison and carried out an analysis of genuineness. The statement given in the paper 'Electronic signatures and trusted archival services' (Jos Dumortier and Sofie Van Den Eynde) that a certification of a trusted third person that the indicated signatory signed a previous version of a reformatted record, which is identical with respect to the content, has at best the value of a certified copy, is obviously wrong.

Consequently, two alternatives may be discussed as a starting point for the future legal and administrative solutions:

1. preserving the signed bit pattern and the information needed to verify its authenticity on a long-term basis, supplemented with the result of an initial verification and other meta data. The result will be many file formats and certification authorities involved;
2. initial verification combined with secure routines to convert the record to a format, uniform for the receivers' archives, and to register the result of the verification and put the authority's digital signature on the record to secure its authenticity on a long-term basis. The result will be no mixture of formats and procedures.

Closing lines

A certificate that a signed electronic record has been verified and found authentic and that the conversion to another format, within a secure system, has been carried out correctly is a completely different situation than a certification of a paper-based copy. The electronic signature has initially been securely verified and the correctness of the converted version may be certified and signed electronically in accordance with special security routines, adapted to the level of trust needed in the present application. A demand on authorities and companies to preserve different formats on a long-term basis for security reasons may in practice show to be hard to combine with limited technical and administrative capacity and the need for reduction of costs. Standardisation may solve some problems, but may probably not be enough to avoid expensive third party services.

Requirements regarding electronic substitutes for signed paper documents and other instruments equipped with a high value as evidence cannot be satisfied by technical standards alone. The legal and administrative demands have to be adapted, from case to case, on the one hand to the level of protection needed and on the other to the fact that complicated technical security measures may be hard to describe and evaluate in legal proceedings. These obstacles should be compensated by clear and distinct routines and understandable interfaces adopted to users, supervisors and court or arbitration proceedings.

Further, the discussion regarding trusted archival services has to be combined with the estimation of costs. In the public sector archival services are often apprehended as 'low-price services', lack-

ing the foundation for costly development work and administration by groups of experts. Also in the private sector the introduction of electronic signatures and PKI may be held back if far-reaching demands regarding archiving will counteract the effects of IT-based rationalisations.

Almacenamiento digital: la necesidad de una solución global que respete los distintos ámbitos del Derecho

Per Furberg

Conviene tener en cuenta la necesidad de encontrar soluciones equilibradas para el almacenamiento permanente y garantizar la demostración posterior de la autenticidad de los archivos. Las normas técnicas no bastan de por sí para satisfacer las condiciones que deben cumplir los sustitutos electrónicos de los documentos en papel firmados y otros instrumentos de gran valor probatorio. Los criterios jurídicos y administrativos deben adaptarse individualmente en función, por una parte, del nivel de protección necesario y, por otra, de la idea de que las medidas de protección técnicas complejas pueden ser difíciles de describir y evaluar en el marco de un procedimiento jurídico. Estos obstáculos deben ser compensados con la instauración de rutinas claras y distintas y de interfaces comprensibles, adaptadas a los usuarios, a los supervisores y a los procedimientos judiciales o de arbitraje. Además, el debate sobre los servicios de archivos terceros de confianza (Trusted Archival Services) debe asociarse a una estimación de los costes. Por lo que respecta al sector público, los servicios de archivo suelen considerarse como servicios de bajo coste, que no permiten trabajos de desarrollo costosos ni la administración por grupos de expertos. Del mismo modo, la introducción de las firmas digitales y de las infraestructuras de clave privada (ICP) puede retrasarse en el sector privado si unas exigencias demasiado elevadas en materia de archivo contrarrestasen las soluciones de racionalización que recurren a las tecnologías de la información.

Digitale Speicherung – Die Notwendigkeit einer umfassenden Lösung unter Beachtung verschiedener Rechtsbereiche

Per Furberg

Dem Erfordernis ausgewogener Lösungen für die Langzeitsicherung und den zukunftssicheren Nachweis der Authentizität ist Rechnung zu tragen. Anforderungen an elektronische Substitute für signierte Papierdokumente und andere Schriftstücke mit hohem Beweiswert lassen sich nicht allein durch technische Standards erfüllen. Die rechtlichen und verwaltungstechnischen Ansprüche müssen je nach Fall zum einen an das erforderliche Schutzniveau und zum anderen daran angepasst werden, dass komplizierte technische Sicherheitsmaßnahmen in Gerichtsverfahren schwer zu beschreiben und zu bewerten sind. Diese Hindernisse sollten durch eindeutige Routinen und verständliche Schnittstellen für Benutzer, Aufsichtsstellen sowie Gerichts- oder Schiedsverfahren kompensiert werden.

Außerdem muss die Diskussion über vertrauenswürdige Archivdienste mit einer Kostenabschätzung kombiniert werden. Im öffentlichen Sektor werden Archivleistungen oft als „Billigdienste“ begriffen, denen die Grundlagen für kostspielige Entwicklungsarbeiten

und eine Verwaltung durch Sachverständigengruppen fehlen. Im privaten Sektor kann zudem die Einführung von elektronischen Signaturen und PKI zurückgestellt werden, wenn weitreichende Ansprüche an die Archivierung den Wirkungen IT-basierter Rationalisierungsmaßnahmen entgegenstehen.

Stockage numérique: la nécessité d'une solution globale respectant les différents domaines du droit

Per Furberg

Il convient de bien voir la nécessité de trouver des solutions équilibrées pour le stockage permanent et de sécuriser la démonstration ultérieure de l'authenticité des archives. Des normes techniques à elles seules ne peuvent pas satisfaire aux conditions requises des substituts électroniques aux documents «papier» signés et d'autres instruments revêtant une grande valeur probatoire. Les critères juridiques et administratifs doivent être adaptés au cas par cas en fonction, d'une part, du niveau de protection nécessaire, de l'autre, de l'idée que des mesures de sécurisation techniques complexes peuvent être difficiles à décrire et à évaluer dans le cadre d'une procédure juridique. Ces obstacles doivent être compensés par la mise en place de routines claires et distinctes et d'interfaces compréhensibles, adaptées aux utilisateurs, aux vérificateurs et aux procédures judiciaires ou d'arbitrage. En outre, le débat sur les services d'archives tiers de confiance *(trusted archival services)* doit être associé à une estimation des coûts. En ce qui concerne le secteur public, les services d'archives sont souvent perçus comme des services à faible coût, qui ne permettent pas des travaux de développement coûteux et une gestion d'experts. De même, l'introduction des signatures numériques et des infrastructures de clé privée (ICP) peut être retardée dans le privé lorsque des exigences trop poussées en matière d'archivage vont à l'encontre des solutions de rationalisation faisant appel aux technologies de l'information.

PLENARY CLOSING SESSION

Wednesday 8 May 2002

Chair: Erik Norberg (Sweden)
Hans Hofmann
(European Commission)

George MacKenzie

George MacKenzie is Keeper of the Records of Scotland (head of the National Archives of Scotland), since January 2001. Before that he was director of external relations in the National Archives. He worked for the International Council on Archives as Deputy Secretary-General in their Paris office for two years, 1995–96, where he was responsible for the professional programme. He continued to work with ICA on an honorary basis until 2000. He has carried out consultancy work on archives for ICA, Unesco, and the World Bank.

Report on the DLM-Forum 2002: parallel sessions 3 and 4

George MacKenzie

Two parallel sessions looked at content management and delivery and organising records and archives. The first concentrated on establishing and maintaining archival systems, best practice and easy to use access, while the second looked at metadata, standardisation and model requirements.

The two sessions took a wide perspective, from the theoretical to the practical, covering both public and private sector organisations. Speakers and participants were from a variety of backgrounds, ranging from commercial firms to government archives.

The papers provided an overview of electronic records management in Europe. Philip Lord, an electronic records management (ERM) consultant with a background in the pharmaceuticals industry, pointed out that there were few true digital archives and that the private sector was rather behind the public sector in developing ERM. The exceptions, he thought, were in industries where there was a strong awareness of records, such as pharmaceuticals.

Lluis-Esteve Casellas Serra gave a snapshot of the management of electronic records in Catalonia, though much of what he said, was of relevance in the wider European context. He stressed that electronic records were being widely generated now, though they are not always recognised as such. He talked of a dissociation or a mismatch, between the production of electronic records, and their management. This is a simple, but important point for DLM — narrowing that gap must be one of our aims. The study gave some valuable background statistics of the use of e-systems and Internet in Catalonia. It concluded that Internet and e-mail were already prevalent in both public and private sectors, and widely used by individual citizens, but that e-commerce, meaning interactive transactions between customers and suppliers, was not yet a reality. Although the use of electronic records alone was growing, it was not yet at the corporate level, but remained at the intermediate level. It was clear that strategies and procedures for appraisal and preservation of electronic records were needed, and urgently. The key issues were the legal framework as a driver to change, the influence of technological factors and the use of document management as a business tool, rather than something for record managers. The contribution of the archivist to the e-records process however, remains vital, in particular the methodological approach, the ability to distinguish the important from the unimportant, the long-term value from the transitory. The public sector can show the way to the private sector. Allied to this, there is a need for specialised, continuous training for those involved in managing electronic records.

The motives for adopting electronic record systems were also dealt with in the sessions. Philip Lord argued that where systems have been set up, this is not because managers care about history, but because they care about reducing risk and containing costs. The importance of electronic records management for business, rather than cultural reasons, was emphasised by Celsa Poza from the Catalan Water Agency. She described the electronic document management system devised to support the business objectives of a major public service agency. The aim of this integrated electronic system is quite simply to increase efficiency and improve the service to the customer.

The sessions also identified the role of external regulation in developing an awareness of electronic records issues and stimulating change. Philip Lord argued that the US Food and Drug Administration Regulation CFR 21 Part II, will have a major influence on private and public sector organisations, obliging them to improve the way they manage their electronic documents and records. It will affect not only organisations that deal directly in this market, but others too, due to the size and importance of the sector. The influence of regulation, this time from a national government, was also seen in the paper by Jan Sorenson of the Danish National Archives. He described their system for identifying, transferring and preserving electronic records, which appears to be one of the most proactive approaches by a national archive in Europe.

A continuing theme in the sessions was the importance of metadata or contextual information. In order to support migration of data as the key means of long term preservation, the Danish archives

have developed the concept of the 'archival version', of a digital product. This includes the data itself, metadata about the data, context information, and additional archival information about the provenance of the data. Probably some archivists would bundle the last three elements together as generic 'metadata' but this is a useful way of distinguishing the different types of contextual information, that enable us to understand and use the content of an electronic archive.

Everyone involved in electronic records management is aware of the importance of metadata, but not everyone understands this complex subject. It was therefore valuable to have a guided tour from Johannes (Hans) Hofman of the Netherlands Archives. He stressed that in understanding metadata, we have to understand that metadata serve different purposes. The creator, the custodian and the user of information may all be different. When this happens, they will have different perspectives and consequently, different metadata requirements.

Among the issues to be dealt with in metadata schemes are organising structures, identification, access and interoperability, the existence of different domains and terminologies, and persistent linking of digital objects. OAIS, Dublin Core, the RLG preservation metadata scheme, the Australian SPIRT record-keeping metadata project, have all produced metadata schemes. In the current work on ISO 15489, the technical committee working group have recognised the confusion of so many different schemes and have decided that another standard would not be much use. Instead they aim to demystify what already exists and enable users to make an intelligent choice. Johannes Hofman argued that coordination and cooperation are needed across the different domains, and that archives should work with libraries and information scientists. He also introduced the idea of meta level, arguing that we may need further levels of metadata.

The importance of contextual information was underlined by Margaret Hedstrom of the University of Michigan in her report on the Camileon project. This asks the simple, but vital question: what are the significant properties of digital objects? Such properties are defined at the moment by encoding schemes, document type definitions (DTDs) and formal descriptions, as well as through software and human judgments. The project notes that different properties will be important in different contexts. It has developed an inventory of over 800 types of significant property. It also makes an interesting definition of complex digital objects, which are those that combine two or more component types. Component types include text, images, video, graphics etc. The project is referencing its findings against the OAIS model, and trying to distinguish structural properties from semantic ones.

Contextual information is also central to the Collate project for historical film research discussed by Ulrich Thiel. This is experimenting with techniques to present information in context to the user. It creates a digital repository of record material about films, and adds to this two elements of metadata: commentaries from film experts on the material, and the views of users. This is a classic application of knowledge management techniques to make archive information more widely available and more useful.

As well as considering the conceptual planning needed in advance of adopting electronic systems, the sessions also gave practical advice to private and public organisations on how to develop an electronic records management strategy. Martin Waldron, a consultant specialising in helping organisations to electronic working introduced the concept of EDRMS, or electronic documents and records management strategy, in place of EDM or ERM. His definition added workflow, forms processing and web services to the usual document and records management. Again reflecting the view that commercial motives are most important in electronic management of records and documents, he argued for an approach that identified the business need, and focused on the business benefits of any strategy or system. The benefits have to include both monetary and non-monetary ones. When identifying costs, the longer term running costs of any system have to be taken into account. He also advocated risk management to evaluate different options. Among the strategic drivers of an EDRMS are: to support electronic business, to comply with new legislation (freedom of information and data protection for example), or as part of a wider initiative such as modernising government.

The case study of Eurocontrol, a public agency responsible for air traffic management across most of Europe, gave further backing to this view. Jean-Claude Van Lancker explained how an earlier document management system had failed, mainly because the users could not or would not agree on their requirements. Based on this experience, the critical success factors in the new project were to understand the different components such as workflow and document management. In other

words they had to let the business lead the project, not the IT specialists. They also had to implement the project in phases. User satisfaction was important, and the ability to scan in documents was one simple yet important step to acceptance. There was less success in managing the document repository, and blockages occurred.

When government agencies are planning and procuring systems it is important that they follow the correct procedures, and Martin Waldron pointed out that the model requirements from the EC (MoReq) are already showing their value in this respect. Discussion about MoReq identified a need to translate it to further languages. Discussion also focused on the need for training and for greater awareness as keys to smoothing change.

One practical implementation of standards is the United Kingdom Public Record Office (PRO) functional requirements, described by Malcolm Todd in the last paper of session 4. He explained that the factors influencing them were the modernising government initiatives and information legislation. The aim of changing to new electronic systems is to enable new ways of working, to help better government, and specifically to make the life of the citizen easier. At present a citizen typically has to select, from a bewildering variety of agencies, which one will provide the service he or she needs. Instead, the citizen should be able to select a service, and let the agencies work out the complex means of delivering it to the citizen.

The PRO functional requirements are arranged in a virtual pyramid, ranging from core needs at the top to additional supporting ones and then on to non-functional requirements. The further up the pyramid, the more important records management becomes. He stressed that the PRO are unique in Europe in offering formal evaluation of mandatory requirements; the only other agency doing this in the world is the US Department of Defence. However, the PRO do not certify or accredit, they merely provide assurance that a product met the generic requirement in an independent test. Approval is given to products, not suppliers and approvals last two years.

Malcolm Todd concluded that there is much common ground between the PRO functional requirements and MoReq. Both aim to be practical and to operate at a lower level than the new international standard on records management, ISO 15489. Overall, there are two very positive signs. First, there is greater maturity and compliance in the market and second, acceptance by major software suppliers of the functional requirements is growing.

Conclusions

The two parallel sessions provided a number of general conclusions. First the factors promoting ERM are not archival, they are not to do with better records, they are not about better sources for historians of the future. They are much more basic and immediate. Organisations install ERM to support their business, to reduce risks and to contain costs, and to improve service to their customers. The role of external regulation is also important, especially information legislation and other statutory systems.

Second, the public sector is generally more advanced in the management of electronic records than the private sector. The one exception is in specific areas such as pharmaceuticals where there is a long-term business need and where there is strong external regulation. The public sector can and should use its position to show best practice and encourage development in the private sector.

Third, electronic records, like other records, are a corporate asset to an organisation, whether public or private, and should be managed as such. Being in electronic form means that records can be more widely shared than paper ones can be. A true electronic records system opens the possibility of a single instance, a single occurrence of a record, which many users can share, rather than having multiple copies owned or used by multiple users.

Fourth, the development of external standards and documents, such as the MoReq, are of considerable value in helping to establish needs and in the procurement process. Several of the case studies described using MoReq and the similar functional requirements guidelines by the UK Public Record Office. Their value lies in their being external and neutral, and being recommendations, rather than mandatory.

Fifth, there is a need for further and continuous training for personnel involved in managing electronic records. This should be addressed by the sectors involved in training records managers and archivists.

Informe sobre el Foro DLM 2002 de Barcelona: sesiones paralelas 3 y 4

George Mackenzie

Las sesiones paralelas 3 y 4 se centraron en la gestión y presentación de contenidos, y en la organización de registros y archivos.

Estas sesiones pusieron de manifiesto una amplia producción y utilización de los archivos electrónicos en Europa, pero se prestó menos atención a su gestión. La reducción de esa diferencia debe ser uno de los objetivos del Foro DLM. Existen pocos archivos digitales auténticos, y el sector privado está mucho más atrasado que el sector público en el desarrollo de la gestión de archivos electrónicos.

Un tema recurrente en las sesiones fue la importancia de los metadatos o información contextual. Quizás las cuestiones más básicas sean: ¿Cuáles son las propiedades significativas de los objetos digitales? ¿Qué necesitamos conservar en los metadatos, en nuestros sistemas de conservación de archivos? Esto se está abordando en parte en los proyectos de investigación.

Junto a estas cuestiones, existe un interés generalizado y un mayor uso de normas tales como los modelos de requisitos o MoReq elaborados por la Comisión Europea, y los requisitos funcionales elaborados por la Public Record Office del Reino Unido.

También se distingue cada vez más entre la gestión de documentos electrónicos (EDM) y la funcionalidad adicional de la gestión de archivos electrónicos (ERM). Un ponente introdujo el concepto de estrategia de gestión de archivos y documentos electrónicos, lo que añade el flujo de trabajo, el tratamiento de impresos y los servicios web a la gestión de archivos.

De las sesiones se obtuvieron las siguientes conclusiones:

- las organizaciones instalan sistemas de archivos electrónicos en apoyo de su actividad empresarial, en vez de para crear archivos históricos;
- la función de la normativa exterior también es importante, especialmente la legislación sobre información y otros sistemas legales;
- el sector público está generalmente más avanzado en la gestión de archivos electrónicos que el sector privado, excepto en ámbitos específicos como el sector farmacéutico, donde existe una necesidad comercial a largo plazo o una fuerte normativa exterior;
- el sector público puede y debería utilizar su posición para mostrar las mejores prácticas y fomentar el desarrollo en el sector privado;
- los archivos electrónicos constituyen un activo para las organizaciones, ya sean públicas o privadas, y deberían gestionarse como tal;
- el desarrollo de normas exteriores y neutras, tales como los MoReq, tienen un gran valor para contribuir a determinar las necesidades y en el proceso de adquisición;
- existe una necesidad de formación adicional y continua para el personal que participa en la gestión de documentos electrónicos

Berichterstattung auf dem DLM-Forum, Barcelona 2002: Parallelsitzungen 3 und 4

George MacKenzie

Die parallelen Arbeitsgruppen 3 und 4 befassten sich mit Content Management und Delivery sowie mit der Archiv- und Schriftgutverwaltung.

Ergebnis war, dass die Erstellung und Nutzung von elektronischen Archiven in Europa weit verbreitet ist, ihrer Verwaltung jedoch weniger Aufmerksamkeit geschenkt wird. Eine Verringerung dieser Lücke muss eines der Ziele des DLM-Forums sein. Es gibt nur wenig wirkliche elektronische Archive, und der Privatsektor hinkt dem öffentlichen Sektor bei der Entwicklung von Systemen zur Verwaltung solcher Archive hinterher.

Ein durchgehendes Thema der Arbeitsgruppen war die Bedeutung von Metadaten oder Kontextinformationen. Die wahrscheinlich wichtigsten Fragen jedoch waren: Was sind die wesentlichen Eigenschaften digitaler Objekte? Was müssen wir in den Metadaten in unseren Archivaufbewahrungssystemen bewahren? Dies wird teilweise in den Forschungsprojekten behandelt.

Daneben gibt es ein weitverbreitetes Interesse an und eine verstärkte Nutzung von Normen wie den Musteranforderungen oder Moreq der Europäischen Kommission und den funktionalen Anforderungen, die vom britischen Public Record Office entwickelt wurden.

Auch wird immer mehr unterschieden zwischen der elektronischen Schriftgutverwaltung (EDM) und der zusätzlichen Funktionalität der elektronischen Archivgutverwaltung (ERM). Ein Teilnehmer stellte das Konzept einer Strategie für elektronische Schrift- und Archivgutverwaltung vor, die die Archivgutverwaltung durch Workflow, Formularbearbeitung und Webdienste ergänzt.

Die Arbeitsgruppen kamen zu folgenden Ergebnissen:

- Organisationen verwenden elektronische Archivierungssysteme, um ihre Arbeit zu unterstützen, und nicht, um historische Aufzeichnungen anzulegen;
- die Rolle einer externen Regelung ist ebenfalls wichtig, insbesondere Gesetze oder andere gesetzliche Bestimmungen im Bereich der Information;
- der öffentliche Sektor ist bei der Verwaltung von elektronischen Archiven generell fortschrittlicher als der Privatsektor, außer in besonderen Bereichen, z. B. im Arzneimittelbereich, wo es einen langfristigen Bedarf an strengen externen Vorschriften gibt;
- der öffentliche Sektor kann und sollte seine Position deutlich machen, um bewährte Verfahren darzulegen und die Entwicklung im Privatsektor zu fördern;
- elektronische Archive sind ein Aushängeschild für jede öffentliche oder private Organisation und sollten als solches behandelt werden;
- die Entwicklung externer, neutraler Normen wie Moreq sind von erheblichem Wert, um den Bedarf festzustellen, sowie beim Beschaffungsprozess.
- notwendig ist eine weitere und anhaltende Ausbildung der Mitarbeiter im Bereich Verwaltung von elektronischen Archiven.

Rapport sur le Forum DLM 2002, Barcelone: sessions parallèles 3 et 4

George Mackenzie

Les sessions parallèles 3 et 4 ont essentiellement porté sur la gestion et la livraison de contenu ainsi que sur l'organisation des documents et des archives.

Elles ont mis en évidence la production et l'utilisation très répandues des documents électroniques en Europe, mais elles ont également révélé qu'une attention moins grande était accordée à leur gestion. Le Forum DLM doit notamment viser à combler ce fossé. Les archives véritablement numériques sont encore rares, et le secteur privé accuse un certain retard par rapport au secteur public en ce qui concerne la gestion de ce type de documents.

L'importance des métadonnées ou de l'information contextuelle a également constitué l'un des fils conducteurs des deux sessions. Les questions les plus fondamentales sont probablement les suivantes: Quelles sont les caractéristiques significatives des objets numériques? Que devons-nous veiller à conserver dans les métadonnées, dans nos systèmes d'archivage? Des projets de recherche abordent, fût-ce partiellement, ces diverses problématiques.

L'intérêt croissant et l'utilisation de plus en plus systématique de normes telles que les spécifications types MoReq produites par la Commission européenne et les exigences fonctionnelles développées par le PRO britannique (Public Record Office) ont également été soulignés.

On note, par ailleurs, une meilleure perception de la distinction entre la gestion des documents électroniques et la fonctionnalité supplémentaire que constitue la gestion des archives électroniques. L'un des orateurs a introduit le concept de plate-forme commune de gestion de documents et d'archives électroniques, qui ajoute à la gestion des archives des aspects tels que le déroulement des opérations (workflow), le traitement des formats et des images et les services basés sur le web.

Les sessions ont abouti aux conclusions suivantes:

— les organisations mettent en place des systèmes de documents électroniques dans la perspective d'une plus grande efficacité plutôt que d'une relation historique;

— la réglementation externe joue, elle aussi, un rôle important (législations et autres dispositions réglementaires relatives à l'information);

— le secteur public est, de manière générale, plus avancé que le secteur privé dans la gestion des documents électroniques, hormis dans certains secteurs spécifiques tels que l'industrie pharmaceutique, où existe un besoin à long terme des entreprises ou une réglementation externe rigoureuse;

— le secteur public peut — et devrait — tirer avantage de sa position pour montrer l'exemple de bonnes pratiques, et en favoriser le développement dans le secteur privé;

— les documents électroniques sont un actif pour l'organisation concernée, qu'elle soit publique ou privée, et devraient être gérés comme tels;

— l'élaboration de normes externes et neutres, telles que les spécifications types MoReq, constitue une précieuse contribution à la détermination des besoins et au processus de passation des marchés;

— une formation continue du personnel participant à la gestion des documents électroniques est impérative.

Thekla Kluttig

Thekla Kluttig was born in 1968 and currently works for the Archives Department of the Ministry of Interior of Saxony, in Dresden, Germany. After studying history and political sciences in Bonn and Hamburg (1987–92) she was awarded a PhD with her dissertation on the main party of the former GDR. From 1995 to 1997 she completed professional training as an archivist at the Hauptstaatsarchiv Dresden, the Archival Institute in Marburg and the National Archives Koblenz. After graduating in archival studies, she began her professional career at the Hauptstaatsarchiv Dresden, Saxony. Since 2000 she is working for the Ministry of Interior of Saxony, Department 44 (Archives), where she is among other things responsible for appraisal and electronic records.
Thekla Kluttig participates in committes and working groups concerning electronic records management and information technology in public administrations in Germany. Since 2002 she gives professional training courses for archivists on electronic records at the Archival Institute in Marburg.

Report on the DLM-Forum 2002: parallel sessions 1 and 6

Thekla Kluttig

Information technology has been introduced into European local and national authorities and will be increasingly used. All archivists who are in charge of a location that is required to make its collections accessible and who are responsible for the appraisal of its documents must — if they feel bound by the archivist ethos — be involved in archiving electronic records. In the terminology of the Archival Law of Saxony, archiving consists of capturing, appraising, recording, preserving and describing, rendering usable and evaluating archival records. The question of which organisational and technical infrastructure is suitable for the recording and long-term preservation of electronic records has only been partially resolved. Nevertheless, we can and should begin with the capture and appraisal of documents. In addition, we should be actively involved when 'our' administration is planning the introduction of IT procedures, e.g. procedures for implementing IT-supported file processing. In order to avoid high costs later, it is absolutely necessary for disposal and archiving functions to be taken into account at the early stages of planning and the introduction of file processing. The inclusion of professional archivist knowledge is also in the administration's interest. The DLM-Forum offers an important opportunity for the communication and coordination of various approaches, as the sessions introducing in the following pages clearly demonstrated.

Session I focused on the 'Memory of the information age: preservation, migration and long-term availability'. It was chaired by Luciana Duranti and Francisco Barbedo and consisted of six presentations which approached this topic from various perspectives, ranging from strategic approaches to reports on archives in practice.

The session was introduced by Richard Blake, Project Manager in the Electronic Records Management Development Unit in the Public Records Office, who presented the strategic considerations accompanying the implementation of an 'Electronic document and records management' (EDRM) in the United Kingdom up to 2004. Although it is costly, the implementation of an EDRM offers medium-term advantages such as continual access to reliable, authentic information and the integration of various business information. These advantages can only be obtained when information is preserved in the long term and is available at all times for business purposes. Richard Blake referred to the area of conflict between records and technology on the one hand and users and the organisation on the other. Records and technology require stability, while the communication of information and user needs require flexibility and further development. Thus an appropriate balance needs to be achieved between stability and change. The keyword is sustainability. Ensuring sustainability is a central task for administration and archives.

In the printed programme and in the summary of the abstracts you have been given is a presentation by Kevin Ashley of the University of London Computer Centre, on developing practical procedures for data preservation following the example of the Arts and Humanities Data Service. Unfortunately the project results could not yet be presented. Instead, Kevin Ashley briefly described the activities of the National Digital Archive of Datasets and referred to some fundamental problems in the development and use of procedures for archiving electronic records. One example is that it is considerably easier to implement procedures within organisations than from outside. The speaker regretted that, while much has been published on archiving strategies and principles, concrete, usable reports from practical experience were rare. I can only agree with this observation — and find it even more regrettable that the presentation originally announced could not yet be given.

Frank Brady of the Secretariat-General of the European Commission offered an insight into a work in progress. He reported on the largest IT project ever since the introduction of IT at the Commission, the modernisation of records management. Important elements are support at the decision making level, central access to information from all divisions of the organisation, the inclusion of all types of documents and IT support of the whole life cycle of documents. The project's overall organisational criteria were laid down at the beginning of 2002. Frank Brady's paper illustrated very impressively how many factors must be taken into consideration in the implementation of a project of this size.

In the regional administration in Saxony, I am repeatedly confronted with the low status accorded to document management. So I would like to refer particularly to the following statements by Brady: information is obtained from documents, results are stored in documents. Information is the administration's most valuable property — and not only the administration's! Thus a corresponding value should be given to high-performance records management. This is intended as a tip for those who consider the management and archiving of documents as dusty and totally outdated. Records management is knowledge management!

The presentations which have been introduced so far dealt with our theme — the memory of the information age — primarily from the perspective of the users and managers of records. In her fascinating lecture on 'Practical experiences of the digital preservation test bed', Jacqueline Slats introduced a project of the Dutch Government which is focused on the possibilities of long-term archiving. Using predefined tasks, three procedures for preserving collections were tested in parallel: migration, emulation and XML. Experiments were carried out using different types of records, e.g. text documents, emails and databases. In each case, efficiency, limits and costs were tested and compared as well as potential uses. Jacqueline Slats was already able to present interesting results. Like the session participants, I hope that the final result will also be published in an English language version.

The British Office for National Statistics has not opted for migration, emulation or XML: For the long-term preservation of survey data from the 2001 census good old microfilm was chosen — like the Sleeping Beauty woken with a kiss. The project manager, Amanda Blunden reported on the implementation of the census from an organisational technical perspective in her clear and to-the-point lecture. Twenty different forms with a circulation of approximately 30 million copies and over 300 million pages had to be produced, protected and digitally processed. The 300 million images were copied onto 16 mm microfilm with the assistance of 12 digital archive writers from Kodak. The Office for National Statistics decided to use microfilm, as it is easy to look at and may be scanned when the results are made available to the public — in 2012, in accordance with British law. Thus the original forms may be destroyed after digitalisation and simultaneous microfilming.

Best practices and solutions — our colleagues from the Danish National Archive in Copenhagen supplied a large number of practical solutions and experiences. The National Archive has been successfully incorporating and protecting electronic records for over 30 years. Birgit Hansen and Jan Danielsen's lecture on the preservation of the digital cultural heritage as a question of IT protection and quality assurance was an extremely stimulating report from archival practice. The Danish National Archive has developed a number of complementary organisational and technical measures. Through regular evaluation and necessary updating they ensure the preservation of the digital cultural heritage. This includes rules which must be observed in the implementation or fundamental revision of IT procedures, requirements in the transmission of electronic records (e.g. to data documentation and data holders) as well as regular testing and monitoring of data. The National Archive is operating within very favourable legal conditions. The presentation of one of the National Archive's test programmes developed to facilitate the automatic control of the data to be archived was very interesting.

Which results in Session I are particularly worthy of note? The presentations were very varied and there are several starting points for further reflection. Our topic was the 'Memory of the information age'. The concept of memory describes the capacity to store information in such a way that it may be retrieved and reproduced. Does the memory of the information age need to reflect human memory more closely? We have a short-term and a long-term memory and not all the information stored in our short-term memory is transmitted to the long-term memory. Meaningless and unimportant information tends to be forgotten. What people's brains do for them — not always to their satisfaction! — must be accomplished by document managers, archivists and information specialists for society. I would like to say another few words on this matter after my report on the sixth session.

Session VI dealt intensively with questions in the area of information technology. Competently chaired by Yannis Ioannidis and Tom Quinlan, the session's five speakers dealt with the 'Capture and transformation of information: automatic technologies for indexing and introducing retrieval

solutions'. As an archivist, my knowledge in this area is that of an amateur. But my task was made easier by the accessibility of the presentations.

The first speaker was Bernhard Lindgens of the German Federal Treasury Department with a concrete case study. His lecture was a practical report on the fight against tax fraud in electronic transactions, using modern technology. Several suppliers of products and services on the Internet are not registered for tax, which leads to tax avoidance and distortions in competition. Conventional search strategies and engines proved unequal to the task. So the Federal Treasury Department is using intelligent software which is knowledge-based for the periodic and automated registration of enterprises supplying over the Internet which are liable to German taxation and for automating other steps necessary for countering tax fraud. For further details, see Industry White Paper 1: Intelligent Capture, Indexing and Auto-categorisation of Information.

Hiromichi Fujisawa of the Central Research Laboratory, Hitachi Ltd gave a paper with an interesting overview of the research and developments in document recognition and retrieval technologies over the last 10 years. How can errors caused by OCR be reduced, if the text generated using OCR is used for retrieval? How can errors be discovered? How can harmless errors be tolerated? These are central questions. Hiromichi Fujisawa focused on three different solutions which could lead to a greater integration of the processes of document recognition and retrieval. For the archivists among you let it be said that the approaches presented do not deal with handwritten documents.

Marcus Junker and Andreas Dengel of the German Research Centre for Artificial Intelligence are also looking for innovative solutions for the automatic categorisation and indexing of electronic documents. Markus Junker gave an introduction to the history of machine learning processes in document categorisation which have been around for about 10 years. The use of machine learning processes for the categorisation of documents was clearly explained. Rule learning procedures, statistical learning procedures and neuronal networks may be used for automation. An exciting question is whether and how these approaches may be used in the area of document management, libraries and archives. In German archives, description is traditionally carried out not at the level of documents but on files so that the approaches presented cannot be used without adaptation. Moreover, their use assumes that documents can be scanned in and transformed using OCR.

Matthias Meusch of the Main State Archive of North Rhine Westphalia in Düsseldorf gave a paper which was very relevant to archival practice. The subject was the DFG (German Research Community) project on the development of tools for the retroconversion of archivists' finding aids. Techniques and strategies for automated recording, categorisation and indexing of information are of great interest for archives. Only with their help can old finding aids, which are often to a large extent in an unsatisfactory form, be retroconverted and so made generally available. Several state archives and the Main State Archive of North Rhine Westphalia as well as the Federal Archive and the Economic Archive of Westphalia are involved in the project. Matthias Meusch introduced in detail the informational/technical procedures. Particularly worthy of note is XML which has already been developed. It is a 'document type definition' which may be used as an exchange format by different archives. After developing the necessary software — the retroconversion tool — approximately 300 finding aids in typewritten or word processing documents were converted. We are very curious to hear how the results will be evaluated.

Ineke Deserno of the COI and Montserrat Canela of the UNHCR reported on the successful introduction of electronic document and records management systems in the UN High Commission for Refugees and the International Olympic Committee. The extensive and indispensable use of e-mail in both organisations was a catalyst for the project. In their paper, based on practical experience, both speakers presented the aims and requirements of EDRM, the steps involved in introducing the system and the organisational and technical challenges. They particularly emphasised the necessity of support from management level, the need for good file plans and ongoing staff training.

Since the sixth speaker, Gennadi Zalaev, unfortunately could not take part in the session, there can be no report on his paper. I refer you here to the abstract in your documentation.

I would like to conclude my report with a few brief observations. DLM — the acronym has been re-interpreted as 'Document life-cycle management'. In the interests of society we have to take

over this management. We: by this I mean the records managers, archivists and information technicians of the private and public sectors together.

In my view it is very important to impress the necessity and usefulness of good records management on those who are politically responsible. Otherwise the word 'e-government' is nothing more than an empty shell and laws on freedom of information will serve no purpose. We need professional document managers. I heard that the use and implementation of metadata standards was the subject of lively discussion in Workshop 1. MoReq and the new ISO Records Management Standard were recognised as valuable instruments for the successful implementation of EDRMS. Practical solutions were introduced based on these standards. These also constitute future tasks for records managers!

One of an archivist's original tasks is that of appraisal. Archivists are experts in structure, form and content — there is no question of that. But the most special qualification for archivists, their most responsible task is that of appraisal: deciding which records deserve to be stored permanently. We archivists will have to thoroughly address the question of appraisal and its transparency when dealing with the long-term archiving of electronic data.

On the other hand, as an archivist I would not wish to have to decide which format is up to date now. That is a specific matter for information specialists. By developing suitable IT solutions, they make an indispensable contribution to the improvement of access and the preservation of electronic information.

In both of the sessions which I had the honour to introduce, practical solutions to pressing problems were offered. I would particularly like to recall the papers given by Jacqueline Slats, Birgit Hansen and Jan Danielsen, Matthias Meusch, Ineke Deserno and Montserrat Canela. They demonstrated how successful a professional and close cooperation between records managers, archivists and information technicians can be. Professionalism, cooperation and work sharing — these were central issues for this DLM-Forum and I am sure that the same will be true of the next forum.

Informe sobre el Foro DLM 2002 de Barcelona: sesiones paralelas 1 y 6

Thekla Kluttig

Thekla Kluttig, Jefa de servicio del Ministerio de Interior del Estado federado de Sajonia (Alemania), es ponente de las comunicaciones que se presentaron en las 1 y 6 sesiones paralelas. La sesión paralela 1 tuvo como título «La memoria de la era de la información: conservación, migración y disponibilidad a largo plazo» y constaba de seis comunicaciones, que abordaron estos temas bajo distintos ángulos y cubrieron un abanico que va desde los enfoques estratégicos hasta la presentación de algunas prácticas archivísticas. La sesión paralela 6, por su parte, se dedicó a cuestiones de contenido más científico, agrupadas bajo el título general «Conseguir y transformar la información: tecnologías automáticas de indexación y aplicación de soluciones de recuperación».

El ponente insistió como conclusión en la idea de que la sesión paralela 1 proporcionaba distintos puntos de partida para proseguir la reflexión. Por lo que respecta a la *memoria de la era de la información,* ¿convendría quizás interesarse más por la memoria humana? No toda la información retenida en la memoria a corto plazo pasa a la memoria a largo plazo. Mucha información inútil y desprovista de sentido se olvida. Incumbe a los gestores de documentos, a los archiveros y a los informáticos realizar por cuenta de la sociedad lo que el cerebro

humano realiza en cada individuo. Para ser más concretos: las comunicaciones de las dos sesiones presentadas, que se complementan mutuamente, pusieron de manifiesto que la gestión del ciclo de vida de los documentos sólo pueden realizarla conjuntamente los gestores de documentos, los archiveros y los informáticos —del sector público y del sector privado— en interés de la colectividad.

Thekla Kluttig hizo hincapié en el hecho de que los responsables políticos deberían tomar más conciencia de la necesidad y el interés de una buena gestión de los documentos, pues, de lo contrario, la administración electrónica no será más que una cáscara vacía y las distintas leyes sobre libertad de información serán papel mojado. La administración necesita gestores de documentos profesionales. El ponente remitió a estos efectos al taller 1, que permitió abordar intensivamente la cuestión de la utilización y la aplicación de normas sobre metadatos. Las especificaciones modelo MoReq y la nueva norma de gestión de documentos ISO se consideran a estos efectos instrumentos preciosos para lograr una correcta aplicación de los sistemas de gestión de documentos y archivos electrónicos.

En conclusión, el ponente señala que una de las primeras misiones de los archiveros es evaluar el valor archivístico de los documentos. Los archiveros son también expertos por lo que respecta a la estructura, el contexto y el contenido. Pero la cualificación especial del archivero, su responsabilidad primordial, es decidir qué documentos deben conservarse a largo plazo. El ponente exhorta pues a los archiveros, por lo que se refiere al archivo a largo plazo de los documentos electrónicos, a que se preocupen activamente por las cuestiones de su evaluación y su transparencia. Por el contrario, no incumbe a los archiveros decidir acerca del formato más actual. Esta cuestión incumbe específicamente a los informáticos que, al desarrollar soluciones informáticas adecuadas, contribuyen de forma indispensable a la mejora de la accesibilidad y de la conservación de la información digital.

Berichterstattung auf dem DLM-Forum, Barcelona 2002: Parallelsitzungen 1 und 6

Thekla Kluttig

Die Parallelsitzung 1 befasste sich mit dem *Gedächtnis des Informationszeitalters. Bewahrung, Migration und Langzeit-Verfügbarkeit* und umfasste sechs Beiträge, die sich dem Thema aus verschiedenen Perspektiven näherten und einen Bogen schlugen von strategischen Ansätzen bis hin zur Darstellung archivischer Praxis. Die Parallelsitzung 6 widmete sich dagegen verstärkt informationswissenschaftlichen Fragestellungen unter dem Rahmenthema *Erfassung und Umwandlung von Informationen: Automatische Technologien zur Indexierung und Einführung von Retrieval-Lösungen*.

Die Berichterstatterin betonte zusammenfassend, dass die Sektion I verschiedene Ansatzpunkte für weiterführende Überlegungen bot. Das *Gedächtnis des Informationszeitalters* – vielleicht muss es sich zunehmend am Gedächtnis des Menschen orientieren? Nicht alle Informationen, die im Kurzzeitgedächtnis gespeichert werden, werden auch in das Langzeitgedächtnis überführt. Sinnlose und unwichtige Informationen werden eher vergessen. Was sein Gehirn für den einzelnen Menschen erledigt, müssen Schriftgutverwalter, Archivare und Informatiker für die Gesellschaft übernehmen. Genauer: Die sich ergänzenden Beiträge der zwei vorgestellten Sitzungen haben gezeigt, dass das *Document Lifecycle Management (DLM)* von Schriftgutverwaltern, Archivaren und Informatikern – des öffentlichen und des privaten Sektors – im Interesse der Gesellschaft nur gemeinsam übernommen werden kann.

Thekla Kluttig betonte, dass die Notwendigkeit und der Nutzen einer guten Schriftgutverwaltung stärker im Bewusstsein der politisch Verantwortlichen verankert werden muss – sonst bleibt *e*-Government nur eine leere Worthülse und Informationsfreiheitsgesetze laufen ins Leere. Die Verwaltung braucht professionelle Schriftgutverwalter. Dabei verwies die Berichterstatterin auf den Workshop 1, in dem intensiv über die Nutzung und Umsetzung von Metadaten-Standards diskutiert wurde. MoReq und der neue ISO Records Management Standard wurden dort als wertvolle Instrumente für die erfolgreiche Einführung von EDRMS anerkannt.

Die Berichterstatterin stellt abschließend fest, dass eine originäre Aufgabe der Archivare die Bewertung von Unterlagen auf ihre Archivwürdigkeit hin ist. Archivare sind auch Experten für Struktur, Kontext und Inhalte. Aber die ganz besondere Qualifikation der Archivare, ihre verantwortungsvollste Aufgabe ist die Bewertung: Die Entscheidung darüber, welche Aufzeichnungen es wert sind, dauerhaft aufbewahrt zu werden. Die Berichterstatterin fordert daher die Archivare auf, sich im Rahmen der Langzeitarchivierung elektronischer Unterlagen intensiv mit der Frage der Bewertung und ihrer Transparenz zu beschäftigen. Dagegen sollten Archivare nicht darüber befinden müssen, welches Format gerade aktuell ist. Denn dies ist eine spezifische Qualifikation von Informatikern. Sie leisten durch die Entwicklung geeigneter IT-Lösungen einen unverzichtbaren Beitrag für die Verbesserung der Zugänglichkeit und der Erhaltung elektronischer Informationen.

Rapport sur le Forum DLM 2002, Barcelone: sessions parallèles 1 et 6

Thekla Kluttig

La session parallèle 1 avait pour titre «La mémoire de l'âge de l'information — Conservation, migration et disponibilité à long terme» et comportait six communications qui ont approché cette thématique sous divers angles et ont couvert un éventail allant des approches stratégiques à la présentation de certaines pratiques archivistiques. La session parallèle 6 était quant à elle consacrée à des questions au contenu plus scientifique, regroupées sous le titre général «Saisie et conversion des informations — Technologies d'indexation automatique et solutions pour optimiser la recherche documentaire».

Le rapporteur a insisté en conclusion sur l'idée que la session parallèle 1 fournissait divers points de départ pour poursuivre la réflexion. Au sujet de la *mémoire de l'âge de l'information,* peut-être convient-il de s'intéresser davantage à la mémoire humaine? Les informations qui sont mémorisées dans la mémoire à court terme ne passent pas toutes dans la mémoire à long terme. De nombreuses informations dépourvues de sens et inutiles sont oubliées. Il incombe aux gestionnaires de documents, aux archivistes et aux informaticiens d'assurer pour le compte de la société ce que le cerveau humain assume au niveau de chaque individu. Pour être plus précis encore: les communications des deux sessions présentées, qui se complètent mutuellement, ont montré que la gestion du cycle de vie des documents ne peut être prise en charge que conjointement par les gestionnaires de documents, les archivistes et les informaticiens — du secteur public et du secteur privé — dans l'intérêt de la collectivité.

Thekla Kluttig a insisté sur le fait que les responsables politiques devaient prendre davantage conscience de la nécessité et de l'intérêt d'une bonne gestion des documents, faute de quoi l'administration électronique ne sera qu'une coquille vide et les différentes lois sur la liberté d'information resteront lettre morte. L'administration a besoin de gestionnaires de docu-

ments professionnels. Le rapporteur a renvoyé en l'espèce à l'atelier I, qui a permis d'aborder de façon intensive l'utilisation et l'application de normes de métadonnées. Les spécifications types MoReq et la nouvelle norme de gestion de documents ISO ont été saluées à cette occasion comme des instruments précieux pour réussir l'introduction des systèmes de gestion de documents et d'archives électroniques.

En conclusion, le rapporteur constate que l'une des missions premières des archivistes est d'évaluer la valeur archivistique des documents. Les archivistes sont également des experts pour ce qui concerne la structure, le contexte et les contenus. Mais la qualification tout à fait particulière de l'archiviste, sa responsabilité primordiale, est de décider quels documents doivent être conservés à long terme. Le rapporteur exhorte donc les archivistes, en ce qui concerne l'archivage à long terme des documents électroniques, à se préoccuper activement des questions de leur évaluation et de leur transparence. En revanche, il n'appartient pas aux archivistes de décider du format le plus actuel. Cette question est du ressort spécifique des informaticiens, qui, en développant des solutions informatiques appropriées, fournissent une contribution indispensable à l'amélioration de l'accessibilité et de la conservation des informations numériques.

Report on the DLM-Forum 2002: parallel sessions 2 and 5

Catherine Dhérent

Catherine Dhérent

Catherine Dhérent graduated at the École nationale des Chartes, Paris in 1981. She obtained a doctorate from the University of Paris-I-Panthéon-Sorbonne, specialising in medieval history in 1993.
She is general curator of national heritage at the Directorate-Generale of the Archives de France.
Previously she was deputy curator at the Archives départementales du Nord from 1981 to 1987, and director of the Archives départementales du Pas-de-Calais from 1987 to 1996, responsible for planning a national archives centre in Rheims from 1996 to 1999, policy officer for the Director of the Archives de France with responsibility for IT matters from 1999 to 2002.
Since January 2002 she is responsible for the technological innovation and standardisation department of the Archives de France Directorate.

I will attempt in this report to organise the points and proposals made in sessions 2 and 5 on using information and improving access to knowledge, even though I found the exercise somewhat risky — in particular since the prevailing theme is not always easy to identify even within a single session.

A recurrent issue in both of these sessions was providing access to increasingly relevant and personalised information: a task which involves a variety of professionals — computer scientists, document creators, online publishers, librarians and a new generation of archivists — all with increasingly adapted and personalised training. New technologies, which can connect anyone to a borderless universe, have also encouraged each individual to expect resources that match his or her specific needs and level of understanding of a topic, whatever that topic might be.

With the aim of continuously improving their service to the general public, authorities such as the Danish Government — mentioned in the example presented by Hermann Weidemann — are showing great interest for new and more efficient digital products that can provide accurate responses to any query. The user is now central to all the research carried out jointly by the public and private sectors. Attempts are being made to develop a new generation of digital information objects based on artificial intelligence technology in order to make things easier for users and facilitate their interaction with the public authorities, following a dynamic approach tailored to each individual case. Administrative checks — which used to be carried out at random on samples — are now executed automatically, with the result that each individual is treated in the same way.

The 'directory services' now being made available reflect a similar concern to improve user access to information within the framework of a wide and varied range of services. Directory services rely on specialist databases and offer very rich content going far beyond a mere alphabetical list of persons and contact details. They were presented by Felix von Bredow: a single administration tool allows registered users — whether internal or external, known or anonymous — to gain access to information stored in various departments of an institution with a single password. It also allows data to be secured and lowers the cost of managing it and entering it into the systems. Overlying the directories implemented with standard tools, metadirectories can be established to connect and synchronise them. The distribution of information among several locations enables it to be transferred very quickly in a manner that is totally transparent to the user.

New generation tools automatically present data according to the queries and profiles of web surfers, allowing each of them to put together their own knowledge clusters. It is therefore becoming vital to preserve, in the longer term, the memory of these pools of knowledge that are building up every day. This preservation must support the dynamic creation of new digital objects. The purpose now is not merely to safeguard data, which is something we have been able to do for several decades. It is also to safeguard information — i.e. multiple combinations of the data — and, more importantly still, knowledge.

Such is the aim of the project that Ken Thibodeau's team have been working on at NARA since 1998. It has yielded considerable advances in techniques, along with the conclusion that the achievements can be applied in many other areas. NARA and its partners (in particular defence authorities, universities and the InterPARES group) are developing an information management architecture that is independent of the technology used. The purpose is to meet the general public's demand for information to be delivered across a wide variety of computer platforms. The new approach and the establishment of an 'architecture for persistent archives' are more comprehensive and provide for a coherent strategy.

NARA's Electronic Records Archives Programme aims not only to deliver all types of electronic document within their cognitive context but also, eventually, to manage their integration and preservation using Internet technology.

This programme for the preservation of access to knowledge has convinced the US Executive Office of the President of the highly important role of archive services, resulting in very significant financing for NARA (an 800 % increase in the space of a year).

The administrative production that archive services are confronted with is one of the most complex there is. It is therefore far from certain that the Open Archives Initiative can be used in this context, in spite of what its name would appear to indicate, for 'archive' here means any digital repository of digital information. The Open Archives Initiative was launched in the United States in 1999; it was first implemented for electronic publications (in fields such as physics, psychology and computer science), then expanded to include libraries and virtual museums. Its aim is to open up access to increasingly vast digital resources, at little cost and for a broad audience, thanks to an interoperability protocol for metadata harvesting. This protocol would be easy to implement and would also provide for usage rights administration. The Open Archives Forum, a project submitted by Donatella Castelli under the fifth framework programme, has recently resulted in the establishment of a European web site that is to disseminate information on the Open Archives Initiative and provide a database on the institutions that have implemented it, a glossary, a newsletter on issues relating to intellectual property rights and a compilation of the software already available or being developed in this area. The Forum will help to identify the professional communities for which open archives could be relevant and useful.

Among the data which give meaning to other data but also modify them, there is the dual (public/private) key signature. This key guarantees, to a certain extent, the integrity of content. Andrew Waugh is a former participant in the Australian Victorian Electronic Record Strategy project, and is now responsible for developing the project for the Victoria government. He outlined the constraints and weaknesses of this authentication tool, that is used to detect the slightest modification in electronic documents. We cannot, at this point in time, depend too much on digital signatures to identify signers for nothing can guarantee the confidentiality of the private key. Only biometric techniques will provide an effective means of overcoming this disadvantage. Verification methods have remained pragmatic and still involve comparing several documents signed with the same public key. While it is possible to detect whether a document has been modified, it is not possible to determine the part that has been changed. Furthermore, any change to a digitally signed document, however minor (such as the addition of metadata or the correction of a spelling mistake), breaks the signature. The document thus loses its protection, opening the way to further undetectable modifications.

These shortcomings, and the fact that signature algorithms are constantly evolving, could also make it necessary to resign documents at regular intervals. The solution might be to use layers of signatures, or to use signatures to lock only specific parts of a document. To ensure long-term access to the signed document, it is also crucial to have an independent validation system for it can be impossible to verify a signature in a digital environment other than that in which the signature was created. In this as in other areas of the digital universe, the systems used must be totally independent from all hardware and software platforms.

All these presentations on the uses of information highlighted the need to move away from proprietary tools, ensure interoperability and further increase standardisation.

In session 5, Wendy Duff demonstrated how difficult it still is today — despite the above attempts to improve long-term dissemination of information among the general public — for ordinary people and professional researchers alike to find what they are looking for using the search tools supplied by archive services in a web environment.

Networks have raised many hopes among researchers of all kinds, for they erase geographical and temporal distances and hold out the promise of delivering information and even primary documents straight to the desktop. But these researchers often end up lost and disappointed. Web surfers frequently lack the specific skills required to understand the complexity of archival organisation and assess the authenticity of what they are getting on their screen. They are missing the mediation of archivists, who remain irreplaceable and have themselves traditionally developed their knowledge of their records and search tools by responding to and assisting the general public.

But preservation also gives archivists a better understanding of the needs of the document originators, another category of user. They must reconcile the search and access needs of these very different groups.

Academic researchers began looking into the use and users of archives only some 20 years ago, whereas the use and users of libraries have been the focus of research since the 1940s. How and why do people look for information in an archival environment? Searching can be formal or informal, and will usually depend on the nature of the records; it can be conducted by means of bibliographical references or by calling on archivists in the case of historians, or can be based on a name, a date and geographical data in the case of genealogists … Archivists must devise systems which will accommodate all these search modes and which will automatically ask web surfers to clarify and detail queries that are too imprecise. Web sites must cater for users with very diverse levels of expertise and knowledge. They must also dispense them as much as possible from having to resort to those responsible for the records and collections.

In the opinion of most users, search tools are never accurate enough. This is even more the case in a web environment, where there are no archivists to call upon except via e-mail.

It must also be said that making search tools available on the web increases public demand for copies as well as for information. Archive services are constantly having to address these new needs.

The web can also be a means for archivists to inform users about the main principles and problems of archiving and about the concepts of authenticity and reliability of a given document, and to teach users to exercise their judgment.

To improve the understanding of these issues and try to resolve them, Wendy Duff suggested the idea of creating a network for research on archive users. This would require close collaboration between archive professionals and teachers in the field of archival studies. Why not envisage a potentially very beneficial link with a project of the DLM network? The proposal was made during the session and I submit it to you for consideration. The content would be a single form, and the findings could be shared by all archive services with a view to identifying the major trends in user demand and practices. This network could be in touch with training services.

It is vital for archive professionals, and all those responsible for documents, to review — and upgrade — their practices with respect not only to users but also to document management. Archivists must become involved in the information processing chain at an earlier stage. The recently adopted records management standard has highlighted the need to establish a continuum in the document life cycle and ensure perfect control over documents — a need that is all the more important for digital objects. This context has fuelled demand for training on the part of document management teams, whose abilities are constantly being challenged. Thijs Laeven outlined what these abilities should be (a blend of skills, attitudes and knowledge, professional or not, combined with standards, ambitions and a collective or individual self-image), and presented evaluation matrixes of the resulting profiles. These matrixes have been made available on a web site launched in January 2002 on the initiative of the Dutch Home Affairs Ministry and a Dutch cultural association responsible for archiving. They allow all individuals to assess their ongoing training needs and the skills they need to acquire in order to make the best possible contribution to the strategy of their institution. The web site could also encourage the development of standards for document processing skills that would be valid and necessary at European level.

The web site is to have close relations with training provision programmes. It has already forged links with the programme described below. The idea of developing specific courses to teach electronic document management had been envisaged at a meeting of European experts in June 1997, at a time when seminars on these issues were already being organised at the Marburg Archives School and the Netherlands Archives School. The course of the Netherlands Archives School was eventually selected as basic model for the E-TERM project, which was awarded Community financing under the Leonardo programme up to December 2001.

This project, which is now operational and is run by seven universities and archives schools from the United Kingdom, Italy, Germany, Finland, Portugal and the Netherlands, was presented by

Peter Horsman. The programme supplies a curriculum and teaching materials (theory-based texts, case studies, multilingual terminology and a web site). The curriculum comprises five modules that can be tailored to local needs. These modules are designed to bridge the gap which all too often exists between theory and practice. The learners prompt the teachers to find, by means of problem analysis, relevant responses to their questions that are adapted to their particular circumstances. The courses can be delivered using traditional classroom methods, but in a fully digital environment which also facilitates international cooperation. This cooperation is all the more necessary since it is extremely difficult to provide an optimised response to the personalised queries of the learners.

The programme could possibly lead to an even more ambitious virtual European centre for advanced training in archival studies. It is in any case expected to forge links with other distance archival training partners and initiatives, such as the one being set up by the international association of French-speaking countries (*Association internationale de la Francophonie*, or AIF).

This type of programme requires very specific teaching skills. As pointed out by Mr Dekeyser and Mr Boccacio on behalf of the business community and by Ms Danbury on behalf of universities, the current situation is not yet satisfactory and is not yet quite adapted to the virtual sphere, even though this form of distance training is increasingly popular. In particular, factors such as the speed of responses, the quality of results, the efficiency and simplicity of establishing standardised solutions, the ability to convert e-learning into mobile learning, etc., must be taken into account. Furthermore, it is also necessary to develop learning management systems that can help learners and teachers who do not have access to the same instruments to develop their own content, tools and procedures on the basis of online content. The Italian regional public administration school has set up such a project, whose results indicate that it is responding very well to learner expectations.

However, these services are very cumbersome to establish and maintain on a daily basis in as interactive a manner as possible. A number of archivists consider that updating them with appropriate and sufficiently rich data is too time-consuming.

That is why some Spanish archivists and computer scientists are working on an evaluation and training web site based on artificial intelligence techniques. The system presented by Paltin Sturdza enables users to determine, by means of a diagnostics tool, what, when and how to learn, without requiring assistance from a teacher. These significant developments are based on language processing methods which provide for the dynamic modelling of knowledge in a way which I believe is quite similar to the implementation of ontologies. One of the objectives is also to explain while doing, and in particular to explain 'why', 'how' and 'why not'. Thanks to the knowledge base, the user can thus determine the precise way to proceed in response to a query.

There still remains much to do and a lot of difficult ground to cover before we can deliver systems which would automatically establish rules without destroying the richness of a knowledge that is highly fluctuating and constantly being renewed.

Nonetheless, thanks to the ongoing cooperation between professions that are very diverse yet are all concerned, to some degree and within a very broad community, with document production, maintenance and long-term accessibility, we can undoubtedly be optimistic about our ability to upgrade our practices.

It is also comforting to see, in exhibitions such as the one organised on the occasion of this event, that the business community is gradually responding to our changing needs and coming up with solutions.

Nevertheless, mediation and face-to-face contact do have specific advantages, even within a digital environment that is more dynamic and more finely tuned to individual requirements. Most participants agreed that machines cannot — and should not — resolve all problems. For at least some time to come, the machine's role will remain to supplement ever more reactive and intelligent direct assistance by human beings, provided of course direct assistance is feasible.

Informe sobre el Foro DLM 2002 de Barcelona: sesiones paralelas 2 y 5

Catherine Dhérent

En esta acta de las sesiones paralelas 2 y 5 consagradas a la utilización de la información y a la mejora del acceso al conocimiento, quisimos ligar las distintas intervenciones y hacer hincapié en la necesidad para los archiveros y gestores de documentos de permitir el acceso de todos los tipos de lectores a una información cada vez más pertinente y personalizada.

Berichterstattung auf dem DLM-Forum, Barcelona 2002: Parallelsitzungen 2 und 5

Catherine Dhérent

In diesem Bericht über die Sitzungen 2 und 5, die parallel durchgeführt wurden und der Nutzung von Informationen sowie der Verbesserung des Zugangs zu Wissen gewidmet waren, haben wir uns bemüht, den Zusammenhang zwischen den einzelnen Beiträgen herzustellen und besonders auf das Bedürfnis der Archivare und Dokumentenverwalter einzugehen, allen Kategorien von Lesern den Zugang zu immer sachdienlicheren und individuelleren Informationen zu ermöglichen.

Rapport sur le Forum DLM 2002, Barcelone: sessions parallèles 2 et 5

Catherine Dhérent

Dans ce compte rendu des sessions parallèles 2 et 5 consacrées à l'utilisation de l'information et à l'amélioration de l'accès à la connaissance, on s'est efforcé de faire le lien entre les diverses interventions et d'insister sur le besoin pour les archivistes et les gestionnaires de documents de permettre l'accès pour tous types de lecteurs à une information de plus en plus pertinente et personnalisée.

Closing speech

Eric Ketelaar

Eric Ketelaar (1944) is Professor of Archivistics in the Department of Book, Archives and Information Studies of the University of Amsterdam (since 1997). His current teaching and research are concerned mainly with the social and cultural contexts of records creation and use.
In 2000/2001 he was the Netherlands Visiting Professor at the University of Michigan (School of Information). He was General State Archivist (National Archivist) of The Netherlands from 1989–97. From 1992–2002 he held the chair of archivistics in the Department of History of the University of Leiden. For 20 years he served the International Council on Archives (ICA) in different capacities. In 2000 ICA elected him Honorary President. He conducted seminars on archival legislation, appraisal and archival management in central Africa, central and eastern Europe and elsewhere. He wrote some 200 articles in Dutch, English, French and German and he wrote or co-authored several books, including two general introductions on archival research and a handbook on Dutch archives and records management law. Since 1986 he is editor of a multi-author loose-leaf handbook on archives and records management methodology and practice (now more than 1 500 pages). In 1997 The archival image, a collection of his essays in English, French and German, was published. He is one of the three editors-in-chief of Archival science. International Journal on Recorded Information._

The archive as a time machine

The ICT industry and public-sector partnership: to promote the preservation and accessibility of the European archival heritage

Eric Ketelaar

In the past two days we have been introduced to best practices and solutions for access and preservation of electronic information. In a remarkable way the three rapporteurs have summarised what has been achieved in the parallel sessions. The conclusions the chairman just read, are a concise summary of the DLM-Forum 2002 achievements. The papers presented were a source of inspiration and will continue to stimulate and enrich theory, methodology and practice. We likewise benefited from the demonstrations in the exhibition. We now depart to return to our research, development, and implementation in different institutional, national and international settings. But before we adjourn I ask your attention for the broader cultural framework of the partnership between the public sector at all levels and the ICT industry. A public-private partnership reinforced at this DLM-Forum 2002, with the common goal of the preservation and accessibility of the European archival heritage.

When, more than 30 years ago, I started as a civil servant in the Ministry of Culture of the Netherlands, I was taught that, before setting to work on a case, one had to ask for the file to be retrieved from the archives. The file contained the minutes, memoranda, record copies and other records of the history of the matter. It was a source of useful precedents. The archive, the memory of the ministry, formed by generations of predecessors, had to be consulted before one could undertake any new action. In those days, one phoned to the records office downstairs and after a few minutes a messenger or a records official knocked on the door, bringing the file.

But with the passage of time, in the 1970s, the interval between the phone call to the records office and the moment the file arrived on the civil servant's desk, became longer and longer. When I began my career, the norm was three to five minutes, late on this increased to 10, 15 minutes, even more. It was then (but also later) that I found a pattern which I will now present to you as 'Ketelaar's law'. We measure the interval between the request for the file from the archives and the moment the file can be actually consulted. If that interval averages 20 minutes, then — according to Ketelaar's law — civil servants start creating their personal file. By that I mean what the Germans call *Handakten* — literally 'handy records_, near at hand. In French they are called 'papiers personnels de fonction', in English 'working copies', 'semi-official records' or 'convenience files' — according to the *Dictionary of archival terminology* of the International Council on Archives: 'documents or copies thereof, papers and/or publications kept by or for officials for their private or personal use, relating directly or indirectly to their official duties.'

This definition makes it clear that in a semi-official file are assembled all sorts of papers which are not regarded as official records. A semi-official file does not belong to the official organisational memory, but to a twilight zone: it is neither strictly private, nor strictly official.

But practice is more stubborn than doctrine. Some official records do not even end up in the official file, but in the convenience file kept near at hand for use by the civil servant. The official file is thus less complete and less authentic. In the convenience file official records can be found among documents which archival theory does not consider official records, and with documents that archivists and records managers regard as non-record material. Creators and users of the file do not bother about the official definition of records. The convenience file with its varied content is an extension of a person's memory, forming the link between his or her memory and the organisational memory.

Sometimes the semi-official files have become so important that the official files can no longer give an authentic and reliable representation of what happened. During the recent investigation (commissioned by the Dutch Government) of the 1995 tragedy in Srebenica (where Dutch soldiers under UN command had the 'mission impossible' to control a safe area), researchers came across

the semi-official convenience files on the whole case, kept by a high-ranking official in the Dutch Ministry of Foreign Affairs. This collection yielded essential information not to be found elsewhere.

Daily practice in public and private organisations shows a blurring of the demarcation between records and non-record material. Secondly, it shows a blurring of the boundaries between personal and institutional memory. Ketelaar's law is an expression of a widely felt need to have one's paper memory close at hand. Each of these factors influences the creation, management, and use of the archival heritage of individuals, communities, societies — indeed, the European archival heritage.

'Everything is archive' is a chapter in *Je pense donc j'archive* (I think, therefore I archive), a recent book by a French archivist, Marie-Anne Chabin ([1]). What is an archive as perceived by society? The popular perception is that archives are cold, musty, dusty places that hold old records. Old records or, as in Webster's dictionary, public records or historical documents. But archives are not only history. As articles in the press clearly demonstrate these days, society considers everything an archive — books, papers, artifacts, sound, images, geological samples — that is serious and reliable information, put in storage to be retrieved when you need it: a backup that saves what may be of value in the future.

Ask any search engine the term 'archives' and it will yield millions of hits (Google on 16 August, 2001 24.3 million, on 21 April 2002 33.2 million — in eight months an increase with 37 %!), most of which are no archives or records in the archivists' terminology, but which are an expression of the value society attaches to keeping account of its present for its future.

Archiving, then, is not about history, but about the future. People want to get hold of the future. For many people, Anthony Smith argues 'the only guarantee of preservation of some form of identity is in the appeal to 'posterity', to the future generations ... only the appeal to a collective posterity offers hope of deliverance from oblivion' ([2]).

This explains why in our archiving society we see the compulsive creation of private records and archives, connecting the living history of individuals and families to national history ([3]). Records and archives in a broad sense, just as in a convenience file. More memory, than archive: the record as memory — 'O! that record is lively in my soul', to quote Shakespeare (Twelfth Night, Vi,1, 256). Marie-Anne Chabin proposes to distinguish archives by birth from archives by baptism, the former corresponding to records and archives in the archivists' terminology, the latter meaning those documents having no primary record status or value, which have survived and are recognised as having a value to retain a memory (or: memories) ([4]).

A remarkable programme in the United States, sponsored by the National Endowment for the Humanities and the White House Millennium Council, promotes the creation of public archives by private people, connecting their history to that of America. Collective memory, Susan Crane wrote, is ultimately 'located not in sites but in individuals. All narratives, all sites, all texts remain objects until they are 'read' or referred to by individuals thinking historically' ([5]).

This fits in with the conception of the man who can be regarded as the father of the notion of collective memory, Maurice Halbwachs. According to Halbwachs there is no individual memory dissociated from collective memory. He was also the first sociologist to stress 'that our conceptions of the past are affected by the mental images we employ to solve present problems, so that collective memory is essentially a reconstruction of the past in the light of the present' ([6]).

In collective memory it is not the facts that count, but remembrance of the past: historical facts are being transformed into myths ([7]). Earlier memory metaphors, like 'photographic memory', assumed that perceptions are stored in memory as immutable traces. More recently, people use the computer as a metaphor for human memory. Both the computer and human memory allow for replacing old data by new information and for altering stored information. The human memory does not store an exact reproduction, but filters incoming information which is coded into a representation of reality. Memory not as passive storage, but as an active power. In this respect collective memory acts just like an individual's memory.

Until now I referred to 'the' collective memory, but evidently there are as many collective memories as there are collectives and social groups. Even within one community there is interaction

([1]) Chabin, M.-A., *Je pense donc j'archive* (L'Harmattan, Paris and Montréal 1999).
([2]) Smith, 'A. D., Towards a global culture?', Featherstone, E. (ed.), *Global culture. nationalism, globalisation and modernity* (Sage, London, Thousand Oaks and New Delhi 1990) p. 182, Smith Anthony D., 'Towards a global culture?', *Theory, culture and society* 7 (1990) p. 182.
([3]) Nora, P., 'Entre mémoire et histoire. La problématique des lieux,' Nora, P. (ed.), *Les lieux de mémoire. I. La République* (Gallimard, Paris 1984), XXVII-XXVIII (English editions: Pierre Nora, 'Between memory and history. Les lieux de mémoire,' *Representations* 26 (Spring 1989) 13–14; *Realms of memory: rethinking the French past* (on the jacket: *The construction of the French past) I. Conflicts and divisions* (Columbia University Press, New York 1996) 8).
([4]) Chabin, *Je pense donc j'archive,* pp. 67–68.
([5]) Crane, S. A., 'Writing the individual back into collective memory', *American historical review* 102 (1997) 1381.
([6]) Coser l. A., (ed.), *Maurice Halbwachs, On collective memory* (University of Chicago Press, Chicago 1992) 34.
([7]) Assmann, J., *Das kulturelle Gedächtnis. Schrift, Erinnerung und politische Identität in frühen Hochkulturen* (Beck, C. H., München 2000, 3. Aufl.).

between various different memories ([8]). What we regard as collective memory, is what the members of a group, an organisation or a society want to remember. That is more than what the elite of that group appropriates as collective memory or what it enforces the group, through 'politics of memory', to view as collective memory ([9]).

The past is recreated and re-lived in commemorative practices, monuments, ceremonies and other 'theatres of memory', like traditional sports and pastimes, costume parades, 'retrochic' and the heritage industry — so beautifully parodied in Julien Barnes' *England, my England*. ([10]) What they all have in common is that they allow the individual to relate directly to what he or she regards as 'the' past: a personal immediacy.

But what is the past? Halbwachs opposed the mémoire collective to the mémoire historique. For Halbwachs history begins where living collective memory ends. But in history the past also lives, because, as the Dutch historian Johan Huizinga said, the past is not photographed, but re-imagined ([11]). History is a social construction, it is, in fact, a special form of collective remembrance. Instead of opposing memory and history, Aleida Assmann distinguishes two modes of remembering: *Funktionsgedächtnis* and *Speichergedächtnis* — functional memory and storage memory ([12]). The storage memory forms the perspective background for the functional memory: 'It holds additional knowledge which as memory of memories can ensure that really existing functional memories can be evaluated critically and, when necessary, be renewed or changed.' ([13])

The collective storage memory, according to Aleida Assmann, does not constitute collective identity as such, as functional memory does. The ever expanding storage memory keeps more information and different information which may be taken out by functional memory, and restructured and recomposed into stories, into meaning, constituting collective identities. Storage memory can be regarded as a reservoir for future functional memories, as a corrective for current functional memories. Therefore, Assmann writes, it is necessary to keep the boundaries between the two memories permeable, allowing the storage memory to act 'as a context for different functional memories, more or less as their outer horizon from where the narrowed perspectives of the past may be analysed, criticised, and, not the least, changed'. ([14])

It is time to draw a preliminary conclusion. Neither collective memory nor history are petrified fossils, but active forces, driven by preoccupations with the present. Italo Svevo used a beautiful metaphor when he wrote that the present is the conductor of an orchestra which is the past ([15]). This can easily be grasped by current post-modern society, sometimes even more easily than by some historians and archivists. In our time people no longer believe in all-embracing ideologies and grand narratives, but in a kaleidoscope of pluralistic stories. Stories replacing histories.

But what about the archival heritage, you may ask. Halbwachs is rather vague about the question how collective memory is transmitted from one generation to another. Paul Connerton, on the other hand, deals with this question in his book *How societies remember*. He says that what Halbwachs calls social memory, is in fact communication between individuals: 'To study the social formation of memory is to study those acts of transfer that make remembering in common possible.' ([16])

Transfer is not the first thing that comes to mind when thinking about an archive, with its immediate connotation of storage, a connotation reinforced by Aleida Assmann's concept of 'storage memory'. Nevertheless archiving — all the activities from creation and management to use of records and archives — has always been directed towards transmitting human activity and experience through time and, secondly, through space. A storage memory transmits information to some later point in time. It is this quality of the archive as a time machine, that I want to stress because it is so essential in promoting the preservation and accessibility of the European archival heritage

For many people, keeping records is identical to cleaning one's desk, sorting and throwing away and putting one's papers in a folder or a box. People who work on a computer save their files and make back-ups. But why do, and did, people keep their papers and computer files? Letters, bills, insurance papers and contracts all reflect an activity. In a business or organisation, paper documents and digital records are used to support the management and control of the work process.

([8]) Confino, A., 'Collective memory and cultural history: problems of method', *American historical review* 102 (5) (December 1977), pp. 1386-1403, esp. 1399–1403.
([9]) Confino, A., 'Collective memory', pp. 1393–1399. Unlike Confino, I do not restrict 'politics of memory' to politics: any power in any community uses politics of memory.
([10]) Samuel, R., *Theatres of memory. Volume I: Past and present in contemporary culture* (Verso, London/New York 1994); Lowenthal, D., *The heritage crusade and the spoils of history* (Free Press, New York 1996).
([11]) Huizinga, J., *Verzamelde werken* VII (H.D. Tjeenk Willink and zoon, Haarlem 1950) 12.
([12]) Assmann, A., *Erinnerungsräume. Formen und Wandlungen des kulturellen Gedächtnisses* (Beck, C. H., München 1999), pp. 130–142.
([13]) Assmann, A., *Erinnerungsräume*, 136: 'Es hält ein Zusatzwissen bereit, welches als Gedächtnis der Gedächtnisse dafür sorgen kann, daß real existierende Funktionsgedächtnisse kritisch relativiert und gegebenenfalls erneuert oder verändert werden können.'
([14]) Assmann, A., *Erinnerungsräume*, 141: 'als Kontext der verschiedenen Funktionsgedächtnisse gewissermaßen deren Außenhorizont, von dem aus die verengten Perspektiven auf die Vergangenheit relativiert, kritisiert, und nicht zuletzt: verändert werden können.'
([15]) Assmann, A., *Erinnerungsräume*, 17.
([16]) Connerton, 39.

Instructions for man and machine, the enhancement of the organisation's products and services, reporting — they all form part of the 'process-bound information', as it is called in archival science. It is the information generated by work processes that are all connected. The information is structured and recorded by these processes in such a way that it may be accessed from the context of the work processes ([17]).

Records are also used as a basis to account for the results of a business. They document transactions and relations — between supplier and client, editor and author or between the committee of a society and its members. They are kept to serve as an account of these transactions or relations, and as evidence. Evidence not only in the legal sense, but also from a historical point of view, to demonstrate what has been.

These two qualities of records fit into Aleida Assmann's scheme of the two modes of memory. The functional memory contains process bound information, while the storage memory serves the evidential function: 'Archives can be organised as functional or as storage memory. The former contains those documents and records that safeguard the legitimizing basis of existing power relations, the latter stores potential sources forming the basis of the historical knowledge of a culture.' ([18])

Archives and records as potential sources of historical knowledge, as a bridge to yesteryear or the past. That bridge is built equally by records in the professional records manager's sense, as by other 'memory-retaining objects'. Both are to be found in the civil servant's convenience file, or in any person's memory box. Everything is archive — not something old and forgotten, but something of value in the present and in the future. Shouldn't we try to connect to that public perception, rather than to impose on society what archivists believe to be true? I am not suggesting to neglect the specificity of records and archives as intrinsic process information or to dispose of the evidential quality of records and archives in favour of their cultural value. But we have to take into consideration that many of the artefacts that in the public perception are considered to be archives function in societal processes of accountability and evidence, just like records and archives-proper. Archives by birth and archives by baptism are not opposites, but should be seen — like Assmann's modes of memory — in a perspective view, where both, at the foreground and in the background, serve to understand the past.

And what about Ketelaar's law? Is it still applicable in a digital environment? Civil servants continue to create semi-official records which are not part of the organisation's record keeping system. Personal computers, laptops and notebooks throughout an organisation store databases, e-mail and other applications kept by staff members for their personal use while carrying out functions for their employer. Several of these systems are created and used by one individual only, just as in the paper age the official kept his or her personal notes, drafts and documentation in a convenience file. Many systems on these PCs and laptops are outside the purview of any organisational systems management and control. Many are, however, not private, but official records. So, Ketelaar's law may be expanded with the proposition — yet to be tested — that if a civil servant (or any administrator) sitting at his or her desktop, needs to click seven times before getting access to a centrally maintained application, he or she will start creating a personal record keeping system. In doing so, the organisational memory will be further fragmented and individualised. What the risks are of an organisational memory being dependent on what individuals have stored on their laptops and PCs, has been dramatically shown by the 11 September tragedies.

What are the solutions? Here is some food for thought while you are preparing your journey back home: six building blocks for possible strategies.

In the first place it is essential that organisational record-keeping systems take into account semi-official record-keeping by individual members of the organisation: record-keeping in the broader sense as 'everything is archive' ([19]). Individual and organisational memories blending together into one distributed memory.

Secondly, we have to rethink what the use of records and archives, as functional and storage memories, really is. In the past, staff kept the file close at hand, for immediate access. Access to the file meant: having the file in your hands. But in the digital age physical access is replaced by virtual access.

([17]) Thomassen, T., 'A first introduction to archival science', to be published in *Archival science* 2 (2002); Ketelaar, F.,'The benefit of archives', *Annual report 2001 Koninklijke Brill NV* (Brill, Leiden 2001), pp. 59–60.

([18]) Assmann, A., *Erinnerungsräume*, 409: 'Archive können sowohl als Funktions- wie als Speichergedächtnis organisiert sein; im einen Falle enthalten sie jene Dokumente und Beweisstücke, die die Legitimationsgrundlage bestehender Machtverhältnisse absichern, im anderen Falle bergen sie potentielle Quellen, die die Grundlage des historischen Wissens einer Kultur ausmachen.'

([19]) Finholt, T. A., 'The electronic office', Cooper, C., Rousseau, D. M., (eds.), *Trends in organisational behaviour* 4 (Wiley, Chichester 1997), pp. 29–41.

In his book *The age of access*, Jeremy Rifkin points to the transformation from a goods-producing to a services performing and experience-generating economy ([20]). No longer do we buy a product, we buy access to services. 'Services are being reinvented as long-term multifaceted relationships between servers and clients.' ([21]) Information technologies are used as relationship technologies. A cellular phone is given away for free, as an inducement to use the telecom services. The physical container becomes secondary to the unique services contained in it. The Oxford English Dictionary or the Encyclopedia Britannica have dematerialised into an online service. Books and journals on library shelves are giving precedence to access to services via the Internet ([22]).

Likewise, in record-keeping systems the files may conceptually be replaced by providing a service, that is: access to a distributed archival memory. Thus making the records manager into a records relations manager or a 'memory relations manager'.

Thirdly, archival institutions are moving from providing physical documents to providing access ([23]), from counting visits to the search room to counting hits on their web site, from issuing a reader's card as a ticket to enter the search room to issuing a customer's card as the start of a multi-faceted relationship between client and service-providing institution. At Amazon.com's web site I am welcomed with recommendations for books and CDs, based upon my browsing history and customer profile. Why doesn't the web site of my library offer such a service? Why don't the archives use their users' statistics and the Internet proactively to keep researchers informed — individually and tailormade — of new acquisitions, new finding aids, new publications, discounts and services which may interest them?

But consumers don't ask themselves as often 'What do I want to have that I don't have already'; they are asking instead, 'What do I want to experience that I have not experienced yet?' ([24]) In such an 'experience economy', as described by Joseph Pine and James Gilmore ([25]), companies produce memories, not goods ([26]). 'Memory has become one of the objects of the consumer society that sell well', wrote Jacques le Goff 20 years ago ([27]).

This leads to my fourth proposition. Preserving and making accessible the archival heritage means: offering the possibility to create memories ([28]). Archival institutions should present these memories as a historical experience, evoking a historical sensation. A small museum in Terhorne, in Friesland, The Netherlands, calls itself *Ervarium*, 'experiencium'. There and elsewhere — I am thinking of the Yarvik museum in York where the visitor even experiences the smell of the past — a historical sensation is evoked: re-enacting, re-living the past ([29]). The archive is — or should be — a place for such a re-enactment as well, a place where the *gout d'archive* (the taste of archive) ([30]) functions as a Proustian *madeleine.*

Archives are — and this is my fifth point — time machines enabling man to carry his thoughts, experiences and achievements through time. The archives of British Columbia present a 'BC time machine' on the Internet, but it only allows a ride back into the past. The archival time machine which I propose, rides into the future and back from the future.

To enable a time travel back and forth, archivists have to care about the future. Not only archivists: the DLM-Forum — from the first one in 1996 throughout this DLM-Forum 2002 — demonstrates clearly that the concern for digital longevity is shared by archivists and administrators in the public and the private sector. We can only achieve this effectively and efficiently through industry-standard hardware and software. That is why the contribution of the ICT industry — called to arms by the DLM-Forum in 1999 — is so essential to ensure travel into future's 'deep time'.

Mankind requires such time machines because they offer an escape from a world in which, as Rifkin writes, 'the marketing experts, advertisers, and cultural intermediaries will be ready and waiting at the gateways, offering up access to all sorts of meaningful new cultural commodities and lived experiences for the price of admission. They will prospect local cultures for fresh fragments of cultural experience that can be mined and commodified. They will make their way back into history in search of story lines for creating exciting and entertaining new experiences.' ([31])

To such commodification and usurpation of private and public memories, a countervailing power must be found. Access to information and knowledge will be that countervailing power. And this

([20]) Rifkin, J., *The age of access. The new culture of hypercapitalism where all of life is a paid-for experience* (Tarcher/Putnam, New York 2000).
([21]) Rifkin, J., *The age of access*, 85.
([22]) Rifkin, J., *The age of access*, 76-93, 100.
([23]) Menne-Haritz, A., 'Access — the reformulation of an archival paradigm', *Archival science* 1 (2001), pp. 57–82.
([24]) Rifkin, J., *The age of access*, p. 145.
([25]) Pine, J., Gilmore, J., *The experience economy. Work is theatre and every business a stage* (Harvard Business School Press, Cambridge MA 1999).
([26]) Rifkin, J., *The age of access*, p. 145.
([27]) Le Goff, J., *Histoire et mémoire* (Gallimard, Paris 1988) 170: 'la mémoire étant devenue un des objets de la société de consommation qui se vendent bien.'
([28]) Menne-Haritz, A., 'Access', p. 59.
([29]) Walsh, K.,'Simulating the past can have negative effects as well', *The representation of the past. Museums and heritage in the post-modern world* (Routledge, London and New York 1992).
([30]) Farge, A., *Le goût de l'archive* (Editions du Seuil, Paris 1989).
([31]) Rifkin, J., *The age of access*, p. 217.

is my sixth and final point. Access to the collective knowledge and wisdom, as a prerequisite for civil education which is 'an essential tool in re-establishing a balanced ecology between culture and commerce.' ([32]) This means: equal access to information, in the context of the cultural diversity of our world, empowering citizens, enabling people to use the archive as a time machine in the present, the past and the future.

This DLM-Forum 2002 was the third in a series, starting in 1996. I believe it is appropriate to recognise the outstanding achievement of the man who inspired and organised the three DLM-Forums: Dr Hans Hofmann.

At the opening of the second DLM-Forum, 18 October 1999, the audience was challenged by the Secretary-General of the European Commission 'to continue seeking ways to preserve past, present and future information, electronic and on multimedia, in a viable and accessible form' ([33]).

This is as true today, as it was three years ago. It is as true, as the opening words of the first DLM-Forum, 18 December 1996, that I repeat now as my final conclusion: 'The digital era could well depend for its development and innovation on the memory which you've brought here today and intend to carry forward to the future' ([34]).

([32]) Rifkin, J., *The age of access*, pp. 254–255.
([33]) *Proceedings of the DLM-Forum on electronic records. European citizens and electronic information: the memory of the Information Society. Brussels, 18–19 October 1999, INSAR Supplement* IV (2000) 15.
([34]) *Proceedings of the DLM-Forum on electronic records. Brussels, 18–20 December 1996, INSAR Supplement* II (1997) 16.

Los archivos: una máquina del tiempo

La asociación entre las tecnologías de la información y de las telecomunicaciones y el sector público: promover la conservación del patrimonio archivístico europeo y su accesibilidad

Eric Ketelaar

Conservar el patrimonio archivístico y hacerlo accesible significa proponer al público la posibilidad de crear recuerdos y de conocer el pasado utilizando los archivos como una máquina del tiempo.

En esta «era del acceso», los sistemas de archivo y los servicios de archivos abandonan cada vez más el suministro de documentos físicos en favor del suministro de acceso a la memoria colectiva. Las fronteras entre los archivos y el material que no es de archivo, así como entre la memoria personal e institucional, se desdibujan en la mayoría de las organizaciones. De ello deriva una fragmentación y una individualización de la memoria organizativa. Los trágicos acontecimientos del 11 de septiembre de 2001 pusieron de manifiesto cuáles eran los riesgos de una memoria organizativa dependiente de lo que los individuos hayan introducido en la memoria de sus PC y de sus ordenadores portátiles. Asimismo, los sistemas de archivo de las organizaciones deben garantizar que se mezclen las memorias de los individuos con la memoria de la organización.

Numerosos instrumentos que el público considera archivos intervienen en procesos colectivos de rendimiento de cuentas y pruebas, tal como ocurre con los documentos y archivos. Los archivos «de nacimiento» y los archivos «bautizados» no son conceptos contrarios, sino que deben considerarse en una perspectiva desde la que permitan comprender el pasado.

Para poder hacer recorridos en el tiempo, los archiveros y los administradores de los sectores público y privado deben preocuparse por el futuro, en colaboración con los sectores de las tecnologías de la información y de las telecomunicaciones (TIC).

Das Archiv als Zeitmaschine

Die Partnerschaft der IKT-Industrie mit dem öffentlichen Sektor: Förderung der Bewahrung und Zugänglichkeit des europäischen archivalischen Erbes

Eric Ketelaar

Bewahrung und Zugänglichmachung des archivalischen Erbes bedeutet, den Menschen die Möglichkeit zu geben, durch Nutzung des Archivs als Zeitmaschine Erinnerungen zu erzeugen und „die Vergangenheit" zu erleben.

In unserem „Zeitalter des Zugangs" vollzieht sich eine Wandlung bei Schriftgutverwaltungssystemen und Archiveinrichtungen, die nicht mehr nur Dokumente als solche bereitstellen, sondern Zugang zum kollektiven Gedächtnis verschaffen. In den meisten Organisationen verschwimmen die Grenzen zwischen Archivgut und Nichtarchivgut sowie zwischen dem persönlichen und dem institutionellen Gedächtnis, was zu einer Fragmentierung und Individualisierung des Organisationsgedächtnisses führt. Welche Risiken entstehen, wenn sich ein Organisationsgedächtnis darauf stützt, was einzelne Mitarbeiter auf ihren Laptops speichern, haben die tragischen Ereignisse des 11. September aufs Deutlichste veranschaulicht. Deshalb sollten organisationseigene Datenverwaltungssysteme dafür sorgen, dass individuelle und Organisationsgedächtnisse zusammengeführt werden.

Viele der Artefakte, die in der öffentlichen Wahrnehmung als Archive angesehen werden, funktionieren in gesellschaftlichen Prozessen der Nachweisfähigkeit und Nachweispflicht ebenso wie Unterlagen und Archive im eigentlichen Sinne. „Richtige" und „vermeintliche" Archive bilden keinen Gegensatz, sondern sollten in einer Perspektive betrachtet werden, in der beide – im Vordergrund und im Hintergrund – dazu dienen, die Vergangenheit zu verstehen.

Um eine Zeitreise hin und zurück möglich zu machen, müssen sich Archivare und Verwalter im öffentlichen und privaten Sektor in Zusammenarbeit mit der IKT-Industrie um die Zukunft kümmern.

Les archives: une machine à remonter le temps

Le partenariat entre les TIC et le secteur public: promouvoir la conservation du patrimoine archivistique européen et son accessibilité

Eric Ketelaar

Conserver le patrimoine archivistique et le rendre accessible signifie de proposer aux gens la possibilité de se forger des souvenirs et de connaître le passé en utilisant les archives comme une machine à remonter le temps.

En cet «âge de l'accès», les systèmes d'archivage et les services d'archives abandonnent de plus en plus la fourniture de documents physiques au profit de la fourniture d'un accès à la mémoire collective. Les frontières entre archives et matériels «non d'archives» et entre la mémoire personnelle et institutionnelle s'estompent dans la plupart des organisations. Il en découle une fragmentation et une individualisation de la mémoire organisationnelle. Les événements tragiques du 11 septembre 2001 ont montré quels étaient les risques d'une mémoire organisationnelle dépendante de ce que les individus ont mis en mémoire dans leurs portables et dans leurs PC. Aussi les systèmes d'archivage des organisations doivent-ils être la garantie d'un mélange entre mémoire de l'individu et mémoire de l'organisation.

De nombreuses créations humaines, que le public considère comme étant des archives, interviennent dans les processus collectifs de transparence et d'administration de preuve, à l'instar des documents et des archives. Les archives «nées» ou «baptisées» comme telles ne s'opposent pas, mais doivent être vues dans une perspective où elles permettent, au premier plan comme à l'arrière-plan, de comprendre le passé.

Pour pouvoir faire des allées et venues dans le temps, les archivistes et les administrateurs du public et du privé doivent se soucier du futur, en collaboration avec les industries des TIC.

Conclusions and follow-up

Erik Norberg

Erik Norberg, born 1942, graduated in history at the University of Stockholm in 1971. He was trained as archivist at the Royal Military Archives, worked with the United Nations (UNIDO, Unesco), was Director of the Royal Military Archives 1982–91 and Director-General of the National Archives from 1991. His research covers contemporary political and military history. He has served with the ICA and other international bodies in different capacities and was chair of the European Board of ICA 1996–2000.

Conclusions and follow-up: towards a European DLM-network on electronic archives

Erik Norberg

There is an old tradition of looking upon archives as a national heritage only and to erect repositories in the shape of national monuments. But reality leads us in a different direction. Modern archival science and current development in archives and records management is transnational. In the electronic era the borders between records creators are established in a functional way instead of being remnants from a historical development. Now the region is the European Union and the context is a developing partnership elaborated during the third DLM-Forum. The conclusions of the DLM-Forum 2002 bear evidence of a fruitful and target-oriented cooperation in the archival field and I will present some short comments on the major items.

At the first DLM-Forum in Brussels 1996 10 follow-up points were adopted. Among the subjects mentioned were better knowledge of the relationship between public administration and archives services, the need of guidelines and best practices for machine readable data, the need of further investigation of legal implications. The primary message of the second DLM-Forum 1999, also in Brussels, was the need for closer links with the ICT industry. The background was a deep common interest: a fast growing market for electronic documents and records management in the archival field, in libraries and museums and furthermore in the entire public and private sectors. The result was a DLM message to the ICT industry, developing several major fields of future cooperation.

The conclusions of the third DLM-Forum are now giving clear evidence of an emerging and important partnership in the field of electronic document and records management, particularly with the ICT industry. The message of 1999 gave a positive answer and six Industry White Papers were presented now in 2002. In his preface to these papers, Erkki Liikanen, European Commissioner for Enterprise and the Information Society, states the aim of the DLM-Forum 'as a crucial requirement for preserving the information society as well as improving business processes for more effective government'. The content of the White Papers received added value by the presentations at the conference, and the DLM-Forum and entire archival community can only welcome the intention of the ICT industry to continue with programmes in this field.

A basic requirement for electronic records management for archivists, public administration and industry is a specification like the European Model Requirements (MoReq). This standard for model requirements and metadata has in general received a positive response. It deserves support from the DLM community and can be developed into a European standard.

The question of education and training has been a prioritised item in all DLM-Forums. E-learning is a new but indispensable tool. Training modules on electronic records management have been elaborated and are now available. We should go on strengthening the cooperation between archive schools in Europe. Another item of highest priority covers the legal aspects related to access, data protection, authenticity, copyright and other issues. Ongoing technical development is followed by continuous changes in legislation and directives on all levels. The task of the archival community must be to analyse and take action in favour of a harmonised approach to access to archives.

And so to the future of DLM itself, or rather the future of a forum for all activities included in electronic archiving. The DLM-Forum 2002 made clear the need of a strong and lasting European platform, where the new partners from the public sector, industry and other bodies could join their efforts. Consequently the DLM-Forum proposed the launch of a European DLM network of excellence on electronic archives. This network aims at the creation of a network of expertise in the field of digital document life-cycle management and access to public information. The DLM-Forum has so far been a success, but now we need further efforts to strengthen the cooperation, to bring all stakeholders together and share the expertise with all interested public or private bodies in Europe. It is now our task to form an interdisciplinary working alliance of institutions, universities, and

archive schools, administrations, organisations, associations, ICT companies and other professionals who are active in the field of electronic archives in Europe.

To develop this initiative of the DLM-Forum, the Scientific Committee of the DLM-Forum has established a DLM Monitoring Committee. One of the first steps will be the preparation of an application for the sixth framework programme of the European Union.

Thus the DLM-Forum 2002 in Barcelona is a further important step in the development of electronic archiving. The Forum has been successfully organised by the Catalan Government in association with other national and international institutions and with great support from the European Union Presidency, the European Commission and the representatives from the ICT industry. It is my pleasure to convey to all those responsible for the conference the appreciation and thanks from the participating archival community.

Conclusiones y seguimiento: hacia una «red DLM» europea para los archivos electrónicos

Erik Norberg

Existe una vieja tradición que consiste en considerar los archivos únicamente como un patrimonio nacional, y en elevarlos al rango de monumento nacional. Ahora bien, la realidad nos lleva por otra dirección. La archivística moderna y la evolución actual de los archivos y de la gestión documental revisten un carácter transnacional. En la era de la electrónica, las fronteras entre los creadores de documentos se trazan de manera funcional, en vez de ser reminiscencias de una evolución histórica. La Unión Europea constituye ahora la región, y el contexto es una asociación evolutiva, elaborada en el tercer Foro DLM. Las conclusiones del Foro DLM 2002 dan prueba de una cooperación fructífera y específica en el ámbito de los archivos. Haré unos breves comentarios sobre los principales elementos de esta cooperación.

En el primer Foro DLM, celebrado en Bruselas en 1996, se adoptaron diez medidas de seguimiento. Entre las cuestiones abordadas figuran las siguientes: un mejor conocimiento de los vínculos entre la administración pública y los servicios de archivo, la necesidad de directrices y buenas prácticas en materia de datos legibles por máquina, y la necesidad de analizar con mayor profundidad las implicaciones jurídicas. El principal mensaje del segundo Foro DLM, organizado también en Bruselas en 1999, fue la necesidad de establecer vínculos más estrechos con el sector de las tecnologías de la información y de las telecomunicaciones (TIC). El contexto, que reviste un interés común, era la evolución rápida del mercado de documentos electrónicos y de la gestión documental en el ámbito de los archivos, en las bibliotecas y museos y, más generalmente, en los sectores público y privado. El resultado fue un mensaje dirigido al sector de las TIC, que desarrolló varios grandes ámbitos para la cooperación futura.

Las conclusiones del tercer Foro DLM constituyen una prueba clara de una nueva asociación importante en el ámbito de la gestión de documentos y archivos electrónicos, en particular con el sector de las TIC. El mensaje de 1999 recibió una respuesta positiva, y en 2002 se han presentado seis Libros Blancos del sector. En su prólogo, Erkki Liikanen, Comisario europeo responsable de Empresa y Sociedad de la Información, considera el objetivo del Foro DLM como «una obligación crucial para preservar la sociedad de la información y para mejorar los procedimientos de trabajo con vistas a una administración más eficaz». El contenido de los Libros Blancos se vio consolidado con el valor añadido de las presentaciones de la conferencia, y el Foro DLM, así como toda la comunidad archivística, no pueden sino congratularse por la intención de los profesionales de las TIC de proseguir los programas en este ámbito.

Unos requisitos básicos, a imagen de las especificaciones modelo europeas (MoReq), constituyen una condición fundamental de la gestión de archivos electrónicos para los archiveros, la administración pública y el sector. Esta norma de especificaciones modelo y metadatos ha tenido en general una acogida favorable. Esta iniciativa merece ser apoyada por la comunidad de DLM y podría dar lugar a una norma europea.

Todos los Foros DLM han hecho hincapié en la educación y la formación. El aprendizaje electrónico es una herramienta nueva e indispensable. Se han desarrollado y ya están disponibles algunos módulos de formación sobre gestión de documentos electrónicos. Es necesario reforzar la cooperación entre las escuelas de formación sobre archivos en Europa. Los aspectos jurídicos relativos al acceso, la protección de datos, la autenticidad, los derechos de reproducción y demás, constituyen también una cuestión de primera importancia. La permanente evolución técnica va seguida automáticamente de una evolución del Derecho en todos los niveles. La misión de la comunidad de archiveros debe ser analizar y tomar medidas en favor de un enfoque armonizado del acceso a los archivos.

Lo mismo reza para el futuro de los DLM en sí, o más bien para el futuro de todas las operaciones incluidas en el archivo electrónico. El Foro DLM 2002 puso de relieve la necesidad de contar con una plataforma europea sólida y duradera, a la que podrán unir sus esfuerzos los nuevos socios del sector público, la industria y otros organismos. En consecuencia, el Foro DLM propuso lanzar una red DLM europea especializada en el ámbito de los archivos electrónicos. El proyecto tiene por objeto crear una red de conocimientos sobre la gestión del ciclo de vida de los documentos digitales y sobre el acceso a la información pública. El Foro DLM ha sido un éxito hasta ahora, pero en la actualidad debemos intensificar nuestros esfuerzos para reforzar la cooperación, para reunir a todos los protagonistas y para poner en común sus conocimientos técnicos con todos los organismos públicos o privados interesados en Europa. Es ahora nuestro deber crear una asociación de trabajo interdisciplinar entre las instituciones, las universidades, las escuelas de formación, las administraciones, las organizaciones, las asociaciones, las empresas de TIC y otros profesionales que trabajan en el sector de los archivos electrónicos en Europa.

Con el fin de desarrollar esta iniciativa del Foro, el Comité científico del Foro DLM creó el Comité de seguimiento DLM. Una de las primeras etapas consistirá en preparar una candidatura al sexto programa marco de investigación de la Unión Europea.

El Foro DLM 2002 de Barcelona constituye un importante paso en el desarrollo de los archivos electrónicos. Las autoridades de Cataluña han organizado con éxito el encuentro, en colaboración con otras instituciones nacionales e internacionales y con el apoyo de la Presidencia de la Unión Europea, la Comisión Europea y los representantes del sector de las TIC. Con mucho gusto transmito a todos los responsables de esta conferencia la enhorabuena y el agradecimiento de los profesionales de los archivos aquí presentes.

Schlussfolgerungen und Folgemaßnahmen: Einrichtung eines europäischen DLM-Kompetenznetzwerks für elektronische Archive

Erik Norberg

Einer alten Tradition nach werden Archive ausschließlich als nationales Erbe betrachtet und Archivgebäude in der Form nationaler Monumente errichtet. Die Realität weist jedoch in eine andere Richtung. Die moderne Archivistik und die aktuellen Entwicklungen in der Archiv-

und Schriftgutverwaltung sind grenzüberschreitend. Im Elektronikzeitalter ergeben sich die Grenzen zwischen Urhebern von Unterlagen funktionell und sind keine Überbleibsel einer historischen Entwicklung mehr. Das Erfassungsgebiet ist nunmehr die Europäische Union, und der Kontext ist eine sich herausbildende Partnerschaft, wie sie auf dem 3. DLM-Forum ausgearbeitet wurde. Die Schlussfolgerungen des DLM-Forums 2002 zeugen von einer fruchtbaren und zielgerichteten Zusammenarbeit im archivischen Bereich, und ich möchte hier kurz auf die wichtigsten Punkte eingehen.

Auf dem ersten Brüsseler DLM-Forum im Jahre 1996 wurden zehn Folgemaßnahmen verabschiedet. Erwähnt wurden dabei unter anderem bessere Kenntnisse über die Beziehungen zwischen öffentlicher Verwaltung und Archivdiensten, die Notwendigkeit von Leitlinien und Beispielen bewährter Verfahren für maschinenlesbare Daten sowie die Notwendigkeit weiterer Untersuchungen zu den rechtlichen Aspekten. Die wichtigste Erkenntnis des DLM-Forums 1999, das ebenfalls in Brüssel stattfand, war das Gebot engerer Verbindungen zur IKT-Industrie vor dem Hintergrund starker gemeinsamer Interessen: ein rasch wachsender Markt für elektronische Dokumenten- und Schriftgutverwaltung im Archiv-, Bibliotheks- und Museumsbereich und darüber hinaus im gesamten öffentlichen und privaten Sektor. Aus diesem Anliegen heraus wurde eine DLM-Botschaft an die IKT-Industrie mit verschiedenen Feldern einer künftigen Zusammenarbeit formuliert.

Die Schlussfolgerungen des 3. DLM-Forums zeigen nunmehr klar und deutlich eine sich entwickelnde wichtige Partnerschaft im Bereich elektronische Dokumenten- und Schriftgutverwaltung insbesondere mit der IKT-Industrie. So fand die Botschaft von 1999 eine positive Resonanz, und es wurden jetzt im Jahre 2002 sechs Industrie-Weißbücher vorgelegt. Im Vorwort zu diesen Weißbüchern bezeichnet Erkki Liikanen, Mitglied der Europäischen Kommission für Unternehmen und Informationsgesellschaft, das Ziel des DLM-Forums als „maßgebliches Erfordernis für die Erhaltung der Informationsgesellschaft sowie für die Verbesserung der Geschäftsprozesse im Sinne eines effektiveren Staates“. Aufgewertet wurde der Inhalt der Weißbücher durch Präsentationen auf der Konferenz, und das DLM-Forum sowie das gesamte Archivwesen können die Absicht der IKT-Industrie, die Programme in diesem Bereich fortzuführen, nur begrüßen.

Eine Grundvoraussetzung der elektronischen Schriftgutverwaltung für Archivare, die öffentliche Verwaltung und die Industrie sind Spezifikationen wie die European Model Requirements (MoReq). Dieser Standard für Musteranforderungen und Metadaten ist generell auf ein positives Echo gestoßen. Er verdient die Unterstützung der DLM-Gemeinschaft und kann zu einer europäischen Norm weiterentwickelt werden.

Die Frage der Aus- und Weiterbildung spielte bei allen DLM-Foren eine herausragende Rolle. e-Learning ist hier ein neues, doch nicht mehr wegzudenkendes Instrument. Es wurden Schulungsmodelle für die elektronische Schriftgutverwaltung erarbeitet, die nunmehr zur Verfügung stehen. Jetzt geht es darum, die Zusammenarbeit zwischen den Archivschulen Europas auszubauen. Ein weiteres Thema höchster Priorität sind die rechtlichen Aspekte in Bezug auf Zugang, Datenschutz, Authentizität, Urheberrecht und andere Fragen. Der ständigen technischen Weiterentwicklung folgt ein ständiger Wandel der Rechtsvorschriften auf allen Ebenen. Aufgabe des Archivwesens muss es sein, die Situation zu analysieren und Maßnahmen zugunsten eines harmonisierten Konzepts für den Zugang zu Archiven zu ergreifen.

Nun zur Zukunft von DLM selbst, oder vielmehr zur Zukunft eines Forums für alle Tätigkeiten im Zusammenhang mit elektronischer Archivierung. Das DLM-Forum 2002 machte die Notwendigkeit einer starken und dauerhaften europäischen Plattform deutlich, auf der die neuen Partner aus dem öffentlichen Sektor, der Industrie und anderer Gremien ihre Anstrengungen bündeln können. Daher schlug das DLM-Forum die Einrichtung eines europäischen DLM-Kompetenznetzwerks für elektronische Archive vor. Dabei geht es um die Schaffung eines Netzes von Kompetenzen in Bereich Lebenszyklusmanagement für digitale Dokumente und Zugang zu öffentlichen Informationen. Das DLM-Forum war bisher ein Erfolg, doch bedarf es jetzt weiterer Anstrengungen, um die Zusammenarbeit noch mehr zu

stärken, alle Interessengruppen zusammenzubringen und die Erfahrungen mit allen interessierten öffentlichen oder privaten Gremien in Europa zu teilen. Unsere Aufgabe ist es nunmehr, ein interdisziplinäres Arbeitsbündnis von Institutionen, Universitäten und Archivschulen, Verwaltungen, Organisationen, Vereinen, IKT-Unternehmen und anderen Berufsgruppen zu schmieden, die auf dem Gebiet elektronischer Archive in Europa tätig sind.

Zur Ausgestaltung dieser Initiative des DLM-Forums hat der Wissenschaftliche Ausschuss des DLM-Forums einen DLM-Monitoringausschuss gebildet. Eine der ersten Aktionen wird die Erarbeitung eines Antrags für das 6. Rahmenprogramm der Europäischen Union sein.

Somit stellt das DLM-Forum 2002 in Barcelona einen weiteren wichtigen Schritt bei der Entwicklung der elektronischen Archivierung dar. Das Forum wurde von der katalanischen Regierung in Zusammenarbeit mit anderen nationalen und internationalen Einrichtung und mit großer Unterstützung durch den Ratsvorsitz der Europäischen Union, die Europäische Kommission und Vertreter der IKT-Industrie erfolgreich organisiert. Es ist mir eine große Freude, allen für die Konferenz Verantwortlichen die Wertschätzung und den Dank der anwesenden Vertreter des Archivwesens zu übermitteln.

Conclusions et suivi: pour un «DLM Network» européen sur les archives électroniques

Erik Norberg

Il existe une vieille tradition qui consiste à ne considérer les archives que comme un patrimoine national et à les ériger au rang de monument national. Or, la réalité nous incite à porter un autre regard. L'archivistique moderne et l'évolution actuelle des archives et de la gestion documentaire revêtent un caractère transnational. À l'ère de l'électronique, les frontières entre les créateurs de documents sont tracées de façon fonctionnelle au lieu d'être les survivances d'un développement historique. L'Union européenne constitue désormais la région, et le contexte est celui d'un partenariat évolutif, élaboré lors du troisième Forum DLM. Les conclusions du Forum DLM 2002 témoignent d'une coopération fructueuse et ciblée dans le domaine des archives. Je formulerai quelques brefs commentaires sur les principaux éléments de cette coopération.

Lors du premier Forum DLM, à Bruxelles en 1996, des mesures de suivi en dix points avaient été adoptées. Les sujets abordés étaient les suivants: une meilleure connaissance des liens entre l'administration publique et les services d'archives, la nécessité de lignes directrices et de bonnes pratiques pour les données lisibles par machine, la nécessité d'analyser davantage les implications juridiques. Le message premier du deuxième Forum DLM, organisé en 1999 également à Bruxelles, avait été de nouer des liens plus étroits avec le secteur des TIC. Le contexte, d'un intérêt commun, était celui de l'évolution rapide du marché des documents électroniques et de la gestion documentaire dans le domaine des archives, dans les bibliothèques et les musées et, plus généralement, dans le public tout entier et le secteur

privé. Un message avait alors été adressé aux industriels des TIC, développant plusieurs grands domaines de coopération future.

Les conclusions du troisième Forum DLM font désormais la démonstration claire et nette d'un nouveau partenariat majeur dans le domaine de la gestion de documents et d'archives électroniques, en particulier avec le secteur des TIC. Le message de 1999 a reçu une réponse positive, et six livres blancs du secteur sont soumis aujourd'hui, en 2002. Dans sa préface, Erkki Liikanen, commissaire européen chargé des entreprises et de la société de l'information, considère l'objectif du présent Forum DLM «comme une obligation cruciale pour préserver la société de l'information et pour améliorer les processus de travail en vue d'une administration plus efficace». Le contenu des livres blancs s'est vu conforté par la valeur ajoutée des présentations lors de la conférence, et le Forum DLM ainsi que toute la communauté des archives ne peuvent que se féliciter de l'intention des professionnels des TIC de poursuivre les programmes dans ce domaine.

Un cahier des charges, à l'image des spécifications types européennes (MoReq), constitue une condition fondamentale de la gestion d'archives électroniques pour les archivistes, l'administration publique et l'industrie. Cette norme de spécifications types et de métadonnées a reçu dans l'ensemble un accueil favorable. Cette initiative mérite d'être soutenue par le monde des DLM et pourrait donner lieu à une norme européenne.

Tous les Forums DLM ont mis en exergue l'éducation et la formation. L'apprentissage électronique est un nouvel outil, indispensable. Des modules de formation à la gestion des documents électroniques ont été développés et sont désormais proposés. Il importe de renforcer la coopération entre les écoles de formation aux archives en Europe. Les aspects juridiques concernant l'accès, la protection des données, l'authenticité, les droits de reproduction et autres constituent également une question de toute première importance. Les évolutions techniques permanentes sont automatiquement suivies d'une évolution du droit à tous les niveaux. La mission de la communauté des archivistes doit être d'analyser et de prendre des mesures en faveur d'une approche harmonisée de l'accès aux archives.

Il en va de même pour l'avenir des DLM en soi, ou plutôt pour l'avenir de toutes les opérations incluses dans l'archivage électronique. Le Forum DLM 2002 a mis en évidence la nécessité d'une plate-forme européenne solide et durable, grâce à laquelle les nouveaux partenaires du secteur public, de l'industrie et d'autres organismes pourront unir leurs efforts. En conséquence, le Forum DLM a proposé de lancer un réseau DLM européen d'excellence dans le domaine des archives électroniques. Le projet vise à créer un réseau d'expertise dans la gestion du cycle de vie des documents numériques et dans l'accès à l'information publique. Le Forum DLM a jusqu'à présent été une réussite, mais il nous faut maintenant intensifier nos efforts pour renforcer la coopération, pour réunir tous les acteurs partie prenante et mettre en commun leur savoir-faire avec tous les organismes publics ou privés intéressés en Europe. Il est maintenant de notre devoir de constituer un partenariat de travail interdisciplinaire entre les institutions, les universités, les écoles de formation, les administrations, les organisations, les associations, les industriels des TIC et avec d'autres professionnels travaillant dans le secteur des archives électroniques en Europe.

Afin de développer cette initiative du Forum, le comité scientifique du Forum DLM a mis en place le comité de suivi DLM. L'une des premières étapes consistera à préparer une candidature au sixième programme-cadre de recherche de l'Union européenne.

Le Forum DLM 2002 de Barcelone est donc une nouvelle étape majeure dans le développement des archives électroniques. Les autorités de Catalogne ont organisé avec succès la rencontre, en collaboration avec d'autres institutions nationales et internationales et avec le soutien de la présidence de l'Union européenne, de la Commission européenne et des représentants du secteur des TIC. C'est avec plaisir que je transmets à tous les responsables de cette conférence les félicitations et les remerciements des professionnels des archives présents.

Conclusions

* The DLM-Forum 2002 was organised by the Secretariat for Telecommunications and Information Society of the Catalan Government together with other Catalan institutions and departments of the Spanish central Government. They benefited from the support of the European Union Presidencies of Sweden (first half of 2001), Belgium (second half of 2001) and Spain (first half of 2002), the European Commission (Secretariat-General, Information Society DG) and representatives of the ICT industry. The DLM-Forum 2002 was organised in close cooperation with the EU Member States and regions.

CONCLUSIONS OF THE EUROPEAN DLM-FORUM 2002 Access and preservation of electronic information: Best practices and solutions

Barcelona 6–8 May 2002

The successful DLM-Forum 2002 (*) confirmed its role as the leading event for archivists, records and information managers of the European public sector. Since the mid-1990s the DLM-Forum has evolved as a platform for multidisciplinary cooperation between archivists, public administration, research and the ICT industry, with a view to identifying and promoting best practices and concrete solutions. This third DLM-Forum extended the participating community and thereby enriched the exchange of information and expertise on electronic document- and records-management. This event brought a new quality of partnership, in particular with the ICT industry. The Forum also launched the European DLM-network initiative. The acronym DLM was newly interpreted as 'document life-cycle management' to reflect the broadened scope of this initiative.

1. The ICT industry brings forward best-practice solutions for electronic document and records management

The message to industry of the DLM-Forum 1999 asked for support and advice on how to handle the challenge of document and information management in the digital age. The ICT industry gave a positive answer (*INSAR, European archives news, No 8*; http://europa.eu.int/historical_archives/dlm_forum) and provided a first series of six Industry White Papers about the following topics of interest:

1. Capture, indexing and auto-categorisation
2. Conversion and document formats
3. Content management
4. Access and protection
5. Availability and preservation
6. Education, training and operation

In the preface to these Industry White Papers, Erkki Liikanen, European Commissioner for Enterprise and the Information Society, states: 'The importance of providing public access and long-term preservation of electronic information is seen as a crucial requirement to preserve the memory of the information society as well as improving business processes for more effective government'. The content of the Industry White Papers was enhanced with the experiences of end-users given in presentations at the conference and shown in the exhibition.

The DLM-Forum welcomes the initiative by the ICT industry to continue with further events, seminars, additional publications and guidelines. The DLM-Forum underlines the necessity of a joint effort to deliver solutions.

2. The European model requirements (MoReq) — an emerging standard

The MoReq specification on model requirements and metadata is now positioned to make key contributions to effective electronic records management. They are already used in several European countries for tenders and implementations. The use of the MoReq specification should be further promoted to archivists, administration, industry and service providers.

The positive response to MoReq creates the possibility to establish it as a standard in Europe. This opportunity should be examined in more detail and supported by the DLM community.

3. Progress on professional education and training on electronic documents and records management

The DLM-Forum 2002 confirmed the essential role of education, training and e-learning in the field of document life-cycle management. The first training modules together with the pedagogic material of the European training programme on electronic records management (E-TERM) are now available to interested Member States and candidate countries. The DLM-Forum 2002 urges that the DLM-network of excellence pays special attention to forging stronger links between archive schools and other specialised training bodies in this field.

The DLM-Forum 2002 recommends the development of competencies, standards and knowledge profiles as instruments to enhance skills, implement solutions and support recruitment. The DLM-Forum stresses that special attention should be given to programmes to promote greater public awareness of the value of archival information and the necessity of the modern management of that information.

4. The challenge of constant change — legal aspects of modern archive management

The Forum notes that ongoing changes in legal regulations and standards at the regional, national and European levels, with regard to e-government, e-citizenship and e-commerce generate a need for a harmonised approach to create and make available information for short- and long-term use. The requirements of document life-cycle management make it necessary to examine in depth existing and forthcoming legislation which effect archives with respect to access, privacy, data protection, authenticity, copyright, electronic signatures and related issues.

The task is to define and propose adequate lines of action.

5. The launch of the European DLM-network initiative

One of the major results of the DLM-Forum 2002 is the broad acceptance and support to establish a strengthened and enlarged European platform with emphasis on standardisation, legal aspects, education, user access, long-term preservation and DLM events. During the Forum new partners from the public sector, industry, and other specialised bodies joined the initiative to launch the European DLM-network of excellence on electronic archives. It is envisaged that the DLM-network will organise the next DLM-Forum.

The Scientific Committee of the DLM-Forum 2002 established the DLM-Network Monitoring Committee to further develop and evaluate this DLM-network of excellence. The tasks include the preparation of an application for the sixth framework programme of the European Union, further integration of new partners and presenting a progress report on the evolving initiative.

The participants at the DLM-Forum 2002 stressed that the high quality of the papers presented at the forum merit publication. The European Commission is therefore asked to rapidly publish and distribute the proceedings of the DLM-Forum 2002.

CONCLUSIONES DEL FORO EUROPEO DLM 2002
Acceso y preservación de la información electrónica: mejores prácticas y soluciones

Barcelona 6-8 de mayo de 2002

El éxito del Foro DLM 2002 ([1]) ha confirmado su papel de acontecimiento de extraordinaria importancia para archiveros, gestores de documentos de archivo e información del sector público europeo. Desde mediados de la década de los noventa, el Foro DLM se ha convertido en plataforma para la colaboración multidisciplinar entre los archiveros, la administración pública, el mundo de la investigación y la industria de las tecnologías de la información y de las comunicaciones (TIC), con el objeto de identificar y fomentar las mejores prácticas y soluciones concretas. Este tercer Foro DLM ha ampliado la comunidad de profesionales que en él han participado, lo cual ha contribuido a enriquecer el intercambio de información y experiencias sobre la gestión de documentos y de documentos de archivo electrónicos. Ello ha supuesto, asimismo, una nueva calidad de colaboración, en concreto con la industria de las TIC. El Foro ha lanzado también la iniciativa llamada European DLM-Network (red europea de DLM). Las siglas DLM significan «Document Lifecycle Management» (gestión del ciclo de vida de los documentos), con el fin de reflejar el mayor alcance de esta iniciativa.

1. **La industria de las TIC ofrece soluciones de mejores prácticas para la gestión de documentos y de documentos de archivo electrónicos**

   El mensaje a la industria del Foro DLM de 1999 pedía apoyo y consejo sobre cómo hacer frente al reto de la gestión de los documentos y la información en la era digital. La industria de las TIC dio una respuesta positiva (*INSAR, European Archives News,* nº 8; http://europa.eu.int/historical_archives/dlm_forum) y proporcionó una primera serie de seis libros blancos y de la industria sobre los temas de interés siguientes:

   1) captura, indexación y autocategorización;
   2) conversión y formatos de documentos;
   3) gestión de contenidos;
   4) acceso y protección;
   5) disponibilidad y preservación;
   6) educación, formación y preservación.

   En el prefacio a estos libros blancos de la industria, Erkki Liikanen, Comisario europeo responsable de Empresa y Sociedad de la Información, afirma: «La importancia de ofrecer acceso público y preservación a largo plazo de la información electrónica se considera hoy requisito crucial para conservar la memoria de la sociedad de la información y, asimismo, mejorar los procesos de negocios de cara a un gobierno más efectivo». El contenido de los libros blancos de la industria se vio enriquecido con las experiencias de los usuarios finales aportadas en intervenciones realizadas durante el congreso y demostradas en la exposición.

   El Foro DLM da la bienvenida a la iniciativa planteada por la industria de las TIC de continuar con ulteriores actos, seminarios y publicaciones y directrices adicionales. El Foro DLM desea destacar la necesidad de realizar un esfuerzo conjunto con el fin de ofrecer nuevas soluciones.

2. **El *European Model Requirements* (MoReq) (modelos de requisitos europeos): una norma emergente**

   Las especificaciones del MoReq sobre modelos de requisitos y metadatos aparecen ahora como elemento clave para conseguir una gestión eficaz de los datos de archivo electró-

([1]) El Foro DLM 2002 fue organizado por la Secretaria de Telecomunicacions i Societat de la Informació de la Generalitat de Catalunya junto con otras instituciones catalanas y ministerios del Gobierno central español. Contaron, asimismo, con el apoyo de las Presidencias de Suecia (primera mitad de 2001), Bélgica (segunda mitad de 2001) y España (primera mitad de 2002) de la Unión Europea; de la Comisión Europea (Secretaria General, Dirección General de la Sociedad de la Información), y representantes de la industria de las TIC. El DLM-Forum 2002 fue organizado en estrecha colaboración con los Estados miembros y regiones de la Unión Europea.

nicos. Ya se están utilizando en varios países europeos para licitaciones y aplicaciones. El uso de las especificaciones MoReq debería fomentarse aún más entre los archiveros, la administración, la industria y los proveedores de servicios.

La respuesta positiva a MoReq crea la posibilidad de establecerlo como norma en Europa. Dicha oportunidad debería ser examinada más detenidamente y contar con el apoyo de la comunidad del DLM.

3. Avances en la educación y formación profesional sobre gestión de documentos y documentos de archivo electrónicos

El Foro DLM 2002 ha venido a confirmar el papel esencial desempeñado por la educación, la formación y el aprendizaje electrónico en el campo de la gestión del ciclo de vida de los documentos. Los primeros módulos de formación junto con el material pedagógico del programa de formación europeo sobre gestión electrónica de documentos de archivo electrónicos (e-term) se hallan actualmente a disposición de los Estados miembros interesados y países candidatos. El Foro DLM 2002 insta a la red de excelencia DLM a que preste una atención especial a la creación de vínculos más fuertes entre las escuelas de archiveros y otras entidades de formación especializadas en este campo.

El Foro DLM 2002 recomienda el desarrollo de competencias, normas y perfiles de conocimiento como instrumentos para potenciar aptitudes, aplicar soluciones y servir de apoyo al reclutamiento. El Foro DLM hace hincapié en que debería prestarse una atención especial a los programas encaminados a fomentar una mayor concienciación por parte del público sobre el valor de la información archivística y la necesidad de una gestión moderna de dicha información.

4. El reto del cambio constante: aspectos legales de la gestión moderna de los archivos

El Foro destaca que los cambios actuales que se están produciendo en las normativas y estándares legales a nivel regional, nacional y europeo, con respecto al gobierno electrónico, la ciudadanía electrónica y el comercio electrónico, plantean la necesidad de un enfoque armonizado para crear y poner a disposición la información para su uso a corto y largo plazo. Los requisitos de la gestión del ciclo de vida del documento hacen que resulte necesario examinar en profundidad la legislación existente y futura que afecte a los archivos con respecto a temas de acceso, privacidad, protección de los datos, autenticidad, derechos de autor, firmas electrónicas y cuestiones afines.

La tarea consiste en definir y proponer unas líneas de acción adecuadas.

5. El lanzamiento de la iniciativa «red europea DLM»

Uno de los principales resultados del Foro DLM 2002 es la amplia aceptación y apoyo conseguidos para establecer una plataforma europea reforzada y ampliada que haga hincapié en la estandarización, los aspectos legales, la educación, el acceso de los usuarios, la conservación a largo plazo y los actos del DLM. A lo largo del Foro una serie de colaboradores nuevos procedentes del sector público, la industria y otros organismos especializados se unieron a la iniciativa de lanzar la «red europea de excelencia DLM» sobre archivos electrónicos. Está previsto que la red DLM organice el próximo Foro DLM.

El comité científico del Foro DLM 2002 estableció el Comité de control de la red DLM (DLM-Network Monitoring Commitee) con el fin de fomentar el desarrollo y la evaluación de esta «red europea de excelencia DLM». Las tareas a realizar incluyen la preparación de una solicitud para el sexto programa marco de la Unión Europea, una ulterior incorporación de nuevos socios colaboradores y la presentación de un informe de los avances realizados en esta iniciativa en fase de evolución.

Los participantes en el Foro DLM 2002 destacaron el hecho de que la alta calidad de los documentos en él presentados merecía su publicación. Por esta razón, se ha solicitado a la Comisión Europea que publique y distribuya las actas del Foro DLM 2002 lo antes posible.

SCHLUSSFOLGERUNGEN DES EUROPÄISCHEN DLM-FORUMS 2002
Zugänglichkeit und Aufbewahrung elektronischer Daten: bewährte Verfahren und Lösungen

Barcelona, 6.-8. Mai 2002

Das erfolgreich verlaufene DLM-Forum 2002 ([1]) ist seiner Rolle als führende Veranstaltung für Archivare sowie für Dokumenten- und Informationsmanager des europäischen öffentlichen Dienstes gerecht geworden. Seit Mitte der 90er Jahre hat sich das DLM-Forum zu einer Plattform für die fächerübergreifende Zusammenarbeit zwischen Archivaren, öffentlicher Verwaltung, Forschung und IKT-Industrie mit dem Ziel entwickelt, bewährte Verfahren und konkrete Lösungen festzulegen und zu fördern. Bei diesem 3. DLM-Forum wurde die Zahl der Teilnehmer erweitert, und somit der Austausch von Informationen und Fachwissen im Hinblick auf elektronisches Dokumenten- und Records Management bereichert. Diese Veranstaltung führte zu einer neuen Qualität der Partnerschaft, insbesondere mit der IKT-Industrie. Auf dem Forum wurde auch die Europäische DLM-Netzinitiative ins Leben gerufen. Das Akronym DLM wurde neu ausgelegt als „Document Lifecycle Management" (Lebenszyklusverwaltung von Dokumenten), um den erweiterten Umfang dieser Initiative widerzuspiegeln.

1. IKT-Industrie fördert bewährte Lösungen für elektronisches Dokumenten- und Records Management

1999 war die Industrie in der Botschaft des DLM-Forums um Unterstützung und Rat in der Frage gebeten worden, wie die Herausforderung des Dokumenten- und Informationsmanagements im digitalen Zeitalter bewältigt werden kann. Die IKT-Industrie reagierte positiv (INSAR, Europäische Archivnachrichten, Nr. 8; http://europa.eu.int/historical_archives/dlm_forum/) und stellte sechs Weißbücher zu folgenden Schwerpunktthemen bereit:

1. Erfassung, Katalogisierung und automatische Kategorisierung
2. Konvertierung und Dokumentenformate
3. Content Management
4. Zugang und Schutz
5. Verfügbarkeit und Erhaltung
6. Aus- und Fortbildung und Betrieb

Im Vorwort zu diesen Weißbüchern der Industrie erklärte Erkki Liikanen, das für Unternehmen und Informationsgesellschaft zuständige Mitglied der Kommission: „Die Gewährleistung der öffentlichen Zugänglichkeit und der langfristigen Erhaltung von elektronischen Informationen ist notwendige Voraussetzung für die Bewahrung des Gedächtnisses der Informationsgesellschaft sowie für die Verbesserung der Geschäftsabläufe im Sinne effizienterer Regierungsmaßnahmen". Der Inhalt der Industrie-Weißbücher wurde aufgrund der auf der Konferenz und in der Ausstellung dargebotenen Erfahrungsberichte der Anwender erweitert.

([1]) Das DLM-Forum 2002 wurde veranstaltet vom Katalanischen Sekretariat für Telekomunikation und Informationsgesellschaft in Zusammenarbeit mit anderen katalanischen Institutionen und Ministerien der spanischen Zentralregierung sowie mit Unterstützung des EU-Ratsvorsitzes Schwedens (1. Halbjahr 2001), Belgiens (2. Halbjahr 2001) und Spaniens (1. Halbjahr 2002), der Europäischen Kommission (Generalsekretariat, Generaldirektion für die Informationsgesellschaft) und Vertretern der IKT-Industrie. Das DLM-Forum 2002 wurde veranstaltet in enger Zusammenarbeit mit den Mitgliedstaaten und den Regionen der Europäischen Union.

Das DLM-Forum begrüßt die Initiative der IKT-Industrie, weitere Veranstaltungen und Seminare durchzuführen sowie zusätzliche Schriften und Leitlinien zu veröffentlichen. Das DLM-Forum unterstreicht die Notwendigkeit gemeinsamer Lösungsbemühungen.

2. Die Europäischen Modellanforderungen (MoReq) – ein sich entwickelnder Standard

Die MoReq-Spezifikation über Modellanforderungen und Metadaten ist mittlerweile in der Lage, wichtige Beiträge für ein effizientes Management elektronischer Dokumente und Aufzeichnungen zu leisten. Sie werden bereits in mehreren europäischen Ländern für Ausschreibungen und Anwendungen benutzt. Archivare, Verwaltung, Industrie und Dienstleister sollten die MoReq-Spezifikation stärker nutzen.

Die positiven Reaktionen bieten die Möglichkeit, dass die MoReq in Europa zur Norm werden. Diese Möglichkeit sollten genauer geprüft und von der DLM-Gemeinschaft unterstützt werden.

3. Fortschritte bei der Aus- und Fortbildung im elektronischen Dokumenten- und Records Management

Das DLM-Forum 2002 bestätigte die entscheidende Rolle der Aus- und Fortbildung und des *e*-Learning im Bereich des Document Lifecycle Managements. Die ersten Lernmodule sowie das didaktische Material des Europäischen Fortbildungsprogramms für elektronisches Dokumenten- und Records Management (E-TERM) stehen nun für interessierte Mitgliedstaaten und Bewerberländer zur Verfügung. Das DLM-Forum 2002 drängt darauf, dass sich das DLM-Network of Excellence insbesondere darauf konzentriert, den Austausch zwischen Archivschulen und anderen einschlägigen Fachausbildungseinrichtungen zu intensivieren.

Das DLM-Forum 2002 empfiehlt die Entwicklung von Kompetenzen, Standards und Wissensprofilen als Instrumente zur Verbesserung von Fertigkeiten, zur Umsetzung von Lösungen und zur Förderung von Neueinstellungen. Nach Auffassung des DLM-Forums ist besonderes Augenmerk zu richten auf Programme zur Förderung einer stärkeren Sensibilisierung der Öffentlichkeit für den Wert von Archivinformationen und die Notwendigkeit eines modernen Managements dieser Informationen.

4. Die Herausforderung des ständigen Wandels – rechtliche Aspekte des modernen Archivmanagements

Das Forum stellt fest, dass der ständige Wandel der Rechtsvorschriften und Standards auf regionaler, nationaler und europäischer Ebene im Hinblick auf *e*-Government, *e*-Citizenship und *e*-Commerce einen Harmonisierungsbedarf schafft, um Informationen für die kurz- und langfristige Nutzung zu erstellen und zur Verfügung zu stellen. Die Anforderungen des Document Lifecycle Managements machen eine genaue Prüfung der bestehenden und zukünftigen Rechtssetzung für das Archivwesen im Hinblick auf den Zugang, den Schutz der Privatsphäre, den Datenschutz, die Authentizität, das Urheberrecht, elektronische Signaturen und damit verbundene Fragen notwendig.

Die Aufgabe besteht darin, entsprechende Maßnahmen auszuarbeiten und vorzuschlagen.

5. Die Gründung der Europäischen DLM-Netzinitiative

Eines der wichtigsten Ergebnisse des DLM-Forums 2002 ist die einhellige Unterstützung für die Einrichtung einer stärkeren und erweiterten europäischen Plattform mit den Schwerpunkten Standardisierung, rechtliche Aspekte, Aus- und Fortbildung, Anwenderzugang, langfristige Erhaltung und DLM-Veranstaltungen. Während des Forums beteiligten sich neue Partner aus dem öffentlichen Bereich, der Industrie und anderen Fachgremien an der Initiative zur Gründung des Europäischen DLM-Network of Excellence für elektronische Archive. Geplant ist, dass das DLM-Netz das nächste DLM-Forum veranstaltet.

Der Programmausschuss des DLM-Forums 2002 gründete das DLM-Network Monitoring Committee, um dieses DLM-Network of Excellence weiterzuentwickeln und zu bewerten. Zu den Aufgaben gehören die Erarbeitung eines Antrags für das 6. Rahmenprogramm der Europäischen Union, die weitere Einbeziehung neuer Partner und die Vorlage eines Fortschrittsberichts über die sich entwickelnde Initiative.

Die Teilnehmer des DLM-Forums 2002 betonten, dass die auf dem Forum gehaltenen Vorträge aufgrund ihrer hohen Qualität eine Veröffentlichung verdienen. Die Europäische Kommission wird daher gebeten, die Unterlagen des DLM-Forums 2002 rasch zu veröffentlichen und zu verteilen.

CONCLUSIONS DU FORUM DLM 2002
@ccès et préservation de l'information électronique: meilleures pratiques et solutions

Barcelone, 6-8 mai 2002

Le Forum DLM 2002 (*), qui a été couronné de succès, a confirmé son rôle de manifestation de premier plan pour les responsables de l'information et des archives du secteur public européen. Depuis le milieu des années 90, le Forum DLM est devenu une plate-forme de coopération pluridisciplinaire entre archivistes, représentants de l'administration publique, milieux scientifiques et industrie des TIC, en vue de recenser et de promouvoir les meilleures pratiques et des solutions concrètes. Le troisième Forum DLM a accueilli de nouveaux participants, enrichissant ainsi l'échange d'informations et de connaissances sur la gestion des documents et des archives électroniques. Cette manifestation a permis d'établir des partenariats d'une nature nouvelle, en particulier avec l'industrie des TIC. Elle a également servi de cadre au lancement du réseau DLM européen. La signification de l'acronyme DLM a été modifiée et renvoie désormais à «document lifecycle management» (gestion du cycle de vie des documents), reflétant la portée plus vaste de cette initiative.

(*) L'organisation du Forum DLM 2002 a été assurée par le secrétariat des télécommunications et de la société de l'information de la généralité de Catalogne, conjointement avec d'autres institutions catalanes et des services du gouvernement central espagnol. Cette manifestation a bénéficié du soutien des présidences suédoise (premier semestre 2001), belge (second semestre 2001) et espagnole (premier semestre 2002) de l'UE, de la Commission européenne (secrétariat général, direction générale de la société de l'information) et de représentants de l'industrie des TIC. Elle a été organisée en étroite collaboration avec les États membres et les régions de l'UE.

1. L'industrie des TIC propose des solutions performantes pour la gestion des documents et des archives électroniques

Dans son message à l'industrie, le Forum DLM de 1999 sollicitait aide et conseils sur la manière de relever le défi lié à la gestion des documents et de l'information à l'ère numérique. L'industrie des TIC a répondu positivement (INSAR, *Courrier européen des archives,* n° 8; http://europa.eu.int/historical_archives/dlm_forum/) et présenté une première série de six livres blancs sur les thèmes suivants:

1) la saisie, l'indexation et la catégorisation automatique;
2) la conversion et les formats des documents;
3) la gestion du contenu;
4) l'accès et la protection;
5) la disponibilité et la conservation;
6) la formation et l'application.

Dans la préface de ces livres blancs de l'industrie, M. Erkki Liikanen, commissaire européen chargé des entreprises et de la société de l'information, indique: «Garantir l'acces-

sibilité et la conservation à long terme de l'information électronique est une condition essentielle pour préserver la mémoire de la société de l'information et améliorer les processus de travail pour une administration plus efficace.» Les expériences présentées par les utilisateurs finals lors de la conférence et de l'exposition sont venues conforter le contenu des livres blancs de l'industrie.

Le Forum DLM se félicite de l'initiative de l'industrie des TIC d'organiser d'autres manifestations et séminaires et d'élaborer de nouvelles publications et orientations. Il souligne la nécessité d'entreprendre des efforts conjoints pour trouver des solutions.

2. Les exigences européennes types (MoReq), une norme naissante

Les exigences types et les métadonnées sont désormais à même de contribuer efficacement à la gestion des archives électroniques. Plusieurs pays européens les utilisent déjà pour des appels d'offres et des applications. Il convient de promouvoir leur utilisation auprès des archivistes, de l'administration, de l'industrie et des prestataires de services.

Compte tenu de l'accueil favorable qui leur a été réservé, les exigences types pourraient devenir une norme en Europe. Cette possibilité doit être examinée de manière plus approfondie et soutenue par la communauté DLM.

3. Progrès accomplis dans l'enseignement et la formation professionnels concernant la gestion des documents et des archives électroniques

Le Forum DLM 2002 a confirmé l'importance de l'enseignement, de la formation et de l'apprentissage électronique dans le domaine de la gestion du cycle de vie des documents. Les premiers modules de formation ainsi que le matériel didactique du programme européen de formation en gestion des archives électroniques (E-TERM) peuvent désormais être mis à la disposition des États membres et des pays candidats intéressés. Le Forum DLM 2002 insiste pour que le réseau d'excellence DLM forge des liens plus étroits entre les écoles d'archives et d'autres organismes de formation spécialisés dans ce domaine.

Il préconise que soient développées les compétences, les normes et les connaissances afin d'améliorer les qualifications, de mettre en œuvre les solutions et d'épauler les recrutements. Le Forum DLM souligne la nécessité de privilégier les programmes visant à sensibiliser davantage le public à la valeur des informations archivées et au besoin d'une gestion moderne de ces informations.

4. Le défi posé par l'évolution permanente: aspects juridiques d'une gestion moderne des archives

Le Forum constate que l'évolution permanente des réglementations et des normes aux niveaux régional, national et européen dans les domaines de l'administration en ligne, des applications électroniques ayant trait à la vie civique et du commerce électronique rend nécessaire une approche harmonisée de création et de diffusion d'informations à court et à long terme. Les exigences liées à la gestion du cycle de vie des documents requièrent d'examiner en profondeur la législation existante et à venir touchant aux archives, en termes d'accès, de protection de la vie privée, de protection des données, d'authenticité, de droits d'auteur, de signature électronique et d'autres questions connexes.

Il convient de définir et de proposer des lignes d'action adéquates.

5. Le lancement du réseau DLM européen

Un des principaux résultats du Forum DLM 2002 est d'avoir recueilli l'adhésion et le soutien des participants à l'établissement d'une plate-forme européenne renforcée et élargie,

axée sur la normalisation, les aspects juridiques, l'éducation, l'accès des utilisateurs, la conservation à long terme et les manifestations DLM. Lors du Forum, de nouveaux partenaires du secteur public et de l'industrie ainsi que d'autres organismes spécialisés se sont associés à l'initiative visant à lancer le réseau d'excellence DLM européen sur les archives électroniques. Il est prévu que le réseau DLM organise le prochain Forum DLM.

Le comité scientifique du Forum DLM 2002 a mis en place le comité de suivi du réseau DLM afin de développer et d'évaluer le réseau d'excellence DLM. Parmi les tâches à accomplir figurent la préparation d'un dossier de candidature au titre du 6e programme-cadre de l'Union européenne, l'intégration plus étroite des nouveaux partenaires et la présentation d'un rapport d'étape sur l'évolution de l'initiative.

Les participants au Forum DLM 2002 ont souligné qu'en raison de leur grande qualité, les contributions présentées lors du Forum méritaient d'être publiées. La Commission européenne est donc invitée à publier et à diffuser rapidement les actes du Forum DLM 2002.

Index of speakers

Participants of the DLM-Forum 2002

Àngels Adroer i Pellicer
Arxui Central Administratiu de la Delegació del Govern de la Generalitat a Girona
Illa Formentera, 42
17007 Girona
Spain
Tel. (34) 972 22 69 72
Fax (34) 972 22 39 95
E-mail: jjuher@yahoo.es

Ferran Agelet Ordobàs
Arxiu Municipal de Barcelona
Ciutat, 9, ent. 1ª
08002 Barcelona
Spain
Tel. (34) 934 19 89 55
Fax (34) 933 22 21 87
E-mail: arxivers@upcnet.es

Ramon Alberch i Fugueras
Arxiu Municipal
Ajuntament de Barcelona
Ciutat, 9, pral. 1a
08002 Barcelona
Spain
Tel. (34) 934 02 76 06
Fax (34) 934 02 76 38
E-mail: ralberch@mail.bcn.es;
arxcap@intercom.es

Rosa Almuzara Roca
AHC Mataro
D'en Palau 32
08301 Mataro
Spain
Tel. (34) 937 58 24 59
Fax (34) 937 58 24 59

Amparo Alonso García
Archivo General de Andalucía
C/ Almirante Apodaca No 4
41003 Sevilla
Spain
Tel. (34) 955 02 45 00
Fax (34) 955 02 45 12
E-mail: aalonso@ccul.junta-andalucia.es

José Alberto Alonso Martínez
Nou Silva Equips
Av. Prat de la Riba, 21, 2n 2a
43001 Tarragona
Spain
Tel. (34) 977 24 96 38
Fax (34) 977 24 96 38
E-mail: silvadoc@tinet.org

Mercedes Altés Santos
Arxiu Multijurisdiccional de Barcelona
C. Roger de Flor 62-68
08013 Barcelona
Spain
Tel. (34) 933 16 45 29
E-mail:
arxiu.multijuris.barcelona@xij.gencat.es

Miquel A. Amutio Gomez
Ministerio de Administraciones Públicas
Mariá de Molina 50
28071 Madrid
Spain
Tel. (34) 915 86 29 90
Fax (34) 915 86 19 20
E-mail: miquel.amutio@sgci.dgopti.map.es

Jordi Andreu Daufí
Universitat de Barcelona
Melcior de Palau, 140
08014 Barcelona
Spain
Tel. (34) 934 03 57 81
Fax (34) 934 03 57 72
E-mail: andreu@fbd.ub.es

Lluis Anglada i de Ferrer
Consorci de Biblioteques Universitaries de Catalunya
Gran Capita, 2–4, edifici Nexus
08034 Barcelona
Spain
Tel. (34) 932 05 64 64
Fax (34) 934 01 58 31
E-mail: langlada@cbuc.es

Carme Anglès Villaseca
Dept. d'Ensenyament Generalitat
Via Augusta, 202–226
08021 Barcelona
Spain
Tel. (34) 934 00 69 00-3863
Fax (34) 934 00 69 94
E-mail: r120@ense.gencat.es

Silvia Arango Docio
National Digital Archive of Datasets —
ULCC—
20 Guilford Street
WC1N 1DZ London
United Kingdom
Tel. (44-20) 76 92 74 12
E-mail: s.arango-docio@ulcc.ac.uk

Foteini Aravali
University of Glasgow
11 University Gardens
G12 8QQ Glasgow
United Kingdom
Tel. (44) 14 13 30 36 35
E-mail: a@student.gla.ac.uk

Tiit Arumae
National Archives of Estonia
Maneezi 4
15019 Tallinn
Estonia
Tel. (372) 693 80 09
Fax (372) 661 62 30
E-mail: tiit.arumae@ra.ee

Kevin Ashley
ULCC
20 Guilford Street
WC1N 1DZ London
United Kingdom
Tel. (44-20) 76 92 13 79
Fax (44-20) 76 42 12 34
E-mail: k.ashley@ulcc.ac.uk

Torsti Astrén
Finnish Defence Forces
Pl 9
00861 Helsinki
Finland
Tel. (358) 918 14 65 04
E-mail: torsti.astren@milnet.fi

Edward Atkinson
Bank for International Settlements
Centralbahnplatz 2
4002 Basel
Switzerland
Tel. (41) 612 80 80 61
E-mail: edward.atkinson@bis.org

Rosa Maria Aubesart Martinez
Departament de Medi Ambient
Còrsega, 594
08025 Barcelona
Spain
Tel. (34) 934 55 94 71
Fax (34) 934 55 94 71
E-mail: toni.rosa@menta.net

Ann Aurén
Ministry of Culture
Drottninggatan 16
103 33 Stockholm
Sweden
Tel. (46) 84 05 34 91
Fax (46) 84 05 13 12
E-mail: ann.auren@culture.ministry.se

Steve Bailey
Joint Information Systems Committee
JISC Office, 3rd Floor, Strand Bridge House,
138-1
WC2R 1HH London
United Kingdom
Tel. (44) 709 23 08 50
E-mail: steve.bailey@bristol.ac.uk

Francesc Balada I Bosch
Arxiu Nacional de Catalunya
Carrer de Jaume I, 33–51
08190 Sant Cugat del Vallés
Spain
Tel. (34) 935 89 77 88
Fax (34) 935 89 80 35

Maria Antònia Balsells Font
Generalitat de Catalunya / Conselleria
d'Interior
Via Laietana, 69
08003 Barcelona
Spain
Tel. (34) 934 84 00 16
Fax (34) 934 84 04 16
E-mail: mbalsells@correu.gencat.es

Assumpció Banqué Molas
Convergència Democràtica de Catalunya
Còrsega, 333
08037 Barcelona
Spain
Tel. (34) 932 30 31 00
Fax (34) 932 36 31 10
E-mail: a.banque@convergencia.org

Francisco Barbedo
Arquivo Distrital do Porto
R. das Taipas, 90
4050–598 Porto
Portugal
Tel. (351) 23 39 51 70
Fax (351) 23 39 51 79
E-mail: fbarbedo@mail.telepac.pt

Montserrat Beltran i Morales
Ajuntament de Barcelona
Masferrer, 29–31
08028 Barcelona
Spain
Tel. (34) 932 91 64 32
Fax (34) 932 91 64 68
E-mail: mbeltran@mail.bcn.es

Bo Berg
Krigsarkivet
Banergatan 64
115 88 Stockholm
Sweden
Tel. (46) 87 82 69 94
Fax (46) 87 82 69 76
E-mail: bo.berg@krigsarkivet.ra.se

Anders Bergkvist
The Regional Archives in Visby/Landsarkivet i Visby
Visborgsgatan 1
SE-621 57 Visby
Sweden
Tel. (46) 498 21 05 14
Fax (46) 498 21 29 55
E-mail:
anders.bergkvist@landsarkivet-visby.ra.se

Angels Bernal I Cercos
Arxiu Nacional de Catalunya
Carrer Jaume I, 33-51
08190 Sant Cugat del Vallés
Spain
Tel. (34) 935 89 77 88
Fax (34) 935 89 80 35
E-mail: abernal@correu.gencat.es

Ewa Berndtsson
National Archives of Sweden
Box 12541
S-102 29 Stockholm
Sweden
Tel. (46) 87 37 63 50
Fax (46) 87 37 64 74
E-mail: ewa.berndtsson@riksarkivet.ra.se

Peter Berninger
European Commission
Office: C100 02/23
1049 Brussels
Belgium
Tel. (32-2) 296 74 02
E-mail: peter.berninger@cec.eu.int

Maike Bielfeldt
DE-CODA GmbH
Breite Straße 29
10178 Berlin
Germany
Tel. (49) 302 03 80 40 25
Fax (49) 302 03 08 40 30
E-mail: caspari@de-coda.de

Richard Blake
Records Management Dept
Public Record Office
TW9 4DU Kew, Surrey
United Kingdom
Tel. (44-20) 83 92 53 66
Fax (44-20) 88 78 89 05
E-mail: stephen.harries@pro.gov.uk

Jean-François Blanchette
Centre de recherche en droit public
Faculté de droit, Université de Montréal
C. P. 6128, succ. centre-ville
H3C 3J7 Montréal (Qc)
Canada
E-mail: blanc@lexum.umontreal.ca

Carme Blanco Cortés
Arxiu central Audiència Provincial Girona
Pl. Catedral,2
17004 Girona
Spain
Tel. (34) 972 20 99 00
E-mail: arxiu.audiencia.girona@xij.gencat.es

Pilar Blesa Herrero
MNAC
Palau Nacional. Parc de Montjuïc
08038 Barcelona
Spain
Tel. (34) 936 22 03 60
Fax (34) 936 22 03 74
E-mail: pilar.blesa@mnac.es

Amanda Blunden
Office National Statistics
Segensworth Road
Titchfield, Farnham
PO14 2DT Hants
United Kingdom
E-mail: amanda.blunden@ons.gov.uk

Joan Boadas Rasset
Associació d'Arxivers de Catalunya
Arxiu Municial de Girona
Placeta de l'institut vell, 1
17004 Girona
Spain
Tel. (34) 972 22 15 45
Fax (34) 972 20 26 94
E-mail: jboadas@ajgirona.org

Michele Boccaccio
Comunicando società per azioni
via Galermo 97
95123 Catania
Italy
Tel. (39) 09 57 31 50 47
E-mail:
michele.boccaccio@comunicandospa.net

Sven Bodin
Sundsvall miniápal
Norrmanmgatan 4
85185 Sundsvall
Sweden
Tel. (46) 601 91 18 76
Fax (46) 60 19 21 36
E-mail: svenbodin@sundsvall.se

Christer Bogefeldt
National Archives of Sweden
Box 12541
102 29 Stockholm Stockholm
Sweden
Tel. (46) 87 37 63 34
Fax (46) 86 57 95 64
E-mail: christer.bogefeldt@riksarkivet.ra.se

Javier Bolea Aguarón
Instituto de Estudios Altoaragoneses
C/ Parque 10
22002 Huesca
Spain
Tel. (34) 974 29 41 20
Fax (34) 974 29 41 22
E-mail: cipca@iea.es

François Bordes
Archives Municipales de Toulouse
Rue des Archives 2
31500 Toulouse
France
Tel. (33) 561 61 63 33
Fax (33) 561 26 46 10
E-mail: archives@mairie-toulouse.fr

Antoni Borfo Bach
Universitat Autònoma de Barcelona
Edifici A Rectorat
08193 Cerdanyola del Vallès
Spain
Tel. (34) 935 81 20 27
Fax (34) 935 81 20 00
E-mail: Antoni.Borfo@uab.es

Karin Borgkvist Ljung
National Archives
Box 12541
102 29 Stockholm
Sweden
Tel. (46) 87 37 63 44
Fax (46) 87 37 64 74
E-mail: karin.borgkvist@riksarkivet.ra.se

Joaquim Borràs I Gòmez
Universitat Pompeu Fabra
Passeig de Circumval.lació, 8
08003 Barcelona
Spain
Tel. (34) 935 42 17 94
Fax (34) 935 42 24 51
E-mail: joaquim.borras@suni.upf.es

Nuria Bosom Palau
Ajuntament De Barcelona
Avda. Bogatell 17 bis
08005 Barcelona
Spain
Tel. (34) 932 21 94 44
Fax (34) 932 21 94 21
E-mail: nbosom@mail.bcn.es

Filip Boudrez
City Archives of Antwerp
Venusstraat 11
2000 Antwerpen
Belgium
Tel. (32) 32 06 94 15
Fax (32) 32 06 94 10
E-mail: filip.boudrez@sd.antwerpen.be

Martin Bould
National Archives of Norway
Postboks 4013 Ullevål Stadion.no
0806 Oslo
Norway
Tel. (47) 22 02 26 00
Fax (47) 22 23 74 89
E-mail: martin.bould@riksarkivaren.dep.no

David V. Bowen
Audata Ltd
30 Salisbury Road
CT2 7HH Canterbury
United Kingdom
Tel. (44) 12 33 66 02 43
Fax (44) 12 33 64 06 90
E-mail: bowenvt@surfaid.org

Frank Brady
European Commission
Office: JECL 3/44
1049 Brussels
Belgium
Tel. (32-2) 295 20 62
Fax (32-2) 296 10 95
E-mail: frank.brady@cec.eu.int

Joan Bravo I Pijoan
Dursi
Via Laietana 33
08003 Barcelona
Spain
Tel. (34) 933 06 54 00
Fax (34) 933 06 54 32
E-mail: jbravo@correu.gencat.es

Mª Berta Bravo Serrano
Consejería De Las Artes
C/ Amaniel, 31
28015 Madrid
Spain
Tel. (34) 915 80 96 07
Fax (34) 915 80 96 46
E-mail: berta.bravo@madrid.org

Manfred Rolf Brech
Europäische Zentralbank
Kaiserstrasse 29
60311 Frankfurt am Main Stadt Frankfurt
Deutschland
Tel. (49) 69 13 44 79 52
Fax (49) 69 13 44 66 40
E-mail: manfred.brech@ecb.int

Lars Brolén
Landsarkivet i Östersund
Arkivvägen 1
831 31 Östersund
Sweden
Tel. (46) 63 10 84 85
E-mail:
lars.brolen@landsarkivet-ostersund.ra.se

Gordon Brown
HATII
George Service House University Gardens
G12 8QQ Glasgow
United Kingdom
Tel. (44) 14 13 30 36 35
Fax (44) 14 13 30 37 88
E-mail: 01026336@student.gla.ac.uk

Nils Brübach
Archivschule Marburg
Bismarckstraße 32
35037 Marburg
Germany
Tel. (49) 642 11 69 71 23
Fax (49) 642 11 69 71 10
E-mail: Bruebach@mailer.uni-marburg.de

Martin Bryan
The Diffuse Project
29 Oldbury Orchard
GL3 2PU Churchdown
United Kingdom
Tel. (44) 14 52 71 40 29
E-mail: mtbryan@diffuse.org

Mark Buesser
Informatique-MTF SA
Route du Bleuet 1
1762 Givisiez
Switzerland
Tel. (41) 264 60 66 66
Fax (41) 264 60 66 60
E-mail: sandra.broillet@imtf.ch

Hartmut Axel Burghard
OPOCE
25 rue du 9 mai 1944
2112 Luxembourg
Luxembourg
E-mail: Hartmut.Burghard@cec.eu.int

Carlota Bustelo Ruesta
Inforárea, S.L.
Avda. Brasil, 17, 13ºI
28020 Madrid
Spain
Tel. (34) 915 97 13 83
Fax (34) 915 97 05 82
E-mail: correo@inforarea.es

Josep Maria Cabau Mateus
Migdia, 24
25141 Torregrossa
Spain
Tel. (34) 667 48 34 40
Fax (34) 973 17 07 43
E-mail: pmateus@eresmas.com

Teresa Cabré I Serrat
Diputacío De Barcelona
Rambla Catalunya, 126
08008 Barcelona
Spain
Tel. (34) 934 02 22 22
E-mail: cabrest@diba.es

José Miguel Cacho Herrera
OCS Software, S.A.
C/Estébanez Calderón, 3
28020 Madrid
Spain
Tel. (34) 915 79 08 00
Fax (34) 915 79 11 35
E-mail: jmcacho@ocs.es

Roser Calpe Andreo
Arxiu Municipal Sant Vicenç Dels Horts
Barcelona 16–18
08620 Sant Vicenç dels Horts
Spain
Tel. (34) 936 76 81 91
Fax (34) 936 56 45 64
E-mail: arxiu@svh.es

Arturo Camarero Gonzalez
Sociedad General de Archivos, S.A.
Pablo Alcover, 75–77
08017 Barcelona
Spain
Tel. (34) 932 12 35 58
Fax (34) 932 12 10 00
E-mail: javier.z@sga.es

Jesús Camarero Santamaria
Consejo Económico y Social
C/Huertas, 73
28014 Madrid
Spain
Tel. (34) 914 29 00 18
Fax (34) 913 69 05 55
E-mail: jesus.camarero@ces.es

Ian Campbell-Grant
CEN/Information Society Standardisation System
15, East Stratton Close
Forest Park
RG12 OXY Bracknell, Berkshire
United Kingdom
Tel. (44) 13 44 41 11 55
E-mail: ian.campbell-grant@icl.com

Isabel Campos Gonzalez
Generalitat de Catalunya
Plaça Sant Jaume, 4
08002 Barcelona
Spain
Tel. (34) 934 02 48 14
E-mail: icampos@presidencia.gencat.es

Francisco Campuzano Juan
Archivo Del Reino De Valencia
Paseo De La Alameda 22
46019 Valencia
Spain
Tel. (34) 963 18 45 33
Fax (34) 963 18 45 27
E-mail: fcampuzano.areino@cult.gva.es

Perrine Canavaggio
International Council on Archives
6n rue des Francs-Bourgeois
75003 Paris
Francia
Tel. (33) 140 27 61 10
Fax (33) 142 72 20 65
E-mail: canavaggio@ica.org

Montserrat Canela
Archives and Records Section, UNHCR
94, rue de Montbrillant
CP 2500
1211 Geneve-2
Switzerland
Tel (22) 89 42/22 or 739 78 41
Fax (22) 739 73 65
E-mail: canelaga@unhcr.ch

M. Carme Capdevila i Guarro
Generalitat de Catalunya
C. Ricard Calvet, s/n
25191 Lleida
Spain
Tel. (34) 973 20 66 57
Fax (34) 973 21 21 69
E-mail: fmata@presidencia.gencat.es

Emília Capell Garriga
Ajuntament de Barcelona
Pça Sant Jaume
08002 Barcelona
Spain
Tel. (34) 934 02 73 39
Fax (34) 933 17 01 39
E-mail: ecapell@mail.bcn.es

Mercè Carafí Rosell
Arxiu Ajuntament De Castellbisbal
Av. Pau Casals,9
08755 Castellbisbal
Spain
Tel. (34) 937 72 02 25
Fax (34) 937 72 13 07
E-mail: arxiu@castellbisbal.org

Teresa Cardellach I Gimenez
Ajuntament De Terrassa
Raval de Montserrat, 14-20
08221 Terrassa
Spain
Tel. (34) 937 39 70 00
Fax (34) 937 39 70 71
E-mail: Teresa.cardellach@Terrassa.org

Montserrat Carmona Fernández
Generalitat de Catalunya
Sepúlveda, 148-150
08011 Barcelona
Spain
Tel. (34) 932 28 57 57
Fax (34) 932 28 57 45
E-mail:
montserrat.carmona@correu.gencat.es

María Dolores Carnicer Arribas
Subdirección General de los Archivos Estatales
Plaza del Rey No 1, planta 0
28004 Madrid
Spain
Tel. (34) 917 01 71 27
E-mail: dolores.carnicer@dglab.mcu.es

Carme Carreño Pombar
Instituto del Teatro
Plaza M.Xirgu s/n
08004 Barcelona
Spain
Tel. (34) 932 27 39 00
E-mail: carrenopc@diba.es

Joaquim Carreras I Barnes
Serveil d'Arxius
Portaferrissa, 1
08002 Barcelona
Spain
Tel. (34) 933 16 27 54
Fax (34) 933 16 27 41
E-mail: jcarrerasb@correu.gencat.es

Susana Carrillo
Pintor Togores 39 1 3
08111 Cerdanyola del Valles
Spain
Tel. (34) 936 91 42 48
E-mail: susanacarmor@hotmail.com

Miquel Casademont Donay
Universitat De Girona
Plaça de Sant Domènec, 3
17071 Girona
Spain
Tel. (34) 972 41 80 74
Fax (34) 972 41 80 31
E-mail: miquel.casademont@pas.udg.es

Lluis-Esteve Casellas Serra
Arxiu Municipal de Girona
Plaça del vi, 1
17004 Girona
Spain
Tel. (34) 972 41 94 34
Fax (34) 972 41 94 01
E-mail: lecasellas@ajgirona.org

Paola Casini
European Economic and Social Committee
Rue Ravenstein, 2
1000 Brussels
Belgium
Tel. (32-2) 546 93 87
Fax (32-2) 513 48 93
E-mail: paola.casini@esc.eu.int

Amelia Castan I Ranch
Arxiu Nacional de Catalunya
Carrer Jaume I, 33-51
08190 Sant Cugat del Vallés
Spain
Tel. (34) 935 89 77 88
Fax (34) 935 89 80 35

Gustavo Castañer Marquardt
Archivo General de la Administración
Paseo de Aguadores, 2
28804 Alcalá de Henares
Spain
Tel. (34) 918 89 29 50 ext. 60
Fax (34) 918 82 24 35
E-mail: chembryl@hotmail.com

Donatella Castelli
Italian National Research Council (CNR)
Area della Ricerca CNR di Pisa
Via G. Moruzzi, 1
56124 Pisa
Italy
Tel. (39) 05 03 15 29 02
Fax (39) 05 03 15 28 10
E-mail: castelli@iei.pi.cnr.it

Lluís Cermeno I Martorell
Servei d'Arxius
Portaferrissa 1
08002 Barcelona
Spain
Tel. (34) 933 16 27 54
Fax (34) 933 16 27 41
E-mail: lcermeno@correu.gencat.es

Gäel Chumiatcher
Parlement européen
Service des archives, ADG 01A005
Plateau du Kirchberg
2929 Luxembourg
Luxembourg
Tel. (352) 430 02 43 52
Fax (352) 43 94 93
E-mail: auxadg4@europarl.eu.int

Alexander Church
EBRD — European Bank for Reconstruction and Development
1 Exchange Square
EC2A 2JN London
United Kingdom
Tel. (44-20) 73 38 73 16
Fax (44-20) 73 38 79 38
E-mail: burninge@ebrd.com

Enric Cobo Barri
Departament Política Territorial i Obres Públiques
Av.Tarradellas, 2–8
08029 Barcelona
Spain
Tel. (34) 935 67 47 52
E-mail: WECOBO@correu.gencat.es

Joan Comasòlivas i Font
Arxiu Històric de Sabadell
Indústria, 32–34
08202 Sabadell
Spain
Tel. (34) 937 26 87 77
Fax (34) 937 27 57 03
E-mail: jcomasolivas@ajsabadell.es

Immaculada Comes Bosch
Ajuntament de Manresa
Pl. Major, 1
08240 Manresa
Spain
Tel. (34) 938 78 23 00
Fax (34) 938 78 23 07
E-mail: icomes@manresa.org

Montserrat Compte I Farrés
Diputació De Barcelona
Mejía Lequerica, 1
08028 Barcelona
Spain
Tel. (34) 934 02 24 46
Fax (34) 934 02 25 53
E-mail: comptefm@diba.es

Xavier Conchillo Roca
Arxiu Municpal Castell-Platja d'Aro
Madrid 75
Sant Feliu de Guixols
Spain
Tel. (34) 972 32 13 90
Fax (34) 972 82 65 39
E-mail: xconc000@yahoo.es

María Luisa Conde Villaverde
Archivo General de la Administración
Paseo de Aguadores, 2
28801 Alcalá de Henares
Spain
Tel. (34) 918 88 02 96
E-mail: aga@cult.mec.es

Alicia Conesa Santamaria
Televisió de Catalunya, S.A.
de TV3
08970 Sant Joan Despí
Spain
Tel. (34) 934 99 94 26
Fax (34) 934 73 14 38
E-mail: aconesa.g@tvcatalunya.com

Mariona Corominas i Noguera
Parlament de Catalunya
Parc de la Ciutadella, s/n
08003 Barcelona
Spain
Tel. (34) 933 04 65 00
Fax (34) 933 04 65 49
E-mail: mcorominas@parlament-cat.es

Piero Corsini
IBM
New Square
Bedfont Lakes
TW14 8HB Feltham, Middlesex
United Kingdom
Tel. (44) 20 88 18 52 44
E-mail: piero_corsini@uk.ibm.com

Immaculada Costa i Viarnés
Arxiu Històric de Girona
Plaça de Sant Josep, 1
17004 Girona
Spain
Tel. (34) 972 22 55 00
Fax (34) 972 22 75 77
E-mail: immaviar@grn.es

Lucia Costanilla Baquedano
Archivo General UPNA
Campus Arrosadía s/n
Pamplona
Spain
Tel. (34) 948 16 89 35
E-mail: lucia.costanilla@unavarra.es

Montse Cots Martínez
Milcap media group
Rambla del Poblenou, 75
08005 Barcelona
Spain
Tel. (34) 933 00 47 08
E-mail: montse.cots@milcap.es

Juan Carlos Cruellas Ibarz
Universidad Politécnica Catalunya
Jordi Girona, 1-D6-103
08034 Barcelona
Spain
Tel. (34) 934 01 67 90
Fax (34) 934 01 70 55
E-mail: cruellas@ac.upc.es

Jose Ramon Cruz Mundet
Universidad Carlos III
Madrid 126
28903 Getafe
Spain
Tel. (34) 916 24 92 44
Fax (34) 916 24 92 12
E-mail: jrcmar@uc3m.es

Ian Cumming
European Training Foundation
Viale Settimio Severo 65
10133 Turin
Italy
Tel. (39) 011 63 02 22 99
E-mail: ian.cumming@etf.eu.int

James Currall
University of Glasgow
James Watt North Building
University Avenue
G12 8QQ Glasgow
United Kingdom
Tel. (44) 14 13 30 48 55
Fax (44) 14 13 30 48 08
E-mail: J.Currall@compserv.gla.ac.uk

Noemi Dachs Lecina
Ajuntament de Barcelona
Creu Coberta 104
08014 Barcelona
Spain
Tel. (34) 932 91 63 05
Fax (34) 932 91 63 33
E-mail: ndachs@mail.bcn.es

Adela Dalós-Moner Vilar
DOC6
Mallorca 272–276
08006 Barcelona
Spain
Tel. (34) 932 15 43 13
Fax (34) 934 88 36 21
E-mail: aalos@doc6.es

Elizabeth Danbury
School of Library, Archive and Information
Studies
University College London
Gower Street
WC1E 6BT London
United Kingdom
Tel. (44-20) 76 79 24 73
Fax (44-20) 73 83 05 57
E-mail: uczceda@ucl.ac.uk

Jan Danielsen
Danish National Archives
Rigsdagsgarden 9
1218 Copenhagen
Denmark
Tel. (45) 33 92 23 58
Fax (45) 33 95 42 11
E-mail: Jd@rasa.dk

Rosa M Datsira Bonet
Prat de la Riba 94
08211 Castellar del Vallés
Spain
Tel. (34) 937 14 45 23
E-mail: pompeia@wanadoo.es

Robert Davies
MDR Partners
2b Fitzgerald Road
SW14 8HA Mortlake
United Kingdom
Tel. (44-20) 88 76 31 21
E-mail: rob.davies@mdrpartners.com

Richard Davis
University of London Computer Centre
20 Guilford Street
WC1N1DZ London
United Kingdom
Tel. (44-20) 76 92 13 50
Fax (44-20) 76 92 12 34
E-mail: r.davis@ulcc.ac.uk

Martine De Boisdeffre
Direction des archives de France
56, rue des Francs-Bourgeois
75141 cedex 03 Paris
France
Tel. (33) 140 27 67 00
Fax (33) 140 27 66 06

Jesús Javier De Diego Erles
Instituto de Estudios Altoaragoneses
C/ Parque 10
22002 Huesca
Spain
Tel. (34) 974 29 41 20
Fax (34) 974 29 41 22
E-mail: cipca@iea.es

Xavier De Hennin
EADP
Franklin Delano Roosvelt 127
1050 Brussels
Bélgica
Tel. (32-2) 646 30 60
Fax (32-2) 646 36 37
E-mail: mailbox@eadp.org

Pilar De Luna Polo
Ajuntament de Barcelona
Ronda Guinardó 49
08024 Barcelona
Spain
Tel. (34) 932 91 67 26
Fax (34) 932 91 67 67
E-mail: pluna@mail.bcn.es

Elisa De Santos Canalejo
Subdirector General de los Archivos Estatales
Plaza del Rey 1, 4a Planta
28071 Madrid
Spain
Tel. (34) 917 01 72 64
Fax (34) 915 21 05 08
E-mail: Elisa.desantos@dglab.mcu.es

Mercedes de Vega
Sria. de Relaciones Exteriores
Ricardo Flores Magón 1
06995 Tlatelolco
México
Tel. (52) 51 17 21 21
Fax (52) 53 27 31 83
E-mail: mdevega@sre.gob.mx

Pilar de Vicente Coll
Departament de Política Territorial i Obres Públiq
Av.Josep Tarradellas, 2-6
08029 Barcelona
Spain
Tel. (34) 935 67 48 21
E-mail: wvicente@correu.gencat.es

Bruno Degaichia
European Commission, Agriculture DG
European Commission, L. 130. 3/241
1049 Brussels
Belgium
Tel. (32-2) 295 19 88
Fax (32-2) 295 01 30
E-mail: bruno.degaichia@cec.eu.int

Rocío del Casar Ximénez
Ruiz de Alarcón, 23
28014 Madrid
Spain
Tel. (34) 913 30 29 35
Fax (34) 913 30 28 50
E-mail: rocio.delcasar@prado.mcu.es

Dominique Delaunay
Institut Universitaire européen
Badia Fiesolana
50016 San-Domenico-Di-Fiesole
Italy
Tel. (39) 05 54 68 56 07
E-mail: dominique.delaunay@iue.it

Diane Denham
Foreign and Commonwealth Office
HP 71/007 Hanslope Park
MK19 7BH Milton Keynes
United Kingdom
Tel. (44) 19 08 51 50 83
Fax (44) 19 08 51 56 92
E-mail: diane.denham@fco.gov.uk

Ineke Deserno
International Olympic Committee
1, Quai d'Ouchy
1001 Lausanne
Switzerland
E-mail: ineke.deserno@olympics.org

Montserrat Deu i Prats
Escola Universitària Salesiana de Sarrià
Joan Maragall, 138 1er. 1a.
08980 St. Feliu de LLobregat
Spain
Tel. (34) 936 85 50 87
E-mail: mdeu@euss.es

Catherine Dhérent
Direction des archives de France
Rue des Francs-Bourgeois, 56
75141 cedex 03 Paris
France
Tel. (33) 140 27 62 65
Fax (33) 140 27 66 36
E-mail: Catherine.dherent@culture.gouv.fr

Maria Dias Andrade
CP — Caminhos de Ferro Portugueses
Av. da Republica, 66 — 9º-
1010-096 L Lisboa
Portugal
Tel. (351) 217 98 04 76
E-mail: mnandrade@mail.cp.pt

Maria José Diaz Martinez
Sony España,S.A.
Pol.Ind.Can Mitjans, S/N
08232 Viladecavalls
Spain
Tel. (34) 937 00 82 84
Fax (34) 937 00 81 06
E-mail: mariajose.diaz@eu.sony.com

Viktoras Domarkas
Lithuanaian archives department
Mindaugo 8
2009 Vilnius
Lithuania
Tel. (370) 265 22 49
Fax (370) 265 23 14
E-mail: v.domarkas@archyvai.lt

Joan Domingo Basora
Departament de Cultura/Generalitat de Catalunya
Rbla. de Sta. Mònica, 8
08002 Barcelona
Spain
Tel. (34) 933 16 27 05
Fax (34) 933 16 27 01
E-mail: jdomingo@correu.gencat.es

Raquel Dominguez
Hospital de Bellvitge-Documentació Clínica i Arxiu
Feixa Llarga, S/N.
08907 L'Hospitalet de Llobregat
Spain
Tel. (34) 932 60 75 24
Fax (34) 932 60 74 51
E-mail: mdsenra@terra.es

Peter Doorn
Head of Information Technology
P.O. Box 95110
1090 HC Amsterdam
The Netherlands
Tel. (31) 204 62 86 07
Fax (31) 206 68 50 79
E-mail: Peter.doorn@niwi.knaw.nl

Karl-Magnus Drake
National Archives of Sweden
P.O.Box 12541
10229 Stockholm
Sweden
Tel. (46) 87 37 63 50
Fax (46) 87 37 64 74
E-mail: karl-magnus.drake@riksarkivet.ra.se

Vladimir Drobnjak
Pokrajinski arhiv Koper
Krozna 4
6000 Koper
Slovenia
Tel. (38) 656 28 40 68
E-mail: vladimir.drobnjak@guest.arnes.si

Nestor Dubikin
Soluziona
Travesía del tornado, 6 1ºD
28760 Tres Cantos
Spain
Tel. (34) 918 04 17 13
E-mail: ndubikin@softpoint.com.uy

Wendy Duff
Faculty of Information Studies, University of Toronto
140 St. George St.
M5S 3 G6 Toronto, Ontario
Canada
Tel. (416) 978 31 52
Fax (416) 971 13 99
E-mail: duff@fis.utoronto.ca

Jos Dumortier
Interdisciplinary Centre for Law and Information Technology
Catholic University of Leuven
Tiensestraat 41
3000 Leuven
Belgium
Tel. (32) 16 32 51 49
Fax (32) 16 32 54 38
E-mail: jos.dumortier@law.kuleuven.ac.be

Ana Duplá Del Moral
Consejería de las Artes
Pza. de España, 8
28008 Madrid
Spain
Tel. (34) 915 80 24 84
Fax (34) 915 80 24 81
E-mail: ana.dupla@madrid.org

Luciana Duranti
School of Library, Archival and Information studies
831-1956 Main Mall
University of British Columbia
B.C. V6T 1Z1 Vancouver
Canada
Tel. (604) 822 25 87
Fax (604) 822 60 06
E-mail: luciani@unixg.ubc.ca

Ann Durniat
Engineering and Physical Sciences Research Council
Polaris House, North star avenue,
SN5 1ET Swindon
United Kingdom
Tel. (44) 17 93 44 42 78
E-mail: ann.durniat@epsrc.ac.uk

Niklas Eden
Sundsvalls kommun
Norrmalmsgatan 4
85185 Sundsvall
Sweden
Tel. (46) 60 19 14 78
Fax (46) 60 59 17 78
E-mail: niklas.eden@sundsvall.se

Myriam Eibar de la Encina
Generalitat valenciana
c/colon, 66
46004 Valencia
Spain
Tel. (34) 961 96 10 64
E-mail: eibar_myr@gva.es

Maria Eka
Ministry of Culture
Drottninggatan 16
103 33 Stockholm
Sweden
Tel. (46) 84 05 45 75
Fax (46) 84 05 13 12
E-mail: maria.eka@culture.ministry.se

Maribel Elizalde Frez
Fòrum Universal de les Cultures Barcelona 2004
Rambla del Poblenou, 75, àtic 1a
08005 Barcelona
Spain
Tel. (34) 933 00 47 08
E-mail: melizalde@barcelona2004.org

Frölander Elke
AstraZeneca AB
Byggn. 410
15185 Södertälje
Sweden
Tel. (46) 855 32 32 49
Fax (46) 855 32 86 90
E-mail: elke.frolander@astrazeneca.com

Myriam Olga Erwin
Archives fédérales suisses
Rue de la Prulay 7
1217 Meyrin
Suisse
Tel. (41) 227 82 11 32
Fax (41) 227 82 11 32
E-mail: m.o.erwin@gmx.net

Pilar Faci I Lacasta
AH Lleida
Sant Antoni Maria Claret 5
25002 Lleida
Spain
Tel. (34) 973 27 08 67
Fax (34) 973 27 31 60

Nuri Faig i Vila
c/ Sant Hipòlit, 29-31, 1er A
17003 Girona
Spain
Tel. (34) 972 21 17 83
E-mail: nfaigvila@yahoo.com

Júlia Fajardo García
Diputació de Barcelona
Rambla de Catalunya, 126
08008 Barcelona
Spain
Tel. (34) 934 02 22 48
Fax (34) 934 02 22 92
E-mail: fajardogj@diba.es

Joan Farré I Viladrich
AHC Balaguer
Plaça Mercadal 1, 4rt
25600 Balaguer
Spain
Tel. (34) 973 44 52 00
Fax (34) 973 44 86 21
E-mail: ahcbal@tinet.org

Emmy Ferbeek
Gemeentearchief Amsterdam
Maasstraat 77III
1078 HE Amsterdam
The Netherlands
Tel. (31) 206 70 81 95
Fax (31) 206 75 05 96
E-mail: erik.tempels@wxs.nl

Josep Fernandez I Trabal
Arxiu Nacional de Catalunya
Carrer Jaume I, 33-51
08190 Sant Cugat del Vallés
Spain
Tel. (34) 935 89 77 88
Fax (34) 935 89 80 35

Victoria Fernández Mera
Congreso de los Diputados
Floridablanca s/n
28071 Madrid
Spain
Tel. (34) 913 90 63 39
Fax (34) 913 90 60 59
E-mail: victoria.fernandez@sgral.congreso.es

Teresa Ferrer Fontanet
Generalitat de Catalunya
Via Laietana, 2
08002 Barcelona
Spain
Tel. (34) 932 30 37 11
Fax (34) 932 30 46 54
E-mail: arxiu1.civil.bcn@gencat.es

Johan Florén
EMT Interbase AB
Bryggavägen 121-125
178 21 Ekerö
Sweden
Tel. (46) 856 03 77 00
Fax (46) 856 03 77 99
E-mail: johanf@emt.se

Assumpta Fluvià i Figueras
Organisme de Gestió Tributària de la
Diputació de Barcelona
c/ Mejía Lequerica, 1
08013 Barcelona
Spain
Tel. (34) 934 02 26 79
E-mail: fluviaas@diba.es

José Fonseca
National Archives of Sweden
Sandslån 3140
87052 Nyland
Sweden
Tel. (46) 61 25 05 80
Fax (46) 61 25 00 18
E-mail: jose.fonseca@svar.ra.se

Pilar Frago Pérez
Arxiu Nacional de Catalunya
Casp 114, esc. B, 6è, 1a.
08013 Barcelona
Spain
Tel. (34) 932 45 46 22
E-mail: pmarti13@pie.xtec.es

Marc Fresko
Cornwell Affiliates plc
Home Barn Court
The Street
KT24 5LG Effingham, Surrey
United Kingdom
Tel. (44) 13 72 45 60 86
Fax (44) 13 72 45 09 50
E-mail: marc.fresko@cornwell.co.uk

David Fuegi
MDR Partners
11 Sanders Drive
CO3 3SE Colchester
United Kingdom
Tel. (44) 12 06 57 74 36
E-mail: david.fuegi@mdrpartners.com

Hiromichi Fujisawa
Central Research Laboratory, Hitachi Ltd
1-280 Higashi-koigakubo
Kokubunji
185-8601 Tokyo
Japan
Tel. (81) 423 23 11 11-2027
Fax (81) 42 327 77 00
E-mail: fujisawa@crl.hitachi.co.jp

Per Furberg
Setterwalls Advokarbyrå AB
Kunstorget 2
PO Box 11235
404 25 Gothenborg
Sweden
Tel. (46) 317 01 17 00
Fax (46) 317 01 17 01
E-mail: Per.Furberg@setterwalls.se

Pierre Fuzeau
Serda
24, Rue de Milan
75009 Paris
France
Tel. (33) 144 53 45 00
Fax (33) 144 53 45 01
E-mail: pierre.fuzeau@serda.com

Tomas Galano
Union de l'Europe Occidentale
Rue de l'Association Num 15
1000 Brussels
Belgique
Tel. (32-2) 500 44 80
Fax (32-2) 500 44 70
E-mail: ueo.registry@skynet.be

Olga Garcia I Cruz
Diputació de Barcelona
Mejía Lequerica, 1
08028 Barcelona
Spain
Tel. (34) 934 02 24 46
Fax (34) 934 02 25 53
E-mail: Garciaco@diba.es

Soraya Garcia Merino
EBRD — European Bank for Reconstruction and Development
1 Exchange Square
EC2A 2JN London
United Kingdom
Tel. (44) 20 73 38 73 16
Fax (44) 20 73 38 79 38
E-mail: garcias@ebrd.org

Alfred Garcia Puig
Departament Industria, Comerç I Turisme
Passeig de Gracia 105
08022 Barcelona
Spain
Tel. (34) 934 84 97 23
Fax (34) 934 84 98 22
E-mail: agarciap@correu.gencat.es

Peter Garrod
University of London Computer Centre
20 Guilford Street
WC1N 1DZ London
United Kingdom
Tel. (44) 20 76 92 13 53
Fax (44) 20 76 92 12 34
E-mail: p.garrod@ulcc.ac.uk

Magnus Geber
National Archives
P.O. Box 12541
10229 Stockholm
Sweden
Tel. (46) 702 56 13 17
Fax (46) 87 37 64 74
E-mail: magnus.geber@riksarkivet.ra.se

Mario Gene Amoros
Ajuntament de Balaguer
Spain

Estefania Genovès Salvadó
Audiència Provincial de Tarragona
Av. Lluís Companys 10
43005 Tarragona
Spain
Tel. (34) 977 23 92 58
Fax (34) 977 23 10 89
E-mail: trg.arxiuaudiencia@xij.gencat.es

Alison Gibney
Cimtech Ltd
45 Grosvenor Rd
AL1 3AW St Albans
United Kingdom
Tel. (44) 17 27 81 36 51
Fax (44) 17 27 81 36 49
E-mail: a.gibney@herts.ac.uk

Maria José Gil
Onoma — Gabinete de Traduções, Lda
Rua Coelho da Rocha, nº 71 — 2º Dto
1350-073 Lisbon
Portugal
Tel. (351) 213 92 95 90
Fax (351) 213 92 95 99
E-mail: geral@onoma.pt

Rosa Maria Gil Tort
Collegi d'Arquitectes Catalunya
Pl. Catedral, 8 Pia Almoina
17004 Girona
Spain
Tel. (34) 972 41 28 92
Fax (34) 972 21 41 51
E-mail: arxiuh.gir@coac.es

Mercè Giménez Caparrós
Convergència Democràtica de Catalunya
Còrsega, 333
08037 Barcelona
Spain
Tel. (34) 932 36 31 00
Fax (34) 932 36 31 10
E-mail: m.gimenez@convergencia.org

Vicent Giménez Chornet
Arxiu del Regne de València
Passeig Albereda, 22
46010 València
Spain
Tel. (34) 963 18 45 50
Fax (34) 963 18 45 27
E-mail: vgimenez.areino@cult.gva.es

Natalija Glazar
Archives of the Republic of Slovenia
Zvezdarska 1, p.p.21
1127 Ljubljana
Slovenia
Tel. (386) 12 41 42 12
Fax (386) 12 41 42 69
E-mail: Natalija.Glazar@gov.si

Isabel Gómez de León Contreras
Cámara de Cuentas de Andalucía
Av. Ramón y Cajal, 35
41005 Sevilla
Spain
Tel. (34) 955 00 96 57
Fax (34) 955 00 96 42
E-mail: isabelgomez@ccuentas.es

Ma. Pau Gómez I Ferrer
AHC Vall D'Aran
Major s/n
25537 Arrós
Spain
Tel. (34) 973 64 25 69
Fax (34) 973 64 25 69

Paloma Gómez Pastor
Banco de España
Alcalá 522
28027 Madrid
Spain
Tel. (34) 913 38 64 12
Fax (34) 913 38 68 82
E-mail: rrhh.desarrollo@bde.es

Víctor González Monedero
Cortes de Aragón
Palacio de la Aljafería sn
50071 Zaragoza
Spain
Tel. (34) 976 28 95 96
Fax (34) 976 28 96 32
E-mail: vgonzale@cortesaragon.es

José Manuel González Sánchez
Sub. Gral. Archivos Estatales
Plaza del Rey, 1
28071 Madrid
Spain
Tel. (34) 917 02 70 00-32909
Fax (34) 917 01 71 32
E-mail: jmanuel.gonzalez@dglabr.mcu.es

Angeles Gonzalez Vera
Ayuntamiento de Viladecans. Archivo Administrativo
Guifré el Pelós, 32-36
08840 Viladecans
Spain
Tel. (34) 936 37 37 84
Fax (34) 936 59 15 03
E-mail: agonzalezv@aj-viladecans.es

Sandrine Gorgon
Permanent representation of Luxemburg to the European Union
75 avenue de Cortenbergh
1000 Brussels
Belgium
Tel. (32-2) 737 56 30
Fax (32-2) 737 56 10
E-mail: sandrine.gorgon@rpue.etat.lu

Hans Goutier
Ministry of Transport, Public Works and Water Management
P.O. Box 20906
2500 EX The Hague
The Netherlands
Tel. (31) 703 51 85 36
E-mail: hans.goutier@cend.minvenw.nl

Claes Gränström
Swedish National Archives
Riksarkivet
PO Box 12541
10229 Stockholm
Sweden
Tel. (46) 87 37 63 97
Fax (46) 87 37 64 74
E-mail: Claes.granstrom@riksarkivet.ra.se

Rosa María Grau Guadix
Congreso de los Diputados
Floridablanca s/n
28071 Madrid
Spain
Tel. (34) 93 90 70 02
Fax (34) 91 390 67 65
E-mail: rosa.grau@sgral.congreso.es

Frances Grey
House of Lords Record Office
House of Lords
SW1A 0PW London
United Kingdom
Tel. (44-20) 72 19 01 00
Fax (44-20) 72 19 25 70
E-mail: greyf@parliament.uk

Suzana Grubesic
Croatian State Archives
Marulicev Trg 21
10000 Zagreb
Croatia
Tel. (385) 14 80 19 29
Fax (385) 14 82 90 00
E-mail: suzana@arhiv.hr

Magnus Gudmundsson
University of Iceland
Adalbygging
101 Sundurgata
Iceland
Tel. (354) 525 43 71
Fax (354) 552 13 31
E-mail: magnusg@hi.is

Maria Guercio
Universitá degli studi di Urbino
Piazza Verbano 8
00199 Rome
Italy
Tel. (39) 06 85 30 20 04
Fax (39) 06 85 30 20 04
E-mail: mc9468@mclink.it

Mercedes Guijarro Anton
Universidad de Alicante
Apartado de Correos,99
03080 Alicante
Spain
Tel. (34) 965 90 37 44
Fax (34) 965 90 37 35
E-mail: M.Guijarro@ua.es

Margarida Guijarro Margalef
Institut Català del Sòl
Còrsega, 289
08008 Barcelona
Spain
Tel. (34) 932 28 60 17
Fax (34) 932 28 60 58
E-mail: m_guijarro@correu.gencat.es

Rolf Hagstedt
City Archives of Stockholm
P.O. Box 220 63
104 22 Stockholm
Sweden
Tel. (46) 850 82 83 02
Fax (46) 850 82 83 01
E-mail: rolf.hagstedt@ssa.stockholm.se

Anttila Hannele
Ministry of Agriculture and Forestry
Box 30
00023 Government
Finland
Tel. (358) 916 01
Fax (358) 91 60 42 02
E-mail: hannele.anttila@mmm.fi

Birgit Hansen
Danish State Archives, Rigsarkivet
Rigsdagsgaarden 9
1218 København K
Denmark
Tel. (45) 33 92 23 59
E-mail: Bih@ra.sa.dk

Johanna Hartman
Ministry of Social Affairs and Health
PO BOX 33
0023 G Helsinki
Finland
Tel. (358) 916 07 49 90
Fax (358) 916 07 41 26
E-mail: johanna.hartman@stm.vn.fi

Jon Atle Haugen
National Archives of Norway
Postboks 4013 Ullevål stadion
0806 Oslo
Norway
Tel. (47) 22 02 26 73
Fax (47) 22 23 74 89
E-mail: jon.haugen@riksarkivaren.dep.no

Jim Hayes
Action Point Software
Nymphenburger str. 136
München
Germany
Tel. (49) 17 25 11 58 17
E-mail: jhayes@actionpoint.com

Margaret Hedstrom
School of Information, University of Michigan
3082 West Hall
550 East University
MI 48109 Ann Arbor
United States
Tel. (1-734) 647 35 82
Fax (1-734) 764 24 75
E-mail: hedstrom@umich.edu

Cecilia Henriques
Instituto dos Arquivos Nacionais
Alameda da Universidade
1649-010 Lisboa
Portugal
Tel. (351) 217 81 15 00
Fax (351) 217 93 72 30
E-mail: cecilia@iantt.pt

Pekka Henttonen
Military Archives
PB 54
00581 Helsinki
Finland
Tel. (358) 918 12 65 02
Fax (358) 918 12 65 12
E-mail: pekka.henttonen@sota-arkisto.fi

Ivar Herø
Telenor Focus
Engaveien 31
2002 Lillestrøm
Norway
Tel. (47) 91 30 68 90
E-mail: ivar-melhuus.hero@telenor.com

Susana Herreros Lopetegui
Archivo Genral de Navarra
Avda. Carlos III, 2
31002 Pamplona
Navarra
Tel. (34) 948 42 19 28
Fax (34) 948 42 13 28
E-mail: sherrerl@cfnavarra.es

Helen Victoria Heskes
Rijksarchiefinspectie
Postbus 11583
2502 AN Den Haag
The Netherlands
Tel. (31) 703 33 91 35
Fax (31) 703 33 91 30
E-mail: hheskes@rijksarchiefinspectie.nl

Heather Heywood
International Telecommunication Union
Place des Nations
1211 Geneva 20
Switzerland
Tel. (41) 227 30 67 71
Fax (41) 227 30 53 26
E-mail: heather.heywood@itu.int

Barbara Hoen
Baden-Württemberg
Eugenstraße 7
70182 Stuttgart
Germany
Tel. (49) 42 31 64 62 26
Fax (49) 42 30 64 62 26
E-mail: Hoen@lad-bw.de

Johannes Hofman
Ministry of the Interior of the Netherlands
PO BOX 20011
2500 EA Den Haag
The Netherlands
Tel. (31) 704 26 61 41
Fax (31) 704 26 76 00
E-mail: Hans.hofman@minbzk.nl

Hans Hofmann
European Commission
Office JECL 3/38
1049 Brussels
Belgium
Tel. (32-2) 295 20 53
Fax (32-2) 296 10 95
E-mail: hans.hofmann@cec.eu.int

Christopher Holgate
Viale Settimio Severo 65
10133 Turin
Italy
Tel. (39) 01 16 30 23 27
E-mail: christopher.holgate@etf.eu.int

Torbjörn Hörnfeldt
The National Archives of Sweden
Box 12541
102 29 Stockholm
Sweden
Tel. (46) 87 37 64 93
Fax (46) 87 37 64 74
E-mail: torbjorn.hornfeldt@riksarkivet.ra.se

Ann Hörsell
RAÄ, National Heritage Board of Sweden
Box 5405
114 84 Stockholm
Sweden
Tel. (46) 851 91 80 53
Fax (46) 851 91 80 88
E-mail: ann.horsell@raa.se

Peter Horsman
Archiefschool, Netherlands Institute for
Archival Education and Training
Weesperzijde, 190
1097 DZ Amsterdam
The Netherlands
Tel. (31) 205 95 19 00
Fax (31) 205 95 19 20
E-mail: p.j.horsman@ai.hva.nl

Montserrat Hosta I Rebes
AH Girona
Plaça de Sant Josep, 1
17004 Girona
Spain
Tel. (34) 972 22 55 00
Fax (34) 972 22 75 77

Christian Ihle
Ministery of Finance
Hintere Zollamtstrasse 4
1030 Viena
Austria

Yannis Ioannidis
Department of Informatics and
Telecommunications, University of Athens
TYPA Buildings
Panepistimioupolis
157 84 Ilissia, Athens
Greece
Tel. (30) 17 27 52 24
Fax (30) 17 27 52 14
E-mail: yannis@di.uoa.gr

Stephane Ipert
Centre de Conservation du Livre
18 rue de la Calade
13200 Arles
France
Tel. (33) 490 49 99 89
Fax (33) 490 49 66 11
E-mail: s.ipert@libertysurf.fr

Javed Iqbal
The British Council
Aleje Jerozolimskie 59
00-697 Warsaw
Poland
Tel. (48) 225 85 10 66
Fax (48) 225 85 10 64
E-mail: javed.iqbal@britishcouncil.pl

Salvatore Italia
Direzioni generale degli archivi
Ministero beni e attivitá culturali
via Gaeta 8a
00185 Rome
Italy
Tel. (39) 06 4 94 14 64
Fax (39) 064 88 23 58
E-mail: sitalia@archivi.beniculturali.it

Jozo Ivanoviæ
Croatian State Archives
Maruliæev trg 21
10000 Zagreb
Croatia
Tel. (385) 14 80 19 47
E-mail: jivanovi@arhiv.hr

Santiago Izquierdo Gonzalez
Ayto. S.S. De Los Reyes
Pza. Constitucion, 1
28700 Reyes
Spain
Tel. (34) 916 59 71 20
Fax (34) 916 54 58 48
E-mail:
sizquierdog@ayuntamiento.sanse.info

Cathy Jager
Algemeen Rijksarchief
Prins Willem Alexanderhof 20
2595 Be Den Haag
The Netherlands
Tel. (31) 703 31 55 03
Fax (31) 703 31 55 40
E-mail: cathy.jager@ara.archief.nl

Per Jansson
National Archives of Sweden
Box 12541
102 29 Stockholm
Sweden
Tel. (46) 87 37 63 50
Fax (46) 87 37 64 74
E-mail: per.jansson@riksarkivet.ra.se

Moulin Jean
Service d'Information S&T
Bd de l'Empereur, 4
1000 Brussels
Belgium
Tel. (32-2) 519 56 56-5667
Fax (32-2) 519 56 45
E-mail: jean.moulin@stis.fgov.be

Maria Angels Jiménez
Ipsen Pharma S.A.
Laureà Miró, 395
08980 Sant Feliu de Llobregat
Spain
Tel. (34) 936 85 81 00
Fax (34) 936 85 10 11
E-mail:
m-angels.jimenez@beaufour-ipsen.com

Marià Jiménez Albaladejo
Consorci Sanitari de Barcelona
Calàbria, 169
08015 Barcelona
Spain
Tel. (34) 932 26 51 42
Fax (34) 932 26 82 39
E-mail: mjimenez@rsbcn.scs.es

Joan Antoni Jiménez Castillo
Servei Català de la Salut
Travessera de les Corts, 131, edifici Olímpia
08028 Barcelona
Spain
Tel. (34) 934 03 87 35
Fax (34) 934 03 89 00
E-mail: jjimenez@catsalut.net

Per Johansson
Dept. 2110 VBBVN
405 08 Gothenborg
Sweden
Tel. (46) 313 22 05 92
Fax (46) 31 66 59 55
E-mail: Per.johansson@volvo.com

Claire Johnson
University of Glasgow
77-87 Dumbarton Rd. Glasgow University
Archive Services
G11 6PW Glasgow
United Kingdom
Tel. (44) 14 13 30 64 94
Fax (44) 14 13 30 41 58
E-mail: C.Johnson@archives.gla.ac.uk

Rosalind Johnson
CILIP(formerly known as Library Association)
7 Ridgmount
WC1E 7AE London
United Kingdom
Tel. (44-20) 72 55 06 39
Fax (44-20) 72 55 05 01
E-mail: victoria.anibueze@cilip.org.uk

Britta Jonell Ericsson
National Archives of Sweden
Box 12541
102 29 Stockholm
Sweden
Tel. (46) 87 37 63 50
Fax (46) 87 37 64 74
E-mail: britta.jonellericsson@riksarkivet.ra.se

Bjorn Jordell
City Archives of Stockholm
Box 22063
104 22 Stockholm
Sweden
Tel. (46) 850 82 83 07
Fax (46) 850 82 83 01
E-mail: bjorn.jordell@ssa.stockholm.se

Werzinger Jörg
Sächsisches Staatsministerium des Innern
Wilhelm-Buck-Straße 2
01097 Dresden
Germany
Tel. (49) 35 15 64 34 46
Fax (49) 35 15 64 34 09
E-mail: joerg.werzinger@smi.sachsen.de

Helle Jukka
City of Espoo
Virastopiha 2 C
02770 Espoo
Finland
Tel. (358) 981 62 51 52
Fax (358) 981 62 55 64
E-mail: jukka.helle@espoo.fi

Toivo Jullinen
The National Archives of Estonia
J.Liivi 4
50409 Tartu
Estonia
Tel. (372) 514 76 72
Fax (372) 738 75 14
E-mail: Toivo.Jullinen@ra.ee

Markus Junker
Deutsches Forschungzentrum für Künstliche Intelligenz (DFKI)
P.O. Box 2080
67608 Kaiserslautern
Germany
Tel. (49) 63 12 05 34 54
Fax (49) 63 12 05 32 10
E-mail: markus.junker@dfki.de

Janina Jurado
Departament de Justicia Generalitat de Catalunya
Ramon Llull 4-6
Girona
Spain
Tel. (34) 972 41 11 47
Fax (34) 972 22 00 15
E-mail: janina-jurado@yahoo.es

Gulbins Jürgen
Ixos
Kapellenstr. 15
75210 Keltern
Germany
Tel. (49) 70 82 94 82 50
Fax (49) 70 82 94 82 51
E-mail: jg@gulbins.de

Borje Justrell
National Archives of Sweden
P.O. Box 12541
102 29 Stockholm
Sweden
Tel. (46) 87 37 63 88
Fax (46) 87 37 64 74
E-mail: borje.justrell@riksarkivet.ra.se

Max Kaiser
Austrian National Library
Josefsplatz 1
1015 Vienna
Austria
Tel. (43) 153 41 03 70
Fax (43) 153 41 03 40
E-mail: max.kaiser@onb.ac.at

Ulrich Kampffmeyer
Project Consult
Oderfelder Strasse, 17
20149 Hamburg
Germany
Tel. (49) 40 46 07 62 20
Fax (49) 40 460 762 29
E-mail:
Ulrich.kampffmeyer@project-consult.com

Alice Keefer
Universitat de Barcelona
Melcior de Palau 140
08014 Barcelona
Spain
Tel. (34) 934 03 59 74
Fax (34) 934 03 57 72
E-mail: akeefer@fbd.ub.es

Kirsi Keravuori
Finnish Literature Society
PL 259
00171 Helsinki
Finland
Tel. (358) 913 12 32 27
E-mail: kirsi.keravuori@finlit.fi

Eric Ketelaar
Universiteit van Amsterdam
Oude Turfmarkt 141
1012 GC Amsterdam
Netherlands
Tel. (31) 205 25 20 13
Fax (31) 205 25 20 97
E-mail: ketelaar@hum.uva.nl

O.P. (Onno) Keuter
Bacosys b.v.
Krijtwal 29
3432 ZT Nieuwegein
Netherlands
Tel. (31) 306 04 92 51
Fax (31) 306 05 20 62
E-mail: keuter@bacosys.nl

Jaana Kilkki
Military Archives
PL 54
00581 Helsinki
Finland
Tel. (358) 918 12 65 00
E-mail: jaana.kilkki@sota-arkisto.fi

Bengt Kjällman
The Regional Archives in Gothenburg
Box 19035
400 12 Gothenburg
Sweden
Tel. (46) 317 78 68 12
Fax (46) 317 78 68 25
E-mail:
bengt.kjallman@landsarkivet-goteborg.ra.se

Elisabeth Klett
City Archives of Stockholm
Box 22063
104 22 Stockholm
Sweden
Tel. (46) 850 82 83 36
Fax (46) 850 82 83 01
E-mail: elisabeth.klett@ssa.stockholm.se

Thekla, Dr Kluttig
Sächisches Staatsministerium des Innern
Referat 44 (Archivwesen)
01095 Dresden
Germany
Tel. (49) 35 15 64 34 45
Fax (49) 35 15 64 34 09
E-mail: thekla.kluttig@smi.sachsen.de

Almudena Knecht Cristóbal
Biblioteca CEACS, Instituto Juan March de Estudios e Investigaciones
Castelló, 77
28006 Madrid
Spain
Tel. (34) 914 35 42 40
Fax (34) 915 76 34 20
E-mail: a.knecht@ceacs.march.es

Adolf Knoll
National Library of the Czech Republic
Klementinum 190
110 01 Prague 1
Czech Republic
Tel. (420) 221 66 32 74
Fax (420) 222 22 03 70
E-mail: adolf.knoll@nkp.cz

Walter Koch
AIT Angewandte Informationstechnik, Forschungsgesellschaft mbH
Hans-Sachs-Gasse 14/3
8010 Graz
Austria
Tel. (43) 31 68 35 35 90
Fax (43) 316 83 53 59 75
E-mail: Kochw@ait.co.at

Petri Konttinen
Ministry of Social Affairs and Health
P.O. Box 33
00023 Helsinki
Finland
Tel. (358) 916 07 38 01
Fax (358) 916 07 41 26
E-mail: petri.konttinen@stm.vn.fi

Dincho Krastev
Central Library of the Bulgarian Academy of sciences
1040 City of Sofia
Bulgaria
Tel. (359) 29 87 89 66
Fax (359) 29 86 25 00
E-mail: dincho@cl.bas.bg

Göran Kristiansson
National Archives of Sweden
Box 2016
220 02 Lund
Sweden
Tel. (46) 46 19 70 43
E-mail:
goran.kristiansson@landsarkivet-lund.ra.se

Tamara Kukurekovic
EBRD — European Bank for Reconstruction and Development
1 Exchange Square
EC2A 2JN London
United Kingdom
Tel. (44-20) 73 38 73 16
Fax (44-20) 73 38 79 38
E-mail: burninge@ebrd.com

Indrek Kuuben
The National Archives of Estonia
J. Liivi 4
50409 Tartu
Estonia
Tel. (372) 738 75 00
Fax (372) 738 75 10
E-mail: indrek.kuuben@ra.ee

Erika Kylvaja
Defence Forces Material Command
PL 69
33541 Tampere
Finland
Tel. (358) 318 15 53 45
Fax (358) 318 15 59 15
E-mail: erika.kylvaja@dme.mil.fi

Thijs Laeven
Innogration Management Consultants
Emdenmeen 56
3844 ED Harderwijk
TNetherlands
Tel. (31) 341 43 08 31
Fax (31) 341 43 09 48
E-mail: laeven@innogration.nl

Brice Lançon
IFHE
Villa Poirier 4
75015 Paris
France
E-mail: brice.lancon@ifrance.com

Peter B. Lange
IBM
Nymoellevej 91
2800 Lyngby
Denmark
Tel. (45) 45 23 32 79
E-mail: lange@dk.ibm.com

Leeni Langebraun
National Archives of Estonia
J.Liivi 4
Estonia
Tel. (372) 738 75 01
E-mail: Leeni.Langebraun@ra.ee

Anna Larsson
AstraZeneca AB
Byggn. 410
15185 Södertälje
Sweden
Tel. (46) 855 32 52 29
Fax (46) 855 32 86 90
E-mail: anna.larsson@astrazeneca.com

Roser Latorre Tafanell
Ajuntament de Sant Boi
Plaá de l'Ajuntament 1
08830 Sant Boi dell Llobregat
Spain
Tel. (34) 936 35 12 00
Fax (34) 936 30 18 56
E-mail: mrlatorre@stboi.es

Ülle Laur
State Chancellery of the Republic of Estonia
Rahukohtu 3
15161 Tallinn
Estonia
Tel. (372) 693 58 44
Fax (372) 693 59 14
E-mail: Ulle.Laur@rk.ee

Mary Lewis
EBRD — European Bank for Reconstruction and Development
1 Exchange Square
EC2A 2JN London
United Kingdom
Tel. (44-20) 73 38 73 16
Fax (44-20) 73 38 79 38
E-mail: garcias@ebrd.org

Vesa Liikanen
Prime Minister's Office
PL 23
00023 Valtioneuvosto Valtioneuvosto
Finland
Tel. (358) 405 45 34 95
E-mail: vesa.liikanen@vnk.vn.fi

Bernhard Lindgens
Bundesministerium der Finanzen
Graurheindorfer Str. 108
53003 Bonn
Germany
Tel. (49) 22 86 82 13 24
Fax (49) 22 86 82 27 07
E-mail: Bernhard.Lindgens@bmf.bund.de

Anne Marie Lits
European Commission
Rue de la Loi 22
1049 Brussels
Belgium
Tel. (32-2) 295 92 99
Fax (32-2) 296 07 40
E-mail: anne-marie.lits@cec.eu.int

Maria Jesús Llavero Porcel
Departament Justícia — Generalitat de Catalunya
Palau de Justícia. Canyeret s/n
25007 Lleida
Spain
Tel. (34) 973 70 02 79
E-mail: arxiu.lleida@xij.gencat.es

David Lobato Buil
Diputació de Barcelona
Ptge. Moliner, 2 bis, Esc. A, 3r 2a
08720 Vilafranca del Penedès
Spain
Tel. (34) 617 22 26 62
E-mail: david.lobato@aac.upcnet.es

Jan Lohman
European Central Bank
Kaiser strasse 29
60311 Frankfurt am Main
Germany
Tel. (49) 69 13 44 74 66
Fax (49) 69 13 44 61 70
E-mail: jan.lohman@ecb.int

Rosa María López Alonso
Universidad de Salamanca
Francisco de Vitoria 6-16
37008 Salamanca
Spain
Tel. (34) 923 29 45 80
Fax (34) 923 29 45 82
E-mail: rlalonso@usal.es

Luis Lopez Corral
Plaza de la Cancilleria 3, 7°
28008 Madrid
Spain
Tel. (34) 918 85 66 45
Fax (34) 918 85 66 46
E-mail: luis.lopez@uah.es

F. Javier López Jiménez
Archivo General Militar
Avda. del Ejército, 2
19004 Guadalajara
Spain
Tel. (34) 949 21 39 35
E-mail: frajalojim@terra.es

Olga López Miguel
Museu Marítim-Consorci de les Drassanes de Barcelona
Avda. Drassanes, s/n
08001 Barcelona
Spain
Tel. (34) 933 42 99 25
Fax (34) 933 01 95 80
E-mail: lopezmo@diba.es; cd.mmaritim@diba.es

Tomás López Pizcueta
Tecnologías de la Documentación, S.A.
Calle Entenza, 332, 3º-5ª
Barcelona
Spain
Tel. (34) 933 63 87 43
Fax (34) 933 22 00 67
E-mail: tlopez@tecnodoc.com

Philip Lord
2 Wayside Court
Arlington Road
TW1 2BQ Twickenham
United Kingdom
Tel. (44-20) 86 07 91 02 or
(mobile) (44) 77 33 26 75 10
Fax (44-20) 87 44 93 22
E-mail: philiplord@britishlibrary.net

Håkan Lövblad
National archives of Sweden
PO BOX 12541
102 29 Stockholm
Sweden
Tel. (46) 87 37 63 50
Fax (46) 87 37 64 74
E-mail: hakan.lovblad@riksarkivet.ra.se

Alison Macdonald
Secure Sciences Limited
2 Wayside Court, Arlington Road
TW1 2BQ Twickenham
United Kingdom
Tel. (44-20) 87 44 93 22
Fax (44) 70 50 67 50 10
E-mail:
alison.macdonald@securesciences.co.uk

George Mackenzie
National Archives of Scotland
EH1 3YY Edinburgh
United Kingdom
Tel. (44) 13 15 35 13 12
Fax (44) 13 15 35 13 60
E-mail: George.mackenzie@nas.gov.uk

Markku Mäenpää
National Archives of Finland
Rauhankatu 17
00170 Helsinki
Finland
Tel. (358) 922 95 21
Fax (358) 917 63 02
E-mail: markku.maenpaa@narc.fi

Anna Magrinya I Rull
Servei d'Arxius
Portaferrissa 1
08002 Barcelona
Spain
Tel. (34) 933 16 27 54
Fax (34) 933 16 27 41
E-mail: amagrinya@correu.gencat.es

Sari Mäkinen
Department of Information Studies
University of Tampere
Yliopistonkatu 38
33014 Tampere
Finland
Tel. (358) 32 15 69 69
Fax (358) 32 15 65 60
E-mail: sari.makinen@uta.fi

Irene Manclús Cuñat
Universitat de València
Edifici Biblioteca de CC. Socials
46071 València
Spain
Tel. (34) 963 82 89 55
Fax (34) 963 82 89 54
E-mail: Irene.Manclus@uv.es

Núria Mañé I Suñé
Generalitat de Catalunya-Departament de Justícia
Rossend Arús 7,Pral 1ª
08014 Barcelona
Spain
Tel. (34) 934 21 39 18
E-mail: nmanyer@terra.es

Fink Manfred
Österreichisches Staatsarchiv, Archiv der Republik
Nottendorfergasse 2
A-1030 Vienna
Austria
Tel. (43) 179 54 02 50
Fax (43) 179 54 01 09
E-mail: manfred.fink@oesta.gv.at

Agneta Mann
National Archives of Sweden
Box 12541
102 29 Stockholm
Sweden
Tel. (46) 87 37 63 50
Fax (46) 87 37 64 74
E-mail: agneta.mann@riksarkivet.ra.se

Jorge Marquez Pacios
Soluziona
Infanta Mercedes, 31 3ª planta
28020 Madrid
Spain
Tel. (34) 915 71 20 00
Fax (34) 915 71 52 81
E-mail: esancheza@soluziona.com

Pilar Marti Gracia
Cortes de Aragón
Palcio de la Aljafería sn
50071 Zaragoza
Spain
Tel. (34) 976 28 95 49
Fax (34) 976 28 96 31
E-mail: pmarti@cortesaragon.es

Maria Lourdes Martínez de Antoñana
Comité des Régions
Rue Montoyen 92-102
1000 Brussels
Belgium
Tel. (32-2) 282 20 58
Fax (32-2) 282 23 25
E-mail: lourdes.martinez de antonana@cor.eu.int

Betlem Martínez Radua
Generalitat de Catalunya. Departament d'Economia I Finances
Rambla Catalunya, 19–21
08007 Barcelona
Spain
Tel. (34) 933 16 21 72
E-mail: mbmartinez@eif.gencat.es

Maria del Carmen Martínez Rodríguez
Ajuntament de Barcelona
c/Ciutat, 9 pral 1a
08002 Barcelona
Spain
Tel. (34) 934 02 76 31
Fax (34) 934 02 76 38
E-mail: cmartinez@mail.bcn.es

Iñaki Maruri
Iberdrola
Avda. San Adrich 48
48003 Laraskitu
Spain
Tel. (34) 944 15 14 11
E-mail: imaruri@iberdrola.es

Carlos Mas González
c/Duque de Lerma, 2 (Hospital Tavera)
45003 Toledo
Spain
Tel. (34) 925 21 03 54
E-mail: nobleza@correo.mcu.es

Josep Maria Masachs I Suriol
AHC Vilafranca del Penedes
Plaça de l'Oli 20
08720 Vilafranca del Penedes
Spain
Tel. (34) 938 17 23 11
Fax (34) 938 17 23 11

Josep Matas I Balaguer
Arxiu Historic de Girona
Plaça de Sant Josep 1
17004 Girona
Spain
Tel. (34) 972 22 55 00
Fax (34) 972 22 75 77

M. Mercè Mateu Escoda
Parlament de Catalunya
Parc de la Ciutadella, s/n
08003 Barcelona
Spain
Tel. (34) 933 04 65 00
Fax (34) 933 04 65 49
E-mail: mmateu@parlament-cat.es

Alfred Mauri i Martí
Ajuntament de Martorell
C. Lloselles, 46
08760 Martorell
Spain
Tel. (34) 937 74 04 21
Fax (34) 937 74 05 95
E-mail: bnn@heraclit.net

Mireia Mayol Alaber
GPO Ingenieria SA
Aragó, 390, 5º Planta
08013 Barcelona
Spain
Tel. (34) 932 47 00 69
Fax (34) 932 31 69 69
E-mail: gpo@gpo.es

Rafael Méndez Sastre
Banco de España
Alcalá, 522
28027 Madrid
Spain
Tel. (34) 913 38 64 12
Fax (34) 913 38 68 82
E-mail: rrhh.desarrollo@bde.es

Matthias Meusch
NRW Hauptstaatsarchiv
Mauerstr. 55
40476 Düsseldorf
Germany
Tel. (49) 211 94 49 72 12
Fax (49) 211 94 49 70 02
E-mail: Matthias.Meusch@hsa.nrw.de

F.E. (Cisca) Meyer
Bacosys b.v.
Krijtwal 29
3432 ZT Nieuwegein
The Netherlands
Tel. (31) 306 04 92 51
Fax (31) 306 05 20 62
E-mail: meijer@bacosys.nl

María José Miravitlles Jover
Doctroactive
Osio 7/9
08034 Barcelona
Spain
Tel. (34) 932 53 46 23
Fax (34) 932 53 46 31
E-mail:
mjose.miravitlles.doctoractive@finaf92.es

Grant Mitchell
International Federation of Red Cross and
Red Crescent Societies
P.O. Box 372
1202 Geneva
Switzerland
Tel. (41) 227 30 42 39
Fax (41) 227 33 03 95
E-mail: mitchell@ifrc.org

Laura Mitchell
National Archives of Scotland
HM General Register House, Princes Street
EH1 3YY Edinburgh
United Kingdom
Tel. (44) 13 15 35 14 17
Fax (44) 13 15 35 14 30
E-mail: laura.mitchell@nas.gov.uk

Riku Moisio
National Archives of Finland
Laaksotie 7 a 5
02700 Kauniainen
Finland
Tel. (358) 922 85 21
E-mail: riku.moisio@narc.fi

Marta Mollà Català
Institut Català del Sòl
Còrsega, 289, 6È
08008 Barcelona
Spain
Tel. (34) 935 67 09 24
Fax (34) 932 28 60 58
E-mail: m_molla@correugencat.es

Jan Möller
Ministry of Transport, Public Works and
Water Management
PO Box 20906
2500 EX The Hague
Netherlands
Tel. (31) 703 51 94 17
Fax (31) 703 51 94 40
E-mail: jan.moller@cend.minvenw.nl

Iolanda Mombelli
European Parliament
Bureau 01C015 — ADG — Plateau de
Kirchberg
2929 Luxembourg
Luxembourg
Tel. (352) 430 02 21 18
Fax (352) 43 94 93
E-mail: imombelli@europarl.eu.int

Elisa Matilde Monteiro Gomes
Ministério Finanças — Secret.Geral
Rua da Alfândega, 5
1010-096 L Lisbon
Portugal
Tel. (351) 218 85 41 05
Fax (351) 218 85 41 17
E-mail: knabino@hotmail.com

Jill Moorcroft
Skipton House
80 London Road
SE1 6LH London
United Kingdom
Tel. (44-20) 79 72 58 72
Fax (44-20) 79 72 59 94
E-mail: jill.moorcroft@doh.gsi.gov.uk

Stepan Moore
Ipsen Pharma
Laureà Miró, 395
08980 Sant Feliu de Llobregat
Spain
Tel. (34) 936 85 81 00
Fax (34) 936 85 10 11
E-mail: stepan.moore@Beaufour-Ipsen.com

Inma Moratalla Vallis
Ajuntament de Barcelona
Camèlies 36–38
08024 Barcelona
Spain
Tel. (34) 932 85 03 57
Fax (34) 932 85 07 80

Manuel Moreno Chacón
Pere III, Nº 25, 1r 2ª
17600 Figueres
Spain
Tel. (34) 972 67 24 06
E-mail: manuel.moreno@aac.upcnet.es

Marta Moreno i Gironès
Parlament de Catalunya
Parc de la Ciutadella s/n
08003 Barcelona
Spain
Tel. (34) 933 04 65 00
Fax (34) 933 04 65 49
E-mail: mamoreno@parlament-cat.es

Paul E. Murphy
ENTR.D.5
SC15 02/65
1049 Brussels
Belgium
Tel. (32-2) 295 77 66
Fax (32-2) 299 02 86
E-mail: paul-e.murphy@cec.eu.int

Daria Nalecz
Head Office of State Archives
Dluga 6
00-950 Warsaw
Poland
Tel. (48) 228 31 32 06
Fax (48) 228 31 92 22
E-mail: mjanuszewska @archiwa.gov.pl

Mertens Natascha
SSTC — Services Fédéraux des Affaires Scientifiques, Techniques et Culturelles
Rue de la Science 8
1000 Brussels
Belgium
Tel. (32-2) 238 36 87
Fax (32-2) 230 59 12
E-mail: mert@belspo.be

Marko Natri
National Archives of Finland
Rauhankatu 17
00171 Helsinki
Finland
Tel. (358) 922 85 22 35
E-mail: marko.natri@narc.fi

Ferran Navalon Casino
Dept. d'Ensenyament Generalitat
Via Augusta, 202-226
08021 Barcelona
Spain
Tel. (34) 934 00 69 00-3863
Fax (34) 934 00 69 94
E-mail: r120@ense.gencat.es

Bengt Neiss
The Royal Library
Box 5039, Humlegården
102 41 Stockholm
Sweden
Tel. (46) 84 63 40 00
Fax (46) 84 63 40 03
E-mail: bengt.neiss@kb.se

Marc Nelissen
University Archives — K.U.Leuven
Mgr. Ladeuzeplein 21
3000 Leuven
Belgium
Tel. (32) 16 32 46 32
Fax (32) 16 32 47 09
E-mail: Marc.Nelissen@bib.kuleuven.ac.be

Montse Net Marce
CTTI-Generalitat de Catalunya
Avda. del Estadi s/n
08038 Barcelona
Spain
Tel. (34) 934 84 20 84
Fax (34) 934 84 20 22
E-mail: mnet@correu.gencat.es

Annemieke Nijdam-Dekker
Rijksarchiefinspectie
Postbus 11583
2502 AN Den Haag
The Netherlands
Tel. (31) 703 33 91 33
Fax (31) 703 33 91 30
E-mail: anijdam@rijksarchiefinspectie.nl

Pille Noodapera
National Archives of Estonia
Liivi 4
50409 Tartu
Estonia
Tel. (372) 693 80 26
Fax (372) 661 62 30
E-mail: pille.noodapera@ra.ee

Erik Norberg
Riksarkivet
Box 12541
10229 Stockholm
Sweden
Tel. (46) 87 37 63 50
Fax (46) 87 37 64 74
E-mail: erik.norberg@riksarkivet.ra.se

Gertrud Nord
The Swedish Parliament
Riksdagens arkiv
100 12 Stockholm
Sweden
Tel. (46) 87 86 57 92
E-mail: gertrud.nord@riksdagen.se

Gunilla Nordström
The Regional Archives in Gothenburg
Box 19035
40012 Gothenburg
Sweden
Tel. (46) 317 78 68 14
Fax (46) 317 78 68 25
E-mail:
gunilla.nordstrom@landsarkivet-goteborg.ra.se

Lena Nygren
City Archives, Sundsvall
Kulturmagasinet
851 96 Sundsvall
Sweden
Tel. (46) 60 19 18 71
Fax (46) 60 17 04 01
E-mail: lena.nygren@sundsvall.se

Miguel Oppenheimer Salinas
SGAE
c/Fernando VI, 4
28004 Madrid
Spain
Tel. (34) 913 49 96 47
Fax (34) 913 49 97 63
E-mail: moppenheimer@sgae.es

Santiago Oros Murruzabal
Departament de Cultura
Muntaner 221
08036 Barcelona
Spain
Tel. (34) 933 63 28 75
Fax (34) 933 63 28 79
E-mail: soros@correu.gencat.es

Pedro Ortiz Paez
SDI,sl
Gremio de tejedores,22-1º
07009 Palma de Mallorca
Spain
Tel. (34) 971 43 11 21
Fax (34) 971 43 11 44
E-mail: portiz@sdi-gsb.com

Jean-Marie Palayret
Archives historiques del'Union européenne
Villa Il poggiolo, Piazza Edison.11,
50133 Florence
Italy
Tel. (39) 05 54 68 56 26
Fax (39) 055 57 37 28
E-mail: jean.palayret@iue.it

Albert Palleja
CTTI Generalitat de Catalunya
Pl. Pau Vila 1
08039 Barcelona
Spain
Tel. (34) 934 84 20 84
Fax (34) 934 84 20 22
E-mail: apalleja@correu.gencat.es

Antoni Pallès i Soler
Departament de Governació i Relacions Laborals
Via Laietana, 26
08003 Barcelona
Spain
Tel. (34) 935 67 18 08
E-mail: apalles@governacio-ri.gencat.es

Bibiana Palomar i Baró
Departament de Justícia
Carrer Pau Claris,81
08010 Barcelona
Spain
Tel. (34) 933 16 41 97
Fax (34) 933 16 41 70
E-mail: Bpalomar@correu.gencat.es

Angel Parra Garcia
INSS
Lopez de Hoyos 171
28002 Madrid
Spain
Tel. (34) 915 68 88 43
Fax (34) 915 19 62 37
E-mail: miguel.angel.parra@inss.seg-soc.es

Irja Peltonen
Ministry of Finance, Finland
P.O. Box 28
00023 Government Helsinki
Finland
Tel. (358) 405 40 20 66
Fax (358) 91 60 47 59
E-mail: irja.peltonen@vm.vn.fi

Ulla-Maija Peltonen
Finnish Literature Society
PL 259
00171 Helsinki
Finland
Tel. (358) 913 12 32 33
E-mail: ulla-maija.peltonen@finlit.fi

Ramon Peña Izquierdo
Teodora Lamadrid, 37, 3-4
08022 Barcelona
Spain
Tel. (34) 932 11 87 61
E-mail: antonio.ramon.pena@aac.upcnet.es

Stela Pereira
Banco de Portugal
Rua Francisco Ribeiro, 2- 7º
1150-165 Lisbon
Portugal
Tel. (351) 213 13 02 49
Fax (351) 213 12 81 20
E-mail: mrcruz@bportugal.pt

Alcídio Pereira
European Parliament
Bât. ADG 01C006 Plateau du Kirchberg
2929 Luxembourg
Luxembourg
Tel. (352) 430 02 23 29
Fax (352) 43 94 93
E-mail: apereira@europarl.eu.int

Victòria dels Àngels Pérez Arenales
Consorci de l'Auditori i l'Orquestra
C/ Lepant, 150
08013 Barcelona
Spain
Tel. (34) 932 47 93 00
Fax (34) 932 47 93 01
E-mail: vperez@auditori.com

María de Lourdes Pérez González
Universidade de Santiago de Compostela
Bemil 13 — Bastavales
15280 Brión
Spain
Tel. (34) 981 88 88 06
E-mail: <bxeugo@sialpha1.usc.es>

Xavier Pérez I Gómez
AHC Santa Coloma de Farners
Carrer de l'Ave Maria 5 bis
17004 Santa Coloma de Farners
Spain
Tel. (34) 972 84 21 46
Fax (34) 972 84 35 42

Montserrat Pérez i Rigol
Departament de Política Territorial i Obres Públiq
Avda de Josep Tarradellas, 2-6
08029 Barcelona
Spain
Tel. (34) 934 95 80 51
Fax (34) 934 95 81 94
E-mail: wperez@correu.gencat.es

Montse Peris Giner
Ipsen Pharma S.A.
Laureà Miró, 395
08980 Sant Feliu de Llobregat
Spain
Tel. (34) 936 85 81 00
Fax (34) 936 85 10 11
E-mail: Montse.Peris@Beaufour-Ipsen.com

Sabi Peris I Serradell
AHC Reus
Plaça del Castell, 4
43201 Reus
Spain
Tel. (34) 977 34 27 54
Fax (34) 977 34 18 59

Remei Perpinyà Morera
Universitat Autònoma de Barcelona
Bellaterra
08193 Barcelona
Spain
Tel. (34) 934 19 89 55
Fax (34) 933 22 21 87
E-mail: arxivers@upcnet.es

Christine Petillat
Centre des archives contemporaines
2, rue des archives
77300 Fontainebleau
France
Tel. (33) 164 31 73 73
Fax (33) 164 31 73 03
E-mail: christine.petillat@culture.gouv.fr

Maria Pi Maymó
Universitat Autònoma de Barcelona
C/ Santa Maria 42
08290 Cerdanyola del Vallès
Spain
Tel. (34) 661 21 47 87
E-mail: mariapimaymo@jazzfree.com

Beatriz Piera Moreno
Planificación Jurídica
C/Trafalgar, 70;1º2º
08010 Barcelona
Spain
Tel. (34) 933 19 58 20
Fax (34) 933 19 21 19
E-mail: biblio@planificacion-juridica.com

Jordi Pique I Padro
Arxiu Històric de Tarragona
Rambla Vella, 30
43003 Tarragona
Spain
Tel. (34) 977 25 10 10
Fax (34) 977 25 26 11

Jordi Plana i Franch
Dpt Justícia-Generalitat de Catalunya
Rda.Universistat,18
08007 Barcelona
Spain
Tel. (34) 936 22 06 09
Fax (34) 936 22 06 24
E-mail: xaviermireia@menta.net

Ramon Planes I Albets
Servei d'Arxius
Portaferrisa 1
08002 Barcelona
Spain
Tel. (34) 933 16 27 54
Fax (34) 933 16 27 41
E-mail: rplanes@correu.gencat.es

Raimo Pohjola
National Archives of Finland
P.O.Box 258
00171 Helsinki
Finland
Tel. (358) 922 85 22 48
Fax (358) 917 63 02
E-mail: raimo.pohjola@narc.fi

Bill Pollard
Engineering and physical sciences research council
Polaris house, North star avenue
SN2 1ET Swindon
United Kingdom
Tel. (44) 17 93 44 42 87
E-mail: bill.pollard@epsrc.ac.uk

Montserrat Pons Mascarilla
Ajuntament de Lleida
Plaça Paeria, 1
25003 Lleida
Spain
Tel. (34) 973 70 03 52
E-mail: mpons@paeria.es

Josep Maria Porta I Balanya
AHC de Montblanc
Raval de Santa Anna 12
43400 Montblanc
Spain
Tel. (34) 977 86 04 92
Fax (34) 977 86 04 92

Marti Porter Huerre
Generalitat de Catalunya
Plaça Antoni Lopez 5
08002 Barcelona
Spain
Tel. (34) 932 68 44 00
Fax (34) 932 68 49 00
E-mail: mporter@presidencia.gencat.es

Celsa Poza
Agència Catalana de l'Aigua
c/Provença 204-208
08036 Barcelona
Spain
Tel. (34) 935 67 28 00
E-mail: ePoza@correu.gencat.es

Penny Prior
Foreign and Commonwealth Office
Hanslope Park, Milton Keynes
MK19 7BH Milton Keynes
United Kingdom
Tel. (44) 19 08 51 59 81
Fax (44) 19 08 51 56 92
E-mail: penny.prior@fco.gov.uk

Richard Przewozniak
Aventis Pasteur
1541, av. Marcel Mérieux
69280 Marcy l'Etoile
France
Tel. (33) 437 37 97 54
Fax (33) 437 37 39 05
E-mail: richard.przewozniak@aventis.com

Pere Puig Ustrell
Plaça Sant Miquel, 12
08224 Terrassa
Spain
Tel. (34) 937 33 14 96
Fax (34) 937 80 14 33
E-mail: pere.puig@terrassa.org

Ester Puyol Ibort
Instituto de Estudios Altoaragoneses
Parque, 10
2002 Huesca
Spain
Tel. (34) 974 29 41 20
Fax (34) 974 29 41 22
E-mail: biblioteca@iea.es

Julio Quílez Mata
Arxiu Central dels Jutjats Penals de Barcelona
Av. Lluís Companys 1-5
08013 Barcelona
Spain
Tel. (34) 935 67 30 40
E-mail:
arxiu.instruccioipenal.barcelona@xij.gencat.es

René Quillet
Staatsarchiv Basel-Landschaft
Postfach
4410 Liestal
Schweiz
Tel. (41) 619 26 76 15
Fax (41) 619 26 76 77
E-mail: rene.quillet@lka.bl.ch

Tom Quinlan
National Archives of Ireland
Bishop Street
8 Dublin
Ireland
Tel. (353) 14 07 23 85
Fax (353) 14 07 23 85
E-mail: Tquinlan@nationalarchives.ie

Carl Michael Raab
National Archives of Sweden
Box 12541
102 29 Stockholm
Sweden
Tel. (46) 87 37 64 09
Fax (46) 87 37 64 74
E-mail: carlmichael.raab@riksarkivet.ra.se

Pepita Raventós Pajares
Universitat de Lleida
Plaça Víctor Siurana, 1
25003 Lleida
Spain
Tel. (34) 973 70 20 11
Fax (34) 973 70 20 62
E-mail: p.raventos@sg.udl.es

Agnieszka Reczuch
Nicholas Copernicus University Torun
ul.Dekerta 8
87-100 Torun
Poland
Tel. (48) 566 11 47 78
Fax (48) 566 11 47 78
E-mail: ipitar@cc.uni.torun.pl

Päivi Repo-Lehikoinen
National Archives of Finland
Rauhankatu 17
00170 Helsinki
Finland
Tel. (358) 922 85 21
Fax (358) 917 63 02
E-mail: paivi.repo-lehikoinen@narc.fi

Ma. Luz Retuerta I Jimenez
AHC Sant Feliu de Llobregat
Parc de la Torre del Roser s/n
08980 Sant Feliu de Llobregat
Spain
Tel. (34) 936 66 67 03
Fax (34) 936 85 02 62

Antoni Reyes i Valent
Ajuntament de Blanes
Unió, 13
17300 Blanes
Spain
Tel. (34) 972 35 52 68
Fax (34) 972 35 71 35
E-mail: arxiu@blanes.net

Fernanda Ribeiro
Faculdade de Letras da Universidade
do Porto
Via Panorâmica, s/n
4150-564 P Porto
Portugal
Tel. (351) 226 07 71 72
Fax (351) 226 09 16 10
E-mail: rdc00979@mail.telepac.pt

Joan Rigata Biguera
CTTI-Generalitat de Catalunya
Avda. Josep Tarradellas 2-6
08029 Barcelona
Spain
Tel. (34) 934 84 20 84
Fax (34) 934 84 20 22
E-mail: zrigata@correu.gencat.es

Maria Pia Rinaldi Mariani
Ministero beni et attivitá culturali
Via Gaeta 8A
00185 Roma
Italy
Tel. (39) 064 44 06 95
Fax (39) 064 92 22 54 45
E-mail: Pmariani@archivi.beniculturali.it

Cati Rios Buceta
Enresa
Emilio Vargas, 7
28043 Madrid
Spain
Tel. (34) 915 66 82 66
E-mail: crib@enresa.es

Ronny Robert
Ministère des Affaires Etrangères, du
Commerce Extérieur, de la Coopération,
de l'Action Humanitaire et de la Défense
5, rue Notre Dame
2240 Luxembourg
Luxembourg
Tel. (352) 478 23 40
E-mail: ronny.robert@mae.etat.lu

Eva Roca I Marquès
Universitat Pompeu Fabra
Passeig de Circumval.lació, 8
08003 Barcelona
Spain
Tel. (34) 935 42 16 84
Fax (34) 935 42 24 51
E-mail: eva.roca@suni.upf.es

Assumpta Rodón I Valls
Diputació de Barcelona
Mejía Lequerica, 1
08028 Barcelona
Spain
Tel. (34) 934 02 24 46
Fax (34) 934 92 25 53
E-mail: rodonvm@diba.es

Màrius Rodrigo Peiró
Consorci de les Drassanes de Barcelona
Avda. de les Drassanes s/n
08001 Barcelona
Spain
Tel. (34) 933 42 99 22
Fax (34) 933 18 78 76
E-mail: rodrigopm@diba.es

Mª Socorro Rodriguez Heras
Ayuntamiento. Archivo General
Administrativo
C/ San Antón, 72. Bajo
18005 Granada
Spain
Tel. (34) 958 25 55 62
E-mail: msocrodrigue@supercable.es

Markus Rohner
IMTF
Route du Bleuet 1
1762 Givisiez
Switzerland
Tel. (41) 460 66 66
E-mail: markus.rohner@imtf.ch

Xavier Roig Lizandra
LGAI
Apt. Correos, 18
08193 Bellaterra
Spain
Tel. (34) 935 67 20 19
Fax (34) 935 67 20 01
E-mail: xroig@lgai.es

Laurent Roland
Commission européenne
JECL 1/40
1049 Bru Bruxelles
Belgium
Tel. (32-2) 295 00 64
Fax (32-2) 296 11 49
E-mail: roland.laurent@cec.eu.int

Roland Rölker
European Central Bank
Kaiserstr. 29
60311 Frankfurt/Main
Germany
Tel. (49) 69 13 44 74 84
Fax (49) 69 13 44 66 40
E-mail: roland.rolker@ecb.int

Bjorn Ragnolf Ronning
The Parliamentary Archives
Karl Johans gt. 22
0026 Oslo
Norway
Tel. (47) 23 31 36 74
Fax (47) 23 31 38 59
E-mail: bjorn.ronning@stortinget.no

Seamus Ross
Humanities Advanced Technology and
Information Institute
11 University Gardens
George Service House, University of Glasgow
G12 8QQ Glasgow
United Kingdom
Tel. (44) 14 13 30 36 35
Fax (44) 14 13 30 37 88
E-mail: s.ross@hatii.arts.gla.ac.uk

Aranzazu Rovira I Diez
Dursi
Via Laietana 33
08003 Barcelona
Spain
Tel. (34) 933 06 54 00
Fax (34) 933 06 54 32
E-mail: arantxarovira@correu.gencat.es

Stephen Rowe
Getronics Business Solutions
Capellalaan 1
2130 LA Hoofddorp
The Netherlands
Tel. (31) 204 30 63 92
Fax (31) 204 30 60 33
E-mail: Stephen.Rowe@getronics.com

Pilar Rubio Jiménez
Diputacion Provincial
C/ Ollerías S/N
29012 Málaga
Spain
Tel. (34) 952 13 39 76
Fax (34) 952 13 39 75
E-mail: prubioarch@sopde.es

Vetterli Rudolf
Swiss Re
Postfach
8022 Zürich
Switzerland
Tel. (41) 432 85 26 90
E-mail: Rudolf_Vetterli@swissre.com

David Ryan
Public Record Office, UK
Kew
TW9 4DU Richmond
United Kingdom
Tel. (44-20) 83 92 52 57
E-mail: david.ryan@pro.gov.uk

Hans Rydberg
National arhives
Box 12541
102 29 Stockholm
Sweden
Tel. (46) 87 37 63 23
Fax (46) 87 37 63 16
E-mail: hans.rydberg@riksarkivet.ra.se

Ramon Saball Balasch
Sala contenciosa administrativa del Tribunal
Superior de Justícia de Catalunya
Via Laietana, 56
08003 Barcelona
Spain
Tel. (34) 933 44 00 16/933 44 00 27
Fax (34) 933 44 00 19
E-mail: arxiu.contencios@xij.gencat.es

Jan Sahlén
Västernorrland County Counsil
Storgatan 1
871 85 Härnösand
Sweden
Tel. (46) 703 99 23 02
Fax (46) 61 18 01 52
E-mail: jan.sahlen@lvn.se

Tom Sahlén
Sundsvall municipality
851 85 Sundsvall
Sweden
Tel. (46) 60 19 18 77
Fax (46) 60 59 18 77
E-mail: tom.sahlen@sundsvall.se

Raija Salin
Finnish Defence Forces
PL 107
02631 Espoo
Finland
Tel. (358) 91 81 01 11
Fax (358) 918 13 54 89
E-mail: raija.salin@cis-centre.mil.fi

Heikki Salmén
Varma-Sampo Mutual Pension Insurance Company
P.O.Box 1
00098 Helsinki
Finland
Tel. (358) 105 14 32 31
Fax (358) 105 14 32 17
E-mail: heikki.salmen@varma-sampo.fi

Montserrat Salvador i Padrosa
Arxiu Històric de Sabadell
Indústria, 32-34
08202 Sabadell
Spain
Tel. (34) 937 26 87 77
Fax (34) 937 27 57 03
E-mail: ahs@ajsabadell.es

Rune Samdal
ErgoEphorma AS
Oscar Wistings v 19
7020 Trondheim
Norway
Tel. (47) 73 51 53 52
Fax (47) 73 51 63 94
E-mail: rune.samdal@ergoephorma.no

Georgina Sanchez Baños
Adra 33
08320 El Masnou
Spain
Tel. (34) 935 55 00 83
E-mail: ginasb@terra.es

Maria Sanchez Rodriguez
Enresa
Emilio Vargas, 7
28043 Madrid
Spain
Tel. (34) 915 66 82 66
E-mail: msar@enresa.es

Carme Sans I Moya
Dursi
Via Laietana 33
08003 Barcelona
Spain
Tel. (34) 933 06 54 00
Fax (34) 933 06 54 32
E-mail: csans@correu.gencat.es

Josep Ma. Sans I Trave
Arxiu Nacional de Catalunya
Carrer de Jaume I, 33-51
08190 Sant Cugat del Vallés
Spain
Tel. (34) 935 89 77 88
Fax (34) 935 89 80 35

Berenice Santonja Masia
Universidad de Alicante
Apartado de Correos,99
03080 Alicante
Spain
Tel. (34) 965 90 24 09
Fax (34) 965 90 37 35
E-mail: Berenice.Santonja@ua.es

Maria Sardà Vilardaga
Generalitat. Departament Política Territorial
Av. Josep Tarradellas, 2-6
08029 Barcelona
Spain
Tel. (34) 934 95 81 00
Fax (34) 934 95 81 94
E-mail: wmsarda@correu.gencat.es

Günther Schietecat
Royal Meteorological Institute
Ringlaan 3
1180 Brussels
Belgium
Tel. (32-2) 450 03 29
Fax (32-2) 450 03 29
E-mail: gds@oma.be

Klaus Arøe Schmidt
Danish National Archives
Rigsdagsgården 9
1218 Copenhagen
Denmark
Tel. (45) 44 85 14 86
Fax (45) 33 15 56 43
E-mail: kas@ra.sa.dk

Carolien Schönfeld
Amsterdam municipal archives
po box 51140
1007 EC Amsterdam
The Netherlands
Tel. (31) 572 03 33
Fax (31) 572 03 26
E-mail: cschonfeld@gaaweb.nl

Jef Schram
European Commission
JECL 3/35
Rue de la Loi, 200
1049 Brussels
Belgium
Tel. (32-2) 295 67 21
Fax (32-2) 296 10 95
E-mail: Josephus.schram@cec.eu.int

Elizabeth Scott-Wilson
The Stationery Office
51 Nine Elms Lane
SW8 5DR London
United Kingdom
Tel. (44-20) 78 73 83 41
E-mail: elizabeth.scott-wilson@tso.co.uk

Jordi Serchs Serra
Ajuntament de Barcelona
c/ Ciutat, 9. pral 1a.
08002 Barcelona
Spain
Tel. (34) 934 02 76 28
Fax (34) 934 02 76 38
E-mail: jserchs@mail.bcn.es

Erika Serna I Coba
AHC Figueres
Plaça Escorxador Vell s/n
17600 Figueres
Spain
Tel. (34) 972 67 01 78
Fax (34) 972 67 01 78

Jordi Serra Serra
Portaferissa, 1
8002 Barcelona
Spain
Tel. (34) 933 16 28 37
Fax (34) 933 16 2741
E-mail: Jordiserra@correu.gencat.es

Elizabeth Shepherd
School of Library, Archive and Information Studies, University College London
Gower Street
WC1E 6BT London
United Kingdom
Tel. (44-20) 76 79 29 45
Fax (44-20) 73 83 05 57
E-mail: e.shepherd@ucl.ac.uk

Leena Sinerva
Finnish Defence Forces
PL 107
02631 Uusimaa
Finland
Tel. (358) 91 81 01 11
Fax (358) 918 13 54 89
E-mail: leena.sinerva@cis-centre.mil.fi

Trond Sirevag
The National Archives of Norway
PBox 4013 Ulleval stadion
0806 Oslo
Norway
Tel. (47) 22 02 26 00
Fax (47) 22 23 74 89
E-mail: trond.sirevag@riksarkivaren.dep.no

Peter Sivervall
National Archives of Sweden
Box 12541
102 29 Stockholm
Sweden
Tel. (46) 87 37 63 50
Fax (46) 87 37 64 74
E-mail: peter.sivervall@riksarkivet.ra.se

Carina Sjogren
Swedish Broadcasting Resoruces
Central Archives N1B
105 10 Stockholm
Sweden
Tel. (46) 87 84 34 98
Fax (46) 86 60 45 93
E-mail: carina.sjogren@srf.se

Rolf Sjögren
City Archives of Norrköping
City Hall
601 81 Norrköping
Sweden
Tel. (46) 11 15 11 70
Fax (46) 11 15 11 62
E-mail: rolf.sjogren@norrkoping.se

Börje Sjöman
Swedish Broadcasting Resources
Central Archives
105 10 Stockholm
Sweden
Tel. (46) 87 84 51 71
Fax (46) 86 60 45 93
E-mail: borje.sjoman@srf.se

Jacqueline Slats
ICT Uitvoeringsorgaan (ICTU)
Nieuwe Duinweg 24-26
Postbus 84011
2508 AA Den Haag
The Netherlands
Tel. (31) 70 331 5511
Fax (31) 708 88 78 80
E-mail: Jacqueline.slats@ictu.nl

Patricia Sleeman
National Digital Archive of Datasets, University of London Computer Centre
20 Guilford St
WC1 N 1DZ London
United Kingdom
Tel. (44) 20 76 92 13 45
E-mail: p.sleeman@ulcc.ac.uk

Staffan Smedberg
City Archives of Stockholm
Box 22063
10422 Stockholm
Sweden
Tel. (46) 850 82 83 10
Fax (46) 850 82 83 01
E-mail: staffan.smedberg@ssa.stockholm.se

Frans Smit
Municipal Archives of Amsterdam
Amsterldijk 67
1074 HZ Amsterdam
The Netherlands
Tel. (31) 205 72 02 27
Fax (31) 206 75 05 96
E-mail: fsmit@gaaweb.nl

Hans Smit
Stichting ICTU
Postbus 84011
2508 AA Den Haag
The Netherlands
Tel. (31) 708 88 78 02
E-mail: hans.smit@ictu.nl

Bernard Smith
DG Information Society, Cultural Heritage Applications
EUFO 03/3274
Rue Alcide de Gasperi
2920 Luxembourg
Luxembourg
Tel. (352) 430 13 41 95
Fax (352) 430 13 35 30
E-mail: Bernard.Smith@cec.eu.int

José Luís Sobreda Antunes
Rua Prof. Moisés Amzalak, 11-7ºB
1600-648 Lisboa
Portugal
Tel. (351) 218 82 32 45
E-mail: jla@dgaeri.pt

Mª Nieves Sobrino Garcia
Consejería de las Artes
Pza. de España, 8
28008 Madrid
Spain
Tel. (34) 915 80 24 84
Fax (34) 915 80 24 81
E-mail: nieves.sobrino@madrid.org

Fina Solà
Consorci de les Drassanes de Barcelona
Av. de les Drassanes, s/n
08001 Barcelona
Spain
Tel. (34) 933 42 99 20
Fax (34) 933 18 78 76
E-mail: solagf@diba.es

Santiago Soler I Simon
AH Girona
Plaça de Sant Josep, 1
17004 Girona
Spain
Tel. (34) 972 22 55 00
Fax (34) 972 22 75 77

Jan Dalsten Sørensen
Danish State Archives, Rigsarkivet
Rigsdagsgarden 9
1218 København K
Tel. (45) 33 92 23
Fax (45) 33 95 42 11
E-mail: jds@ra.sa.dk

Josep Maria Soteras Pons
Germans Maristes
C/ Evarist Arnús 44
08014 Barcelona
Spain
Tel. (34) 934 90 81 46
Fax (34) 934 90 40 67
E-mail: fms@maristes-cat.es

Sebastien Soyez
Services Fédéraux des Affaires Scientifiques, Techniques et Culturelles (SSTC)
Rue de la Science, 8
1000 Bruxelles
Belgium
Tel.(32-2) 238 36 86
Fax (32-2) 230 59 12
E-mail: soye@belspo.be

Chistoph Spilling
Parlement Européen
Service des Archives, ADG 01C004
Plateau de Kirchberg
2929 Luxembourg
Luxembourg
Tel. (352) 430 02 02 99
Fax (352) 43 94 93
E-mail: auxjdg4@europarl.eu.int

Anki Steen
Luftfartsverket
Vikboplan 11
601 79 Norrköping
Sweden
Tel. (46) 11 19 21 39
E-mail: anki.steen@lfv.se

Charlotte Steinmark
Mondo A/S
Lyngbyvej 32F
2100 Copenhagen
Denmark
Tel. (45) 39 13 02 00
E-mail: cst@mondo.dk

Birger Stensköld
Swedish National Archive
Box 12541
10229 Stockholm
Sweden
Tel. (46) 87 37 64 46/706 56 63 28
E-mail: birger.stenskold@riksarkivet.ra.se

Bernt Stigsson
Swedish Broadcasting Resources
RH-BVB
105 10 S Stockholm
Sweden
Tel. (46) 87 84 97 92
E-mail: bernt.stigson@srf.se

Susanne Strangert
City Archives of Stockholm
P O Box 22063
104 22 Stockholm
Sweden
Tel. (46) 850 82 83 08
Fax (46) 850 82 83 01
E-mail: susanne.strangert@ssa.stockholm.se

Lia Strik
Ministry of Transport, Public Works and Water Mana
Po Box 20901
2500 EX The Hague
The Netherlands
Tel. (31) 703 51 94 06
E-mail: lia.strik@cend.minvenw.nl

Oscar Struijve
Education for Change Ltd
Bow House, 153 — 159 Bow Road
E3 2SE London
United Kingdom
Tel. (44-20) 898 08 18
Fax (44-20) 898 08 18
E-mail: o.struijve@efc.co.uk

Paltin Sturdza
Departemento de Ciencas de la Computación
Universidad de Alcala
28871 Alcala de Henares
E-mail: paltin.sturdza@uah.es

Martin Stürzlinger
Wiener Stadt- und Landesarchiv
Rathaus
1082 Wien
Austria
Tel. (43) 14 00 08 48 23
Fax (43) 140 00 72 38
E-mail: stm@m08.magwien.gv.at

Imma Subirats Coll
Generalitat de Catalunya, Departament de Política
Av. Josep Tarradellas, 2-6
08029 Barcelona
Spain
Tel. (34) 934 95 84 21
Fax (34) 934 95 81 94
E-mail: wisubirats@correu.gencat.es

Rutger J. Suir
Provincie Noord-Brabant
Ockeghemstraat 23
5262 HP Vught
The Netherlands
Tel. (31) 736 80 83 92
Fax (31) 736 80 82 02
E-mail: rsuir@brabant.nl

Perez Sylvia
Josep Trueta, 115
08224 Terrassa
Spain
Tel. (34) 608 73 76 24
Fax (34) 937 80 48 45
E-mail: syp@cedefop.eu.int

John Symon
AIIM International
Chappel House
The Green Datchet
SL3 9EH Berkshire
United Kingdom
Tel. (44) 17 53 59 27 69
Fax (44) 17 53 59 27 70
E-mail: Jsymon@aiim.org

Kirsten Taarnskov
Danmarks Nationalbank
Havnegade 5
1093 København K
Denmark
Tel. (45) 33 63 60 51
Fax (45) 33 63 71 03
E-mail: kt@nationalbanken.dk

Antero Taimiaho
Ministry of Social Affairs and Health
PO BOX 33
0023 Government Helsinki
Finland
Tel. (358) 916 07 38 87
Fax (358) 916 07 31 17
E-mail: antero.taimiaho@stm.vn.fi

Juha Tallinen
Defence Forces Materiel Command
P. Box 69
33541 Tampere
Finland
Tel. (358) 318 15 50 36
Fax (358) 318 15 58 88
E-mail: juha.tallinen@dme.mil.fi

Xavier Tarraubella Mirabet
Ajuntament de Barcelona
Carrer de Sta. Llucia 1
08002 Barcelona
Spain
Tel. (34) 933 18 11 95
Fax (34) 933 17 83 27
E-mail: xtarraubella@mail.bcn.es

Xavier Tarres Chamorro
Informàtica Ajuntament de Sabadell (I.A.S.)
C. de Sant Pere, 4, 2n pis
08201 Sabadell
Spain
Tel. (34) 937 45 32 23
Fax (34) 937 45 32 26
E-mail: xtarres@ajsabadell.es

Antoni Tarrés Rossell
Ajuntament de Barcelona
c/Ciutat, 9. pral 1a.
08002 Barcelona
Spain
Tel. (34) 934 02 71 95
Fax (34) 934 02 76 38
E-mail: ttarres@mail.bcn.es

Albert Taulé i Tello
Arxiu Històric de Sabadell
Indústria, 32-34
08202 Sabadell
Spain
Tel. (34) 937 26 87 77
Fax (34) 937 27 57 03
E-mail: ahs@ajsabadell.es

Jean-Pierre Teil
Centre des Archives Contemporaines
Rue des Archives, 2
77300 Fontainebleau
France
Tel. (33) 164 31 74 75
Fax (33) 164 31 74 66
E-mail: Jean-pierre.teil@culture.gouv.fr

Jaume Teixo I Montala
Servei d'Arxius
Portaferrissa 1
08002 Barcelona
Spain
Tel. (34) 933 16 27 54
Fax (34) 933 16 27 41
E-mail: jteixidom@correu.gencat.es

Susanne Terkildsen
A/S GEA Farmaceutisk Fabrik
Holger Danskes Vej 89
2000 Frederiksberg
Denmark
Tel. (45) 38 34 42 42
Fax (45) 38 34 15 72
E-mail: terkilds@gea-pharm.dk

Giuliano Terzuoli
Archives Historiques des Communautés Européennes
Piazza Edison, 11
50133 Firenze
Italy
Tel. (39) 05 54 68 56 33
Fax (39) 055 57 37 28
E-mail: giuliano@iue.it

Kenneth Thibodeau
National Archives and Records Administration
8601 Adelphi Rd. 1540
20740-6001 College Park, MD
United States
Tel. (1-301) 713 73 30 ext 293
Fax (1-301) 713 74 99
E-mail: ken.thibodeau@nara.gov

Ulrich Thiel
Fraunhofer Institut für integrierte Publikations- und Informationssysteme (IPSI)
Dolivostr. 15
64293 Darmstadt
Germany
Tel. (49) 61 51 86 98 55
Fax (49) 615 18 69 69 68
E-mail: thiel@ipsi.fraunhofer.de

Juhani Tikkanen
PO Box 258
171 Helsinki
Finland
Tel. (358) 922 85 22 41
Fax (358) 917 63 02
E-mail: Juhani.tikkanen@narc.fin

Nathalie Tinjod
European Space Agency
8-10 Rure Mario Nikis
75738 Paris-Cedex
France
Tel. (33) 153 69 75 67
Fax (33) 153 69 76 01
E-mail: nathalie.tinjod@esa.int

Malcolm Todd
Electronic Records Management Development Unit, Records Management Dept
Public Record Office
TW9 4DU Kew
United Kingdom
Tel. (44-20) 83 92 53 40
Fax (44-20) 83 92 52 83
E-mail: malcolm.todd@pro.gov.uk

Peter Toebak
Toebak GmbH / Université de Lausanne
Spitzackerstrasse 7
4410 Liestal
Switserland
Tel. (41) 797 06 24 39
Fax (41) 61 921 89 93
E-mail: toebak@toebak.ch

Jordi Torner Planell
AHC Manresa
Via Sant Ignasi, 40
08240 Manresa
Spain
Tel. (34) 938 73 65 25
Fax (34) 938 73 65 25

Rafael Tornero
SCN Sistemas S.L.
Igarsa S/N
28860 Paracuellos de Jarama
Spain
Tel. (34) 916 58 03 25
Fax (34) 916 58 41 68
E-mail: rtornero@scn-sistemas.com

Maria Jose Tornero Pla
Ayuntamiento de Sedavi
C/ Sagunta,3 Bajo
46910 Sedavi
Spain
Tel. (34) 963 18 24 77
Fax (34) 963 96 38 79
E-mail: sedavi_arx@gva.es

Dan Tørning
Danish State Archives, Rigsarkivet
Rigsdagsgarden, 9
1218 København K
Denmark
Tel. (45) 33 92 21 81
Fax (45) 33 92 45 11
E-mail: dt@ra.sa.dk

Agustín Torreblanca López
Ministerio de Hacienda
Alcalá 9
28071 Madrid
Spain
Tel. (34) 915 95 88 70
Fax (34) 915 95 83 71
E-mail: agustin.torreblanca@sgt.minhac.es

Marta Tort Pascual
Consorci de Biblioteques Universitaries
de Catalunya
Gran Capita, 2-4, edifici Nexus
08034 Barcelona
Spain
Tel. (34) 934 01 58 31
Fax (34) 934 01 58 31
E-mail: mtort@cbuc.es

Pirjo Turunen
National Agency for Medicines
Mannerheimintie 166
00301 Helsinki
Finland
Tel. (358) 947 33 42 32
Fax (358) 971 44 69
E-mail: pirjo.turunen@nam.fi

Cristobal Urbano Salido
Universitat de Barcelona
Melcior de Palau 140
08014 Barcelona
Spain
Tel. (34) 934 03 57 74
Fax (34) 934 03 57 72
E-mail: urbano@fdb.ub.es

Ricard Urgell Hernández
Arxiu del Regne de Mallorca
C/Anselm Turmeda, 10 6è A
07010 Palma
Spain
Tel. (34) 971 76 19 66
E-mail: ricard.urgell@ono.com

Maarten Willem Van Boven
Rijksarchiefdienst
Prins WIllem Alexanderhof 20
2595 BE The Hague
The Netherlands
Tel. (31) 703 31 55 17
Fax (31) 703 31 55 55
E-mail: maarten.van.boven@rad.archief.nl

Cor van der Meer
NIWI / Steinmetz Archive
Muyskenweg 25
1096 CJ Amsterdam
The Netherlands
Tel. (31) 204 62 86 24
E-mail: cor.van.der.meer@niwi.knaw.nl

Hans Van Der Neult
Algemeen Rijsarchief
Prins Willem Alexanderhof 20
2595 Be Den Haag
Netherland
Tel. (31) 703 31 55 37
Fax (31) 70 33 11 55 40
E-mail: hans.van.der.neut@ara.archief.nl

Guus Van Domburg
European Central Bank (ECB)
Kaiserstrasse 29
60311 Frankfurt am Main
Germany
Tel. (49) 69 13 44 85 22
Fax (49) 69 13 44 66 40
E-mail: guus.van_domburg@ecb.int

Dominique Van Laeys
Canal (Belgique
656 Chaussée de Louvain
1030 Brussels
Belgium
Tel. (32-2) 730 02 43
Fax (32-2) 730 02 02

Jean-Claude Van Lancker
Eurocontrol
Rue de la Fusée 96
1130 Brussels
Belgium
Tel. (32-2) 729 30 85
E-mail:
jean-claude.van-lancker@eurocontrol.be

Maria Van Loon
Stichting ICTU
Postbus 84011
2508 AA Den Haag
The Netherlands
Tel. (31) 708 88 78 24
E-mail: riet.vanloon@ictu.nl

Nanning Van Til
Gemeentearchief Amsterdam
Postbus 51140
1007 EC Amsterdam
The Netherlands
Tel. (31) 205 72 03 13
E-mail: nvantil@gaaweb.nl

Leena Vanhanen
Prime Minister´s Office
PB 23
000 23 Valtioneuvosto
Finland
Tel. (358) 01 60 40 01
Fax (358 01 60 21 65
E-mail: leena.vanhanen@vnk.vn.fi

Willem Vanneste
Stadsarchief Antwerpen
Venusstraat 11
2000 Antwerpen
België
Tel. (32) 32 06 94 19
Fax (32) 32 06 94 10
E-mail: willem.vanneste@sd.antwerpen.be

Manuel Vaquerizo Gil
Consejería de Cultura, Turismo y Deporte.
Gobierno de Cantabria
Gravina, 4.
39001 Santander
Spain
Tel. (34) 942 24 15 60
Fax (34) 942 24 15 64
E-mail: arhispro@cultura-cantabria.org

Josep Vela Albesa
Fihoca
Avinguda de Roma, 25-27
08029 Barcelona
Spain
Tel. (34) 934 10 90 00
Fax (34) 934 10 30 81
E-mail: jvela@fihoca.com

Susanna Vela Palomares
Govern d'Andorra
C/ Prada Casadet, 8-12
00000 Andorra la Vella
Andorra
Tel. (376) 86 74 85
Fax (376) 86 86 45
E-mail: sacultura.gov@andorra.ad

Josep M. Ventura
Institut Municipal d'Hisenda
Avda. Litoral. 30
08005 Barcelona
Spain
Tel. (34) 934 02 37 70
Fax (34) 934 02 38 50
E-mail: jventura@mail.bcn.es

Lucie Verachten
Ruisbroekstraat, 2
1000 Brussels
Belgium
Tel. (32-2) 513 76 80
Fax (32-2) 513 76 81
E-mail: verachten@arch.be

Joao Vieira
Direccao Geral Edificios E Monumentos
Nacionales
Praça do Comercio
Lisbon
Portugal
Tel. (351) 214 15 61 33
Fax (351) 214 15 64 87

Maria Eliana Vieira Santos
Institut d'Estudis Territorials
Pg. Circumval lacio 8
08003 Barcelona
Spain
Tel. (34) 935 42 24 52
Fax (34) 935 42 25 99
E-mail: eliana.vieira@ietcat.org

Eila Vila i Delclos
Departament d'Industria, Comerç I Turisme
Passeig de Gracia 105, 2º
08003 Barcelona
Spain
Tel. (34) 934 84 95 06
Fax (34) 934 84 98 22
E-mail: evilad@correu.gencat.es

Pedro Manuel Vilaça
Banco de Portugal
Rua Francisco Ribeiro, No 2 7º planta
1150-165 Lisboa
Portugal
Tel. (351) 213 13 02 49
Fax (351) 213 12 81 20
E-mail: mrcruz@bportugal.pt

Montserrat Vilalta Porta
Universitat Politecnica de Catalunya
Jordi Girona, 31 — Edifici CH
08028 Barcelona
Spain
Tel. (34) 934 01 07 43
Fax (34) 934 01 79 62
E-mail: montserrat.vilalta@upc.es

Teresa Vilàs Barbé
Junta de Residus
Dr Roux, 80
08017 Barcelona
Spain
Tel. (34) 935 67 33 00
Fax (34) 935 67 32 94
E-mail: tvilas@correu.gencat.es

Glòria Vilella Garcia
c. Baró de Maials, 81 3º 2ª
25005 Lleida
Spain
Tel. (34) 636 44 15 11
E-mail: gvilella@uoc.edu

Hans Visser
Municipal Archives Amsterdam
Amsteldijk 67
1074 HZ Amsterdam
The Netherlands
Tel. (31) 205 72 03 66
Fax (31) 206 75 05 96
E-mail: hvisser@gaaweb.nl

Jörg Vogler
Zeutschel
Heerweg 2
72070 Tübingen
Germany
Tel. (49) 707 19 70 60
Fax (49) 70 71 97 06 44
E-mail: info@zeutschel.de

Felix Von Bredow
Project Consult GmbH
Oderfelder Strasse, 17
20149 Hamburg
Tel. (49) 40 46 07 62 20
Fax (49) 40 46 07 62 29
E-mail: corinna.busch@project-consult.com
& felix.v.bredow@project-consult.com

Rik Vooges
European Commission
rue de la Loi 200
1049 Brussels
Belgium
Tel. (32-2) 296 81 08
Fax (32-2) 296 70 28
E-mail: rik.vooges@cec.eu.int

Stephanie Waeyenbergh
VUB-Archiefdienst
Pleinlaan 2
1050 Brussel
Belgium
Tel. (32-2) 629 24 34
E-mail: Stephanie.Waeyenbergh@vub.ac.be

Gerhard K. Wagner
Generalsekretär
Postfach 273
1092 Wien
Austria
Tel. (43) 19 61 01 02
Fax (43) 19 61 01 03
E-mail: Gkwagner@via.at

Hubert Wajs
Central Archives of Historical Records
Dluga 7
00-263 Warsaw
Poland
Tel. (48) 228 31 15 25
Fax (48) 228 31 16 08
E-mail: mjanuszewska @archiwa.gov.pl

Martin Waldron
In-form Systems Limited
Weathervane House
Old Shire Lane
WD3 5WP Chorleywood
United Kingdom
Tel. (44) 19 23 28 36 94
E-mail: mw@informsystems.co.uk

Jonas Walker
Västernorland County Council
Storgatan 1
871 85 Härnösand
Sweden
Tel. (46) 702 15 69 51
Fax (46) 61 18 01 52
E-mail: jonas.walker@lvn.se

Lisa Wallace
University of Glasgow
77-87 Dumbarton Road
G12 7PQ Glasgow
United Kingdom
Tel. (44) 14 13 30 20 15
E-mail: l.wallace@archives.gla.ac.uk

Andrew Waugh
CSIRO Mathematical and Information Sciences
723 Swanston Street
3053 Carlton, VIC
Australia
Tel. (61) 395 45 84 55
Fax (61) 395 45 80 80
E-mail: Andrew.waugh@csiro.au

Sheila Webb
EBRD — European Bank for Reconstruction and Development
One Exchange Square
EC2A 2JN London
United Kingdom
Tel. (44-20) 73 38 73 16
Fax (44-20) 73 38 79 38
E-mail: garcias@ebrd.org

Hartmut Weber
Bundesarchiv
Potsdamer Straße 1
56075 Koblenz
Germany
Tel. (49) 261 50 52 00
Fax (49) 261 50 52 95
E-mail: h.weber@barch.bund.de

Hermann Weidemann
Naestved Kommune
Naestved Raadhus
Teatergade 8
4700 Naestved
Denmark
Tel. (45) 55 78 45 10
E-mail: weide@naeskom.dk

Karl Wessbrandt
Statskontoret (Swedish Agency for Public Management)
Box 2280
103 17 Stockholm
Sweden
Tel. (46) 84 54 47 12
Fax (46) 84 54 46 93
E-mail: karl.wessbrandt@statskontoret.se

Paul Wheatley
University of Leeds (Camileon Project)
28 Dene House Court
LS2 9BS Leeds
United Kingdom
Tel. (44) 77 53 80 35 18
E-mail: p.r.wheatley@leeds.ac.uk

George John White
European Commission, Agriculture DG
L. 130. 3/241
1049 Brussels
Belgium
Tel. (32-2) 295 88 29
Fax (32-2) 295 01 30
E-mail: george.white@cec.eu.int

Georges Willemin
Comité international de la Croix-Rouge
19, avenue de la Paix
1202 Genève
Switzerland
Tel. (41) 227 30 20 53/227 30 28 32
Fax (41) 227 33 20 57
E-mail: gwillemin.gva@icrc.org

Mary Wise
P.M.A. sas
Via Dante 1
20090 Limito
Italy
Tel. (39) 02 92 16 34 69
Fax (39) 02 92 16 30 82
E-mail: marywise@tin.it

Werner Wolfgang
Rheinisches Archiv- und Museumsamt
Ehrenfriedstr. 19
50259 Pulheim
Germany
Tel. (49) 223 49 85 43 43
Fax (49) 221 82 84 29 80
E-mail: w.werner@lvr.de

Rafael Yanguas Laborda
Institut Municipal de Hisenda
Avda. Litoral, 30
08005 Barcelona
Spain
Tel. (34) 933 09 86 40
Fax (34) 933 09 80 74
E-mail: ryanguas@mail.bcn.es

Javier Zabala Jordana
Sociedad General de Archivos, S.A.
Pablo Alcover, 75-77
08017 Barcelona
Spain
Tel. (34) 932 12 35 58
Fax (34) 932 12 10 00
E-mail: javier.z@sga.es

Gennadi Zalaev
Russian State Archives, Scientific and Technical Documentation
Profsouznaja str., 82
117389 Moscow
Russian Federation
Tel. (7) 09 53 34 23 22
Fax (7) 09 53 35 00 95
E-mail: zalaev@rgantd.ru

Jaume Enric Zamora Escala
Diputació de Barcelona
Comte d'Urgell, 187
08036 Barcelona
Spain
Tel. (34) 934 02 25 70
Fax (34) 934 02 28 25
E-mail: zamoraej@diba.es

Karolina Zernicka
Nicholas Copernicus University
ul. Dekerta 8
87-100 Torun
Poland
Tel. (48) 566 11 47 78
Fax (48) 566 11 47 78
E-mail: sb4zeka@cc.uni.torun.pl

Iolanda Zuil Escobar
Departament d'agricultura, ramaderia i pesca
lepant 222A
08203 Sabadell
Spain
Tel. (34) 933 04 67 00
Fax (34) 933 04 67 07
E-mail: aizuies@correu.gencat.es

Bronislaw Zurawski
Nicholas Copernicus University, Torun
Dekerta 8
87-100 Torun
Poland
Tel. (48) 566 11 47 78
Fax (48) 566 11 47 78
E-mail: bbz@cc.uni.torun.pl

European Commission

Proceedings of the DLM-Forum 2002

@ccess and preservation of electronic information: best practices and solutions

Barcelona, 6–8 May 2002

Luxembourg: Office for Official Publications of the European Communities

2002 — 641 pp. — 21 x 29.7 cm

ISBN 92-894-4415-0

Venta • Salg • Verkauf • Πωλήσεις • Sales • Vente • Vendita • Verkoop • Venda • Myynti • Försäljning
http://eur-op.eu.int/general/en/s-ad.htm

BELGIQUE/BELGIË

Jean De Lannoy

Avenue du Roi 202/Koningslaan 202
B-1190 Bruxelles/Brussel
Tél. (32-2) 538 43 08
Fax (32-2) 538 08 41
E-mail: jean.de.lannoy@infoboard.be
URL: http://www.jean-de-lannoy.be

**La librairie européenne/
De Europese Boekhandel**

Rue de la Loi 244/Wetstraat 244
B-1040 Bruxelles/Brussel
Tél. (32-2) 295 26 39
Fax (32-2) 735 08 60
E-mail: mail@libeurop.be
URL: http://www.libeurop.be

Moniteur belge/Belgisch Staatsblad

Rue de Louvain 40-42/Leuvenseweg 40-42
B-1000 Bruxelles/Brussel
Tél. (32-2) 552 22 11
Fax (32-2) 511 01 84
E-mail: eusales@just.fgov.be

DANMARK

J. H. Schultz Information A/S

Herstedvang 12
DK-2620 Albertslund
Tlf. (45) 43 63 23 00
Fax (45) 43 63 19 69
E-mail: schultz@schultz.dk
URL: http://www.schultz.dk

DEUTSCHLAND

Bundesanzeiger Verlag GmbH

Vertriebsabteilung
Amsterdamer Straße 192
D-50735 Köln
Tel. (49-221) 97 66 80
Fax (49-221) 97 66 82 78
E-Mail: vertrieb@bundesanzeiger.de
URL: http://www.bundesanzeiger.de

ΕΛΛΑΔΑ/GREECE

G. C. Eleftheroudakis SA

International Bookstore
Panepistimiou 17
GR-10564 Athina
Tel. (30-1) 331 41 80/1/2/3/4/5
Fax (30-1) 325 84 99
E-mail: elebooks@netor.gr
URL: elebooks@hellasnet.gr

ESPAÑA

Boletín Oficial del Estado

Trafalgar, 27
E-28071 Madrid
Tel. (34) 915 38 21 11 (libros)
913 84 17 15 (suscripción)
Fax (34) 915 38 21 21 (libros),
913 84 17 14 (suscripción)
E-mail: clientes@com.boe.es
URL: http://www.boe.es

Mundi Prensa Libros, SA

Castelló, 37
E-28001 Madrid
Tel. (34) 914 36 37 00
Fax (34) 915 75 39 98
E-mail: libreria@mundiprensa.es
URL: http://www.mundiprensa.com

FRANCE

Journal officiel

Service des publications des CE
26, rue Desaix
F-75727 Paris Cedex 15
Tél. (33) 140 58 77 31
Fax (33) 140 58 77 00
E-mail: europublications@journal-officiel.gouv.fr
URL: http://www.journal-officiel.gouv.fr

IRELAND

Alan Hanna's Bookshop

270 Lower Rathmines Road
Dublin 6
Tel. (353-1) 496 73 98
Fax (353-1) 496 02 28
E-mail: hannas@iol.ie

ITALIA

Licosa SpA

Via Duca di Calabria, 1/1
Casella postale 552
I-50125 Firenze
Tel. (39) 055 64 83 1
Fax (39) 055 64 12 57
E-mail: licosa@licosa.com
URL: http://www.licosa.com

LUXEMBOURG

Messageries du livre SARL

5, rue Raiffeisen
L-2411 Luxembourg
Tél. (352) 40 10 20
Fax (352) 49 06 61
E-mail: mail@mdl.lu
URL: http://www.mdl.lu

NEDERLAND

SDU Servicecentrum Uitgevers

Christoffel Plantijnstraat 2
Postbus 20014
2500 EA Den Haag
Tel. (31-70) 378 98 80
Fax (31-70) 378 97 83
E-mail: sdu@sdu.nl
URL: http://www.sdu.nl

PORTUGAL

Distribuidora de Livros Bertrand Ld.ª

Grupo Bertrand, SA
Rua das Terras dos Vales, 4-A
Apartado 60037
P-2700 Amadora
Tel. (351) 214 95 87 87
Fax (351) 214 96 02 55
E-mail: dlb@ip.pt

Imprensa Nacional-Casa da Moeda, SA

Sector de Publicações Oficiais
Rua da Escola Politécnica, 135
P-1250-100 Lisboa Codex
Tel. (351) 213 94 57 00
Fax (351) 213 94 57 50
E-mail: spoce@incm.pt
URL: http://www.incm.pt

SUOMI/FINLAND

**Akateeminen Kirjakauppa/
Akademiska Bokhandeln**

Keskuskatu 1/Centralgatan 1
PL/PB 128
FIN-00101 Helsinki/Helsingfors
P./tfn (358-9) 121 44 18
F./fax (358-9) 121 44 35
Sähköposti: sps@akateeminen.com
URL: http://www.akateeminen.com

SVERIGE

BTJ AB

Traktorvägen 11-13
S-221 82 Lund
Tlf. (46-46) 18 00 00
Fax (46-46) 30 79 47
E-post: btjeu-pub@btj.se
URL: http://www.btj.se

UNITED KINGDOM

The Stationery Office Ltd

Customer Services
PO Box 29
Norwich NR3 1GN
Tel. (44) 870 60 05-522
Fax (44) 870 60 05-533
E-mail: book.orders@theso.co.uk
URL: http://www.itsofficial.net

ÍSLAND

Bokabud Larusar Blöndal

Skólavördustig, 2
IS-101 Reykjavik
Tel. (354) 552 55 40
Fax (354) 552 55 60
E-mail: bokabud@simnet.is

SCHWEIZ/SUISSE/SVIZZERA

Euro Info Center Schweiz

c/o OSEC Business Network Switzerland
Stampfenbachstraße 85
PF 492
CH-8035 Zürich
Tel. (41-1) 365 53 15
Fax (41-1) 365 54 11
E-mail: eics@osec.ch
URL: http://www.osec.ch/eics

BÅLGARIJA

Europress Euromedia Ltd

59, blvd Vitosha
BG-1000 Sofia
Tel. (359-2) 980 37 66
Fax (359-2) 980 42 30
E-mail: Milena@mbox.cit.bg
URL: http://www.europress.bg

CYPRUS

Cyprus Chamber of Commerce and Industry

PO Box 21455
CY-1509 Nicosia
Tel. (357-2) 88 97 52
Fax (357-2) 66 10 44
E-mail: demetrap@ccci.org.cy

EESTI

Eesti Kaubandus-Tööstuskoda

(Estonian Chamber of Commerce and Industry)
Toom-Kooli 17
EE-10130 Tallinn
Tel. (372) 646 02 44
Fax (372) 646 02 45
E-mail: einfo@koda.ee
URL: http://www.koda.ee

HRVATSKA

Mediatrade Ltd

Pavla Hatza 1
HR-10000 Zagreb
Tel. (385-1) 481 94 11
Fax (385-1) 481 94 11

MAGYARORSZÁG

Euro Info Service

Szt. István krt.12
III emelet 1/A
PO Box 1039
H-1137 Budapest
Tel. (36-1) 329 21 70
Fax (36-1) 349 20 53
E-mail: euroinfo@euroinfo.hu
URL: http://www.euroinfo.hu

MALTA

Miller Distributors Ltd

Malta International Airport
PO Box 25
Luqa LQA 05
Tel. (356) 66 44 88
Fax (356) 67 67 99
E-mail: gwirth@usa.net

NORGE

Swets Blackwell AS

Hans Nielsen Hauges gt. 39
Boks 4901 Nydalen
N-0423 Oslo
Tel. (47) 23 40 00 00
Fax (47) 23 40 00 01
E-mail: info@no.swetsblackwell.com
URL: http://www.swetsblackwell.com.no

POLSKA

Ars Polona

Krakowskie Przedmiescie 7
Skr. pocztowa 1001
PL-00-950 Warszawa
Tel. (48-22) 826 12 01
Fax (48-22) 826 62 40
E-mail: books119@arspolona.com.pl

ROMÂNIA

Euromedia

Str.Dionisie Lupu nr. 65, sector 1
RO-70184 Bucuresti
Tel. (40-1) 315 44 03
Fax (40-1) 312 96 46
E-mail: euromedia@mailcity.com

SLOVAKIA

Centrum VTI SR

Nám. Slobody, 19
SK-81223 Bratislava
Tel. (421-7) 54 41 83 64
Fax (421-7) 54 41 83 64
E-mail: europ@tbb1.sltk.stuba.sk
URL: http://www.sltk.stuba.sk

SLOVENIJA

GV Zalozba

Dunajska cesta 5
SLO-1000 Ljubljana
Tel. (386) 613 09 1804
Fax (386) 613 09 1805
E-mail: europ@gvestnik.si
URL: http://www.gvzalozba.si

TÜRKIYE

Dünya Infotel AS

100, Yil Mahallessi 34440
TR-80050 Bagcilar-Istanbul
Tel. (90-212) 629 46 89
Fax (90-212) 629 46 27
E-mail: aktuel.info@dunya.com

ARGENTINA

World Publications SA

Av. Cordoba 1877
C1120 AAA Buenos Aires
Tel. (54-11) 48 15 81 56
Fax (54-11) 48 15 81 56
E-mail: wpbooks@infovia.com.ar
URL: http://www.wpbooks.com.ar

AUSTRALIA

Hunter Publications

PO Box 404
Abbotsford, Victoria 3067
Tel. (61-3) 94 17 53 61
Fax (61-3) 94 19 71 54
E-mail: jpdavies@ozemail.com.au

BRESIL

Livraria Camões

Rua Bittencourt da Silva, 12 C
CEP
20043-900 Rio de Janeiro
Tel. (55-21) 262 47 76
Fax (55-21) 262 47 76
E-mail: livraria.camoes@incm.com.br
URL: http://www.incm.com.br

CANADA

Les éditions La Liberté Inc.

3020, chemin Sainte-Foy
Sainte-Foy, Québec G1X 3V6
Tel. (1-418) 658 37 63
Fax (1-800) 567 54 49
E-mail: liberte@mediom.qc.ca

Renouf Publishing Co. Ltd

5369 Chemin Canotek Road, Unit 1
Ottawa, Ontario K1J 9J3
Tel. (1-613) 745 26 65
Fax (1-613) 745 76 60
E-mail: order.dept@renoufbooks.com
URL: http://www.renoufbooks.com

EGYPT

The Middle East Observer

41 Sherif Street
Cairo
Tel. (20-2) 392 69 19
Fax (20-2) 393 97 32
E-mail: inquiry@meobserver.com
URL: http://www.meobserver.com.eg

MALAYSIA

EBIC Malaysia

Suite 45.02, Level 45
Plaza MBf (Letter Box 45)
8 Jalan Yap Kwan Seng
50450 Kuala Lumpur
Tel. (60-3) 21 62 92 98
Fax (60-3) 21 62 61 98
E-mail: ebic@tm.net.my

MÉXICO

Mundi Prensa México, SA de CV

Río Pánuco, 141
Colonia Cuauhtémoc
MX-06500 México, DF
Tel. (52-5) 533 56 58
Fax (52-5) 514 67 99
E-mail: 101545.2361@compuserve.com

SOUTH AFRICA

Eurochamber of Commerce in South Africa

PO Box 781738
2146 Sandton
Tel. (27-11) 884 39 52
Fax (27-11) 883 55 73
E-mail: info@eurochamber.co.za

SOUTH KOREA

**The European Union Chamber of
Commerce in Korea**

5th Fl, The Shilla Hotel
202, Jangchung-dong 2 Ga, Chung-ku
Seoul 100-392
Tel. (82-2) 22 53-5631/4
Fax (82-2) 22 53-5635/6
E-mail: eucck@eucck.org
URL: http://www.eucck.org

SRI LANKA

EBIC Sri Lanka

Trans Asia Hotel
115 Sir Chittampalam
A. Gardiner Mawatha
Colombo 2
Tel. (94-1) 074 71 50 78
Fax (94-1) 44 87 79
E-mail: ebicsl@slnet.ik

T'AI-WAN

Tycoon Information Inc

PO Box 81-466
105 Taipei
Tel. (886-2) 87 12 88 86
Fax (886-2) 87 12 47 47
E-mail: euitupe@ms21.hinet.net

UNITED STATES OF AMERICA

Bernan Associates

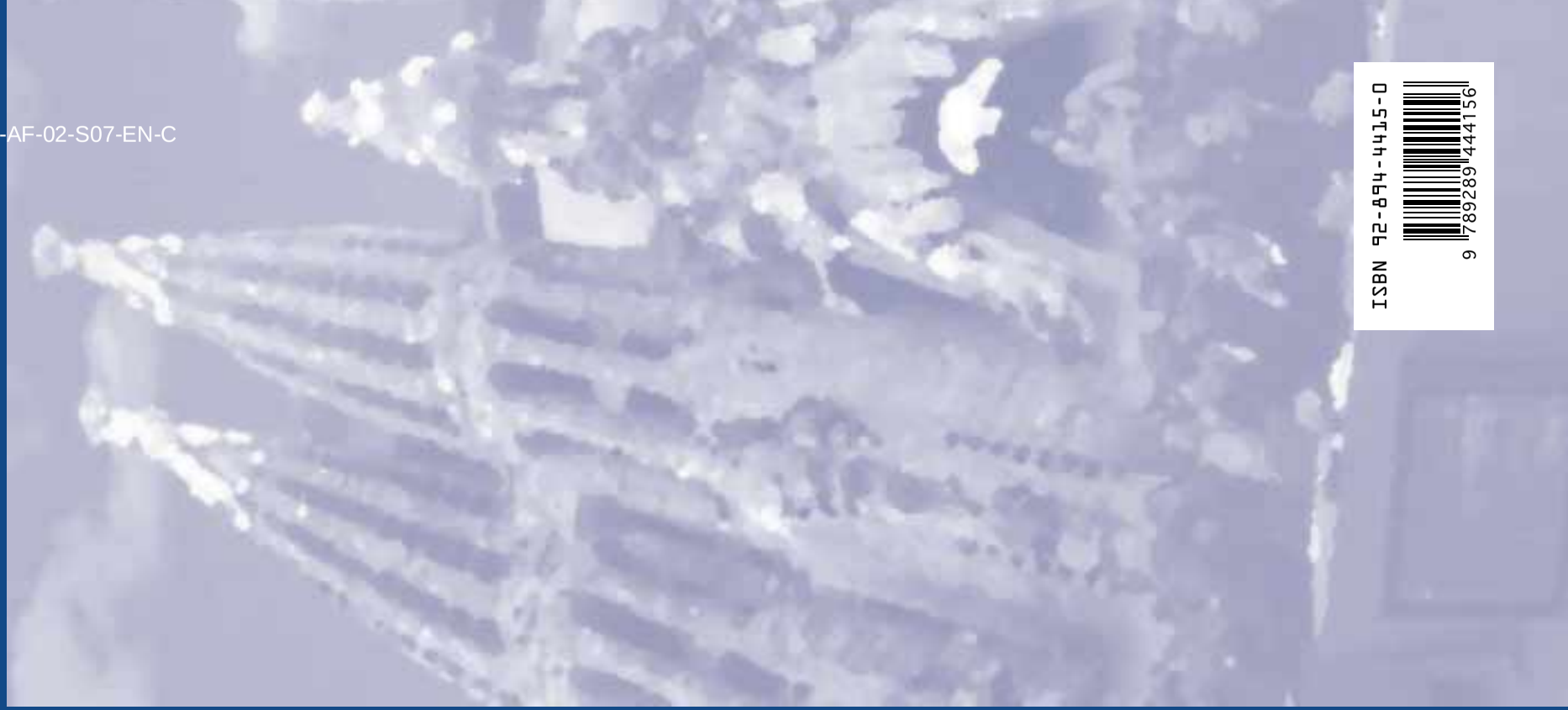
4611-F Assembly Drive
Lanham MD 20706-4391
Tel. (1-800) 274 44 47 (toll free telephone)
Fax (1-800) 865 34 50 (toll free fax)
E-mail: query@bernan.com
URL: http://www.bernan.com

ANDERE LÄNDER
OTHER COUNTRIES
AUTRES PAYS

Bitte wenden Sie sich an ein Büro Ihrer Wahl/Please contact the sales office of your choice/Veuillez vous adresser au bureau de vente de votre choix

Office for Official Publications of the European Communities
2, rue Mercier
L-2985 Luxembourg
Tel. (352) 29 29-42455
Fax (352) 29 29-42758
E-mail: info-info-opoce@cec.eu.int
URL: publications.eu.int

2/2002



ISBN 92-894-4415-0
9 789289 444156

Publications Office