



## BUILDING A METRICAL ONTOLOGY AS A MODEL TO LINK DIGITAL POETIC REPERTOIRES

González-Blanco, Elena

Universidad Nacional de Educación Distancia, España

Seláf, Levente

ELTE University, Budapest

Del Rio Riande, María Gimena

Secrit-CONICET, Argentina

Martínez Cantón, Clara Isabel

Universidad Nacional de Educación Distancia, España

Martos Pérez, María Dolores

Universidad Nacional de Educación Distancia, España

Category: Long Paper

Session: 4

Date: 2014-07-10

Time: 09:00:00

Room: 414 - Amphimax

### 1. Introduction

#### 1.1. Overview

The technology of semantic web seems to be a suitable environment to offer solutions for linking poetic repertoires that belong to different European traditions and literatures (González-Blanco & Seláf 2013)<sup>[1]</sup>. The problem of linking poetic repertoires is not simple, as there are not only technical issues involved, but also conceptual and terminological problems: each repertoire belongs to its own poetical tradition and each tradition has developed its own analytical terminology for years in a different and independent way. The result of this uncoordinated evolution is a bunch of varied terminologies to explain analogous metrical phenomena through the different poetic systems whose correspondences have been hardly studied.

#### 1.2. Methodology

The aim of this paper is to present a model able to serve as a uniform solution for terminological issues in order to build a solid semantic structure as a basis to link the different poetic systems. This structure will be used to publish repertoires on the web in a structured format and using open standards in order to build an open-source and collaborative platform based on a poetic ontology which lets interoperability among the different European metrical repertoires with different applications, such as faceted searches based on SPARQL or different kinds of visualizations, very helpful for comparative analysis.

The first step to organize and manage repertoires and database systems was the construction of conceptual schema to define their basic entities and relationships. The ER (Entity-Relationship) data model is the most commonly used for this purpose, together with the data model based on records for the logical implementation (Elmasri & Navathe 2011, 27-ss)<sup>[2]</sup>, which is also widely accepted.

To implement this conceptual model, the project ReMetCa (Digital Repertoire on Medieval Spanish Poetry: [www.uned.es/remetca](http://www.uned.es/remetca)) has tested different systems (commercial, free, open-source, and proprietary). The final decision, after experimenting with Oracle Express Edition (González-Blanco & Rodríguez 2013)<sup>[3]</sup>, has been MySQL combined with a XML tagging using the TEI-verse module. The relationship between ontological models and TEI is being taken into consideration very seriously in the last years, as it is shown by the activity of the SIG ontologies group [wiki.tei-c.org/index.php/SIG:Ontologies](http://wiki.tei-c.org/index.php/SIG:Ontologies) and the specific papers published on this topic (Eide & Ore 2007)<sup>[4]</sup>. There are also projects that have applied these techniques to the study and analysis of medieval documents (Ciula, Spence & Vieira 2008)<sup>[5]</sup>.

### 2. Getting Started

From the three levels described (conceptual, logical and physical), this paper will focus on the first layer: the semantic description with the design of the semantic ontology, whose elements will be extensible and reusable for its application to other poetic repertoires. The conceptual model, designed on the basis of ReMetCa, will be transferred to the semantic Web as Linked Open Data. The abstraction of this initial model is prepared to be amplified with the necessary fields and terms to define metrical phenomena which are not shown in the Spanish poetic system or in the other repertoires which have been taken into account to design this first version of the semantic prototype. In order to enlarge its horizons, structure, description and contents, datasets of the following corpora have also been taken into account:

The Cantigas de Santa Maria Database: [csm.mml.ox.ac.uk](http://csm.mml.ox.ac.uk)

Analecta Hymnica Digitalia: database on Medieval Latin poetry: [webserver.erwin-rauner.de/crophius/Analecta\\_conspectus.htm](http://webserver.erwin-rauner.de/crophius/Analecta_conspectus.htm)

Bibliografia Elettronica dei Trovatori: [w3.uniroma1.it/bedt/BEdT\\_03\\_20](http://w3.uniroma1.it/bedt/BEdT_03_20)

Le Nouveau Naetebus: database on French narrative Medieval poetry : [www.nouveaunaetebus.elte.hu](http://www.nouveaunaetebus.elte.hu)

Répertoire de la Poésie Hongroise Ancienne (RPHA) : Repertoire on Medieval Hungarian poetry: [rpha.elte.hu](http://rpha.elte.hu)

MedDB: Lírica Profana Galego-Portuguesa [www.cirp.es/pls/bdo2/f?p=MEDDB2](http://www.cirp.es/pls/bdo2/f?p=MEDDB2)

Corpus rhythmorum musicum (IV-IX seculo): database on Latin Medieval poetry accompanied with music [www.corimu.unisi.it](http://www.corimu.unisi.it)

Skaldic poetry of Scandinavian Middle Ages: <https://www.abdn.ac.uk/skaldic/db.php>

English Broadside Ballad: [ebba.english.ucsb.edu/](http://ebba.english.ucsb.edu/)

Dimev: Digital index of medieval English verse: [www.cddc.vt.edu/host/imev/record.php?recID=6768](http://www.cddc.vt.edu/host/imev/record.php?recID=6768)

To implement the conceptual model, the project uses one of the most recognized standards for semantic Web description: the Ontology Web Language (OWL), developed by W3C as an extension of RDFS. OWL is used to define the different classes, their properties and the instances of classes. It integrates sets of predefined metadata using namespaces. The set joins not only traditional well-known initiatives, such as Dublin Core, MARC or TEI, but also local proposals such as those used by some of the digital poetic repertoires that serve as a basis for this project. The TEI-Verse module<sup>[6]</sup> plays also an important role, due to the use that several repertoires have made of it, such as Henrik Ibsens (<http://www.ibsen.uio.no>) or the project of Lyrik des Hohen Mittelalters, (whose web access is not public yet), or ReMetCa itself by the addition of a XMLType field to its relational database.

Primer verso del poema :	Andemos, señora, andemos
Número de versos :	18
Número de estrofas :	5
FormulaMétrica :	8,8,8,8,8,8,8,8,8,8,8,8,8,8,8,8,8,8
FormulaRimática :	zzabbaaz*z*zcddcz*z*z
Código TEI :	<pre> &lt;lg type="estribillo" subtype="pareado" asonancia="consonante" met="8,8" rhyme="zz"&gt;   &lt;l n="63" rhyme="emos"&gt;Andemos, señora, andemos.&lt;/l&gt;   &lt;l n="54" rhyme="emos"&gt;o si manda, descansemos.&lt;/l&gt; &lt;/lg&gt;  &lt;lg type="estrofa" subtype="redondilla" asonancia="consonante" met="8,8,8,8" rhyme="abba"&gt;   &lt;l n="55" rhyme="ón"&gt;No me carga mi zurrón.&lt;/l&gt;   &lt;l n="56" rhyme="ado"&gt;no he menester ni cayado.&lt;/l&gt;   &lt;l n="57" rhyme="ado"&gt;que de Dios soy consolado.&lt;/l&gt;   &lt;l n="58" rhyme="ón"&gt;libre de toda pasión.&lt;/l&gt; &lt;/lg&gt;  &lt;lg type="estrofa" real="consonante" met="8,8,8,8" rhyme="az*z*z"&gt;   &lt;l n="59" rhyme="ón"&gt;pues que nuestra redención&lt;/l&gt;   &lt;l n="60" rhyme="emos"&gt;con nosotros la traemos.&lt;/l&gt;   &lt;lg type="estribillo" subtype="pareado" asonancia="consonante" met="8,8" rhyme="z*z*z"&gt;   &lt;l n="61" rhyme="emos"&gt;Andemos, señora, andemos.&lt;/l&gt;   &lt;l n="62" rhyme="emos"&gt;o si manda, descansemos.&lt;/l&gt; &lt;/lg&gt;  &lt;lg type="estrofa" subtype="redondilla" asonancia="consonante" met="8,8,8,8" rhyme="cddc"&gt;   &lt;l n="63" rhyme="ero"&gt;El descanso verdadero&lt;/l&gt;   &lt;l n="64" rhyme="oso"&gt;es nuestro hijo precioso.&lt;/l&gt;   &lt;l n="65" rhyme="oso"&gt;Este es Dios poderoso.&lt;/l&gt;   &lt;l n="66" rhyme="ero"&gt;Este es el manso cordero.&lt;/l&gt; &lt;/lg&gt; </pre>

Fig. 1: ReMetCa database screen with XML-type field

The software used to build the collaborative ontology is Webprotege (Tudorache et al. 2011)<sup>[7]</sup>, initially combined with Poolparty to create and organize vocabularies. It has been installed at ReMetCa server and opened via web in order to let participation of researchers with similar projects in the field of metrical repertoires. This system presents a light and intuitive interface but solid enough to develop a complex ontology with OWL. An important advantage is that it offers multilingual edition, which is very important for the development of such an international proposal. Once the model had been set, a metadata system has been designed to link the conceptual and logical levels based on a global abstract classification (schema), in which the different particular embodiments of each poetical tradition will be progressively included. This proposal shows both the consistency of this general language purpose and the benefits that can be obtained from the application of this model to the different local projects using a collaborative and open work system, which is essential for this new paradigm.

There are a few studies which deal obliquely with some of the above mentioned aspects (Bootz & Szoniecky 2008<sup>[8]</sup> and Zöllner-Weber 2009<sup>[9]</sup>), but there is not yet a conceptual model of ontology referred to metrics and poetry. The closest related works to this topic are probably the conceptual model of CIDOC ([www.cidoc-crm.org](http://www.cidoc-crm.org)), the vocabularies of the Getty Museum, as they are designed to express relations and artistic manifestations in the field of humanities (<http://www.getty.edu/research/tools/vocabularies/>), the controlled vocabularies of English Broadside Ballad Project <http://ebba.english.ucsb.edu/> and the linked data relations offered by the Library of Congress (<http://id.loc.gov/>), which do not offer a deep information on metrics vocabulary.

To sum up, this project of a poetic and metrical ontology intends be much more than a repository of datasets, thesauri or controlled vocabularies. It aims to create a semantic standardized structure to describe, analyze and develop logical operations through the different poetic digital repertoires and their related resources. Its final objective is to interconnect, reuse and locate the data disseminated through poetic databases in order to get interoperability among projects, to perform complex searches and to make the different resources “talk” to each other following a unique but adaptive model.

## References

1. **González-Blanco, E. & L. Seláf** (2013), *Megarep: A comprehensive research tool in Medieval and Renaissance poetic and metrical repertoires*, Humanitats a la xarxa: món medieval / Humanities on the web: the medieval world, edited by L. Soriano, M. Coderch, H. Rovira, G. Sabaté & X. Espluga., Oxford, Bern, Berlin, Bruxelles, Frankfurt am Main, New York, Wien, Peter Lang (forthcoming).
2. **Elmarsí, R. & S. B. Navathe** (2011), *Fundamentos de Sistemas de Bases de Datos*, Madrid, Pearson, Addison Wesley.
3. **González-Blanco, E. & J. L. Rodríguez** (2013), *ReMetCa: a TEI based digital repertory on Medieval Spanish poetry*, at The Linked TEI: Text Encoding in the Web, Book of Abstracts - electronic edition. Abstracts of the TEI Conference and Members Meeting 2013: October 2-5, Rome edited by Fabio Ciotti & Arianna Ciula, DIGILAB Sapienza University & TEI Consortium, Rome, 178-185. [digilab2.let.uniroma1.it/teiconf2013/abstracts/](http://digilab2.let.uniroma1.it/teiconf2013/abstracts/). Accessed 30-10-2013.
4. **Eide, Ø. & C.-E. Ore** (2007), *From TEI to a CIDOC-CRM Conforming Model: Towards a Better Integration between Text Collections and Other Sources of Cultural Historical Documentation*, paper presented at the DH conference 2007. Abstract available at: [www.edd.uio.no/artiklar/tekstkoding/poster\\_156\\_eide.html](http://www.edd.uio.no/artiklar/tekstkoding/poster_156_eide.html)
5. **Ciula, A., P. Spence & J. M. Vieira** (2008), *Expressing complex associations in medieval historical documents: the Henry III Fine Rolls Project*, Literary and Linguist Computing, 23 (3): 311-325.
6. **Burnard, L. & S. Bauman**, eds., *TEI P5: Guidelines for Electronic Text Encoding and Interchange. Ver. 2.5.0*. Accessed 30-10-2013. [www.tei-c.org/release/doc/tei-p5-doc/en/html/](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/)
7. **Tudorache, T., C. I. Nyulas, N.F. Noy & M.A. Musen** (2011), *WebProtégé: A Collaborative Ontology Editor and Knowledge Acquisition Tool for the Web*, Semantic Web Journal, IOS Press. [www.semantic-web-journal.net/content/webprot%C3%A9g%C3%A9-distributed-ontology-editor-and-knowledge-acquisition-tool-web](http://www.semantic-web-journal.net/content/webprot%C3%A9g%C3%A9-distributed-ontology-editor-and-knowledge-acquisition-tool-web). Accessed: 30/10/2013.
8. **Bootz, P. & S. Szoniecky** (2008), *Towards an ontology of the field of digital poetry*, paper presented at Electronic Literature in Europe. Full text available at [elmcip.net/node/415](http://elmcip.net/node/415)
9. **Zöllner-Weber, A.**, “Ontologies and Logic Reasoning as Tools in Humanities?”, DHQ 2009, 3: 4. [www.digitalhumanities.org/dhq/vol/3/4/000068/000068.html](http://www.digitalhumanities.org/dhq/vol/3/4/000068/000068.html) Accessed: 30/10/2013.
9. **Zöllner-Weber, A.** (2009), *Ontologies and Logic Reasoning as Tools in Humanities?*, DHQ, 3: 4. [www.digitalhumanities.org/dhq/vol/3/4/000068/000068.html](http://www.digitalhumanities.org/dhq/vol/3/4/000068/000068.html) Accessed: 30/10/2013.