Validation of a metadata application profile domain model

Mariana Curado Malta, Helena Bermúdez-Sabel, Ana Alice Baptista, Elena González-Blanco

Agenda

- Context - European Poetry
- Methodology
- Validation of the Domain Model
- Final remarks: where we are and future work
European Poetry

- Publish poetry data as LOD

- Digital repertoires: a catalog that gathers information about:
  - metrical and rhythmical schemes of poetical tradition
  - the period and the school
  - sometimes the text of the poem
  - information related to authors, manuscripts, editions, music, etc

European Poetry: starting point

- Starting point: digital repertoires of a community of researchers
- Not interoperable
- https://goo.gl/O0mqhI
Agenda

- Context - European Poetry

- Methodology

  - Validation of the Domain Model
  - Final remarks: where we are and future work

Methodology

- Me4MAP: MAP-EP (use-case to validate Me4MAP)
- Total: 23 repertoires
- 17 provided database structures
  - Reverse engineering process
  - From logical models to conceptual
- 5 → Website analysis → informational needs
- survey to final users
- Iterative process, building up knowledge:
  - v0.1; v0.2; v0.3 → v1.0
Agenda

- Context - European Poetry
- Methodology
- Validation of the Domain Model
- Final remarks: where we are and future work

Validation#1

- Workshop with 9 delegates of the 11 repertoires
- All delegates were researchers in philology
- The delegates received a dossier with:
  - A UML diagram with the conceptual model of their own database
  - A spreadsheet file with a mapping between the logical model of the database of the delegate and the conceptual model of the database
  - A UML class diagram of the Domain model
  - A spreadsheet with i) a list of the classes of the Domain Model and description ii) a list of the attributes of the Domain Model with description iii) a list of the relations, with domain and range
- A testing sheet (in Excel) was used to execute the validation
### Table 1: Property Label and Range

<table>
<thead>
<tr>
<th>Property Label</th>
<th>Range</th>
<th>Instance Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>isRealisedThrough</td>
<td></td>
<td>Redaction</td>
</tr>
<tr>
<td>isPreservedIn</td>
<td></td>
<td>AlternativeSource</td>
</tr>
<tr>
<td>hasCreator</td>
<td></td>
<td>Person</td>
</tr>
<tr>
<td>isPartOf</td>
<td></td>
<td>Ensemble</td>
</tr>
<tr>
<td>isEditedIn</td>
<td></td>
<td>BibliographicSource</td>
</tr>
<tr>
<td>isReferred</td>
<td></td>
<td>BibliographicSource</td>
</tr>
<tr>
<td>comesFrom</td>
<td></td>
<td>Place</td>
</tr>
<tr>
<td>isOrganisedIn</td>
<td></td>
<td>CatalogEntry</td>
</tr>
</tbody>
</table>

### Table 2: Property Label and Range

<table>
<thead>
<tr>
<th>Property Label</th>
<th>Range</th>
<th>Instance Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>allTitle</td>
<td>0-1</td>
<td></td>
</tr>
<tr>
<td>betaReferenceID</td>
<td>0-1</td>
<td></td>
</tr>
<tr>
<td>date</td>
<td>0-M</td>
<td></td>
</tr>
<tr>
<td>dateNote</td>
<td>0-M</td>
<td></td>
</tr>
<tr>
<td>domain</td>
<td>0-1</td>
<td></td>
</tr>
<tr>
<td>buttonReferenceID</td>
<td>0-1</td>
<td></td>
</tr>
<tr>
<td>genre</td>
<td>0-M</td>
<td></td>
</tr>
</tbody>
</table>

---

Curado Malta et al. | LINHD, UNED - Spain | CEOS.PP, Polytechnic of Porto - Portugal | Algoritmi Center, University of Minho - Portugal
### Validation of a Metadata Application Profile Domain Model

**Instance Label:** OP1

<table>
<thead>
<tr>
<th>Property Label</th>
<th>Range</th>
<th>Card.</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>allTitle</td>
<td>0-1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>betaReferenceID</td>
<td>0-1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>date</td>
<td>0-M</td>
<td></td>
<td>814</td>
</tr>
<tr>
<td>dateNote</td>
<td>0-M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>domain</td>
<td>0-1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>duttonReferenceID</td>
<td>0-1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>genre</td>
<td>0-M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>incipitumCarminumLatinorumReferenceID</td>
<td>0-1</td>
<td>32</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Property Label</th>
<th>Range</th>
<th>Instance Label</th>
<th>Range</th>
<th>Instance Label</th>
<th>Range</th>
<th>Instance Label</th>
<th>Range</th>
<th>Instance Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>isRealisedThrough</td>
<td>Redaction</td>
<td>R1</td>
<td></td>
<td>R2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isPreservedIn</td>
<td>AlternativeSource</td>
<td>PER2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hasCreator</td>
<td>Person</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isPartOf</td>
<td>Ensemble</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isEditedIn</td>
<td>BibliographicSource</td>
<td>ED1</td>
<td></td>
<td>ED2</td>
<td>ED3</td>
<td></td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>isReferred</td>
<td>BibliographicSource</td>
<td>BIB12</td>
<td></td>
<td>BIB13</td>
<td>BIB14</td>
<td></td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>isOrganisedIn</td>
<td>CatalogEntry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>comesFrom</td>
<td>Place</td>
<td>PL3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Delegates were asked to register the issues that arose during the validation process.

The delegates were also asked to fill in an on-line form with the following questions:

- Could you describe all your data with the available elements? If not, please refer the difficulties.
- Did you have any difficulty in particular to describe your data? Were there any ambiguities?
- Is there anything else you want to add?

We used all the comments to build v0.2 of the Domain Model.
Agenda

- Context - European Poetry
- Methodology
- Validation of the Domain Model
- Final remarks: where we are and future work

Validation#2

- Three additional resources that were not on the repertoires used
- Address at a certain point the general scope of the model (other contexts, in the same community of practice)
- Done by a master student that did not participate in the process of construction of the model

Process:
- 6 different repertoires, 9 poetic resources chosen randomly (9 use-cases)
- Populate the Domain Model v0.3 with the data from the poetic resources
- Evaluate and report
Populate the Domain Model v0.3

- Description of the Domain Model v0.3 in XML
- Schema files for the use-cases that validate their contents against the DM:
  - restricted classes, attributes and relations, so any elements that were not contained in the Domain Model could not be added.
  - prevented the repetition of labels that identified the different instances of each class so to avoid ambiguities.
  - controlled the relations between the different instances of class: except for the instance of class Opus, every instance of any other class had to be the range of at least one relation.
- Modelling of the use-cases → describe the resource in XML using the classes, attributes and relations of the Domain Model.

- The construction of the use-cases affected the contents of the DM
- New Informational need identified → added to the Domain Model
- We had an updated version to validate against the use-cases
- The XML provided as representative of Version 0.3 represents a previous stage of the Domain Model than the use-cases
Populate the Domain Model v0.3

- The construction of the use-cases affected the contents of the DM
- New informational needs identified → added to the Domain Model
- We had an updated version to validate against the use-cases
- The XML provided as representative of Version 0.3 represents a previous stage of the Domain Model than the use-cases

Agenda

- Context - European Poetry
- Methodology
- Validation of the Domain Model
- Final remarks: where we are and future work
This process of validation resulted in the first stable version of the Domain Model of the MAP-EP

The importance of having a good DM

Follows (according to Me4MAP):
- Vocabulary alignment
- Constraints matrix
- Coding of MAP-EP
- …

This work is also a use-case for the validation of Me4MAP