Europeana and (many) linked open vocabularies

Valentine Charles
18th June 2012, Madrid
Europeana is a service that aggregates data from the cultural heritage sector in Europe.

- Libraries, museums, archives and audio-visual archives
- [http://www.europeana.eu/](http://www.europeana.eu/)

Provides a portal for users to access that data

- Metadata, previews and links to source

Will make the metadata freely available for anyone to re-use

- Under Creative Commons Zero (CC0) - public domain dedication

Enriches data, provides tools

- Link to data from other sites, embed on wikipedia, API

Makes data available as Linked Open Data (experimentally)

- [http://data.europeana.eu/](http://data.europeana.eu/)
Solving a problem of data integration

Europeana is aggregating data from the cultural heritage sector in Europe.

- Each domain is using its specific metadata standards and vocabularies
  - Athena for museums with Lido
  - APEnet for Archives with EAD

Until now.. the Europeana Semantic Elements (ESE)

- Represents lowest common denominator for object metadata
- Forces interoperability
- Flat model, mostly with text string values
- One-to-one principle violated
- Major drawback: richness of the original metadata is lost
First steps for providing semantics enabled services

- Enrichment of the data with selected vocabularies and datasets:
  - DBpedia
  - Geonames
  - GEMET
  - Enrichment process based on a selection of Dublin Core elements

- Europeana needs to have access to open resources
First steps for providing semantics enabled services
Re-building the service on rich metadata format

→ The Europeana Data Model (EDM) provides a new environment
  • Provides an open environment for Cultural Heritage data

→ Re-use OAI ORE for organization of metadata about an object
  • Distinguish between the real world **object** (painting, book, program) and its **digital representation**
  • And the **object** and the **metadata record** describing the object
  • Supports several statements per resources

→ EDM gives support for contextual resources (semantic layer)

→ Re-use existing standards
  • OAI-ORE
  • Dublin Core
  • SKOS…
Requirements for Linked Open Vocabulary
Access and maintenance

- Vocabularies should be technically available and documented
  - Organisational support is key

- Vocabularies should be well-connected together
  - Key to avoid duplication and redundancy

- Elements sets should address specific requirements instead of semantic specifications

- The necessity of a map of the vocabulary landscape
  - Linked Open vocabularies: http://labs.mondeca.com/datasets/lov
  - Open Metadata Registry: http://metadataregistry.org
Requirements for Linked Open Vocabulary Tools

→ MINT: http://mint.image.ece.ntua.gr/
Requirements for Linked Open Vocabulary

Tools

⇒ AMALGAME: http://semanticweb.cs.vu.nl/amalgame/

Analyse

<table>
<thead>
<tr>
<th>mapping</th>
<th>sources</th>
<th>targets</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact on altLabel</td>
<td>281 (8%)</td>
<td>299 (0.2%)</td>
<td>313</td>
</tr>
<tr>
<td>Exact on prefLabel</td>
<td>723 (16%)</td>
<td>703 (0.4%)</td>
<td>723</td>
</tr>
<tr>
<td>1-1 matches</td>
<td>683 (15%)</td>
<td>683 (0.4%)</td>
<td>683</td>
</tr>
<tr>
<td>n-m matches</td>
<td>40 (1%)</td>
<td>20 (0.01%)</td>
<td>40</td>
</tr>
</tbody>
</table>


Thank you

Valentine Charles

valentine.charles@kb.nl