

data.europeana.eu **The Europeana Linked Open Data Pilot**

Bernhard Haslhofer
Cornell Information Science, USA
University of Vienna, Austria
bernhard.haslhofer@cornell.edu

Antoine Isaac
Europeana, The Hague, The Netherlands
Vrije Universiteit Amsterdam, The Netherlands
aisaac@few.vu.nl

Abstract

data.europeana.eu is an ongoing effort of making Europeana metadata available as Linked Open Data on the Web. It allows others to access metadata collected from Europeana data providers via standard Web technologies. The data are represented in the Europeana Data Model (EDM) and the described resources are addressable and dereferencable by their URIs. Links between Europeana resources and other resources in the Linked Data Web will enable the discovery of semantically related resources. We developed an approach that allows Europeana data providers to opt for their data to become Linked Data and converts their metadata to EDM, benefiting from Europeana efforts to link them to semantically related resources on the Web. With that approach, we produced a first Linked Data version of Europeana and published the resulting datasets on the Web. We also gained experiences with respect to EDM, HTTP URI design, and RDF store performance and report them in this paper.

Keywords: Europeana; linked data; open data; metadata mapping

1. Introduction

Linked Data is a data publishing technique that uses common Web technologies to connect related data and make them accessible on the Web. It relies mainly on identifying resources with (HTTP) Uniform Resource Identifiers (URI), and, using standards such as the Resource Description Framework (RDF), on providing data about these resources and connecting them to other resources on the Web (Berners-Lee, 2009). Soon after the first Linked Data sources went online, libraries and cultural institutions, such as the Library of Congress and the Swedish Union Catalogue, followed the Linked Data principles to publish their data. As a result, we can now access library resources and their descriptive metadata simply by dereferencing HTTP URIs, which facilitates data access and reuse.

Europeana (<http://europeana.eu>) is the European Union's flagship digital library project. It enables search and discovery in more than 17 million items by collecting metadata from approximately 1,500 cultural institutions (*data providers*) across Europe. Becoming part of the Linked Data architecture is a clear goal in Europeana (Gradmann, 2010) and with the Europeana Data Model (EDM) a suitable data model for publishing and linking Europeana metadata is currently being developed (Europeana, 2010). We believe that applying a Linked Data publishing strategy directly at Europeana not only contributes a large dataset to the community at once, but also relieves the participating data providers from implementing their own Linked Data publication infrastructure. In the past, several issues prevented the adoption of Linked Open Data (LOD) in the Europeana production system: (i) lack of metadata expressed in EDM, (ii) missing links to other sources, and, most importantly, (iii) missing data provider agreements.

As a solution, we decided to build the *Europeana Linked Data Pilot*. It is technically decoupled from the Europeana production system and allows those data providers, who want to make their data available as Linked Open Data, to opt for their metadata to be openly published on the Web. To bootstrap the adoption of the EDM, we implemented a simple procedure that converts the current Europeana data to EDM. We also include links that are generated from

Europeana's effort of enriching existing metadata records with links to semantically related resources. We have published a first sub-set of the Europeana data set, which is also available for download at <http://data.europeana.eu> and accessible via a SPARQL endpoint. Since the data publication process is transparent and the applied tools are open source, we expect input or even contributions from the community.

The major contributions of this project can be summarized as follows: we (i) provide one of the largest *open* dataset released in the libraries, archives and museums area so far and (ii) implemented the Europeana Data Model, and thereby also the Open Archives Object Reuse and Exchange (OAI-ORE) model,¹ in a large dataset. In this paper, we report on the experiences we made when transitioning from the current Europeana Semantic Elements (ESE)² to EDM and making the Linked Data pilot operational. We believe that these are important findings for others who want to adopt a similar strategy for their data, as well as for the further development of Web-based object and data aggregation models.

In the following, we first summarize related work in the area of Linked Data and library data. Then, in Section 3, we give an overview of the organizational processes and technical properties of the Europeana Linked Data Pilot. The conversion from the legacy ESE format to EDM is described in Section 4 and the lessons we learned so far are summarized in Section 5.

2. Related Work

The idea of using the Web infrastructure for making metadata accessible in an interoperable way is not new. Protocols such as OpenURL (Van de Sompel et al., 2001), OAI-PMH,³ SRU-SRW,⁴ already indicated the transition from proprietary, library-centric metadata representation and exchange mechanisms to more open, broadly applicable, and Web-based standards.

For a detailed explanation of Linked Data and a general state-of-the art survey on existing data sources and solutions we refer to (Heath et al., 2011). In the following, we give a rather representative than complete overview of Linked Data activities in the digital library and cultural heritage community.

The Library of Congress (LoC) was one of the early adopters of the Linked Data approach in the library domain (Summers et al., 2008). The service went into production (<http://id.loc.gov>) in 2009 and exposed approximately 260,000 authority records. Later the LoC extended its authority and vocabulary service by the *Thesaurus of Graphic Materials*, the *MARC Code List for Relators*, and other smaller vocabularies. The Swedish Union Catalogue (<http://libris.kb.se>) is another early Linked Data Service (Malmsten, 2008). It exposes millions of records from about 175 libraries describing various types of resources including persons, books, authors, subjects, etc. The records were made available by building a simple RDF wrapper on top of the integrated library system.

A SKOSified version of RAMEAU, which is the main subject vocabulary at the French national library, was exposed as Linked Data in 2009 (Van der Meij et al., 2010). It contains approx. 160,000 concepts including common nouns and geographic names. The concepts are interlinked with the LCSH concepts based on 60,000 manual mappings available from the MACS project (Landry, 2009). In the same year, OCLC announced the Dewey Decimal Classification (DDC) to be published as Linked Data (Panzer, 2008). At <http://dewey.info> one can access the top three levels of the DDC in eleven languages and the Abridged Edition 14 in three languages. The STW Thesaurus for Economics (<http://zbw.eu/stw>) is another important vocabulary source and publishes RDF representations of approximately 6,000 standardized subject headings and 18,000 entry terms on the Web.

¹ <http://www.openarchives.org/ore/>

² <http://version1.europeana.eu/web/guest/technical-requirements/>

³ <http://www.openarchives.org/pmh/>

⁴ <http://www.loc.gov/standards/sru/>

Another relevant project is VIAF, the Virtual International Authority File (<http://viaf.org/>). It is a joint project of more than ten national libraries, implemented and hosted by OCLC. It has the goal to match and link authority files of national libraries and then making that information available on the Web. In a similar manner, the Gemeinsame Normdatei (GND, <http://wiki.d-nb.de/display/LDS>) unifies German authority files, which are currently used for authority control and subject indexing in German, Austrian, and Swiss libraries and library networks, including personal names (PND), corporate bodies (GKD) and subject headings (SWD).

The Open Library project (<http://openlibrary.org/>) has the ambitious goal to create one Web page for every book. At the moment, it has gathered around 20 million records and provides a resource-centric architecture for manipulating its objects. Finally, the Hungarian National Library catalog (NSZL, http://nektar.oszk.hu/wiki/Semantic_web) also exposes its OPAC catalogue and authority data. For a more complete overview and further Linked Data sets in the library area we refer to the CKAN Library Linked Data packages available at <http://ckan.net/group/lld>.

3. Overall Approach

Our aim was to organize the technical processes behind the Europeana Linked Open Data Pilot as open and transparent as possible. All required ESE2EDM conversion scripts and configuration files for setting up <http://data.europeana.eu> are open source and available at <https://github.com/behase/ese2edm>. They can freely be re-used or extended.

From a technical perspective, as shown in Fig. 1, the Europeana LOD Pilot runs in parallel to the Europeana production system but is based on the same ESE-formatted metadata basis.

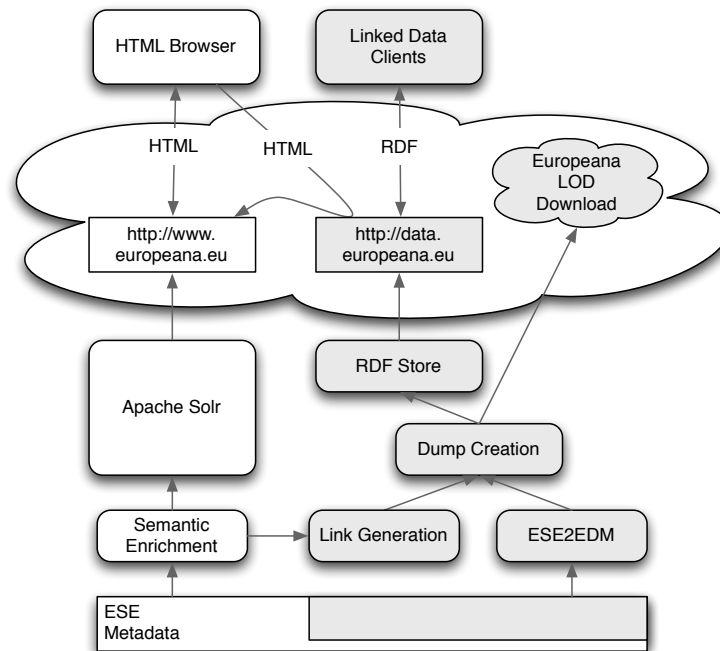


Fig. 1. data.europeana.eu - Technical Architecture

In the first step, we extract a subset of the ESE source metadata based on these providers who express the wish to become part of the pilot. We then convert this dataset into the Europeana Data Model (EDM) and assign HTTP URIs to the resulting objects (aggregations, proxies, resource maps). The mapping between ESE and EDM, the details of which will be given in the next section, are implemented in an XML stylesheet and executed via XSLT 1.0. As a result, we obtain an RDF/XML representation of each data provider's metadata.

Third, we fetch semantic annotations created at the Europeana Office prior to this project, using the Annocultor tool (<http://annocultor.eu/>). This enrichment has been obtained by matching values of selected metadata fields—e.g., Dublin Core's *dc:spatial*—with labels of resources from

controlled vocabularies.⁵ We make available around 5 million such links to the Geonames gazetteer (<http://geonames.org>) and 7 million links to an adhoc time ontology. We will subsequently include *subject* links to the GEMET thesaurus (<http://www.eionet.europa.eu/gemet>) and *person* links to DPpedia (<http://dbpedia.org>), which Europeana is creating at the time of writing.

We also link to existing LOD services maintained by our own Europeana partners. This applies to the already mentioned National Library of Hungary as well as SOCH, the Swedish culture aggregator (<http://www.ksamsok.se/>). As Europeana's mission is to connect cultural items, it must duly acknowledge already published data. This is also a way to alleviate the *URI proliferation* issue on the Web of Data. Linked data publishers should indeed try to re-use existing URIs for the entities they describe, or if not possible—in our case the data we serve is quite different from our providers' LOD—they should indicate co-reference links. To handle our two cases, we opted for an adhoc process, in which *owl:sameAs* statements are obtained from *dc:identifier* values that institutions provided to Europeana for their items. Later on, however, the Europeana ingestion workflow should allow providers to explicitly indicate the existence of LOD URIs for their items.

We take all the resulting RDF/XML files and the supplied/generated links and generate data dumps, which are then (i) made available as Europeana EDM download at <http://data.europeana.eu/download/> and (ii) ingested into an RDF store.

According to established recipes (Sauermaun et al., 2008), incoming HTTP requests from Linked Data Clients, which carry an RDF-specific Internet media type (e.g., *application/rdf+xml*) in the HTTP *Accept* header field, are answered by an RDF store. All other HTTP requests sent to <http://data.europeana.eu> are redirected to the Europeana portal at <http://www.europeana.eu>.

4. From Europeana Semantic Elements to the Europeana Data Model

The Europeana Semantic Elements (ESE) schema is the current metadata format used in the European production system. It is based on simple unqualified Dublin Core,⁶ adding Europeana-specific fields such as *dataProvider* or *country*. Fig. 2 gives an example of a (test) ESE record.

As a basic solution to Europeana interoperability problems, ESE suffers from many issues. First, it is a “flat” model that uses simple string values for the metadata fields. This prevents properly linking an item ingested by Europeana to other objects (e.g., a series of portraits), or to contextual entities as represented by complex resources, e.g., a creator with many naming variations, or a concept that is more specific than another one, all of which could help improving access to Europeana items (Schreiber et al., 2008). Also, ESE aggregates in one and the same record metadata fields that can apply to different entities, breaking the “one-to-one” principle (Miller, 2010) and causing great confusion: for example, some providers use rights- or date-related fields to give information for the “real-world” object they hold, while others use the same fields for data on the digital representations of these items.

In recent years, the RDF-based Europeana Data Model (EDM) has been developed as an alternative to the ESE schema, aiming at solving these shortcomings. It should facilitate Europeana's transition from a closed data repository to an open information space that integrates with the Web architecture and the Linked Data principles for identifying and exposing resources on the Web. Key features of EDM are the ability to distinguish a “provided object” (real-world item such as paintings, books) from its digital representation(s), the possibility of several descriptions of a same item to co-exist in Europeana, and the support for complex item representation, with sequence or partition information (Europeana, 2010).

⁵ <http://europeanalabs.eu/wiki/EDMPrototypingTask21Annocultor> [2011.04.30]

⁶ <http://dublincore.org/documents/dcmi-terms/>

ESE Field	Value
dc:identifier	02-Lemba-Lakkous-Lady.tif
dc:title	Stone figurine
dc:subject	Cyprus--Antiquities
dc:description xml:lang="en"	Stone vessel from Choirokoitia with embossed decoration...
dc:publisher	Cyprus Archaeological Museum
dc:type	Image
dc:format	image/tiff
dcterms:temporal	mid 3rd millenium B.C.
dc:rights	Cyprus Archaeological Museum
dc:title	Λίθινο ειδώλιο
dc:description	Λίθινο ειδώλιο από τη θέση Λέμπα ...
europaana:provider	Cyprus Library
europaana:type	IMAGE
europaana:rights	http://creativecommons.org/licenses/by/3.0/nl/
europaana:dataProvider	Cyprus Archaeological Museum
europaana:isShownBy	http://www.mcw.gov.cy/mcw/da/da.nsf/All/8308A7AB9CA6CB5EC22574CC003FF5E1/\$file/02-Lemba-Lakkous-Lady.jpg?OpenElement
europaana:uri	http://www.europaana.eu/resolve/record/00000/E2AAA3C6DF09F9FAA6F951FC4C4A9CC80B5D4154
europaana:country	cyprus
europaana:language	gr

xmlns:dc = "http://purl.org/dc/elements/1.1/"
xmlns:europaana = "http://www.europaana.eu"

Fig. 2. Sample Europeana ESE record.

Rather than systematically introducing new elements, EDM re-uses and links to existing reference vocabulary elements. The core data structure of EDM, for instance, is based on the Open Archive Object Reuse and Exchange Model (OAI-ORE), a reference model for the description and exchange of aggregations of Web resources.⁷ ORE *aggregations* are used to represent a data provider's contribution to Europeana, which consists of the “provided item” together with its digital “view(s)” (modeled as *web resources*). Each of them can come with its own metadata, using a mix of Dublin Core properties and adhoc Europeana ones, many adapted from the CIDOC Conceptual Reference Model (CRM) (Doerr et al., 2008). As an improvement over existing ESE metadata, most properties can now be used to link to fully-fledged “contextual resources”, e.g., *places*, *agents* or *concepts*, using existing vocabularies. Concepts, for instance, should be described using the W3C SKOS model for Knowledge Organization Systems.⁸ The ORE *proxy* mechanism allows several descriptions for a same provided item to co-exist. This is especially useful to track metadata with different provenance, i.e., when different data providers contribute different descriptions of a same book, or when Europeana itself performs semantic enrichment, hence creating metadata that differ from a data provider's original contribution.

We defined a mapping between ESE and EDM that covers the creation of the EDM entities (items, aggregations, proxies), the assignment of dereferencable HTTP URI identifiers to these entities, as well as the attachment of metadata fields to these entities. In the following, we informally explain the mapping by using the above ESE record to illustrate how the metadata values from legacy ESE records are distributed over the various EDM entities, and new ones are created. For a detailed explanation we refer to the full mapping specification⁹ as well as the EDM Primer and the OAI-ORE documentation.

Fig. 3 shows the skeletal structure of the EDM record that results from the conversion of our sample ESE record. The resource in the center represents the real-world item, i.e., a stone figure

⁷ <http://www.openarchives.org/ore/>

⁸ <http://www.w3.org/2004/02/skos/>

⁹ <http://europeanalabs.eu/wiki/EDMPrototypingTask15> [2011.04.30]

from the Cyprus Archaeological Museum. It is described by metadata coming from one data provider and also by Europeana-specific metadata. The data provider's description, attached to one *ore:Proxy*, is aggregated in a single data provider *ore:Aggregation*. Europeana aggregates this aggregation, as well as a Europeana-specific *ore:Proxy* resource, into a single *ens:EuropeanaAggregation*. We assign dereferencable HTTP URIs to each of these objects.

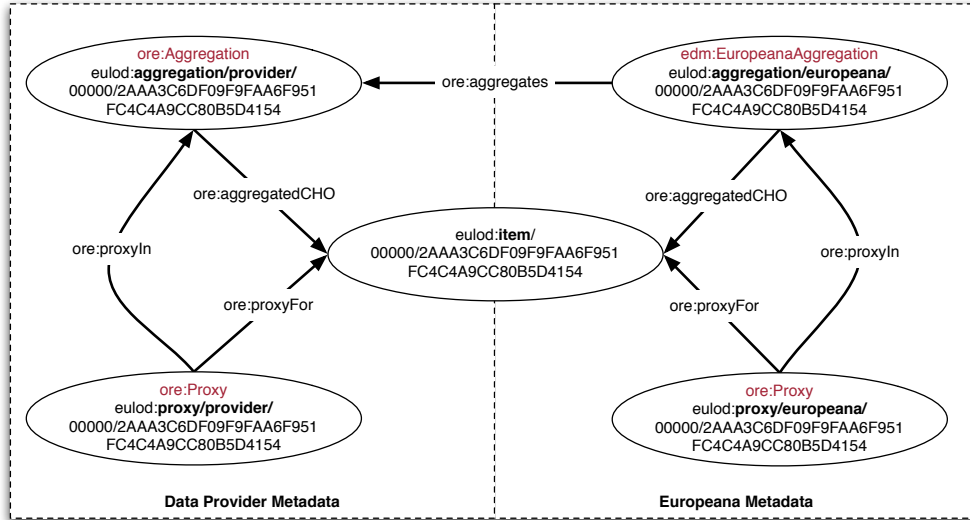


Fig. 3. Basic structure of EDM networked resources.

Fig. 4 shows that the data provider and Europeana aggregations build the bridge between the human- and the machine-readable Europeana Web content. The *ore:Aggregation* resource references the data provider's website about the stone figure, the *ens:EuropeanaAggregation* points to the Europeana HTML Web page of that item.

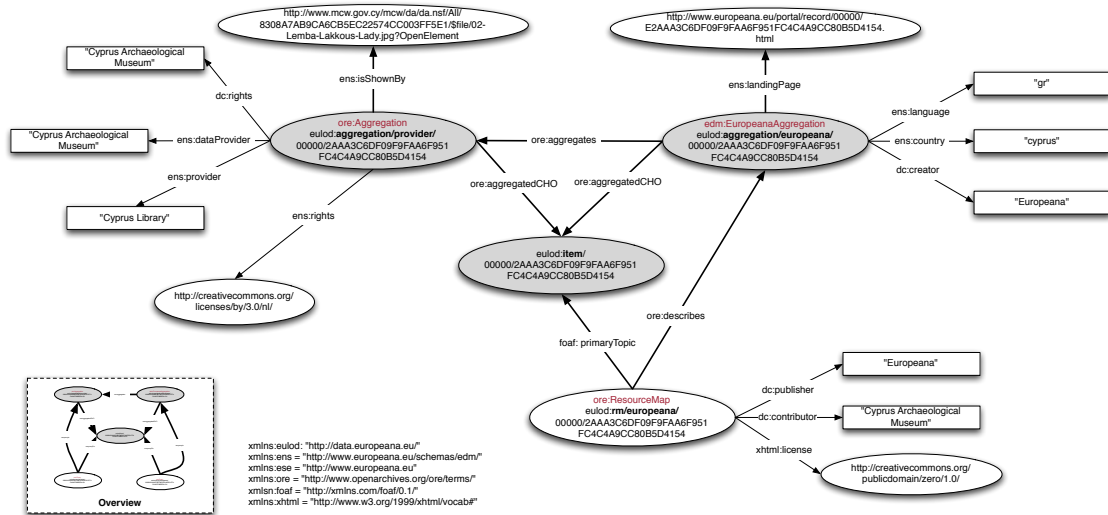


Fig. 4. EDM aggregations.

The descriptive metadata for the real-world stone figure are, as depicted in Fig. 5, attached to the data provider's *Proxy*, if this provider directly contributed these data. Data produced by Europeana, either as part of the several data normalization and quality enhancement procedures in the data ingestion process, or from subsequent semantic enrichment, are attached to Europeana's proxy for the item.

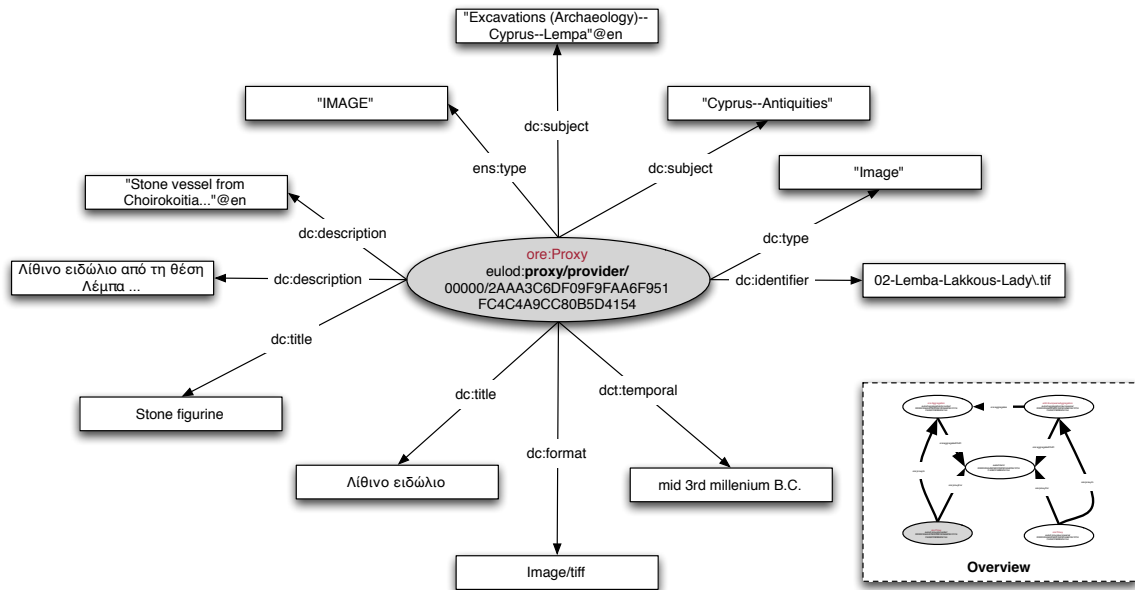


Fig. 5. EDM provider proxy.

5. Lessons Learned and Discussion

Version 1.0 of the Europeana Linked Open Data Pilot operates 3.5M items harvested from 9 direct Europeana providers and almost 300 indirect data providers, from 16 European countries. Without links to other data sources this results in a dataset containing 185M triples.

In the following we discuss some issues we encountered when converting ESE data to EDM and assigning URIs to the resulting resources. We also summarize the results of a first RDF store performance evaluation and other lessons we learned while developing our pilot.

5.1. Mapping ESE to EDM in the context of linked data practices

The major difference between ESE and EDM is that ESE records are flat lists of property value pairs while EDM data are networked resources. Converting ESE data to EDM thus requires distributing ESE elements over EDM resources, each playing a different role in the EDM model. The major issue here was to identify the target EDM resource for a given ESE property since the one-to-one mapping principle does not apply. In many cases, indeed, providers have not strictly followed the ESE guidelines, and use values that apply to digital representations in fields that in theory would apply to the provided “real-world” item.

Another problem concerns the structure of EDM data itself. The requirement for distinguishing different data sources, and data that applies to different resources, leads to creating complex networks of aggregations, proxies and other resources. While this has many benefits, it also raises the barrier to data access and consumption. Beyond adding extra complexity to the RDF graphs published, the proxy pattern, which was introduced because of the lack of support for named graphs in RDF, is indeed quite a counter-intuitive necessary evil for linked data practitioners—including the authors of this paper. It is especially confusing to find statements, which are about, say, a painting, but are attached to a resource that is not strictly speaking standing for that painting. We were tempted to make the work of linked data consumers easier, at least by copying the statements attached to the provider and Europeana proxies onto the “main” resource for the provided item, so as to allow *direct access* to these statements—*i.e.*, not mediated through proxies. We decided against it, trying to avoid such data duplication. Feedback from data consumers may yet cause us to re-consider this decision. On the longer term, also, we hope that

W3C will soon standardize “named graphs” for RDF.¹⁰ This mechanism would allow EDM to meet the requirements for tracking item data provenance without using proxies.

In order to enhance navigability between resources, we were also tempted to add extra links to compensate for problems EDM inherits from OAI-ORE. For example, the main access point to Europeana item data is the URI of the item itself.¹¹ But ORE has no property for linking *from* this resource *to* the aggregations of the proxies connected to it. This makes accessing these resources more difficult, if one assumes a simple “follow-your-nose” approach to data publication and consumption, *i.e.*, only publishing *outbound* RDF statements¹² for a given resource. Fortunately we did not have to turn to this option, as the RDF store we use (Openlink Virtuoso) also returns *inbound* links when asked for the data on a given resource. Practitioners should however be aware that other implementation choices might have dictated advanced RDF store configuration, or, worse, technical additions to the data model itself—*e.g.*, coining a new *ex:hasProxy* property, inverse of the existing *ore:ProxyFor* one, to enable direct navigation from an item to the proxies related to it.

5.2. Disseminating meta-metadata

An important requirement for our service is to communicate meta-level information about the data it publishes. Europeana's missions indeed include becoming a trusted source of information, while encouraging more open data circulation in the culture sector. To this end, provenance and licensing information are crucial—be they about the cultural items being accessed, or about the metadata on these items.

However, Linked Data technology still lacks a fully standardized suite to express such meta-level information. Keeping an eye on relevant ongoing efforts—the above-mentioned W3C RDF Group's *named graphs*, the W3C Provenance Working Group (<http://www.w3.org/2011/prov/>), or DCMI's Provenance Task Group (<http://dublincore.org/groups/provenance/>)—we tried to combine various *existing* solutions together.

The base for our meta-level data is OAI-ORE's *resource maps*. These entities, a close encounter to named graphs in the ORE context, serve to reify the very data that describe ORE aggregations. We connect each Europeana aggregation and its corresponding item to a specific resource map, which carries the required licensing and provenance data. The former follows the CCrel recommendations (http://wiki.creativecommons.org/CC_REL). The latter uses DC elements: *dc:publisher* indicates that Europeana is responsible for the data dissemination, while *dc:contributor* acknowledges the Europeana partners who contributed the data.

Our choice is similar to the ones of other institutions. For example, the New York Times' linked data service (<http://data.nytimes.com/>) reifies the RDF files it serves and attaches to them the appropriate meta-metadata. We however hope to get more feedback from our community, which will help us realize a fully consensual approach.

5.3. HTTP URI design

The transition from Europeana URIs to dereferencable HTTP URIs for EDM aggregations and proxies was a major challenge in the conversion process. The existing Europeana URIs follow a pattern¹³ that identifies records rather than resources representing real-world objects. Since the main Europeana production system and the Europeana Linked Open Data Prototype are still two distinct systems and we somehow needed to create a bridge between the identification

¹⁰ For progress on named graph standardization, see the RDF Working Group site, <http://www.w3.org/2011/rdf-wg/>

¹¹ *E.g.*, <http://data.europeana.eu/item/00000/E2AAA3C6DF09F9FAA6F951FC4C4A9CC80B5D4154>

¹² Cf. the Concise Bounded Description data publication pattern, <http://www.w3.org/Submission/CBD/>.

¹³ <http://www.europeana.eu/resolve/record/{collectionID}/{itemID}>, where *collectionID* is a number attributed to a collection by Europeana at the time of data ingestion, and *itemID* a key uniquely identifying an object within that collection.

mechanisms in place, we reused the Europeana URI pattern for the dereferencable resources in the Europeana LOD prototype. Yet we face persistence issues: while building this pilot some Europeana collections have been re-harvested, and the URIs of their items have been changed. While we could solve such issues prior to our first release, both the LOD infrastructure and the underlying Europeana identification mechanism will have to find better strategies in the future.

Another issue was the articulation between provenance and typing for our identifiers. According to established practices for minting URIs, *e.g.*, (CTO UK Council, 2009), the type of entities should be reflected in their URIs. The assumption is that consistently giving URIs of the form <http://data.europeana.eu/item/{itemID}> to all Europeana *items* will help data consumption and access, especially in the context of a RESTful architecture (Fielding, 2000). But giving such importance to resource types somewhat clashes with granting high visibility to collection-specific information spaces. A pattern like <http://data.europeana.eu/{collectionID}/item/{itemID}> would allow Europeana to later build a data service that provides direct access to all resources related to the collection *collectionID*, irrespective of their type. This would be especially useful for providers who want to get access to data on one collection, in line with Europeana's mission of data (re-)distributor for cultural institutions. As such a scenario is not precisely articulated now, we however decided to opt for the more classical URI minting strategy.

5.4. RDF Storage and Scalability

To answer the evergreen question on the performance and scalability of RDF stores for data volumes in the size of the Europeana dataset, we conducted a survey of existing RDF stores (Haslhofer et al., 2011), compared existing solutions, and conducted a quantitative load and query response time performance analysis using the complete Europeana RDF dataset. The result of this study was that stores that can load data sets in the size of Europeana are already available. When running on high-memory machines, they provide acceptable performance for simple SPARQL SELECT and DESCRIBE queries. At the moment we are using RDF stores for read-only purposes after an initial dump import. We haven't gained any experience how RDF stores perform for other metadata management tasks in a larger scale.

6. Conclusions and Next Steps

With data.europeana.eu we are making a first version of the Europeana metadata set available as Linked Open Data. It includes data from those providers who have opted for their data to be published in the public domain. Anyone can now obtain descriptive metadata about Europeana resources by downloading the dumps, dereferencing HTTP URIs, or executing SPARQL queries against a dataset that follows the Europeana Data Model (EDM). The tools and scripts we developed for this purpose are open source and can freely be re-used or extended by others.

We are seeking community support for adding new datasets, define linking specifications on a per-institution basis, and improve the overall quality of data. Linking from Europeana metadata to other Linked Data sets is especially crucial. We need to further evaluate the alignments currently included in our pilot, and extend them to other data sources, such as DBpedia, VIAF and other relevant initiatives from our community. We will also investigate the possibility to integrate automated linking frameworks such as Silk (Volz et al., 2009) into the dump creation process. Refining the mapping from ESE to EDM and trying to find solutions for the issues mentioned in the previous section is also part of our future work. If the community positively reacts to this pilot, we will finally propose strategies for providing periodic updates and better synchronizing the published dataset with the data hosted in the Europeana production service.

Another area of future work will be the integration of Europeana data with other initiatives that promote the availability of structured metadata on the Web, such as schema.org. We believe that this can increase the visibility of Europeana on the Web.

Acknowledgements

The authors would like to thank Mary Rowlett, Cesare Concordia, Jan Stankovsky, Adam Horvath, Victor de Boer, Jacco van Ossenbruggen, Jan Wielemaker, Steffen Henicke, Carlo Meghini, Tsuyoshi Sugibuchi, Go Sugimoto, and Stefan Gradmann for their support in infrastructure setup, data model design, communication with data providers, and continuous lobbying to make this pilot happen. The work has partly been supported by the European Commission as part of the eContentplus program (EuropeanaConnect) and by a Marie Curie International Outgoing Fellowship (SciLink) within the 7th Europeana Community Framework Programme.

References

- Berners-Lee, Tim. (2009) Linked Data - Design Issues. Retrieved April 30, 2011, from <http://www.w3.org/DesignIssues/LinkedData.html>.
- CTO Council UK. (2009). Designing URI Sets for the UK Public Sector. Technical report, Chief Technology Officer Council, UK, October 2009. Retrieved April 30, 2011, from <http://www.cabinetoffice.gov.uk/resource-library/designing-uri-sets-uk-public-sector>.
- Doerr, Martin, and Dolores Iorizzo. (2008). The dream of a global knowledge network—a new approach. *J. Comput. Cult. Herit.*, 1:5:1–5:23, June 2008.
- Europeana. (2010). Europeana Data Model Primer. Technical report. August 2010. Retrieved April 30, 2011, from <http://version1.europeana.eu/web/europeana-project/technicaldocuments/>.
- Fielding, Roy. (2000). Architectural Styles and the Design of Network-based Software Architectures. PhD thesis, University of California, Irvine, 2000. Retrieved April 30, 2011 from: <http://www.ics.uci.edu/fielding/pubs/dissertation/top.htm>
- Gradmann, Stefan. (2010). Knowledge = Information In Context. On the Importance of Semantic Contextualiation in Europeana. Technical report, Berlin School of Library and Information Science, Humboldt University, April 2010. Retrieved April 30, 2011, from <http://www.scribd.com/doc/32110457/Europeana-White-Paper-1>
- Haslhofer, Bernhard, Elaheh Momeni Roochi, Bernhard Schandl, and Stefan Zander. (2011). Europeana RDF Store Report. Technical report, University of Vienna. Retrieved April 30, 2011, from: <http://eprints.cs.univie.ac.at/2833/>
- Heath, Tom, and Christian Bizer. (2011). Linked Data: Evolving the Web in a Global Data Space. *Synthesis Lectures on the Semantic Web: Theory and Technology*. Morgan & Claypool, 1 edition.
- Landry, Patrice. (2009). Multilingualism and subject heading languages: how the MACS project is providing multilingual subject access in Europe. *Catalogue & Index: Periodical of the Chartered Institute of Library & Information Professionals (CILIP) Cataloguing & Indexing Group*, 157.
- Malmsten, Martin. (2008). Making a library catalogue part of the semantic web. In *Proc. of the Intl. Conf. on Dublin Core and Metadata Applications (DC2008)*, Berlin, Germany.
- Miller, Steven J. (2010). The One-To-One Principle: Challenges in Current Practice. In *Proc. of the Intl. Conf. on Dublin Core and Metadata Applications (DC2010)*, Pittsburgh, Pennsylvania.
- Panzer, Michael. (2008). Cool URIs for the DDC: towards web-scale accessibility of a large classification system. In *Proc. of the Intl. Conf. on Dublin Core and Metadata Applications (DC2008)*, Berlin, Germany.
- Sauermann, Leo and Richard Cyganiak. (2008). Cool URIs for the Semantic Web. Technical report, November 2008. W3C Interest Group Note. Retrieved April 30, 2011, from <http://www.w3.org/TR/cooluris/>.
- Schreiber, Guus, Alia K. Amin, Lora Aroyo, Mark van Assem, Viktor de Boer, Lynda Hardman, Michiel Hildebrand, Borys Omelayenko, Jacco van Ossenbruggen, Anna Tordai, Jan Wielemaker, and Bob J. Wielinga. (2008). Semantic annotation and search of cultural-heritage collections: The Multimedial e-culture demonstrator. *J. Web Sem.*, 6(4): 243–249.
- Summers, Ed, Antoine Isaac, Clay Redding, and Dan Krech. (2008). LCSH, SKOS and Linked Data. In *Proc. of the Intl. Conf. on Dublin Core and Metadata Applications (DC2008)*, Berlin, Germany.
- Van de Sompel, Herbert, and Oren Beit-Arie. (2001). Open linking in the scholarly information environment using the OpenURL Framework. *New Review of Information Networking*, 7(1): 59–76.
- Van der Meij, Lourens, Antoine Isaac, and Claus Zinn. (2010). A Web-Based Repository Service for Vocabularies and Alignments in the Cultural Heritage Domain. In *Proc. of the 7th Extended Semantic Web Conf., ESWC 2010*, pages 394–409, Springer.
- Volz, Julius, Christian Bizer, Martin Gaedke, and Georgi Kobilarov. (2009) Silk - a link discovery framework for the web of data. In *2nd Linked Data on the Web Workshop (LDOW)*.