

The POSTDATA Network of Ontologies for European Poetry

María Luisa Díez Platas^a, Salvador Ros^{a*}, Elena González-Blanco^b, Oscar Corcho^d, Helena Bermúdez^c, Javier de la Rosa^a, Alvaro Pérez^a

^a *POSTADATA Project. SSCC. Escuela Técnica Superior de Informática, UNED, Madrid, Spain*

^b *Coverwallet. Madrid, Spain*

^c *Section des Sciences du Langage et de l'Information, Université de Lausanne, Suisse*

^d *Ontology Engineering Group. ETSI Informáticos, Universidad Politécnica de Madrid, Madrid, Spain*

Abstract.

A relevant line of work in Digital Humanities is related to standardization processes to describe traditional concepts using computer-readable languages. In regard to Literary studies, poetry is a particularly complex domain due to, among other aspects, its special use of language. This paper presents a network of ontologies for capturing domain knowledge about poetry, including the poetic work itself, as well as its structural and prosodic components. The network also includes an ontology about dates, focusing on the special representational needs of literary works. This work is one of the results of the POSTDATA (Poetry Standardization and Linked Open Data) European Research Council project, which aims to provide a means for researchers on European poetry to publish and consume semantically-enriched data.

Keywords: European Poetry, Standardization, Network of Ontologies, Interoperability, Linked Open Data

1. Introduction

The need for standardization has increased significantly in different research fields, as a way of understanding and exchanging information in the corresponding domain. Many scientific disciplines have established formal protocols and languages, which have been quickly adopted and adapted to their particular problems. Some humanities and cultural disciplines have followed, however, an independent path in which creativity and tradition play an essential role. Literature, and especially poetry, is a clear reflection of this idiosyncrasy.

From the philological point of view, there is no uniform academic approach to analyse, classify or study

the different poetic manifestations, and the divergence of theories is even bigger when comparing poetry schools from different languages and periods. One of the most significant conceptual and terminological problems is that, even when a set of poetic works is formalized under a repertoire, each repertoire belongs to its poetical tradition, and each tradition has independently developed its analytical terminology, in some cases over the course of centuries [1]. The result of this uncoordinated evolution is a great variety of terminologies to explain similar metrical phenomena through the different poetic systems whose correspondences have been hardly studied. For example, the same quatrain of dodecasyllables can be encoded

* Corresponding author. E-mail: sros@scc.uned.es.

in different ways depending on the philological tradition:

(i.e. 12A12A12A12A or 4x(7pp+7p) or 4aaaa) or even named with a different meaning: “alexandrine” means a 14-syllable line in Spanish but only 12-syllables in French [2].

As a result, if a researcher were to look for quatrains of dodecasyllables in different traditions, it would be necessary to visit each database independently and then carry out different searches adapting the query to the conventions of the resources.

There is one additional drawback: research in this field is usually conducted in an individual and isolated manner, and there is a certain lack of communication with other areas of knowledge.

There are also significant technical issues, as these repertoires were created in different periods, and stand-alone collected databases drive most of them [3–8]. Interoperability among all these collections would be useful to perform comparative studies and to move a step forward beyond the modern philological state-of-the-art, to explain phenomena like the origins of vernacular poetry or the evolution from accentual to syllabic rhythmical patterns.

Although the current technical infrastructures are prepared to harvest such collections and provide access to them by search engines, it is necessary to standardize metadata and vocabularies at a philological level to be able to climb up the semantic layer and link data among different traditions [9–11].

In this context, the use of technologies applied to poetry is ground-breaking, as this way of representing distributed literary collections as machine-readable repositories will open the door to pose new research questions and to perform comparative philological analysis between heterogeneous poetic corpora with different formats.

All these difficulties and problems of access to poetic resources and, in short, the impossibility of having the means to process this information in a complete and efficient way have been the origin and the incentive for the conception of an ontology network of poetry [9,12].

For this purpose, we have extracted from a set of repertoires of different poetic traditions and periods, [13,14], the concepts and relationships necessary to

achieve the representation of a universal and complete poetry domain. From this study, we have identified areas of knowledge that are complementary to the core of poetry knowledge. We have modelled each of these areas as complementary ontologies. The result of the whole process has led us to the development of a network of ontologies for European poetry.

This paper presents the methodology carried out to build such a network of ontologies and the most significant ontologies of this network. This work is part of the results of the POSTDATA (Poetry Standardization and Linked Open Data) ERC project, which aims to provide a means for poetry researchers to publish and consume semantically-enriched data in the context of European poetry.

The document is structured as follows. In section §2, we identify some previous results related to ontologies in the literature, especially in the domain of poetry. Section §3 presents a description of the methodology used for the ontology network development. Section §4 presents a detailed description of the most relevant ontologies in this network. Section §5 presents potential applications of the network of ontologies. Finally, §6 outlines the conclusions and future work.

2. Related works

The progressive transformation of Humanities into “Digital Humanities” was accompanied by the creation of new standards, such as the Text Encoding Initiative TEI-XML¹, Dublin Core² or CIDOC-CRM³, among others, to describe traditional concepts with computer-readable languages. These systems are developing fast in several areas, such as digital text editions, libraries, or archives, and a significant number of projects exists that are working with them, such as TextGrid⁴, OpenEdition⁵ or Scholar Digital Editions (SDE)⁶.

Semantic web technologies have had great success in archives, libraries and museums (group known as LODLAM⁷). However, the application of these technologies to poetry is still limited [15,16]. As far as we know, there is no ontology yet that models all the necessary concepts about poetry and metrics.

¹ <https://tei-c.org/>

² <https://www.dublincore.org/specifications/dublin-core/>

³ <http://www.cidoc-crm.org/>

⁴ <https://textgrid.de/>

⁵ <https://www.openedition.org/>

⁶ <http://www.sd-editions.com/>

⁷ <http://lodlam.net/>

The first attempt to build a poetry ontology can be found in the ReMetca project [5], which defined a conceptual model for poetry and participated in the definition of TEI-Verse module. However, this ontology would need to be expanded and completed to reflect the different possibilities of poetic properties and relationships. The next closest work related to this topic is the CIDOC Conceptual Reference Model (CIDOC-CRM)⁸, an ontology that formally describes the concepts and relationships used to document cultural heritage. This model is more focused on the representation of museums' heritage works, although it contains concepts for the representation of entities such as people and places associated with the works. Other related ontologies are the Functional Requirements of Bibliographic Records (FRBR)⁹ and FRBRoo¹⁰. FRBR offers a perspective on the structure and relationships of bibliographic and authority records [17]. The most significant entities in this ontology are Work, Expression, Manifestation, and Item, which represent the different ways of conceiving a literary work as a text or physical resource. FRBRoo is an object-oriented version of FRBR combined with the CIDOC-CRM model, thus harmonizing information from museums, archives, and libraries as information relating to cultural heritage entities. These ontologies can cover the descriptive aspects of the works and their forms of expression and manifestation, but they do not cover structural aspects, literary analysis or prosody.

Another ontology aligned with FRBR is FaBio¹¹ (FRBR-aligned Bibliographic Ontology) that is conceived for the description of entities that are published or that are potentially publishable.

All these ontologies focus on bibliographic aspects of the works but do not model information that may come from the analysis of textual features.

Another form of semantic enrichment of literary and poetic texts is their annotation by means of markup languages. The Text Encoding Initiative (TEI) is one of the most relevant approaches in this field [19–21]. TEI has a verse module [18] with the elements needed to annotate forms and structures of poetic works. Although the relationship between ontological models and TEI has been taken into consideration in recent years [22–24], however, it is an open

field that we are still working because the description of poetry in TEI has to improve yet.

In regard to the analytical aspect of the works, we can find other ontologies; for example, the Lexicon Model for Ontologies (Lemon)¹² designed for modelling machine-readable dictionaries and lexicons. Lemon covers aspects of lexical decomposition, sentence structure, syntax, variation, morphology, and mapping of lexical ontology. The Gold¹³ ontology is a complete ontology for descriptive linguistics and formally describes the most basic categories and relationships used in the scientific description of human language. It tries to solve the problems of linguistic data tagging. One of the important features of this ontology is that it is applicable to all languages. The Rhetorical Annotation Ontology Project (RAOP)¹⁴, a specific domain ontology, is built for the annotation of speech figures and the rhetorical aspects of written and oral texts. Thus, it can be mapped to represent the structures of a rhetorical system. This project is one of the possible approaches that have been taken into account for the digitization of speech figures through the use of semantic technologies. RAOP covers rhetorical aspects that are considered in the model that has been built in the POSTDATA project.

Beyond this analysis of well-known ontologies in the areas of Digital Humanities, we have also made a general search for ontologies that may be relevant in usual ontology repositories, such as Linked Open Vocabularies¹⁵, Open Metadata Registry¹⁶, and Basel Register of Thesauri Ontologies & Classifications¹⁷.

This review shows the convenience to tackle the development of a new ontology network on poetry that includes all the entities and properties required to fully represent and annotate poetic works, while reusing some of the concepts and relationships defined in the aforementioned ontologies. Even more, we have proposed specializations or refinements to express more clearly the unique semantic features of the poetic works and their expressions.

⁸ <http://www.cidoc-crm.org/cidoc-crm>

⁹ <http://purl.org/vocab/frbr/core#>

¹⁰ <http://www.cidoc-crm.org/frbroo/home-0>

¹¹ <https://sparontologies.github.io/fabio/current/fabio.html>

¹² <https://www.lemon-model.net/index.php>

¹³ <http://purl.org/linguistics/gold/>

¹⁴ <http://bakulf.github.io/raop/>

¹⁵ <https://lov.linkeddata.es/>

¹⁶ <http://metadataregistry.org>

¹⁷ <http://www.bartoc.org/>

3. Ontology development

For the development of the ontology we have considered four scenarios of the Neon methodology [25]: scenario 2 for reusing and re-engineering non-ontological resources, scenario 3 that contemplates the reuse of ontological resources, scenario 4 for the reengineering of ontological resources, and scenario 7 for the reuse of ontological design patterns.

The first step for tackling this work was to build a conceptual domain model of European poetry, trying to have a precise picture of the domain.

For this purpose, we analysed a set of twenty-five repertoires described in [26], most of them available on the web. These repertoires were selected because they represent different poetry traditions, languages, prosodic systems, and cultures [27]. Most of these repertoires are the result of research projects and hence contain information gathered or generated by experts, what improves the reliability and robustness of information, categories, and structures. They are implemented using a variety of tools and formats, such as MySQL dumps, XSD and XML files, Perl scripts, and spreadsheets. To tackle this task, we used a twofold approach. We focused on using already existing models for data integration but arising problems related to consistency and disagreement [9,11] that we solved by the study of the poetry domain.

The obtained conceptual model covers both the descriptive and bibliographic aspects of poetic works. It includes details about textual transmission, as well as aspects related to prosody, literary and rhetorical analysis, the structures identified in the poems, significant elements for publications, and its relationships with music, as summarized in Figure 1. The result is a European Poetry Domain Model (DM-EP) with 40 entities, 494 attributes, and 409 relationships. A detailed description of the methodology used to construct the DM-EP can be found in [13,14].

This domain model was the starting point for the development of our ontology network. Due the complexity of the poetry domain, we decided to build a network of ontologies [28], where each module has a high degree of cohesion. That is, each module in the ontology network contains all the classes and

properties needed to define the relationships between its internal classes. In this way, the modules are self-contained, preserve the relationships within the module and avoid strong coupling with other modules.

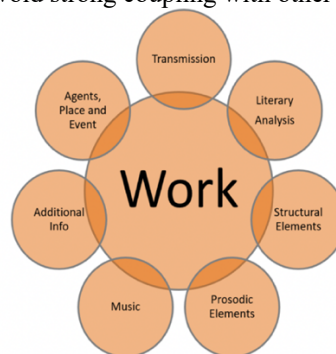


Figure 1. Areas of knowledge in the domain model.

The development process was carried out through an iterative-incremental model. Each ontology was built by the premise of reusing existing ontologies, aligning vocabularies and properties to facilitate its development, improving the semantic understanding of entities and facilitating interoperability. A general overview of the ontology network is presented in Figure 2.

After each iteration, the OWL implementation of each ontology was obtained. Besides, a graphical representation was created using a diagram of classes, data properties, and object properties.

It is important to mention that we have placed particular emphasis on defining explicitly both the domains and the ranges of all properties, so as to reduce ambiguity as much as possible, as well as including as many cardinality, universal and existential restrictions for classes as possible. For the definition of the ranges of some properties we have identified a set of controlled vocabularies (see Annex I), which have been specified using the data model of the Simple Knowledge Organization System (SKOS)¹⁸. These controlled vocabularies allow establishing standard terminologies that can be shared by the scientific community

¹⁸ <https://www.w3.org/2004/02/skos>.

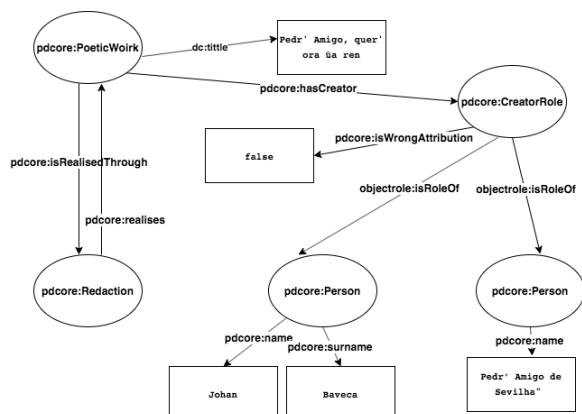


Figure 4 Example of authorship in a Poetic Work

- Line represents a unit of verse that usually ends in a visual or typographic break, and is generally characterized by its length and meter.
- Stanza represents a group of lines. Usually, this grouping forms the basic recurring metrical unit of a poem.
- Word represents words as list of syllables.
- Punctuation represents the punctuation symbols.
- Syllable represents a single unit of speech sound as written or spoken.

According to the semantics of the Ordered List ontology, these classes are subclasses of `olo:Slot` since all them are considered a slot of an ordered list. But also the three first are subclasses of `olo:OrderedList`, since they are considered a list of ordered elements, as shown in Figure 5.

4.3. *postdata-prosodicElements ontology (pdprosodic)*

This ontology contains the classes and properties needed to structure the information extracted in a prosodic analysis of poetic work. The complete ontology, published in [31], contains 10 new classes, 52 data properties, and 40 object properties.

The prosodic analysis of a poetic work contains information about the metrical patterns of a poem. These metrical patterns are defined in three levels: poem, stanza, and line. This ontology imports the *postdata-structuralElements* ontology because the information it provides is related to the metrical patterns of the line, the stanza, and the poem.

We have defined three classes, subclasses of `pdprosodic:Pattern`, that conceptualize the metric patterns that are analyzed in poetic works:

- `pdprosodic:LinePattern` models the metric pattern of the Line. Some important properties related to the line pattern are:
 - `pdprosodic:accentedVowels` to represent stressed vowels in the order in which they occur in the text.
 - `pdprosodic:countingMetricalScheme` to represent the metrical scheme according to the number of syllables.
 - `pdprosodic:grammaticalStressPattern` to represent patterns based on the position of the expected stresses according to grammar rules including the distribution of weak and strong positions.
- `pdprosodic:StanzaPattern` summarizes some specific properties of the stanza as `pdprosodic:rhymeScheme`. It represents the rhyme scheme, the pattern of rhymes at the end of each line of the stanza. One of the most common conventions is to use letters to indicate which lines rhyme.
- `pdprosodic:WorkPattern` shares some of the properties defined in `LinePattern` and `StanzaPattern` but define more specific properties, such as `pdprosodic:presentRhymeMatching` that allows the categorization of the poetic work according to the degree in which the different rhyming sounds match (i.e., assonant, consonant). This property also applies to the `StanzaPattern` class.

These ontologies have been completed with more classes used to represent prosodic analysis data. Some of these classes are:

- `pdprosodic:Rhyme` to represent the repetition of similar sounding words occurring at the end of lines in poems or songs.
- `pdprosodic:Foot` to represent the unit of poetic meter in most Indo-European poetic traditions, including the English syllabic verse and the quantitative compass of Greek and ancient Latin classical poetry.

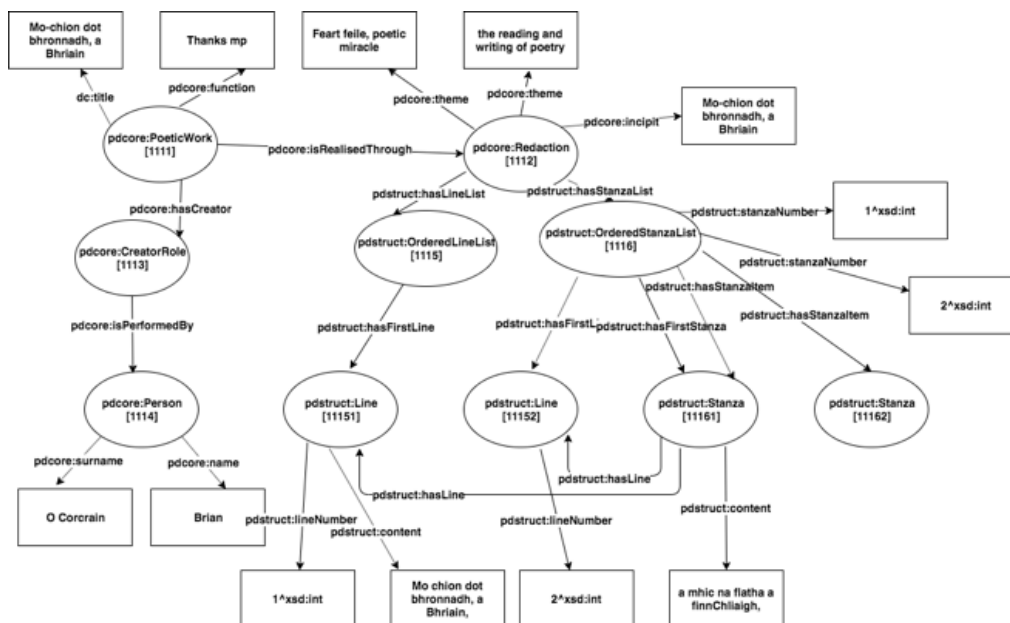


Figure 5 Example of list of lines and stanzas in structural ontology

- `pdprosodic:metricalEncoding` to represent the notation employed to represent a metrical pattern, for example, the plus sign to encode the strong positions of syllables.

Finally, in poetry it is common to find poetic licenses that affect the number of syllables, for example, dieresis, syneresis, synalepha, or hiatus. These figures consist of altering the writing or pronunciation of words without altering their meaning. The `pdprosodic:Metaplasm` class allows knowing the type of prosodic figure, where it is allocated, and which words are involved in the phenomena.

4.4. postdata-dates ontology

Depending on the composition period, it may be difficult to date with exactitude a poetic work or its manifestations. Moreover, in ancient or anonymous publications, it is not always possible to find a particular date, and it may be necessary to establish ranges or make a suggestion on the likelihood of a date. This problem is present when the form of transmission or preservation does not facilitate the trace of a date of composition. For this reason, we have proposed an independent and reusable ontology, which may be applied not only in the literary domain but also in any other cultural heritage domain, which covers special dating needs. In this ontology, two classes are provided:

- `pddates:DateEntity` to represent a temporal entity associated with a poetic work, its manifestations, or an event.
- `pddates:DateExpression` is the base of a class hierarchy. This class and its sub-classes provide several modes to represent a date related to the creation or whatever event associated with an entity.

The ontology, published in [32], contains 7 classes, 7 data properties, and 2 object properties, as shown in Figure 6. .

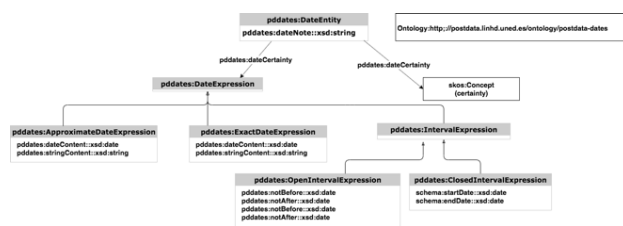


Figure 6 Postdata-dates Ontology.

In Annex II we present an example of how our ontology network can be used for the song "Mais nos faz Santa Maria a seu Fillo perdõar", written by Alfonso X el Sabio.

5. Potential applications of the ontologies

Besides the development of the ontology network that has been presented in this paper, the POSTDATA project is also developing the PoetryLab. PoetryLab is an extensible open-source toolkit for poetry analysis. At this moment it can perform the syllabification, scansion (extraction of stress patterns), enjambment detection (syntactical units split into two lines), rhyme detection, and medieval named entity recognition for Spanish poetry. This toolkit achieves state of the art performance in the tasks for which reproducible alternatives exist. Moreover, it is designed to aggregate as many tools as needed for poetry analysis.

PoetryLab uses this network of ontologies as its underlying data model. It uses RDF triples according to this ontology network as source and generates new RDF triples as results. So, the ontology network is currently the data persistence and interoperability base of PoetryLab.

Another of the objectives of the POSTDATA project is to share poetry knowledge as linked data. We have started to populate the ontology. For this purpose, we are currently mapping some poetry databases or repertories: *MedDB Base de Datos da Lirica Profana Galego-Portuguesa* [7], *Bardic Poetry database* [33], *ReMetCa*, a medieval Spanish metrical repertory [21], *Musisque Deoque* repository [34], and *Pede Certo*, a digital Latin metre database [6].

To improve the use of the network of ontologies, we will also deploy an OMEKA server. This action will allow including poetry ibjetc by using the network of ontologies. Omeka has been selected for its ease of use and its high adoption in the digital humanities community.

6. Conclusions and future work

In this paper, we have presented the most significant ontologies from the network of ontologies for European Poetry developed in the framework of the POSTDATA ERC Project: `postdata-core`, `postdata-prosodic`, `postdata-structural`, and `postdata-dates`. This network of ontologies is composed of four more ontologies, which are currently under construction: `postdata-literaryAnalysis`, `postdata-transmission`, `postdata-music`, and `postdata-additionalContent`.

To build this network, we have started from the development of a domain model for European poetry. This model was developed for three years, based on the analysis of 25 repertories. This domain model is the most complete data model specific of poetry ever done.

The ontologies have been developed using state-of-the-art ontology engineering methodologies and published according to best practices and recommendations for Linked Data vocabulary publishing.

The ontologies described in this work are Postdata core ontology, prosodic ontology, structural ontology, and date ontology. The first three are related to the description of the poetic work and its essential properties such as the prosodic elements. The fourth is an ontology for capturing the dating knowledge that in literature entails an especial complexity.

These ontologies are the semantic and data persistence base of PoetryLab, an extensible open-source toolkit for poetry analysis developed in the POSTDATA Project as well.

Finally, we are mapping different poetry databases and repertories to the ontologies with the main aim of populating the ontologies and sharing the information in an interoperable RDF format. This is a complex task because building such knowledge graph requires creating RML mappings. We are currently working on two aspects: the revision and publication of the rest of the ontologies that form part of the network and the definition of the identified controlled vocabularies.

The ontologies we are currently working on are: The ontology of literary analysis that covers the aspects necessary to describe the information that is extracted from the literary analysis of a poetic work. It focuses especially on rhetorical figures or figures of speech, such as tropes and schematics. The Postdata-additionalFeatures ontology will cover the aspects that increase the expressiveness of the works or add contexts that appear regularly, such as paratexts and illustrations and the transmission ontology that covers the classes and properties related to the transmission of the works and other bibliographic aspects such as primary sources, bibliographic sources and the location of the work under study in a bibliographic element.

The ontology network will also take into account a feature that is present in many poetic works, which is the presence of musical accompaniment. In the music ontology, we will represent the features related to music that can enrich the text and play an important role as complementary information.

7. Acknowledgements

The authors acknowledge the support of the Starting Grant research project Poetry Standardization and Linked Open Data: POSTDATA (ERC-2015-STG-679528). This project is funded by the European Research Council (ERC) under the research and innovation program Horizon2020 of the European Union.

8. References

- [1] E. González-Blanco, L. Seláf, Megarep: A comprehensive research tool in medieval and renaissance poetic and metrical repertoires, *Humanit. Xarxa Món Mediev. Web Mediev. World.* (2014) 321–332.
- [2] E. González-Blanco García, M. Manailescu, S. Ros, From syllables, lines and stanzas to linked open data: standardization, interoperability and multilingual challenges for digital humanities, (2016). <http://e-spacio.uned.es/fez/view/bibliuned:363-Egonzalez11> (accessed December 5, 2018).
- [3] L. Leonardi, Repertorio della tradizione poetica italiana dai Siciliani a Petrarca, (n.d.). <http://www.mirabileweb.it/> (accessed September 15, 2019).
- [4] L. Grijp, Dutch Song Database, (n.d.). <http://www.liederenbank.nl/index.php?lan=en> (accessed September 15, 2019).
- [5] E. González-Blanco García, J.L. Rodríguez, ReMetCa: a TEI based digital repertory on Medieval Spanish poetry, (2013). <http://e-spacio.uned.es/fez/view/bibliuned:365-Egonzalez6> (accessed September 15, 2019).
- [6] M. Colombi, Pedecerto, Pedecerto Metrica Lat. Digit. (n.d.). <http://www.lutessa.it/pedecerto/public/> (accessed September 15, 2019).
- [7] M. Brea, MedDB - Inicio, MedDB Base Datos Lírica Profana Galego-Port. (n.d.). <http://bernal.cirp.gal/ords/f?p=MEDDB3:2> (accessed September 15, 2019).
- [8] S. Asperti, F. Zinelli, Bibliografia Elettronica dei Trovatori - v.2.5 - 2012, (n.d.). http://www.bedt.it/BEdT_04_25/index.aspx (accessed September 15, 2019).
- [9] H. Bermúdez-Sabel, M.L. Díez Platas, S. Ros Muñoz, E. González-Blanco, Towards a common poetry model: challenges and solutions, in: *Utrecht*, 2019. <https://doi.org/10.5281/zenodo.3335509>.
- [10] E. González-Blanco, G. del Rio Riande, C. Martínez Cantón, Linked Open Data To Represent Multilingual Poetry Collections. A Proposal To Solve Interoperability Issues Between Poetic Repertoires, in: *Proc. 5th Workshop Linked Data Linguist.*, ELRA, Paris, 2016: pp. 77–80. <https://doi.org/10.5281/zenodo.60767>.
- [11] H. Bermúdez-Sabel, M. Curado Malta, E. González-Blanco, Towards Interoperability in the European Poetry Community: The Standardization of Philological Concepts, in: J. Gracia, F. Bond, J.P. McCrae, P. Buitelaar, C. Chiarcos, S. Hellmann (Eds.), *Lang. Data Knowl.*, Springer International Publishing, 2017: pp. 156–165.
- [12] E. Gonzalez-Blanco, S. Ros, POSTDATA – Poetry Standardization and Linked Open Data, *Postdata.* (2016). <http://postdata.linhd.es/> (accessed August 1, 2018).
- [13] M. Curado Malta, E. Gonzalez-Blanco, C. Martínez Cantón, G. del Rio Riande, A Common Conceptual Model for the Study of Poetry in the Digital Humanities, in: *Digit. Humanit. 2017 Conf. Abstr.*, McGill University / Université de Montréal, Montréal, Canada, 2017: pp. 210–212. <https://dh2017.adho.org/abstracts/148/148.pdf>.
- [14] M. Curado Malta, E. González-Blanco, C.I. Martínez Cantón, G. del Rio Riande, Digital repertoires of poetry metrics: towards a linked open data ecosystem, in: E. William De Luca, P. Bianchini (Eds.), *Digit. Humanit. Digit. Curation Proc. First Workshop Digit. Humanit. Digit. Curation Co-Located 10th Conf. Metadata Semant. Res. MTSR 2016*, 2016. <http://ceur-ws.org/Vol-1764/1.pdf>.
- [15] P. Bootz, S. Szonieczky, Toward an ontology of the Field of Digital Poetry, (2008). <https://hal-univ-paris8.archives-ouvertes.fr/hal-01011128> (accessed September 15, 2019).
- [16] DHQ: Digital Humanities Quarterly: Ontologies and Logic Reasoning as Tools in Humanities?, (n.d.). <http://www.digitalhumanities.org/dhq/vol/3/4/000068/000068.html> (accessed September 15, 2019).
- [17] D.B. Tillett, What is FRBR? A conceptual model for the bibliographic universe, *Aust. Libr. J.* 54 (2005) 24–30. <https://doi.org/10.1080/00049670.2005.10721710>.
- [18] 6 Verse - The TEI Guidelines, (n.d.). <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/VE.html> (accessed September 5, 2019).
- [19] Henrik Ibsens skrifter: Forside, (n.d.). <https://www.ibsen.uio.no/> (accessed September 15, 2019).

- [20] Lyrik des deutschen Mittelalters (LDM), (n.d.). <http://www.ldm-digital.de/> (accessed September 15, 2019).
- [21] ReMetCa: Repertorio Digital de la Métrica Medieval Castellana, Remetca. (n.d.). <http://proyectorremetca.weebly.com/> (accessed September 15, 2019).
- [22] C.-E. Ore, Ø. Eide, TEI and cultural heritage ontologies: Exchange of information?, *Lit. Linguist. Comput.* 24 (2009) 161–172. <https://doi.org/10.1093/lit/fqp010>.
- [23] Ø. Eide, Ontologies, Data Modeling, and TEI, *J. Text Encoding Initiat.* (2014). <https://doi.org/10.4000/jtei.1191>.
- [24] Ontologies SIG – TEI: Text Encoding Initiative, (n.d.). <https://tei-c.org/activities/sig/ontologies/> (accessed September 20, 2019).
- [25] M.C. Suárez-Figueroa, NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse, phd, Facultad de Informática (UPM), 2010. <http://oa.upm.es/3879/> (accessed September 27, 2018).
- [26] Postdata ERC project, TOWARDS A NETWORK OF ONTOLOGIES FOR THE EUROPEAN POETRY - POSTDATA, (n.d.). <http://postdata-prototype.linhd.uned.es/methodology.php> (accessed September 17, 2019).
- [27] Postdata ERC project, POSTDATA Repertories, Google My Maps. (n.d.). https://www.google.com/maps/d/viewer?mid=15MAs3IVHIOk-eWUfBWXPB_prHbE (accessed September 21, 2019).
- [28] Postdata ERC project, Network of ontologies - POSTDATA, (n.d.). <http://postdata-prototype.linhd.uned.es/ontology.php> (accessed September 17, 2019).
- [29] Postdata ERC project, postdata-core ontology, Postdata-Core Ontol. (n.d.). <http://postdata.linhd.uned.es/ontology/postdata-core/documentation/index.html> (accessed September 17, 2019).
- [30] Postdata ERC project, postdata-structural ontology, (n.d.). <http://postdata.linhd.uned.es/ontology/postdata-structuralElements/documentation/index-en.html> (accessed October 25, 2019).
- [31] Postdata ERC project, postdata-prosodic ontology, (n.d.). <http://postdata.linhd.uned.es/ontology/postdata-prosodicElements/documentation/index-en.html> (accessed September 17, 2019).
- [32] Postdata ERC project, postdata-dates ontology, (n.d.). <http://postdata.linhd.uned.es/ontology/postdata-dates/documentation/index-en.html> (accessed October 25, 2019).
- [33] K. Simms, Bardic Poetry Database, (n.d.). <https://bardic.celt.dias.ie/> (accessed October 26, 2019).
- [34] Musisque Deoque, (n.d.). <http://mizar.unive.it/mqdg/public/> (accessed October 26, 2019).

ANNEX I

As a result of the design of the ontology network, we have identified a set of controlled vocabularies used as ranges of the properties described in table 1.
(See <https://github.com/linhd-postdata/skos-codelists>)

Table 1. Controlled Vocabularies detected in the network of ontologies

Ontology	Class	Property
postdata-core ontology (pdcore)	PoeticWork	- genre - poeticType - authorEducationLevel
	Redaction	- typeOfTextualElement
	Person	- gender, - literaryPeriod - school - socialStatus - religiousAffiliation - birthPlaceCentainty
	Role	- roleFunction - typeOfCharacter
	CreatorRole	- typeOfDesignation
postdata-structuralElements ontology (pdstruct)	Stanza class	- typeOfStanza - typeOfStanzaEdition
	Word	- partOfSpeech
	Syllable	- nucleusType
postdata-prosodicElements ontology (pdprosodic)	Patterns	- clausulaSchemeType - metricalCategory - metricalComplexity - metricalContext - rhymeDispositionType - versificationType
	Line Stanza	- feetType - metricalType
	Foot FootDivision	- footType - footUnitType - clausula
	RhymeMatching	- typeOfRhymeMatching
	Metaplasm	- typeOfMetaplasm

ANNEX II

The example presented in Listing 1, corresponds to the song "Mais nos faz Santa María a séu Fillo perdõar", written by Alfonso X, el Sabio. The analyzed redaction has been extracted from "Cantigas de Santa Maria for Singers" (<http://www.cantigasdesantamaria.com>), created by Andrew Casson. The RDF implementation represents the Poetic Work whose author is Alfonso X and the redaction prepared by Andrew Casson. The date is shown as an open interval, since it is not known exactly. An extract of the structure in stanzas and lines is also presented. A characteristic that should be highlighted is that this poetic composition is a song, with a refrain. The refrain is the first stanza, but it is repeated after all the others. For this reason, the refrain has no number assigned to it and its lines are indicated as part of a stanza. Finally, the patterns of the Work, the stanzas and the lines, and the rhyme are also represented.

The relationship between the Poetic Work and the Redaction is represented with the object property `pdcore:is-RealisedThrough`. The object property `pdcore:hasCreator` is used to represent both the author of the Poetic Work and the creator of the Redaction, which are both of type `Agent (Person)`. All of them are related with `pdcore:CreatorRole` objects.

So, the author of the Poetic Work is a `Person` who has:

- Name (`pdcore:name`)- "Alfonso X"
- Nickname (`pdcore:addName`) - "el Sabio"
- A sentence linked to the name indicating its position (`pdcore:nameLink`) – "Rey de Castilla"
- A link to a Virtual International Authorities File (VIAF) (`schema:url`) – "http://viaf.org/viaf/66476694/#Alfonso_X,_Rey_de_Castilla"

On the other hand, the creator of the Redaction is a `Person` who has:

- Name (`pdcore:name`) - "Andrew"
- Surname (`pdcore:addName`) - "Casson"
- A Link to the personal page (`schema:url`) – <https://independent.academia.edu/AndrewCasson>

The date of the Poetic Work is shown as an open interval, since it is not known exactly and it is associated with the Poetic Work through the object property `date` that refers a `pddates:DateEntity`, expressed as an `pddates:OpenIntervalExpression`, with the following properties:

- not Before – "1270-01-01"
- not After – "1282-12-31"

In addition, because this work is a piece of medieval literature with a difficult dating, a note is created about how and who identified the date:

- Note (`pddate:dateNote`) – "Dates set by Walter Mettmann....."

The redaction, in addition to the `url` and the creator, is described by the structure of the text that it represents.

In this sense, the structure of the redaction is a list of stanzas (`pdstruct:OrderedStanzaList`) related to the redaction through the property `pdstruct:hasStanzaItem` from the ontology of structural elements.

A characteristic to emphasize is that this poetic composition is a song, with a refrain. The refrain is the first stanza, but it is repeated after all the others. For this reason, the refrain has no number assigned to it and its lines are indicated as part of a refrain.

Therefore, the stanzas are described by:

- The stanza number, a positive integer indicating its position in the list (`pdstruct:stanzaNumber`).
- If a stanza is a refrain, then
 - Is Refrain (`pdstruct:isRefrain`) – "true"
- If a stanza is not a refrain but must be followed by the refrain of which only one instance will be created

- Is Refrain Omitted (`pdstruct:isRefrainOmitted`) – “true”
- The next stanza (`pdstruct:nextStanza`)
- The first and last line of the stanza (`pdstruct:hasFirstLine`, `pdstruct:hasLastLine`).

The lines that contain the text and other features that are listed below:

- The text of the line (`pdstruct:content`) - "a séu Fillo perdõar,"
- The position of the line (`pdstruct:lineNumber`) - "1"
- If the line is part of a refrain it is indicated (`pdstruct:belongsToRefrain`) – “true”
- The first and last line of the stanza (`pdstruct:hasFirstLine`, `pdstruct:hasLastLine`).

Prosodic features are covered at different levels. First, the prosodic analysis of the redaction, stanza and line with the property (`pdprosodic:isAnalysedThrough`) to refer a `pdprosodic:WorkPattern` object, a `pdprosodic:StanzaPattern` object and a `pdprosodic:LinePattern` object, respectively. The patterns are described as follows:

- `pdprosodic:WorkPattern`
 - If the redaction has a refrain (`pdstruct:hasRefrain`) – “true”
 - The metric scheme (`pdprosodic:countingMetricalScheme`) - "77777777"
 - The rhyme scheme (`pdprosodic:rhymeScheme`) - "ABAB/cdcdcdcdB"
 - The number of stanzas (`pdprosodic:numberOfStanzas`) – “5”
- `pdprosodic:StanzaPattern`
 - The metric scheme (`pdprosodic:countingMetricalScheme`) - "7’ 7 7’ 7"
- `pdprosodic:LinePattern` (different for each line)
 - The metric scheme (`pdprosodic:countingMetricalScheme`) - "7"
 - Phonetic transcription (`pdprosodic:phoneticTranscription`) - "mais nos fadzsanta maRi.a".

The rhyme is also represented by an object of the `pdprosodic:Rhyme` class object associated with the line.

Some of its properties are the following:

- The label (`pdprosodic:label`) – “A”
- The associated phonemes (`pdprosodic:ending`) – “i.a”

Listing 1. Example using the network of ontologies

```

@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix pdstruct: <http://postdata.linhd.uned.es/ontology/postdata-structuralElements#> .
@prefix pdcore: <http://postdata.linhd.uned.es/ontology/postdata-core#> .
@prefix pdprosodic: <http://postdata.linhd.uned.es/ontology/postdata-prosodicElements#> .
@prefix pddates: <http://postdata.linhd.uned.es/ontology/postdata-dates#> .
@prefix pdcore: <http://postdata.linhd.uned.es/ontology/postdata-core#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix schema: <http://schema.org/> .
@prefix objectrole: <http://www.ontologydesignpatterns.org/cp/owl/objectrole.owl#> .

<http://postdata.linhd.uned.es/resource/PoeticWork/Cantiga3> a pdcore:PoeticWork ;
    dc:title "Cantiga 3: Mais nos faz Santa Maria" ;
    pdcore:hasCreator <http://postdata.linhd.uned.es/resource/CreatorRole/Cantiga3-1> ;
    pdcore:isRealisedThrough <http://postdata.linhd.uned.es/resource/Redaction/Cantiga3-1> ;
    pdcore:date <http://postdata.linhd.uned.es/resource/DateEntity/Cantiga3> .

<http://postdata.linhd.uned.es/resource/Redaction/Cantiga3-1> a pdcore:Redaction ;
    schema:url <http://www.cantigasdesantamaria.com/csm/3> ;
    pdcore:sourceNote "The analyzed redaction has been taken from Cantigas de Santa Maria for Singers
(http://www.cantigasdesantamaria.com), a complete edition of the lyrics and music of the 13th century
Cantigas de Santa Maria of Alfonso X El Sabio, specially prepared for singers and instrumentalists. The
analyzed cantiga (http://www.cantigasdesantamaria.com/csm/3) has a refrain that is presented at the
beginning and repeated behind each stanza" ;
    pdcore:hasCreator <http://postdata.linhd.uned.es/resource/CreatorRole/Cantiga3-2> ;
    pdstruct:hasStanzaList <http://postdata.linhd.uned.es/resource/OrderedStanzaList/Cantiga3-1> ;
    pdstruct:hasFirstStanza <http://postdata.linhd.uned.es/resource/Stanza/Cantiga3-1> ;
    pdstruct:hasLastStanza <http://postdata.linhd.uned.es/resource/Stanza/Cantiga3-5> ;
    pdprosodic:isAnalysedThrough <http://postdata.linhd.uned.es/resource/WorkPattern/Cantiga3-1> ;
    pdprosodic:isAnalysedThrough <http://postdata.linhd.uned.es/resource/StanzaPattern/Cantiga3-1> ;.

<http://postdata.linhd.uned.es/resource/CreatorRole/Cantiga3-1> a pdcore:CreatorRole ;
    objectrole:isRoleOf <http://postdata.linhd.uned.es/resource/Person/Alfonso_X> .

<http://postdata.linhd.uned.es/resource/CreatorRole/Cantiga3-2> a pdcore:CreatorRole ;
    objectrole:isRoleOf <http://postdata.linhd.uned.es/resource/Person/Casson> .

<http://postdata.linhd.uned.es/resource/Person/Alfonso_X> a pdcore:Person ;
    pdcore:name "Alfonso X " ;
    pdcore:addName "El Sabio" ;
    pdcore:nameLink "Rey de Castilla" ;
    schema:url <http://viaf.org/viaf/66476694/#Alfonso_X,_Rey_de_Castilla> .

<http://postdata.linhd.uned.es/resource/Person/Casson> a pdcore:Person ;
    pdcore:name "Andrew" ;
    pdcore:surname "Casson" ;
    schema:url <https://independent.academia.edu/AndrewCasson> .

<http://postdata.linhd.uned.es/resource/DateEntity/Cantiga3> a pddates:DateEntity ;
    pddates:isExpressedAs <http://postdata.linhd.uned.es/resource/OpenIntervalExpression/Cantiga3> ;

```

pddates:dateNote "Dates set by Walter Mettmann. Reference: W. Mettmann, Introduction to his ed. by Alfonso X, Cantigas de Santa María, Madrid: Castalia, 1986, I, p. 24." .

<http://postdata.linhd.uned.es/resource/OpenIntervalExpression/Cantiga3> a pddates:OpenIntervalExpression ;
pddates:notBefore "1270-01-01"^^xsd:date ;
pddates:notAfter "1282-12-31"^^xsd:date .

<http://postdata.linhd.uned.es/resource/WorkPattern/Cantiga3-1> a pdprosodic:WorkPattern ;
pdstruct:hasRefrain "true"^^xsd:boolean ;
pdprosodic:countingMetricalScheme "77777777";
pdprosodic:rhymeScheme "ABAB/cdcddcdcdB";
pdprosodic:numberOfStanzas "5"^^xsd:positiveInteger.

<http://postdata.linhd.uned.es/resource/StanzaPattern/Cantiga3-1> a pdprosodic:StanzaPattern ;
pdprosodic:countingMetricalScheme "7' 7 7' 7".

<http://postdata.linhd.uned.es/resource/OrderedStanzaList/Cantiga3-1> a pdstruct:OrderedStanzaList ;
pdstruct:hasStanzaItem <http://postdata.linhd.uned.es/resource/Stanza/Cantiga3-1> , <http://postdata.linhd.uned.es/resource/Stanza/Cantiga3-2> , <http://postdata.linhd.uned.es/resource/Stanza/Cantiga3-3> , <http://postdata.linhd.uned.es/resource/Stanza/Cantiga3-4> , <http://postdata.linhd.uned.es/resource/Stanza/Cantiga3-5> , <http://postdata.linhd.uned.es/resource/Stanza/Cantiga3-6> ;
pdcore:notes "Because of the musical features of this expression or redaction, this strophe is a refrain, which is repeated after each other stanza. For this reason, a stanza number has not been assigned, and this stanza is omitted in all places where it should appear according to the source being analyzed".

<http://postdata.linhd.uned.es/resource/Stanza/Cantiga3-1> a pdstruct:Stanza ;
pdstruct:hasFirstLine <http://postdata.linhd.uned.es/resource/Line/Cantiga3-1> ;
pdstruct:hasLastLine <http://postdata.linhd.uned.es/resource/Line/Cantiga3-4> ;
pdstruct:isRefrain "true"^^xsd:boolean .

<http://postdata.linhd.uned.es/resource/Stanza/Cantiga3-2> a pdstruct:Stanza ;
pdstruct:stanzaNumber "1"^^xsd:positiveInteger ;
pdstruct:nextStanza <http://postdata.linhd.uned.es/resource/Stanza/Cantiga3-3>;
pdstruct:hasFirstLine <http://postdata.linhd.uned.es/resource/Line/Cantiga3-5> ;
pdstruct:hasLastLine <http://postdata.linhd.uned.es/resource/Line/Cantiga3-12>;
pdstruct:isRefrainOmitted "true"^^xsd:boolean .

The remaining stanzas are omitted due to the extension of the code

<http://postdata.linhd.uned.es/resource/Line/Cantiga3-1> a pdstruct:Line ;
pdstruct:content "Mais nos faz Santa María" ;
pdstruct:lineNumber "1"^^xsd:positiveInteger ;
pdstruct:belongsToRefrain "true"^^xsd:boolean ;
pdprosodic:isAnalysedThrough <http://postdata.linhd.uned.es/resource/LinePattern/Cantiga3-1> ;
pdprosodic:presents <http://postdata.linhd.uned.es/resource/Rhyme/Cantiga3-1> .

<http://postdata.linhd.uned.es/resource/Line/Cantiga3-2> a pdstruct:Line ;
pdstruct:content "a séu Fillo perdôar," ;
pdstruct:lineNumber "2"^^xsd:positiveInteger ;
pdstruct:belongsToRefrain "true"^^xsd:boolean ;
pdprosodic:isAnalysedThrough <http://postdata.linhd.uned.es/resource/LinePattern/Cantiga3-2> ;
pdprosodic:presents <http://postdata.linhd.uned.es/resource/Rhyme/Cantiga3-2> .

The remaining lines are omitted due to code extension


```
<http://postdata.linhd.uned.es/resource/LinePattern/Cantiga3-1> a pdprosodic:LinePattern ;  
    pdprosodic:countingMetricalScheme "7" ;  
    pdprosodic:phoneticTranscription "mais nos fadzsanta maRi.a".  
# The remaining line patterns are omitted due to the extension of the code
```

```
<http://postdata.linhd.uned.es/resource/Rhyme/Cantiga3-1> a pdprosodic:Rhyme ;  
    pdprosodic:label "A" ;  
    pdprosodic:ending "i.a".
```

```
<http://postdata.linhd.uned.es/resource/Rhyme/Cantiga3-2> a pdprosodic:Rhyme ;  
    pdprosodic:label "B" ;  
    pdprosodic:ending "aR".  
# The remaining lines are omitted due to the extension of the code
```