

Harmonization Efforts in Digital Archiving in Europe



Zoltán Szatucsek

The 5th International Conference on Electronic Records Management & Technology

Taipei, 13.12.2016

1. Guidelines, specifications...
Research on why they are not having the desired effect?
2. Make better use of existing Groups like DLM-Forum (Digital Records)
3. Dissemination and engagement of the ~~U.E.~~ E.A.G on existing electronic record initiatives such as E-Ark.
4. Participation on wider perspectives on Digital Preservation (p.e. MSEG on Digital Preservation).

DLM and the European Commission

COUNCIL AND MINISTERS OF CULTURE
RESOLUTIONS of 14 November 1991 (91/C 314/02)

- set up **a group of experts** for the purpose of examining to what extent greater coordination of archives policy and practice within the Community is desirable
- **progress report** before the end of 1992

ARCHIVES IN THE EUROPEAN UNION 1994 'BLACK BOOK'

- Warranted coordination and cooperation on a European level was the need for adequate management and storage of computerised archives.

COUNCIL OF THE EUROPEAN UNION
CONCLUSIONS of 17 June 1994

- concerning greater cooperation in the field of archives (94/C 235/03)
- **organizing 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**



Evolution

Name

1994 Documents Lisibles par Machine (Machine readable documents)
2002 Document Lifecycle Management

First activities

1996 - Electronic Records - Co-operation Europe-Wide
1999 - European citizens and electronic information: the memory of the Information Society, Cooperation Europe-wide
2002 - Access and preservation of electronic information: best practices and solutions

Organization

2002 DLM Network – Monitoring Committee
2003 DLM Network EEIG
2010 DLM Forum Foundation



Goals

- to establish a strategic network to bring together individuals and organisations from both the public sector involved in Document Lifecycle Management
- to provide a formal and sustainable entity to support the foundation members in fulfilling their roles and responsibilities and to raise awareness in the field of DLM
- to provide technology and knowledge transfer
- to provide information services
- to provide education and and skills development
- to provide research
- to provide independent standards, appropriate practical guidelines and benchmark indicators to government and industry



44 members included 20 NAs

AIT Austrian Institute of Technology GmbH
 ARMA International
 City of Rotterdam
 Estonia Dept of Information
 Head of Archives and Records European Central Bank
 ES Solution
 Finnish Municipal RM Association
 Gimmel LLC
 GORDIC spol s.r.o
 HP Autonomy Systems Ltd.
 INESC-ID
 Lulea University of Technology
 Ministry of Interior of Czech Republic
 Mikkeli University of Applied Sciences
 National Archive of Austria
 National Archive of Denmark
 National Archives of Estonia
 National Archives of Finland
 National Archives of Germany
 National Archives of Hungary
 National Archives of Ireland
 National Archives of Latvia

National Archives of Luxembourg
 National Archives of Malta
 National Archives of Norway
 National Archives of Poland
 National Archives of Scotland
 National Archives of Slovenia
 National Archives of Spain
 National Archives of Sweden
 National Archives of The Netherlands
 National Archives of the UK
 National Records & Archives Authority
 Dr Peter Faester Nielsen
 Office of Chief Archivist of Lithuania
 Open Text Corporation
 Rasmus Greve
 RecordPoint Software
 RSD
 Russian Records Managers Guild
 State Enterprise Centre of Registers
 Preservica
 Torben Juul
 University of Brighton
 Ministry of Justice, The Netherlands

DLM Events

Member meetings/AGMs

Every six months/Presidency of European Commission

Triennials

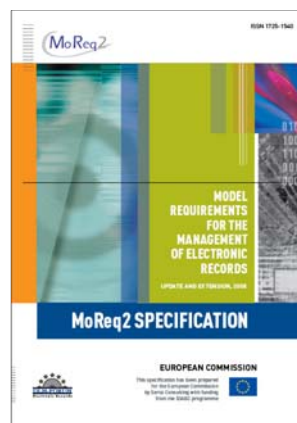
- 1996 Brussels
- 1999 Brussels
- 2002 Barcelona
- 2005 Budapest
- 2008 Toulouse
- 2011 Brussels
- 2014 Lisbon
- 2017 Brighton



Modular Requirements for the Management of Electronic Records



Moreq



Moreq2 (2008)



Moreq 2010

E-ARK

- Consortia of DLM members and other „digital preservation minded“ partners
 - DLM Forum Foundation
 - Digital Preservation Coalition
 - 5 national archives (Denmark, Estonia, Norway, Slovenia, Hungary)
 - 4 institutes / universities (incl. University of Portsmouth as project coordinator)
 - 3 software providers (digital repository and EDRMS)
 - 2 additional public sector institutions as „end-users“
- E-ARK = European Archival Records and Knowledge Preservation
- Content in short: Proposing pan-european standardisation of export, transfer, ingest, preservation and access workflows and formats for born-digital records

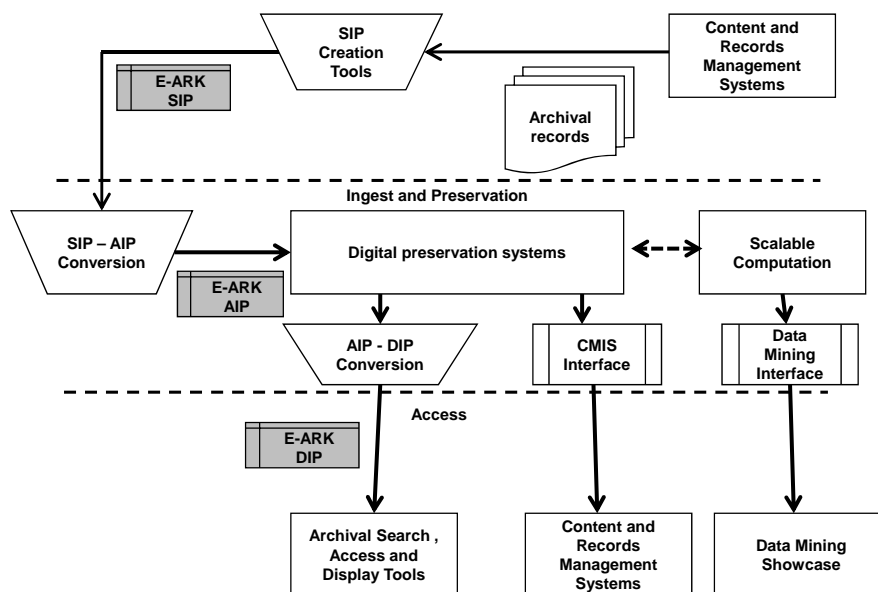


E-ARK: Background

- Budapest: DLM Members' Meeting discussion on harmonization
- Copenhagen: DLM vision document, What does successful Information Governance across Europe look like? What role should the DLM Forum take?
 - 7.2 Establishing the DLM Forum as an organisation that bridges across the different sub-disciplines of Information Governance
 - 7.5 Participating in European projects related to Information Governance. „DLM will seek to contribute to the EU Digital Agenda for Europe and to participate in relevant EU projects“
 - 7.7 Building on the legacy of MoReq2010 DLM Forum establishes leadership, authority and referenceability over the broader field of Information Governance.
- Zagreb: Call for partners issued at DLM Members' Meeting
- Dublin: Presentation about the proposal

Consortium members

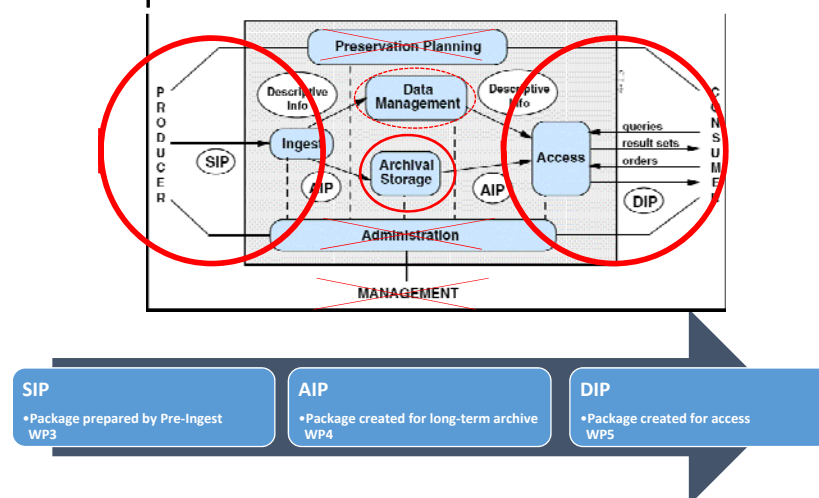
- UNIVERSITY OF PORTSMOUTH/UNIVERSITY OF BRIGHTON - United Kingdom
- Austrian Institute of Technology – Austria
- Ministrstvo za izobraževanje, znanost in sport – Slovenia
- DLM Forum Foundation – Europe
- Statens Arkiver – Denmark
- Digital Preservation Coalition - United Kingdom
- Universitaet zu Koeln – Germany
- Instituto Superior Tecnico – Portugal
- National Archives of Hungary – Hungary
- Rahvusarhiv – Estonia
- Riksarkivet – Norway
- ES Solutions – Sweden
- Magenta ApS – Denmark
- KEEP Solutions – Portugal
- Agencia Para a Modernizacao Administrativa IP – Portugal
- Ministerio de Hacienda y Administraciones Publica – Spain



Timeline

- February 2014 – kick off in Lisbon
- August 2014 – best practice report
- January 2015 – first drafts of SIP/AIP/DIP specifications
- January 2016 – end of pilots and final specifications
- January 2017 – implementation of tools
- January 2017 – end of the 36 months project

E-ARK: Scope



E-ARK: the ambition

Vision: All digital preservation systems receive, store and provide access to information regardless of its size, type or format according to a set of agreed principles which allow systems to identify, verify and validate the information in a uniform way

Goal: Interoperability between data sources, archives and reuse environments is improved to a point where digital preservation tools can be reused across borders and institutions. This opens up new possibilities for collaboration and limits the need for development resources for any single institution



E-ARK: the ambition

The scope: All digital archiving systems: receive, store and provide access to data

The data: regardless of its: size, type or format

Harmony: a set of agreed **principles** which allow systems to identify, verify and validate the data in a uniform way



Outcomes

Standardisation of available best-practices

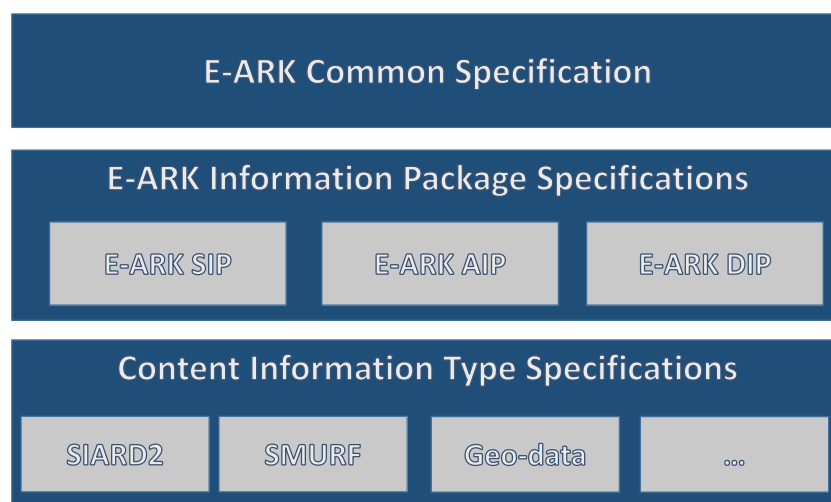
- Common OAIS SIP, AIP and DIP format specifications
- Pre-ingest, ingest and access workflows

Open source tools

- to implement the specifications
- to be scalable, modular, robust, computational, and adaptable
- to be implemented individually or as an integrated reference implementation

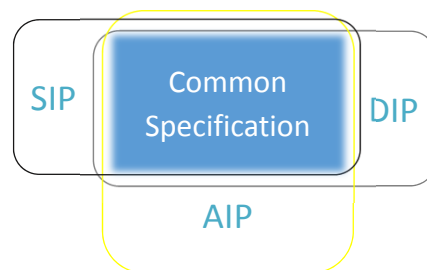


E-ARK Specifications



E-ARK IPs

- E-ARK SIP
 - Administrative metadata for ingest and pre-ingest processes
 - Transfer agreement
- E-ARK AIP
 - Preservation metadata
 - Additional structural layer to allow keeping the original SIP untouched
- E-ARK DIP
 - Administrative metadata (log the creation of the DIP)
 - Description of complex access environments in PREMIS

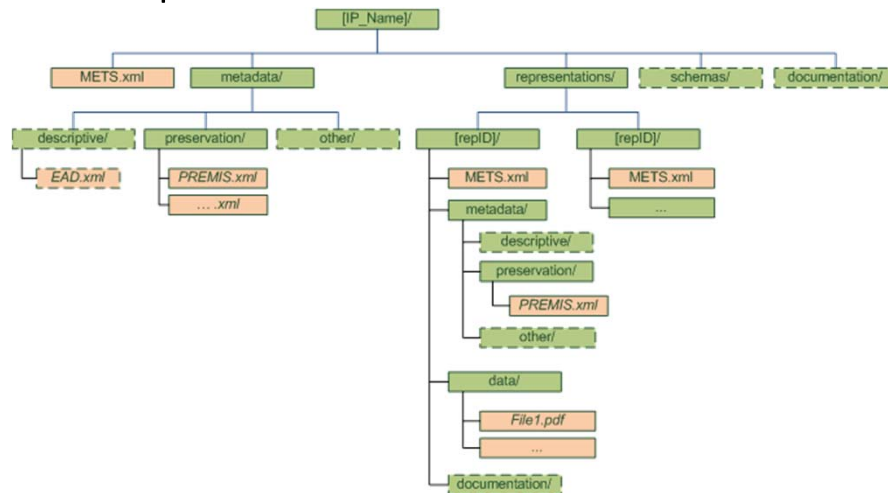


Common Specification

- **E-ARK Common Specification for Information Packages**
 - Provides core principles for any IP
 - Most crucial: a common structure
 - Influenced by the Swedish CS
 - Aims to support automation of high-level package identification and validation (integrity, technical validity, etc.)
 - Is built on widely used XML standards (METS, PREMIS)
 - Concentrates on structural metadata, includes elements of administrative and technical metadata
 - Is flexible enough to be used for any type of data
 - Allows for further specification and localisation if needed



Common Specification: structure



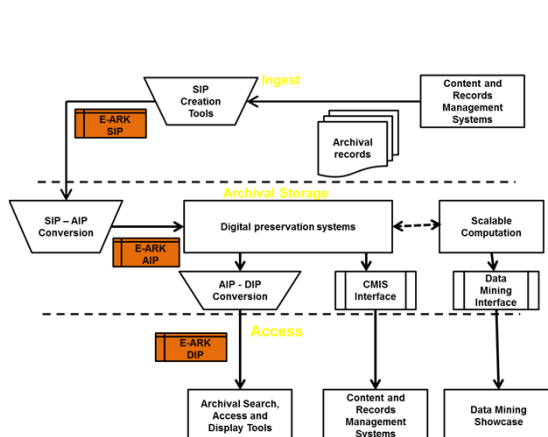
E-ARK Content Information Types

- E-ARK Content Information Types allow description of the details of the package content
 - Descriptive metadata (e.g. sub-specification of EAD, MARC, CIDOC-CRM)
 - Technical metadata (e.g. specific requirements for digital image metadata)
 - Internal structure of the data folder (e.g. the physical structure of series and records in an ERMS-derived Information Package)
 - Specific representation information (e.g. define the composition of files which constitute an archived relational database and the appropriate viewer)
- Content Information Types can be
 - Data specific (for databases, websites etc.)
 - Metadata specific (for EAD, MARC, Dublin Core etc.)

E-ARK Content Information Types

- E-ARK provides three data specific content type specifications
 - SIARD2
 - Specification for relational databases
 - Updates the original SIARD format in regard to improved scalability and standards support
 - In collaboration with the Swiss Federal Archives
 - SMURF (Semantically Marked-Up Records Format)
 - Specification for ERMS-exported and „loose“ records
 - Extends EAD with records management elements
 - Partly synchronised with MoReq2010
 - Geo-data
 - Specification to define a basic descriptive layer
 - Defines the structure and location of geo-info elements within the Information Package
- E-ARK provides a metadata specific content type for EAD
 - Possible to be used within any Information Package

E-ARK: tools



Sustainability

- E-ARK ends on 31st of January 2017 ..
- Maintenance of specifications ensured through the DLM Forum
 - Specific workgroup on standards currently being set up
 - Non-members of DLM welcome to join!
- Discussions to ensure the sustainability of tools in collaboration with the Open Preservation Foundation (OPF)



The Advisory Boards

- These boards will serve to **assess contributions to and from the project**
- The boards will be **open** to interested parties
- Meetings of these groups (annual face-to-face meeting with teleconference in between) will effectively **benchmark comments received** and provide a '**gateway review**' at various points in the meeting.
- These three external groups will be used to **adjudicate on conflicting views of the community should these arise – ensuring that the project listens to the views of its constituents**
- The information gathered from all these consultations via normal knowledge elicitation methods (questionnaires, in depth surveys etc.) will form the basis of a **comprehensive knowledge base** that will be kept by the DLM Forum as an ongoing resource and single point of reference

Archival Advisory Board

- Archives of Emilia-Romagna, Italy
- Directorate-General of the Book, of Archives & of Libraries, Portugal
- EC Archives & Records Management
- EC Historical Archives
- German Federal Archives
- National Archives of Bulgaria
- National Archives of Finland
- National Archives of France
- National Archives of Sweden
- National Archives of the Netherlands
- Polish Data Archive
- Queensland State Archives
- Swiss Federal Archives
- UK National Archives
- UK Parliamentary Archives

Commercial / Technical Advisory Board

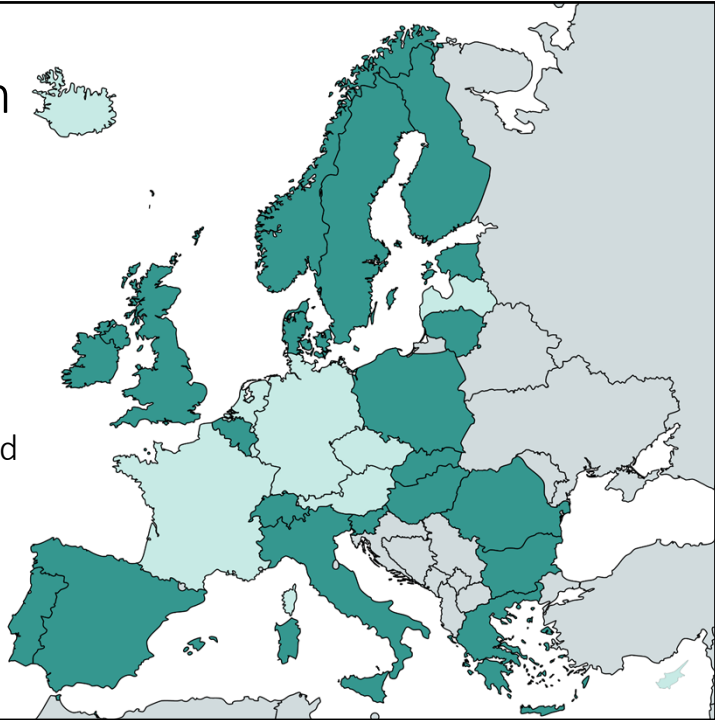
- Arkivum
- ARMA Europe
- DigitalForever
- Discovery Garden
- Microsoft Research
- Open Preservation Foundation
- Open Text Initiative
- Preservica
- Versity

Data Provider Advisory Board

- Danish Agency for Digitisation
- Estonian Ministry of Economic Affairs & Communication
- Estonian Unemployment Insurance Fund
- James Lappin, RM Consultant

DLM report to EAG on Digital Archiving

- Conducted in autumn 2016
- Finished on 25 of November
- 21/30 EAG members responded
- Current projects, activities, goals
- Focuses on challenges



DLM report - Current challenges

- Archives are more and more involved in e-government initiatives
- In government everything is digital by now.
- Open and reusable public data is regarded by Digital Single Market strategy as a driver for economic growth
- Centralized government services and open data portals are emerging.
- Digital preservation is beyond its infancy.
- Archives are part of information industry
- Budget constraints are not in line with this



DLM report - Recommendations

- R.1. Change how people think about archives
- R.2. Automate more
- R.3. Collaborate more
- R.4. Educate more
- R.5. Win the argument for funding



Thank you for your
attention!

Zoltán Szatucsek
chair, DLM Forum Foundation
szatucsek.zoltan@mnl.gov.hu

