



European
Research
Council



Towards interoperability in the European poetry community

Shaping Data in Digital Humanities
University of Copenhagen, 20/04/2018

Helena Bermúdez Sabel

helena.bermudez@linhd.uned.es

Laboratorio de Innovación en Humanidades Digitales, UNED, Spain

Member of the project Poetry Standardization and Linked Open Data -

POSTDATA



- Presentation
- Rationale
- Methodology
- Concepts standardisation
- Conclusion and future work

- **Presentation**
- Rationale
- Methodology
- Concepts standardisation
- Conclusion and future work

POSTDATA axes



POSTDATA

Poetry Standardization
and Linked Open Data

Semantic Web Infrastructure



Linked
Open Data
(LOD)

Virtual Research Environment



Digital
Scholarly
Editing

Poetry Lab



Natural
Language
Processing

- Presentation
- **Rationale**
- Methodology
- Concepts standardisation
- Conclusion and future work

Western literature



POSTDATA

Poetry Standardization
and Linked Open Data



And yet...



POSTDATA

Poetry Standardization
and Linked Open Data

- Diversity of paradigms
- Fragmentary access to poetic resources

And yet...



POSTDATA

Poetry Standardization
and Linked Open Data

- Diversity of paradigms
- Fragmentary access to poetic resources
- Partiality of methods (and results)

A workaround



POSTDATA

Poetry Standardization
and Linked Open Data

- The key concepts: standardisation and interoperability
- The most advantageous paradigm: Linked Open Data

LOD cloud

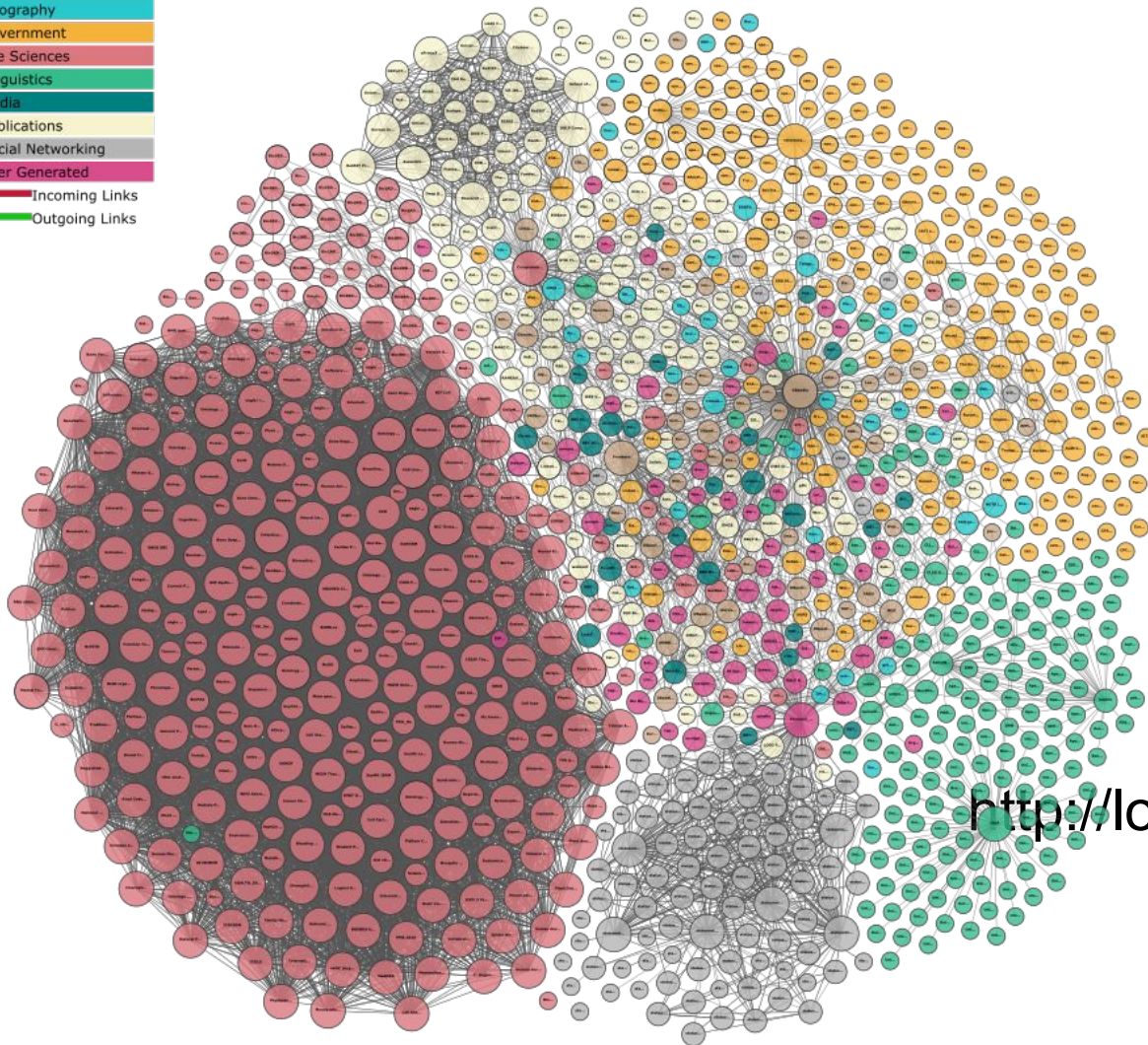


POSTDATA

Poetry Standardization
and Linked Open Data

Legend

Cross Domain
Geography
Government
Life Sciences
Linguistics
Media
Publications
Social Networking
User Generated
Incoming Links
Outgoing Links



<http://lod-cloud.net/>

- A huge database (worldwide) on the Web of Data
- Since the data is open, the community can rely also on other sources of data
- Big advantage over other approaches

Goal



POSTDATA

Poetry Standardization
and Linked Open Data

- Enhance interoperability between existent poetic resources
- Facilitate the creation of new resources

- Presentation
- Rationale
- **Methodology**
- Concepts standardisation
- Conclusion and future work

Main task



POSTDATA

Poetry Standardization
and Linked Open Data

- Defining a Metadata Application Profile (MAP)
for European Poetry

- Main task: Defining a Metadata Application Profile (MAP) for European Poetry

(MAP: a semantic model in the context of LOD)

- Main task: Defining a Metadata Application Profile (MAP) for European Poetry
- Milestone: to define a Domain Model

- Main task: Defining a Metadata Application Profile (MAP) for European Poetry
- Milestone: to define a Domain Model

(Domain Model: a model that presents the concepts of the context and its relations)

- Repertoires of European Poetry and Poetry Metrics
 - Data models
 - Functionalities of the Websites
- Report of data needs of the community

Starting point



POSTDATA

Poetry Standardization
and Linked Open Data

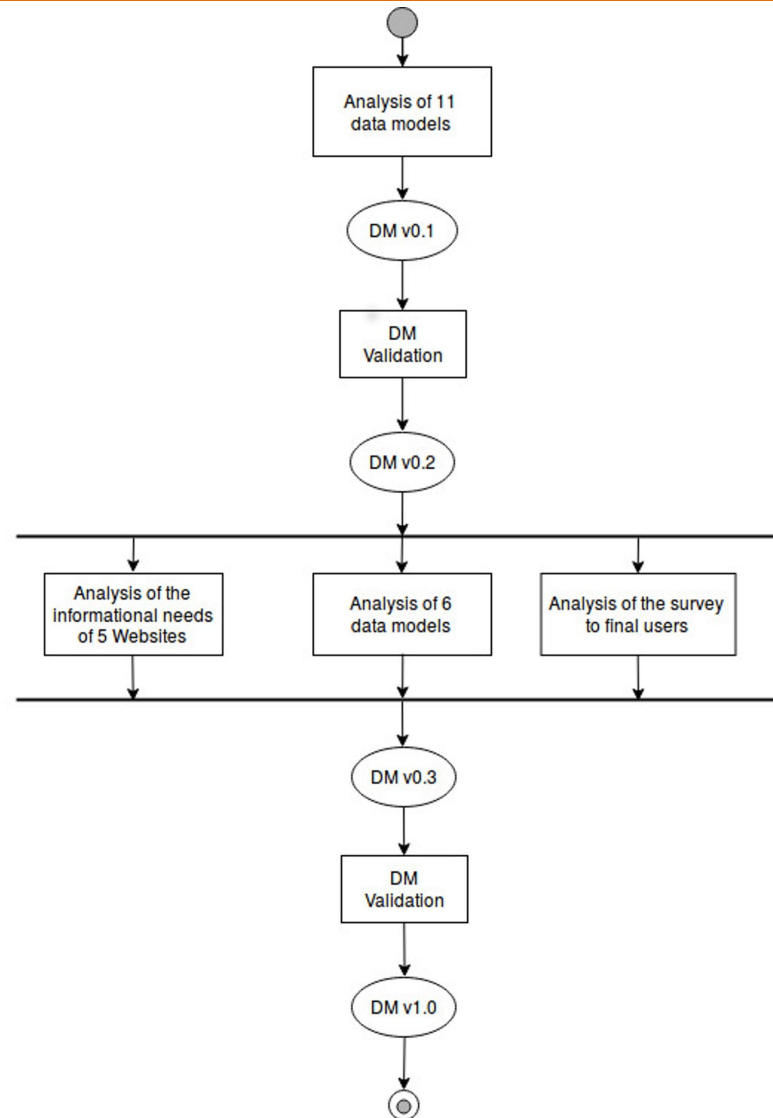


Workflow



POSTDATA

Poetry Standardization
and Linked Open Data



- Analysis of the data model of a representative sample of EP databases
- Analysis of a survey addressed to the final users of the repertoires in order to understand the data needs of the users of poetry databases.

Analysis of the data models



POSTDATA

Poetry Standardization
and Linked Open Data

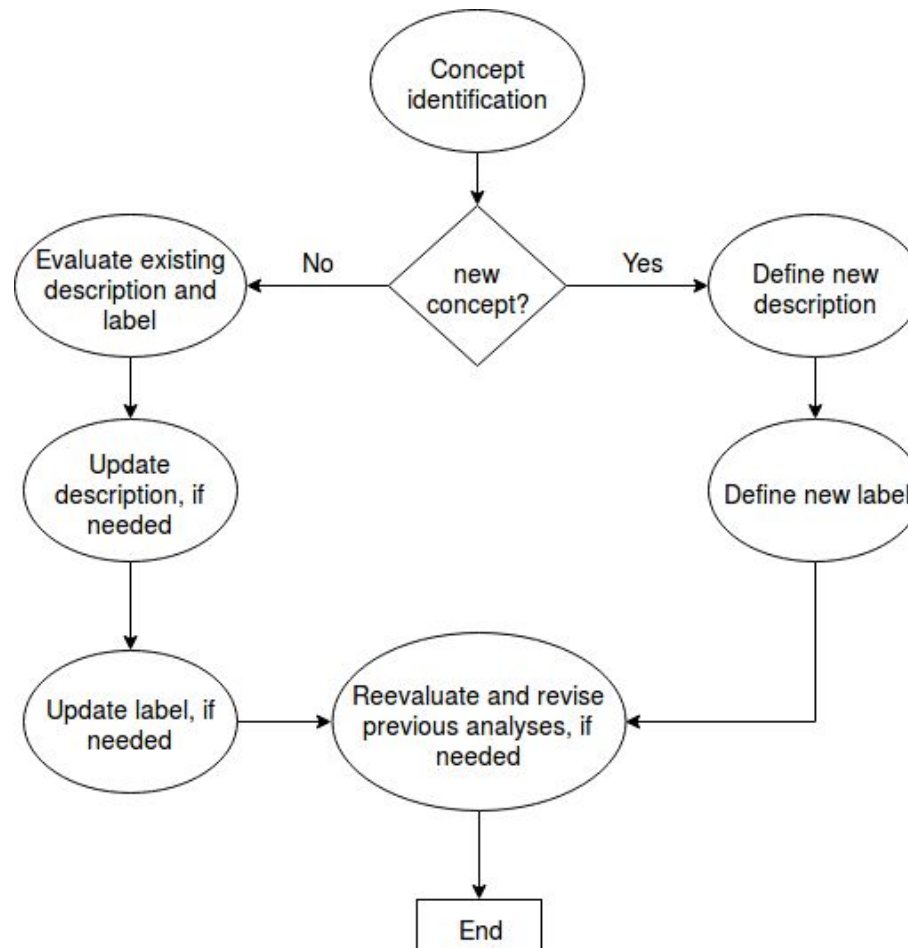
- Analysis of the file sent by each partner with the structure of the database
- Study of the structure of the database
- Identification of every concept of the database, as well as the properties that characterize that concept
- Identification of the relationships between concepts

Concept analysis process



POSTDATA

Poetry Standardization
and Linked Open Data



Types of modeling

- **egoistic modeling:** to express specific research ideas in cases where data is being created to support the creator's own research needs
- **altruistic modeling:** to serve as an interchange format for some types of users and user communities where data is typically being created and modeled with someone else's needs in mind

(Jannidis & Flanders 2013)

Importance of data modeling



POSTDATA

Poetry Standardization
and Linked Open Data

- (Effective) semantic interoperability
- Data needs of the community are met

Looking for balance

The tension between semantics and interoperability

- Level of abstraction tends to increase (and names are changed retroactively)
- Semantics may be lost in favor of interoperability gain

- Presentation
- Rationale
- Methodology
- **Concepts standardisation**
- Conclusion and future work

- Multi-linguistic data

Eg. “poetic line”

- Diachronic corpora
- Different philological traditions

Eg. “dieresis”

Same term in different technical vocabularies

- Different terms depending on the theoretical background

for similar concepts

Eg. “apparatus”

- Multi-linguistic data

- Multi-linguistic data
 - Eg. “poetic line”

Issues



POSTDATA

Poetry Standardization
and Linked Open Data

- Multi-linguistic data
- Diachronic corpora

- Multi-linguistic data
- Diachronic corpora
 - Eg. “line group as a metrical unit”

- Multi-linguistic data
- Diachronic corpora
- Different philological traditions
 - Same term in different technical vocabularies
 - Eg. “diaeresis”

- Multi-linguistic data
- Diachronic corpora
- Different philological traditions
 - Same term in different technical vocabularies
 - Different terms depending on the theoretical background for similar concepts
 - Eg. “apparatus”

Semantic efficiency



POSTDATA

Poetry Standardization
and Linked Open Data

- Evaluation of which concepts should be integrated
- Evaluation of which concepts should be merged
- Evaluation of which concepts should be distinguished

Depending on the prosodic tradition:

- we count syllables
- we discern the distribution of stresses
- we analyse the length of the syllables

Metrical scheme

- “*syllabicMetricalScheme*”
- “*accentualMetricalScheme*”

- We might create a rigorous model from a semantic and philological point of view, but lack the acceptance of the community
- The conceptualization of certain elements might be perceived as “intrusive”

- Presentation
- Rationale
- Methodology
- Concepts standardisation
- **Conclusion and future work**

Research possibilities



POSTDATA

Poetry Standardization
and Linked Open Data

- Questions about the contents
- Questions about the structure

Research possibilities



POSTDATA

Poetry Standardization
and Linked Open Data

- Questions about the contents
- Questions about the structure
- Conclusion: a broader contextualization and a more accurate picture

Work in progress



POSTDATA

Poetry Standardization
and Linked Open Data

- Vocabulary alignment: match each concept with a term of a RDF vocabulary
- Encode the semantic model
- Report on how to enrich the repertoires' data with links to existent resources
- Develop documentation: Manuals, HowTos, Use case examples



helenabermudez@
linhd.uned.es

<http://postdata.linhd.es/>
<http://linhd.uned.es>
@linhduned

