Specification for E-ARK Content Information Type Specification for Archival Information (CITS Archival Information)

A proper front page will be created for the publication occurring after implementation of review comments.

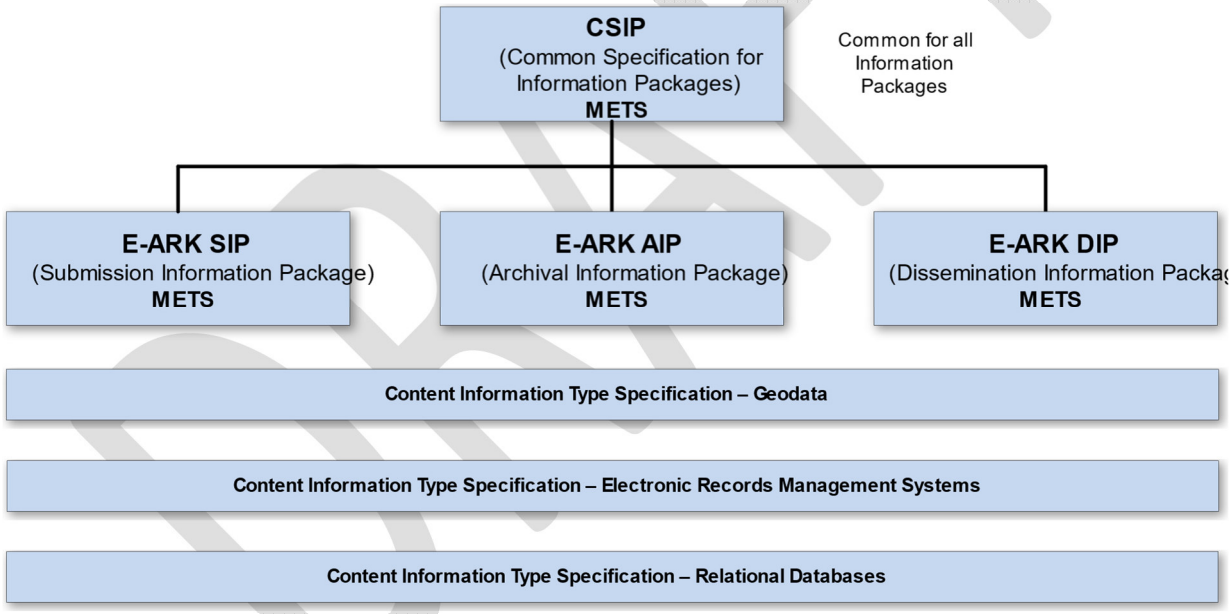
1 Preface

The correct preface will be inserted for the publication occurring after implementation of review comments.

1.1 Aim of the specification

This E-ARK specification is part of a family of specifications that provide a common set of requirements for packaging digital information. These specifications are based on common, international standards for transmitting, describing and preserving digital data. They have been produced to help data creators, software developers and digital archives tackle the challenge of short-, medium- and long-term data management and reuse in a sustainable, authentic, cost-efficient, manageable and interoperable way.

The foundation for these specifications is the Reference Model for an Open Archival Information System (OAIS) which has Information Packages at its core. Familiarity with the core functional entities of OAIS is a prerequisite for understanding the specifications. A visualisation of the current specification network can be seen here:



The E-ARK specification dependency hierarchy

Specification	Aim and Goals
Common Specification for Information Packages	This document introduces the concept of a Common Specification for Information Packages (CSIP). Its three main purposes are to: • Establish a common understanding of the requirements which need to be met in order to achieve interoperability of Information Packages. • Establish a common base for the development of more specific Information Package definitions and tools within the digital preservation community.

Specification	Aim and Goals
	• Propose the details of an XML-based implementation of the requirements using, to the largest possible extent, standards which are widely used in international digital preservation. Ultimately the goal of the Common Specification is to reach a level of interoperability between all Information Packages so that tools implementing the Common Specification can be adopted by institutions without the need for further modifications or adaptations.
E-ARK SIP	The main aims of this specification are to: • Define a general structure for a Submission Information Package format suitable for a wide variety of archival scenarios, e.g. document and image collections, databases or geographical data. • Enhance interoperability between Producers and Archives. • Recommend best practices regarding metadata, content and structure of Submission Information Packages.
E-ARK AIP	The main aims of this specification are to: • Define a generic structure of the AIP format suitable for a wide variety of data types, such as document and image collections, archival records, databases or geographical data. • Recommend a set of metadata related to the structural and the preservation aspects of the AIP as implemented by the reference implementation eArchiving ToolBox (formerly earkweb). • Ensure the format is suitable to store large quantities of data.
E-ARK DIP	The main aims of this specification are to: • Define a generic structure of the DIP format suitable for a wide variety of archival records, such as document and image collections, databases or geographical data. • Recommend a set of metadata related to the structural and access aspects of the DIP.
Content Information Type Specifications	The main aim and goal of a Content Information Type Specification is to: • Define, in technical terms, how data and metadata must be formatted and placed within a CSIP Information Package in order to achieve interoperability in exchanging specific Content Information. The number of possible Content Information Type Specifications is unlimited.

1.2 Organisational support

This specification is maintained by the Digital Information LifeCycle Interoperability Standards Board (DILCIS Board). The DILCIS Board (http://dilcis.eu/) was created to enhance and maintain the draft specifications developed in the European Archival Records and Knowledge Preservation Project (E-ARK project) which concluded in January 2017 (http://eark-project.com/). The Board consists of eight members, but there is no restriction on the number of participants in the work. All Board documents and specifications are stored in GitHub (https://github.com/DILCISBoard) while published versions are made available on the Board webpage. Since 2018 the DILCIS Board has been responsible for the core

specifications in the Connecting Europe Facility eArchiving Building Block (https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eArchiving).

1.3 Authors

A full list of contributors to this specification, as well as the revision history can be found in Appendix 1.

TABLE OF CONTENT

LIST OF TABLES

TABLE OF FIGURES

1 Context

1.1 Purpose

The purpose of this document is to describe the Content Information Type Specification (CITS) for Archival Information. The term 'archival information' includes documents such as the finding aid and the creator information.

The specification is used for both the transfer to archives and for information exchange between different systems that require access to the Archival Information about the transfer. This specification is supported by several different XML-schemas, Schematron documents and by ontologies provided by the different endorsed profiles. Schematron is an XML-document accompanying the different XML-schemas which includes rules that the XML-schema cannot validate.

1.2 Scope

The scope of this specification is to add an Archival Information to the transfer of an information package.

2 Archival Information

Archival Information is divided into different categories:

Archival Information or Finding Aid: This document describes the structure of the archival material itself. The document follows a structure developed by the International Council of Archives (ICA) called the General International Standard Archival Description (ISAD-G) (https://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition). ISAD-G does not provide a transfer format but uses the Encoded Archival Description (EAD) (http://www.loc.gov/ead/index.html) maintained by the Technical Subcommittee on Encoded Archival Standards (https://www2.archivists.org/governance/handbook/section7/groups/Standards/TS-EAS).

Archival Creator: This document provides information about the person or body that created the archival collection. The document follows a structure developed by the International Council of Archives (ICA) called the International Standard Archival Authority Record for Corporate Bodies, Persons and Families (ISAAR(CPF)) (https://www.ica.org/en/isaar-cpf-international-standard-archival-authority-record-corporate-bodies-persons-and-families-2nd). ISAAR(CPF) does not provide a transfer format, but uses the Encoded Archival Context for Corporate Bodies, Persons and Families (EAC-CPF) https://eac.staatsbibliothek-berlin.de/ maintained by the Technical Subcommittee on Encoded Archival Standards (https://www2.archivists.org/governance/handbook/section7/groups/Standards/TS-EAS).

Archival Institution: This document provides information about the institution or body that maintains the archival holdings. The document follows a structure developed by the International Council of Archives (ICA) called the International Standard for Describing Institutions with Archival Holdings (ISDIAH) (https://www.ica.org/en/isdiah-international-standard-describing-institutions-archival-holdings). ISDIAH does not provide a transfer format but uses the Encoded Archival Guide (EAG) (http://wiki.archivesportaleurope.net/index.php/EAG2012) maintained by the Archives Portal Europe Foundation (http://www.archivesportaleuropefoundation.eu/index.php).

Functions or Activities of the Archival Creator: This document provides information regarding the functions or activities that the person or body that created the archival collection has performed. The document follows the structure developed by the International Council of Archives (ICA) called the International Standard for Describing Functions (ISDF) (https://www.ica.org/en/isdf-international-standard-describing-functions). ISDF does not provide a transfer format, although work on a format is being undertaken by the Technical Subcommittee on Encoded Archival Standards (https://www2.archivists.org/governance/handbook/section7/groups/Standards/TS-EAS).

Records in Context: The International Council of Archives (ICA) Expert Group on Archival Description (EGAD) (https://www.ica.org/en/about-egad) is creating a conceptual model for archival description called Records in Context (RiC). This consists of a model described in a textual form (https://www.ica.org/en/egad-ric-conceptual-model) and an ontology (https://www.ica.org/en/records-in-contexts-ontology). The first version is planned to be published in November 2020. It will be possible to include documents following RiC in an information package.

3 Metadata and Mapping

The document or documents forming the archival description are created during export from an archival description system. There is much software available which can automatically export in the different transfer formats for archival description without requiring mapping. The commonly implemented profiles for archival descriptions in European systems follow the Archives Portal Europe profiles. These should be used if national profiles are not utilised.

Note that the profiles used for the Archival Information in the transmission need to be recorded in the submission agreement.

3.1 Archives Portal Europe

The Archives Portal Europe is a portal for gathering European archival descriptions. All the profiles can be found here, http://wiki.archivesportaleurope.net/index.php/Category:Standards. Exports supporting these profiles have been implemented in the common European Systems meaning that an export can be used to transfer the Archival Information in the Information Package.

Note that the METS profile used by Archives Portal Europe is for describing a collection of digital objects in the Archival Information, not as an information package like CSIP.

3.2 Local archival description profiles

Many national profiles of the Archival Information use different transfer formats for transferring information. For example, Sweden has national profiles for Archival Information that use both the current and previous versions of EAD and EAC-CPF (see FGS Arkivredovisning at this URL https://riksarkivet.se/faststallda-kommande-fgser). When national profiles are available, and the submission agreement states that the national profiles for Archival Information are used, these can be placed in the information package.

4 Transfer of Archival Information only

In some cases, only archival Information will be transferred. This could occur, for example, during the transfer of a large number of analogue archives with no digital content, as might arise in a national archive. In such instances, the Archival Information documents are treated as metadata objects only in the transfer and will not be preserved in the archival system.

5 Metadata in the Information Package

Considering the two previous use cases of Archival Information transfer:

- Transfer of Archival Information together with data objects.
- Transfer of several different Archival Information seen as metadata objects in a transfer.

Irrespective of the use case they are inserted in the package in the same manner.

5.1 Using the Content Information Type Specification for Archival Information together with the Common Specification for Information Packages (CSIP)

When the Archival Information documents have been created, they can be transmitted in a package following the requirements described in the chapter about Use of the METS descriptive metadata section (element dmdSec) in the CSIP specification (http://earkcsip.dilcis.eu/).

Each Archival Information type file in the transfer has one “Descriptive metadata” section. Requirements pertaining to the file such as mimetype, size and equals are not described in the following tables.

5.1.1 Specific fields to use in CSIP for the Archival Information or finding aid

The Archival Description is encoded with the help of EAD as previously described.

Table 1: Specific fields to use in CSIP for an EAD document

Element name	METS path	Value
Descriptive metadata	mets/dmdSec	n/a One dmdSec for each file
Status of the descriptive metadata	mets/dmdSec/@STATUS	SUPERSEDED or CURRENT
Type of metadata	mets/dmdSec/mdRef/@MDTYPE	EAD

5.1.2 Specific fields to use in CSIP for the archival creator

The archival creator is encoded with EAC-CPF as previously described.

Table 2: Specific fields to use in CSIP for an EAC-CPF document

Element name	METS path	Value
Descriptive metadata	mets/dmdSec	n/a One dmdSec for each file
Status of the descriptive metadata	mets/dmdSec/@STATUS	SUPERSEDED or CURRENT
Type of metadata	mets/dmdSec/mdRef/@MDTYPE	EAC-CPF

5.1.3 Specific fields to use in CSIP for the archival institution

The archival institution is encoded with EAG as previously described.

Table 3: Specific fields to use in CSIP for an EAG document

Element name	METS path	Value
Descriptive metadata	mets/dmdSec	n/a One dmdSec for each file
Status of the descriptive metadata	mets/dmdSec/@STATUS	SUPERSEDED or CURRENT
Type of metadata	mets/dmdSec/mdRef/@MDTYPE	OTHER
Type of metadata	mets/dmdSec/mdRef/@OTHERMDTYPE	EAG

5.1.4 Specific fields to use in CSIP for the functions or activities of the archival creator

The functions and activities are encoded in a way described in the submission agreement.

Table 4: Specific fields to use in CSIP for a function and activity document

Element name	METS path	Value
Descriptive metadata	mets/dmdSec	n/a One dmdSec for each file
Status of the descriptive metadata	mets/dmdSec/@STATUS	SUPERSEDED or CURRENT
Type of metadata	mets/dmdSec/mdRef/@MDTYPE	OTHER
Type of metadata	mets/dmdSec/mdRef/@OTHERMDTYPE	FUNCTIONS

5.1.5 Specific fields to use in CSIP for a description following Records in Context

Records in Context is encoded with RiC as previously described.

Table 5: Specific fields to use in CSIP for a RiC document

Element name	METS path	Value
Descriptive metadata	mets/dmdSec	n/a One dmdSec for each file
Status of the descriptive metadata	mets/dmdSec/@STATUS	SUPERSEDED or CURRENT
Type of metadata	mets/dmdSec/mdRef/@MDTYPE	OTHER
Type of metadata	mets/dmdSec/mdRef/@OTHERMDTYPE	RIC

5.2 Placement of Archival Information metadata documents in a CSIP Information Package

All the documents that provide Archival Information for the data objects in the transfer are placed in the "metadata" section of the IP, as shown in Figure 1 using EAD as the example.

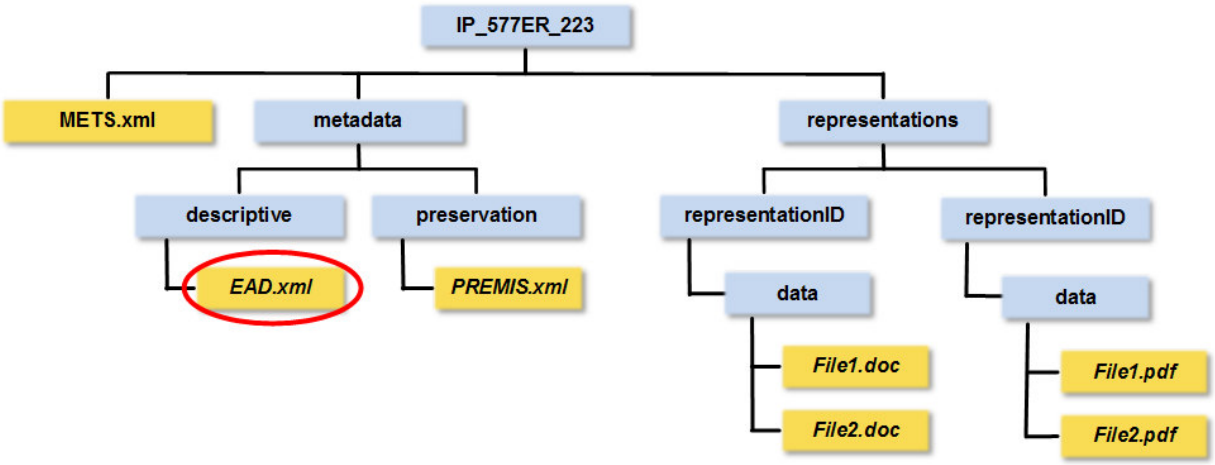


Figure 1: Example of placement of EAD in the IP.

6 Postface

Fill these parts. Keep all for now, there might be parts removed.

AUTHOR(S)	
Name(s)	Organisation(s)
Karin Bredenberg	Kommunalförbundet Sydarkivera
Jaime Kaminski	Highbury Research and Development

REVIEWER(S)	
Name(s)	Organisation(s)
Kerstin Arnold	Archives Portal Europe Foundation
[Name]	[Affiliation]
[Name]	[Affiliation]

Project co-funded by the European Commission within the ICT Policy Support Programme Dissemination Level		
P	**Public**	x
C	**Confidential, only for members of the Consortium and the Commission Services**	

REVISION HISTORY AND STATEMENT OF ORIGINALITY

Submitted Revisions History

Revision No.	Date	Authors(s)	Organisation	Description
[Version]	[Date]	[Who]	[Affiliation]	[What]
0.1	2020-03-16	Karin Bredenberg	SYD	CITS created

Statement of originality:
This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.