

1 Preface

1.1 Aim of the specification

This specification is one of several related specifications. The single most important aim of all of these specifications is the provision of a common set of specifications for packaging digital information for archiving purposes. The 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 to 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 upon which the specifications are built is the Reference model for an Open Archival Information System (OAIS) (OAIS Reference model) which has Information Packages as its basis. 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:

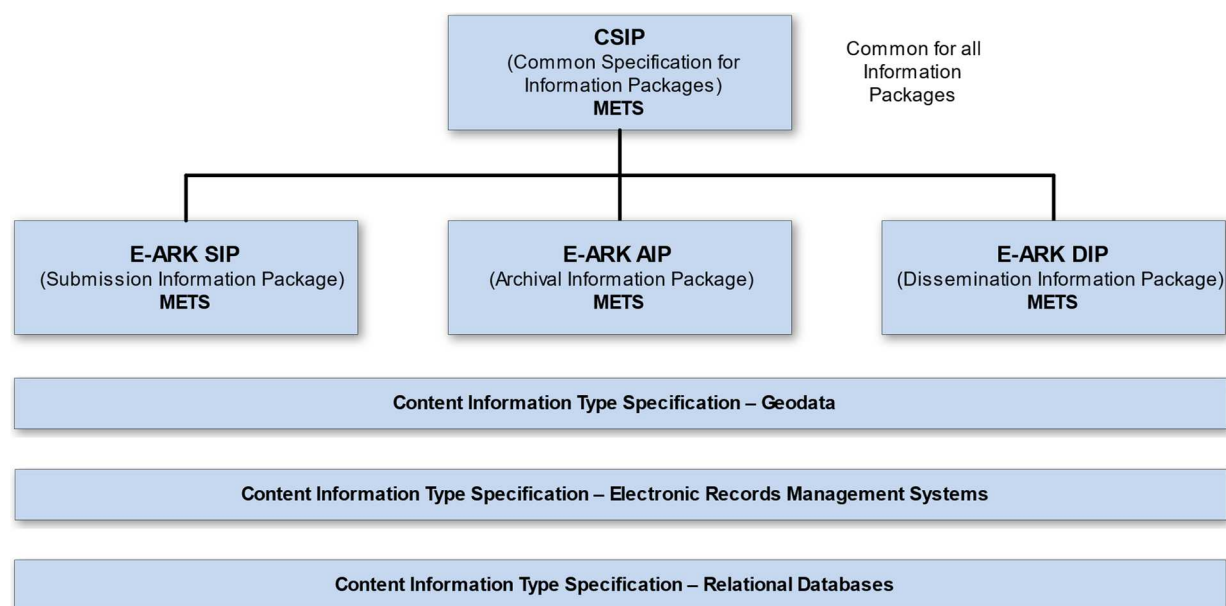


Diagram showing E-ARK specification dependency hierarchy.

Specification	Aim and Goals
Common Specification for Information Packages	<p>This document introduces the concept of a Common Specification for Information Packages (CSIP). Its three main purposes are to:</p> <ul style="list-style-type: none"> Establish a common understanding of the requirements which need to be met in order to achieve interoperability of Information Packages.

Specification	Aim and Goals
	<ul style="list-style-type: none"> • Establish a common base for the development of more specific Information Package definitions and tools within the digital preservation community. • 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. <p>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.</p>
E-ARK SIP	<p>The main aims of this specification are to:</p> <ul style="list-style-type: none"> • 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	<p>The main aims of this specification are to:</p> <ul style="list-style-type: none"> • 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 (eakweb). • Ensure the format is suitable to store large quantities of data.
E-ARK DIP	<p>The main aims of this specification are to:</p> <ul style="list-style-type: none"> • 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	<p>The main aim and goal of a Content Information Type Specification is to:</p> <ul style="list-style-type: none"> • 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. <p>The number of possible Content Information Type Specifications is</p>

Specification	Aim and Goals
	unlimited. For at list of existing Content Information Type Specifications see, and read more about Content Information Type Specifications in the Common Specification for Information Packages

1.2 Organisational support

This specification is maintained by the Digital Information LifeCycle Interoperability Standards Board (DILCIS Board, <http://dilcis.eu/>). The DILCIS Board was created to enhance and maintain the draft specifications developed in the European Archival Records and Knowledge Preservation Project (E-ARK project, <http://eark-project.com/>) which concluded in January 2017. 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

1	Context.....	5
1.1	Purpose and Scope.....	5
1.2	Methodology based upon the E-ARK project.....	6
1.3	Limitations.....	7
2	Layered Data Model.....	7
3	Metadata and Mapping.....	9
3.1	Moreq2010 Entities and Metadata for the ERMS Specification	10
3.2	Translating MoReq2010 Class and Aggregation Values.....	15
3.2.1	Mapping between ERMS and Archive Aggregations	17
3.3	Using the ERMS specification together with the Common Specification for Information Packages (CSIP).....	18
3.3.1	Specific fields to use in CSIP.....	19
3.3.2	Placement of data in a CSIP Information Package.....	19
4	Glossary.....	19
5	Metadata.....	22
5.1	Model picture.....	22
5.2	Metadata for the Control element.....	23
5.3	Metadata for ERMS Records	23
5.4	Metadata for ERMS Aggregation.....	25
5.5	Example of use of own elements	30
5.6	Value other in value lists	31
6	Appendix 1	32

LIST OF TABLES

Table 1: Specific fields to use in CSIP.....	19
Table 2: Glossary	19
Table 3: ERMS use of metadata for Records.....	23
Table 4: ERMS use of metadata for Aggregations.....	26

LIST OF FIGURES

Figure 1: Data Model Structure.....	8
Figure 2: Components	11
Figure 3: Hierarchical classes.....	15
Figure 4: Non-hierarchical classes.....	16
Figure 5: Aggregation	16
Figure 6: Components of the ERMS XML-format	22

1 Context

1.1 Purpose and Scope

The purpose of this document is to describe the content information type specification for ERMS (Electronic Records Management Systems). The specification is designed to be used for the transfer to archives as well as for records exchange between different ERMS systems. This specification is supported by an XML-schema and a Schematron document which includes rules that the XML-schema cannot validate.

There are two options for extracting information from a producer's system:

1. Extracting data in a relational database structure

Extracting data from a relational database into a long-term database preservation format (SIARD) that preserves the properties of the relational database so that the data can be further imported into a relational database management system (RDBMS) at time of access. Access can happen through database queries or via a search field. The main access use cases are:

- a. The producer wishes to retrieve their data for business purposes and/or re-use.
- b. The consumer wishes to consult the data for research purposes.
- c. The archivist wishes to retrieve the data for professional treatment: to check and, if necessary, perform preservation actions, etc.

More information about this option is available in the SIARD 2.0 Profile Specification.¹

2. Extracting data and metadata as records

Extract the records from the system and normalise them to a standard XML format. This means that the records are semantically marked up using metadata. Being technically valid and complying with this specification makes them directly accessible for validation, data management, indexing and searching. The structured semantic metadata description is explicit rather than hidden inside a RDBS. The main advantages over the RDBS representation are that:

- a. Records from different sources can be merged.
- b. Search and access is possible across all records from all sources.
- c. Records can be managed and accessed uniformly.
- d. The original database/records system software does not need to be licensed and preserved.

It is this particular case (i.e. specifying the semantically marked-up metadata profile) that will be discussed and described in the remainder of this ERMS specification.

This specification is expected to be implemented in tools that:

¹ The SIARD 2.0 specification together with a Content Information Type Specification for SIARD represents the SIP profile for the relational databases content type. <https://github.com/DILCISBoard/SIARD>

-
- Extract metadata and data from the native producer systems.
 - Validate that the metadata and data:
 - conform to the specification
 - are complete, and
 - are internally consistent.
 - Receive the metadata and data in another producer system.
 - Create a Submission Information Package (SIP) package from the extracted data and metadata.
 - Transfer the SIP to the archive.
 - Receive the SIP in the archive.
 - Create an Archival Information Package (AIP) from the SIP.
 - Validate that the AIP:
 - conforms to the specification,
 - is complete, and
 - is internally consistent.
 - Ingest the AIP into archival storage.
 - Manage AIPs within the archive.

1.2 Methodology based upon the E-ARK project

The specification is based on the work by the E-ARK project partners during 2014-2017 of specific use cases and requirements to be implemented by export tools for electronic records systems. During the process of developing the specification, decisions were made about:

- Which metadata elements to support?
- Which additional ones were needed?
- How they are to be implemented?

The choice about which entities and metadata elements to adopt in the ERMS specification was made according to two broad criteria. First, only accepted metadata standards that were in frequent use were adopted for the required functions in the ERMS and archive communities. Second, not every entity or metadata element defined in these standards was adopted. Adoption was limited to those that were relevant for the ERMS scenario in order to meet submission use cases and were:

- in use in all archives in the E-ARK project or
- in use in most archives in the E-ARK project or
- required by national regulation and legislation or, to a lesser extent, required by policy decisions within the national archives and related institutions. The former results in mandatory data entities, metadata elements and processes.

From this it was possible to identify which requirements, processes, entities and metadata elements were mandatory for every use of the ERMS specification.

Rather than adopting any particular metadata standard we agreed to adapt existing ones if and as necessary. For example, the mandatory MoReq requirements for metadata elements

were relaxed if they could not be supplied in practice. We also defined extension points so that other metadata elements can be added to support local needs.

Note: the ERMS metadata and data validates correctly with the standard supplied ERMS-schema.

1.3 Limitations

The following are out of scope of this document:

- Proprietary extraction formats, even if they were accompanied by their extraction schema, and functional or records system specification. These sorts of formats have different:
 - use cases resulting in different metadata needs (for original users who want to use the records again in the same form in which they were submitted; for archive management; for future users with new access patterns and content use),
 - amounts of metadata associated with them,
 - degrees of authenticity,
 - dependencies on knowledge about the functionality of the system in which the record can be used.
- The ERMS XML schema is accompanied by a Schematron document with extra rules for conformance. Schematron and how to apply the rules are not described in this document.

To simplify the analysis, the sections of this document do not discuss optimisations with respect to packaging and storage. The data model and metadata element definitions only discuss what information is needed, not how it is packaged, stored and optimised for handling.

2 Layered Data Model

This chapter introduces the structure of the data model which is based on a layered approach for information package definitions (Figure 1). The Common Specification for Information Packages (CSIP) forms the outermost layer. The general SIP, AIP and DIP specifications add, respectively, submission, archiving and dissemination information to the CSIP specification. These two layers are not part of this document. The third layer of the model represents specific content type specifications, such as this ERMS specification. Additional layers for business-specific specifications and local variant implementations of any specification can be added.

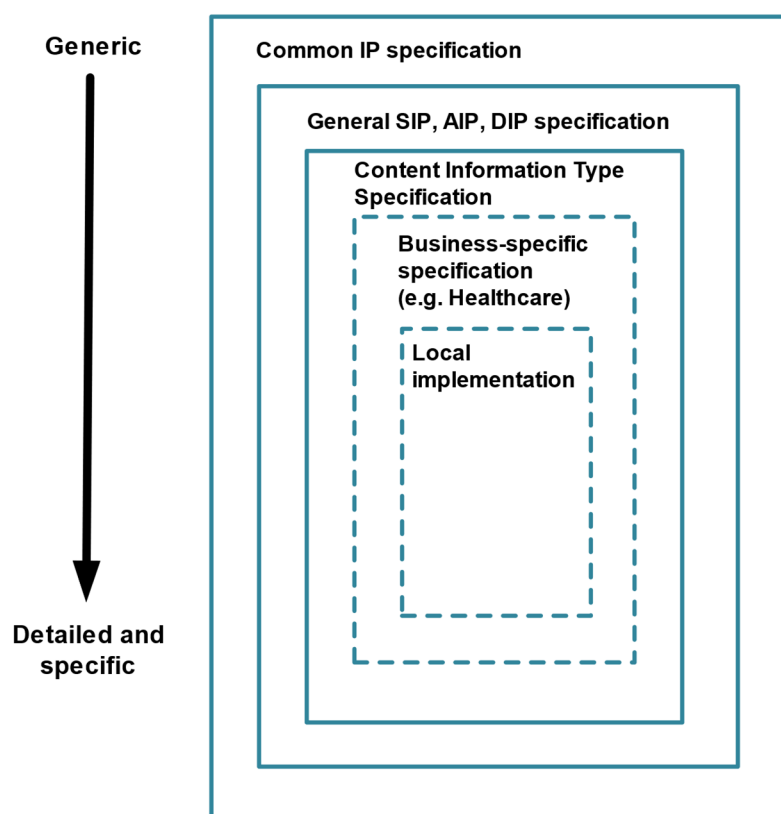


Figure 1: Data Model Structure

The ERMS specification omits all information that is specific to a business area (such as social security) or a specific content-type (such as healthcare). However, these specific types of information may be needed by users of the ERMS specification. This need is addressed by providing extension mechanisms in the ERMS specification so that local (e.g. national) extensions to accommodate local requirements can be added by users.

Every level inherits metadata entities and elements from the higher levels. In order to increase adoption a flexible schema has been developed. This will allow for extension points where the schema in each layer can be extended to accommodate additional information on the next specific layer until, finally, the local implementation can add specific entities or metadata elements to satisfy very specific local needs. Extension points can be implemented via:

- Embedding foreign extension schemas (in the same way as supported by METS² and PREMIS³). These support both increasing the granularity of existing metadata elements by using more detailed data structures as well as adding new types of metadata.
- Single extra metadata elements (as supported by using MoReq contextual metadata elements) without the need to define foreign extension schemas.

² <http://www.loc.gov/standards/mets/>

³ <http://www.loc.gov/standards/premis/>

The structure allows the addition of more detailed requirements for metadata entities for example by:

- Increasing the granularity of metadata elements by using more detailed data structures, or
- Adding local controlled vocabularies.

For consistency, design principles are reused between layers as much as possible.

3 Metadata and Mapping

Metadata can be obtained in several ways that are not mutually exclusive:

- automatically from the source system;
- extracted from the content;
- added manually during submission agreement or ingest.

Ideally metadata should be created or captured as close to its source as possible where it can be most easily or exclusively obtained.

The balance of manual versus automated creation of metadata, as well as the origin of metadata (producer versus archive) varies greatly because of different best-practices and legal environments at the local level. However, in most cases the metadata is a mixture of metadata created manually and in automated fashion, by both the archive and the producer.

In the case of an ERMS export, the specification builds mainly on the MoReq2010 metadata and export service. However, there are some differences between the MoReq export and an archive transfer service:

- There are entities and metadata elements in the MoReq export schema that are not needed for archive export, because:
 - Most existing production systems are not MCRS⁴ and may not be in a position to export according to the semantics and syntax described in the MoReq export schema.
 - The archive does not support the full functionality for a MCRS, including records creation and corresponding workflows, and does not support original technical access restriction management, or manage retention periods.
- There are entities and metadata elements that are needed for archive export (SIP export) but not in the MoReq export, because:
 - Archives may have additional functionality. For example, they may wish to merge records from multiple sources. They then need to map the disparate local producer implementations to a normalised archive implementation. Metadata is needed for this.
 - Archives need additional metadata to address long-term preservation, in particular technical and additional provenance metadata.

⁴ MoReq Compliant Records System (MCRS)

Therefore, the MoReq export schema⁵ is used as an inspiration for a transfer service (the MoReq export schema is not being adopted but built upon).

Note: Not all the metadata is a part of the current specification. Only the use of the metadata related to ERMS specification will be explained in the following metadata sections.

3.1 Moreq2010 Entities and Metadata for the ERMS Specification

In MoReq2010 each core service manages entities belonging to a specified number of entity types⁶ and each entity must belong to only one of the entity types. The MoReq2010 core services refer to the following entity types:

- **Aggregations⁷**

What is the entity?

Aggregation is a core entity in MoReq2010 and defines aggregations as “...accumulations of related record entities that, when combined, may exist at a level above that of a single record”. Note that MoReq2010 does not distinguish between the archival terms Fonds, Sub-fonds, Series, File and Sub-files. These are all categorised as aggregations at various, specifiable levels. They all can be mapped to the MoReq2010 aggregation entity.

How are aggregations dealt with in the ERMS specification?

Different institutions use various combinations and patterns of values for this Entity type. Also, some partners are obliged by law to use specific terms for aggregation levels. Therefore, the vocabulary for titles of the aggregation entity is not controlled by the ERMS but can be freely chosen by the users. It should be controlled locally in the organisation.⁸ We recommend using ISAD-G (in the form of EAD) for contextual descriptions in the archive. See section 3.2 for more details.

- **Class⁹**

What is the entity?

Class is a core entity in MoReq2010 and in all E-ARK partner implementations. Class is defined in MoReq2010 as “a unit of classification that may be associated with an aggregation or a record”¹⁰.

⁵ XML Export Schema. See <http://www.moreq.info/index.php/specification>

⁶ The following entity types appear in the MoReq2010 core services: Aggregations, Classes, Components, Disposal Holds, Disposal Schedules, Entity types, Events, Function Definitions, Groups, Metadata Element Definitions, Records, Roles, Templates and Users [MoReq2010, Glossary].

⁷ Aggregation of individual records or higher level aggregation of aggregations of records [page 249, MoReq2010 Specification].

⁸ Aggregations used in agencies/ERMS are not necessarily the same ones as required/wished for by the archive.

⁹ Business classification applied to records and aggregations of records [page 250, MoReq2010 Specification].

How is class dealt with in the ERMS specification?

The vocabulary for titles of the Class entity is not controlled by this specification but can be freely chosen by users.

- **Component**

What is the entity?

In MoReq2010, a record can have more than one discrete resource making up its content, and these different resources may even be stored in different locations. MoReq2010 defines component as “a part of a record that represents a discrete item of content”.

The association between a record and its content is provided by component entities. Each record can have one or more components. Each referenced component is a single item of content [MoReq2010, Glossary]. A component can either be electronic (referring to a digital resource such as a datafile) or physical (referring to a real-world object such as paper document or DVD).

How is component dealt with in the ERMS specification?

The metadata for this entity type is presented in appendices as a part of a record. It is important to note that each component must belong to only one record (Figure 2) as stated in MoReq2010.

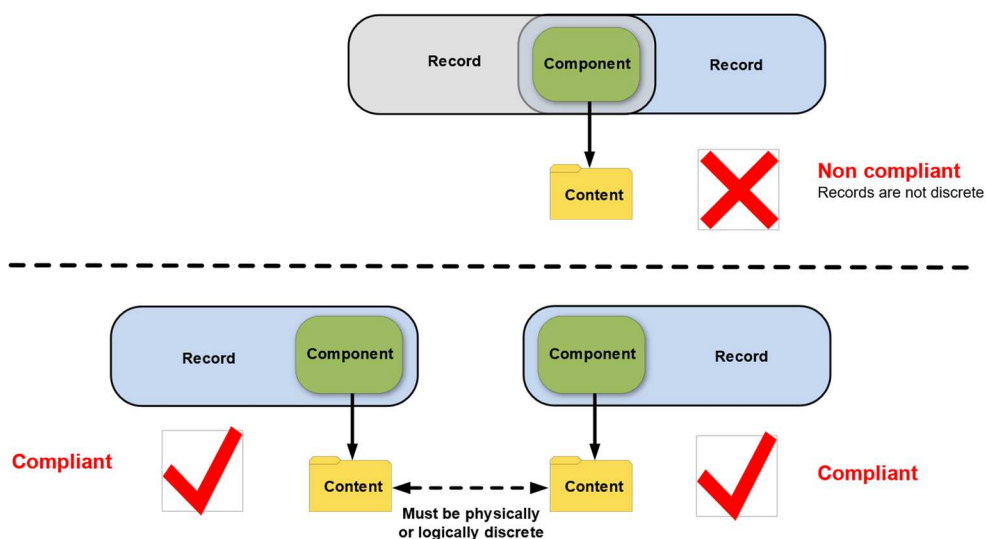


Figure 2: Components

¹⁰ In a somewhat circular definition, MoReq2010 defines classification as “the act of associating a class from a classification scheme to an aggregation or record.” A unit of classification is not defined.

- **Metadata element definitions¹¹**

What is the entity?

In MoReq2010 a metadata element definition sets out the semantics of a metadata element with a list of the required properties of the element. MoReq2010 allows for specialised subtypes of this entity type and divides metadata element definitions into:

- system metadata element definitions.
- contextual metadata element definitions.

How are metadata element definitions dealt with in the ERMS specification?

While specialised elements are not within the scope of this ERMS specification they might still be valid for use by systems as well as archives. Possible use of specialised metadata elements is something every individual needs to consider and describe in Submission Agreements.

- **Disposal holds¹²**

What is the entity?

MoReq2010 defines disposal hold as “a legal or other administrative order preventing the destruction of records”. According to MoReq2010, if the disposal hold is associated with an individual record, it prevents the destruction of that record while the disposal hold remains active.

How are disposal holds dealt with in the ERMS?

Since disposal holds can apply to records in archive custody as well, metadata about disposal holds need to be included in the extraction XML.

- **Disposal schedules¹³**

What is the entity?

According to MoReq2010, disposal schedules are critical to managing records because a record in an MCRS may only be destroyed as part of a disposal process governed by the disposal schedule applicable to that record. A record’s disposal schedule determines how long a record is retained and how it is subsequently disposed of at the end of its retention period.

How are disposal schedules dealt with in the ERMS specification?

Since disposal schedules do not possess the same importance for the archive’s records preservation activities, they will only be archived as metadata and their ERMS

¹¹ Definition of the properties of a contextual metadata element [page 252, MoReq2010 Specification].

¹² A disposal hold is a legal or other administrative order that interrupts the normal disposal process and prevents the destruction of some of an organisation’s records while the disposal hold is in place [MoReq2010, Glossary].

¹³ Disposal schedules are used to manage the life cycles of records in all MCRS solutions [MoReq2010, Glossary].

functionality is not supported in the archive. When using the specification for transfer between different ERMS systems, this information needs to be included in an extension point.

- **Events¹⁴**

What is the entity?

Events are defined in MoReq2010 as “an entity that is generated by performing a function”. Events are not independent entities insofar as all entities, except access control lists and events, will have an associated event history in the MCRS, consisting of a description of all the events in which the entity has participated.

How are events dealt with in the ERMS?

To simplify the MoReq2010 model and make it easier to understand events the descriptions are stored with its record entity, instead of as a separate entity.

- **Function definitions¹⁵**

What is the entity?

Function definitions are used to define operational functions and are represented as entities. Function definitions are used for both access control (roles, users, groups) and in events that are generated by performing functions. When events are generated, the function definition of the function that was performed is included in the event.

How are function definitions dealt with in the ERMS specification?

Functions are described as part of events in another entity description (instead of being a stand-alone entity) as function definitions only define functions which can be performed with an entity by a user in MCRS.

- **Groups,¹⁶ Roles¹⁷ and Users¹⁸**

What is the entity?

¹⁴ A description of the outcome of a function that was performed previously and is retained to show the history of an entity [page 256, MoReq2010 Specification].

¹⁵ A definition of function that can be performed with an entity by a user [page 257, MoReq2010 Specification].

¹⁶ An entity type that usually represents a team or business unit within the organisation and has various user entities as members [MoReq2010, Glossary].

¹⁷ An entity representing a set of function definitions. Granting a role to a user or group in relation to an entity enables that user, or any member of that group, to perform that role on the entity and its descendants. Roles are generally constructed to mirror the tasks of a staff member filling a particular position within the organisation. For example, different roles may be constructed around each of the following usage types: office clerk, local records officer, senior records manager, personnel manager, sales representative, auditor, external contractor, guest or office temp, executive personal assistant, senior executive officer, etc. [MoReq2010, Glossary].

¹⁸ A person or system with an account which enables access to and use of an MCRS. The user does not have to be a human and could be another business system. Users must be authenticated before they can use an MCRS [MoReq2010, Glossary].

In MoReq2010 these are separate entity types but for ERMS purposes we have chosen to describe their use in the ERMS specification together. MoReq2010 allows for specifying individual users who participated in events, as well as their roles. Different use cases require keeping different kinds of information about such entities. Role-based event information may suffice for records of archival value. User-level event information is needed if archived materials are used for legal reasons (legal deposit, other legal scenarios such as discovery orders).

How does the ERMS specification deal with these entities?

Groups should be exported by MoReq as individual Users because the Group Entity type as a functional entity is not supported in all implementations. Roles and Users will only be archived as metadata (if they are related to some actions/events), but their functionality is not supported in the archive.

- **Record¹⁹**

What is the entity?

A record is a core entity in MoReq2010 and is defined as any “information created, received and maintained as evidence and information by an organisation or person, in pursuance of legal obligations or in the transaction of business” (ISO 15489-1:2001, 3.15).

How is record dealt with in the ERMS specification?

The record entity, being at the heart of MCRS functionality and of archival holdings is described fully by this specification. The metadata for this entity is set out in Table 3: .

- **Service²⁰**

What is the entity?

There is an export service sub-type of the MoReq2010 service entity type that specifies the process of exporting records and metadata from a MoReq compliant records system to another MCRS.

How is service dealt with in the ERMS specification?

The export service is used as a basis for defining the format to be used when exchanging records between the ERMS and the archives.

- **Template**

What is the entity?

¹⁹ Record of a business transaction made up of one or more components that are managed atomically [page 260, MoReq2010 Specification].

²⁰ A logical subset of the total functionality of an MCRS that focuses on managing only one or a small group of entity types. For example, the disposal scheduling service only manages disposal schedules [MoReq2010, Glossary].

According to MoReq2010 a template is a set of contextual metadata element definitions that can be used to add contextual metadata elements to entities at creation or later.²¹

How is template dealt with in the ERMS specification?

The ERMS specification does not use the MoReq2010 metadata templates, as such metadata, if it exists in the MCRS, it will be recorded by other means: for example in an extension.

3.2 Translating MoReq2010 Class and Aggregation Values

As well as general descriptive and administrative metadata, there are two important metadata entities in MoReq2010 which can be successfully incorporated into the export document to maintain contextual information needed for provenance and authenticity. MoReq2010 specifies both *class* and *aggregation* as entity types used for managing and accessing records in a MCRS. *Class* is a unit of classification that can be associated with a record or an aggregation and is used to relate records and aggregations to the business activity (functions, activities, transactions, etc.) which produced the records. Although class values can usually be organised hierarchically (Figure 3), it is not mandatory (and sometimes unnecessary) to do so (Figure 4).

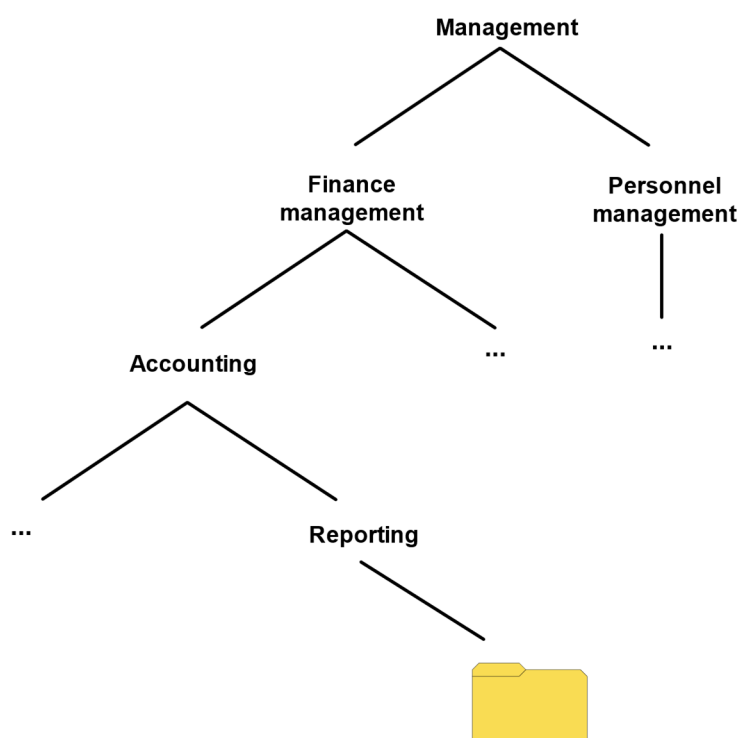


Figure 3: Hierarchical classes

²¹ See page 263, MoReq2010 Specification. Contextual metadata is defined as “metadata that is not mandated by MoReq2010 but is created within an MCRS in a local context to support the local business needs and operations of an organisation” [Page 203, MoReq2010 Specification].

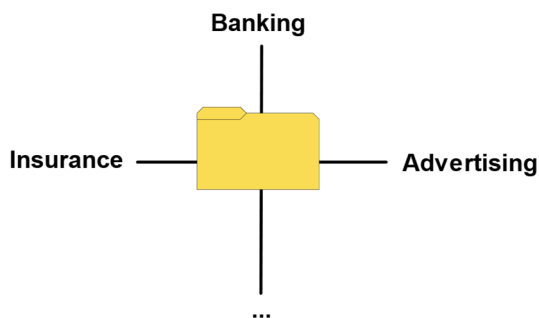


Figure 4: Non-hierarchical classes

An *aggregation* is any accumulation of record entities at a level above the record object (e.g. folder, series, fonds, etc., see Figure 5²²).

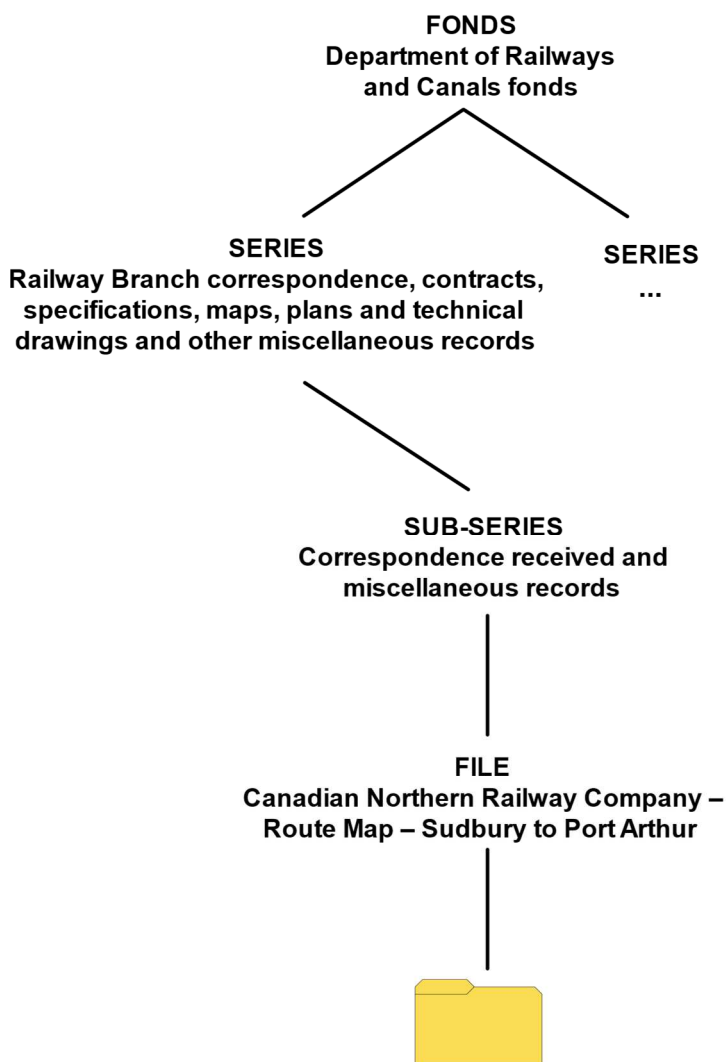


Figure 5: Aggregation

²² The example is based on ISAD(G) General International Standard Archival Description. See <https://www.ica.org/en/isadg-general-international-standard-archival-description-second-edition>

It is important to note that classification is not a way of structuring records but is a way of categorising records for management and access purposes. In contrast, aggregation is a way of structuring records to place them into the context of their creation and use. Because the records in aggregations arise out of business activities, information about both the aggregation to which a record belongs and the business activity which caused the record to be created is required to fully understand the context of a record. Such metadata must accompany the SIP and be incorporated into the systems in use by the receiving archive.

Chapter 3 above specifies that class hierarchies and aggregation structures are to be represented using ERMS metadata. Many of the aspects of the submissions from producers are governed by law or existing constraints. Producers may, for example, not be able to submit complete aggregation information with a record, or may not be able to choose whether they submit a single record or a whole series, or may be obliged to record information from several classification schemes. In these cases it is not possible to specify a mandatory requirement for implementing tools in one specific way or specify rigid metadata structures that are mandatory.

It is not possible to anticipate all possible data dependencies between the producer and archive representations and our data model, metadata profile, workflows and requirements therefore cannot implement all possibilities. We are trying to capture the most common ones found in our stakeholder analysis. The aim is not to specify a catch-all solution but provide guidelines for the most critical issues.

3.2.1 Mapping between ERMS and Archive Aggregations

When producer class and aggregation values are received in a SIP they can be used by the archive in a number of ways:²³

- Incorporated as raw values into the Encoded Archival Description (EAD) record for the AIP
- Mapped and translated into the archive's EAD profile
- Incorporated into the archive's EAD profile by extending the EAD profile.
- Archived as an ERMS document containing the class and aggregations values referenced in the archival description or EAD profile.

EAD uses aggregation values as the "level" attribute on the elements <archdesc> and <c>, to specify the aggregation level at which description belongs (Example 1).

²³ The ERMS specification presents only one mechanism (using ERMS metadata) for how MCRS aggregation values can be translated to archival aggregation values, and do not restrict the use of any other method (i.e. adding all relevant class values as keywords to each individual record).

Example 1:

```

<archdesc level="fonds">
  ...
  <dsc>
    <c level="series">
      ...
      <c level="file">
        Records and computer files
      </c>
    </c>
  </dsc>
</archdesc>

```

The names of aggregation levels depend on the agreements between data producers and archives. EAD3 has defined a set of values (class, collection, file, fonds, item, otherlevel, recordgrp, series, subfonds, subgrp, subseries) for that purpose, but it allows other values to be used as well if they are defined as “otherlevel” (Example 2).

Example 2:

```

<archdesc level="collection">
  ...
  <dsc>
    <c level="series">
      ...
      <c otherlevel="case"> <!--A new aggregation level-->
        Records and computer files
      </c>
    </c>
  </dsc>
</archdesc>

```

3.3 Using the ERMS specification together with the Common Specification for Information Packages (CSIP)

When an ERMS document has been created, it can be transmitted in a package following the principles and requirements described in the CSIP specification, <http://earkcsip.dilcis.eu/>.

3.3.1 Specific fields to use in CSIP

Table 1: Specific fields to use in CSIP

Element name	METS path	Value
General content type	mets/@TYPE	Dataset
Specific content type	mets/@csip:CONTENTINFORMATIONTYPE	ERMS
Specific content type	fileGrp/@csip:CONTENTINFORMATIONTYPE When the FileGrp describes a Representation	ERMS

3.3.2 Placement of data in a CSIP Information Package

The ERMS document is placed as a representation file following the instructions in CSIP.

4 Glossary

Table 2: Glossary

Name	Description
Aggregation	Aggregations of records are accumulations of related record entities that, when combined, may exist at a level above that of a single record. Aggregations of records may reflect relationships such as shared characteristics or attributes, or the existence of sequential relationships between related records [MoReq2010, v 1.1].
AIP	Archival Information Package
Class	A unit of classification that may be associated with an aggregation or a record. In MoReq2010®, classes always have a default disposal schedule, which is inherited by any record they classify, in accordance with the principle in ISO 15489 that 'Classification of business activities acts as a powerful tool to assist the conduct of business and in many of the processes involved in the management of records including ... determining appropriate retention periods and disposition [i.e. disposal] actions for records' (ISO 15489-1:2001, 9.5.1) [MoReq2010, v 1.1].
Component	A part of a record that represents a discrete item of content. For completeness, a record, including all its components and their content, must be managed atomically.
Contextual Metadata	Metadata that is not mandated by MoReq2010 but is created within an MCRS in a local context to support the local business needs and operations of an organisation.
Contextual	The contextual metadata element definition is the definition of a contextual

metadata element definition	metadata element. Contextual metadata element definitions must be exported whenever contextual metadata is exported to ensure that an MCRS that imports the export data can interpret the metadata element and represent it correctly.
DIP	Dissemination Information Package
EAD	<p>Encoded Archival Description. A non-proprietary de facto standard for the encoding of finding aids for use in a networked (online) environment based on ISAD(G). Finding aids are inventories, indexes, or guides that are created by archival and manuscript repositories to provide information about specific collections. While the finding aids may vary somewhat in style, their common purpose is to provide detailed description of the content and intellectual organization of collections of archival materials. EAD allows the standardisation of collection information in finding aids within and across repositories.</p> <p>http://www.loc.gov/ead/eadabout.html [accessed on May 17, 2019]</p>
EAC-CPF	<p>Encoded Archival Context – Corporate bodies, Persons, and Families (EAC-CPF). A non-proprietary de facto standard for encoding the names of creators of archival materials and related information. EAC-CPF is based on ISAAR(CPF). See http://eac.staatsbibliothek-berlin.de/ [accessed on May 17, 2019]</p>
Entity	<p>Entities represent individual and discrete units of information within an information system. In an MCRS, each entity must be of a particular entity type and have some, or all, of the following:</p> <ul style="list-style-type: none"> • system metadata, • contextual metadata, • access control list, • event history. <p>The system metadata, and sometimes the contextual metadata, link the entity to other entities, forming relationships [MoReq2010, v 1.1].</p>
ERMS	Electronic Records Management System
IP	Information Package
MCRS	MoReq Compliant Records System
METS	<p>Metadata Encoding and Transmission Standard. A de-facto standard for describing information packages. See http://www.loc.gov/standards/mets/ [accessed on May, 17, 2019]</p>
MoReq2010	<p>MoReq2010: Modular Requirements for Records Systems. See https://www.moreq.info/files/moreq2010_vol1_v1_1_en.pdf [Accessed on May</p>

	17, 2019]
PREMIS	PREservation Implementttation Strategies. A de-facto standard for preservation metadata. See http://www.loc.gov/standards/premis/ [accessed on May 17, 2019]
Record	<p>Any ‘information created, received and maintained as evidence and information by an organisation or person, in pursuance of legal obligations or in the transaction of business (ISO 15489-1:2001, 3.15)’. In MoReq2010, a record may be further characterised as follows.</p> <ul style="list-style-type: none"> • It has an extensible set of metadata that describe it. • It has one or more components that represent its content. • It is classified with a business classification. • It has a disposal schedule that describes explicitly if, how and when it will be disposed of or destroyed. • It belongs to an aggregation of records. • Access to it is controlled and limited to authorised users. • Its destruction may be prevented by a disposal hold. • It may be exported to another MCRS while retaining all of the characteristics listed above [MoReq2010, v 1.1].
SIP	Submission Information Package

5 Metadata

5.1 Model picture

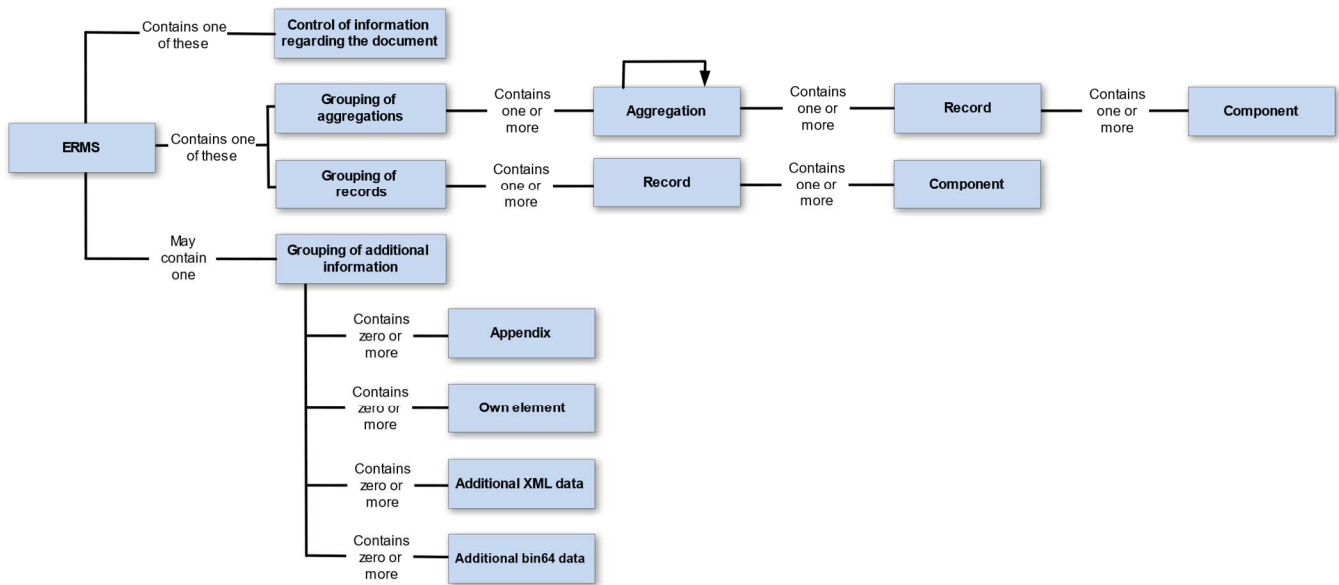


Figure 6: Components of the ERMS XML-format

5.2 Metadata for the Control element

Each XML document containing ERMS information has a mandatory “Control” element to use in an ERMS transfer. The element describes information regarding the ERMS file.

The table will be present in the next version of the specification.

5.3 Metadata for ERMS Records

The following tables contain elements to be used in an ERMS transfer.

The XPath is not yet entered in the table.

Table 3: ERMS use of metadata for Records

ID	Name and location	Description and usage	Cardinality
	Created	Date and time the entity was created, set by the system	1..1
	Originated	Date and time of origin of a record or other entity which may vary from the creation date of the entity in the system	0..1
	Title	The identifying name or title of the entity	1..1
	Description	A description of the entity	0..1
	ParentAggregationId	Parent aggregation for a child aggregation or record	0..1
	DisposalScheduleId	An element for documenting decisions and actions related to assessing the archival value and disposition of the materials being described.	0..1
	RetentionStart	System generated date calculated from the record’s disposal schedule indicating the start of its retention period	0..1
	DisposalAction	Code describing the action to be taken on disposal of the record	0..1
	DisposalActionDue	The calculated date that the record is due for disposal	0..1
	LastReviewedComment	Comment made by the user who last reviewed the record explaining the disposal decision made by that review	0..1
	LastReviewed	System set date and time indicating when the last review was completed	0..1
	Transferred	System set date and time indicating when the transfer of the record was confirmed	0..1
	Duplicates	Reference to another entity that has been created by duplicating the record, component or event, and is an exact	0..1

ID	Name and location	Description and usage	Cardinality
		copy up to the event of duplication, with an identical provenance	
	TransformationEntry	An element for recording an event of a transformed record	0..n
	SystemId	Universally unique identifier for an entity that is generated automatically by the system and stays with the entity forever	1..1
	Extrald	Any external identifier that is used by an ERMS system or is required in a country (use localtype and label attributes for type or name of the ID that is used in this field)	0..n
	Notes	Notes	0..n
	Subject	Subject of the record – use localtype, relator attributes on subject element to mark keyword source type (free text described by creator or ontology subject related by archivist)	0..n
	Keywords	Keywords describing the content	0..n
	GeographicalLocations	List of geographical locations related to the content	0..n
	FindingAid	Information about any finding aids that the repository or records creator may have that provide information relating to the context and contents of the unit of description	0..n
	Description Source	References to publications and other materials used for description	0..n
	Creator	An entity primarily responsible for making the content of the resource; an entity primarily responsible for making the resource (examples of a Creator include a person, an organisation, or a service)	0..n
	Owner	Owner of the record	0..n
	Administrator	Administrator of the record	0..n
	Reader	Everyone who should be able to read the contents of the record (in the source ERMS system)	0..n
	Sender	Sender of the record	0..n
	Editor	Person(s) who could edit the record (including adding) in the source ERMS system	0..n
	Recipient	Recipient of the record	0..n

ID	Name and location	Description and usage	Cardinality
	Other	Other persons/organisations related to the record	0..n
	RecordType	Type of the record	0..n
	RecordLevelName	Level in archival hierarchy	0..n
	MediumType	Structured statements describing the type of the physical medium of the materials	0..n
	ExtentUnit	The unit used to describe the extent of the record (e.g. MB, pages, number of files/components)	0..n
	ExtentValue	The extent of the record expressed in the unit in the previous field	0..n
	RelatedRecord	Related record and type of relation	0..n
	ArchivalHistory	Information on the history of the unit of description that is significant for its authenticity, integrity and interpretation	0..n
	OwnTextElement	Any additional metadata	0..n
	MainSignatureDate	Date of main signature	0..1
	MainSigner	Name of responsible person who signed the record	0..1
	MainSignerRole	Main signing person role	0..n
	OtherSignatureDate	Date of other signature	0..n
	OtherSigner	Other person signing the record	0..n
	OtherSignerRole	Other signing person role	0..n
	DispatchDate	Date of dispatch of the record	0..1
	Dispatcher	Person responsible for dispatching the record	0..1
	Addressee	Original addressee of the record	0..n
	DispatchMode	Mode of dispatching of the record	0..n

5.4 Metadata for ERMS Aggregation

The following tables contain elements to be used in an ERMS transfer.

The XPath is not yet entered in the table

Note: The following table contains guidelines for most common cases.

Table 4: ERMS use of metadata for Aggregations

ID	Name and Location	Description and usage	Cardinality
	Created	System set date and time showing when the entity was created	1..1
	Originated	Date and time of origin of a record or other entity which may vary from the creation date of the entity in the system	0..1
	FirstUsed	System generated date and time indicating when an entity was first used; generally taken as the last time it can be modified or deleted without formally destroying it	0..1
	LastAddition	System set date and time indicating when the most recent record or child aggregation was added to the parent aggregation	0..1
	ClassId	An ID of the file plan	0..n
	Title	The identifying name or title of the entity	1..1
	Description	A description of the entity	0..1
	ScopeNotes	An element that provides information about the nature of and activities reflected in the described materials	0..1
	Closed	System set date and time indicating when the aggregation was closed	0..1
	Destroyed	System set date and time indicating when the entity was destroyed	0..1
	MaxLevelsOfAggregation	The maximum number of levels of aggregation allowed below a root aggregation	0..1
	ParentAggregationId	Parent aggregation for a child aggregation	0..1
	SystemId	Universally unique identifier for an entity that is generated automatically by the system and stays with the entity forever	1..1
	Extrald	Any external identifier that is used by an ERMS system or is required in a country (use localtype and label attributes for type or name of the ID that is used in this field)	0..n
	Notes	Notes	0..n
	Subject	Subject of the record (free text described by creator or ontology subject related by archivist)	0..n

	Keywords	Keywords	0..n
	GeographicalLocations	List of related geographical locations	0..n
	FindingAid	Information about any finding aids that the repository or records creator may have that provide information relating to the context and contents of the unit of description	0..n
	Publication	Publications that are about or are based on the use, study, or analysis of the unit of description	0..n
	Description Source	References to publications and other materials used for description	0..n
	Creator	An entity primarily for making the content or the resource; an entity primarily responsible for making the resource (examples of a Creator include a person, an organisation, or a service)	0..n
	Owner	Responsible person or role	0..n
	Editor	Person(s) who can edit the aggregation (including adding)	0..n
	Administrator	Administrator of the aggregation	0..n
	Reader	Everyone who should be able to read the contents of the aggregation	0..n
	Other	Other persons/organisations related to the aggregation	0..n
	AggregationType	Type of the aggregation (e.g. case file, subfile etc.)	0..n
	MovedRecords	Information about records that have been moved to other aggregations	0..n
	DeletedRecords	Explanation of the fact that the record has been deleted by the administrator or has been destroyed due to technical errors.	0..n
	AggregationStatus	Status of the aggregation	0..n
	Decision	Decision about the aggregation	0..n
	DecisionDate	Date of the decision about the aggregation	0..n
	DecisionDeadline	Deadline of decision about the aggregation	0..1
	ArchivalHistory	Information on the history of the unit of description that is significant for its authenticity, integrity and interpretation	0..n
	ReceivedDate	Date and time when the aggregation was received	0..1

	ClassificationDate	Date of classification	0..1
	OwnershipStartDate	Date when ownership started	0..1
	PhysicalLocation	Physical or logical placement of the aggregation	0..n
	ExtentUnit	The unit used to describe the extent of the aggregation (e.g. MB)	0..n
	ExtentValue	The extent of the aggregation expressed in the unit in the previous field	0..n
	SystemOfArrangement	Information on the internal structure of the aggregation, the order and/or the system of contents	0..n
	RelatedAggregation	ID of a related aggregation	0..n
	OwnTextElement	Any additional metadata	0..n
	Security classification	Security classification of the records in the aggregation	0..n
	Access restriction ExplanatoryText	Textfield for describing the restriction	0..n
	Regulation	An element for indicating the statutorily defined status of the aggregation	0..n
	RestrictionStartDate	Start date of the restriction	0..n
	RestrictionEndDate	End date of the restriction	0..n
	RestrictionDuration	Duration of the restriction	0..n
	RestrictionType	Type of the restriction	0..n
	AppraisalDate	Date of appraisal	0..n
	AppraisalDecision	Appraisal decision	0..n
	AppraisalDescription	Appraisal description	0..n
	IPROwner	Copyright owner	0..n
	Reproduction conditions	IPR condition description	0..n
	IPRDuration	Restriction duration	0..n
	IPRType	Reference to IPR type according legislative act	0..n
	CopyrightNotice	Information about copyrights	0..n
	LoanDate	Date of loan	0..n

	Borrower	Borrower	0..n
	AuthorizingPerson	Person authorising the loan	0..n
	TakeBackDate	Date of take back	0..n
	ResponsiblePerson	Person responsible for taking back after loan	0..n
	ArchiverName	Person responsible for inhouse archiving	0..n
	ArchivingDate	Date of inhouse archiving	0..n
	DisposalDate	Disposal date	0..n
	DisposingPerson	Person carrying out the disposal action	0..n
	SupervisingPerson	Person supervising the disposal	0..n
	TransferDate	Date of transfer to the archive	0..n
	Deliverer	Person responsible for the delivery to the archive	0..n
	Recipient	Person responsible for receipt in the archive	0..n

5.5 Example of use of own elements

It is possible in this specification to add single extra elements following these examples.

```
<OwnElement>
  <OwnElementDescription>Own element used for detailing accounting
information</OwnElementDescription>
  <OwnElement Name="Responsible unit" DataType="String" Format="Used accounting system">
    <Value>3456/206/86176</Value>
    <Property>
      <Attribute Name="Accounting information">
        <Value>Se-1234-3214-444</Value>
      </Attribute>
    </Property>
  </OwnElement>
</OwnElement>
```

```
<OwnElement>
  <OwnElementDescription>Own element used for detailing accounting information and value
representing the accountant</OwnElementDescription>
  <OwnElement Name="Responsible unit" DataType="String" Format="Used accounting system">
    <Value>3456/206/86176</Value>
    <Property>
      <Attribute Name="Accounting information">
        <Value>Se-1234-3214-444</Value>
      </Attribute>
    </Property>
    <OwnElement Name="Accountant" DataType="String" Format="Username">
      <Value>MARJAAS</Value>
    </OwnElement>
  </OwnElement>
</OwnElement>
```

```
<OwnElement>
  <OwnElementDescription>Comments regarding the system</OwnElementDescription>
  <OwnElement Name="Comment" DataType="String">
    <Value>System comment 1</Value>
  </OwnElement>
  <OwnElement Name="Comment" DataType="String">
    <Value>System comment 2</Value>
  </OwnElement>
</OwnElement>
```

5.6 Value other in value lists

In the value lists for the attributes there is always a value “OTHER” or “OWN” present to accommodate the possibility to use values used in one’s own system. When the value is selected, the use of an attribute with the same name and the prefix “OTHER” is validated with the Schematron rules. The use of the value “OTHER” or “OWN” needs to be stated in a transmission agreement, as well as which values that can be used.

6 Appendix 1

AUTHOR(S)

Name(s)	Organisation(s)
Karin Bredenberg	National Archives of Sweden (NAS)

REVIEWER(S)

Name(s)	Organisation(s)
Jaime Kaminski	Highbury IVS
Janet Anderson	Highbury IVS
DILCIS Board	DILCIS Board

Project co-funded by the European Commission

Policy Support Programme

within the ICT

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
0.1	22 April 2015	Angela Dappert	DLM	Draft outline.
0.2	28 April 2015	Angela Dappert	DLM	Draft outline slightly updated.
0.3	14 August 2015	Angela Dappert	DLM	Incorporate issues from ERMS meetings.
0.35	16 October 2015	Tarvo Kärberg	NAE	Reorganising, incorporating feedback.
0.4	10 November 2015	Tarvo Kärberg	NAE	Updating the content, incorporating feedback.
0.5	12 November 2015	Andrew Wilson	UPHEC	Updating the content. Adding new information about SIP to AIP transformation.
0.6	16 November 2015	João Cardoso	IST	Updating the content.
0.7	17 November 2015	Levente Szilágyi	NAH	Updating the content. Adding information about metadata tables.
0.71	19 November 2015	Tarvo Kärberg	NAE	Incorporating feedback, cleaning the text, merging the content.
0.8	30 November 2015	Tarvo Kärberg	NAE	Merging the content.
0.9	13 January 2016	Tarvo Kärberg	NAE	Merging the content.
0.91	15 January 2016	Levente Szilágyi	NAH	Updating the content. Adding information about metadata tables.
0.92	19 January 2016	Jože Škofljanec	SNA	Updating the content related to SFSB records.
0.93	21 January 2016	Gregor Završnik	SNA	Updating the content related to geodata.
0.94	22 January 2016	Levente Szilágyi	NAH	Updating the content related to EAD tables.
0.95	26 January 2016	Alex Thirifays	DNA	Quality assurance and proof-reading.
0.96	28 January 2016	Kuldar Aas	NAE	Quality assurance and proof-reading.
0.97	29 January 2016	Andrew Wilson	University of Brighton	Quality assurance and proof-reading.
1.0	29 January 2016	Tarvo Kärberg	NAE	Final version (part of D3.3).

1.1	20 July 2016	Tarvo Kärberg	NAE	Incorporating additional feedback from Andrew Wilson and Advisory Board.
1.2	30 September 2016	Tarvo Kärberg	NAE	Incorporating agreements made in the Common Specification work group.
1.3	18 November 2016	Tarvo Kärberg	NAE	The ERMS specification was split in two. This specification contains information about ERMS only from this point forward.
1.4	23 November 2016	Tarvo Kärberg	NAE	Updating appendices II and III.
1.5	05 January 2017	Levente Szilágyi	NAH	Adding MoReq2010 based EAD XML.
1.6	09 January 2017	Levente Szilágyi	NAH	Updating aggregation elements.
1.7	12 January 2017	Tarvo Kärberg	NAE	Finalising the document.
1.8	23 November 2018	Jaime Kaminski	DLM	Quality assurance and proof reading.
1.9	25 November 2018	Karin Bredenberg	NAS	Update according to the new schema for ERMS. Tables not ready. Draft for review, E-ARK4ALL project.
2.0	31 May 2019	Karin Bredenberg	NAS	Update after review.
2.0	31 May 2019	Janet Anderson	DNA	Final proof read
2.0	31 May 2019	DILCIS Board	DILCIS Board	Release of version 2.0

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.