



eArchiving: 10 years and still going strong

Reflections, achievements, milestones and challenges

Janet Anderson, 29th May 2024



Photo Baltimore castle, Ireland. J. Anderson

Overview

Reflections, achievements, milestones and challenges in:

- The E-ARK project
- The eArchiving Building Block
- The eArchiving Initiative

European **A**rchival **R**ecords and **K**nowledge Preservation

www.eark-project.eu

The E-ARK Project



E-ARK in a nutshell <http://www.eark-project.eu/>



- **FP7 Pilot B PSP CIP**, Grant no. 620998, 3 years, Feb 2014-Jan 2017
- **5 National Archives**: Estonia, Slovenia, Denmark, Norway, Hungary, (Sweden)
- **4 research institutions**: Universities: **Portsmouth then Brighton (Coordinator)**, Köln, Austrian Institute of Technology (AIT), Technical University Lisbon (INESC-ID)
- **3 SMEs**: KEEPS, ESS, Magenta
- **2 Government Home Offices**: AMA Portugal, MINHAP Spain
- **2 Membership organisations**: DPC, DLM

E-ARK in a nutshell <http://www.eark-project.eu/>



- **Project Officer:** Alina Senn, **DG CNECT, Luxembourg**
- **Technical Coordinator:** **Kuldar Aas, National Archives of Estonia**
- **3 Advisory Boards:** Technical, Commercial, Archival.
- **7 pilots in 6 countries**
- **Website** still in existence! <https://www.eark-project.com/>

E-ARK Website <https://www.eark-project.com/>



NOTICE: Historic website, not updated.

Search ... **GO**



E-ARK A European "Showcase" Project [Site Map](#)

HOME ABOUT NEWS EVENTS COMMUNITY PARTNERS RESOURCES PILOTS USE CASES STORIES

E-ARK

In co-operation with commercial systems providers, E-ARK created and piloted a pan-European methodology for electronic document archiving, synthesising existing national and international best practices, that will keep records and databases authentic and usable over time.

[Find out more](#)

NOTICE: This page is now static and will not have further updates. It's archived, available only for reference.

[@EARKProject on Twitter](#)

For the latest EARK news follow us on Twitter [@EARKProject](#)

[Follow @EARKProject](#)

Welcome to the E-ARK Project

(European Archival Records and Knowledge Preservation)

E-ARK was a multinational CIP FP7 pilot project that improved the methods and technologies of digital archiving, in order to achieve consistency on a Europe-wide scale.

Tackling a range of problems associated with independent record-keeping technologies, systems and practices, E-ARK benefited the development of internationally accessible archives through: the provision of technical specifications and tools, the development of an integrated archiving infrastructure, the demonstration of improved availability, access and use, and the rigorous analysis of aggregated sets of archival data.

Running from 1st February 2014 to 31st January 2017 E-ARK was co-funded by the European Commission under its ICT Policy Support

Latest News

E-ARK Rated "Excellent" by the EC - "A European Showcase Project" 03-04-2017

The European Commission's Year 3 Review of the E-ARK Project has rated it as "Excellent" and have...

E-ARK Year 3 Summary Report Published 16-03-2017

The E-ARK summary report of project activities in the final year of the project can be downloaded by...

E-ARK Partners



Athens Meeting June 2016



E-ARK mission

The E-ARK project goal is to **pilot** archival services to keep digital records authentic and usable based on current **best practices**. These will address the three main endeavours of an archive: **acquiring, preserving and enabling re-use of information**. The potential benefits for public agencies, citizens and business will be demonstrated by **providing easy and efficient access to the archived records**. Archival processes at a pan-European level will be **harmonized, supported by guidelines and recommended practices** that will cater for a range of data from different types of source including record management systems and databases. The E-ARK approach will be **public-facing, open source, robust, replicable and scalable**, and will address a wide range of organisations, taking full account of legal constraints.



E-ARK: goals

We have developed practical **interoperability specifications** for data preparation, ingest, preservation and access

We have developed a set of **interoperable tools** which implement the specifications and are reusable across borders



E-ARK: core benefits

*Extended possibilities for
collaboration in tool development*

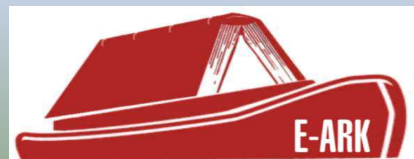
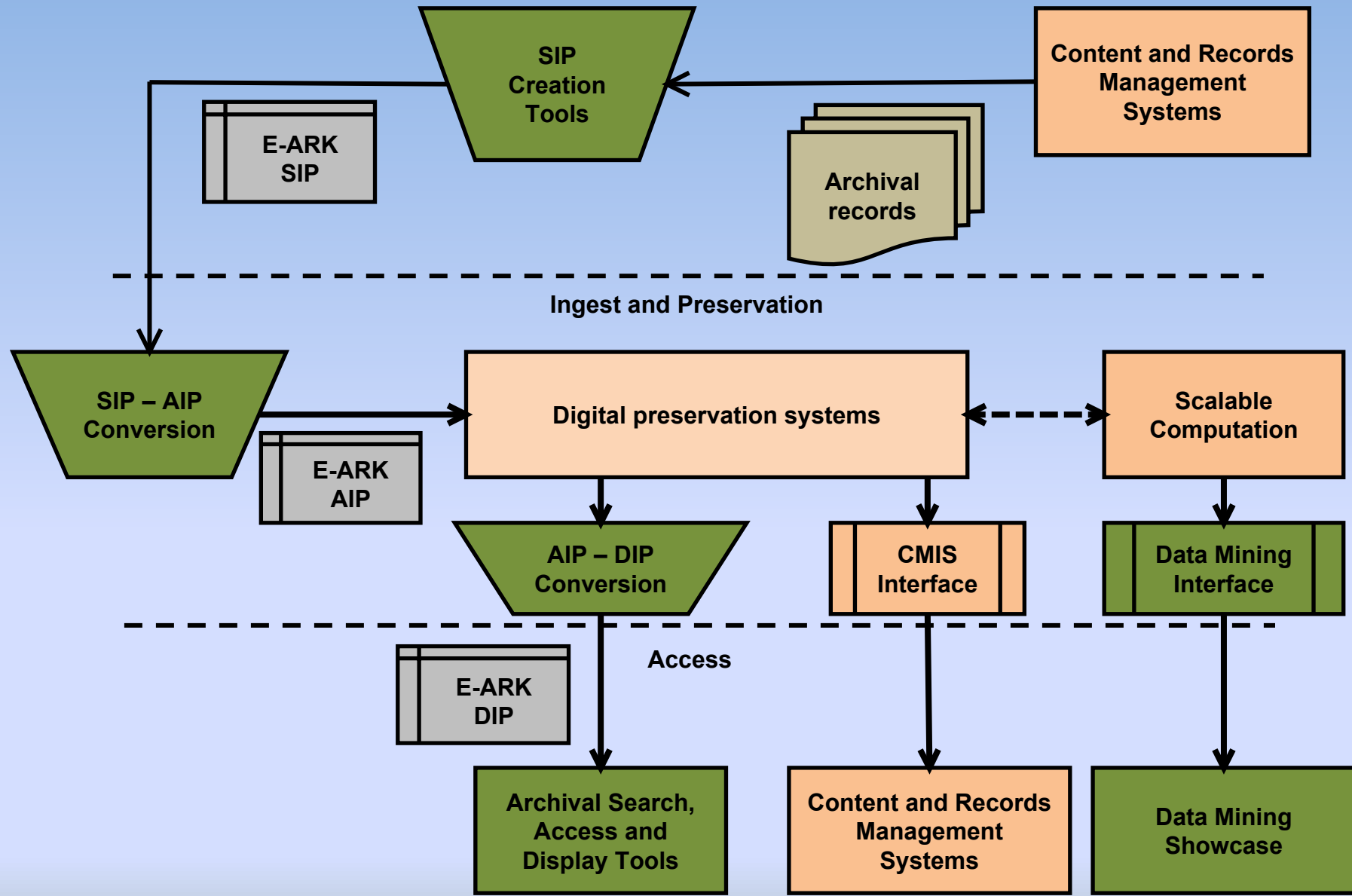


*Shared development = less cost
for each individual partner*



*Digital preservation is affordable
for everyone!*





Final Review March 2017



E-ARK Outcome Highlights and Achievements

- The **open source, digital archiving framework**, complete with accompanying metadata and other standards, has been thoroughly **tested** and has made a **significant impact on the institutions who carried out the pilots**
- The **Data Mining Showcase**: Data warehousing, Big Data
- We used a detailed **impact questionnaire** as the basis for carrying out interviews at the pilot sites

D6.3 Data Mining Showcase Popular

Published on 27 January 2017 • Modified on 22 December 2017 • By Clive Billenness • 1154 downloads

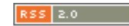
E-ARK has developed the Integrated Platform Reference Implementation (IPRIP) which is a software prototype for package creation, search, and access using data-centric technologies. The underlying concepts and technologies of the system have been detailed in past deliverables (D6.1 and D6.2). The main components of the IPRIP e-archiving environment include a scalable packaging infrastructure for creating OAIS Information Packages and a data-centric repository for searching and accessing data items on demand.

[Download](#)
(pdf, 3.12 MB)

620998 E-ARK D6.3 updated.pdf

This deliverable focuses on recent developments that support the deployment of the IPRIP in different configurations at E-ARK stakeholder sites. A flexible packaging mechanism combined with a standalone backend implementation enables custom single-server deployments on demand. The scalable, Hadoop-based backend implementation has been ported to the latest CDH (Cloudera Distribution Hadoop) distribution in order to support a recent technology stack, advanced data mining concepts, and enterprise demands. Showcases that dealt with the application of text mining approaches, the extraction and visualization of geographical information, the implementation of a database archiving workflow, and mass document ingest have been implemented and deployed at stakeholder sites in Hungary and Slovenia.

The E-ARK Project has now reached the end of its Research Phase On 31 January 2017, after 3 years, the...



[@EARKProject on Twitter](#)

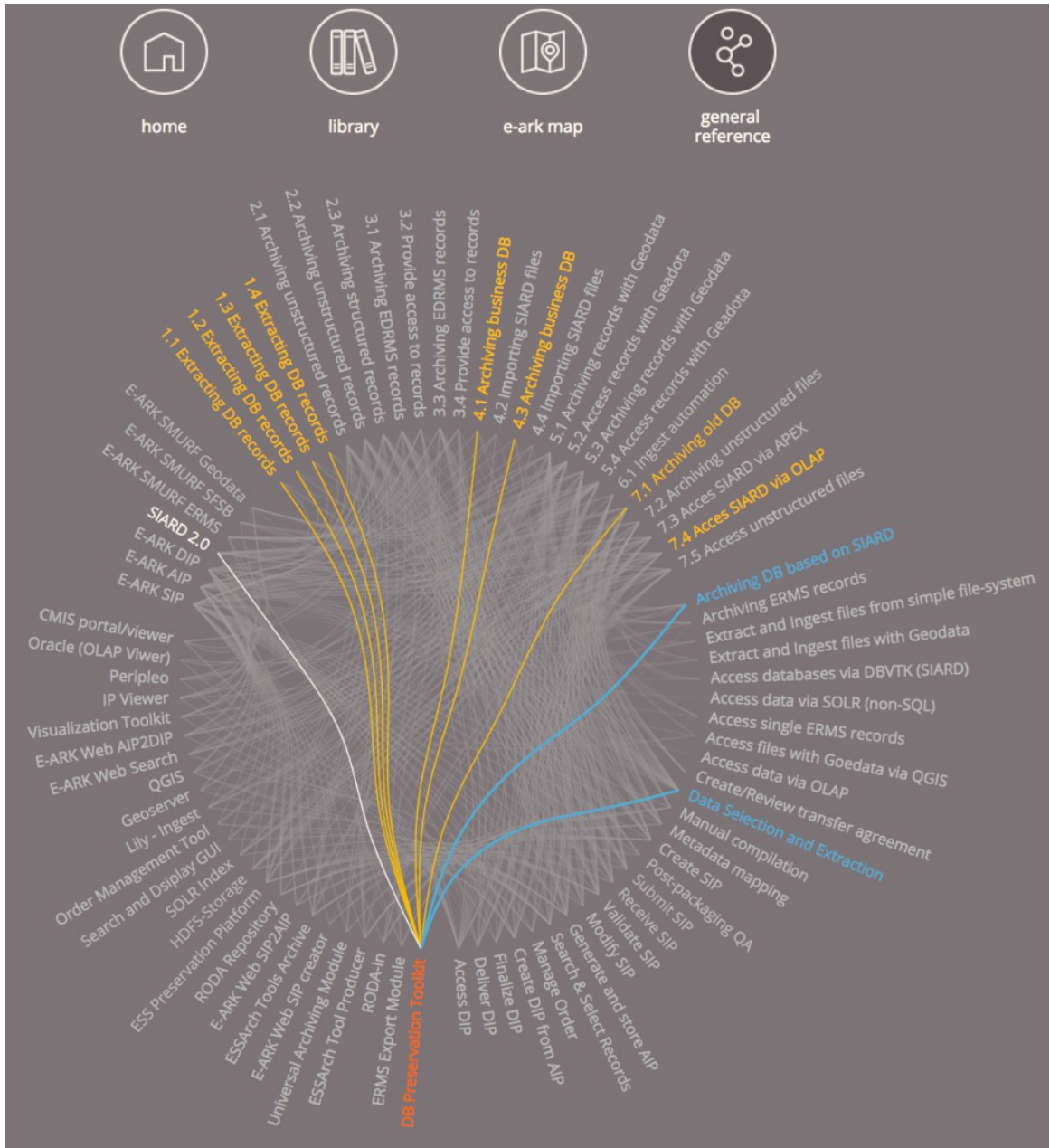
For the latest EARK news follow us on Twitter [@EARKProject](#)

[Follow @EARKProject](#)

E-ARK Outcome Highlights and Achievements

Highlights from this **impact analysis** include:

- **major savings** in the cost providing pre-ingest tools due to increased competition;
- **harmonisation of geo-spatial data** archiving practices which facilitate comparison of Natura 2000 sites across Europe;
- the benefits of using the **earkweb tool**, together with advanced search and data mining facilities, in national, regional and local archives;
- **robust, common standards and tools that can truly be used interchangeably across Europe.**



E-ARK General Model v2.0



The eArchiving Building

Block

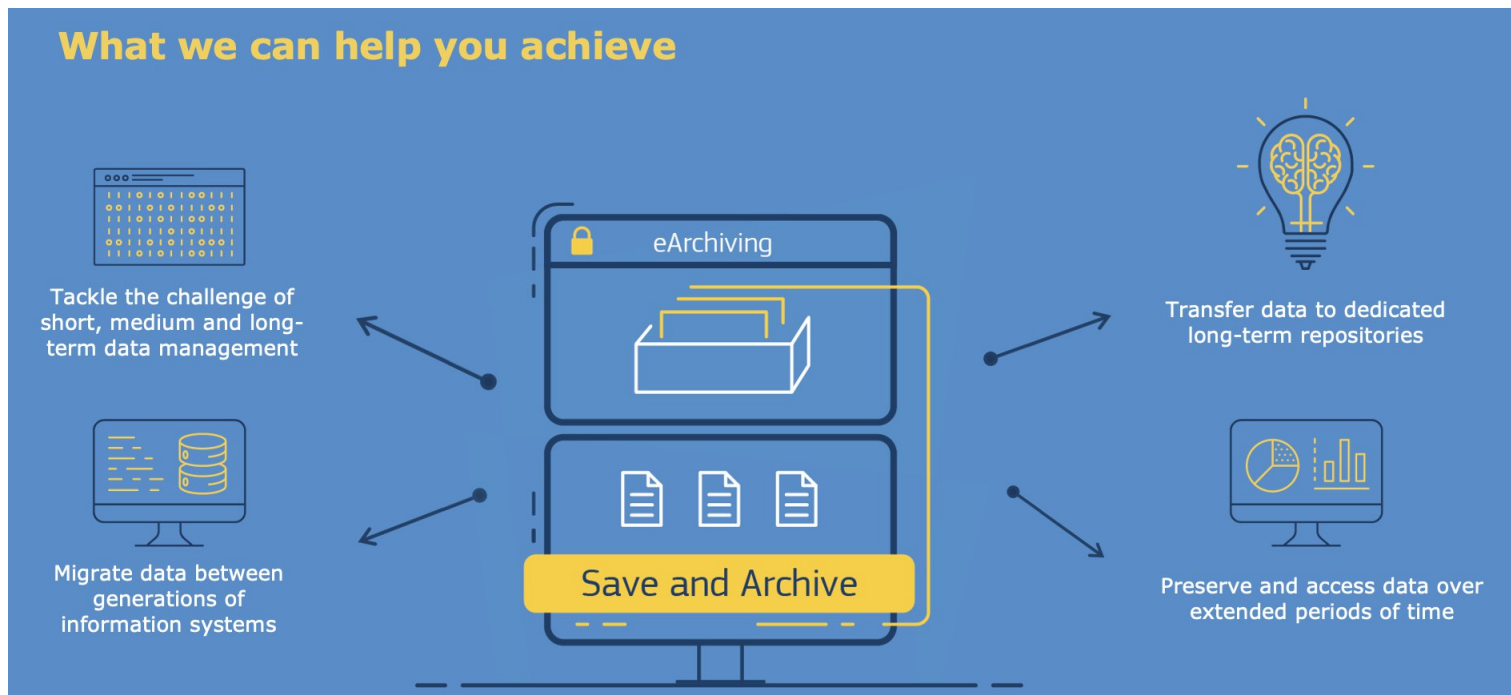
E-ARK4ALL 1st June 2018 - 31st May 2019

E-ARK3 1st November 2019 - 31st October 2021

eArchiving Building Block in a nutshell:



- Covered by two grants: E-ARK4ALL and E-ARK3
- Policy officers: **Manuela Speiser, Adelina Cornelia Dinu, Fulgencio Sanmartín**
- Expert Group: **European Archives Group (EAG)**



eArchiving Building Block in a nutshell: E-ARK4ALL



- eArchiving Building Block, E-ARK4ALL Agreement No. LC-00921441 CEF-TC-2018-15, part of the **Connecting Europe Facility Building Blocks**
 - **2 National Archives:** Denmark, Sweden
 - **1 Research Institution:** Austrian Institute of Technology (AIT)
 - **1 SME:** Keep Solutions
 - **1 Membership organization:** DLM Forum

DLM Project Members:

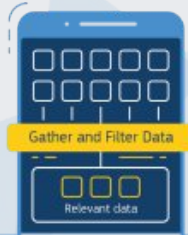
National Archives of Estonia, Norway, Slovenia; Federal Archives of Germany, Switzerland OPF, ES Solutions; Geoarh, Gregor Završnik, s.p.; Easy Lean OÜ; Poliphon; Mikkeli University of Applied Sciences; Instituto de Engenharia de Sistemas e Computadores (INESC-ID); Highbury Associates IVS; Gabinete UMBUS SL

Since Dec 2018



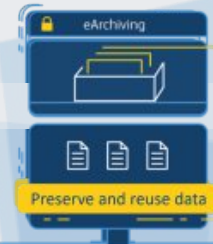
Big Data Test Infrastructure

Explore and experiment with big data for improved performance and decision making



Context Broker

Analyse, manage and share data, in real time, at the right time, throughout Europe



eArchiving

Facilitate the preservation, migration, reuse and trust of your data



European Blockchain Services Infrastructure

Harness the power of a European-wide network of blockchain services, increasing trust through data security, privacy and transparency



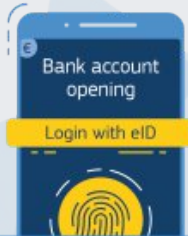
Once Only Principle

Re-use data held by Public Administrations.



eInvoicing

Promote the implementation of the European standard for electronic invoicing across borders



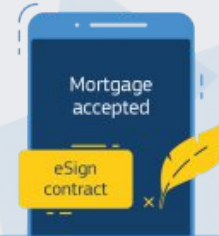
eID

Securely identify European citizens



eDelivery

Exchange online data and documents reliably and securely



eSignature

Create and verify electronic signatures between businesses and EU citizens



eTranslation

Offer machine translation to translate your documents and web content into any official EU language, Norwegian or Icelandic

Mapping of the eArchiving governance bodies to the CEF Governance

CEF POLICY

CEF Coordination Committee

CEF IT STRATEGY & COORDINATION

Architecture Management Board

CEF Telecom Expert Group

CEF DSIs IMPLEMENTATION

EC (DIGIT & CNECT)

E-ARK4ALL

DSI OPERATIONS
Operational Management Board

- DSI Owner
- DSI Solution Providers
- (Users: Commission Services)

DSI TACTICS
DSI Expert Group

- Policy Expert Group
- User Communities

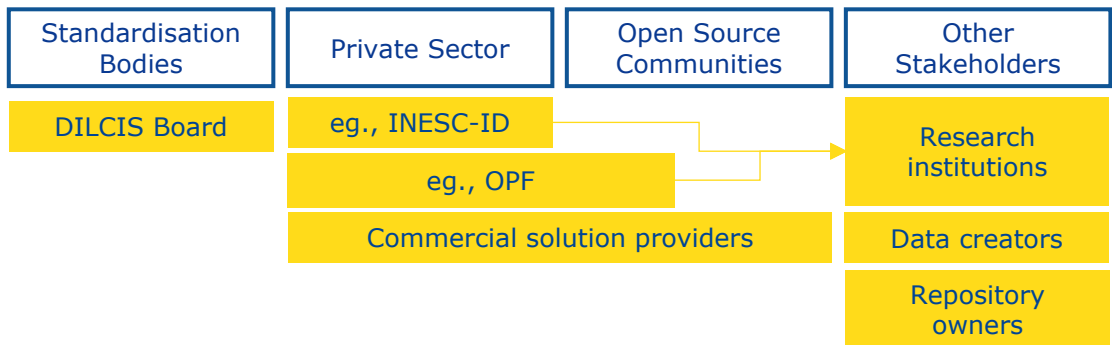
EAG

CEF DSIs STAKEHOLDERS

LEGEND

 eArchiving

 European Commission

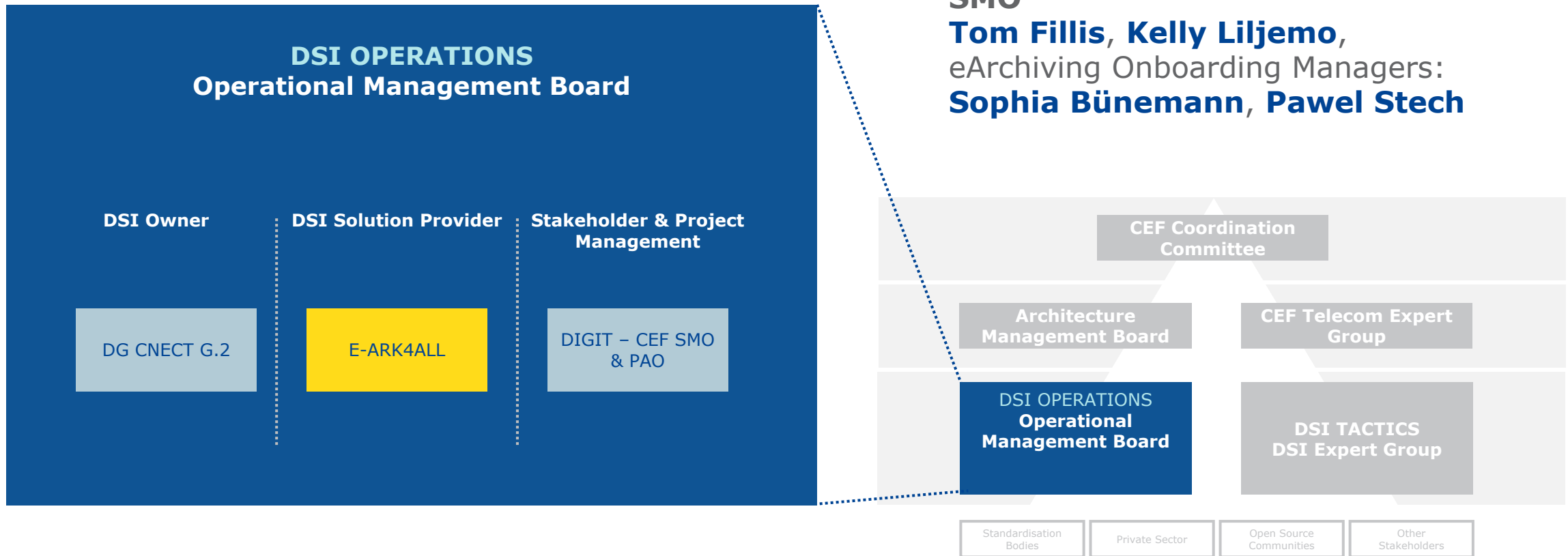


Proposed composition of the DSI Operational Management Board

LEGEND

 eArchiving

 European Commission

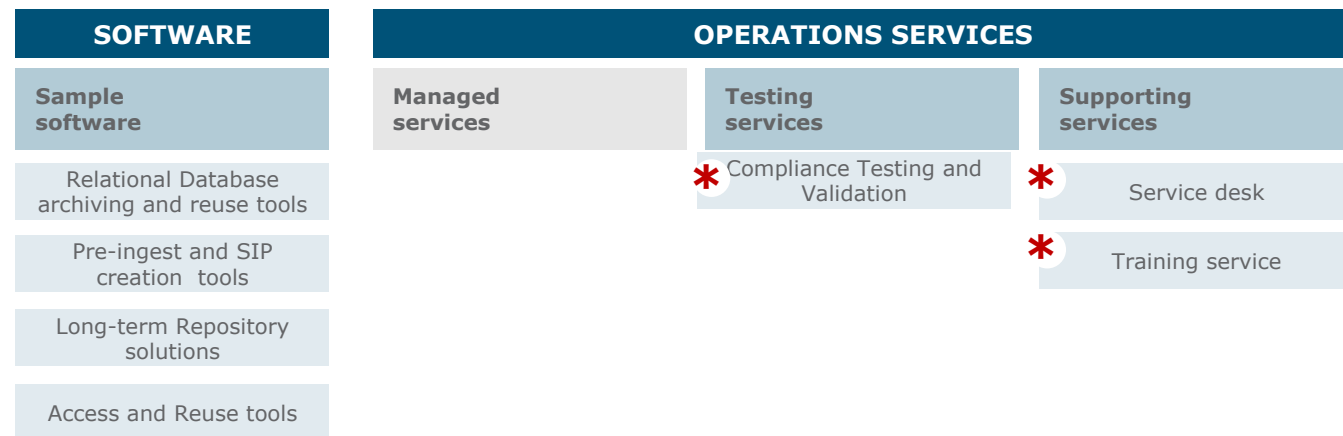


eArchiving Service Offering TO-BE (scenario proposed by CEF)

* This service is being developed in E-ARK4ALL

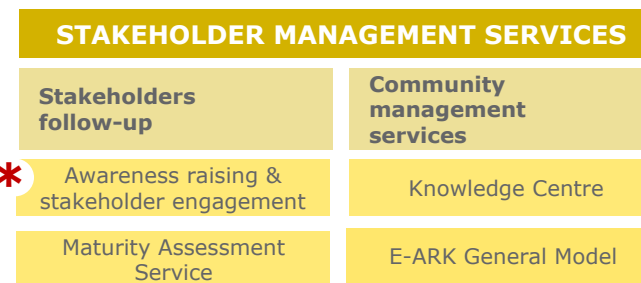
ENABLING SERVICES

enable the adoption of the core services

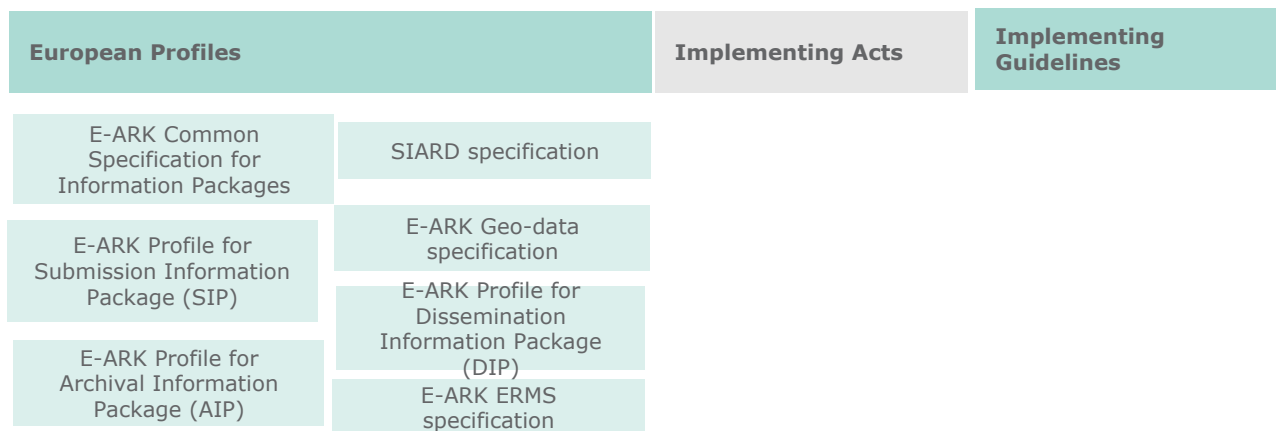


ENHANCING SERVICES

enhance customer experience



TECHNICAL SPECIFICATIONS



STANDARDS OF ESOs



AUDIENCE

- R Repositories
- D Data Creators
- S Solution Providers

CORE SERVICES

facilitate cross-border/ cross-sector technical interoperability among heterogeneous information systems

eArchiving Building Block in a nutshell: E-ARK3

<https://e-ark4all.eu/> eArchiving Building Block, E-ARK3 Agreement No. LC-01390244 CEF-TC-2019-3,



4 National Archives: Denmark, Estonia, Slovenia, Norway

3 research institutions: Austrian Institute of Technology (AIT),
Technical University Lisbon (INESC-ID), South-Eastern Finland
University of Applied Sciences (XAMK)

6 SMEs: KEEP Solutions, ES Solutions, PIQL, Highbury, Georh,
Gabinete,

2 Membership organisations: DLM, OPF

DLM Project Members:

National Archives of Finland, Federal Archives of Switzerland,
Poliphon, Sydarkivera

eArchiving Service Offering Canvas

ENABLING SERVICES *enable the adoption of the core services*

SOFTWARE		OPERATIONS SERVICES	
Sample software		Testing services	Supporting services
Relational Database archiving and reuse tools	Long-term Repository solutions	* Compliance Testing and Validation	Service desk
Pre-ingest and SIP creation tools	* Access and Reuse tools		* Training service

ENHANCING SERVICES *enhance customer experience*

STAKEHOLDER MANAGEMENT SERVICES	
Stakeholders follow-up	Knowledge Exchange
Stakeholders onboarding	Knowledge Centre
* Maturity Assessment Service	* Reference Architecture

E-ARK TECHNICAL SPECIFICATIONS

Specifications		
E-ARK Common Specification for Information Packages		
E-ARK Profile for Submission Information Package (SIP)	E-ARK Profile for Archival Information Package (AIP)	E-ARK Profile for Dissemination Information Package (DIP)
E-ARK SIARD CITS specification	E-ARK Geospatial CITS specification	E-ARK ERMS CITS specification
* E-ARK eHealth1 CITS specification	* E-ARK eHealth2 CITS specification	* E-ARK PREMIS CITS specification
* E-ARK GIS CITS specification	* E-ARK Archival Description CITS specification	

RELATED SPECIFICATIONS & STANDARDS

ISO 14721:2012 (OAIS)	EAD3
PREMIS	ISAAR-CPF
METS	EAC-CPF
eCH SIARD	ISDIAH
HL7 Family	EAG
ISAD-G	INSPIRE Directive
EAD	

LEGEND

* Services under development

AUDIENCE CONCERNED BY ALL SERVICES:

- A Archives
- P Data Producers
- S Solution Providers

CORE SERVICES *facilitate cross-border/cross-sector technical interoperability among heterogeneous archival information systems*

eArchiving Building Block: Outcome Highlights and Achievements

- Four new **Generic Services projects** in new fields such as Holocaust studies and Blockchain, bringing dependencies and responsibilities.
- **Collaboration** with eSignature, eDelivery, CEN TC/468 etc.
- eArchiving BB recommended in **Digital Services Act 2020**
- Participation in **high profile CEF events** which promoted eArchiving.

The ICTD - TFD recommendations

Recommendation 1

Use the CEF **eSignature** building block or its associated managed service, **EU Sign** to create, verify and extend electronic signatures and seals

Recommendation 2

Use the CEF **eDelivery** building block or its associated managed service, **EU Send** to exchange data and documents in a secure way

Recommendation 3

Use the CEF **eID** building block or its associated managed service, **EU Login** to enable electronic authentication of users

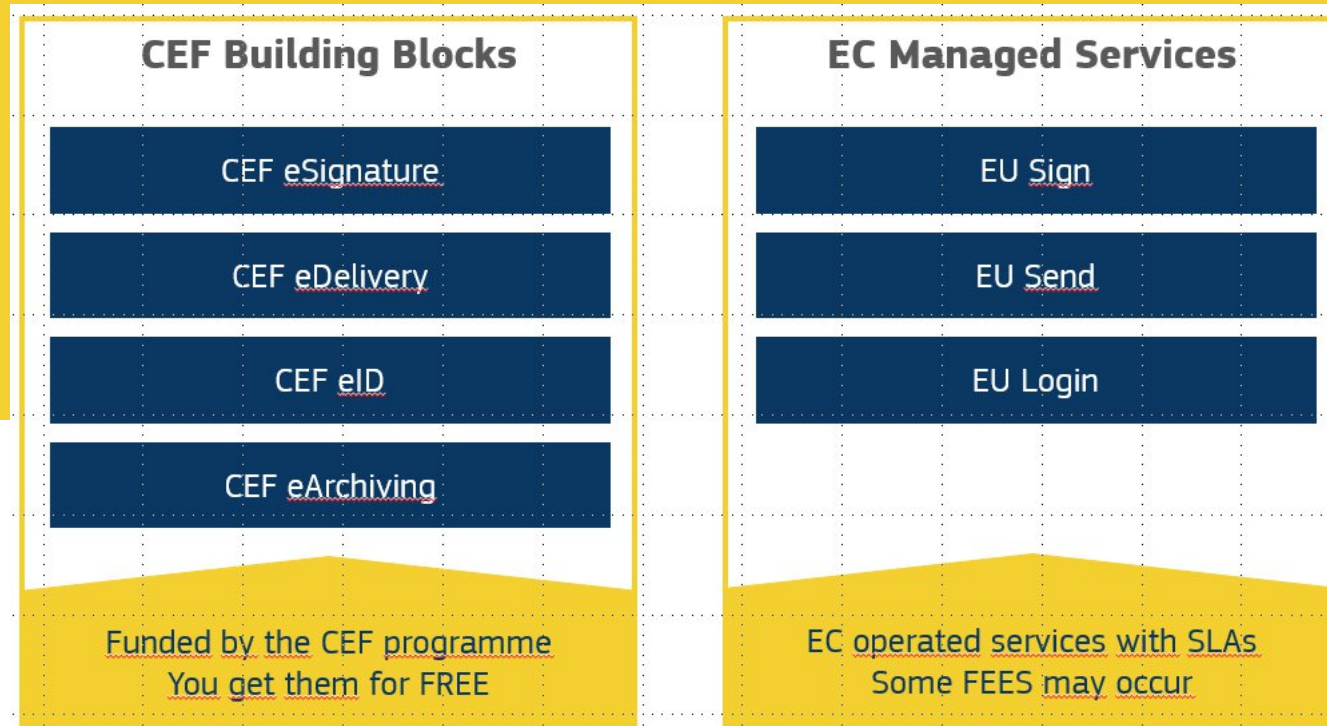
Recommendation 4

Use the CEF **eArchiving** building block to store and preserve digital information in the long-term



The EC Reusable Solutions are here to assist you in this process

“CEF building blocks have allowed the Member States to achieve an impressive level of interoperability. The success of EUIBAs in this same field is of benefit to the respective organisations, the Member States and the Union as a whole.”



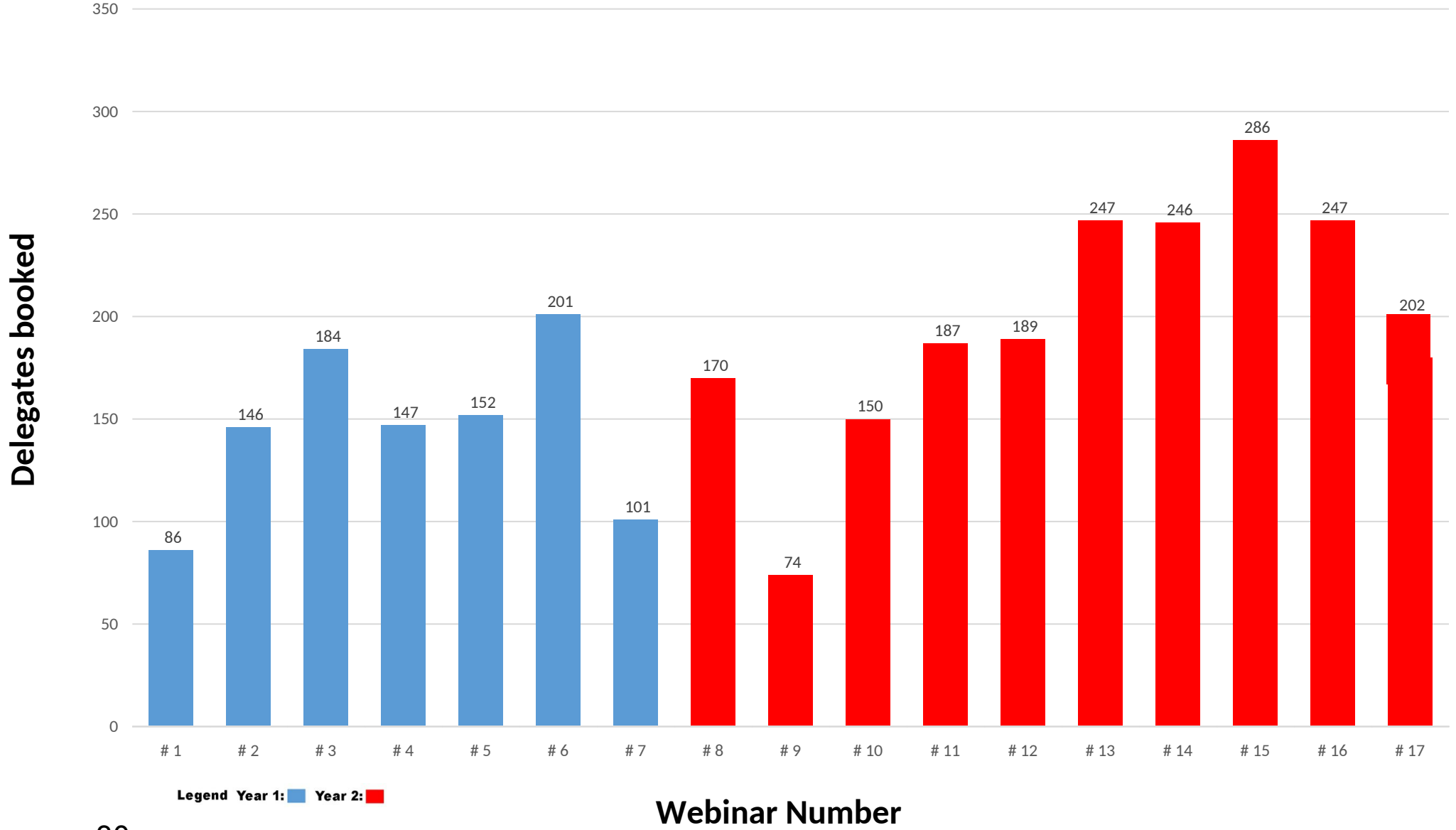
eArchiving Building Block: Outcome Highlights and Achievements

- Training webinars, wide range of topics, large, loyal audience
- Onboarding – large user community
- Success stories
- Reference Architecture



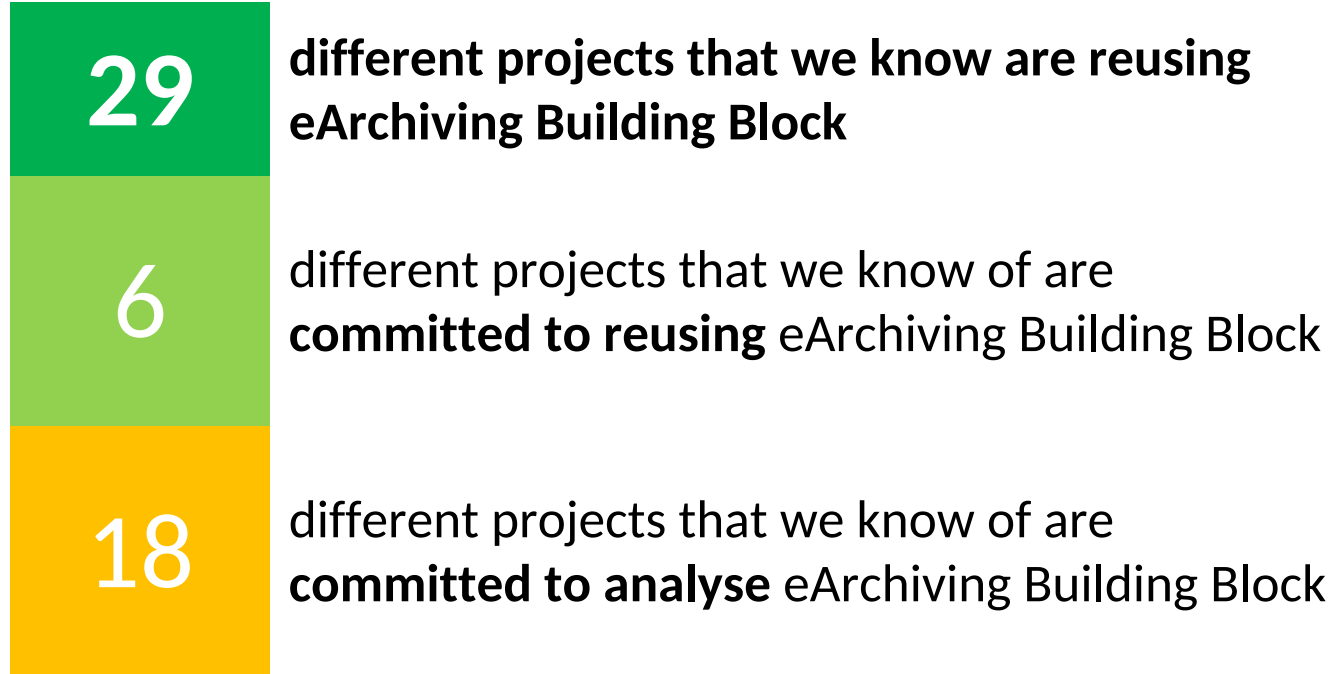
©Photo courtesy of the Central State Archive, Italy

E-ARK3 Training webinar bookings



User Engagement

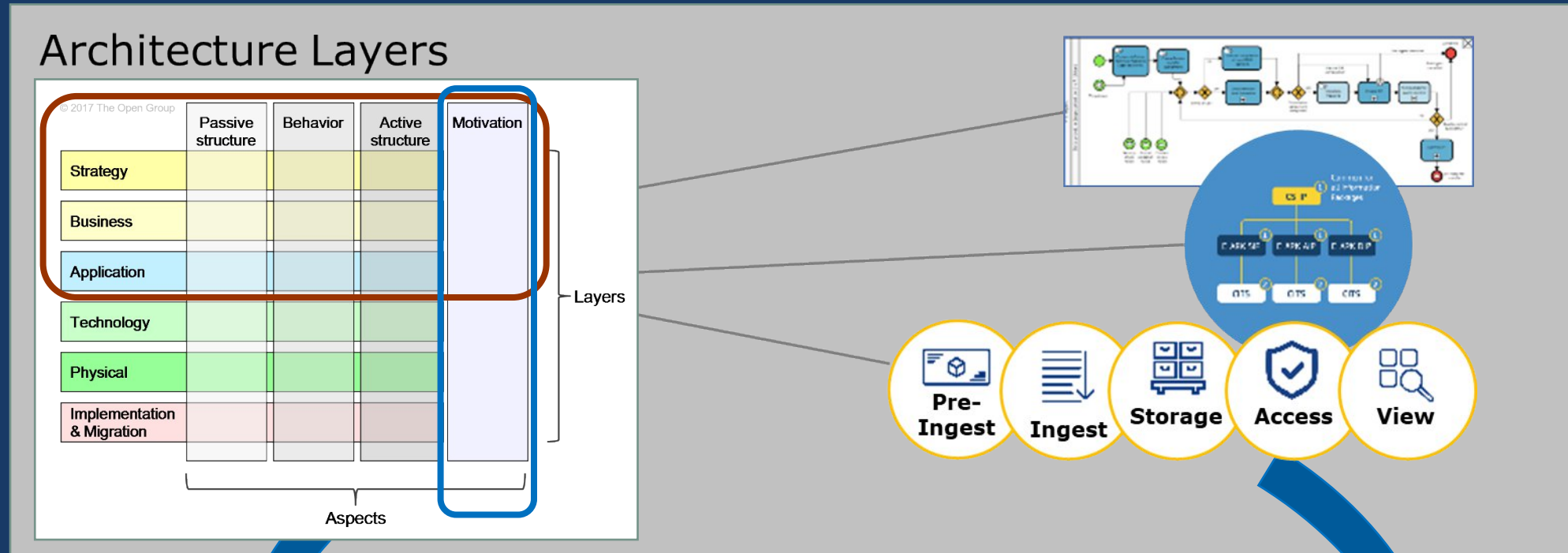
Reuse dashboard



(Numbers until 31st October 2021. Official dashboard 2021 Q1 updated by the Consortium)



Reference Architecture <https://kc.dlmforum.eu/earchiving-ra20/>





The eArchiving Initiative



The eArchiving Initiative in a nutshell

Service Contract: eArchiving Common Services Platform - CNECT/LUX/2021/OP/0077 NUMBER – LC-01905904

- **1 Research Institution: Austrian Institute of Technology (AIT)**
- **3 SME: Keep Solutions, Highbury, Gabinete**
- **1 Membership organization: DLM Forum**

AIT Subcontractors

AAFB, Easy Lean, Ever Team, TrustScape, KMD, Prisma.

DLM Project Members:

National Archives of Estonia, Finland, Norway, Slovenia, Swiss Federal Archives, OPF, ES Solutions; Geoarh, Poliphon, XAMK, INESC-ID, SERDA, Sydarkivera, Penwern, Digital Lifecycle Management.



The eArchiving Initiative in a nutshell

- Policy officers: **Fulgencio Sanmartín, Árpád Welker, DG CNECT**
- Head of Unit DG CNECT: **Rehana Schwinninger-Ladak**
- Expert Group: **European Archives Group (EAG)**



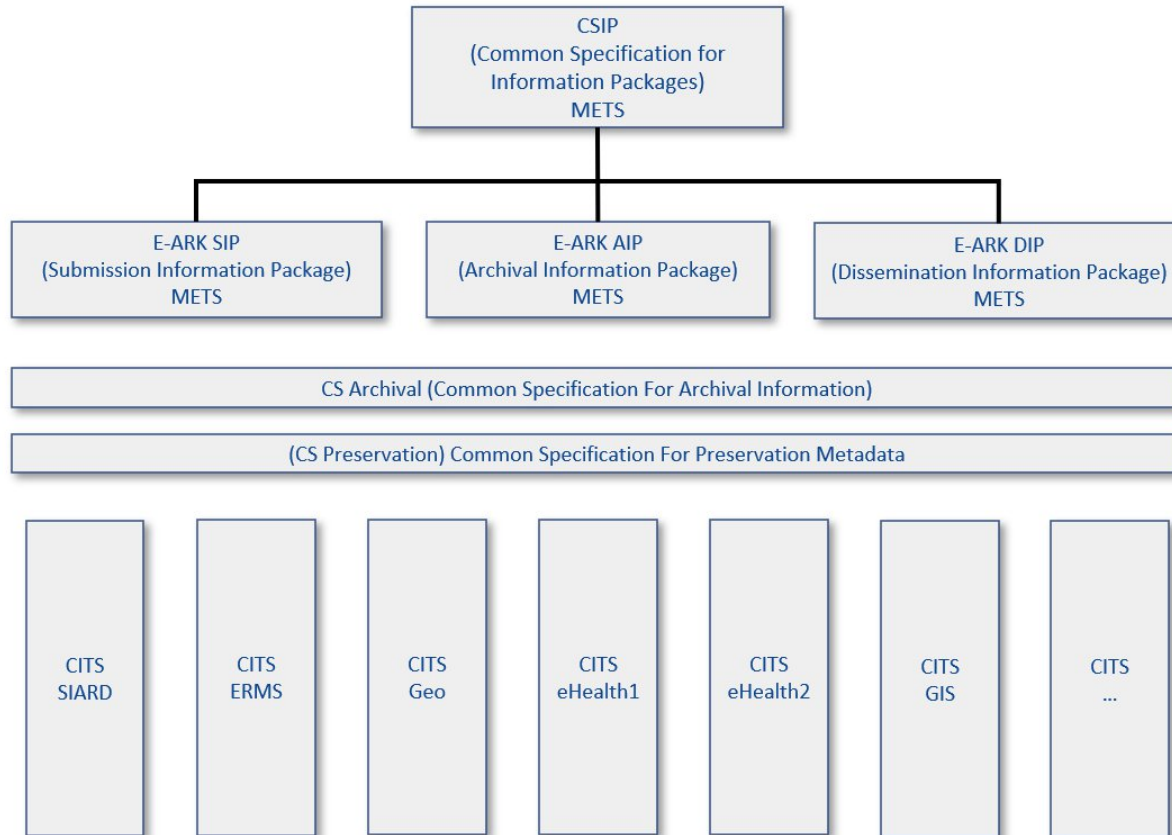
eArchiving Vision

- Wide adoption of the E-ARK specifications across a broad range of sectors, domains and countries, evidenced by real impact and engagement, culminating in being the **official de facto standard**.
- **Sustainability** in all its forms is a key part of our mission. Developing an **eArchiving curriculum** and **Conformance Seal** is part of this drive.
- We are focussing on **Capacity Building** and **Support** to deepen our relationships with our **user communities**.



Highlights

Specifications – new version!



Content-Specific Validators

E-ARK Specification Test Coverage

Test coverage for the E-ARK Information Package Specifications.

E-ARK CSIP METS Profile V2.0.4 2020-06-12T09:00:00

- 133 requirements.
- 101 test cases.
- 147 validation rules.
- 85/353 packages.

WIP

E-ARK SIP METS Profile 2.0 SIPV2.0.4 2020-06-12T09:00:00

- 35 requirements.
- 8 test cases.
- 18 validation rules.
- 0/40 packages.

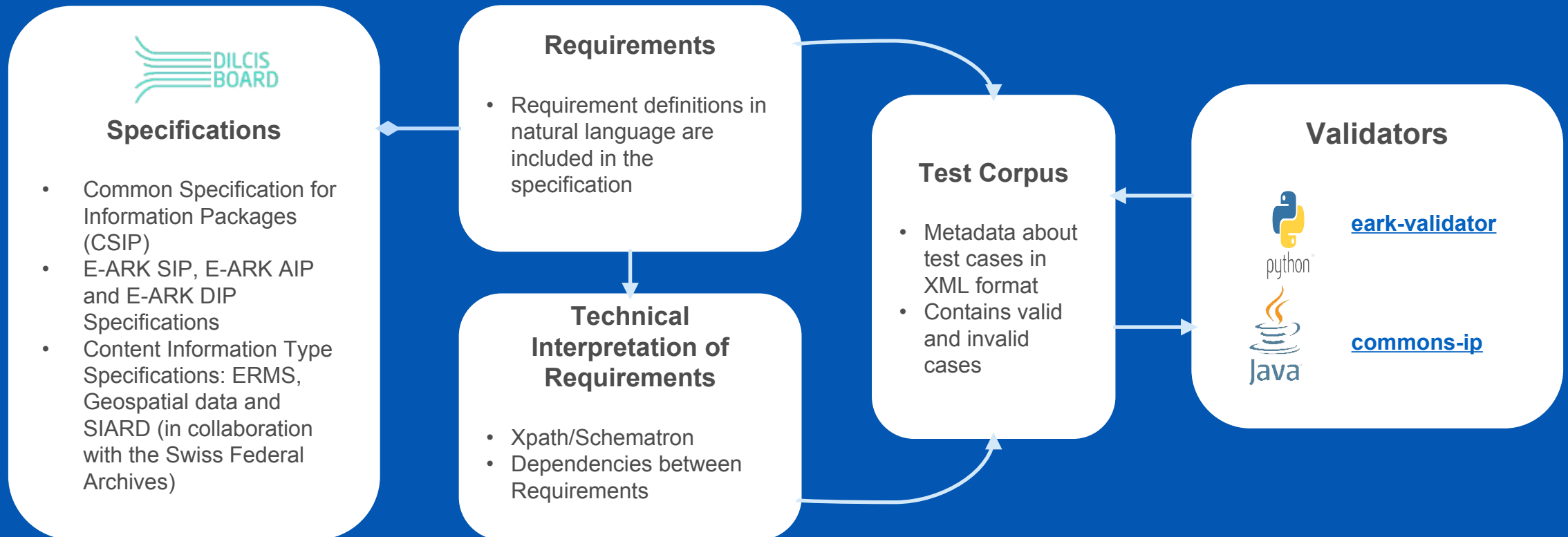
WIP

E-ARK DIP METS Profile DIPV2.0.4 2020-06-12T09:00:00

- 4 requirements.
- 0 test cases.
- 0 validation rules.
- 0/0 packages.

WIP

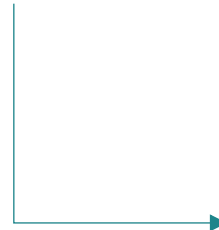
Validation Framework



Highlights

eArchiving Reference Architecture

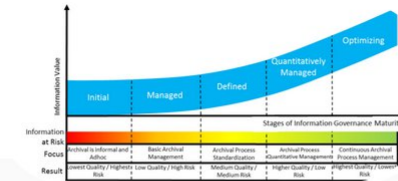
Archiving by design



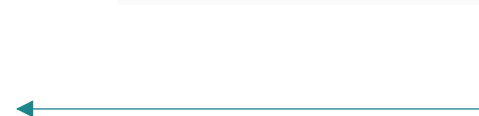
eArchiving Reference Architecture



CEN/TC 468 - MANAGEMENT AND PRESERVATION OF DIGITAL CONTENT



eArchiving Capability and Maturity Model



Highlights

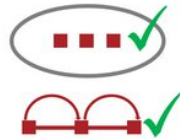
eArchiving Conformance Seal



1. Validation

Validation is the process of **testing a single Information Package (IP)** file for conformance. You can start by verifying if individual information packages of your digital repository or archival solutions are valid by using the **Validation Service** to test if your information packages are conformant with E-ARK specifications.

Validation Service



2. Self-assessment conformance

eArchiving conformance of a digital archive or archiving solution includes **to meet the E-ARK requirements for ingest and processing of Information Packages**, in addition to the validation of a representative sample of information packages.

Self-Assessment



3. Conformance application review

A brief interview with the eArchiving Team is arranged to review the evidence provided and verify the implementation of **E-ARK requirements for ingest and processing of Information Packages**.

Apply for a Review



4. Seal

After successful review the **eArchiving Conformance Seal** is awarded by the European Commission.

Highlights

eArchiving EC Website

<https://digital-strategy.ec.europa.eu/en/activities/earchiving>

eArchiving Initiative

The eArchiving Initiative provides core specifications, software, training and knowledge to help people store information for longer.



What is eArchiving?

Important information should be kept accessible and reusable for years to come, regardless of the system used to store it. **eArchiving provides core specifications, software, training and knowledge** to help people preserve and reuse information over the long-term.



Interoperability by default

Following international standards and specifications for packaging digital information enables organisations to transmit documents and information across borders. It also supports the storage, access and reuse of information regardless of the platform used.



Openness and transparency

Having a common set of open specifications for packaging and archiving digital information, promotes a high level of transparency and confidence among all participants in the information lifecycle.



Sustainability and legal compliance

With eArchiving, digital archival systems can implement reusable modular components that are compliant with various national legal requirements.

Featured links

[Technical specifications >](#)

[Conformance Seal >](#)

[Knowledge Centre >](#)

[Training >](#)

[Support desk >](#)

[Get involved >](#)

Follow the latest progress and learn more about getting involved.

[E-ARK on Twitter](#)

To know more about the Initiative and other advantages before start using it, check the other [benefits](#) eArchiving has to offer.

Highlights

The Academic Curriculum

Working with the eArchiving community

- ~ Based on iterative engagement with the community.
- ~ Initial input considerable (160 responses to our questionnaire).
- ~ Positive feedback at Salamanca.
- ~ Interest from academic institutions.



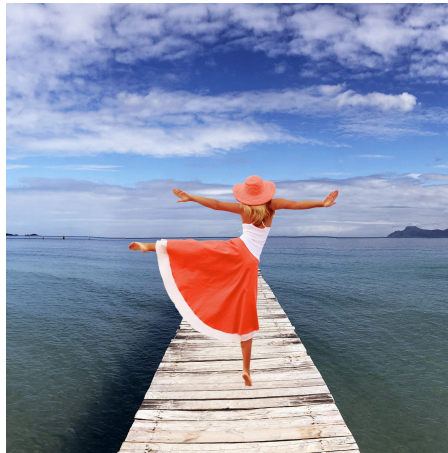
The eArchiving Academic Curriculum

Questions for the panellists

- **Jože Q1.** You, Kuldar Aas from the National Archives of Estonia and Jan Dalsten from the Danish National Archives were the drivers behind E-ARK. What were your motivations for wanting E-ARK? What needs were you trying to address?
- **Alina Q1.** You had not been involved with digital archiving before you encountered E-ARK. What were your first impressions?
- **Adrian Q1.** You have written widely on Digital Preservation. When you were appointed as one of the E-ARK project reviewers in 2014, how important was the topic of digital archiving in your opinion?
- **John Q1.** We first met you when the E-ARK outputs were presented at the DLM meeting in Oslo in 2016. How important do you think were those first E-ARK outputs?
- **Tom Q1.** You had not been involved with digital archiving before you encountered E-ARK. What were your first impressions?
- **Anders Bo Q1.** The DNA has been a key player in E-ARK from the beginning, especially as the Coordinator for the eArchiving Building Block phase. What did you hope to gain for the DNA and for the wider community when you first joined E-ARK?

Questions for the panellists

- **All Q2.** From your standpoint, what has been the biggest achievement/milestone/challenge for E-ARK/eArchiving?
- **All Q3.** What is your hope for eArchiving going forward?



Thank you



© European Union 2020

Unless otherwise noted the reuse of this presentation is authorised under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license. For any use or reproduction of elements that are not owned by the EU, permission may need to be sought directly from the respective right holders.

Slide xx: [element concerned](#), source: [e.g. Fotolia.com](#); Slide xx: [element concerned](#), source: [e.g. iStock.com](#)

