Project Acronym: Europeana v2
Grant Agreement number: 270902
Project Title: Europeana Version 2

D5.2: Up-to-date Guidelines

Revision	FINAL
Date of submission	30 May 2014
Author(s)	Robina Clayphan, Interoperability Specialist, Europeana Foundation.
Dissemination Level	[only for members of the consortium and the Commission Services]

REVISION HISTORY AND STATEMENT OF ORIGINALITY

Revision History

Revision No.	Date	Author	Organisation	Description
Draft 0.1	08/01/2014	Robina Clayphan	Europeana	First draft
Draft 0.2	14/01/2014	Robina Clayphan	Europeana	Edited and extended after review
Final	22/05/2014	Robina Clayphan	Europeana	Finalised after recent document updates.

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise.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

Table of Contents

1. Introduction and Scope

1.1. Introduction

"D5.2 is the deliverable for Task 5.3: Support for providers and aggregators.

Development of the portal will consist of a series of releases over the three year term of the project to include changes with impact on the data supplied by providers. This task will ensure on-going interoperability through the periodic update and enhancement of resources to support Data Providers and Aggregators. This will include producing and updating documentation and guidance materials relating to the data ingestion workflow and data validation tools such as the Content Checker. Production of further specifications, recommendations and guidelines based on standards and best practice regarding the contribution of data to Europeana including the provision of a Technical FAQ. Specification of requirements for improving the ingestion workflow tools and procedures. D5.2 Up-to-date guidelines."[1]

The development and implementation of the Europeana Data Model has been a learning process for all involved: the R&D and development staff at Europeana had to take a theoretical model and translate this into a functioning implementation; the ingestion team had to understand the practical implications of harvesting and processing the data; the data providers and aggregators had to understand the relationship of their local data to the new and complex model. Documentation and web resources have been created to support all the participants in this data supply chain. There has been an on-going process of revision and updating as the realities of the undertaking revealed themselves.

1.2. Scope

Up-to-date guidance material has been produced throughout the course of the project both as documentation and as web resources. This report summarises the production of the guidance material indicating the intended audience for each item and showing how it has reflected the changes in data and processes over the course of that time. A separate document, D5.3 from Task 5.4, focuses on changes made in the Europeana Data Model to support developments in the portal and repository and arising from the R&D work in WP7.

2. Documentation and web resources – summary

2.1. List of Data Model Documentation
- The EDM Factsheet [2]
- The EDM Definition – the full specification of the Europeana Data Model (EDM)[2]
- The EDM Primer – the narrative explanation of EDM[2]
- The EDM presentation[2]
- The EDM Guidelines – guidance for creating EDM data[2]
- The EDM Object Templates[3]
- The EDM XML Schema and Validation resource[2]
- The EDM Case Studies[2]
- The final version of the Europeana Semantic Elements (ESE) documentation aligned to EDM[4]

---

[1] Europeana Version 2 Description of Work, WP5, Task 5.3 Description,
[2] http://pro.europeana.eu/ edm-documentation
[3] http://europeanalabs.eu/wiki/EDMObjectTemplatesProviders

2.2. The Provide Data web pages in Europeana Professional[5]

- Who provides data?
- How to provide data?
- Legal requirements
- Technical requirements
- Providers' FAQs

2.3. Other resources

Slideshare account for presentations

Further information about each of these is given in the following sections.

3. Data Model Documentation

Europeana V2 started in October 2011 and at that time the data model in use was ESE. During the course of the project the transition was made to EDM. This started as a process of prototyping with a few partners while most partners continued to supply ESE. EDM is fully backward compatible with ESE and such data is still accepted but is now deprecated. The two metadata formats were running in parallel for the course of the project so both ESE and EDM documentation has been produced and updated.

3.1. The EDM Factsheet

This is a two page factsheet giving a high level overview summarising the rationale and expected benefits of EDM. It is intended for non-technical people who may need information about the new model.

3.2. The EDM Definition

The *EDM Definition* is the formal specification of the classes and properties that could be used in Europeana. Version 5.1 was released early in 2010 and comprehensively reviewed and revised through to May 2011 resulting in v5.2.2 as part of Europeana V1.

Updates contained in version 5.2.3 dated 24/02/2012
- changes resulting from the process of moving from theory to implementation internally e.g. addition of the ProvidedCHO class
- changes needed for on-going Europeana developments and initiatives e.g. the addition of edm:ugc for projects gathering objects from users
- changes resulting from requirements put forward by other projects e.g. introduction of "3D" as an edm:type for CARARE.

Updates contained in version 5.2.4 dated 14/07/2013.
- ongoing amendments and additions as implementation progressed technically e.g. addition of edm:europeanaProxy, edm:begin, edm:end, edm:preview edm:collectionName
- amendments to various aspects of properties, such as the definitions or constraints, as usage demonstrated the need
- generalisation of property definitions to make EDM usable outside Europeana. For example, the Digital Public Library of America wished to adopt many of our terms but

---

[4] http://pro.europeana.eu/ese-documentation
[5] http://pro.europeana.eu/provide-data

> needed the definitions to be less Europeana-specific (Country, DataProvider and Provider)
> - adjustments recommended by Task Forces e.g. edm:isNextInSequence made repeatable following the recommendations from the hierarchical objects Task Force
> - major editorial changes to rationalise EDM and ESE documentation pending official transition to EDM and deprecation of ESE.

Updates contained in version 5.2.5 dated 22/05/2014. This is the current version.
- edm:datasetName property added and collectionName deprecated
- dcat:Dataset class added in "Relevant classes…" section and "dcat" added as a namespace

3.3. The EDM Primer

It is not easy to understand how the data model works only from reading the specification so this Primer was written as a companion document and gives a narrative explanation. First released early in 2010 the Primer was significantly re-organised after expert review in October 2011. This version also took account of the changes made to the EDM Definition described above. The current version was issued in July 2013 with minor updates.

3.4. The EDM presentation

A PowerPoint presentation was created in 2010 and widely used to introduce partners to the data model and offer a more visual route to understanding it. It has been regularly updated following feedback from partners and to keep it abreast of changes in the model. It was last revised in September 2013 to incorporate examples of real data and images that are now available in the portal.

3.5. The EDM Mapping Guidelines

The EDM Mapping Guidelines document is the key publication for data providers and aggregators wanting to learn how to create data in the EDM format for submission to Europeana. It specifies the subset of EDM currently being implemented. It groups the metadata properties according to the required classes and contains definitions of the properties, information about the data types that can be used as values and the obligation level of each property. Importantly, the document also has an example of original provider data (in LIDO format), the same data converted to EDM (in RDF/XML) and diagrams giving a visual representation of the metadata distributed between the classes.

Version 1.0 dated 27/10/2011
EDM was (and is) being implemented incrementally so only a subset of the classes and properties defined in the model specification were available for use in the first instance. To avoid confusion by appearing to differ markedly from the EDM specification, and to indicate the full future potential of EDM, all the EDM classes and properties were included in the document with the ones not yet in use marked accordingly.
- properties grouped by class in simple table layout with key information
- a model example of real data before and after mapping
- a section showing data provider options for "Moving from ESE to EDM"

Version 1.0.1 dated 24/02/12
- Edits to align with changes in the EDM Definition

Version 2.0 dated 14/07/2013.
After considerable prototyping work with partners able to supply EDM data and technical developments internally, many more properties were opened up for use. The guidelines were comprehensively edited to reflect the changes already described in the section relating

to the EDM Definition.
- removal of "unavailable" marking on most properties
- changes to reflect the developments in the model during prototyping
- removal of "Moving from ESE to EDM" section as EDM became the metadata standard in use by Europeana. The section is still available for reference on the web pages as Europeana has committed to support ESE for the immediate future.
- Simplification changes reflecting comment from providers and the Ingestion Team as data was created and harvested in EDM (explicit "Mandatory" tables, etc.)

V2.1 dated 22/05/2014. This is the current version.
Further changes to refelect feedback from users and requirements for additional data properties.
- dc:creator added to WebResource
- place of birth and death properties added to Agent. Mapping table from original RDA namespace properties to new ones included for information.
- addition of explanatory section on language codes
- addition of explanatory section on Identifiers for resources in the data
- addition of link to EDM case studies and inclusion of Schematron rules
- addition of rule about one conceptual class for multiple CHO records
- additional terms defined in "Terminology"

3.6. The EDM Object Templates, the XML Schema and Validation resource

These three resources are for people looking for information and tools on the technical side of providing data.
- The EDM object templates are two simple wiki-listings of the provider and the internal views of the model. They are working documents for the R&D and development staff to quickly record and share changes as they arise. Being a wiki they are also a means of allowing interested (and probably more technically orientated) partners to watch the progress of the work as it evolves
- The EDM XML Schema for the *Europeana Data Model* (EDM) allows for automatic validation of EDM metadata represented in EDM. The EDM mapping rules are defined against the official release of this schema.
- The *EDM Validation* document explains how to make use of the validation rules with the Oxygen XML editor.

3.7. The EDM Case Studies

Recognising the need to provide models and share experience the case studies were produced. Most of them result from the early prototyping work with data providers. The studies show implementations of EDM in different domains and reports from task forces investigating optimum ways to represent certain aspects of data. It aims to demonstrate how EDM can be used for richer data provision, how the main EDM specifications can be specialised to support the needs of specific domains and how EDM is suitable for the publication of Linked Data. Studies were added as they became available during the course of the project:
- MIMO and EDM
- Europeana Libraries and EDM for Libraries
- Europeana Linked Open Data pilot
- Mapping EAD data to EDM
- The Polymath Virtual Library and EDM
- CARARE and EDM
- Recommendations for the representation of hierarchical objects in Europeana

3.8. Final Europeana Semantic Elements (ESE) documentation

Throughout Europeana V2 the ESE documentation had to updated in line with any changes in EDM and the current implementation of the portal that may impact data providers who were continuing to provide ESE data. This was quite an editing load as all documents had others referencing them and it was important to maintain consistency across the board during the transition period. The ESE documentation had the same pattern as the EDM documentation: a formal ESE Specification document and the Mapping and Normalisation Guidelines.

Versions of both dated 24/02/2012
- Both documents were edited to incorporate new properties and values specified in the EDM Definition of that date.

Version dated 14/07/2013
- Merging of the two ESE documents into one. The impending official transition to EDM as the preferred format meant that it was unlikely that the ESE documentation would need to be updated in future. The documents were therefore merged into a single document containing all the useful information from both. A link is provided to the earlier separate documents for those who require them for reference.

4. The Provide Data web pages in Europeana professional

The principal web resource for partners or any other cultural institutions wishing to submit data to Europeana is the "Provide Data" section of the Europeana Professional website. These pages attempt to give comprehensive guidance about the "Who", the "What" and the "How " of becoming a provider. The introductory page has a general introduction containing an up-to-date presentation of the value of being a provider and a contact form.

4.1. Who provides data?

This page explains the concept of the Europeana Network and gives an overview of the aggregation model adopted.

4.2. How to provide data?

This page gives a high level view of the steps involved in becoming a data provider. It starts with the administrative requirements of contact and discussion of options for participating then gives an outline of and pointers to the legal and technical requirements.

4.3. Legal requirements

There is a comprehensive set of pages explaining the legal context in which Europeana works. It is principally concerned with IPR issues and the licensing of both content and metadata.

4.3.1. Rights and EDM

It is in the section on "Available Rights Statements" that the legal requirements intersect with EDM.   Assigned rights statement need to be entered in the appropriate EDM property and to conform to the standards specified in order for rights-based functionality to work in the portal and API. Both EDM and these pages have been updated in line with the evolving strategies for providing clear rights information to users. Further changes are anticipated as EDM is adapted to handle the temporal aspects of IPR .

4.4. Technical requirements

These are the pages that give access to the EDM and ESE documentation already discussed. They give an overview of the available documentation and are updated to ensure that providers access only the latest versions. In addition to data format information there is a page giving information about aspects of multilinguality in the metadata.

4.5. Providers' FAQs

The providers' faqs attempt to answer the questions that most commonly arise about the metadata standard and the processes around submitting data. This page is updated periodically as the need arises.

5. Other EDM-related resources

5.1. Slideshare account[6]

Europeana has a presence on the Slideshare platform where presentations about all aspects of Europeana are uploaded. EDM presentations from many occasions and for different audiences can be found there.

6. Conclusion

Documentation and web resources relating to the data formats and work processes have been produced and updated throughout the course of Europeana v2. The project marked the transition from ESE to EDM for metadata and the streamlining of processes for ingesting data into Europeana. Feedback is still being received and it is anticipated that further changes will be made to support providers and aggregators going forward.

---

[6] http://www.slideshare.net/Europeana