



europaana  
think culture

---

# Europeana Dataset Profile

April 2016

**Europeana**



co-funded by the European Union

## 1 Introduction and Scope

This profile defines the elements that will be used to represent datasets ingested by Europeana. This profile will be mainly used to disseminate dataset level information via the API.

It is envisaged that the value for many of the properties described below will originate from various part of the Europeana's workflow such as the Customer Relationship Management system (SugarCRM) or the Mongo database.

Some of these properties will be mapped from the CRM fields to EDM classes accordingly or will be generated from the system.

The profile is based on the dataset description elements defined in:

- Data Catalog Vocabulary (DCAT) <http://www.w3.org/TR/vocab-dcat/>
- The VOID vocabulary <http://www.w3.org/TR/void/>
- Asset Description Metadata Schema (ADMS) <http://www.w3.org/TR/2013/NOTE-vocab-adms-20130801/>

## 2 Terminology

The following terms will be used in this document.

**Dataset:** A dataset for Europeana in an Information Package (as defined by OAIS<sup>1</sup>) or a collection of data. A Europeana dataset can be about a certain topic, originate from a certain source or process and is aggregated by a certain custodian<sup>2</sup>.

In some cases, a Europeana dataset can be the representation of one existing collection held by an institution. In other cases the correspondence between datasets and collections won't be one-to-one.

**Provider:** Organization that collects, formats and manages data from Data Providers before submitting data to Europeana.eu

**Data Provider:** Organization that submits data to Europeana.eu. Data providers may submit data directly to Europeana but are more likely to use the services of a Provider.

**Contact Person:** the person recorded as being the contact for Europeana.

---

<sup>1</sup> [http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=57284](http://www.iso.org/iso/catalogue_detail.htm?csnumber=57284)

<sup>2</sup> VOID Dataset definition at <https://www.w3.org/TR/void/#dataset>

### 3 Overview of the Dataset profile

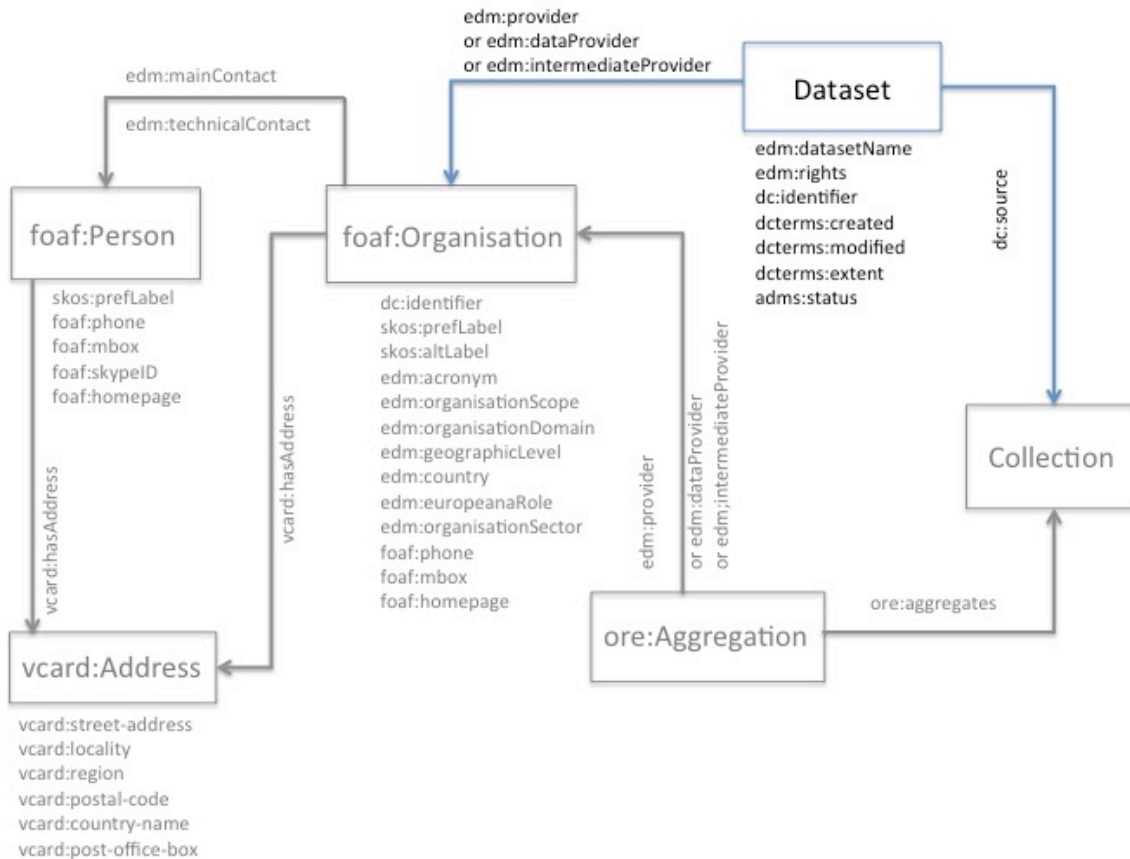


Fig. 1 Overview of Dataset class, its properties and relationships

### 4 Description Elements

Property	Label	Proposed definition	Cardinality	Value type
edm:datasetName <sup>3</sup>	Dataset Name	This property holds the identifier given to the dataset in Europeana.	<b>Mandatory</b> 1....1	Literal
edm:dataProvider	Data Provider	The name or identifier of a data provider of the dataset (i.e. an organization providing the dataset to an aggregator). Note that a dataset can have multiple	<b>Mandatory</b> 1....n	Literal or ref

<sup>3</sup> The edm:collectionName property originally recommended, is deprecated and replaced by edm:datasetName.

		data providers.		
edm:intermediateProvider		The name or identifier of the intermediate organization that selects, collates, or curates data from a Data Provider that is then aggregated by a Provider from which Europeana harvests. The Intermediate Provider must be distinct from both the Data Provider and the Provider in the data supply chain.	<b>Optional</b> 0....n	Literal or ref
edm:provider	Provider	The name or identifier of the provider of the dataset (i.e. the organization providing datasets directly to Europeana).	<b>Mandatory</b> 1....1	Literal or ref
edm:rights	Rights	This property defines the rights applied to digital objects.	<b>Recommended</b> 1....n	Reference
dc:identifier	Identifier	An unambiguous reference to the dataset identifier within the context of the Europeana datasets database. The value is derived from edm:CollectionName.	<b>Mandatory</b> 0....n	Literal
dc:source	Source	A related resource from which the dataset is derived (for instance a dataset is derived from a formal collection).	<b>Optional</b> 0....n	Reference
dcterms:created	Submission date	Date of creation of the dataset in the Europeana database	<b>Mandatory</b> 1....1	Literal
dcterms:extent	Dataset extent	The number of ProvidedCHOs in the dataset	<b>Optional</b> 0....n	Literal
dcterms:modified	Modification date	Date of the last publication update of the dataset	<b>Mandatory</b> 0....n	Literal
adms:status	Europeana Status	The value should come from a controlled list. Current list is OAI-PMH testing- OAI-PMH test sent to organisation-Ready for harvesting- Mapping and normalisation- Ready for replication-Ongoing scheduled updates-	<b>Mandatory</b> 1....1	Literal

		Ingestion complete- Disabled and replace		
--	--	--	--	--

## 5 Future work

- Publisher : Europeana publication statement
- IsPartOf relation : To link a subset to a full dataset.
- Digital Object Type: Type of the WebResource
- Enrichment statistics: Type of enrichment performed on the dataset

### Document History

Version	Author	Updated	Comments
V0.1	Valentine Charles	21/10/2013	First version
V0.2	Valentine Charles	20/01/2014	<ul style="list-style-type: none"> <li>• dcterms extent changed into repeatable.</li> <li>• dc:identifier removed from future work and added to the profile.</li> </ul>
V0.3	Valentine Charles	04/04/2016	<ul style="list-style-type: none"> <li>• edm:collectionName deprecated and replaced by edm:datasetName.</li> <li>• edm:intermediateProvider added to the profile.</li> </ul>

## 6 Acknowledgements

Thanks are due to the following for their contribution to these guidelines: Valentine Charles, Robina Clayphan, Antoine Isaac, Dimitra Atsidis, Marie-Claire Dangerfield and Cecile Devarenne.