ECP-2007-DILI-517009

EuropeanaLocal

D3.1 – Central parallel test execution environment: demonstrator for first review

Deliverable number	*D3.1*
Dissemination level	*PP; Restricted to other programme participants (including Commission services and project reviewers)*
Delivery date	*June 2009*
Status	*Final*
Author(s)	*Lizzy Komen, Robina Clayphan, Julie Verleyen*

*e***Content***plus*

---

[1] OJ L 79, 24.3.2005, p. 1.

Contents

1 Introduction

EuropeanaLocal has the task of making available digital content from local and regional institutions through Europeana. To make this process manageable this involves the establishment of a harvestable network of OAI-PMH compliant metadata repositories to aggregate this content. Within EuropeanaLocal a central, scalable test execution environment has been established to facilitate the validation of content for the Europeana Prototype service.

EuropeanaLocal will be the first project that will make use of the (content) test environment provided by EDL Foundation and it can therefore be seen as a real 'testbed' for any future test environments for Europeana related projects. In summer 2010 the operational service of Europeana will be released, with the aim of ten million digital items accessible online. EuropeanaLocal plays an important role in reaching this aim by providing a large amount of content. The EuropeanaLocal test environment will assist in the process of making content accessible through Europeana.

2 Scope

This report forms part of D3.1 and summarises the current state of the central test environment for EuropeanaLocal, to be specified, developed and implemented by EDL Foundation. The test environment complies with the requirements specified by Europeana.Net and responds to the results of the partner survey (D2.1), to aggregate and index metadata and/or content harvested from the network of EuropeanaLocal repositories. D3.1 is the actual demonstrator for the first review, while this report explains what the test environment encompasses.

3 EuropeanaLabs - a central test environment

The EuropeanaV1.0 project has the objective of setting up an Open Source Infrastructure called "EuropeanaLabs" as shown in Image 1.
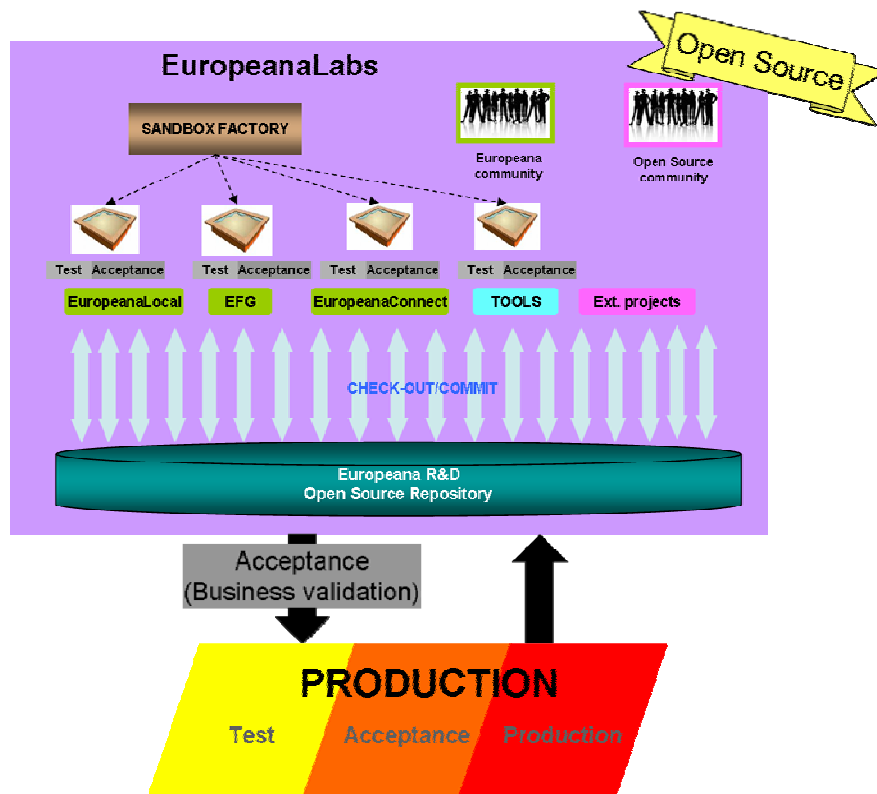
Image 1: EuropeanaLabs

EuropeanaLabs is the overall infrastructure where individual test areas can be established and allocated to participants in the development of Europeana. These test areas are called "Sandboxes".

A sandbox is a test and validation environment for both content and technologies. Therefore, the EDL Foundation distinguishes two types of sandboxes;

1. "Content Sandbox", the place where partners will harvest, validate and test ingestion of content with the assistance of tools provided by Europeana
2. "Application Sandbox", for validating and testing prototypes/technologies developed and proposed by partners for possible inclusion in Europeana operational service.

EuropeanaLabs is an Open Source infrastructure and will also store a repository of code and data and make this available for reuse by the Open Source community.

A Content Sandbox can be used to validate results (content or technology) proposed for inclusion in Europeana. It does not validate the tools that participants may use to locally create their input, for example the tools used to prepare any content contributed to Europeana. Sandboxes are not designed to be a a development platform: in EuropeanaLabs they are the places where content and technology can be tested and demonstrated for validation purposes before being brought into the production system.

4   EuropeanaLocal Sandbox

EuropeanLocal has been assigned a "Content Sandbox" and this will be fully installed towards the end of June 2009. The Sandbox is based on the Europeana (beta) Prototype that was launched in November 2008. The EuropeanLocal partners will be able to import harvested, normalised and mapped metadata into the Sandbox to validate it for inclusion in Europeana.

The Multi-regional training workshops organised in April and May 2009 by the technical partners in EuropeanaLocal, addressed the basic processes required to contribute local content via harvesting to Europeana. One of the main goals was to set up OAI repositories at the regional aggregation levels, necessary for the inclusion of metadata into the sandbox for testing and validation.

The Sandbox includes two components: a dashboard to manage the import of content and a copy of the Europeana Prototype that will run on this content. The demonstator of the EuropeanaLocal Content Sandbox will show what is available at the time of the first review. Work on the content sandbox is ongoing, while proper testing on e.g. usage and content ingestion will need to take place in the months following the first review.

4.1   Sandbox technologies

The Sandbox comprises of the following elements:

*   Linux based Operating System
*   Application server to deploy Java web programmes
*   Database Management System
*   The Europeana software system (Digital Library Mangement System)

4.2   Sandbox usage specifications

Content can be imported in XML format into the sandbox for testing and validation.

The XML Schema is based on Europeana Sementic Elements V3.1 (ESE).
http://www.version1.europeana.eu/c/document_library/get_file?uuid=5fdb5df9-bde5-42c4-8df5-7c9e488e7cf7&groupId=10128

4.3   Europeana Metadata & Normalisation Guidelines for the Europeana Prototype

The Europeana Semantic Elements V3.1 (ESE) is the metadata set developed for the prototype version of Europeana launched in November 2008. It is a Dublin Core-based application profile[1] providing a generic set of terms that can be applied to heterogeneous materials thereby providing a baseline to allow contributors to take advantage of their existing rich descriptions.

To provide metadata in the ESE format, it is necessary for EuropeanaLocal contributors to map elements from their own metadata format to ESE. In addition to the mapping it is necessary for a normalisation process to be carried out on some values to enable machine readability

The Europeana office has produced the "Metadata & Normalisation Guidelines for the Europeana Prototype" to assist content providers with the preparation of their data. The Guidelines document goes into more detail about the use of the elements and arranges them to support the decision-making process for mapping. The Guidelines forms Annexe A of this report.

4.4   Content ingestion workflow EuropeanaLocal to Europeana

Part of the work of EuropeanaLocal is to identify the best technical route for each partner to supply their metadata to Europeana. This must be seen in the context of two main constraints: the desire of Europeana to streamline its own organisational activity, and the technical and organisational situation with each partner. In discussion between Europeana and EuropeanaLocal, models of how this could be achieved were produced and partners were asked to consider these from the perspective of their own organisation. This section uses the slides produced for the Technical Workshops in April and May 2009 to explain the options.

4.4.1 The current aggregation model in Europeana

The current position of content provision to Europeana is very mixed – as shown in Image 2. The Europeana office deals with aggregators and also directly with individual content providers. As the number of providers expands the management overhead of maintaining this model will become prohibitive.
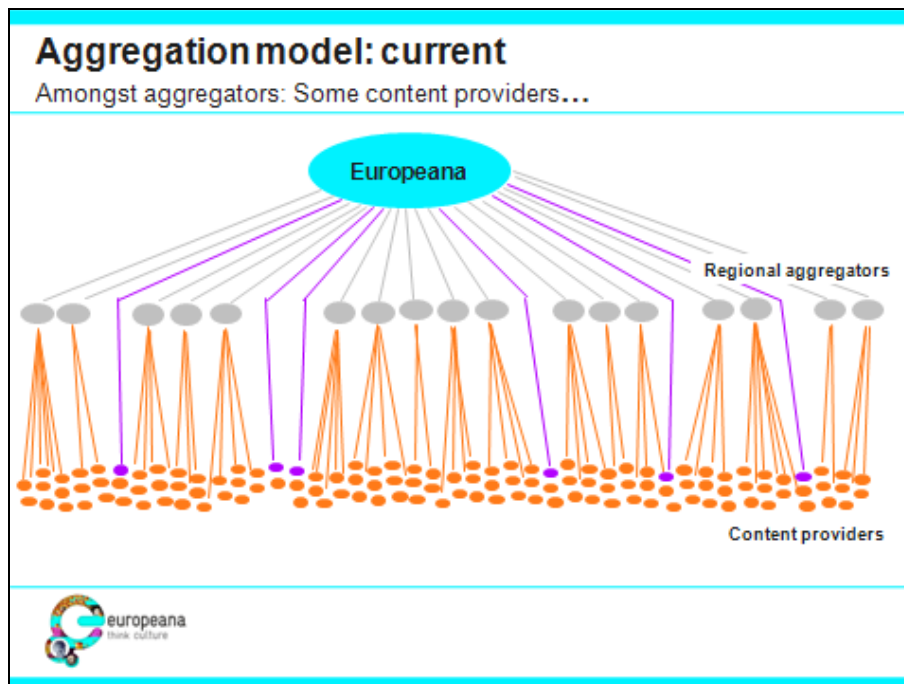
---

[1] http://dublincore.org/documents/

Image 2: Aggregation model: current

The preferred model for the future is for Europeana to deal only with aggregators unless there is a strategic necessity to deal directly with an individual provider. The desired position is shown in Image 3.
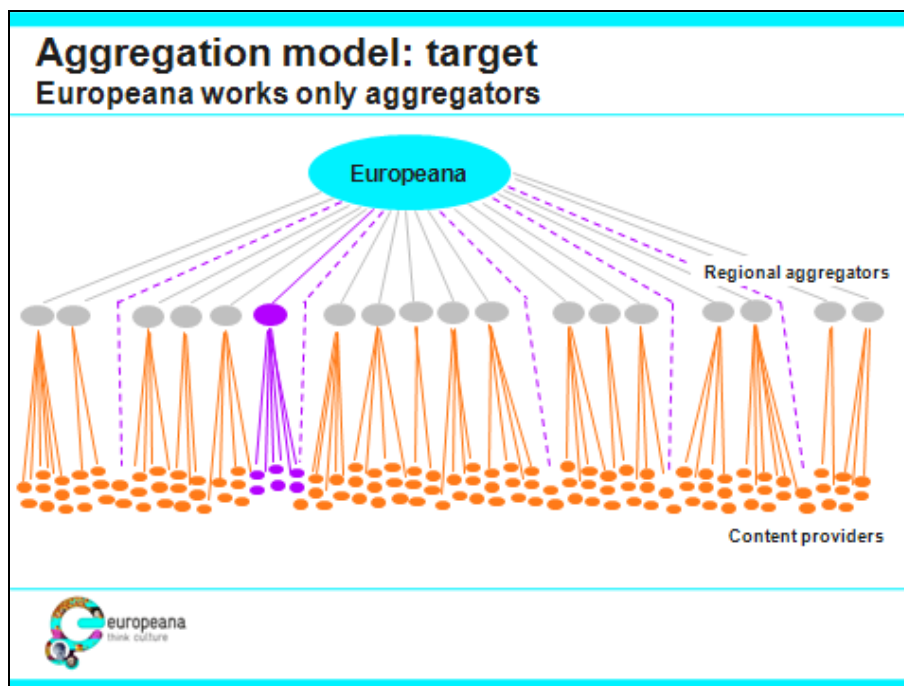

Image 3: Aggregation model: target

Deliverable *D2.3 Existing local/regional repositories and aggregations,* provides further information about the current status of aggregation and repositories of EuropeanaLocal partners.

4.4.2 Ingestion Workflows in EuropeanaLocal

As the first of the Europeana-related projects to try to supply data to Europeana since the initial launch, many aspects of how the process will work are still undefined. EuropeanaLocal is seeking to define the optimal workflows for a range of different types of provider. To do this, four activity models were devised that show alternative scenarios for **Who** would do **What** activity and **Where** it would take place. The diagrams not only show what happens in the Content Sandbox, but show all the processes required for content ingestion from the content provider right through to Europeana. The UML (Unified Modeling Language) diagrams represent:

- The sequence of activities through which the data flows from one point to another
- The different Actors who are responsible for a part of the activities of the flow
- The different System locations / Environments where activities take place

As shown in Image 4 the Diagrams are divided into four columns showing roles and locations. On the left is column 1, the content provider who holds the original digital objects and provides the source data. Column 2 is the EuropeanaLocal Regional Aggregator who collects data from the provider and delivers it to Europeana. The third column represents the Content Sandbox where testing and verification will take place. The fourth column, on the right, represents the Europeana Administrator who will ingest the content into the Europeana system.
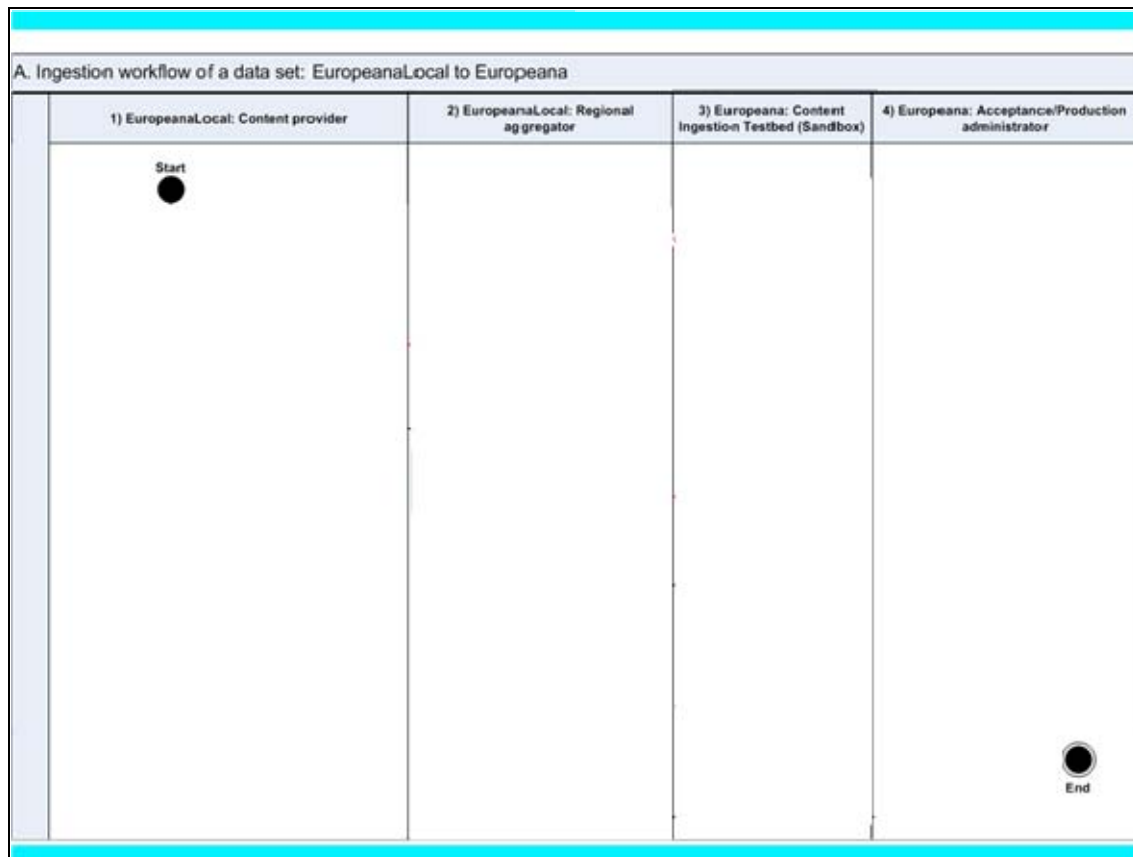

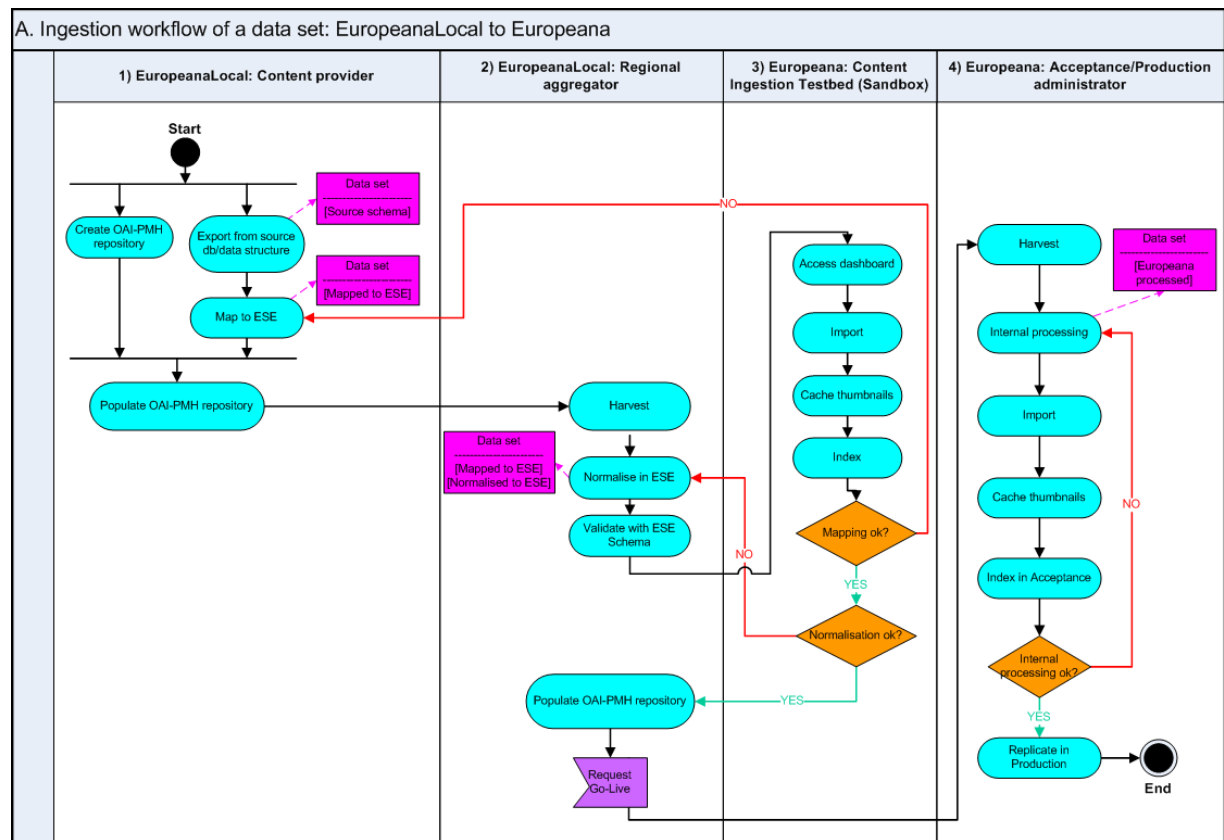**Image 4: Workflow for content ingestion**

Image 5: Model A

In Model A (Image 5) it is assumed that:

1. the content provider will install an OAI repository. They will extract the data from their local system, map and convert it to ESE and load it into their repository ready for harvesting.
2. The Regional aggregator will harvest the data, carry out the first normalisation activity and validate the data against the ESE schema before loading into the sandbox.
3. Once in the sandbox the thumbnails will be cached, the data will be indexed and the mapping and normalisation verified.
4. The data is then returned to the aggregator's OAI repository ready to be harvested by Europeana. Here it goes through internal Europeana processes and is integrated into the system.
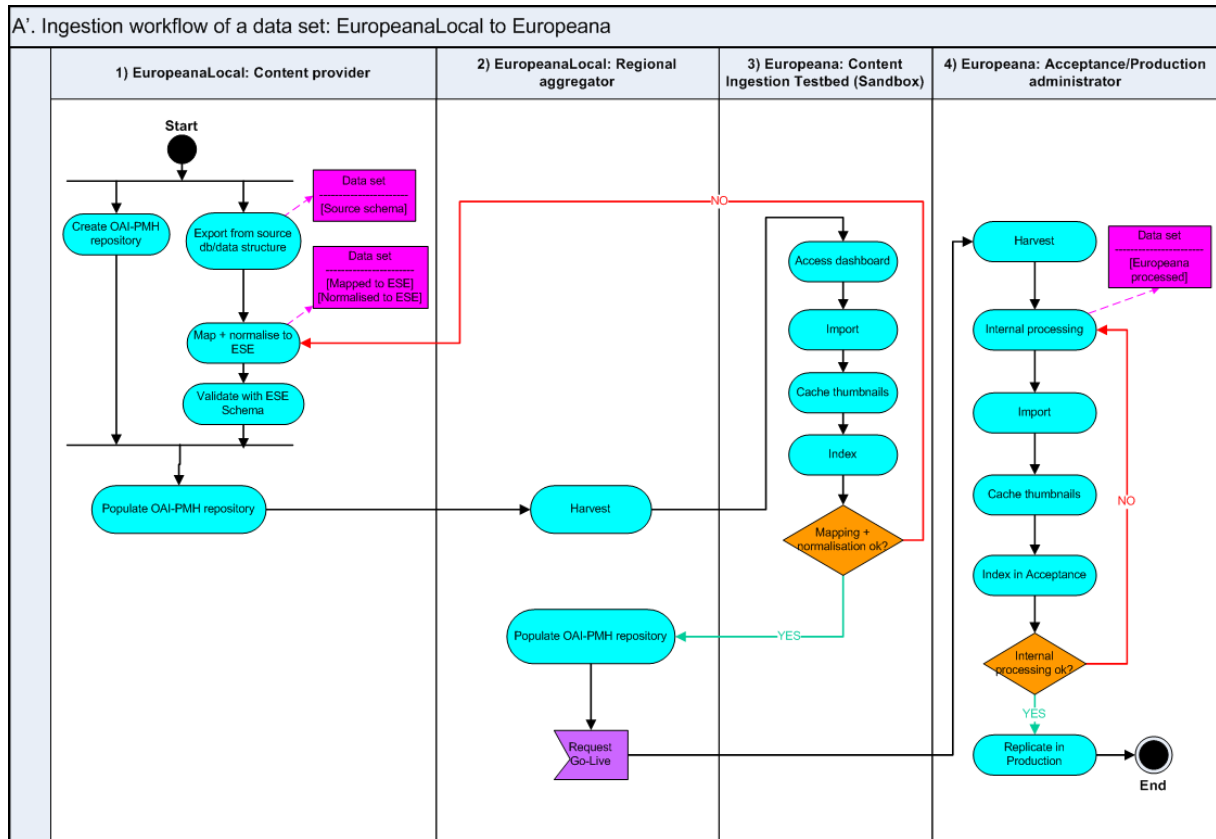
Image 6: Model A'

In Model A' (Image 6) the tasks of carying out the first normalisation and validation have passed back from the aggregator to the content provider.
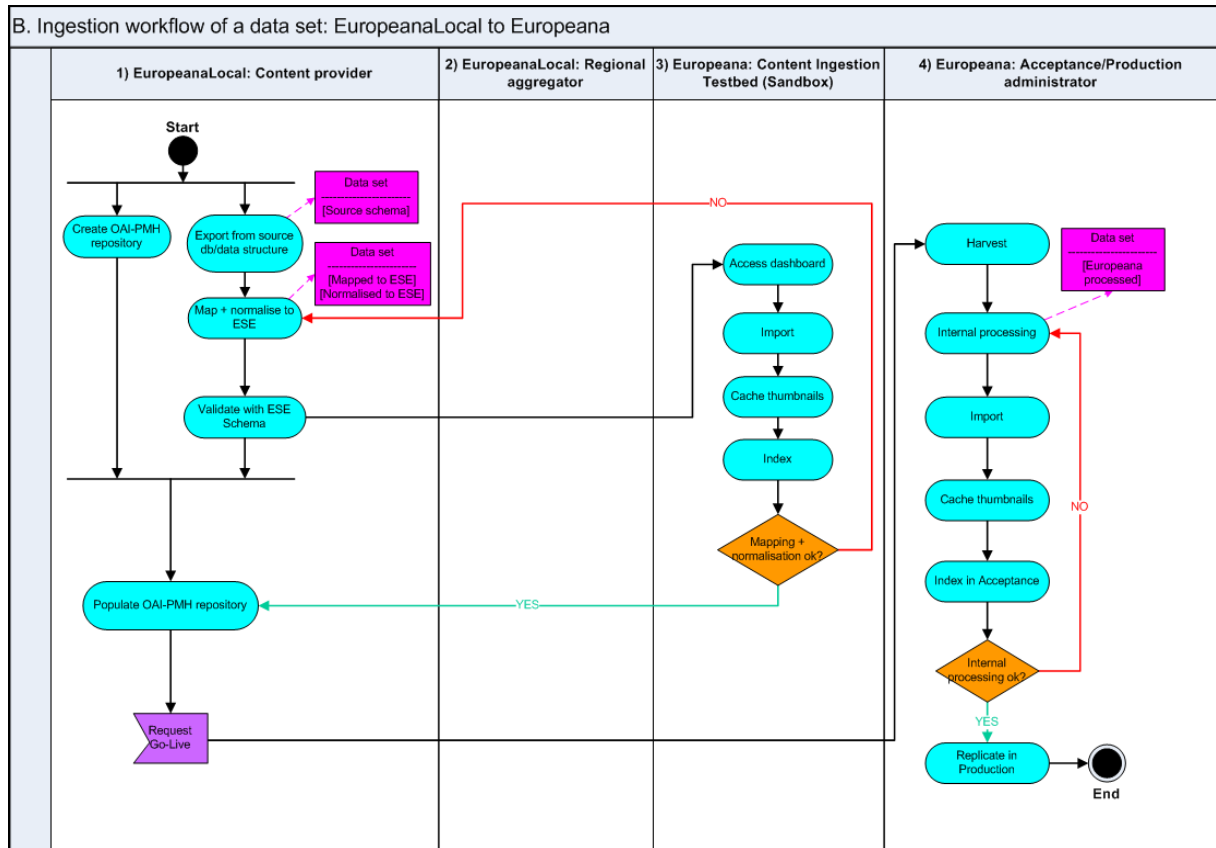
Image 7: Model B

In Model B (Image 7) the content provider is also the aggregator and can carry out all the roles that will enable the data to be harvested to Europeana.
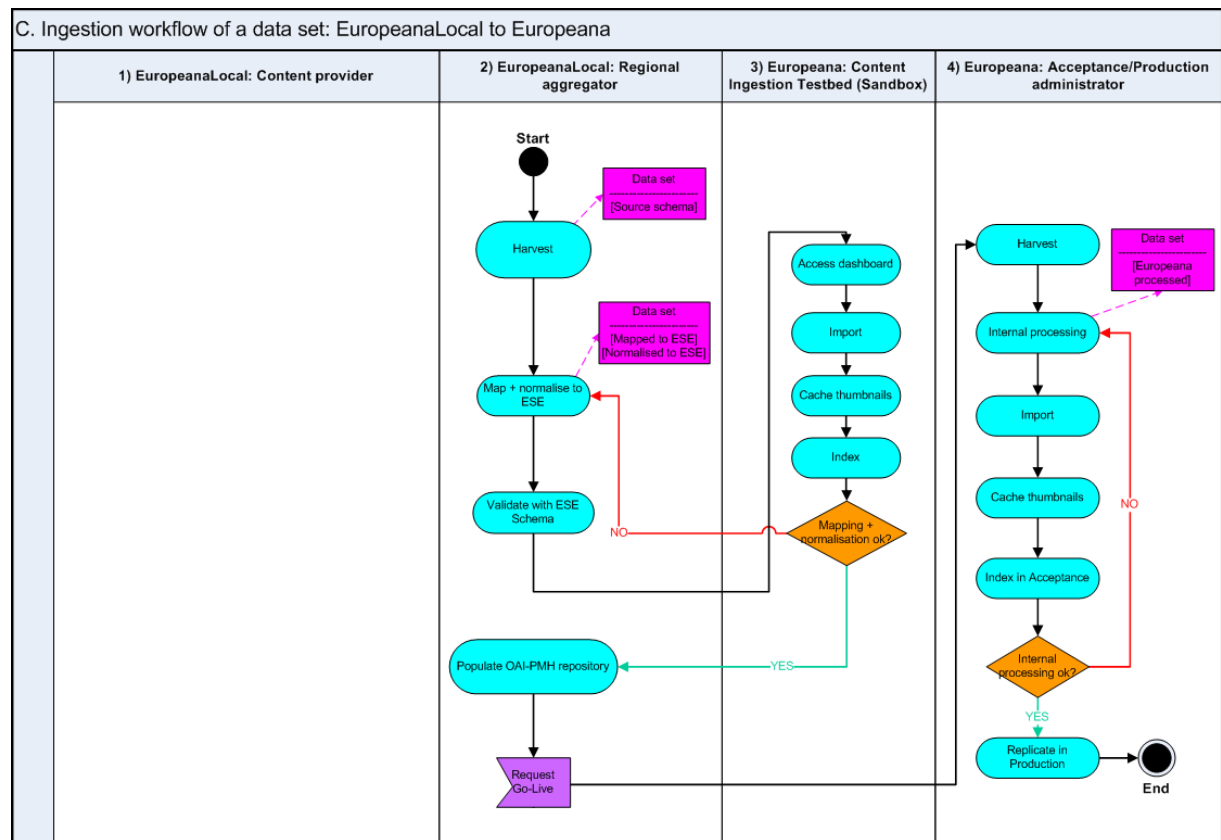
Image 8: Model C

In Model C (Image 8) the content provider has no role in the described workflow. The aggregator is probably an established aggregator taking data in a variety of local formats and transforming them on behalf of the content provider as part of its existing role. All the activity formerly in column 1 is therefore moved to column 2.

5   Conclusion

This report summarises the current state of the Content Sandbox for EuropeanaLocal, specified, developed and implemented by the EDL Foundation. The Content Sandbox is embedded in EuropeanaLabs; the overall Open Source infrastructure for Europeana where individual test areas for validation of content and technology are established.

While this Deliverable aims to portray the current state of the EuropeanaLocal Content Sandbox in terms of functionality, technologies and usage specifications, additional information is given about the Europeana Guidelines for Mapping & Normalisation (to assist providers to prepare their data) and about the different Content Ingestion Workflows depicting how content flows from the EuropeanaLocal content provider right through to Europeana.

D3.1 delivers a central test execution environment, with a demonstrator for the first review. The next steps would be to test the harvesting of EuropeanaLocal content, to identify any initial problems and requirements for adjustment in either the central test environment or local

repositories. This initial testing will take place in conjunction with several content co-ordinators from countries with existing or well-advanced OAI-PMH repositories. Ultimately, the harvested and approved EuropeanaLocal content will be indexed and available via the Europeana portal, disclosing the great diversity of cultural heritage held at local and regional level throughout Europe.

Annexe A- Mapping & Normalisation Guidelines for the Europeana Prototype

Version 1.0
28 May 2009

Contents

6 1. Introduction

Europeana provides integrated access to digital objects from the cultural heritage organisations of all the nations of the European Union. It encompasses material from museums, libraries, archives and audio-visual archives with the aim of making Europe's multicultural and multilingual riches discoverable together in a common on-line environment.

To do this Europeana harvests and idexes the descriptive metadata associated with the digital objects. As there is no one universal metadata standard applied across the participating domains, a set of metadata elements has been developed that will allow a common set of information to be supplied to support the functionality desired by the user and needed for the operation of the underlying system.

The Europeana Semantic Elements V3.1 (ESE)[1] is the metadata set developed for the prototype version of Europeana launched in November 2008. It is a Dublin Core-based application profile[2] providing a generic set of terms that can be applied to heterogeneous materials thereby providing a baseline to allow contributors to take advantage of their existing rich descriptions.

To provide metadata in the ESE format, it is necessary for contributors to map elements from their own metadata format to ESE. In addition to the mapping it is necessary for a normalisation process to be carried out on some values to enable machine readability. In the initial implementation of the Europeana prototype much of the mapping and normalisation was carried out centrally in the Europeana Office. This work will increasingly be passed to content providers or aggregators and these Guidelines are provided to support those tasks.

It is anticipated that further domain-specific examples will be provided as the ESE is applied by providers and aggregators.

An XML Schema has also been produced[3] as a further tool to assist providers in ensuring compliance with ESE. Some details of the schema are explained later in this document.

 ESE v3.1 is a sub-set of the metadata initially defined in the Europeana Metadata Requirements described in the EDLnet deliverable D2.5 "Europeana Outline Functional Specification"[4]

The mapping and normalising methodologies described in this document are specific to the Europeana service.

---

[1] http://dev.europeana.eu/public_documents/Specification_for_metadata_elements_in_the_Europeana_prototype.pdf

[2] http://dublincore.org/documents/

[3] http:// …..

[4] http://dev.europeana.eu/public_documents/EDLnet_D2.5_Outline_Functional_Specifications20080820_version_1.2_commentfree.pdf

7   2. Scope of this document

As a Dublin Core (DC) application profile the ESE incorporates both DC terms and some locally coined terms which have been added specifically to support functionality in Europeana. A full alphabetic declaration of these terms can be found in the ESE V3.1 Specification. This Guidelines document goes into more detail about the use of the elements and arranges them to support the decision-making process for mapping. They have been grouped according to the value they are considered to have in terms of searching with the addition of a separate group for the europeana elements which are used for specific functionality. There is also a discussion of various issues encountered in mapping heterogeneous materials and recommendations to resolve these.

When making mapping decisions, providers are also asked to consider how their data will perform in response to "who, what, where and when" queries.

7.1   2.1 List of elements

Strongly recommended	Recommended	Additional elements	Europeana elements
dc:title	dc:coverage	dc:format	europeana:country
dcterms:alternative	dcterms:spatial	dcterms:extent	europeana:hasObject
dc:creator	dcterms:temporal	dcterms:medium	europeana:isShownAt
dc:contributor	dc:description	dc:identifier	europeana:isShownBy
dc:date	dcterms:isPartOf	dc:rights	europeana:language
dcterms:created	dc:language	dcterms:provenance	europeana:object
dcterms:issued	dc:publisher	dc:relation	europeana:provider
	dc:source	dcterms:conformsto	europeana:type
	dc:subject	dcterms:hasFormat	europeana:unstored
	dc:type	dcterms:isFormatOf	europeana:uri
		dcterms:hasVersion	europeana:usertag
		dcterms:isVersionOf	europeana:year
		dcterms:hasPart	
		dcterms:isReferencedBy	
		dcterms:references	
		dcterms:isReplacedBy	
		dcterms:replaces	
		dcterms:isRequiredBy	
		dcterms:requires	
		dcterms:tableOfContents	

8   3. Mapping guidelines

This chapter explains the general principles of mapping practice for the Europeana Portal and should be read in conjunction with the ESE Ver 3.1 specification. It also offers some mapping examples and additional guidance for mapping information that is a difficult fit with ESE.

8.1   3.1 General rules

1. Map as many as possible of the original source elements to the available ESE elements.

2. If it is not possible to map the source element to an appropriate ESE element then leave it unmapped or consider using <europeana:unstored>.

3. Preserve all XML attributes in the source data as they may enable future functionality.

4. Insert an xml:lang attribute wherever possible.

5. If possible use one of the more specific <dcterms> refinements (as indicated in the text) but only when the semantic of the source term clearly corresponds to the narrower term.

6. The persistent link to digital object and/or full information page should be given as a URL. These may need to be constructed from metadata values and information external to the metadata.

7. If it is difficult to decide which ESE element to map a source term to, consider how best to meet expectations of the user and the functionality of the system.

8. Where there are multiple values for the same element repeat the element for each instance of the value.

9. Consider how the data would perform in response to "who, what, where and when" queries. This therefore encompasses names, types, places and dates relevant to the object and what it depicts.

10. To ensure that your data will be meaningful when displayed in the new context consider adding a prefix or suffix. As a simple example, "100 x 200" could become "100 x 200**cm**"Such additions are especially important where data from several source elements have been aggregated into one ESE element. In the following (abbreviated) example several source elements have been mapped to <dc:description>:

Description: D. Nath. Matthaus von Wolf wurde 1724 in Konitz/Westpreußen geboren. 1748 Leibarzt des Bischofs von Posen, Leibarzt des Fürsten Lubomirski und Czartorski, Arzt in Dirschau……... Der Naturforschenden Gesellschaft, die dieses Kapitel verwaltete, vermachte er ein Herbarium und ein Naturalienkabinett.
Papier
Radierung
Text unter Abbildung: "D: Nath: Matthaus von Wolf, geb. 1724 d: 28. Jan: zu…..

Gedani 1785." Wolf, Nath. Matthaeus von. Arzt, Astronom, Naturforscher.
Erbaute der Stadt eine Sternwarte. 1724 Konitz/Westpreußen - 1784 Danzig.

The word"Papier"could have been preceded by "Materiel:"and "Radiering"by
"Technik:"as they were taken from those elements in the source data.

8.2   3.2. XML Attributes

Where xml attributes exist in the source metadata please preserve these in the data
transformation. It may be possible to use them in the future to trigger new functionality.

Language is of particular importance to Europeana so it is desirable to preserve or insert an
xml:lang attribute to all elements where it is applicable – for example in title, subject,
description. This attribute is used to indicate the language of the metadata and will allow
Europeana to distinguish the language of the metadata from the language of the object itself
or the language of the provider of the resources.

The value used in the xml:lang attribute should comply to RFC4646[1]. The values should
consist of an ISO 639-1[2] two letter code. In cases where only an ISO639-2 three letter code
exists the three letter value may be used.   ISO 639-3[3] should not be used.

8.2.1

8.2.2 *Examples of using the xml:lang attribute:*

- <dc:subject xml:lang="en">weather</dc:subject> two letter code, not "eng"
- <dc:subject xml:lang="en-US">cop</dc:subject> this form is acceptable in RFC4646
- <dc:subject xml:lang="de">wetter</dc:subject> two letter code not "ger" or "deu"
- <dc:subject xml:lang="ain">   </dc:subject> a local language that only has a
  three letter code

8.3   3.3 Strongly recommended elements

8.3.1 3.3.1   Title

<dc:title>

---

[1] RFC4646 http://www.ietf.org/rfc/rfc4646.txt

[2] List of ISO 639-1 and 639-2 http://www.loc.gov/standards/iso639-2/php/code_list.php

[3] ISO 639-3 http://www.sil.org/iso639-3/

The title or name by which the digital object is known. This is likely to be the title or name also applied to the original physical object in the case of a digitisation. Use the refinement dcterms:alternative for any title variants, translations etc.

Where many objects share the same title, as may be the case where there are many issues of a periodical, consider aggregating the title values with information from another element to give a more useful display.

Example:
<dc:title>A-Z : Luxemburger illustrierte Wochenschrift</dc:title>
<dc:description>1933-12-24 (Numéro 01)</dc:description>

Aggregate the values using a full stop as a separator to give:

<dc:title>A-Z : Luxemburger illustrierte Wochenschrift. 1933-12-24 (Numéro 01)</dc:title>

If possible, indicate the language of the title using an xml:lang attribute.

dc example:
<dc:title xml:lang="en">Eight weeks</dc:title>

MODS crosswalk[1] example
<mods:titleInfo><mods:title>Florida Environments Online<mods:/title></mods:titleInfo>
becomes
<dc:title xml:lang"en">Florida Environments Online</dc:title>
Note that MODS defines several Type attributes and sub-elements for the <titleInfo> element which may more accurately be mapped to <dcterms:alternative>[2].

EAD 2002 XML example:
<unittitle>The Golden Gate entrance to the bay of San Francisco: sunrise</unittitle>

8.3.2   3.3.2   Alternative title

<dcterms:alternative> This is a refinement of dc:title

This can be any alternative title or name by which the digital object is known and will often be the a name also applied to the original physical object in the case of a digitisation. It can include abbreviations or translations of the title.

---

[1] The MODS examples for dc elements are gratefully quoted from http://www.loc.gov/standards/mods/dcsimple-mods.html. The examples for dcterms are not taken from this source.

[2] See also http://www.loc.gov/standards/mods/v3/mods-userguide-elements.html

If possible, indicate the language of the alternative title using an xml;lang attribute.

DC example:
<dcterms:alternative xml:lang="es">Ocho semanas</dcterms:alternative>

8.3.3   3.3.3   Creator

<dc:creator>

The name of the creator or creators of the original physical object or the born digital object. Names can include those of people, organisations or services

Map each name to a separate repeated creator element if possible.

Ideally choose a preferred form of name from an authority source and if your data includes the authority source in an xml attribute this should be retained for future use. If you do not use and authority source, use a consistent form of the name e.g. Shakespeare, William.

DC example:
<dc:creator xsi:type="localnameauthority">Shakespeare, William</dc:creator>

8.3.4   3.3.4   Contributor

<dc:contributor>

The name of contributors to the either the original physical object or the born digital object. Names can include those of people, organisations or services

Map each name to a separate repeated contributor element if possible.

Ideally choose a preferred form of name from an authority source and if your data includes the authority source in an xml attribute this should be retained for future use. If you do not use and authority source, use a consistent form of the name e.g. Shakespeare, William.

DC example:
<dc:contributor scheme="localnameauthority">Frances Bacon</dc:contributor>
MODS crosswalk:
<mods:name><mods:namePart>Florida Geological Society</mods:namePart></mods:name>
<mods:name ><mods:namePart>Florida. Department of Agriculture and Consumer Affairs</mods:namePart></mods:name>
<mods:name><mods:namePart>Agricultural Experiment Station/Extension Service (IFAS)</mods:namePart></mods:name>
becomes:
<dc:contributor>Florida Geological Society</ dc:contributor >
< dc:contributor >Florida. Department of Agriculture and Consumer Affairs</ dc:contributor >

< dc:contributor >Agricultural Experiment Station/Extension Service (IFAS) </ dc:contributor >

Note the MODS role attributes before deciding where to map names.

Museumdat example:
```
<museumdat:museumdat>
    <museumdat:descriptiveMetadata>
      <museumdat:eventWrap>
        <museumdat:indexingEventWrap>
          <museumdat:indexingEventSet>
            <museumdat:eventType>Herstellung</museumdat:eventType>
            <museumdat:indexingActorSet>
              <museumdat:nameActorSet>
                <museumdat:nameActor museumdat:type="personalName">Anders, Albert</museumdat:nameActor>
              </museumdat:nameActorSet>
              <museumdat:roleActor>Hersteller</museumdat:roleActor>
            </museumdat:indexingActorSet>
          </museumdat:indexingEventSet>
        </museumdat:indexingEventWrap>
      </museumdat:eventWrap>
  </museumdat:museumdat>
```

In this museum example the data also contains the role of the contributor "Hersteller" (Manufacturer). Although Roles cannot be directly mapped to ESE, and ESE does not support the associating of names with roles, this information could usefully be mapped to <europeana:unstored>.

8.3.5 3.3.5 Date

<dc:date>

This date element should be used to contain the most significant date in the life of the digital object or the original physical object in the case of a digitisation. The value will be used as the basis for locating the object in the Timeline and the Date facet in the portal. If you are using the more precise date terms of <dcterms:created> or <dcterms:issued > these may also be used for the Timeline and Date facet.

The value in the date element (or one of its refinements), will be used by the europeana normalisation process to generate the value for <europeana:year> element. <europeana:year> will contain a standard four digit year of the Gregorian calendar (e.g. 1523). The value supplied may not be in this standard form but the normalisation process will attempt to identify a four digit year from the value supplied. Ideally therefore the value in this element should contain a year in the form yyyy. For example:
<dc:date>1933-12-24</dc:date> generates
<europeana:year>1933</europeana:year>

<dc:date>1914-1918</dc:date>generates <europeana:year>1914</europeana:year>
<dc:date>19780403</dc:date> generates <europeana:year>1978</europeana:year>

Objects where no <europeana:year> value can be generated will not appear in the Timeline or Date facet.

Before the Common Era, Before Christ or Before Present dates
Currently, the portal cannot use BC, BCE or BP dates but such dates should be retained in the mapped metadata in order to be present for future development of the portal.

Textual date values
Textual time periods will display in a result list but cannot be represented in the Timeline or date facet and should be converted into numeric dates.

Example where the provider has a specific time period defined:
> <localtimeperiod>Roman Britain</localtimeperiod>
> Transform and map as e.g. <dc:date>0043</dc:date> and <dc:date>0410</dc:date>

which will generate
> <europeana:year>0043</europeana:year> and
> <europeana:year>0410</europeana:year>

Example where text is used:
> <localtimeperiod>17th century</localtimeperiod>
> Transform and map as <dc:date>1601</dc:date> and <dc:date>1700</dc:date>

which will generate:
> <europeana:year>1601</europeana:year> and
> <europeana:year>1700</europeana:year>

8.3.6   3.3.6   Date Created

<dcterms:created> This is a refinement of <dc:date>

This is the date of the creation of the digital object or, in the case of a digitsation, the original physical object. See <dc:date> above for fuller information about the form of the date.

8.3.7   3.3.7   Date Issued

<dcterms:issued> This is a refinement of <dc:date>

The date when the digital object was formally issued or published. This is likely to be the date the original physical object was issued in the case of a digitisation. See <dc:date> above for fuller information about the form of the date.

MODS example:
<mods:originInfo><mods:dateIssued
encoding="iso8601">20030331</mods:originInfo><mods:dateIssued>
becomes
<dcterms:issued scheme="iso8601">20030331</dcterms:issued>
Note that the hyphenated form is preferred e.g. 2003-03-31

8.4 3.4 Recommended elements

8.4.1

8.4.2 3.4.1 Coverage

<dc:coverage>

Coverage can be used for either spatial for temporal aspects of the object being described. Values will typically include either a spatial location (place name or geographic co-ordinates) or a temporal period (a date range or period label).

If analysis of the data shows that it contains only spatial or only temporal data then please map to either the spatial or temporal element refinements described below: the greater precision will allow the addition of place and time-based functionality. If values in the source data are mixed or unknown then this more generic dc:coverage element should be chosen for the mapping.

When metadata contains information about multiple objects (e.g. a photograph of a flint arrowhead that has been digitised into a jpeg file) it can be difficult to decide whether to map aspects of the image to <dc:coverage>, <dcterms:spatial> or <dc:subject>. Examples of some interpretations used in the prototype can be found in the section on difficult elements in Section 5 of this document.

8.4.3 3.4.2 Spatial

<dcterms:spatial> A refinement of dc:coverage.

Use this element for information about the spatial characteristics of the digital object or the original physical object in the case of a digitistation i.e. what the resource represents or depicts in terms of space. This may be a place name, a location, spatial co-ordinates or a named adminitstative entity. Values in this element will appear in the Subject line of a full result display.

EAD example:
<controlaccess>
<geogname role="country of coverage" source="tgn">United States</geogname>

```
<geogname role="state of coverage" source="tgn">California</geogname>
<geogname role="city of coverage" source="tgn">San Francisco</geogname>
</controlaccess>
```
Becomes (with preserved attributes)
```
<dcterms:spatial role="country of coverage" source="tgn">United States</dc:spatial>
<dcterms:spatial role="state of coverage" source="tgn">California</dc:spatial>
<dcterms:spatial role="city of coverage" source="tgn">San Francisco</dc:spatial>
```

8.4.4

8.4.5 3.4.3 Temporal

<dcterms:temporal> A refinement of dc:coverage.

Use this element for the temporal characteristics of the digital object or the original physical object in the case of a digitsation. Values in this element will appear in the Subject line of a full result display.

8.4.6

8.4.7 3.4.4 Description

<dc:description>

A prose description of the digital object or the original physical object in the case of a digitisation, elaborating on the information in the metadata. It is recommended to use a language attribute if possible to indicate the language of the description.

ESE has a limited number of elements and this can make it difficult to map some of the richer data elements that exist in provider metadata. Data mapped to <dc:description> will show in the results display. The following are examples of such data that could be mapped to the description element:

- statements relating to a technique applied to an object in terms of technology or craftsmanship e.g. carving, pressing, shoe making, binding
- statements where a technique includes reference to a material e.g. wood carving. When the statement is only about the material (wood, ivory) then <dcterms:medium> should be chosen.)
- statements about an event relating to an object

The europeana:unstored element could be used for this information if it is more appropriate in the provider context, however, the data would not then show in the result display.

See section 5 for further discussion on mapping difficult elements and some interpretations applied in the prototype.

MODS crosswalk example:

<mods:note>Florida Environments Online contains both digital full text materials and research bibliographies about the ecology and environment of Florida. […..] as well as engineering documents created by the UF Engineering and Industrial Experiment Station</mods:note>

becomes

<dc:description xml:lang"en"> Florida Environments Online contains both digital full text materials and research bibliographies about the ecology and environment of Florida. […..] as well as engineering documents created by the UF Engineering and Industrial Experiment Station</dc:description>

EAD 202 XML examples:

<scopecontent>
<p>View from off Point Lobos looking towards San Francisco Bay with Steamship John L. Stephens and clipper ship Flying Cloud just entering the Golden Gate.</p></scopecontent>

becomes

<dc:description xml:lang"en">View from off Point Lobos looking towards San Francisco Bay with Steamship John L. Stephens and clipper ship Flying Cloud just entering the Golden Gate.</dc:description>

<physdesc>
<physfacet type="Materials and Techniques">print:engraving, hand colored</physfacet>
</physdesc>

Becomes

<dc:description xml:lang"en">print: engraving, hand colored</dc:description>

8.4.8   3.4.5   Is Part of

<dcterms:isPartOf> A refinement of <dc:relation>

This element should be used to identify a related resource in which the described resource is physically or logically included. More particularly, use this for the name of the collection of which the digital object is part.

8.4.9   3.4.6   Language

<dc:language>

This element should be used to state the language of the digital object and should be repeated if the object has more than one language. If there is no language aspect to the object (for instance, a photograph) then the element should be ignored.

The use of RFC 4646 is highly recommended. Best practice is to use that ISO 639-1 (two letter code) unless only the ISO639-2 three letter code exists. (ISO 639-3 should not be used.)

Note: the <dc:language> element should not be used to indicate the language of the metadata but an xml attribute can be used for that purpose e.g. <dc:title xml:lang="en">After dark</dc:title>. Additionally, the language of the content provider is given in a normalised <europeana:language> element described later in this docment.

MODS crosswalk example
<mods:language authority="rfc4646">en</mods:language>
becomes
<dc:language scheme="rfc4646">en</dc:language>

8.4.10

8.4.11 3.4.7   Publisher

<dc:publisher>

The name of the publisher of the digital object or the original physical object in the case of a digitisation.

MODS example:
<mods:originInfo><mods:publisher>Oxford University Press</mods:originInfo><mods:publisher>

8.4.12 3.4.8   Source

<dc:source>

Officially this element should be used to indicate a related resource from which the digital object is derived. The element has widely been used for other interpetations of "source" however including reference sources and the names of content providers and Europeana does not differentiate between these uses. They may all continue to be mapped to this dc:source element but the name of the content provider will be held in the normalised <europeana:provider> element.

8.4.13

8.4.14 3.4.9   Subject

<dc:subject>

The subject of the digital object or the original physical object in the case of a digitisation. This can include topics, people and places but consider using the spatial and temporal elements for places and time periods if the source data allows.

Best practice is to use a separate instance of the element for individual subject terms.

MODS example:
<mods:subject><mods:topic>Environmental sciences</mods:topic></mods:subject>
<mods:subject><mods:topic>Geology</mods:topic></mods:subject>
<mods:subject><mods:topic>Agriculture</mods:topic></mods:subject>
becomes
<dc:subject xml:lang="en">Environmental sciences<dc:subject>
<dc:subject xml:lang="en">Geology<dc:subject>
<dc:subject xml:lang="en">Agriculture<dc:subject>

EAD 2002 XML example:
<controlaccess>
<persname role="subject" rules="aacr2">Casey, James P.</persname>
<persname role="subject" rules="aacr2">Cora, Charles</persname>
<subject source="lctgm">Bays</subject>
<subject source="lctgm">Vessels</subject>
</controlaccess>
<controlaccess>
<geogname role="country of coverage"source="tgn">United States</geogname>
<geogname role="state of coverage"source="tgn">California</geogname>
<geogname role="site of coverage" source="tgn">Golden Gate
(channel)</geogname>
</controlaccess>
becomes
<dc:subject">Casey, James P.</dc:subject>
<dc:subject">Cora, Charles</dc:subject>
<dc:subject xml:lang"en">Bays</dc:subject>
<dc:subject xml:lang"en">Vessels</dc:subject>
<dcterms:spatial xml:lang"en">United States</dcterms:spatial>
<dcterms:spatia xml:lang"en"l>California</dcterms:spatial>
<dcterms:spatial xml:lang"en">Golden Gate (channel)</dcterms:spatial>

8.4.15 3.4.10 Type

<dc:type>

The nature or genre of the digital object or the original physical object in the case of a digitisation. This should be used to record the values given in the source data which, ideally, will have been taken from a controlled vocabulary. Typically this element contains values such as photograph, painting, sculpture.

Although the portal needs normalised values to support type-related functions it is desirable to keep the richer local terms. Thus, all these richer terms should be mapped to <dc:type>. A separate <europeana:type> element has been added to contain the normalised value for use in the portal. Use of this element is described in the Section 4 - Europeana elements and normalisation.

EAD 2002 XML example:
<controlaccess>
<genreform source="aat">color lithographs</genreform>
<genreform source="aat">marines (visual works)</genreform>
</controlaccess>
becomes
<dc:type scheme="aat">color lithographs</dc:type>
<dc:type scheme="aat">marines (visual works)</dc:type>

9   3.5 Additional elements

9.1.1   3.5.1   Format

<dc:format>

This element can include the file format, physical medium and dimensions of the original physical object or the digital object. It is recommended to use this element for the file format of digitised or born-digital object. For digital objects it is recommended to use MIME Types[1].

The refinements of <dcterms:extent> and <dc:terms:medium> can be used as appropriate for the more specific information.

It can be difficult to map terms from heterogeneous domains to <dc:format> and its two refinements <dcterms:extent> and <dcterms:medium> . Examples of some interpretations used in the prototype can be found in the section on difficult elements in Section 5 of this document.

MODS crosswalk example:
<mods:physicalDescription><mods:form>text/html</mods:form></mods:physicalDescription>
Becomes
<dc:format>text/html</dc:format>

9.1.2   3.5.2   Extent

<dcterms:extent> A refinement of <dc:format>

Use to record the size or duration of the original physical or digital object. This includes measurements of physical objects that have been digitised. To ensure a meaningful display for the user, please indicate the units of measurement in the value.

EAD example:
<physdesc>
<dimensions unit="metric">42.4 x 68 cm</dimensions>
<dimensions unit="inches">17 inches x 27inches</dimensions>
</physdesc>
Becomes (with preserved attributes)

---

[1] http://www.iana.org/assignments/media-types/

<dcterms:extent unit="metric">42.4 x 68 cm <dcterms:extent>
<dcterms:extent unit="inches">17 inches x 27inches <dcterms:extent>

9.1.3 3.5.3 Meduim

<dcterms:medium> A refinement of <dc:format>

The material or physical carrier of the resource. This refers to the medium of the digital or original physical object such as paper, wood or ivory.

It can be difficult to map terms from heterogeneous domains to <dc:format> and its two refinements <dcterms:extent> and <dcterms:medium> . Examples of some interpretations used in the prototype can be found in the section on difficult elements in Section 5 of this document

9.1.4 3.5.4 Identifier

<dc:identifier>

This element can be used for an identifier of the digital object or the original physical object in the case of a digitisation. Keeping this identifier throughout the ingestion workflow will support the de-duplication process.

The element <europeana:isShownBy> has been added specifically to hold the URL that gives direct access to the digital object. If the URL of the digital object is to be mapped to <dc:identifier> then it should also be mapped to <europeana:isShownBy> or <europeana:isShownAt> as appropriate (see section 4 on Europeana elements).

<mods:identifier type="uri">http://palmm.fcla.edu/feol/</mods:identifier>
Becomes
<dc:identifier>http://palmm.fcla.edu/feol/</dc:identifier>
And
<europeana:isShownBy>http://palmm.fcla.edu/feol/</dc:identifier>

9.1.5 3.5.5 Rights

<dc:rights>

Information about intellectual property rights, access rights or license arrangements for the digital object.

DC Example:

<dc:rights xml:lang="en">Creative Commons Attribution 3.0 License</dc:rights>

9.1.6 3.5.6 Provenance

<dcterms:provenance>

This element is to record a statement of any changes in ownership and custody of the resource since its creation that are significant for its authenticity, integrity and interpretation. This may include a description of any changes successive custodians made to the resource.

For Europeana this relates particularly to the ownership and custody of the original analog or born-digital object.

9.1.7 3.5.7 Relation

<dc:relation>

This element should be used for information about resources that are related to the digital object or the original physical object in the case of a digitisation. It has been used for a wide range of relationships and it is recommended to use one of the several more specific relationship refinements where appropriate. Ideally this value should be a URI but it is recognised that practice varies in this respect.

DC example:
<dc:relation xml:lang="en"> maps.crace.1/33</dc:relation>
This is the shelfmark for the map that was digitised from the Crace collection at the British Library.

9.1.8 3.5.8 Conforms to

<dcterms:conformsto> A refinement of <dc:relation>

Use this element to identify standards to which the described resource conforms.

<dcterms:conformsTo> W3C WCAG 2.0</dcterms:conformsTo> (for an HTML document that conforms to web content accessibility guidelines).

9.1.9 3.5.9 Has format

<dcterms:hasFormat> A refinement of <dc:relation>

Use this element to identify another resource that is substantially the same as the digital object but exists in a different format. Use <dcterms:hasVersion> for changes in content.

Note that the purpose of this element is to give the identifier of the other resource in a different format, not to state the format of the object being described.

<dcterms:hasFormat>http://upload.wikimedia.org/wikipedia/en/f/f3/Europeana_logo.png</dcterms:hasFormat>
A link to another image format of the tiff image file being described.

9.1.10 3.5.10 Is format of

<dcterms:isFormatOf> A refinement of <dc:relation>

Use this element to identify a related resource that is substantially the same as the digital object but in a different format.

9.1.11 3.5.11 Has Version

<dcterms:hasVersion> A refinement of <dc:relation>

Use this element to identify a related resource that is a version, edition or adaptation of the digital object. This implies changes in content rather than just a different format.

9.1.12 3.5.12 Is Version of

<dcterms:isVersionOf> A refinement of <dc:relation>

Use this element to identify a related resource of which the described resource is a version, edition, or adaptation. Changes in version imply substantive changes in content rather than differences in format.

9.1.13

9.1.14 3.5.13 Has Part

<dcterms:hasPart> A refinement of <dc:relation>

Use this element to identify a related resource that is included either physically or logically in the digital object.

9.1.15

9.1.16 3.5.14 Is Referenced by

<dcterms:isReferencedBy> A refinement of <dc:relation>

Use this element to identify a related resource that references, cites, or otherwise points to the digital object.

9.1.17 3.5.15 References

<dcterms:references> A refinement of <dc:relation>

Use this element to identify related resources that are referenced, cited, or otherwise pointed to by the digital object.

9.1.18 3.5.16 Is Replaced by

<dcterms:isReplacedBy> A refinement of <dc:relation>

Use this element to identify a related resource that supplants, displaces, or supersedes the digital object.

9.1.19 3.5.17 Replaces

<dcterms:replaces> A refinement of <dc:relation>

Use this element to identify a related resource that is supplanted, displaced, or superseded by the digital object.

9.1.20

9.1.21 3.5.18 Is Required by

<dcterms:isRequiredBy> A refinement of <dc:relation>

Use this element to identify a related resource that requires the digital object to support its function, delivery or coherence.

9.1.22 3.5.19 Requires

<dcterms:requires> A refinement of <dc:relation>

Use this element to identify a related resource that is required by the digital object to support its
function, delivery or coherence.

9.1.23 3.5.20 Table of contents

<dcterms:tableOfContents> A refinement of dc:description

Use for a list of the sub-units of the digital object or the original physical object in the case of a digitisation.

MODS example:
<mods:tableOfContents>1. Introduction, 2. Contents of Elements, 3. Mapping, 4. Changes Made for MODS Version 3.2</mods:tableOfContents>
Becomes
<dcterms:tableOfContents xml:lang"en">1. Introduction, 2. Contents of Elements, 3. Mapping, 4. Changes Made for MODS Version 3.2</dc:termstableOfContents>

10 4 Europeana elements and normalisation of values

Several of the Europeana terms are designated as mandatory and, in addition to this, they must contain values in standard form. Generally this means the data has to be normalised, either at the time of creating the mapped data or at the point of ingestion to the Europeana system. The normalisation takes place at different points in the process for different elements. This is summarised in the table below and more detail is provided in the following element descriptions.

Element	Obligation	Normalised format	Normalised value	Where is it normalised?
europeana:country	Mandatory	Yes (String)	Yes (Controlled list)	centrally
europeana:hasObject	Mandatory	Yes (Boolean)	Yes (True / False)	centrally
europeana:language	Mandatory	Yes (code)	Yes (Controlled list)	centrally
europeana:provider	**Mandatory**	**Yes (String)**	**Yes (Controlled list)**	**locally**
europeana:type	**Mandatory**	**Yes (String)**	**Yes (Controlled list)**	**locally**
europeana:unstored		Yes (String)	No	locally
europeana:uri	Mandatory	Yes (URI)	No	centrally
europeana:usertag		Yes (String)	No	centrally
europeana:isShownAt or isShownBy	Mandatory	Yes (URL)	No	locally
europeana:object	Mandatory	Yes (URL)	No	locally
europeana:year	Recommended	Yes (4 digits)	No	centrally (from local mapping)

10.1.1 4.1 Country

<europeana:country> This mandatory element is used to support the Country facet in the portal.

This is the name of the country in which the content provider is based or the value is "Europe" in the case of Europe-wide projects. The name of the country is derived automatically from information held in the Europeana content providers authority file. The element is populated automatically during the ingest process and content providers do not need to provide this information separately.

Country names are normalised to the ISO 3166 standard

10.1.2 4.2   Has Object

<europeana:hasobject> This element supports the process of generating an object thumbnail from a URL supplied by the data provider. It is for internal use during the ingestion process and is added centrally as part of that process. Content providers should not map a value to this element.

For information: this element takes a value of True or False and indicates to the Europeana system that there is a URL in the <europeana:object> element identifying an object in a compliant format could be used to create a thumbnail for the digital object. Objects with no URL for a thumbnail will have only the default type icon showing in the display.

10.1.3

10.1.4 4.3   Is Shown At

<europeana:isShownAt> This element will be active in the portal and will provide the link to the digital object in full information context on the provider website.

It is a complementary element to < europeana:isShownBy> (please see below before mapping) and **it is mandatory to provide a URL link in one of these elements.**

Europeana is focuses on giving access to digital objects and this two-fold URL strategy is useful when ESE is not able to properly represent the original metadata model. For example, EAD, an archival metadata format, has complex hierarchical information and cannot be displayed as such in Europeana interface. The isShownAt element allows the Europeana users to access to the EAD hierarchy.

If the digital object is displayed with local metadata, with a header or banner, or if the object is only accessible by clicking another icon on the local page then <europeana:isShownAt> should be used.

Records that do not have a valid link in either <europeana:isShownAt> or <europeana:isShownBy> will be excluded from the portal. The link should be to the object in its best available resolution.

Examples of using this element are provided in Annexe A

10.1.5 4.4   Is Shown By

<europeana:isShownBy> This element will be active in the portal and will provide a link to the digital object on the provider website.

To map to <europeana:isShownBy> the object should be directly accessible by the URL and reasonably independent at that location. The inclusion of, for example, short copyright information or minimal navigation tools would be acceptable in this element. The link should be to the object in its best available resolution.

This is a complementary element to <europeana:isShownAt> (above) and **it is mandatory to provide a URL link in one of these elements**.

(Note that it may be possible to construct URL's to objects that are embedded in HTML by right-clicking on the object and using the metadata identifier. In this case please provide the direct URL in <europeana:isShownBy> as well as the link provided in <europeana:isShownAt>.)

Europeana is focussed on giving access to digital objects and records that do not have a valid link in either <europeana:isShownBy> or <europeana:isShownAt> will be excluded from the portal.

Examples of using this element are given in Annexe A.

10.1.6

10.1.7 4.5   Language

<europeana:language> This mandatory element is used to support the language facet in the portal.

This is the official language of the country in which the content provider is located. The name of the country is derived automatically from information held in the Europeana content providers authority file. The element is populated automatically during the ingest process and content providers do not need to provide this information separately.

Languages are normalised to the ISO 369-1 standard two character code. For countries with more than one official language the value "mul" is used.

10.1.8 4.6   Object

<europeana:object> This mandatory element supports the process of creating thumbnails for the portal and works in conjunction with <europeana:hasObject>. There are three situations:

1) the metadata includes a URL to an existing thumbnail image so this will be mapped to <europeana:object>
2) There is no existing thumbnail (so no URL to it) but a thumbnail image can be generated from the resource the URL in <europeana:isShownBy> links to. In this case the URL mapped to <europeana:isShownBy> is also mapped to <europeana:object>.
3) There is no existing thumbnail (so no URL to it) and a thumbnail image cannot be generated from the resource the URL in <europeana:isShownBy> links to. In this case nothing can be mapped to <europeana:object>, the value in <europeana:hasObject> will be "False" and a default thumbnail will be used in the portal.

Thumbnail images are generated from digital objects whose format (such as PDF, JPEG, and TIFF) is compliant with the Image Magik software used. Image Magic[1]k needs to have direct access to the image to create a thumbnail: it cannot make a thumbnail out of an image embedded in an HTML page.

Europeana creates one thumbnail per record. If a record contains several pages in a PDF, the front page is used to create the thumbnail. If a record contains several image files, the first file that appears is used for the thumbnail.

10.1.9

10.1.10   4.7 Provider

<europeana:provider> This mandatory element supports the Provider facet in the portal.

This element should contain the name of the organisation that holds the digital object. The names of these organisations should be provided in a standard form from a controlled list. The value may be supplied by either the content provider or the aggregator who is supplying data to Europeana but care should be taken to ensure that the name is consistently provided in the same form. (Note: ESE V3.1 states that the Europeana Office will adds this data but that has now changed.)

The Europeana Office has maintained a controlled list of provider names for organisations that have contributed to date, but this may not continue to be the case.

10.1.11   4.8 Type

<europeana:type> This mandatory element is used to support the Type facet in the portal, the categorisation of objects in the result display and assignment of the appropriate default thumbnail icon.

---

[1] http://www.imagemagick.org/script/index.php

Europeana currently handles only four object types: text, image, sound and video. Content providers must map each of the values they use in their local data to one of the four Europeana type values. For the prototype this was achieved by providers supplying a mapping (and a translation) on a spreadsheet to the Europeana Office and the processing took place centrally. This process must now be carried out by the content provider or their aggregator. At some point in the process of extracting data from local databases, transforming and normalising it into ESE, one of the four Europeana type values must be inserted, in upper case, into a <europeana:type> element for each object. The original rich local values should not be discarded but should be mapped to <dc:type>

Example of Type mapping spreadsheet:

Local terms in your metadata <dc:type xml:lang="xx">	Europeana Type Classification	Remarks (<dc:type xml:lang="en">)
Czasopismo	TEXT	periodical
dokument życia społecznego	TEXT	pamphlet, ephemera
druk muzyczny	TEXT	music prints
Fotografia	IMAGE	photography
Grafika	IMAGE	print
Książka	TEXT	book
Pocztówka	IMAGE	postcard

10.1.12   4.9 Unstored

<europeana:unstored> This element is indexed for searching.

This element has been created in order to allow providers to retain all important information that cannot otherswise be mapped to ESE. The contents of this element are indexed and searched but the values do not show in the search result. Care should be taken not to map several fields with similar data to avoid distorting the weighting.

Example:
<europeana:unstored xml:lang="en"> National Gallery, London</europeana:unstored>
Where the National Gallery is the current location of the object but there is no other element in ESE suitable for this information.

10.1.13   4.10 URI

<europeana:uri>   This mandatory element supports the internal functioning of the Europeana system. It is generated by the Europeana system and providers do not need to supply a value for this element.

The value in this element is a unique identifier for each object record in the system. It is generated algorithmically based on an element in the source metadata that provides a unique identifier for the object. Should a provider supply records where duplicated values are found in all the possible identifier elements only the first record will be retained and subsequent ones discarded.

10.1.14

10.1.15　　4.11　User tag

<europeana:usertag> This element is provided to support future functionality in Europeana. It is not currently implemented and providers do not need to supply any values for it.

10.1.16　　4.12　Year

<europeana:year> This element is used to support the Timeline and the Date facet in the portal.

The value in this element is a four digit year (YYYY) from the Gregorian calendar. It is generated during the normalisation process from the value provided in the <dc:date> element (or one of the date refinements as appropriate) by the provider. The value supplied in <dc:date> may not be in this standard form but the normalisation process will attempt to identify a four digit year from the value supplied. Ideally therefore the value in the <dc:date> element should contain a year in the form YYYY.   Objects where no <europeana:year> value can be generated will not appear in the Timeline or Date facet.

See the <dc:date> in section 3.2.5 for suggestions about converting textual dates into a numeric form.

11 5 Difficult elements

When mapping from diverse rich metadata formats to the relatively simple set provided by ESE it can be difficult to decide where an element belongs. This section attempts to summarise some of the interpretations that were made during the mapping for the prototype.

11.1 5.1 Techniques and materials related to the object

It is difficult to map information about the techniques and materials used in relation to an object to ESE. This section summaries information that may already have been stated in a particular element as well.

dc:format should be used if the semantic is not clear enough to distinguish the following cases:

- If data includes information about the craftsmanship or technology related to the object (i.e. pressing, binding, carving, shoe making etc), the preferred mapping is to **dc:description** as the data will then be displayed in the result.
- If data includes information about the physical materials of the object (i.e. ivory, wooden, cast-iron etc), the preferred mapping is to **dcterms:medium**
- If data includes information about both two cases described above, it is recommended to map it to preferably **dc:description**

Examples from the film domain inlcude the following examples. Again **dc:format** is used as the generic element but in some cases the refinements could be applied.

- *FilmManifestation/Duration* - **dcterms:extent**
  (1 min, 4 min, 1 min 30 sec, 2 min, 3 min, 4 min 8 sec)

- *FilmManifestation/CarrierAspect* - **dc:format**
  (1:1,33, 16:9 86 1:1,37, 1:2,35)

- *FilmManifestation/CarrierFormat* - **dc:format**
  (35 mm, DVD, 16 mm, VHS)

- *FilmManifestation/Colour* - **dc:format**
  (Black and White, Colour, tinted)

- *FilmCopy/CarrierType* - **dc:format**
  (Positive, Digital file, n/a, Video tape, Duplicate negative, Reversal positive, Negative)

- *FilmCopy/CarrierMaterial* - **dcterms:medium**
  (Acetate, Polyester, Nitrate)

11.2 5.2 Measurements

The measurement of the object (e.g. 10 cm, 534 kg, 34 minutes, 6 miles, 800x600 pixels, 600dpi, 192kbps etc) is mapped preferably to **dcterms:extent.**

Example
_DIMENSIONS_DE_L_OBJET_-_CM_ has values: L. 84, x l. 68, x ép, L.182 x l.140 diam 1,5 diam, 2.8 diam, 1.6 diam, 1.8, long, 9.4 diam, therefore it is mapped to **dcterms:extent**

11.3 5.3 Events and roles

It is accepted that it is not easy to record event and role data in ESE. The following are suggestions.

If data includes information about events related to the object, it is recommended to map it to preferably **dc:description**

Data may include information about people/actors in events. If you are confident that it is directly related to the object itself, the preferred mapping is to **dc:contributor** If you are not sure, for example,the people/actors may be related to what is depicted in image, use **dc:description** or **europeana:unstored**. **dc:creator** should only be used if you are confident about the relationship between the object and the actor (e.g. maybe an attribute indicates the actor is the creator of the object).

If data includes information about the type of the agent roles in events (e.g. mask maker, builder, actress, archaeologist etc), it is recommended to map it to preferably **europeana:unstored**. The names of agent (roles) have to be mapped to **dc:contributor.** A notable example for agent roles is MARC Relator. According to DCMI, most of MARC Relator should be mapped to **dc:contributor**.
For more information, please see these websites:
Using Dublin Core – Appendix, roles http://dublincore.org/documents/usageguide/appendix_roles.shtml
MARC relator terms and Dublin Core http://dublincore.org/usage/documents/relators/
Relators http://lcweb2.loc.gov/cocoon/loc.terms/relators/dc-relators.html

11.4 5.4 Multiple resources described in one record

In this example, a physical flint arrowhead is depicted in a photograph which is digitized as jpeg.

No	Value	Original field name	Best guess at mapping to DC	Solution adopted
1	Title Neolithic arrowhead from Southern Iraq	Name	dc:title	dc:title
2	Unknown	Actor	dc:creator	dc:creator
3	Neolithic	Period	dc:date	dc:date
4	image/jpeg	Format	dc:format	dc:format

5	1.5 cm X 3cm	Dimension	dcterms:extent	dcterms:extent
6	flint	Material	dcterms:medium	dcterms:medium
7	Iraq	Origin	dcterms:spatial?	**dc:subject**
8	London; BM	Current Location	dcterms:spatial?	**dc:description**

The mappings of items 1 to 6 are imperfect but adequate for searching given the limitations of ESE metadata. However, for items 7 and 8 it does not make much sense to map both values to the same <dcterms:spatial> element, as shown in the "Best Guess" column.

In this case it is suggested that number 7 should be mapped to <dc:subject> as Iraq is one of the subjects in the image of the arrowhead i.e. it appears as the background with the arrowhead in the foreground. The image is not a photograph of the excavation in Iraq which would more readily justify the use of <dc:spatial>. For number 8 where the relationship of the object to the location is less obvious, the generic <dc:description> is adequate in these circumstances.

This suggestion provides a principle for using <dc:coverage> (and its spatial refinement) in these complex cases:
- where an image depicts a place, typically a landscape painting or photograph, use <dc:coverage> or <dc:spatial>
- where the place depicted is incidental to the story of the object depicted, use <dc:subject>
- where the relationship is not obvious the fall-back mapping is <dc:description>

12 Annexe A

12.1 Examples of <europeana:isShownBy> and <europeana:isShownAt>

europeana:isShownBy should contain the the URL that gives a direct link to the digital object. The digital object needs to be reasonably independent and directly accessed by the URL.

europeana:isShownAt should contain a URL where the object is displayed within an information context or is accessed indirectly via another link.
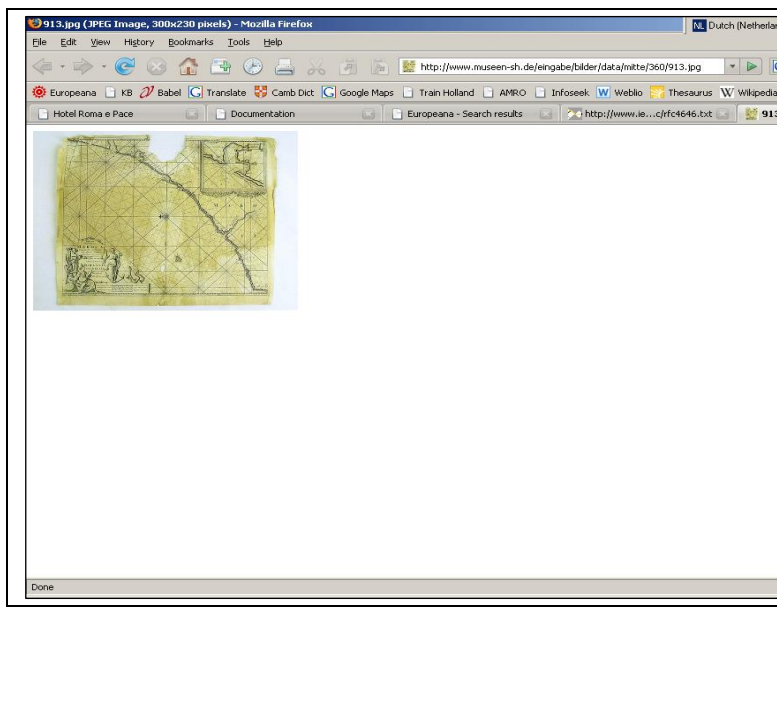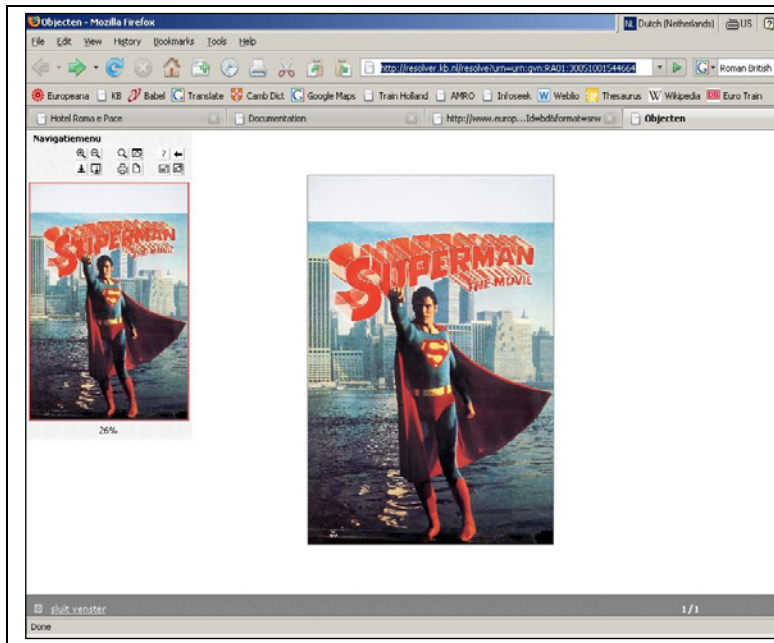

12.1.1

12.1.2

12.1.3 Example 1 - isShownBy

This is clearly <isShownby> because the URL leads to the JPEG image itself (note the "jpg" extension in the URL).
http://www.museen-sh.de/eingabe/bilder/data/mitte/360/913.jpg
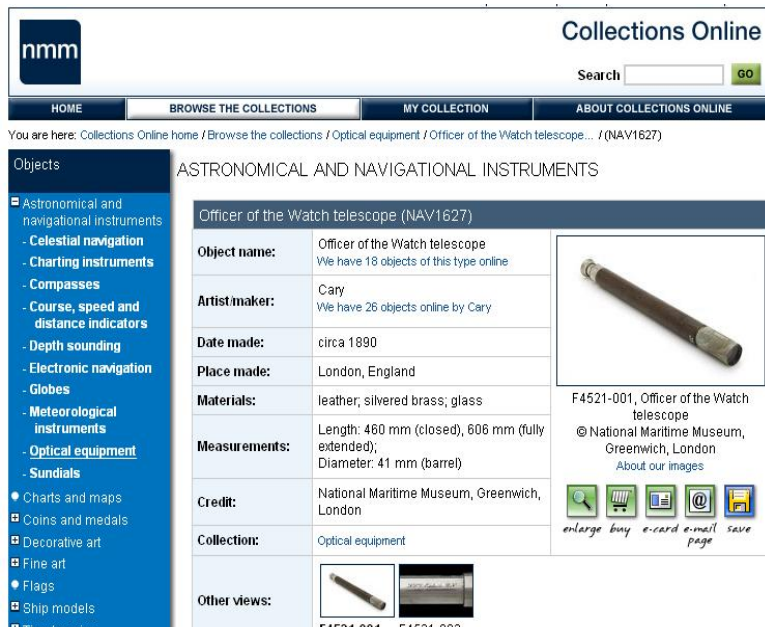
12.1.4

12.1.5 Example 2 - isShownBy

This could be interpreted differently as there is a small set of navigation tools with the image. <isShownBy> is still appropriate however as there is no other accompanying information - no banner or HTML frame etc. Immediate access to the digital object is the main criterion.
http://resolver.kb.nl/resolve?urn=urn:gvn:RA01:3005100154464

12.1.6

12.1.7

12.1.8 Example 3 - isShownAt

This is clearly <isShownAt> because the object is embedded in HTML web page.
http://www.nmm.ac.uk/collections/explore/object.cfm?ID=NAV1627

Consider constructing the URL of the object by right-clicking it to see the identifier. This could then be used with external data to construct a value for <isShownBy>
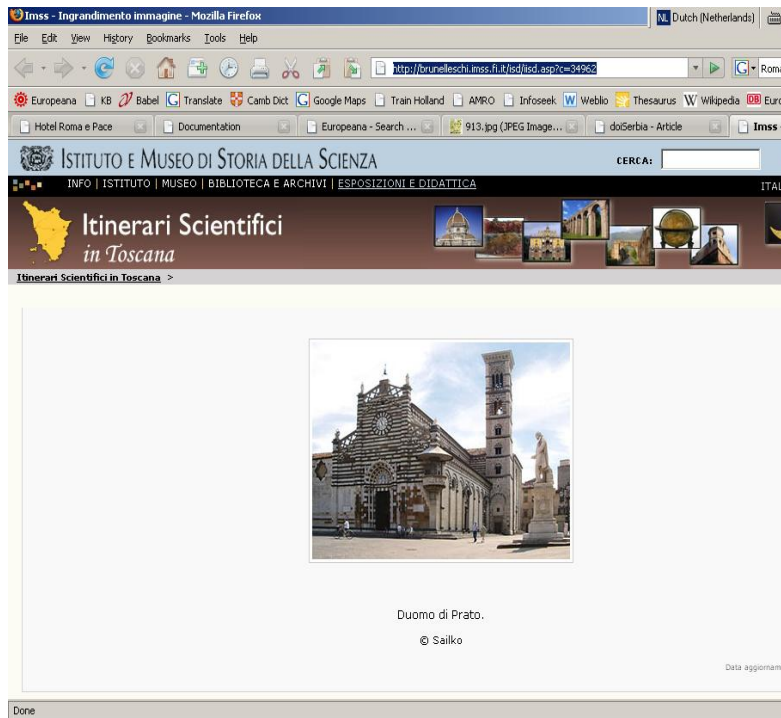
12.1.9

12.1.10

12.1.11   Example 4 - isShownAt

Although the web page is very simple showing only one object (digital photo), this should be <isShownAt> because there is a banner etc http://brunelleschi.imss.fi.it/isd/iisd.asp?c=34962

12.1.12

12.1.13

12.1.14   Example 5 - isShownAt

The red circle shows the "Full Text" hyperlink which allows the users to link to the pdf file. The URL of this website is <isShownAt>, while the URL of PDF is <isShownBy>. http://www.doiserbia.nbs.bg.ac.yu/Article.aspx?ID=0352-51390212887G&AspxAutoDetectCookieSupport=1

13 Document history

Version	Author	Updated	Comments
V0.5	Go Sugimoto	08-04-09	Draft creation
V0.6	Robina Clayphan	21-04-09	Draft modification
V0.7	Robina Clayphan	28-04-09	Draft modification after review
V0.8	Robina Clayphan	01-05-09	Draft modification for limited circulation after partial review
V1.0	Robina Clayphan	28-05-09	Initial release.

14 Acknowledgements

Thanks are due to the following for their contribution to these guidelines:
Robina Clayphan, Catherine Lupovici, Go Sugimoto, Julie Verleyen

Some of the MODS and EAD 2002 XML examples have been copied or adapted from
http://www.loc.gov/standards/mods/dcsimple-mods.html and
http://www.oclc.org/programs/ourwork/past/culturalmaterials/RLG_desc_metadata.pdf
respectively, with grateful thanks.