D.3.3 SMURF (Semantically Marked Up Record Format) Profile

.

DOI: 10.5281/zenodo.1172618

Grant Agreement Number:	620998
Project Title:	European Archival Records and Knowledge Preservation
Release Date:	13th February 2018

Contributors	
Name	**Affiliation**
Tarvo Kärberg	National Archives of Estonia
Angela Dappert	DLM Forum
Andrew Wilson	University of Brighton
Levente Szilágyi	National Archives of Hungary
Jože Skofljanec	National Archives of Slovenia
João Cardoso	Danish National Archive
Gregor Zavrsnik	National Archives of Slovenia
Kuldar Aas	National Archives of Estonia
Alex Thirifays	Danish National Archive
David Anderson	University of Brighton
Janet Anderson	University of Brighton

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

1   Context

1.1   Purpose and Scope

The purpose of this document is to describe SMURF (semantically marked up record format) profile, which includes ERMS (electronic records management systems) and SFSB (simple file-system based) records as described below. When extracting information from a producer's system one has the choice of two generic options:

1. **Extracting data in a relational database structure**
   Extracting data from a relational database into a long-term preservation format (SIARD) that preserves the properties of the relational database so that the data can be imported into a relational database management system (RDBMS) on Access. Access can happen via database queries or via a search field.[1]
   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 purposes of research.
   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 can be read in the SIARD 2.0 Profile Specification.[2]

2. **Extracting data and metadata as records**
   Extract the records and normalise them to a standard E-ARK 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. Their structured semantic metadata description is explicit rather than hidden inside a RDBS.
   The representation of descriptive metadata inside the archive can be in the E-ARK SMURF AIP format and/or another native archive format. The main advantages over the RDBS representation are that:
   o Records from different sources can be merged.
   o Search and access is possible across all records from all sources.
   o Records can be managed and accessed uniformly.
   o The original database / records system software does not need to be licensed and preserved.
   This case (specifying the semantically marked-up metadata profile) will be presented in this SMURF profile document.

SMURF is further divided into two sub-formats:[3]
- ERMS (MoReq-based records)[4] – records that have been exported from an ERMS according to a metadata standard. It is possible to partly reuse this metadata in order to create archival descriptive metadata (i.e. EAD[5]) automatically.
- SFSB (simple file-system based records) – records that contain simple file-system based folders or files, including those originating from content and data management systems, such as SharePoint. SFSBs are for example operating systems' native file systems designed to organize

---

[1] The E-ARK will aim to provide access via a search field, but that is not a must criteria.
[2] The SIARD 2.0 specification (http://www.eark-project.com/resources/specificationdocs/32-specification-for-siard-format-v20/file) represents the SIP profile for the relational databases content type.
[3] Although we support both profiles (ERMS, SFSB) the main focus is on ERMS.
[4] MoReq2010 Specification v1.1, http://www.moreq.info/files/moreq2010_vol1_v1_1_en.pdf
[5] Encoded Archival Description, http://www.loc.gov/ead/

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

files and folders. The data requires manual enrichment with additional descriptive metadata. Additional pre-ingest tools (e.g. EAD editor[6]) may be required to acquire this metadata. Otherwise these records share the metadata specification and workflows of the general ERMS records, with possibly fewer[7] mandatory elements. For example, one specific sub-type of the SFSB profile is geodata which we will use to demonstrate the meaning of the SFSB content format.

This specification is implemented in tools that:

- Extract metadata and data from the native producer systems.
- Create a SIP package from them.
- Validate that metadata and data conforms to the specification, is complete, and internally consistent.
- 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,[8] is complete, internally consistent.
- Ingest the AIP into Archival Storage.
- Manage AIPs within the archive.

1.2 Methodology

The specification is based on the definition of specific use cases and requirements to be implemented by E-ARK tools for electronic records system and simple file-system based records.

The work involved several stakeholders:

- 6 national archives and 2 records' producers to specify the requirements.
- Archival service providers to develop and test the specifications and tools.

Involving so many stakeholders allowed us to learn from actual implementations.

The entities and metadata elements presented in the SMURF profile were inherited from the Common IP[9] and the SIP/AIP specifications, and extracted from standards (MoReq 2010,[10] EAD), and existing implementations in archives.

It was decided in the specifying process:[11]

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

---

[6] EAD editor, https://github.com/ewg118/eaditor
[7] as SFSB does assume manual creation of metadata and thus includes less mandatory elements
[8] E-ARK AIP pilot specification, 2016, http://eark-project.com/resources/project-deliverables
[9] E-ARK Draft Common Specification for Information Packages in the E-ARK project, version 0.13, http://eark-project.com/resources/specificationdocs/50-draftcommonspec-1
[10] MoReq2010 Specification v1.1, http://www.moreq.info/files/moreq2010_vol1_v1_1_en.pdf
[11] It is important to note that the MoReq export schema is intended for inter-records-systems exchange rather than for archive deposit. This is a different use case from the ones we are addressing. Because of this, we needed to identify metadata elements in the export schema that were not preservation-relevant and identify preservation-relevant metadata elements that were missing from the MoReq 2010 schema.

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

How did we decide which entities and metadata elements to adopt in the SMURF profile? Firstly, we adopted metadata standards that were accepted and in frequent use for the needed functions in the ERMS and archive communities. Secondly, we did not adopt every entity or metadata element defined in these standards, but only the ones that were relevant for the ERMS and SFSB scenario to archive submission use cases and

- 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.

This knowledge allowed us to identify which requirements, processes, entities and metadata elements are mandatory for every use of the SMURF profile.

Rather than adopting a standard we agreed to adapt them if necessary. For example, we relaxed mandatory MoReq requirements for metadata elements if they could not be supplied in practice.

We also defined extension points so that metadata elements can be added to support local needs.

Profile-dependent tools need to be used to add functionality to the pre-ingest so that, for example, missing mandatory metadata can be manually added for the simple file-system based records case if it is not submitted with the files (i.e. an SIP cannot consist only of computer files). The tools are intended to be used by the producer and should be implemented in a way friendly to the occasional user.

1.3 Limitations

In this document the following are <u>out of scope</u>:

- Proprietary extraction formats, even if they were accompanied by their extraction schema, functional or records system specification.
  These formats have different:
  - o 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),
  - o Degrees of metadata associated with them,
  - o Degrees of authenticity,
  - o Dependencies on knowledge about the functionality of the system in which the record can be used.
- The XML Schema for the SMURF profile. The schema will be provided in the final SIP specification in 2017.
- The E-ARK AIP format (to be covered by a separate document, the E-ARK AIP specification).[12]

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.

---

[12] E-ARK AIP pilot specification, released January 2016, http://eark-project.com/resources/project-deliverables

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

2   Layered Data Model

This chapter introduces the structure of the data model. E-ARK uses a layered approach for information package definitions (Figure 1). The Common IP Specification for all information packages (SIP, AIP, DIP) forms the outermost layer. The general SIP, AIP and DIP specifications add, respectively, submission, archiving and dissemination information to the E-ARK Common IP specification. These two layers are not part of this document as the SMURF specification is content-type specific and it covers only ERMS and SFSB scenarios as introduced earlier.
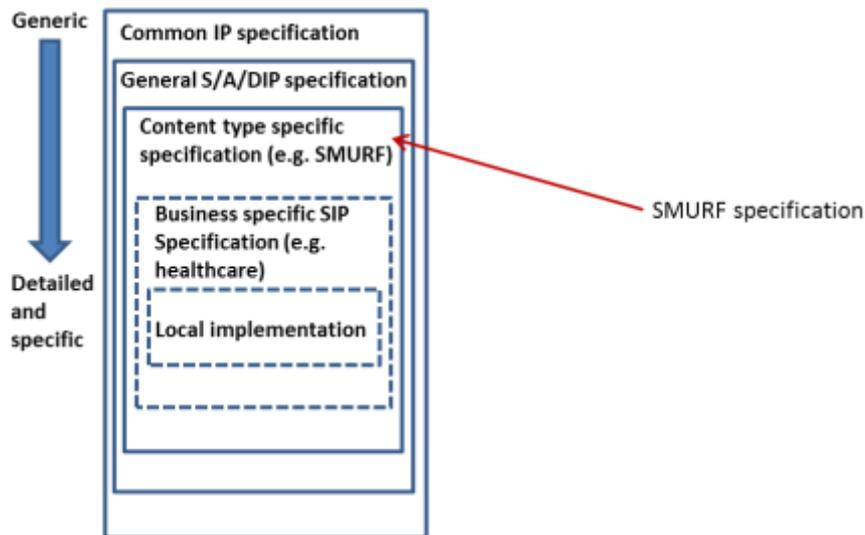

Figure 1: Data Model Structure

The SMURF profile omits all information that is specific to a business area (such as social security) or a specific content-type (such as healthcare).[13] However, these specific types of information may be needed by users of the E-ARK SMURF profile. We address this need by providing extension mechanisms in our IP specifications 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 we create flexible schemas that allow for extension points where the schema on 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 specific local needs. Extension points can be implemented using a built-in extension mechanism via:

- Embedding foreign extension schemas (as supported by METS and PREMIS). They 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.

The structure allows posing more detailed requirements for metadata entities e.g. by:

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

For consistency purposes we reuse design principles between layers as much as possible.

---

[13] Although we do not describe specific content types in this deliverable, we still present geodata as an example of more specific content-oriented SFSB records type.

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

3 Metadata and Mapping

Metadata for a SIP can be obtained in several ways that are not mutually exclusive: automatically from 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 best or exclusively obtained.

The balance of manual vs automated creation of metadata, as well as the origin of metadata (producer vs archive) varies to a great extent due to different local best-practices and legal environments. However, in most cases the full SIP metadata is a mixture of manual, automated, archive and producer created metadata.

In the case of an ERMS export the SMURF profile 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, since
    - Most existing production systems are not MCRS[14] 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 need to support the full functionality for a MCRS, records creation and corresponding workflows, support for original technical access restriction management, managing retention periods.
- There are entities and metadata elements that are needed for archive export but not in the MoReq export, since
    - 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.

Therefore we use the MoReq export schema[15] as an inspiration for an archive transfer service – we are not adopting the MoReq export schema but adapting it. We have chosen parts of the schema for a partial extraction and enrich it with preservation metadata and other features.

The work group has agreed upon the following principles regarding metadata standards in the E-ARK project:

- We allow for use of multiple metadata standards in a single SIP.
- We allow for the duplication of metadata in different standards if it eases information management or if it keeps related metadata elements together (e.g. all metadata about computer files should be recorded by using only one standard), but we keep duplication to a minimum.
- We use EAD standard for descriptive metadata in the IP. Search in the archive can be based on these elements (plus additional ones in the catalogue).

---

[14] MoReq Compliant Records System (MCRS)
[15] XML Export Schema, http://www.moreq.info/index.php/specification

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

- Archival metadata regarding classification, aggregation, disposal schedules and disposal holds will be represented using EAD or MoReq2010.[16]
- We use EAD (mapped from MoReq2010) for descriptive metadata in the SIP for ERMS records.
- The metadata for a single record should be presented in EAD or MoReq but not both in parallel.
- We use EAD for descriptive metadata in the SIP for file system-based records.
- We map MoReq2010 aggregations and classes into EAD for AIP aggregations that are used for search.
- Different archive solutions can expose different subsets of the SIP metadata elements for search in the catalogue (i.e. the SIP format does not care/prescribe which metadata elements are going to be exposed for searching).
- All SIP metadata is to be treated as metadata rather than as content so that it can be made accessible to search in the archive. This means that it is submitted in the "metadata" folder rather than the "content" folder (i.e. we define only binary files as data, everything else is metadata).

For better understandability we categorise the SMURF metadata into 4 groups:[17]

- **Descriptive metadata**
  Implemented in EAD, but MoReq2010 may be required for some specific cases for information originating from the ERMS.
  Implemented only in EAD for information that is SFSB.
- **Structural**
  Implemented in EAD or PREMIS.
- **Administrative / Preservation**
  Implemented in EAD or PREMIS.
- **Uncategorised metadata**
  This category includes metadata support for customising and expanding the E-ARK SMURF SIP profile for local needs. Can be implemented using extension points of the SMURF profile.

NB! It is important to note that not all the metadata is a part of the current specification. Only the use of the metadata related to SMURF profile will be explained in the following metadata sections, but

---

[16] It may be required to keep them in parallel in MoReq and EAD. As an example to keep track of the original (in MoReq) and the archival (in EAD).

[17] There are 3 main types of metadata according to NISO (National Information Standards Organization):

- Descriptive metadata describes a resource for purposes such as discovery and identification. It can include elements such as title, abstract, author, and keywords.
- Structural metadata indicates how compound objects are put together, for example, how pages are ordered to form chapters.
- Administrative metadata provides information to help manage a resource, such as when and how it was created, file type and other technical information, and who can access it. There are several subsets of administrative data; two that are sometimes listed as separate metadata types are:
  - o Rights management metadata, which deals with intellectual property rights,
  - o Preservation metadata, which contains information needed to archive and preserve a resource.

Understanding Metadata, 2004, http://www.niso.org/publications/press/UnderstandingMetadata.pdf

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

some of the metadata is a part of the general SIP specification and the Common Specification for IPs (i.e. METS metadata) as explained on page 8.

3.1   Moreq2010 Entities and Metadata for the Profile

According to MoReq2010 each core service manages entities belonging to a specified number of entity types[18] and each entity must belong to one, and only one, of the entity types (Figure 2).
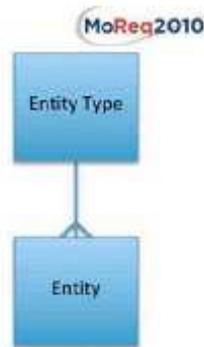
<div align="center">**Figure 2: Entity Types**</div>

The MoReq2010 core services refer to the following entity types:

- **Aggregations[19]**

  What is the entity?
  Aggregation is a core entity in MoReq 2010 and in all E-ARK partner implementations. MoReq 2010 does not distinguish between Fonds, Sub-fonds, Series, File and Sub-files. They are all Aggregations at various, specifiable levels. They all can be mapped to the MoReq 2010 Aggregation Entity.

  How is it dealt with in the SMURF?
  Different E-ARK partners use various combinations and patterns of these Entity types. 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 within E-ARK, but can be freely chosen by the users. It should be controlled locally in the organisation.[20] We recommend using ISAD-G in the archive. See section 3.2 for more details.

  The metadata for this entity is presented in Table 2.

- **Class[21]**

  What is the entity?
  Class is a core entity in MoReq 2010 and in all E-ARK partner implementations.

  How is it dealt with in the SMURF?

---

[18] 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 [MoReq 2010, Glossary].

[19] Aggregation of individual records or higher level aggregation of aggregations of records [page 249, MoReq 2010 Specification].

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

[21] Business classification applied to records and aggregations of records [page 250, MoReq 2010 Specification].

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

The vocabulary for titles of the Class entity is not controlled within E-ARK, but can be freely chosen by chosen by the users. More information about mapping MoReq 2010 classes to archive classes can be can be found in section 3.2.2 and about metadata for classes in

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

Table 3.

- **Component[22]**

  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.
  The association between a record and its content is provided by component entities. Each record can have one or more components. Each component is a reference to a single item of content [MoReq 2010, 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 it dealt with in the SMURF?
  The metadata for this entity type is presented in appendices as a part of a record. It is also important to note that each component must belong to only one record (Figure 3) as stated in MoReq2010.
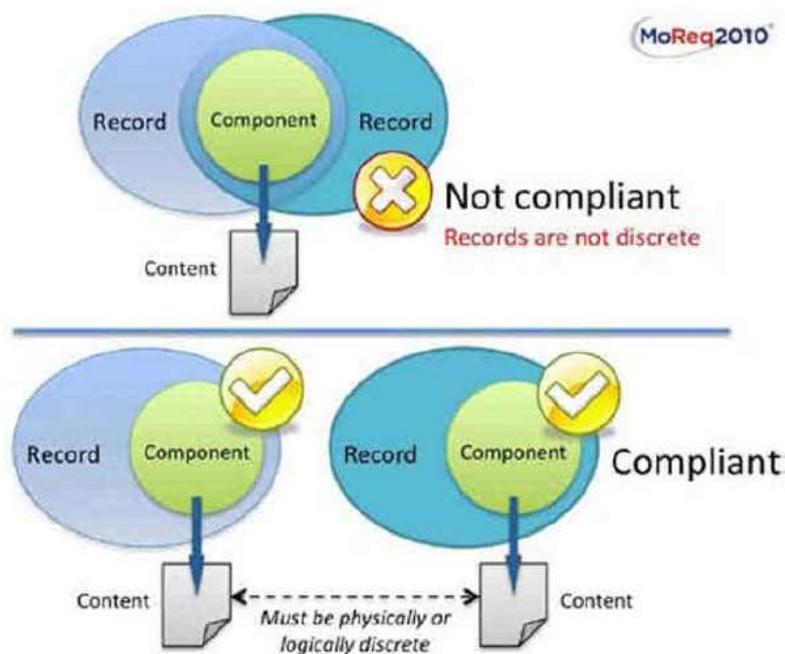
Figure 3: Components

- **Contextual metadata and system metadata definitions[23]**

  What is the entity?
  MoReq2010 allows for specialised subtypes of each entity type and divides metadata element definitions into:
  - o system metadata element definitions.
  - o contextual metadata element definitions.

---

[22] Individual component of a record representing its content [page 251, MoReq 2010 Specification].
[23] Definition of the properties of a contextual metadata element [page 252, MoReq 2010 Specification].

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

How is it dealt with in the SMURF?

While specialised elements are not in the scope of SMURF they might still be very valid for archives. Thus – it's something every individual archive needs to look at and is possibly described in the Submission Agreement.

- **Disposal holds**[24]

  What is the entity?

  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 is it dealt with in the SMURF?

  Therefore, as disposal holds don't possess the same importance for archives then these will only be archived as metadata and their functionality is not supported in the archive.

- **Disposal schedules**[25]

  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 is it dealt with in the SMURF?

  Therefore, as disposal schedules don't possess the same importance for the archive's records preservation activities they will only be archived as metadata and their functionality is not supported in the archive.

- **Events**[26]

  What is the entity?

  Events are not independent entities. All entities, except access control lists and events, will have an associated event history consisting of a series of events.

  How is it dealt with in the SMURF?

  To simplify the MoReq2010 model and make it easier to understand in archives we have decided to store events with some other entity instead of as a separate entity.

- **Function definitions**[27]

  What is the entity?

  Function definition types 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.

---

[24] 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 [MoReq 2010, Glossary].

[25] Disposal schedules are used to manage the life cycles of records in all MCRS solutions [MoReq 2010, Glossary].

[26] Description of the outcome of a function that was performed previously and is retained to show the history of an entity [page 256, MoReq 2010 Specification].

[27] Definition of function that can be performed with an entity by a user [page 257, MoReq 2010 Specification].

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

How is it dealt with in the SMURF?

Functions are described as part of events with some other entity (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,[28] Roles[29] and Users[30]**

How is the entity?

MoReq 2010 allows for specifying individual users as well as roles that participated in events. Different use cases require keeping different kinds of information. For some use cases event information in the producer system does not matter at all. 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 (deposit, legal scenarios).

How is it dealt with in the SMURF?

Groups should be exported by MoReq as individual Users – the Group Entity type as a functional entity is not supported in the archive.

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.

- **Record[31]**

What is the entity?

Record is a core entity in MoReq 2010 and in all E-ARK partner implementations.

How is it dealt with in the SMURF?

The metadata for this entity is presented in Table 1.

- **Service[32]**

What is the entity?

There is an export service that specifies the export from a MoReq compliant records system to another MCRS.

How is it dealt with in the SMURF?

We use the export service as a basis for defining the exchange format from ERMS to archives.

---

[28] An entity type that usually represents a team or business unit within the organisation and has various user entities as members [MoReq 2010, Glossary].

[29] 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. [MoReq 2010, Glossary].

[30] A person or system with an account that enables access to and use of an MCRS. A user does not have to be a human and could be another business system. Users must be authenticated before they can use an MCRS [MoReq 2010, Glossary].

[31] Record of a business transaction made up of one or more components that are managed atomically. [page 260, MoReq 2010 Specification].

[32] 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 [MoReq 2010, Glossary].

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

- **Template**

  <u>What is the entity?</u>
  According to MoReq 2010 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.[33]

  <u>How is it dealt with in the SMURF?</u>
  The SMURF profile does not use the MoReq 2010 metadata templates as the metadata will be organised by other means.

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 SIP and AIP 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 4), it is not mandatory (and sometimes unnecessary) do to so (Figure 5).
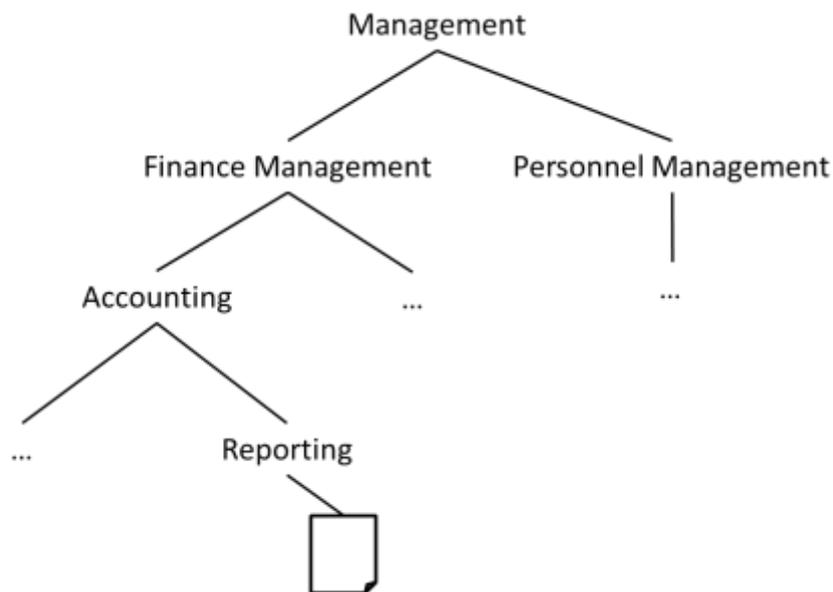
Figure 4: Hierarchical classes

---

[33] Page 263, MoReq 2010 Specification.

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation
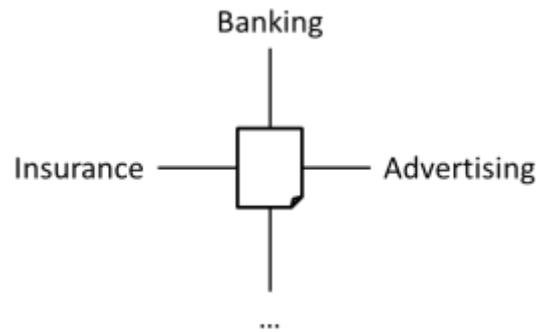
Figure 5: Non-hierachical classes

An *aggregation* is any accumulation of record entities at a level above the record object (e.g. folder, series, fonds, cf. Figure 6[34]).
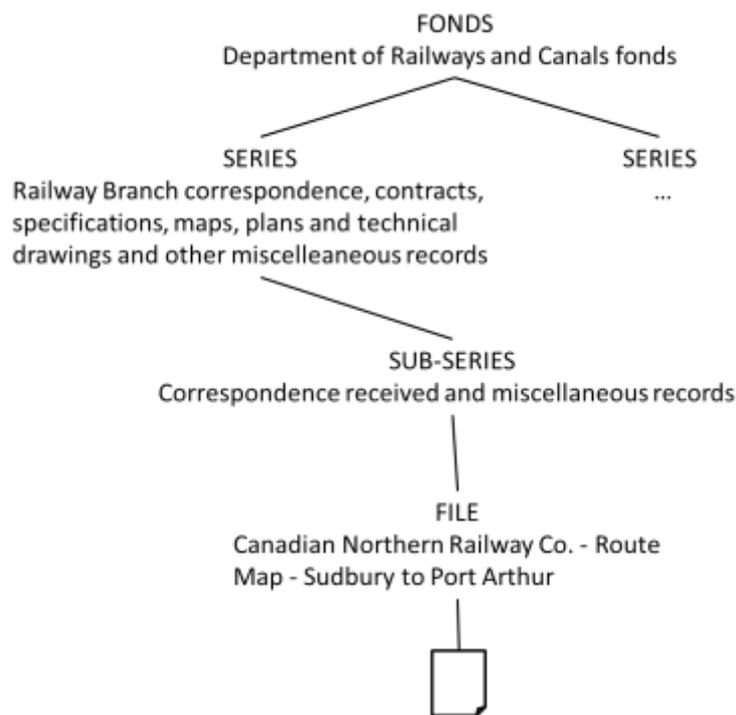
Figure 6: Aggregation

It is important to keep in mind that classification is not a way of structuring records but is a way of categorising records for management and access purposes. Aggregation on the other hand is explicitly 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 necessary 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.

---

[34] The example is based on ISAD(G) General International Standard Archival Description,
http://www.icacds.org.uk/eng/ISAD(G).pdf.

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

In chapter 3 above we specify that class hierarchies and aggregation structures are to be represented using the EAD metadata. This requires translating the MoReq 2010 metadata exported to the SIP from the producer system to the EAD profile used by the receiving archive.

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 this case we may not 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. Instead of specifying a catch-all solution we provide guidelines for the most critical issues and solve some select cases in our pilot implementations. This provides a good starting point to anyone wanting to implement a class and aggregation migration between two systems.

3.2.1 Mapping between ERMS and Archive Aggregations

When producer class and aggregation values are exported into the SIP they can be used by the archive in a number of ways:[35]

- Incorporated as raw values into the EAD record for the AIP
- Mapped and translated into the archive's EAD profile
- Incorporated into the archive's EAD profile by extending the profile

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

3.2.2 Mapping between ERMS and Archive Class Hierarchies

EAD has no direct equivalent to the class entity used in MoReq 2010. The only EAD property that is useful for recording MoReq class values drawn from (hierarchical) classification schemes is <fileplan> (see the example 1 below). In the case of translating MoReq2010 to EAD for use in SIPs and AIPs, the class hierarchy in the MoReq2010 could be reproduced under the <fileplan> property by using an EAD ordered list to record each element of the hierarchical value. However, as with aggregation values the E-ARK SIP profile does not require this translation to be carried out in any prescribed way.

Implementers are free to translate MoReq/ERMS class hierarchies as they see fit. It may even be the case that this metadata is not used in an archive's EAD implementation, although E-ARK does not recommend such a policy.

Example 1:[36]

<fileplan>
    <head>File List</head>

---

[35] The SMURF profile presents only one way (using EAD metadata) how MCRS aggregation values can be translated to archival aggregation values, but does not restrict of using any other method (i.e. adding all relevant class values as keywords to each individual record).

[36] Encoded Archival Description Tag Library, 2002, http://www.loc.gov/ead/tglib/elements/fileplan.html.

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

```
<note><p>The list below outlines the classification system used for the central files of Vice President
      Mondale's office. This structure assigned alpha-numeric codes to primary subjects and to secondary
      and tertiary subdivisions thereunder.</p>
</note>
<fileplan>
  <head>AGRICULTURE (AG)</head>
  <list type="ordered">
    <defitem>
      <label>1</label>
      <item>Home Economics</item>
    </defitem>
    …
  </list>
</fileplan>
<fileplan>
  …
</fileplan>
</fileplan>
```

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

3.3 SFSB Records Metadata

The following section will present the logic of incorporating simple file system based records into a SIP and the minimum metadata needed for the SFSB records from an archival perspective.

The SFSB records scenario can appear in many ways. For example, the producer has a shared drive with a set of computer files with (or without) archival descriptions. In order to include the computer files into the SIP we have to define an atomic intellectual entity. In this document the entity for the SFSB profile is a "record" as for ERMS profile and records can be composed of components as also described in MoReq2010. In the SFSB scenario the components are computer files. We will include them into the data directory as required by the E-ARK Draft Common Specification for Information Packages[37] and make sure that we group the computer files into separate representations if needed.[38] For example, we may have the same content of a record represented by two different file formats (DOC and PDF/A) and we may want to include both computer files into the SIP. In that case we have to place them into two separate representations as seen in Figure 7.
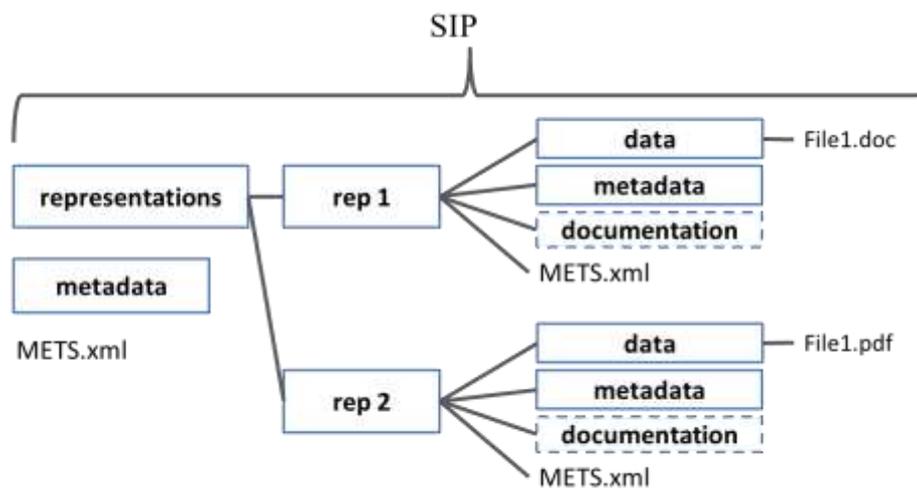
Figure 7: SFSB SIP

The metadata folder under representations will have to contain the archival descriptions if they are not already available on the root level. The metadata that applies to any SFSB record can be found in

---

[37] E-ARK Draft Common Specification for Information Packages in the E-ARK project, version 0.13, page 23, http://eark-project.com/resources/specificationdocs/50-draftcommonspec-1
[38] This step will be performed manually.

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

Appendix II Metadata for SFSB Records.

The <u>documentation</u> folder is not mandatory, but it allows of adding additional information about the content if needed.

<u>METS</u> files represent the metadata for a whole SIP package or for a single representation.

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

Geodata

A specific example and subcase of the SFSB scenario is the archiving of geodata.

Geodata is a combination of graphical representation of objects in space and their descriptions or attributes. Increasingly, geospatial formats include geospatially focused datasets or databases that contain primary information about a geographic location.

Geodata generally comes in two forms, vector or raster, and can be stored as a set of files or a database. Therefore, it is possible to handle geodata as a database (and apply the SIARD 2.0 specification) or as SFSB records. In this document we will explain the latter.

Following the E-ARK Draft Common Specification for Information Packages [39] at least one representation should be put in the representation directory (Figure 8).
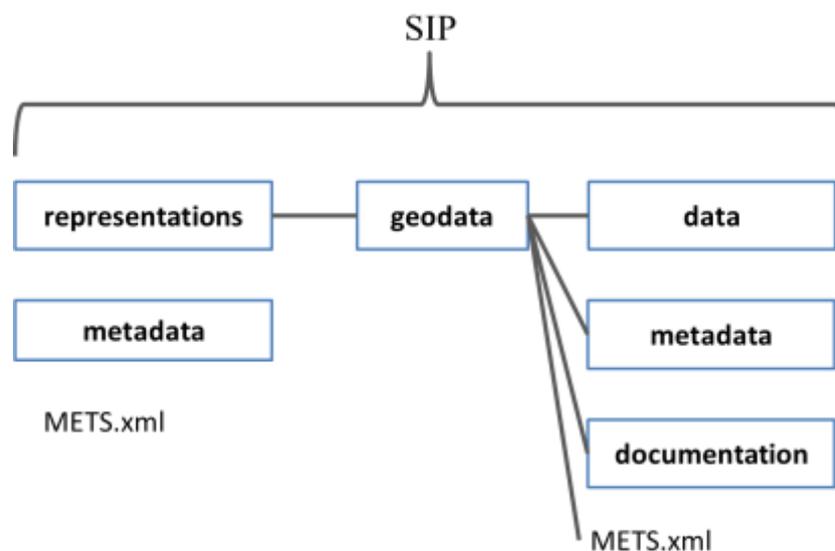
<div align="center">**Figure 8: Geodata SIP**</div>

In the current case, the representation folder will be named 'geodata' (use of any other name is also allowed) and the subfolders are as follows:

- **Data** – geodata in a long term preservation format (GML for vector and GeoTIFF for Raster) and all information that is needed to properly render the information:
  - **Graphical information.** Contains graphical information in vector or raster format. Vector data can be stored in different formats, such as SHP[40], KML[41], DXF[42], GML[43], etc.; the GML format as defined by the ISO19136:2007 standard was chosen as the long term preservation format; for raster geodata it is GeoTIFF with a GML bounding box.
  - **Attribute information.** Is already a part of the GML itself.

---

[39] E-ARK Draft Common Specification for Information Packages in the E-ARK project, version 0.13, page 23, http://eark-project.com/resources/specificationdocs/50-draftcommonspec-1
[40] http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf
[41] http://www.opengeospatial.org/standards/kml
[42] http://www.autodesk.com/techpubs/autocad/acad2000/dxf
[43] http://www.opengeospatial.org/standards/gml

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

- o **Georeferencing information – Coordinate Reference System (CRS) information.** CRS tells us how to locate objects in the geodata on earth's surface. Elements of spatial reference system are projection, geodetic datum, and unit of measure. All the elements can be defined by an EPSG code.[44]
- o **Visualization information.** Many GIS software applications enable export of the symbol definitions to a special file, however they may be in a proprietary XML structure (like the *.qml in QGIS, *.tab in MapInfo…) or even in a binary format (*.lyr in ArcGIS). The only OGC standard for symbols is the OGC Styled Layer Description XML format (sld files).

Metadata – the E-ARK project has identified some specific contextual metadata, based on the EC EC Directive INSPIRE[45] which have been mapped to EAD3 elements (

---

[44] http://www.epsg.org

[45] The Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE),, http://inspire.ec.europa.eu/index.cfm/pageid/3

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

- Appendix III Metadata for Specific Geodata Content Type) in order to enable automatic capture of descriptive metadata from the INSPIRE description of the spatial data sets.
- **Documentation** – additional information that does not fit into the metadata or data directory. Additional documentation helps us to understand the data-set in a wider social context and will provide a better understanding of the meaning, use and structure of the spatial data. The directory could contain:
  - o **Feature Catalogue.** Represents the logical structure of attributes, provides a better understanding of the meaning, use and structure of the spatial data. It provides a unified classification of spatial data in feature types (classes).
  - o **Visualisation and cartographic representation.** A collection of agreed cartographic symbols, which are used in visualization of spatial data sets to display objects in space. Cartographic symbols are shown in the legend, which explains their meaning.
  - o **The logical structure of layers and attribute definition.** Represents the organization of the data layers at the level of logical tables contained in the database or in a connection of unstructured objects and their attributes organised in a GIS application. This information is similar to information contained in the Feature Catalogue, but is described in a different manner.
  - o **Table relations.** Contains descriptions and diagrams of relations between tables in a database.
  - o **Common queries.** A list and a description of most common queries help to understand how the data set was used.
  - o **Other contextual documentation.** Other relevant documentation (e.g. user manuals) or links (e.g. scientific articles) to them.

The current specification does not aim to describe all details related to the geodata specifics as the geodata sub-profile will be developed further to support access scenarios described by WP5 "Archival Records Access Services" in E-ARK.

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

4 Feedback on MoReq2010 Adaptation

This chapter lists the main recommendations to the DLM Forum for MoReq2010 from an archiving perspective.

The recommendations are following:

- Full MoReq2010 metadata should only be used for the transfer from one MCRS to another and not for archiving purposes.
- The export schema (for an archiving transfer) should loosen up the cardinality requirements.
- The weakest cardinality requirement between the MoReq XSD and the archival transfer XSD should be adopted.

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

5   Glossary

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. [MoReq 2010, 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). [MoReq 2010, 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 metadata element definition	The contextual metadata element definition is the definition of a contextual 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	Encoded Archival Description. A non-proprietary de facto standard for the encoding of finding aids for use in a networked (online) environment. 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 standardization of collection information in finding aids within and across repositories.

http://www.loc.gov/ead/eadabout.html, accessed on 19 November 2015 |
| **EAC-CPF** | Encoded Archival Context - Corporate bodies, Persons, and Families (EAC-CPF). http://eac.staatsbibliothek-berlin.de/, accessed on 6 August 2015 |
| **Entity** | 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: |

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

	• system metadata, • contextual metadata, • access control list, • event history. The system metadata, and sometimes the contextual metadata, link the entity to other entities, forming relationships. [MoReq 2010, v 1.1]
ERMS	Electronic Records Management System
IP	Information Package
MCRS	MoReq Compliant Records System
MoReq2010	MoReq 2010; Modular Requirements for Records Systems. http://moreq2010.eu/pdf/moreq2010_vol1_v1_1_en.pdf
Record	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. • 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. [MoReq 2010, v 1.1]
SFSB	Simple file-system based records: records that contain simple file-system based folders or files, including those originating from content and data management systems, such as SharePoint, that are not based on true file systems. They address the submission of computer files or folders from the file producers rather than from an ERMS. They require manual enrichment with additional descriptive metadata
SIP	Submission Information Package
SMURF	**S**emantically **m**arked **u**p **r**ecords **f**ormats SMURF = ERMS (MoReq-based records) + SFSB (simple file-system based records)

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

6 Appendicies

6.1 Appendix I Metadata for MoReq2010 based ERMS Records

The following tables contain elements from EAD3 and MoReq2010 standards. Further information

- on EAD elements can be found in TagLibrary-VersionEAD3.pdf available at http://www2.archivists.org/sites/all/files/TagLibrary-VersionEAD3.pdf
- on MoReq2010 elements, in the MoReq2010 specification available at http://www.moreq.info/index.php/specification.

If no "MoReq2010 ID" is marked in the tables below then the element is recommended by the E-ARK project as an addition to the ERMS metadata.

If an "EAD element" is marked with "/*" then no exact match with the EAD standard has been identified, but the element name in the first column can be considered to be used.

NB! Presenting an XML Schema for the extended EAD metadata is not a part of this deliverable. The schema will be developed before the E-ARK pilots begin and presented in the final SIP specification which will be delivered in 2017.

Table 1: Extended EAD metadata for Records

Element name	EAD element (with path if necessary)	Attributes	Description and usage	Proposed cardinality	MoReq2010 ID
Created	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="created"	System set date and time when the entity was created	1..1	M14.4.9
Originated	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="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	M14.4.61
Title	archdesc/did/unittitle		The identifying name or title of the entity	1..1	M14.4.104
Description	archdesc/did/abstract archdesc/scopecontent		Description of the entity	0..1	M14.4.16
ParentAggregationId	archdesc/did/note@type="aggregation_id"/p/ref		Parent aggregation for a child aggregation or record	0..1	M14.4.63
Aggregated	/*		System set date and time when the child aggregation or record was created in, or added to, its current aggregation	0..1	M14.4.1

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

ClassId	archdesc/fileplan/p/ref		The classification of the entity, used by child aggregations and records to override the class they inherit from their parent aggregation	0..1	M14.4.4
DisposalScheduleId	archdesc/appraisal/*			0..1	M14.4.22
RetentionStart	archdesc/appraisal/*		System generated date calculated from the record's disposal schedule indicating the start of its retention period	0..1	M14.4.93
DisposalAction	archdesc/appraisal/*		Code describing the action to be taken on disposal of the record	0..1	M14.4.18
DisposalActionDue	archdesc/appraisal/*		The calculated date that the record is due for disposal	0..1	M14.4.19
DisposalConfirmationDue	archdesc/appraisal/*		The calculated date by which confirmation of carrying out the disposal action is due	0..1	M14.4.20
DisposalOverdueAlert	archdesc/appraisal/*		System set date and time when an alert was sent because the record was overdue for disposal	0..1	M14.4.21
LastReviewedComment	archdesc/appraisal/*		Comment made by the user that last reviewed the record explaining the disposal decision made by that review	0..1	M14.4.49
LastReviewed	archdesc/appraisal/*		System set date and time of when the last review was completed	0..1	M14.4.50
Transferred	archdesc/acqinfo/chronlist/chronitem/datesingle archdesc/acqinfo/chronlist/chronitem/event		System set date and time indicating when the transfer of the record was confirmed	0..1	M14.4.106
Destroyed			System set date and time when the entity was destroyed	0..1	M14.4.17
Duplicates	archdesc/altformavail/p		Reference to another entity that has been created by duplicating the record, component or event, and is an exact copy up to the event of duplication, with an identical provenance	0..1	M14.4.23
TransformedRecordId	control/representation			-	
TransformationEntry	control/maintenancehistory/maintenanceevent			0..n	
SystemId	archdesc/did/unitid	localtype="system_id"	Universally unique identifier for an entity that is generated automatically by the system and stays with the entity forever	1..1	M14.4.100

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

ExtraId	archdesc/did/unitid	localtype, label	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	control/filedesc/notestmt/controlnote		Notes	0 .. n	
Subject	archdesc/did/abstract/subject/part	localtype, relator	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	archdesc/index/head="keywords" archdesc/index/indexentry/subject		Keywords	0 .. n	
GeographicalLocations	archdesc/index/head="location_index" archdesc/index/indexent ry/geogname		List of related geographical locations	0 .. n	
FindingAid	archdesc/otherfindaid		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	
PublicationStatus	control/publicationstatus	value	The current publishing status of the material	0 .. 1	
Publication	control/filedesc/publicationstmt/publisher control/filedesc/publicationstmt/address control/filedesc/publicationstmt/date		Publications that are about or are based on the use, study, or analysis of the unit of description	0 .. n	
Description Source	control/sources/source		References to publications and other materials used for description	0 .. n	
Creator	archdesc/did/origination/name archdesc/did/origination/corpname archdesc/did/origination/famname archdesc/did/origination/persname	label="creator"	An entity primarily for making the content of the resource; an entity primarily responsible for making the resource (comment: Examples of a Creator include a person, an organization, or a service)	0 .. n	
Owner	archdesc/controlaccess/name archdesc/controlaccess/persname	relator="Owner"	Responsible person or role	0 .. n	
Administrator	archdesc/controlaccess/name archdesc/controlaccess/persname	relator="administrator"	Administrator of the record	0 .. n	
Reader	archdesc/controlaccess/name archdesc/controlaccess/persname	relator="reader"	Everyone who should be able to read the contents of the record (in the source ERMS system)	0 .. n	

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

Sender	archdesc/controlaccess/name archdesc/controlaccess/persname archdesc/controlaccess/corpname archdesc/controlaccess/famname	relator="sender"	Sender of the record	0 .. n	
Editor	archdesc/controlaccess/name archdesc/controlaccess/persname	relator="editor"	Person(s) who could edit the record (including adding) in the source ERMS system	0 .. n	
Recipient	archdesc/controlaccess/name archdesc/controlaccess/persname archdesc/controlaccess/corpname archdesc/controlaccess/famname	relator="recipient"	Recipient of the record	0 .. n	
Other	archdesc/controlaccess/name archdesc/controlaccess/persname archdesc/controlaccess/corpname archdesc/controlaccess/famname	relator	Other persons/organizations related to the record	0 .. n	
RecordType	control/filedesc/notestmt/controlnote	localtype="record_type"	Type of the record	0 .. n	
RecordLevelName	archdesc	level="item" OR level="otherlevel" otherlevel="record"	Level in archival hierarchy	0 .. n	
RecordDirection	control/filedesc/notestmt/controlnote	localtype="record_direction"	Record direction (incoming/outgoing)	0 .. 1	
Language	archdesc/did/langmaterial/language archdesc/did/langmaterial/languageset/language		Language of archived textual content, Examples: "EST", "ENG", "GER"	0 .. n	
TitleLanguage	control/languagedeclaration/language			0 .. n	
ScriptingSystem	archdesc/did/langmaterial/language/script			0 .. n	
RecordStatus	control/filedesc/notestmt/controlnote	localtype="record_status"	Status of the record	0 .. n	
Decision	archdesc/odd/head="decision"		Decision about the record	0 .. n	
DecisionDate	archdesc/odd/head="decision" archdesc/odd/chronlist/chronitem	chronitem@localtype="decision_date"	Date of the decision about the record	0 .. n	
DecisionDeadline	archdesc/odd/head="decision" archdesc/odd/chronlist/chronitem	chronitem@localtype="decision_deadline"	Deadline of decision about the record	0 .. 1	
ArchivalHistory	archdesc/custodhist		Information on the history of the unit of description that is significant for its authenticity, integrity and interpretation	0 .. n	
ReceivedDate	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="received"	Date when the record was received	0 .. 1	

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

Recipient	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/event/name archdesc/odd/chronlist/chronitem/event/persname archdesc/odd/chronlist/chronitem/event/famname archdesc/odd/chronlist/chronitem/event/corpname	chronitem@localtype="dispatched" *name@relator="recipient"	Original recipient of the record	0 .. n	
DeliveryMode	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/event/[text]/*	chronitem@localtype="received"	Delivery mode	0 .. 1	
DeliveryPriority	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/event/[text]/*	chronitem@@localtype="received"	Delivery priority	0 .. 1	
DeliveryId	archdesc/did/unitid	localtype="delivery_id", label	Delivery id	0 .. 1	
OpeningDate	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="opening"	Date of opening of the record	0 .. 1	
OpeningPerson	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/event/name archdesc/odd/chronlist/chronitem/event/persname archdesc/odd/chronlist/chronitem/event/famname	chronitem@localtype="dispatched" *name@relator="opening_person"	Person who opened the record	0 .. 1	
FirstUsed	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="first_used"	Date of first use	0 .. 1	
LastAdditionDate	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="last_addition"	Date of last addition	0 .. 1	
ClassificationDate	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="classification"	Date of classification	0 .. 1	
Prepared	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="prepared"	Date of preparation	0 .. 1	
FinishedDate	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="finished"	Date when the record was finished	0 .. 1	
MainSignatureDate	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="main_signature"	Date of main signature	0 .. 1	
MainSigner	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/event/name archdesc/odd/chronlist/chronitem/event/persname	chronitem@localtype="main_signature" *name@relator="main_signer"	Responsible person who signed the record	0 .. 1	

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

	archdesc/odd/chronlist/chronitem/event/famname				
MainSignerRole	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/event/occupation	chronitem@localtype="main_signature" *name@relator="main_signer_role"	Main signing person role	0 .. n	
OtherSignatureDate	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="other_signature"	Date of other signature	0 .. n	
OtherSigner	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/event/name archdesc/odd/chronlist/chronitem/event/persname archdesc/odd/chronlist/chronitem/event/famname	chronitem@localtype="other_signature" *name@relator="other_signer"	Other person signing the record	0 .. n	
OtherSignerRole	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/event/occupation	chronitem@localtype="other_signature" *name@relator="other_signer_role"	Other signing person role	0 .. n	
Expedited	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="expedited"	Date and time when the record was expedited	0 .. 1	
DispatchDate	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="dispatched"	Date of dispatch of the record	0 .. 1	
Dispatcher	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/event/name archdesc/odd/chronlist/chronitem/event/persname archdesc/odd/chronlist/chronitem/event/famname	chronitem@localtype="dispatched" *name@relator="dispatcher"	Person responsible for dispatching the record	0 .. 1	
Addressee	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/event/name archdesc/odd/chronlist/chronitem/event/persname archdesc/odd/chronlist/chronitem/event/famname	chronitem@localtype="dispatched" *name@relator="addressee"	Original addressee of the record	0 .. n	
DispatchMode	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/event/[text]/*	localtype="dispatched"	Mode of dispatching of the record	0 .. n	

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

OwnershipStartDate	Archival ownership/*		Date and time since the current owner owns the record	0 .. 1	
MediumType	archdesc/did/physdescstructured/physfacet	physdescstructured@physdesc structuredtype ="materialtype"		0 .. N	
ExtentUnit	archdesc/did/physdescstructured/unittype	physdescstructured@physdesc structuredtype ="spaceoccupied"	The unit used to describe the extent of the record (e.g MB, pages, num of files/components)	0 .. n	
ExtentValue	archdesc/did/physdescstructured/quantity	physdescstructured@physdesc structuredtype ="spaceoccupied"	The extent of the record expressed in the unit in the previous field	0 .. n	
RelatedRecord	archdesc/relatedmaterial	localtype=<type of relation>	Related record and type of relation	0 .. n	
OwnTextElement	archdesc/odd	localtype=<Own Element Description>	Any additional metadata	0 .. n	
Notes	See above.		Original notes of the record	0 .. n	
DigitalAdditionalMetaData	/*		Additional metadata in digital signatures	0 .. n	

Table 2: Extended EAD metadata for Aggregation

Element name	EAD element (with path if necessary)	Attributes	Description and usage	Proposed cardinality	MoReq2010 ID
Created	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="created"	System set date and time when the entity was created	1..1	M14.4.9
Originated	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="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	M14.4.61
FirstUsed	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="first_used"	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	M14.4.32

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

LastAddition	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="last_addition"	System set date and time indicating when the most recent record or child aggregation was added to the parent aggregation	0..1	M14.4.48
ClassId	archdesc/fileplan/p/ref			0..n	M14.4.4
Title	archdesc/did/unittitle		The identifying name or title of the entity	1..1	M14.4.104
Description	archdesc/did/abstract archdesc/scopecontent		Description of the entity	0..1	M14.4.16
ScopeNotes	archdesc/scopecontent			0..1	M14.4.97
Closed	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="closed"	System set date and time when the aggregation was closed	0..1	M14.4.5
Destroyed	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="destroyed"	System set date and time when the entity was destroyed	0..1	M14.4.17
MaxLevelsOfAggregation	/*		The maximum number of levels of aggregation allowed below a root aggregation	0..1	M14.4.52
ParentAggregationId	/*		ID of the aggregation this aggregation belongs to	0..1	M14.4.63
SystemId	archdesc/did/unitid	localtype="system_id"	Universally unique identifier for an entity that is generated automatically by the system and stays with the entity forever	1..1	M14.4.100
ExtraId	archdesc/did/unitid	localtype, label	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	archdesc/did/note/p		Notes	0 .. n	
Subject	archdesc/did/abstract/subject/part	localtype, relator	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	archdesc/index/head="keywords" archdesc/index/indexentry/subject		Keywords	0 .. n	
GeographicalLocations	archdesc/index/head="location_index" archdesc/index/indexent ry/geogname		List of related geographical locations	0 .. n	
FindingAid	archdesc/otherfindaid		Information about any finding aids that the repository or records creator may have that provide information relating to the context and contents of	0 .. n	

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

				the unit of description.		
Publication	control/filedesc/publicationstmt/publisher control/filedesc/publicationstmt/address control/filedesc/publicationstmt/date			Publications that are about or are based on the use, study, or analysis of the unit of description	0 .. n	
Description Source	control/sources/source			References to publications and other materials used for description	0 .. n	
Creator	archdesc/did/origination/name archdesc/did/origination/corpname archdesc/did/origination/famname archdesc/did/origination/persname	label="creator"		An entity primarily for making the content or the resource; an entity primarily responsible for making the resource (comment: Examples of a Creator include a person, an organization, or a service)	0 .. n	
Owner	archdesc/controlaccess/name archdesc/controlaccess/persname	relator="Owner"		Responsible person or role	0 .. n	
Editor	archdesc/controlaccess/name archdesc/controlaccess/persname	relator="editor"		Person(s) who can edit the aggregation (including adding)	0 .. n	
Administrator	archdesc/controlaccess/name archdesc/controlaccess/persname	relator="administrator"		Administrator of the aggregation	0 .. n	
Reader	archdesc/controlaccess/name archdesc/controlaccess/persname	relator="reader"		Everyone who should be able to read the contents of the aggregation	0 .. n	
Counterpart	/*				0 .. n	
Other	archdesc/controlaccess/name archdesc/controlaccess/persname archdesc/controlaccess/corpname archdesc/controlaccess/famname	relator		Other persons/organizations related to the aggregation	0 .. n	
AggregationType	archdesc	level, otherlevel		Type of the aggregation (e.g. case file, subfile etc.)	0 .. n	
Language	archdesc/did/langmaterial/language archdesc/did/langmaterial/languageset/language			Language of archived textual content, Examples: "EST", "ENG", "GER"	0 .. n	
MovedRecords	control/maintenancehistory/maintenanceevent/*			Information about records that have been moved to other aggregations	0 .. n	
DeletedRecords	control/maintenancehistory/maintenanceevent/*			Explanation of the fact that the record has been deleted by the administrator or has been destroyed due to technical errors.	0 .. n	
TitleLanguage	control/languagedeclaration/language				0 .. n	
ScriptingSystem	archdesc/did/langmaterial/language/script			Examples: "Latin", "Cyrillic", "Hebrew"	0 .. n	

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

AggregationStatus	control/filedesc/notestmt/controlnote	localtype="aggregation_status"	Status of the aggregation	0 .. n	
Decision	archdesc/odd/head="decision"		Decision about the aggregation	0 .. n	
DecisionDate	archdesc/odd/head="decision" archdesc/odd/chronlist/chronitem	chronitem@localtype="decision_date"	Date of the decision about the aggregation	0 .. n	
DecisionDeadline	archdesc/odd/head="decision" archdesc/odd/chronlist/chronitem	chronitem@localtype="decision_deadline"	Deadline of decision about the aggregation	0 .. 1	
ArchivalHistory	archdesc/custodhist		Information on the history of the unit of description that is significant for its authenticity, integrity and interpretation	0 .. n	
ReceivedDate	archdesc/odd/head="aggregation_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="received"	Date and time when the aggregation (usually case) was received	0 .. 1	
ClassificationDate	archdesc/odd/head="aggregation_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="classification"	Date of classification	0 .. 1	
OwnershipStartDate	Archival ownership/*		Date when ownership started	0 .. 1	
Prepared	archdesc/odd/head="aggregation_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="preparedn"	Date and time when case was prepared.	0 .. 1	
Expedited	archdesc/odd/head="aggregation_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="expedited"	Date and time when the case was expedited	0 .. 1	
PhysicalLocation	archdesc/did/physloc		Physical or logical placement of the aggregation	0 .. n	
ExtentUnit	archdesc/did/physdescstructured/unittype	physdescstructured@physdescstructuredtype ="spaceoccupied"	The unit used to describe the extent of the aggregation (e.g MB)	0 .. n	
ExtentValue	archdesc/did/physdescstructured/quantity	physdescstructured@physdescstructuredtype ="spaceoccupied"	The extent of the aggregation expressed in the unit in the previous field	0 .. n	
SystemOfArrangement	archdesc/arrangement		Information on the internal structure of the aggregation, the order and/or the system of contents	0 .. n	
RelatedAggregation	archdesc/relatedmaterial	localtype=<type of relation>	ID of a related aggregation	0 .. n	
OwnTextElement	archdesc/odd	localtype=<Own Element Description>	Any additional metadata	0 .. n	
Security classification	archdesc/accessrestrict	localtype="security_classification"	Security classification	0 .. n	

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

ExplanatoryText	archdesc/accessrestrict		Textfield for describing the restriction	0 .. n	
Regulation	archdesc/legalstatus		Which regulation and paragraph is used	0 .. n	
RestrictionStartDate	archdesc/accessrestrict/chronlist/chronitem/daterange/fromdate	chronitem@localtype="restriction_period"	Start date of the restriction	0 .. n	
RestrictionEndDate	archdesc/accessrestrict/chronlist/chronitem/daterange/todate	chronitem@localtype="restriction_period"	End date of the restriction	0 .. n	
RestrictionDuration	/*		Duration of the restriction	0 .. n	
RestrictionType	archdesc/accessrestrict	localtype="restriction_type"	Type of the restriction	0 .. n	
AppraisalDate	archdesc/appraisal/chronlist/chronitem/datesingle archdesc/appraisal/chronlist/chronitem/daterange		Date of appraisal	0 .. n	
AppraisalDecision	archdesc/appraisal/chronlist/chronitem/event		Appraisal decision	0 .. n	
AppraisalTitle	archdesc/appraisal/head archdesc/appraisal/chronlist/chronitem/event		Appraisal title	0 .. n	
AppraisalDescription	archdesc/appraisal/p		Appraisal description	0 .. n	
IPPOwner	archdesc/userestrict/*		Copyright owner	0 .. n	
Reproduction conditions	archdesc/userestrict/*		IPR condition description	0 .. n	
IppDuration	archdesc/userestrict/*		Restriction duration	0 .. n	
IppType	archdesc/userestrict/*		Reference to IPP type according legislative act	0 .. n	
CopyrightNotice	archdesc/userestrict/*			0 .. n	
LoanDate	archdesc/odd/head="aggregation_history" archdesc/odd/chronlist/chronitem/daterange/fromdate	chronitem@localtype="loan"	Date of loan	0 .. n	
Borrower	archdesc/odd/head="aggregation_history" archdesc/odd/chronlist/chronitem/event/name archdesc/odd/chronlist/chronitem/event/persname archdesc/odd/chronlist/chronitem/event/famname archdesc/odd/chronlist/chronitem/event/corpname	chronitem@localtype="loan" *name@relator="borrower"	Borrower	0 .. n	
AuthorizingPerson	archdesc/odd/head="aggregation_history" archdesc/odd/chronlist/chronitem/event/name archdesc/odd/chronlist/chronitem/event/persname archdesc/odd/chronlist/chronitem/event/famname	chronitem@localtype="loan" *name@relator="authorising_person"	Person authorizing the loan	0 .. n	
Term	/*		Loan term	0 .. n	
TakeBackDate	archdesc/odd/head="aggregation_history" archdesc/custodhist/chronlist/chronitem/daterange/todate	chronitem@localtype="loan"	Date of take back	0 .. n	

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

ResponsiblePerson	archdesc/odd/head="aggregation_history" archdesc/odd/chronlist/chronitem/event/name archdesc/odd/chronlist/chronitem/event/persname archdesc/odd/chronlist/chronitem/event/famname	chronitem@localtype="loan" *name@relator="responsible_person"	Person responsible for taking back after loan	0 .. n	
ArchiverName	archdesc/odd/head="aggregation_history" archdesc/odd/chronlist/chronitem/event/name archdesc/odd/chronlist/chronitem/event/persname archdesc/odd/chronlist/chronitem/event/famname	chronitem@localtype="archived" *name@relator="archiver"	Person responsible for inhouse archiving	0 .. n	
ArchivingDate	archdesc/odd/head="aggregation_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="archived"	Date of inhouse archiving	0 .. n	
DisposalDate	archdesc/odd/head="aggregation_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="disposed_of"	Disposal date	0 .. n	
DisposingPerson	archdesc/odd/head="aggregation_history" archdesc/odd/chronlist/chronitem/event/name archdesc/odd/chronlist/chronitem/event/persname archdesc/odd/chronlist/chronitem/event/famname	chronitem@localtype="disposed_of" *name@relator="disposing_person"	Disposing person	0 .. n	
SupervisingPerson	archdesc/odd/head="aggregation_history" archdesc/odd/chronlist/chronitem/event/name archdesc/odd/chronlist/chronitem/event/persname archdesc/odd/chronlist/chronitem/event/famname	chronitem@localtype="disposed_of" *name@relator="supervising_person"	Person supervising the disposal	0 .. n	
TransferDate	archdesc/acqinfo/chronlist/chronitem/datesingle	chronitem@localtype="transfer"	Date of transfer to the archive	0 .. n	
Deliverer	archdesc/acqinfo/chronlist/chronitem/event/name archdesc/acqinfo/chronlist/chronitem/event/persname archdesc/acqinfo/chronlist/chronitem/event/famname	chronitem@localtype="transfer" *name@relator="delivering_person"	Person responsible for the delivery to the archive	0 .. n	
Recipient	archdesc/acqinfo/chronlist/chronitem/event/name archdesc/acqinfo/chronlist/chronitem/event/persname archdesc/acqinfo/chronlist/chronitem/event/famname	chronitem@localtype="transfer" *name@relator="receiving_person"	Person responsible for receipt in the archive	0 .. n	

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

Table 3: Extended EAD metadata for Class

Element name	EAD element (with path if necessary)	Attributes	Description and usage	Proposed cardinality	MoReq2010 ID
Created	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="created"	System set date and time when the entity was created	1..1	M14.4.9
Originated	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="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	M14.4.61
FirstUsed	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="first_used"	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	M14.4.32
Title	archdesc/did/unittitle		The identifying name or title of the entity	1..1	M14.4.104
Description	archdesc/did/abstract archdesc/scopecontent		Description of the entity	0..1	M14.4.16
ScopeNotes	archdesc/scopecontent		Guidance to authorised users indicating how best to apply a particular entity and stating any organisational policies or constraints on its use	0..1	M14.4.97
Destroyed	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="destroyed"	System set date and time when the entity was destroyed	0..1	M14.4.17
DefaultDisposalScheduleId	/*			0..1	M14.4.11
HierarchicalParentClassId			The parent class for a hierarchical class	0..1	M201.7.2
SystemId	archdesc/did/unitid	localtype="system_id"	Universally unique identifier for an entity that is generated automatically by the system and stays with the entity forever	1..1	M14.4.100
ExtraId	archdesc/did/unitid	localtype, label	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	control/filedesc/notestmt/controlnote		Notes	0 .. n	
Keywords	archdesc/index/head="keywords" archdesc/index/indexentry/subject		Keywords	0 .. n	
LevelOfDescription	archdesc@level="class"		Level in archival classification hierarchy	1 .. 1	
LastAdditionDate	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="last_adition"	Date when last record was added to this class	0 .. 1	

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

PeriodBegin	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="period_begin"	Start date of period covered by this class	0 .. 1	
PeriodEnd	archdesc/did/unitdate archdesc/did/unitdatestructured	datechar="period_end"	End date of period covered by this class	0 .. 1	
Prepared	archdesc/odd/head="record_history" archdesc/odd/chronlist/chronitem/datesingle	chronitem@localtype="prepared"	Date of preparation	0 .. 1	
OwnTextElement	archdesc/odd	localtype=<Own Element Description>	Any additional metadata	0 .. n	
Security classification	archdesc/accessrestrict	localtype="security_classification"	Security classification	0 .. 1	

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

6.2 Appendix II Metadata for SFSB Records

The following table is based on EAD3 standard. Further information on EAD elements can be found in TagLibrary-VersionEAD3.pdf available at
http://www2.archivists.org/sites/all/files/TagLibrary-VersionEAD3.pdf

Table 4: Metadata for SFSB records

Element name	EAD element	Description and usage	Proposed cardinality	ISAD(g) 2.ed
ID of the Unit	<unitid>	**<unitid>** may contain any alpha-numeric text string that serves as a unique reference point or control number for the described material, such as a lot number, an accession number, a classification number, or an entry number in a bibliography or catalog. <unitid> is primarily a logical designation, which sometimes indirectly provides location information, as in the case of a classification number.	1..1	ISAD(G) 3.1.1
Title of the Unit	<unittitle>	**<unittitle>** is for recording the title statement, either formal or supplied, of the described materials. The title statement may consist of a word or phrase. <unittitle> is used at both the highest unit or <archdesc> level (e.g., collection, record group, or fonds) and at all the subordinate <c> levels (e.g., subseries, files, items, or other intervening stages within a hierarchical description).	1..1	ISAD(G) 3.1.2
Date of the Unit, Structured Date of the Unit	<unitdate>, <unitdatestructured>	**<unitdate>** is for indicating the date or dates the described materials were created, issued, copyrighted, broadcast, etc. <unitdate> may be in the form of text or numbers, and may consist of a single date, a date range, or a combination of single dates and date ranges; **<unitdatestructured>** provides a machine-processable statement of the date or dates the materials described were created, issued, copyrighted, broadcast, etc. <unitdatestructured> must contain one of the following child elements: <datesingle>, <daterange>, or <dateset>. <unitdatestructured> may contain only one child, therefore <dateset> must be used in situations where complex date information needs to be conveyed and requires at least two child elements. A date set may combine two or more <datesingle> and <daterange> elements.	1..*	ISAD(G) 3.1.3
Scope and Content	<scopecontent>	<scopecontent> contains a narrative statement that summarizes the range and topical coverage of the materials. It provides the researcher with the information necessary to evaluate the potential relevance of the materials being described. <scopecontent> may include information about the form and arrangement of the materials; dates covered by the materials; significant organizations, individuals, events, places, and subjects represented in the materials; and functions and activities that generated the materials being described. It may also identify strengths of or gaps in the materials.	0..1	ISAD(G) 3.3.1
Accruals	<accruals>	Used to indicate anticipated additions to the material already held by the repository. May indicate quantity and frequency of the accruals. **<accruals>** may also be used to indicate no additions are expected.	0..*	ISAD(G) 3.3.3
Conditions Governing Access	<accessrestrict>;	Record in **<accessrestrict>** information about the availability of the described materials, whether due to the nature of the information in the materials being described, the physical condition of the materials, or the location of the materials. Examples include restrictions imposed by the donor, legal statute, repository, or other agency, as well as the need to make an appointment with repository staff. May also indicate that the materials are not restricted;	1..*	ISAD(G) 3.4.1

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

Conditions Governing Use	<userestrict>	Use **<userestrict>** for information about any limitations, regulations, or special procedures imposed by a repository, donor, legal statute, or other agency. These conditions may be related to reproduction, publication, or quotation of the described materials after access to the materials has been granted. <userestrict> may also be used to indicate the absence of restrictions, such as when intellectual property rights have been dedicated to the public.	1..*	ISAD(G) 3.4.2
Language of the Material	<langmaterial>	**<langmaterial>** records information about languages and scripts represented in the materials being described. <langmaterial> must contain one or more <language> or <languageset> elements, but cannot contain text.	0..*	ISAD(G) 3.4.3
Related Material	<relatedmaterial>	**<relatedmaterial>** is used to identify associated materials in the same repository or elsewhere. These materials may be related by sphere of activity, or subject matter.	0..*	ISAD(G) 3.5.3
Description of Subordinate Components	<dsc>	Use **<dsc>** to wrap subordinate components in the archival hierarchy of the materials being described. Although <dsc> may repeat, it is recommended to include only a single <dsc> element. Because it is a wrapper element and not an essential part of archival description, <dsc> may be deprecated in future versions of EAD. Avoiding multiple <dsc> elements within an EAD instance will make future migrations simpler.	0..*	/
Component (Unnumbered)	<c>	As a wrapper for a set of elements, **<c>** provides information about the content, context, and extent of a subordinate body of materials. It is always a child or descendant of <dsc> and often a child and/or parent of another <c>. Each <c> identifies a logical section, or level, of the described materials. The physical filing separations between components need not always coincide with the intellectual separations. For example, a <c> that designates dramatic works might end in the same box in which the next <c> begins with short stories. Also, not every <c> directly corresponds to a folder or other physical entity. Some <c> elements simply represent a logical point in a hierarchical description.	0..1	/
Digital Archival Object	<dao>	<dao> is a linking element that uses @href to connect to born digital records or digital representations of the described materials. Digital representations may include graphic images, audio or video clips, images of text pages, and electronic transcriptions of text. The objects can be selected examples, or digital surrogates of all the materials in a collection, fonds, or an individual file.	1..*	/
Descriptive Identification Note	<didnote>	**<didnote>** can encode textual notes within <did> that are not more appropriately encoded in the other available elements.	0..1	ISAD(G) 3.6.1
Other Descriptive Data	<odd>	**<odd>** may be useful in converting legacy finding aids to the EAD format, by designating as "other" information that does not easily map to a more specific element. <odd> may be used when information about the described materials does not correspond to another element's definition, when the information is heterogeneous enough to make a single classification difficult, and when shifting the information to permit more specific content designation would be too costly or burdensome.	0..*	ISAD(G) 3.6.1

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

6.3 Appendix III Metadata for Specific Geodata Content Type

The following table is based on EAD3 standard and INSPIRE Metadata Implementing Rules: Technical Guidelines, based on EN ISO 19115 and EN ISO 19119 version 1.3. Further information on EAD elements can be found in TagLibrary-VersionEAD3.pdf available at http://www2.archivists.org/sites/all/files/TagLibrary-VersionEAD3.pdf and INSPIRE Metadata Implementing Rules at http://inspire.jrc.ec.europa.eu/documents/Metadata/MD_IR_and_ISO_20131029.pdf.

Table 5: Metadata for geodata

Element name	EAD element	Description and usage	Proposed cardinality	ISAD(g) 2.ed	INSPIRE el. No.	INSPIRE el. Name	INSPIRE Explanation
Title of the Unit	\<unittitle>	**\<unittitle>** is for recording the title statement, either formal or supplied, of the described materials. The title statement may consist of a word or phrase. \<unittitle> is used at both the highest unit or \<archdesc> level (e.g., collection, record group, or fonds) and at all the subordinate \<c> levels (e.g., subseries, files, items, or other intervening stages within a hierarchical description).	1..1	ISAD(G) 3.1.2	2.1.1	Resource title	Name by which the cited resource is known.
Scope and Content	\<scopecontent>	\<scopecontent> contains a narrative statement that summarizes the range and topical coverage of the materials. It provides the researcher with the information necessary to evaluate the potential relevance of the materials being described. \<scopecontent> may include information about the form and arrangement of the materials; dates covered by the materials; significant organizations, individuals, events, places, and subjects represented in the materials; and functions and activities that generated the materials being described. It may also identify strengths of or gaps in the materials.	0..1	ISAD(G) 3.3.1	2.2.2	Resource abstract	Brief narrative summary of the content of the resource(s).
Other Record Identifier	\<otherrecordid>	**\<otherrecordid>** can be used to record an identifier that is an alternative to the mandatory identifier provided in \<recordid>. These might include identifiers from systems that were used to generate the EAD instance or that are no longer current but had some part in the history and maintenance of the EAD instance.	0..1	/	2.2.5	Unique resource identifier	Value uniquely identifying an object within a namespace.

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

Related Material	<relatedmaterial>	<relatedmaterial> is used to identify associated materials in the same repository or elsewhere. These materials may be related by sphere of activity, or subject matter.	0..*	ISAD(G) 3.5.3	2.2.6	Coupled resource	Provides information about the datasets that the service operates on.
Language of the Material	<langmaterial>	<langmaterial> records information about languages and scripts represented in the materials being described. <langmaterial> must contain one or more <language> or <languageset> elements, but cannot contain text.	0..*	ISAD(G) 3.4.3	2.2.7	Resource language	Language(s) used within the datasets.
Subject	<subject>	<subject> indicates a topic reflected in the described materials.	1..1	/	2.3.1	Topic category	Main theme(s) of the dataset.
Other Descriptive Data	<odd>		0..1	/	2.3.2	Spatial data service type	A service type name from a registry of services
Index Entry	<indexentry>	A wrapper element that pairs an index term with zero or more linking elements. Each <indexentry> must contain an access element, such as <corpname>, <persname>, <subject>, etc., or <namegrp> to handle multiple access elements. It may also contain <ref>, <ptr>, or <ptrgrp> to identify and/or provide a link to the relevant position in the finding aid. If desired, use controlled vocabulary terms to facilitate access to information within and across finding aid systems.	0..*	/	2.4.1	Keyword value	Commonly used word(s) or formalised word(s) or phrase(s) used to describe the subject.
Geographic Coordinates	<geographiccoordinates>	Use <geographiccoordinates> to express a set of geographic coordinates such as latitude, longitude, and altitude representing a point, line, or area on the surface of the earth.	0..*	/	2.5.1	Geographic bounding box	Western-most coordinate of the limit of the dataset extent, expressed in longitude in decimal degrees (positive east). Eastern-most coordinate of the limit of the dataset extent, expressed in longitude in decimal degrees (positive east) Northern-most coordinate of the limit of the dataset extent, expressed in latitude in decimal degrees

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

							(positive north) Southern-most coordinate of the limit of the dataset extent, expressed in latitude in decimal degrees (positive north).
Date of the Unit, Structured Date of the Unit	<unitdate>, <unitdatestructured>	**<unitdate>** is for indicating the date or dates the described materials were created, issued, copyrighted, broadcast, etc. <unitdate> may be in the form of text or numbers, and may consist of a single date, a date range, or a combination of single dates and date ranges; **<unitdatestructured>** provides a machine-processable statement of the date or dates the materials described were created, issued, copyrighted, broadcast, etc. <unitdatestructured> must contain one of the following child elements: <datesingle>, <daterange>, or <dateset>. <unitdatestructured> may contain only one child, therefore <dateset> must be used in situations where complex date information needs to be conveyed and requires at least two child elements. A date set may combine two or more <datesingle> and <daterange> elements.	1..*	ISAD(G) 3.1.3	2.6.1	Temporal extent	Time period covered by the content of the dataset.
Date of the Unit, Structured Date of the Unit	<unitdate>, <unitdatestructured>	**<unitdate>** is for indicating the date or dates the described materials were created, issued, copyrighted, broadcast, etc. <unitdate> may be in the form of text or numbers, and may consist of a single date, a date range, or a combination of single dates and date ranges; **<unitdatestructured>** provides a machine-processable statement of the date or dates the materials described were created, issued, copyrighted, broadcast, etc. <unitdatestructured> must contain one of the following child elements: <datesingle>, <daterange>, or <dateset>. <unitdatestructured> may contain only one child, therefore <dateset> must be used in situations where complex date information needs to be conveyed and requires at least two child elements.	0..1	ISAD(G) 3.1.3	2.6.2	Date of publication	Reference date for the cited resource – publication.

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

		A date set may combine two or more <datesingle> and <daterange> elements.					
Date of the Unit, Structured Date of the Unit	<unitdate>, <unitdatestructured>	**<unitdate>** is for indicating the date or dates the described materials were created, issued, copyrighted, broadcast, etc. <unitdate> may be in the form of text or numbers, and may consist of a single date, a date range, or a combination of single dates and date ranges; **<unitdatestructured>** provides a machine-processable statement of the date or dates the materials described were created, issued, copyrighted, broadcast, etc. <unitdatestructured> must contain one of the following child elements: <datesingle>, <daterange>, or <dataset>. <unitdatestructured> may contain only one child, therefore <dataset> must be used in situations where complex date information needs to be conveyed and requires at least two child elements. A date set may combine two or more <datesingle> and <daterange> elements.	0..1	ISAD(G) 3.1.3	2.6.3	Date of last revision	Reference date for the cited resource – revision.
Date of the Unit, Structured Date of the Unit	<unitdate>, <unitdatestructured>	**<unitdate>** is for indicating the date or dates the described materials were created, issued, copyrighted, broadcast, etc. <unitdate> may be in the form of text or numbers, and may consist of a single date, a date range, or a combination of single dates and date ranges; **<unitdatestructured>** provides a machine-processable statement of the date or dates the materials described were created, issued, copyrighted, broadcast, etc. <unitdatestructured> must contain one of the following child elements: <datesingle>, <daterange>, or <dataset>. <unitdatestructured> may contain only one child, therefore <dataset> must be used in situations where complex date information needs to be conveyed and requires at least two child elements. A date set may combine two or more <datesingle> and <daterange> elements.	0..1	ISAD(G) 3.1.3	2.6.4	Date of creation	Reference date for the cited resource – creation.

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

Other Descriptive Data	<odd>		0..1	/	2.7.2	Spatial resolution	• Equivalent scale: level of detail expressed as the scale denominator of a comparable hardcopy map or chart • Distance: ground sample distance
Conditions Governing Access	<accessrestrict>;	Record in **<accessrestrict>** information about the availability of the described materials, whether due to the nature of the information in the materials being described, the physical condition of the materials, or the location of the materials. Examples include restrictions imposed by the donor, legal statute, repository, or other agency, as well as the need to make an appointment with repository staff. May also indicate that the materials are not restricted;	1..1	ISAD(G) 3.4.1	2.9.1	Limitations on public access [and use]	Access constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on obtaining the resource. Limitations on public access: - Access constraints - Example: otherRestrictions (limitation not listed). - Other constraints - Example: No limitations - Classification - Example: unclassified
	<accessrestrict>;	Record in **<accessrestrict>** information about the availability of the described materials, whether due to the nature of the information in the materials being described, the physical condition of the materials, or the location of the materials. Examples include restrictions imposed by the donor, legal statute, repository, or other agency, as well as the need to make an appointment with repository staff. May also indicate that the materials are not restricted;	1..1	ISAD(G) 3.4.1	2.9.2	Conditions applying to access and use	Restrictions on the access and use of a resource or metadata.

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

Conditions Governing Use	<userestrict>	Use **<userestrict>** for information about any limitations, regulations, or special procedures imposed by a repository, donor, legal statute, or other agency. These conditions may be related to reproduction, publication, or quotation of the described materials after access to the materials has been granted. <userestrict> may also be used to indicate the absence of restrictions, such as when intellectual property rights have been dedicated to the public.	1..1	ISAD(G) 3.4.1; ISAD(G) 3.4.2; ISAD(G) 3.4.4	2.9.2	Conditions applying to access and use	Restrictions on the access and use of a resource or metadata.
ID of the Unit	<unitid>	**<unitid>** may contain any alpha-numeric text string that serves as a unique reference point or control number for the described material, such as a lot number, an accession number, a classification number, or an entry number in a bibliography or catalog. <unitid> is primarily a logical designation, which sometimes indirectly provides location information, as in the case of a classification number.	1..1	ISAD(G) 3.1.1			
Accruals	<accruals>	Used to indicate anticipated additions to the material already held by the repository. May indicate quantity and frequency of the accruals. **<accruals>** may also be used to indicate no additions are expected.	0..*	ISAD(G) 3.3.3			
Description of Subordinate Components	<dsc>	Use **<dsc>** to wrap subordinate components in the archival hierarchy of the materials being described. Although <dsc> may repeat, it is recommended to include only a single <dsc> element. Because it is a wrapper element and not an essential part of archival description, <dsc> may be deprecated in future versions of EAD. Avoiding multiple <dsc> elements within an EAD instance will make future migrations simpler.	0..*	/			

The E-ARK SMURF Profile

Project Acronym: E-ARK
Grant Agreement Number: 620998
Project Title: European Archive Records and Knowledge Preservation

Component (Unnumbered)	<c>	As a wrapper for a set of elements, **<c>** provides information about the content, context, and extent of a subordinate body of materials. It is always a child or descendant of <dsc> and often a child and/or parent of another <c>. Each <c> identifies a logical section, or level, of the described materials. The physical filing separations between components need not always coincide with the intellectual separations. For example, a <c> that designates dramatic works might end in the same box in which the next <c> begins with short stories. Also, not every <c> directly corresponds to a folder or other physical entity. Some <c> elements simply represent a logical point in a hierarchical description.	0..1	/			
Digital Archival Object	<dao>	<dao> is a linking element that uses @href to connect to born digital records or digital representations of the described materials. Digital representations may include graphic images, audio or video clips, images of text pages, and electronic transcriptions of text. The objects can be selected examples, or digital surrogates of all the materials in a collection, fonds, or an individual file.	1..*	/			
Descriptive Identification Note	<didnote>	**<didnote>** can encode textual notes within <did> that are not more appropriately encoded in the other available elements.	0..1	ISAD(G) 3.6.1			
Other Descriptive Data	<odd>	**<odd>** may be useful in converting legacy finding aids to the EAD format, by designating as "other" information that does not easily map to a more specific element. <odd> may be used when information about the described materials does not correspond to another element's definition, when the information is heterogeneous enough to make a single classification difficult, and when shifting the information to permit more specific content designation would be too costly or burdensome.	0..*	ISAD(G) 3.6.1			