→ **URIs in the context of the Europeana Data Model, v1.0**

1.   Introduction

Within the Europeana Data Model (EDM) a cultural heritage object consists of several sections of metadata represented by core classes and (optional) contextual classes.[1] Each of these sections is considered a "resource". For these resources to be linked together within a RDF/XML document, they need to have a unique identifier in the form of a URI-reference, which is defined in the "rdf:about" statement of the resource.

The URI-references defined in the "rdf:about" statement are typically local identifiers, but can also point to an external resource (HTTP URI). The URI-reference within the "rdf:about" statement of edm:WebResource should however always be a valid HTTP URI, which means that it references a web resource external to the metadata submitted to Europeana.

The use of HTTP URIs within an RDF/XML document can be a bit confusing, as a HTTP URI looks like a web address. Within the RDF/XML document they are just names for entities. When using dereferenceable URIs however, it may be expected that there is something on the web at that address.[2]

2.   What is a URI?

The term URI stands for Uniform Resource Identifier and consist of four basic consecutive parts, two of which are optional:
- a scheme name;
- a hierarchical part;
- a query (optional);
- and a fragment (optional).[3]

As the name implies, a URI identifies a resource. This resource can either be a physical thing, existing in the real world, or a digital resource like a website. The most common form of URI is a URL or Uniform Resource Locator, informally referred to as a web address. Another common form of URI is the URN or Uniform Resource Name, which is often used to identify physical objects, such as books (e.g. urn:isbn:0-486-27557-4).

---

[1] See

http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation//EDM_Mapping_Guidelines_v2.2.pdf

[2] See https://github.com/JoshData/rdfabout/blob/gh-pages/intro-to-rdf.md#

[3] Generic syntax: `<scheme name> : <hierarchical part> [ ? <query> ] [ # <fragment> ]`, see

http://tools.ietf.org/html/std66 and http://tools.ietf.org/html/rfc3986 .

3.   What is a URI-reference?

A URI-reference can be either absolute or relative.[4] In other words, a URI-reference is either a complete URI or just a part of it. These URI-references, taken from Wikipedia[5], show some intuitive examples:

- https://en.wikipedia.org/wiki/Uniform_resource_identifier#Examples_of_URI_references
- https://example.org/absolute/URI/with/absolute/path/to/resource.txt
- //example.org/scheme-relative/URI/with/absolute/path/to/resource.txt
- /relative/URI/with/absolute/path/to/resource.txt
- relative/path/to/resource.txt
- ../../../resource.txt
- ./resource.txt#frag01
- resource.txt
- #frag01
- (empty string)

4.   Recommendations

Recommendations on choosing the right "rdf:about" values within EDM:

- ✓ For edm:ProvidedCHO, create a URI-reference based on a unique internal identifier (e.g. dc:identifier) starting with "#" (e.g. "#UEDIN:214"); or use the uniform resource name (URN) of the object (e.g. "urn:isbn:0-486-27557-4"); it is very important that the provided URI-references are persistent over time, in order to avoid broken links or duplicated data when records are updated in the future[6];
- ✓ For edm:WebResource always use the hyperlink to the resource (image, text, video, html page, etc.);
- ✓ For ore:Aggregation use the URI of the website about the provided CHO; when there is no specific page describing the CHO or more CHOs are described on one page, make sure the URI is unique e.g. by adding the identifier of the object (e.g. "urn:isbn:0-486-27557-4:aggregation");
- ✓ For the contextual classes (edm:Agent, edm:Place, edm:TimeSpan, skos:Concept and cc:Licence) use the URI from an external vocabulary; or create a URI-reference based on a unique local identifier by adding as prefix "#";
- ✓ Make sure that your HTTP URIs (i.e. web resources) are dereferenceable, so that RDF clients can access that page and get additional information.

---

[4] http://www.w3.org/TR/2004/REC-rdf-primer-20040210/#identifiers

[5] http://en.wikipedia.org/wiki/Uniform_resource_identifier

[6] A detailed study on persistent URIs can be found at: http://philarcher.org/diary/2013/uripersistence/.

5. Examples

Below are two examples of how to use URIs within EDM to describe a cultural heritage object (CHO). The first example uses an local identifier for the object (UEDIN:214). The second example uses a URN in the form of a persistent identifier (urn:nbn:nl:ui:13-rvky-mu).

Example 1
Example of well-formed EDM using a unique local identifier for a CHO:

```
<edm:ProvidedCHO rdf:about="#UEDIN:214">
...
</edm:ProvidedCHO>

<edm:WebResource rdf:about="http://www.mimo-
db.eu/media/UEDIN/IMAGE/0032195c.jpg">
...
</edm:WebResource>

<ore:Aggregation rdf:about="http://www.mimo-db.eu/UEDIN/214">
    <edm:aggregatedCHO rdf:resource="#UEDIN:214"/>
    <edm:isShownBy rdf:resource="http://www.mimo-
db.eu/media/UEDIN/IMAGE/0032195c.jpg"/>
</ore:Aggregation>
```

Example 2
Example of well-formed EDM using a URN identifier for a CHO:

```
<edm:ProvidedCHO rdf:about="urn:nbn:nl:ui:13-rvky-mu">
...
</edm:ProvidedCHO>

<edm:WebResource rdf:about="http://www.persistent-
identifier.nl/?identifier=urn:nbn:nl:ui:13-rvky-mu">
...
</edm:WebResource>

<ore:Aggregation rdf:about="urn:nbn:nl:ui:13-rvky-mu:aggregation">
    <edm:aggregatedCHO rdf:resource="urn:nbn:nl:ui:13-rvky-mu"/>
    <edm:isShownAt rdf:resource="http://www.persistent-
identifier.nl/?identifier=urn:nbn:nl:ui:13-rvky-mu"/>
</ore:Aggregation>
```

6.   Transformation rdf:about of edm:ProvidedCHO

To provide a unique URI in the Europeana portal for every cultural heritage object, Europeana uses the rdf:about in edm:ProvidedCHO to create resolvable URIs that lead to the specific object in the portal. Such a URI consists of three parts:
- A base URL: http://data.europeana.eu/item/;
- followed by the Europeana identifier of the collection;
- followed by an identifier derived from rdf:about of edm:ProvidedCHO.

The original value in rdf:about or edm:ProvidedCHO will be transformed during the process if it includes any special characters (i.e. everything that is not a letter or a number). Special characters will be replaced with an underscore.

For example, the transformation of the following CHO:

```
<edm:ProvidedCHO rdf:about="AK-MAK-629"/>
```

will result in this CHO in the portal:

```
<edm:ProvidedCHO
rdf:about="http://data.europeana.eu/item/90402/AK_MAK_629"/>
```

This transformation might have as a result the occurrence of so called derived duplicates. This happens when non-identical URIs become identical after transformation. The values "AK-MAK-629", "AK*MAK*629" and "AK@MAK@629" would for example all be transformed to "AK_MAK_629". Therefore, it is an important thing to keep in mind when providing your identifiers. Providing unique identifiers without special characters will make sure this will not happen.

7.   Document history

Version	Editor	Date	Comments
v1.0	Jeroen Cichy	09/09/2015	• Final version of v1.0 incorporating previous drafts and contributions as well as comments and suggestions from colleagues and reviewers.

8.   Acknowledgements

This document was reviewed and amended by Valentine Charles, Cecile Devarenne, Nuno Freire, Antoine Isaac and Hugo Manguinhas who suggested useful improvements.