Summary of AIIM Industry White Papers on DLM Document Lifecycle Management

Electronic Archives are the Memory of the Information Society

The Information Society impacts in many different ways on the European citizen, the most visible being the provision of access to information services and applications using new digital technologies. Economic competitiveness of Europe’s technology companies and the creation of new knowledge-rich job opportunities are key to the emergence of a true European digital economy. Equally, the Information Society must reinforce the core values of Europe’s social and cultural heritage – supporting equality of access, social inclusion and cultural diversity. One important element in ensuring a sound balance between these economic and social imperatives is cooperation between the information and communication industries and public institutions and administrations.

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.

For this, a series of six ICT Industry White Papers was produced. I am sure that the reader will find these White Papers both relevant and valuable, both as a professional and as a European citizen.

Erkki Liikanen
Member of the Commission for Enterprise and Information Society

Six Industry White Papers on Records, Document and Enterprise Content Management for the Public Sector

The Industry White Papers are published by the DLM-Forum of the European Commission and AIIM International Europe. They address the needs of public administration and archives at the European, national, federal and local level and to educate the public sector throughout Europe about available solutions for archival problems on relevant topics about acquisition, management, long term storage, multilingual access, indexing and training issues. The leading suppliers of Enterprise Content Management technologies participate in this series and focus on electronic archival, document management and records management for the public sector in the European Community. The Industry White Papers are as follows:

(1) Capture, Indexing & Auto Categorisation (Hewlett Packard)
(2) Conversion & Document Formats (FileNET)
(3) Content Management (IBM)
(4) Access & Protection (Kodak)
(5) Availability & Preservation (TRW Systems Europe/University College London/comunicando)
(6) Education, Training & Operation

Take a look at the following overviews – and get your White Paper at the DLM-Forum in May.

John Mancini
President of AIIM International

John L. Symon
Sr. Vice President AIIM International Europe
We are currently faced with an ever-increasing overload of information and must decide how we will go about mastering it. 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. To sum it up, the standard has already been set and recognised as natural language, but where are the systems which have adapted this natural standard?

Table of content

1. Introduction
2. The importance of safe indexing
   2.1 Description of the problem
   2.2 The challenge of rapidly growing document volumes
   2.3 The quality of indexing defines the quality of retrieval
   2.4 The role of metadata for indexing and information exchange
   2.5 The need for quality standards, costs and legal aspects
3. Methods for indexing and auto-categorisation
   3.1 Manual, semi-automatic and automatic indexing and categorisation methods
   3.2 Auto-categorisation methods
   3.3 Extraction methods
   3.4 Handling different types of information and document representations
4. The Role of Databases
   4.1 Database types and related indexing
   4.2 Indexing and Search methods
   4.3 Natural languages
5. Standards for Indexing
   5.1 Relevant standards for indexing and ordering methods
   5.2 Metadata structuring through XML
   5.3 Relevant standardisation bodies and initiatives
6. Best practice Applications
   6.1 Automated distribution of incoming documents: Statistical Office of the Free State of Saxony
   6.2 Knowledge-Enabled Content Management: CHIP Online
7. Outlook
   7.1 Citizen Portals
   7.2 Natural Language based Portals
Glossary
Abbreviations
References
Authoring Company

© DLM-Forum 2002 / AIIM Europe 2002
Hewlett Packard

Conversion & Document Formats

Backfile conversion and format issues for information stored in digital archives

This 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.

Table of content

1. Introduction
2. The Bottleneck of Converting Existing Archives
   2.1 Conquering the document mountain
   2.2 Overcoming perceptions
   2.3 Keeping pace with technology
   2.4 Lost Information
   2.5 Conversion Costs
   2.6 Legality
   2.7 Records management vs. document management
3. Content in the Digital Age
   3.1 Paper original
   3.2 Raster or vector imaging
   3.3 Scanning ‘norms’
   3.4 Compression
   3.5 Eliminating image ‘noise’
   3.6 Raster image characteristics
   3.7 Renditioning
   3.8 Adaptive thresholding
   3.9 Annotations
   3.10 Recognition techniques
   3.11 Colour and greyscale
   3.12 Microfilm, Microfiche
   3.13 Electronic objects
   3.14 Forms
4. Information Capture Technologies and Methods
   4.1 In-house or bureau?
   4.2 Traditional archives
   4.3 COLD/ERM Enterprise Report Management
   4.4 Forms capture
   4.5 Electronic documents
   4.6 Planning a Document Conversion
5. Standards for Formats
   5.1 Tagged Image File Format (TIFF)
   5.2 Joint Photographic Experts Group (JPEG)
   5.3 Graphics Interchange Format (GIF)
   5.4 Moving Pictures Expert Group (MPEG)
   5.5 Audio Video Interleave (AVI)
   5.6 Adobe Acrobat Portable Document Format (PDF)
6. Best Practice Applications
   6.1 Department of Forestry, Baden-Württemberg
   6.2 Case Sanctuary Housing Association
   6.3 Staffordshire County Council
   6.4 Levy Gee
7. Outlook
   7.1 Formats
   7.2 Conversion Strategy
   7.3 Archive Value
   7.4 Technology Developments
Glossary
Abbreviations
References
Authoring Company

© DLM-Forum 2002 / AIIM Europe 2002
FileNET
Content Management
Managing the Lifecycle of Information

This 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 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.

Table of Content
1. Introduction
2. From Archival to Enterprise Content Management
   2.1 Getting a Handle on Managing Information
   2.2 Records Management
   2.3 Archival
   2.4 Content Management
   2.5 Knowledge Management
   2.6 Distributed Solutions
   2.7 Implications of the Internet
   2.8 Content Management Today
   2.9 Enterprise-Wide Compatibility
   2.10 Costs
   2.11 Legal Aspects
3. Content Management Architectures
   3.1 The Anatomy of Content Management
   3.2 Scalability: Expanding a System
   3.3 Performance: Multi-Threading and More
   3.4 High Availability
   3.5 Storage Management: Cost effective data storage
   3.6 LDAP Support
   3.7 Making the Connection: How Content Services Communicate
   3.8 WebDAV
   3.9 Digital Signatures
4. Functionality of Content Management Solutions
   4.1 Version Control
   4.2 Process Management and Workflow
   4.3 Rendition Support for Effective Publication
   4.4 Publishing and Web Publishing
   4.5 Replication Services
4.6 Compound Documents - Complexity with Control
4.7 Record Retention and Document Lifecycle Management
4.8 Openness: Interoperability with Industry Standards
4.9 Security: Multiple Levels
4.10 Additional Security Features
4.11 Audit Trail
4.12 Internationalisation and Localisation
4.13 System Administration
5. Content Management Standards
   5.1 XML as an Interoperability Standard
   5.2 Document Management and Records Management Standards
   5.3 Workflow Industry Standards
   5.4 Document Imaging Industry standards
   5.5 Other Content Management Related Standards and Guidance
   5.6 ANSI/AIIM Guidance
6. Best Practice Applications
   6.1 Online Services via a Web Portal
   6.2 Language Services and Payroll Management
   6.3 Digital Processing for Tax Documents
   6.4 Other Government Applications
7. Outlook – Content Management and e-Government Today
   7.1 e-Government…
   7.2 …and Beyond e-Government
   7.3 e-Merging Technologies
   7.4 The Future of Content Management
Glossary
Abbreviations
References
Authoring Company

© DLM-Forum 2002 / AIIM Europe 2002
IBM

Access & Protection

Managing Open Access & Information Protection

In this White Paper the following key topics for user and information access will be addressed. Issues regarding litigation, privacy protection and networks 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. Which standards are relevant to user access and 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 paper will also describe the main drivers for architectural change.

Table of content

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Information access in Europe</td>
<td>2.2 A common framework for information interchange</td>
<td>3.1 The portal as a paradigm</td>
<td>4.1 Security issues as market drivers</td>
<td>5.1 European Union’s, “Model Requirements for the Management of Electronic Records” (MoReq)</td>
</tr>
<tr>
<td>2.3 Drivers for open access</td>
<td>2.4 Litigatory exposure</td>
<td>3.2 Usage scenarios - corporate, personal, marketplace portals</td>
<td>4.2 Management vs. Retention</td>
<td>5.2 United States’, Department of Defense (DoD) 5015.2 Standard, “Design Criteria Standard for Electronic Records Management Software Applications”</td>
</tr>
<tr>
<td>2.5 Privacy-protection versus behaviour tracking</td>
<td>2.6 Digital rights protection</td>
<td>3.3 Information aggregation</td>
<td>4.3 Emergence of virtual documents and virtual records</td>
<td>5.3 United Kingdom’s, Public Records Office (PRO) “Functional Requirements for Electronic Records Management Systems”</td>
</tr>
<tr>
<td>3. Accessing Public Information</td>
<td>3.4 Web content</td>
<td>3.5 Search and mining</td>
<td>4.4 Audit trails</td>
<td>5.4 Norwegian, National Archives, NOARK-4 , “Functional Description, Requirements and Specifications for Recordkeeping Systems”</td>
</tr>
<tr>
<td>3.9 Metadata</td>
<td>3.10 The role of XML</td>
<td>3.11 The enterprise content management challenge</td>
<td>4.6 Common solution requirements</td>
<td>ISO 11179: Specification and Standardization of Data Elements</td>
</tr>
<tr>
<td>3.12 Presentation support</td>
<td>3.13 Application Services</td>
<td>3.14 Collaboration</td>
<td>4.7 Emerging issues</td>
<td>6. Best Practice Applications</td>
</tr>
<tr>
<td>3.15 Personalisation, strategies and tools</td>
<td>4. Protecting Public Information</td>
<td>6.1 The Open Digital Administration project – cities of Naestved and Skurup</td>
<td>7. Outlook</td>
<td></td>
</tr>
<tr>
<td>4.1 Security issues as market drivers</td>
<td>4.2 Management vs. Retention</td>
<td>6.2 The Keen Project</td>
<td>7.1 Proven strategies</td>
<td></td>
</tr>
<tr>
<td>4.3 Emergence of virtual documents and virtual records</td>
<td>4.4 Audit trails</td>
<td>6.3 Statens Museum for Kunst – The National Danish Art Museum</td>
<td>7.2 Technology benefits</td>
<td></td>
</tr>
<tr>
<td>4.5 Transaction integrity through electronic signatures</td>
<td>4.6 Common solution requirements</td>
<td>7.3 Critical Success Factors</td>
<td>7.3 Trends</td>
<td></td>
</tr>
<tr>
<td>4.7 Emerging issues</td>
<td>4.8 Consolidated and Syndicated Content</td>
<td>7.4 Glossary</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© DLM-Forum 2002 / AIIM Europe 2002
This Industry White 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, analog 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.

Table of content

1. Introduction
2. Electronic Archives Are a Virtual Memory of the Information Society
   2.1 Digital Document Overview
   2.2 Accumulating digital records of an online age
   2.3 Preservation for availability across generations
   2.4 The importance of content and form
   2.5 Balancing processibility and inalterability
   2.6 Digital Preservation with a Reference Archive
   2.7 Current long-term preservation strategies
   2.8 Conclusion
3. Long-Term Availability
   3.1 Reference Archive overview
   3.2 Process overview
   3.3 Application Requirements
   3.4 Benefits of Analogue Rendering
   3.5 Challenges of Analogue Rendering
   3.6 Conclusion
4. Constant Migration of Information
   4.1 Migration overview
   4.2 Process overview
   4.3 Application Requirements
   4.4 Benefits of Migration
   4.5 Challenges of Migration
   4.6 Analysis of digital preservation based solely on migration
   4.7 Conclusion
5. Standards for Reference Archive Media and Reference Archiv Management Software
   5.1 Reference Archive Media
   5.2 Reference Archive Management Software
   5.3 Conclusion
6. Best Practice Applications
   6.1 UK Census 2001 Dual-Path Record Keeping Project of UK Office of National Statistics
   6.2 Records Management of State of Virginia Land Project of Library of Virginia & County Clerk’s Offices
   6.3 Conclusion
7. Outlook
   7.1 Digital Preservation strategy and process overview
   7.2 Benefits of Digital Preservation with Reference Archive protection
   7.3 Future developments
   7.4 Solutions availability
   7.5 Conclusion

© DLM-Forum 2002 / AIIM Europe 2002
TRW / UCL / comunicando

Education, Training & Operation
From the Traditional Archivist to the Information Manager

This White Paper looks at issues of education and training 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 which are propelling the E-learning requirements of the digital economy, discusses some of the potential benefits of E-learning and argues that businesses and corporate institutions in the 21st century must have and implement a learning and training vision.

Table of content

1. Introduction
2. Archivists and Information Managers
   2.1 Organisational and social trends and electronic record keeping
   2.2 International developments and management of electronic records
3. Managing Digital Archives
   3.1 The value of content
   3.2 Digital archives and record systems
   3.3 On-line education and training
   3.4 Challenges of e-Learning for universities and students
   3.5 Content creation and access
4. Education and Training Requirements in the digital world
   4.1 Drivers to e-Learning
   4.2 The learning gap: demand and supply
   4.3 Technology and standards development
   4.4 Reflection around systems integration and investments
5. The role of computer based training and e-Learning
6. Best Practice Applications
   6.1 Computer Based or e-Learning
   6.2 The wider arena of e-Learning
   6.3 Characteristics and benefits of e-Learning
   6.4 Technologies, tools and methodology in content creation and delivery
7. Outlook
   7.1 Need for a strategy
   7.2 From strategy to implementation
   7.3 Change management
Glossary
Abbreviations
References
Authoring Companies

© DLM-Forum 2002 / AIIM Europe 2002
The series of six Industry White Papers is published to address the needs of public administration and archives at the European, national, federal and local level. They are designed to educate the public sector throughout Europe concerning available solutions for archival problems on relevant topics covering acquisition, management, long term storage, multilingual access, indexing and training issues.

DLM-Forum 2002

The current DLM acronym stands for Données Lisibles par Machine (Machine Readable Data). It is proposed that after the DLM-Forum 2002 in Barcelona this definition be broadened to embrace the complete "Document Lifecycle Management". The DLM-Forum is based on the conclusions of the Council of the European Union, concerning greater co-operation in the field of archives (17 June 1994). The DLM-Forum 2002 in Barcelona will be the third multidisciplinary European DLM-Forum on electronic records to be organised. It will build on the challenge that the second DLM-Forum in 1999 issued to the ICT (Information, Communications & Technology) industry to identify and provide practical solutions for electronic document and records management.

The task of safeguarding and ensuring the continued accessibility of the European archival heritage in the context of the Information Society is the primary concern of the DLM-Forum on Electronic Records. The DLM-Forum asks industry to actively participate in the multidisciplinary effort aimed at safeguarding and rendering accessible archives as the memory of the Information Society and to improve and develop products to this end in collaboration with the users.

DLM-Forum 2002
Electronic Records
Scientific Committee Secretariat
European Commission SG.B.3
Office JECL 3/36, Rue de la Loi 200, B-1049 Brussels, Belgium
A/e: dlm-forum@cec.eu.int

AIIM International - The Enterprise Content Management Association

AIIM International is the leading global industry association that connects the communities of users and suppliers of Enterprise Content Management. A neutral and unbiased source of information, AIIM International produces educational, solution-oriented events and conferences, provides up-to-the-minute industry information through publications and its industry web portal, and is an ANSI/ISO-accredited standards developer.

AIIM Europe is member of the DLM-Monitoring Committee and co-ordinates the activities of the DLM/ICT-Working Group.

AIIM International, Europe
Chappell House, The Green, Datchet, Berkshire SL3 9EH, UK
http://www.aiim.org

Industry White Papers on Records, Document and Enterprise Content Management

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Capture, Indexing &amp; Auto Categorisation</td>
<td>ISBN 3-936534-00-4</td>
</tr>
<tr>
<td>(3) Content Management</td>
<td>(FileNET) ISBN 3-936534-02-0</td>
</tr>
<tr>
<td>(4) Access &amp; Protection</td>
<td>(IBM) ISBN 3-936534-03-9</td>
</tr>
<tr>
<td>(5) Availability &amp; Preservation</td>
<td>(Kodak) ISBN 3-936534-04-7</td>
</tr>
<tr>
<td></td>
<td>ISBN 3-936534-07-1</td>
</tr>
</tbody>
</table>

© DLM-Forum 2002 / AIIM Europe 2002