



MILESTONE

Project Acronym: Ev3

Grant Agreement number: 620484

Project Title: Europeana Version 3

MS13: White paper on enhanced Europeana Data Model

Revision	Final
Date of submission	30-05-2015
Author(s)	Valentine Charles (Europeana Foundation) Antoine Isaac (Europeana Foundation)
Dissemination Level	[Public]

Project co-funded by the European Commission within the ICT Policy Support Programme

REVISION HISTORY AND STATEMENT OF ORIGINALITY

Revision History

Revision No.	Date	Author	Organisation	Description
1	18-03-2015	Valentine Charles	Europeana Foundation	First structure and table of contents
2	17-04-2015	Valentine Charles	Europeana Foundation	1 st draft
3	13-05-2015	Valentine Charles	Europeana Foundation	2 nd draft
4	25-05-2015	Antoine Isaac	Europeana Foundation	Comments and edits
5	27-05-2015	Valentine Charles	Europeana Foundation	Final draft
6	28-05-2015	Juliane Stiller Gregory Markus	Humboldt-Universität zu Berlin Netherlands Institute for Sound and Vision	Comments and edits
7	28-05-2015	Antoine Isaac	Europeana Foundation	Conclusion

Statement of originality:

This milestone contains original unpublished work except where clearly indicated otherwise.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

Introduction	4
1. The role of EDM in Europeana Services	4
2. EDM as a living model	6
2.1 The important contributors to the EDM development.....	6
2.2 Maintaining the Europeana Data Model as an open, living standard.....	8
2.3. The role of EDM in the context of the Semantic Web	10
3. Evolution and model refinements for EDM.....	11
3.1 Profiles for representing datasets and organizations	12
3.3 Profile for representing collections.....	13
3.4 Profile for representing technical metadata	14
3.5 Extensions for complex rights	14
3.6 Future work.....	15
4. Re-use and extensions of the EDM model.....	15
4.1 EDM extensions in the Europeana Network	15
4.2 EDM extensions outside the Europeana network.....	17
Conclusion.....	18
References	19

Introduction

From its beginning the Data Model (EDM) has been developed as a collaborative effort. The first requirements have been defined together with representatives from all the domains represented in Europeana: libraries, museums, archives and audiovisual archives. EDM defines now itself as an interoperable framework for describing digital cultural heritage data. The same collaborative effort supports its extensions and refinements to accommodate the subtleties of cultural heritage domain specific data. Europeana facilitates this effort by maintaining EDM as a flexible model, providing the adequate documentation to support communities' specific work.

This White Paper gives an account of the latest developments around EDM and highlights the key requirements that are necessary for the model to keep being a suitable framework for cultural heritage data.

1. The role of EDM in Europeana Services

In five years the Europeana Data Model (EDM) went from the status of a prototype and abstract model to become a mature data model which is supported by the whole Europeana data integration workflow. Europeana now aggregates, processes, enriches and disseminates data using the Europeana Data Model. The model is now a key part of the Europeana platform.

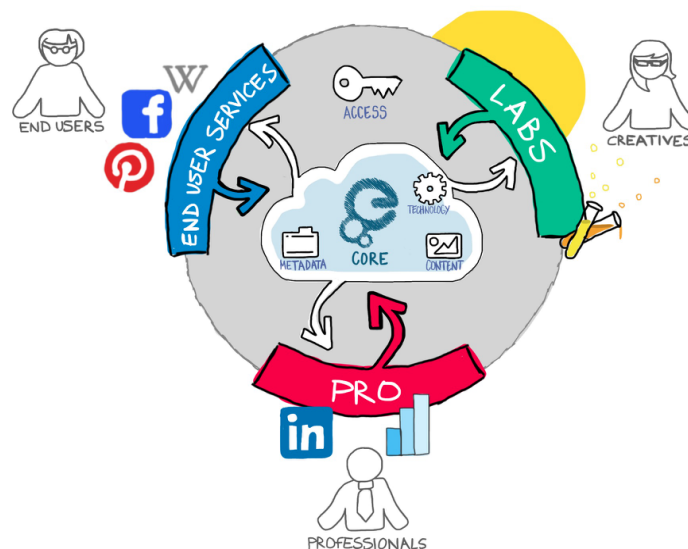


Figure 1: The Europeana Platform from the [Europeana Strategy 2015-2020](#). Visuals by Elco van Staveneren, www.denkschets.nl (CC-BY-SA)

It supports the *Core* of the platform¹ as it is the model used by data providers to describe the data they provide to Europeana. As EDM is built on an open and cross-domain Semantic Web based framework rather than being a community specific standard, it can accommodate the range and richness of particular domain standards such as LIDO² for museums, EAD³ for archives or

¹ Europeana Strategy 2015-2020 available at <http://strategy2020.europeana.eu/>

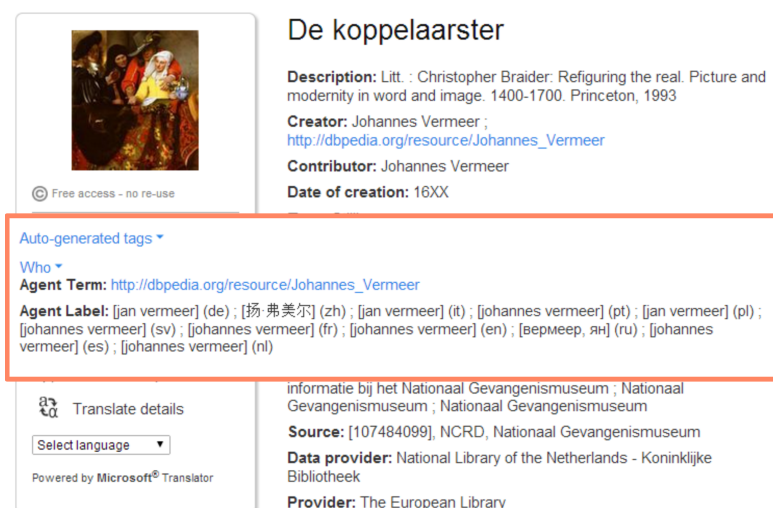
² <http://network.icom.museum/cidoc/working-groups/lido/lido-technical/specification/>

MARC⁴ for libraries. It contributes to the improvement of the overall quality of the Europeana dataset.

EDM allows Europeana to progressively build a network of digital and born-digital cultural heritage objects thanks to the representation of semantic links and relationships between objects. Europeana can for instance support more complex objects such as the ones described in the archives. In order to represent the diversity of definitions around hierarchical entities, Europeana and its Network worked on a series of recommendations on how to represent hierarchical objects in EDM (Charles, V. et al. (2013)). This work allowed Europeana to build a display for hierarchical objects in the Europeana portal. This display is directly built on top of the EDM properties supporting sequential and parent-child relationships.

In addition, Europeana has started to build its “Semantic Layer”, a network of contextual information on top of the cultural heritage objects. It includes concepts from ‘value vocabularies’ like thesauri, authority lists, and classifications, either coming from the network of Europeana’s providers or from third-party data sources. Indeed EDM enables the provision of contextually enriched data, which encourages data provider to send richer data. Europeana receives links to vocabularies as part of the data sent by providers. Europeana developed internally a small enrichment tool in order to ‘dereference’ the vocabularies URIs provided in the data, i.e., fetch all the multilingual and semantic data attached to a given vocabulary concept. This is especially made easy when the target vocabulary is represented with SKOS⁵, which is also the model EDM re-uses for describing concept data.

In the same way, Europeana performs automatic enrichment with other external value vocabularies and datasets such as GEMET⁶, GeoNames⁷ and DBpedia⁸ by creating links to objects in Europeana (Charles, V, Isaac, A, Freire, N. (2014)).



De koppelaarster

Description: Litt. : Christopher Braider: Refiguring the real. Picture and modernity in word and image. 1400-1700. Princeton, 1993

Creator: Johannes Vermeer ; http://dbpedia.org/resource/Johannes_Vermeer

Contributor: Johannes Vermeer

Date of creation: 16XX

Auto-generated tags ▾

Who ▾

Agent Term: http://dbpedia.org/resource/Johannes_Vermeer

Agent Label: [jan vermeer] (de) ; [扬·弗美尔] (zh) ; [jan vermeer] (it) ; [johannes vermeer] (pt) ; [jan vermeer] (pl) ; [johannes vermeer] (sv) ; [johannes vermeer] (fr) ; [johannes vermeer] (en) ; [вермеер, ян] (ru) ; [johannes vermeer] (es) ; [johannes vermeer] (nl)

Translate details

Select language ▾

Powered by Microsoft® Translator

informatie bij het Nationaal Gevangenis­museum ; Nationaal Gevangenis­museum ; Nationaal Gevangenis­museum

Source: [107484099], NCRD, Nationaal Gevangenis­museum

Data provider: National Library of the Netherlands - Koninklijke Bibliotheek

Provider: The European Library

Figure 2: Enrichment of the *dc:creator* field with language variants from DBpedia.

³ <http://www.loc.gov/ead/>

⁴ <http://www.loc.gov/marc/>

⁵ SKOS specifications are available at <http://www.w3.org/TR/skos-reference/>

⁶ <http://www.eionet.europa.eu/gemet/>

⁷ <http://geonames.org>

⁸ <http://dbpedia.org>

EDM allows Europeana to become “a big aggregation of digital representations of culture artefacts together with rich contextualization data and embedded in a linked Open Data architecture” (Gradmann, S. (2010)).

EDM also contributes to the *Access* layer of the platform. EDM data are published via the Europeana API but also as Linked Open Data via data.europeana.eu⁹. It allows cultural heritage to be re-used by third parties in other applications not necessarily connected to cultural heritage.

Finally, EDM enables the development of new *Services* within Europeana. New portal functionalities can be developed based on EDM such as the hierarchical display, which is fully based on the relationships expressed in the data. EDM also enables new browsing features for users answering the ‘Who?’, ‘What?’, ‘When?’, ‘Where?’ questions which contextualize the Europeana content.

2. EDM as a living model

2.1 The important contributors to the EDM development

The development of the EDM wouldn’t have been possible without the support of the data providers and aggregators within the Europeana Network. By raising requirements, highlighting the issues encountered with the previous model (Europeana Semantic Elements¹⁰), they helped in making the model stronger. The technical community, EuropeanaTech¹¹, more specifically contributed to the refinement of the model around some key challenges raised by the community. For instance the Task Force on Hierarchical Objects made recommendations (Charles, V. et al. (2013) on how to represent complex objects such as archival documents. The recommendations specified the use of properties defining hierarchical relationships between objects. The Task Force on Multilingual and Semantic Enrichment (Stiller, Isaac & Petras (eds.) (2014)) conducted an analysis of the controlled vocabularies, collections and metadata fields used and produced for/by enrichment in Europeana. This work is currently followed by a new Task Force aiming at the definition of enrichment frameworks and evaluations methodology for improving automatic enrichment¹².

In general EDM can be seen as an anchor to which various finer-grained models can be attached, ensuring their interoperability at a semantic level. For instance EDM has been aligned to CIDOC-CRM¹³ in its definition of an event-centric model (Isaac, A. (2013)). Further alignment with CIDOC-CRM has been done in the EDM–FRBRoo Application Profile Task Force (EFAP-TF) which defined a mapping of FRBRoo and EDM based on the work of the CIDOC CRM working group (Doerr, M., Gradmann, S., LeBoeuf, P., Aalberg, T., Bailly, R. & Olensky, M. (2013)). These alignments allow the definition of adequate applications profiles that enable the transition from one model to another without hindering the interoperability of the data. In a similar way,

⁹ <http://labs.europeana.eu/api/linked-open-data/introduction/>

¹⁰ <http://pro.europeana.eu/ese-documentation>

¹¹ EuropeanaTech community at <http://pro.europeana.eu/get-involved/europeana-tech>

¹² Task Force on evaluation and enrichment <http://pro.europeana.eu/get-involved/europeana-tech/europeanatech-task-forces/evaluation-and-enrichments>

¹³ <http://www.cidoc-crm.org/>

EDM re-uses properties from the EBUcore model for defining its technical metadata as explained further in section 3.4.

Re-using existing vocabularies and defining mappings with existing standards is crucial for EDM to remain a framework reusable by the cultural heritage domain (Charles V, Isaac, A. (2012)).

EDM does not require changes in the local approaches, although any changes in local practices that increase the cross-domain usefulness of the data is encouraged, such as the use of publicly accessible vocabularies (for persons, places, subjects etc.).

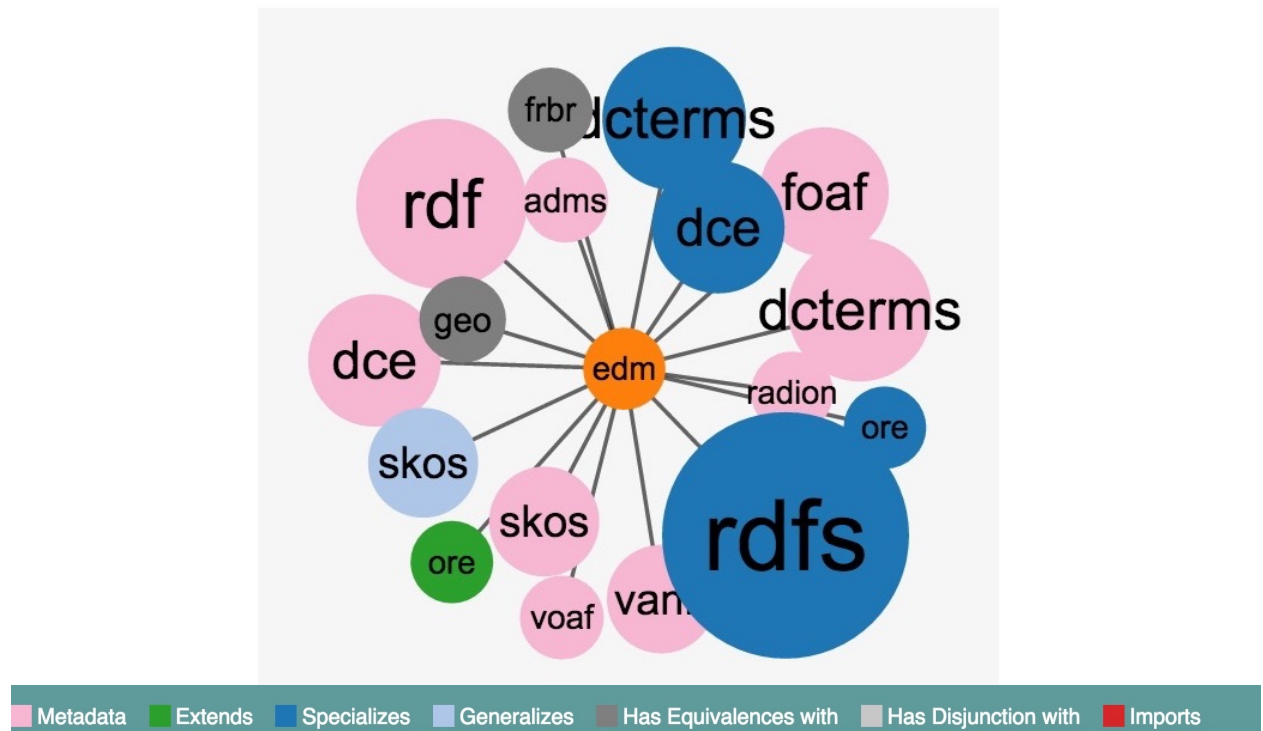


Figure 3: Re-use of different vocabularies in EDM from <http://lov.okfn.org/dataset/lov/vocabs/edm>

EDM is developed in a way that it is optimal for re-use. As EDM becomes more widely known outside Europeana and its family of projects, other international initiatives started to use the model in their own applications. The collaboration between Europeana and these institutions is important as it pushes Europeana to develop EDM in a way it can be re-used by others. That's why Europeana tries to respect the principle of minimal ontological commitment: EDM shouldn't be constrained in a way that would prevent others to re-use it.

The Digital Public Library of America (DPLA)¹⁴ for instance re-uses EDM as the core of its model and continues since then to collaborate with Europeana. Definitions of properties were for instance made more generic to allow their use by DPLA¹⁵. The most recent collaboration involves the re-definition of the *edm:rights* property so that it can be used also in the context of DPLA (DPLA (2015)).

This re-utilisation of the model oversee generates discussions in various groups working on technologies around the Semantic Web. Europeana is currently involved in a DCMI task group working on RDF Application Profiles and uses EDM as one case study for this particular work.

¹⁴ <http://dp.la/>

¹⁵ The definitions of *edm:country*, *edm:provider* and *edm:dataProvider* were generalised for re-use purposes.

2.2 Maintaining the Europeana Data Model as an open, living standard

Another condition to the re-use of EDM is the way it is maintained by the Europeana Foundation. Sharing vocabularies, the way Europeana does with EDM, implies that the model is well documented and openly shared¹⁶.

Best practices in terms of publication of vocabularies on the Web implies that a human readable as well as machine-readable documentation are available.

Different documentation is produced for the different profiles of EDM:

- The EDM model as used by the data providers for data delivery. This documentation includes guidance material relating to the data ingestion workflow made public through the Europeana Professional Website (European Data Model (EDM). Version 2.2 (2014)).
- Lots of efforts have been put in making the EDM documentation the most clear and useful for data provider. For this purpose the section of the Europeana Professional Website about “Data Guidelines” has been completely re-organized when the new site was launched in February 2015¹⁷.
- The EDM model for data enriched, stored and published by Europeana (containing the results of data enrichment for instance) in its API or Linked Data service. The documentation is available on Europeana Labs¹⁸.
- Specific EDM profiles supporting domain specific data or specific functional requirements.

To support these different uses of the model, Europeana maintains the EDM schema in a transparent way: it is accessible at any time in its most recent version on Github¹⁹. Any updates on the schema are systematically documented and made public on the Europeana Professional Website.

The Europeana Data Model is maintained in a controlled and transparent environment specific to the way standards are maintained. While being controlled, the development of EDM is done as openly as possible, to facilitate the exchange and discussions among the community of experts. This flexibility gives the insurance that the needs and requirements from cultural heritage institutions are covered.

¹⁶ <http://www.w3.org/TR/dwbp/#dataVocabularies>

¹⁷ <http://pro.europeana.eu/share-your-data/data-guidelines>

¹⁸ <http://labs.europeana.eu/api/linked-open-data/data-structure/>

¹⁹ <https://github.com/europeana/corelib/tree/master/corelib-edm-definitions/src/main/resources/eu>

<p>The <i>EDM Definition</i> is the formal specification of the classes and properties that could be used in Europeana. This overview provides an idea of the frequency of updates of the model. Note that the schema is updated more frequently than the documentation. An update of the main documentation usually covers more than one change at once.</p>	
<p>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.</p>	
<p>Updates contained in version 5.2.3 dated 24/02/2012</p>	<ul style="list-style-type: none"> - changes resulting from the process of moving from theory to implementation internally e.g. addition of the <i>ProvidedCHO</i> class - changes needed for on-going Europeana developments and initiatives e.g. the addition of <i>edm:ugc</i> for projects gathering objects from users - changes resulting from requirements put forward by other projects e.g. introduction of “3D” as an <i>edm:type</i> for CARARE.
<p>Updates contained in version 5.2.4 dated 14/07/2013.</p>	<ul style="list-style-type: none"> - ongoing amendments and additions as implementation progressed technically e.g. addition of <i>edm:europaProxy</i>, <i>edm:begin</i>, <i>edm:end</i>, <i>edm:preview</i> - 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 needed the definitions to be less Europeana-specific (Country, DataProvider and Provider) - adjustments recommended by Task Forces e.g. <i>edm:isNextInSequence</i> 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.
<p>Updates contained in version 5.2.5 dated 22/05/2014.</p>	<p>This is the current version.</p> <ul style="list-style-type: none"> - <i>edm:datasetName</i> property added and <i>collectionName</i> deprecated - <i>dcat:Dataset</i> class added in “Relevant classes...” section and “dcat” added as a namespace

Figure 4: Updates made on the EDM Definitions from report D5.2: Up-to-date Guidelines²⁰

²⁰

http://pro.europeana.eu/files//EuropeanaPro/Europeana%20Professional/Projects/Project_list/Europeana_Version2/Deliverables?preview=Ev2+D5.2+Updated+EDM+documentation.pdf

2.3. The role of EDM in the context of the Semantic Web

EDM is for Europeana the main foundation for adopting a Linked Open Data vision for its data services. The model now encourages providers to send richer and networked metadata. The support for contextual resources for instance encourages providers to send URIs from linked open vocabularies that will allow Europeana to get richer data.

EDM has also been identified as a suitable model to publish Linked Open Data .The European Library for instance has based its Linked Data model on EDM²¹.

Another interesting aspect is the re-use of EDM as an example of teaching material for information management professionals. The University of Nevada, Las Vegas, for instance, has chosen EDM for training its academic library staff²².

The advantages of EDM are

- the presence of a large documentation about the model and the existence of many use cases,
- the re-use of existing standards such as Dublin Core and Friend Of A Friend (FOAF) that can be also used as an introduction to these other standards,
- the introduction of other concepts relevant to the Semantic Web technologies such as data mapping, URI assignment and controlled vocabularies.

As detailed more in depth in section 3, EDM is a flexible model that can be extended to cover more specific data requirements. The adoption of EDM by more data providers also supports the increased production of mappings, applications profiles and extensions of EDM. These different types of interoperability patterns define various types of rules, including cardinality constraints, functional requirements but also some more general data quality rules. More generally they allow the production of different data “flavours” that need to be managed. The rules behind these flavours can be very complex and their correct interpretation and implementation can be a challenge. It is therefore crucial that they are appropriately documented in order to be readable and interpretable both by humans and machines. The definition of those rules in machine-readable formats is also a condition to the distribution and the re-use of metadata as Linked Data, since they provide a way to validate the data against specific requirements.

Europeana currently works within the DCMI RDF Application Profiles Task Group (RDF-AP)²³ to develop a set of recommendations regarding the proper creation of data models, in particular the proper re-use of existing data vocabularies based on use cases such as the Europeana Data Model. In the context of Web technology, RDF is the first candidate the group is looking at. It is a natural candidate to formalise rules that would enable the validity of the data, that would also provide further provenance and more important would facilitate the sharing and re-use of metadata schemas.

²¹ <http://www.theeuropeanlibrary.org/tel4/access/data/opendata>

²² <http://pro.europeana.eu/learningcontext-edm/>

²³ http://wiki.dublincore.org/index.php/RDF_Application_Profiles/

3. Evolution and model refinements for EDM

EDM is not a static model. As Europeana develops new data features, it refines and extends its model to accommodate new requirements. These extensions of the model allow Europeana to describe metadata at various levels of granularity (Charles, V. & Olensky, M. (2014)).

Europeana refines its model in two ways:

- It creates new set of guidelines or rules applicable to existing classes and properties to specify their use in a given context. It means that the property or class being refined will be used in a narrower, but still compatible sense compared to the already existing properties. Such set of guidelines is also called an Application Profile.
- It creates extensions to the EDM by adding new classes and properties to the original EDM specifications. The new elements sets are declared as specializations of the more generic properties. The specialization of the EDM classes and properties is done by using constructs from RDFS Schema:
 - `rdfs:subClassOf` to state that all the instances of one class are instances of a more general one
 - `rdfs:subPropertyOf` to state that all resources related by one property are also related by a more general one.

This principle of specialization, as shown in Figure 4, allows the co-existence between a generic level (the EDM classes and properties) and a specific level (DM2E classes in the example below). In Figure, *dm2e:writer* is a specialization of the more generic property *dc:creator*.

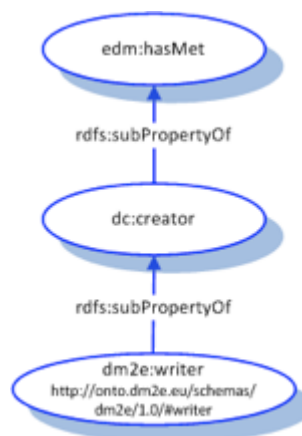


Figure 4: Principle of ontology specialization based on the RDFS properties. This figure is based on some properties available in the Europeana Data Model.

Europeana attempts to not make EDM too complex by offering two levels of formalization:

- a generic level that provide simple semantics.
- a specific level where properties and classes have been specialized. These two levels allow the re-use of EDM by different audiences and allow both complex tasks such as reasoning and more simple operations.

In addition, EDM re-uses as much as possible existing vocabularies from the community in order to increase the interoperability and also to reduce redundancy with existing vocabularies, as described in section 2.1.

The following sections are examples of specialization and extensions created by Europeana. Each extension has a table detailing the motivation or the requirements behind its creation and the vocabularies re-used in EDM to support these requirements. Further details on the classes and properties and how they are defined in EDM can be found at <https://github.com/europeana/corelib/wiki/EDMObjectTemplatesProviders>.

3.1 Profiles for representing datasets and organizations

In addition to the management of data about cultural heritage objects, Europeana also needs to manage information about organizations and the datasets sent by these organizations. Europeana created two extensions to support information about these two entities in a controlled way.

Organization profile

Motivation	<p>An organization in EDM is structured around three classes:</p> <ul style="list-style-type: none"> • An Organization is represented as a <i>foaf:Organization</i> class (a subclass of <i>edm:Agent</i>) and has a “hasAddress” property with the range of an “Address” class, which contains all the details of the postal address. An Organization also has “mainContact” and “technicalContact” properties with the range of a “Person” class (sub-property of <i>edm:Agent</i>) • A Person will be represented as a <i>foaf:Person</i> class (a sub:class of <i>edm:Agent</i>) and will have a “hasAddress” property with the range of an “Address” class, which contains all details of the postal address. • An Address will be represented using <i>vcard:hasAddress</i> syntax (EDM organization profile (2014))
Vocabularies re-used	<p>Friend-Of-A-Friend (FOAF) http://xmlns.com/foaf/0.1/ vCard http://www.w3.org/2006/vcard/ns# Creation of additional properties in the EDM namespace such as <i>edm:organizationSector</i>.</p>

Dataset profile

Motivation	The dataset profile defines the elements that will be used to represent a dataset ingested by Europeana (EDM dataset profile (2014)). In the context of Europeana, a dataset is a collection of data from a given organization. A Europeana dataset can be about a certain topic, originate from a certain source or process and is aggregated by a certain custodian.
Vocabularies re-used	Data Catalog Vocabulary (DCAT) http://www.w3.org/TR/vocab-dcat/ Asset Description Metadata Schema (ADMS) http://www.w3.org/ns/adms#/

The organization and dataset profiles are currently used in Europeana providers and datasets API²⁴. API users can retrieve information about the institutions and datasets currently represented in Europeana. The profiles have however not be made available to data providers.

3.3 Profile for representing collections

In the aggregation model used to submit data to Europeana the totality of a curated collection submitted by a data provider may be difficult to find in a large, aggregated dataset. “Collection structures provide the organizational and intellectual context important to researchers, and collection descriptions provide information needed by users for interpreting the relevance and significance of individual items for their purposes” (Wickett, K.M., Isaac, A., Fenlon, K., Doerr, M., Meghini, C., Palmer, C.L & Jett, J. (2013)). The collection profile has been defined to enable the description of collections in the context of the EDM (EDM collection profile (2014)). It will allow data providers to supply curated descriptions of the identifiable collections they have submitted to Europeana.

Europeana does currently not implement this profile but use cases are collected from data providers (such as the archeology domain) or from other communities such as researchers in the context of Europeana Research²⁵

Motivation	A collection in the context of Europeana is a group of objects gathered together for some intellectual, artistic, or curatorial purpose. In the context of this profile the Collection class represents the original collection of cultural heritage objects (probably physical objects, but also born digital where appropriate).
Vocabularies re-used	Additional EDM classes and properties were created such as the main <i>edm:Collection</i> class. http://www.w3.org/TR/vocab-dcat/ Dublin Core Collection Description Terms http://purl.org/cld/terms/

²⁴ <http://labs.europeana.eu/api/provider/>

²⁵ <http://research.europeana.eu/>

3.4 Profile for representing technical metadata

Europeana currently defines its Content-Re-use Framework that promotes the re-use of the digital representations provided as part of the delivered metadata (EDM profile for technical metadata (2015)). In order to facilitate the re-use of these materials, Europeana builds facets and filters based on the technical metadata of these digital representations (described in EDM by the class *edm:WebResource*. EDM needed therefore to be extended to fit the five media types currently supported by Europeana, namely: Sound, Video, Text, Image and 3D objects.

Motivation	This profile lists the properties that will apply to the <i>WebResource</i> class and an additional class that were defined to support such functionality.
Vocabularies re-used	EBUcore http://www.ebu.ch/metadata/ontologies/ebucore/

3.5 Extensions for complex rights

The new developments around the Content Re-Use Framework provide new requirements for representing rights statements²⁶ for digital representations of cultural heritage objects. Europeana needs indeed a way to structure rights information for different types of resources. This extension of EDM focuses on the creation of “complex” values for the existing property *edm:rights*. The indication of an identifier of a rights statement (e.g. CC-BY) in a unique field (*edm:rights* on the EDM *ore:Aggregation* resource) no longer covers more complex requirements. This profile envisions that different access and re-use conditions can be provided for different views of a cultural object. It therefore allows the representation of individual views with specific rights statements.

An additional requirement came as the result from public-private partnership that requires the specifications of contractual restrictions. For this given type of licence, Europeana needed a new property for storing the end date of a copyright claim, license or other type of restriction on re-use as expressed in the *edm:rights* field of the metadata records.

Motivation	The profile introduces a new <i>cc:Licence</i> class to which additional properties are added such as <i>cc:deprecatedOn</i> to declare an expiration date.
Vocabularies re-used	Creative Commons Rights Expression Language (ccREL) http://creativecommons.org/ns# ODRL http://www.w3.org/ns/odrl/2/

²⁶ <http://pro.europeana.eu/blogpost/extending-the-europeana-licensing-framework>

Europeana continues its work on representing complex rights as part of the collaboration with the Digital Public Library of America (DPLA) aiming at building a common technical infrastructure for rights statements and an extensible framework to host the rights statements at rightsstatements.org. This work requires the definition of new classes and properties to express data for the statements that may be propagated into the EDM specifications²⁷.

3.6 Future work

The Europeana Strategy for 2015-2020 lists plans to further develop EDM. One important item is the development of new services for end-users that will improve their engagement with cultural heritage data. Crowdsourcing is one way to engage end-users and Europeana is currently looking at implementing it under the form of annotation. The EDM will be extended based on the recommendations from the Web Annotation Data Model²⁸. This extension will allow the support of scenarios such as semantic tagging or crowdsourced metadata enrichment.

4. Re-use and extensions of the EDM model

The Europeana Data Model is not only the model used internally by Europeana to describe cultural heritage objects, it also meant to be re-used by the broader cultural heritage domain.

In some cases the EDM as defined by Europeana in its specifications cannot represent the semantics of providers' metadata with sufficient details. Those details matter even more when EDM is re-used in a domain specific context. Data providers need to represent their metadata at different levels of granularity. Cultural heritage institutions are therefore developing extensions in the same way Europeana is doing it to cover its own needs. It is important to note that the elements in these extensions can almost always be mapped to a more generic property in EDM, following the specialization principle mentioned in Section 3. This mechanism allows metadata described for specific applications to remain interoperable with more generic applications such as the main Europeana.eu portal.

The extensions of EDM are for most part based on existing vocabularies but in some cases the institution or project have decided to create their own properties and have therefore declared a new namespace for it.

The examples below present some of the extensions that have been developed within Europeana affiliated projects and organizations, but also by institutions outside Europeana 's direct Network.

4.1 EDM extensions in the Europeana Network

Europeana affiliated projects have started to re-use EDM in their applications rapidly. The Europeana Fashion project (Vanstappen, H. (2012)), the German Digital Library²⁹ (DDB) or the Polymath Virtual Library³⁰ have build their data modeling on EDM and integrated new elements or elements from other metadata standards for addressing their specific needs.

²⁷ <http://pro.europeana.eu/blogpost/developing-and-implementing-a-technical-framework-for-interopera>

²⁸ <http://www.w3.org/TR/annotation-model/>

²⁹ <http://dcevents.dublincore.org/IntConf/dc-2014/paper/download/231/300>

³⁰ <http://pro.europeana.eu/share-your-data/data-guidelines/edm-case-studies/polymath-edm>

We will highlight here two extensions addressing the needs of different communities but also supporting different types of requirements.

DM2E specialization of EDM for manuscripts

The DM2E model allows the description of rich metadata for manuscripts, answering the needs of scholars without hindering the interoperability with Europeana as the DM2E model is built on top of the Europeana one (DM2E (2014)). As for other domains, the data from the manuscripts domain aggregated by DM2E are represented with various schemas (e.g. TEI, METS/MODS, MARC21 or provider-specific schema) that describe the data at a different level of granularity than EDM. DM2E have therefore created specializations of EDM for:

- supporting the description of manuscript specific information (for instance supporting the DM2E model introduces the property *dm2e:incipit*, which is used for representing the opening words of a manuscript.
- describing more granular information for certain type of entities (for instance different types of creators such as writer, author...)
- representing provenance information
- supporting different levels of description (objects or parts of an object like a book, a manuscript or a page).
- representing annotable resources such as image, like PNG or JPEG, or text.

The DM2E model re-uses the following namespaces (many of which "inherited" from EDM):

bibo	http://purl.org/ontology/bibo/
crm	http://www.cidoc-crm.org/cidoc-crm/
dc	http://purl.org/dc/elements/1.1/
dcterms	http://purl.org/dc/terms/
edm	http://www.europeana.eu/schemas/edm/
fabio	http://purl.org/spar/fabio/
foaf	http://xmlns.com/foaf/0.1/
ore	http://www.openarchives.org/ore/terms/
pro	http://purl.org/spar/pro/
rdaGr2	http://rdvocab.info/ElementsGr2/

EDM profile for Sound

A task force as part of the Europeana Sounds project developed an EDM profile for supporting characteristics of sounds objects and the metadata describing them (Charles, V. et al. (2015)). The task force identified the specificities of sounds objects metadata and analyzed the needs that could be met using EDM in its current definition.

The extensions have been taken from existing open data models, namely EBUCore, the Music Ontology³¹ and Dublin Core. These new properties have been declared as specializations of the existing ones.

The EDM profile for Sound supports the following requirements:

- Distinction of the master version among the digital representations available for a Cultural Heritage Object (CHO)
- Duration in sounds CHOs and WebResources
- Track information for sound CHOs and WebResources
- Hierarchical relationships and collections.

The development of such extensions is beneficial to Europeana. It allows the identification of new domain specific requirements that can be implemented in Europeana's own service without hindering the interoperability between data services. The development of the model itself remains a collaborative effort.

4.2 EDM extensions outside the Europeana network

EDM is also been extended and re-use outside of Europeana direct Network. The Smithsonian American Art Museum (SAAM) uses an extension of EDM to publish data on the Web (Szekely, P., Knoblock, C., Yang, F., Zhu, X., Fink, E., Allen, R., & Goodlander, G. (2013)). Using EDM was adopted as a way for to maximise the interoperability with existing museum LOD datasets.

In the same way the Linking History project uses EDM to model the data resulting from a research project in which students linked people and places names of Australia's capital city, Canberra to cultural heritage objects. EDM was also chosen as way to generate Linked Data for this project.

Lastly EDM is also used as a foundation of the data model used by the Digital Public Library of America. The current version re-uses EDM specific classes and properties as well as definitions from the following sources:

- Resource Description Framework (RDF) and the RDF Schema (RDFS)
- OAI Object Reuse and Exchange (ORE)
- Dublin Core namespaces (dc elements, dcterms, and dcmitype)
- The Basic Geo (WGS84 lat/long) Vocabulary

³¹ <http://musicontology.com/>

Conclusion

Five years ago, Europeana's first white paper, by Prof. Stefan Gradmann, advocated the design of an advanced data model, re-using existing pieces of web technology (Gradmann, S. (2010)). The Europeana Data Model has now become a reality we can work with, together with our data partners and data re-users.

Since its original release, EDM has been implemented in the infrastructure and entire portfolio of Europeana services: we ingest, store, enrich and exchange data following a richer, more semantic approach. In the process EDM has been updated or "profiled" to enable new functions, such as representing hierarchical objects. This work continues, for example as Europeana prepares to handle more data enrichment, including coming from user annotations.

As a matter of fact the ongoing maintenance of EDM is an open process, not a one-directional movement. EDM would not exist without the contributions of our data and academic partners! In the past years, task forces within the EuropeanaTech community have played a crucial role; for example for hierarchical objects, sounds, and compliance with other modeling approaches like FRBR.

We do our best to keep the model well documented and flexible. EDM is built for re-use, if just because we try to re-use ourselves as much as possible of existing elements from well-established vocabularies for its design.

Also, in a true linked data fashion, EDM "profiles" can be developed without Europeana having to update the core model anymore. It is perfectly alright to have elements attached to the model, which do not have a corresponding implementation in Europeana's core platform.

This flexible approach has made it easier to extend EDM to meet the data needs of specific domain aggregators, like Europeana Sounds³², and addresses the requirements of new data services and enrichment in Europeana's main platform. EDM is now used by Europeana and several other cultural aggregators, such as DPLA and DDB.

We foresee more extensions and profiles in the future. One challenge then will be how to handle the resulting diversity of profiles in a seamless way. Exploiting the data expressed with these profiles across different systems still requires work. But thanks to the basic principles underpinning EDM, it is no longer impossible to realize the vision where the design of data models is decentralized and tailored to specific applications, while the data created and exchanged with them still will form together a vast, semantically interoperable knowledge environment.

³² <http://www.europeanasounds.eu/>

References

- Charles, V. et al. (2013). Recommendations for the representation of hierarchical objects in Europeana. Retrieved May 25, 2015 from <http://pro.europeana.eu/documents/468623/4a6eb2ec-4cc6-48b1-8824-92a1e564a279>
- Charles, V. et al. (2015). EDM profile for Sound. Retrieved May 25, 2015 from http://pro.europeana.eu/files/Europeana_Professional/EuropeanaTech/EuropeanaTech_taskforce_s/EDMSound//TF_Report_EDM_Profile_Sound_301214.pdf
- Charles, V. & Olensky, M. (2014). Report on Task force on EDM mappings, refinements and extensions. Retrieved May 25, 2015 from http://pro.europeana.eu/files/Europeana_Professional/EuropeanaTech/EuropeanaTech_taskforce_s/Mapping_Refinement_Extension//EDM%20%20Mapping%20refinement%20extension%20Report.pdf
- Charles, V, Isaac, A, Freire, N. (2014). Linking Libraries in The European Library and Europeana. [IFLA 2014 Satellite Meeting Linked Data in Libraries.](#)
- Charles V, Isaac, A. (2012). Europeana and (many) linked open vocabularies. [LOV Symposium: Linking and Opening Vocabularies](#), Universidad Carlos III de Madrid, Spain, June 18, 2012. [Slides](#)
- DM2E (2014). DM2E Model V 1.2 Specification http://dm2e.eu/files/DM2E_Model_V1.2.pdf
- Doerr, M., Gradmann, S., LeBoeuf, P., Aalberg, T., Bailly, R. & Olensky, M. (2013). Final Report on EDM – FRBRoo Application Profile Task Force. Retrieved May 25, 2015 from http://pro.europeana.eu/files/Europeana_Professional/EuropeanaTech/EuropeanaTech_taskforce_s/EDM_FRBRoo//TaskfoApplication%20Profile%20EDM-FRBRoo.pdf
- DPLA (2015). <http://dp.la/info/developers/map>
- European Data Model (EDM). Version 2.2 (2014). http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Mapping_Guidelines_v2.2.pdf
- EDM collection profile (2014). Retrieved May 25, 2015 from http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_profiles/EDM%20Collection%20Profile.doc
- EDM dataset profile (2014). Retrieved May 25, 2015 from http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_profiles/Dataset%20profile.pdf
- EDM organization profile (2014). Retrieved May 25, 2015 from http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_profiles/Organisation%20profile.pdf

EDM profile for technical metadata (2015). Retrieved May 25, 2015 from http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_profiles/Technical%20Metadata%20Properties_20150217.docx

Gradmann, S. (2010). Knowledge = Information in Context: on the Importance of Semantic Contextualisation in Europeana. Retrieved from <http://pro.europeana.eu/publication/knowledgeinformation-in-context>

Isaac, A. (2013). Europeana Data Model Primer. Retrieved May 25, 2015 from http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf

Stiller, Isaac & Petras (eds.) (2014): EuropeanaTech Task Force on a Multilingual and Semantic Enrichment Strategy: final report. Available: http://pro.europeana.eu/files/Europeana_Professional/EuropeanaTech/EuropeanaTech_taskforces/MultilingualSemanticEnrichment//Multilingual%20Semantic%20Enrichment%20report.pdf

Szekely, P., Knoblock, C., Yang, F., Zhu, X., Fink, E., Allen, R., & Goodlander, G. (2013). Connecting the Smithsonian American Art Museum to the Linked Data Cloud. In P. Cimiano, O. Corcho, V. Presutti, L. Hollink, & S. Rudolph (Eds.), *Lecture Notes in Computer Science. The Semantic Web: Semantics and Big Data* (pp. 593–607). Berlin, Heidelberg: Springer.

Vanstappen, H. (2012). Europeana Data Model – Fashion Profile (EDM-fp). Retrieved May 25, 2015 from http://pro.europeana.eu/files/Europeana_Professional/EuropeanaTech/EuropeanaTech_taskforces/Mapping_Refinement_Extension/Case_studies/EuropeanaFashion/EDM-FP%20specification.docx

Wickett, K.M., Isaac, A., Fenlon, K., Doerr, M., Meghini, C., Palmer, C.L & Jett, J. (2013). Modeling cultural collections for digital aggregation and exchange environments. Retrieved May 25, 2015 from <http://hdl.handle.net/2142/45860>