D_PM.11 Annual review and report 2023: Executive Summary

The eArchiving Initiative

Date: 23/02/2024
Doc. Version: 1.0
Template version: 3.0.1

Document Control Information

Settings	Value
Document Title:	D_PM.11 Annual review and report 2023: Executive Summary
Project Title:	The eArchiving Initiative
Document Authors:	Janet Anderson with input from David Anderson, Sven Schlarb, Jaime Kaminski, Carlota Bustelo, Miguel Umlauff, Jože Škofljanec, Karin Bredenberg, Stephen Mackey, Begoña Sanchez-Royo, Luís Faria, István Alföldi.
Project Owner (PO):	Árpád Welker, Fulgencio Sanmartín, DG CNECT
Doc. Version:	1.0
Sensitivity:	Public
Date:	23/02/2024

Document Approver(s) and Reviewer(s):

NOTE: All Approvers are required. Records of each approver must be maintained. All Reviewers in the list are considered required unless explicitly listed as Optional.

Name	Role	Action	Date
Ross King	E-ARK CCO	Review/Approve	
Árpád Welker/Fulgencio Sanmartín	Project Owner	*<Approve / Review>*	

Document history:

The Document Author is authorised to make the following types of changes to the document without requiring that the document be re-approved:

- Editorial, formatting, and spelling
- Clarification

This document is a living document and will be updated as appropriate throughout the life of the contract.

Contact the Document Author or Owner to request a change to this document. Changes to this document are summarised in the following table in reverse chronological order (latest version first).

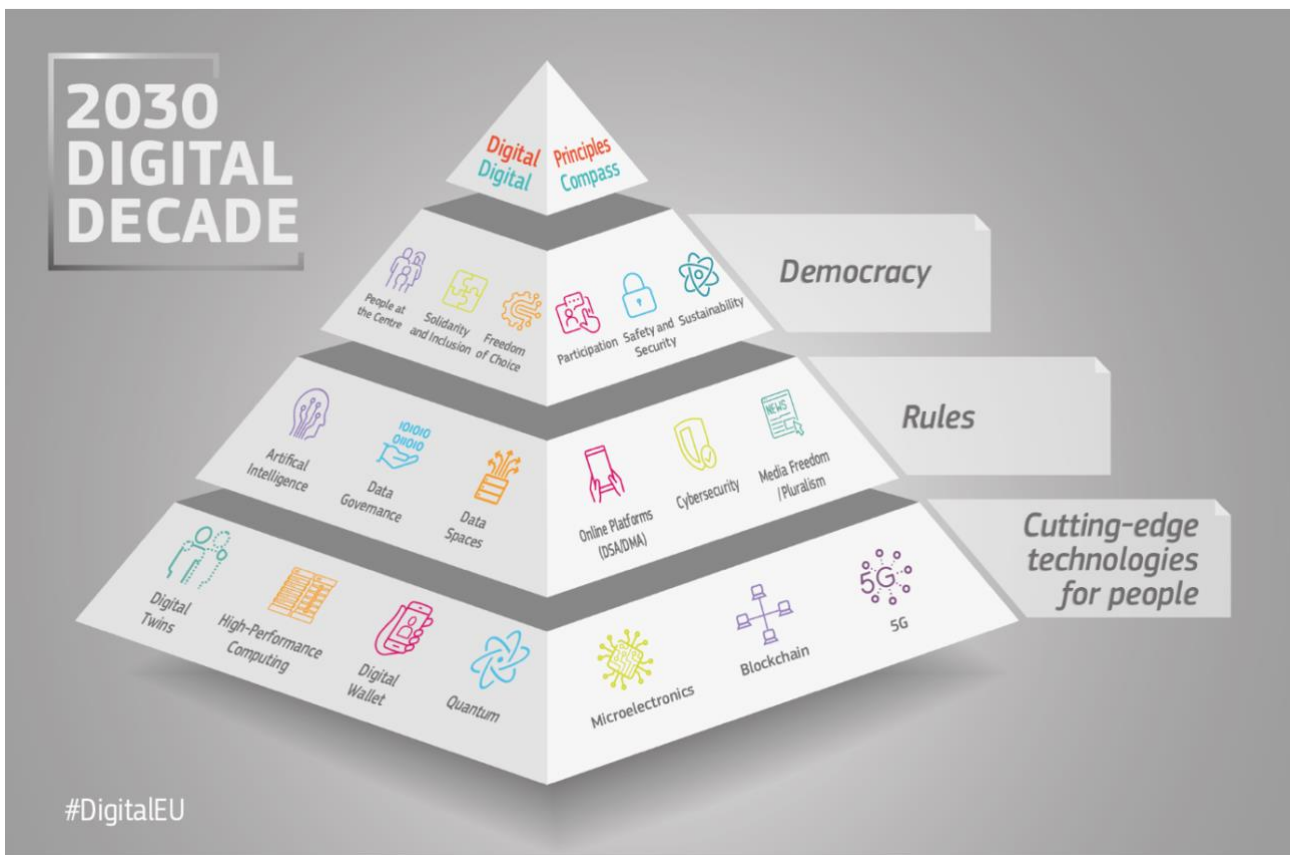| Revision	Date	Created by	Short Description of Changes
1.0	23/10/2023	Janet Anderson	Corrections following review by David Anderson, Ross King, and Sven Schlarb

Statement of originality

EXECUTIVE SUMMARY OF THE EARCHIVING INITIATIVE ANNUAL REPORT 2023

The eArchiving Initiative[1] is part of the European Union's (EU's) Digital Decade Policy Programme[2] (DDPP) whose mission is to provide a "human-centric, sustainable vision for digital society throughout the digital decade to empower citizens and businesses". A key element of this vision is ensuring that data can be sustainably kept and accessed over a variety of timeframes, especially considering the expansion of data in, for example, 3D visualisations such as those found in 3D Digital Cultural Heritage[3]. This leads to a challenge for the DDPP's green digital transition: to not only store increasingly huge amounts of current data, but to also change legacy data from analogue to electronic whilst depending only on systems and solutions created by and for the EU, such as eArchiving. Building these essential digital capabilities gets even more complex with technological backdrop of High Performance Computing; Artificial Intelligence; Cybersecurity; Advanced digital skills; and European digital innovation hubs. And at the other end of the spectrum, any personal, confidential or sensitive data must be securely and legally held. Because all this important information should be kept accessible and reusable for years to come, regardless of the system used to store it, eArchiving provides core specifications, standards, validation, software, training and knowledge to help people preserve and reuse information over the short, medium and long-term.

https://digital-strategy.ec.europa.eu/en/policies/europes-digital-decade

---

[1] https://digital-strategy.ec.europa.eu/en/activities/earchiving

[2] https://digital-strategy.ec.europa.eu/en/policies/europes-digital-decade

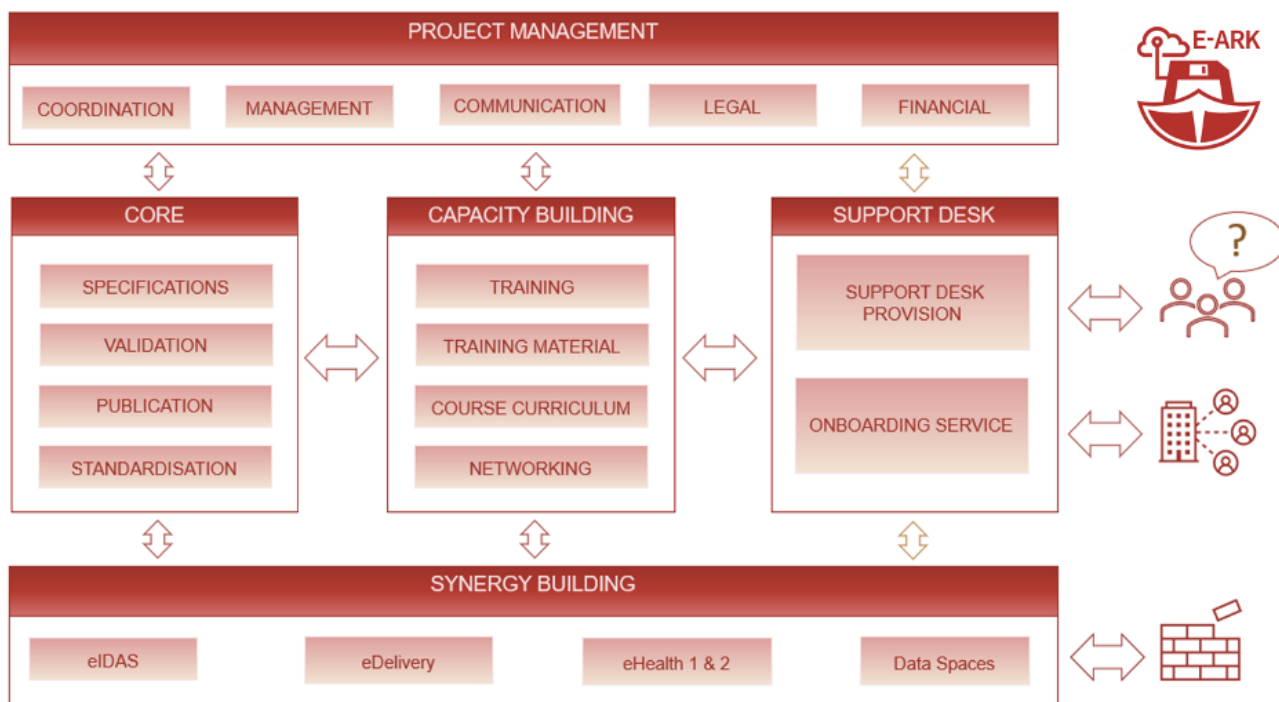[3] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32021H1970, accessed 21st February 2024.

The eArchiving annual report 2023 marks the conclusion of Phase I of the eArchiving Procurement, which commenced in October 2022. This phase builds upon the foundation laid by the eArchiving Building Block[4] and its predecessor, the original E-ARK Project[5]. The eArchiving Initiative is under the stewardship of the E-ARK Consortium, led by the Group Leader AIT Austrian Institute of Technology GmbH[6], with key contributions from the Group Members Highbury R&D Ltd (Ireland), DLM Forum MTÜ (Estonia), Gabinete Umbus SL (Spain), and KEEP Solutions LDA (Portugal). Additionally, fifteen organisations (entities upon whom we rely) actively participate in this initiative through the DLM Forum, and AIT also collaborates with six subcontractors.

The primary objective of this initiative is to promote sustainable eArchiving across Europe and provide access to E-ARK's open-source standards, specifications, tools, validation procedures, training resources, and outreach services. This mission is accomplished through five core activities, as illustrated below.

The eArchiving Initiative is overseen by DG CNECT, ensuring transparent and effective management. Project progress, risks, and financial aspects have been consistently and rigorously monitored through monthly meetings and bi-monthly reports, bolstered by frequent communication via email and online channels. Remarkably, all Key Performance Indicators (KPIs) and deliverables have been not only met but also submitted on schedule, and the procurement process has remained well within its allocated budget.

The journey began with a kick-off meeting on October 14th, 2022, and throughout the initiative's course, an array of internal meetings have been convened within the consortium. These internal meetings include gatherings specific to each activity or task, cross-activity discussions, and fortnightly sessions involving the Activity Leads. Additionally, there have been pivotal gatherings such as an All-Staff hybrid meeting and a Steering Committee meeting, which took place alongside the DLM Forum meetings in Ljubljana in May 2023[7].

---

[4] The Connecting Europe Facility (CEF) eArchiving Building Block was operated via two grants: E-ARK4ALL, Grant Agreement number: LC-00921441 CEF-TC-2018-15 eArchiving (2018-2019), and E-ARK3, Grant Agreement No. LC-01390244 CEF-TC-2019-3 E-ARK3, (2019-2021).

[5] The E-ARK project (2014-2017) was funded by the European Commission under the ICT Policy Support Programme (PSP) within Call 7 of the Competitiveness and Innovation Framework Programme, Grant number 620998.

[6] https://www.e-ark-foundation.eu/consortium-members/

[7] https://www.dlmforum.eu/index.php/all-events/141-dlm-mm-ljubljana-may-2023

The initial six months of the project were dedicated to establishing crucial technical infrastructure in collaboration with DG CNECT. This infrastructure encompassed project and risk management, as well as event organization and communication. By the end of December 2023, the initiative had successfully implemented a meticulously detailed workplan, a robust risk management strategy, and a comprehensive communication plan. These foundational elements were devised with a keen focus on engaging key stakeholders, including archives, data producers, solution providers, and educators.

Furthermore, to enhance communication and accessibility, the initiative launched the DG CNECT eArchiving website at https://digital-strategy.ec.europa.eu/en/activities/earchiving, primarily hosting static content. Additionally, dynamic web pages, including form submissions for the Conformance Seal (see below), were established on the E-ARK Foundation website, accessible at both https://seal.e-ark-foundation.eu and https://www.e-ark-foundation.eu.

Specifications serve as the cornerstone of the eArchiving Initiative, with the E-ARK Specifications, overseen and available through the DILCIS Board[8].

Central to the initiative are three primary specifications:

1. E-ARK Submission Information Package (SIP): This specification facilitates the intake of materials into the digital archive.
2. E-ARK Archival Information Package (AIP): Designed for preserving materials over short, medium, or long-term durations, this specification plays a crucial role in safeguarding digital content.
3. E-ARK Dissemination Information Package (DIP): This specification is pivotal for the retrieval and export of materials from the digital archive.

Each of these three specifications is built upon the foundation of the Common Specification for Information Packages (CSIP).

Furthermore, the initiative encompasses Content Type Information Specifications (CITS) that are tailored to specific data content types. The current spectrum of CITS covers the following domains:

- **E-ARK Content Information Type Specification for Electronic Records Management Systems (CITS ERMS):** Focused on archiving electronic records management systems.
- **E-ARK Content Information Type Specification for digital geospatial data records archiving (CITS Geospatial):** Targeting the archiving of digital geospatial data records.
- **E-ARK Content Information Type Specification for Relational Databases using SIARD[9] (CITS SIARD):** Designed for preserving relational databases through the SIARD format.
- **E-ARK Content Information Type Specification for Archival Information (CITS Archival):** Addressing the specific needs of archiving and preserving essential archival information.
- **E-ARK Content Information Type Specification for Patient Medical Records (CITS eHealth1):** Tailored for the archiving of patient medical records.
- **E-ARK Content Information Type Specification for Cancer Registry export (CITS eHealth2):** Focused on the export and archiving of cancer registry data.
- **E-ARK Content Information Type Specification for Preservation metadata (CITS PREMIS):** Specifically designed to manage and preserve critical preservation metadata.

In this inaugural year of the eArchiving Initiative, the listed specifications, all of which were initially published in August and October 2021, have been diligently adhering to the prescribed revision cycle outlined in the "Guidelines for revision of Common Specifications and Content Information Type Specifications," accessible at https://dilcis.eu/images/procedures/LC-001390244_CEF-TC_2019-3_M23d.pdf.

---

[8] The DILCIS Board: https://dilcis.eu/
[9] https://www.ech.ch/de/ech/ech-0165/1.0

The stipulated revision cycle spans five years, as delineated in the guideline. This means that, at the end of each five-year cycle, a comprehensive evaluation is conducted to determine the necessity of a major revision. Until the onset of the fifth year, scheduled for this set of specifications in 2026, only minor revisions and bug fixes are permitted. The critical aim here is to maintain a phase of specification upkeep where major revisions are avoided, ensuring the ease of use and thereby promoting wider adoption.

In accordance with the guideline, the specifications have undergone minor updates over the past year. These updates primarily involve refining the language used in requirement descriptions to enhance their clarity and comprehensibility. Simultaneously, any detected spelling errors have been corrected. The process of updating the specifications hinges on the active participation of users, who contribute their insights and observations via the GitHub repositories designated for each specification[10].

Throughout the upcoming year within the eArchiving initiative, we anticipate the continuation of this maintenance work, alongside an effort to improve the publication process for information package specifications. The Core Specifications efforts are bolstered by the support of a Knowledge Centre[11], encompassing the E-ARK Reference Architecture[12], the E-ARK Maturity Model, and a comprehensive Glossary. Furthermore, the alignment of all E-ARK services is meticulously maintained through a quarterly Release Management process.

An essential aspect of this initiative is the ability to demonstrate compliance with the E-ARK Specifications, be it for an archive or a software component. Thus, considerable attention has been dedicated to refining the E-ARK validation process during this initial phase. This has entailed ensuring the robustness and comprehensiveness of the validation test corpus, as well as aligning the outputs of the two complementary validators. Notably, significant effort has been invested in the development of the eArchiving Conformance Seal, designed for use by archives and solution providers. The workflow has been defined, and the process is operational on the initiative's website. Valuable feedback on this endeavour has been sought during the DLM May 2023 meeting and through a dedicated webinar held on April 26th, 2023.

---

[10] e.g. https://github.com/DILCISBoard/E-ARK-SIP
[11] http://kc.dlmforum.eu/home
[12] http://kc.dlmforum.eu/earchiving-ra10/

The inaugural User Capacity Building event took place online on April 25th and 26th, 2023, with the primary aim of re-establishing connections with our audience and outlining future plans. Out of the 107 attendees, the event garnered a notable satisfaction rate of 8.73/10.

Subsequently, the first transnational Specialist Network event was held on September 22nd, 2023, under the title "How the eArchiving Initiative can support the implementation of eIDAS electronic archiving trust services." This gathering brought together Archival Solutions Providers and various specialists to collectively explore the potential market impact of the eIDAS Regulation, which is currently under discussion in the European Parliament. The event drew participation from 33 companies and organizations representing 22 European countries.

Furthermore, consortium members delivered several informative presentations on eArchiving at the DLM Meeting in May 2023. These presentations significantly contributed to both user capacity building and networking. They covered topics such as eArchiving geospatial data specifications and their implementation at the Geoforum meeting, eArchiving Conformance, specifications for 3D models (including Building Information Models), and models for Digital Cultural Heritage. Additionally, a general presentation provided an overview of the progress made within the eArchiving Initiative.

In the latter half of this first year, the Capacity Training initiatives gained momentum, and they successfully conducted six webinars encompassing a wide spectrum of eArchiving topics. Effective communication and collaboration between the Initiative and the four Connecting Europe Facility (CEF) Generic Services (GS) projects, which received funding in 2021 (J-ARK, PROHUB, OneClick, and the eArchiving of Engineering and Science Library), have been pivotal. Notably, members of the E-ARK Consortium participated as partners in all four GS projects, with designated representatives facilitating seamless knowledge transfer and support.

Most of the GS projects neared their conclusion in January-February 2023. Consequently, by the time the eArchiving Initiative commenced operations in October 2022, these GS projects were already in the winding-down phase. This emphasised the need for collaborative efforts aimed at comprehending how eArchiving could bolster their activities, encompassing marketing, training, legacy preservation, and technical support.

A key milestone for the year revolved around the development of an eArchiving course curriculum tailored specifically for the higher education community. This approach not only caters to the existing eArchiving user base within the community but also extends its reach to the higher education sector. The creation of a Masters curriculum has been progressing systematically, driven by the Academic Curriculum Working Group, which convenes biweekly to shape this educational initiative. The initial curriculum draft adheres to a modular structure designed to accommodate the diverse academic systems prevalent across Europe. This modular framework offers institutions the flexibility to either adopt the entire curriculum or cherry-pick specific modules that align with their unique needs. The proposal recommends the inclusion of a core set of six mandatory modules emphasizing practical application and providing students with a solid foundational understanding of eArchiving principles. Complementary elective modules will further enhance this core curriculum. The upcoming phase is dedicated to the development of materials for "train the trainer" sessions, scheduled to commence at the beginning of Phase II.

The eArchiving Support Desk became operational in February 2023, and during this initial phase, it received an average of five requests per month. These requests covered a broad spectrum of topics, ranging from the validation of the Building Blocks Generic Services Call projects to issues related to eArchiving events. The Support Desk team anticipates a surge in requests with the launch of the eArchiving Conformance Seal at the upcoming eArchiving Summit.

The primary stakeholder groups within eArchiving encompass Archives, Data Producers, Solution Providers, and Educators. In March 2023, the Initiative inherited a comprehensive list of unique contacts from the previous eArchiving Building Block for its onboarding service. This list facilitates the identification of contacts, their associated organizations, countries of origin, and consent, all in accordance with IDPR guidelines. These contacts represent 626 organizations spanning 66 countries and dependencies, including EU Institutions and

Agencies, other European organisations, and International Organisations. Presently, there are 61 distinct projects committed to reusing or analysing eArchiving components.

eArchiving maintains a robust presence on social media platforms, including a LinkedIn Group boasting 752 members, a LinkedIn Company page with 198 followers, and an active Twitter account.

Throughout its first year of operation, eArchiving has diligently fostered collaborative synergies with relevant Digital Europe projects, particularly those formerly under the CEF Building Blocks umbrella, and these include the Data Spaces Support Centre (DSSC) and Data Spaces, encompassing areas like Tourism and Skills. The Synergy Building team has established strong partnerships with these stakeholders and engaged in efforts to identify opportunities for the long-term preservation of information, with a focus on enhancing accessibility, ensuring data integrity and authenticity, and respecting individual rights concerning access and personal data protection.

Of pivotal significance is the eIDAS Regulation, a landmark development in the realm of electronic identification and trust services within the EU. The eArchiving team has closely monitored its evolution and played a key role in facilitating the integration of eIDAS instruments to preserve digital data and guarantee its accessibility.

Additionally, the eDelivery and eHealth domains are integral components of the Digital Europe programme, aimed at streamlining document exchange and healthcare information sharing among EU member states. The eArchiving Initiative possesses valuable expertise in implementing solutions and crafting specifications within these domains, with a heightened focus on enhancing activities in these areas. This effort is critical for ensuring the secure, reliable, and efficient preservation and access to digital data.

Looking ahead, the key priorities in the upcoming phase will revolve around the continued development of the eArchiving Curriculum and the implementation of the Conformance Seal.

This first year has been exceptionally productive, with all five Activities putting forth substantial efforts, staying well within the allocated budget, and consistently meeting key performance indicators (KPIs) and deadlines. As we transition into phase II, which encompasses the second procurement year from October 1st, 2023, to September 30th, 2024, our primary focus will be centred around two major initiatives. Firstly, we are committed to finalizing the eArchiving Curriculum, ensuring it is a comprehensive and robust resource. Secondly, we are gearing up for the official launch of the eagerly awaited Conformance Seal.

It is worth noting that the ongoing tasks and Activities will continue working towards their respective objectives, as detailed in our previous bi-monthly updates. Moreover, we are excited to announce our plans for a significant Specifications event scheduled to take place in Brussels in May 2024. This event will serve as a platform to gauge the extent of user adoption, gather valuable feedback, and celebrate the continued growth and impact of the eArchiving Initiative.