



Journal of the Text Encoding Initiative

Issue 8 | December 2014 - December 2015
Selected Papers from the 2013 TEI Conference

ReMetCa: A Proposal for Integrating RDBMS and TEI-Verse

Elena González-Blanco and José Luis Rodríguez



Electronic version

URL: <http://journals.openedition.org/jtei/1274>

DOI: 10.4000/jtei.1274

ISSN: 2162-5603

Publisher

TEI Consortium

Electronic reference

Elena González-Blanco and José Luis Rodríguez, « ReMetCa: A Proposal for Integrating RDBMS and TEI-Verse », *Journal of the Text Encoding Initiative* [Online], Issue 8 | December 2014 - December 2015, Online since 23 September 2015, connection on 19 April 2019. URL : <http://journals.openedition.org/jtei/1274> ; DOI : 10.4000/jtei.1274

For this publication a Creative Commons Attribution 4.0 International license has been granted by the author(s) who retain full copyright.

ReMetCa: A Proposal for Integrating RDBMS and TEI-Verse

Elena González-Blanco and José Luis Rodríguez

1. The ReMetCa Project: Definition and Scope

- ¹ [ReMetCa](#)¹ is a computer-based metrical repertoire of Medieval Castilian poetry. It gathers all poetic witnesses, from the very beginnings of Spanish lyrics at the end of the twelfth century through the rich and varied poetic manifestations of the fifteenth and sixteenth century, the “Cancioneros.” When complete, it will include over 10,000 texts and offer a systematic metrical analysis of each poem. ReMetCa is the first digital tool to analyze Medieval Spanish poetry. It enables users to carry out complex searches in its corpus, following the models of other digital resources in the Romance lyrical traditions, such as the Galician-Portuguese, Catalan, Italian, or Provençal. ReMetCa is not merely a metrical repertoire; it combines metrical schemes together with text analysis as well as data sheets with the main philological aspects that characterize the poems. One of its most important values is that it is a born-digital project designed to be interoperable with other existing poetry databases and digital repertoires, as conceptualized within the [Megarep](#)² project, in which Spanish researchers are already working in collaboration with the Hungarian team led by Levente

Seláf (González-Blanco and Seláf 2014). The Spanish repertoire is conceived as an essential tool to complete the digital European poetic puzzle, enabling users to conduct powerful searches in many fields at the same time.

2. The Role of ReMetCa in the European History of Poetic Repertoires

- 2 ReMetCa belongs to the latest generation in a long tradition of metrical repertoires, which can be divided into three significant periods based on the different technologies used for their construction. The first stage, in which repertoires were published as printed books, started at the end of the nineteenth century.³ The second period started after the Second World War, when repertoires became computer-assisted.⁴ Technological advance has made it possible to create a third generation of repertoires, available online and searchable through the web, in which research time has been considerably reduced by richer search capabilities and the accessibility of online interfaces, especially compared with the complexity of printed indexes and lists of metrical schemes. The first online digital poetic repertoire was the *RPHA: Répertoire de la Poésie hongroise ancienne jusqu'à 1600* (Horváth 1991–); Galician researchers created *MedDB: Base de datos da Lírica profana galego-portuguesa* (Brea et al. 1994–); Italian researchers digitalized *BEdT: Bibliografia Elettronica dei Trovatori* (Asperti et al.); the *Nouveau Naetebus* (Seláf), the *Oxford Cantigas de Santa María Database* (Parkinson), the *Analecta Hymnica Digitalia* (Rauner), and the *Dutch Song Database* (Grijp) all appeared later.⁵ Most of these repertoires are built on relational databases; some of them are open source, but the majority are built on proprietary software. Some of them were previously published as books and later as CDs, but now almost every project has developed an online version. The technical description of these resources is not the topic of this paper, but it will be studied in depth in a later stage of the project, as further possibilities for interoperability are analyzed.

3. The State of the art in Spain

- 3 Compared with the European tradition described above, the Spanish panorama looks weak, as we do not have a published poetic repertoire which gathers the metrical patterns of Medieval Castilian poetry (except for the book by Gómez-Bravo [1999], restricted to “Cancionero” poetry,

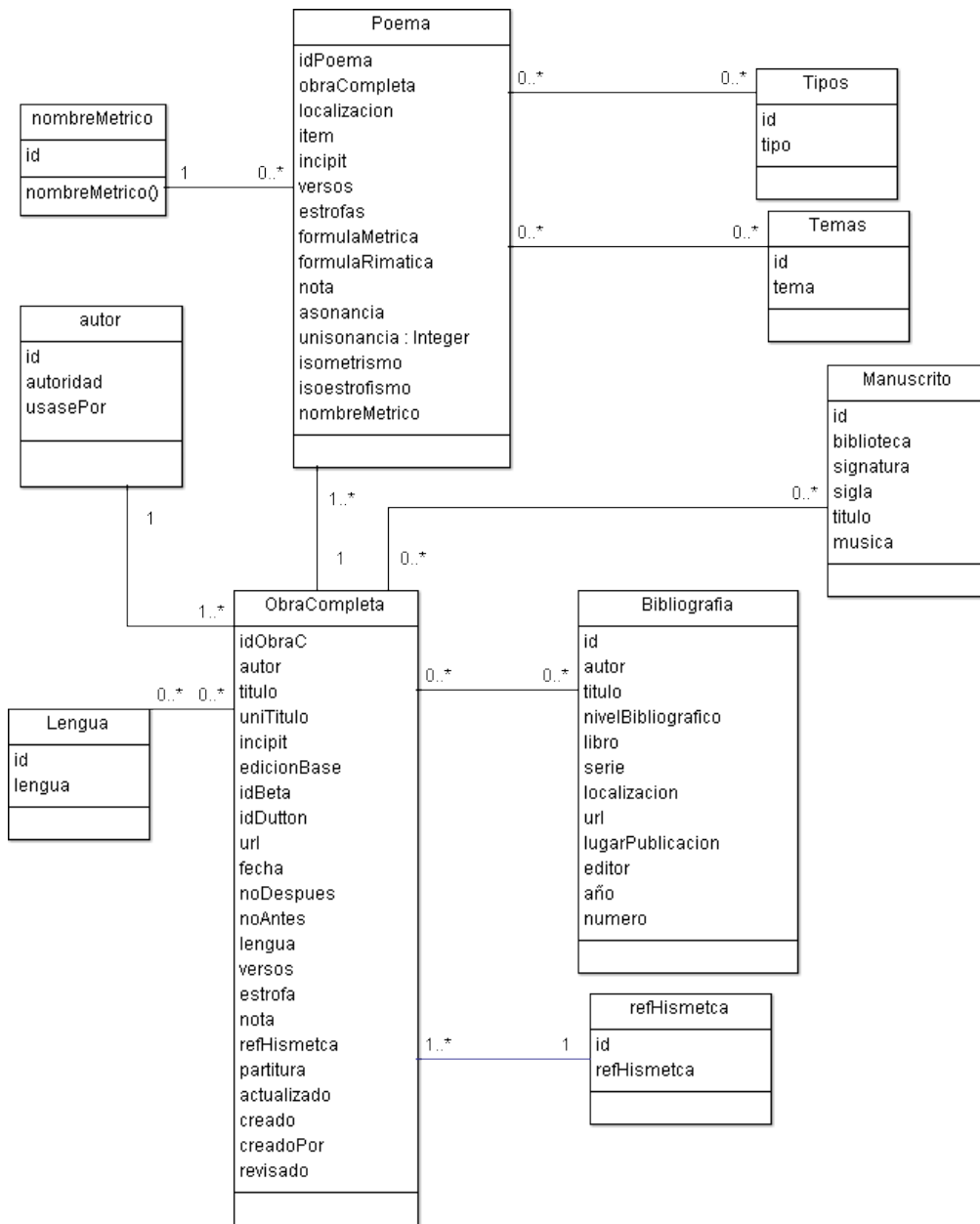
which only covers the fifteenth century, but not the poems written during the thirteenth and fourteenth centuries), and, until now, there was no digital resource available. However, researchers are nowadays more conscious of the importance of metrical studies for the analysis and understanding of Medieval Spanish poetry, as has been recently shown.⁶ On the other hand, metrical studies have flourished thanks to the creation of specialized journals, such as *Rhythmica*, *Revista española de métrica comparada*,⁷ available online, and *Ars Metrica*,⁸ born as an electronic journal whose scientific committee is composed of researchers from universities and centers in different countries.

- 4 At this point, ReMetCa's aim is to fill the Spanish gap with information in order to build a metrical repertoire within this complex puzzle of digital European poetry, but the objective of the project in its future stages is to design a system—to be based on the TEI model, but integrating linked data technologies—which allows interoperability, information exchange, and combined searches across the different repertoires and databases.

4. The Conceptual Schema of ReMetCa

- 5 The description of ReMetCa's structure starts with the definition of a conceptual model based on the domain of medieval metrics. Its global structure is synthesized into a relational database, which will be exportable in the future to full TEI in order to facilitate data interchange and interoperability with other projects and systems. For the moment, TEI is restricted to tagging metrical and poetic phenomena by using the TEI Verse module in only one of the database fields. Although there are many possible ways to structure such a database, our proposed model takes into account the elements we need to describe and analyze our field of study. For the graphical representation we use a light version of UML (Unified Modeling Language). Entity names and their attributes are shown as boxes, and these are connected by lines which represent the existing relationships among them, together with numbers to show cardinality.

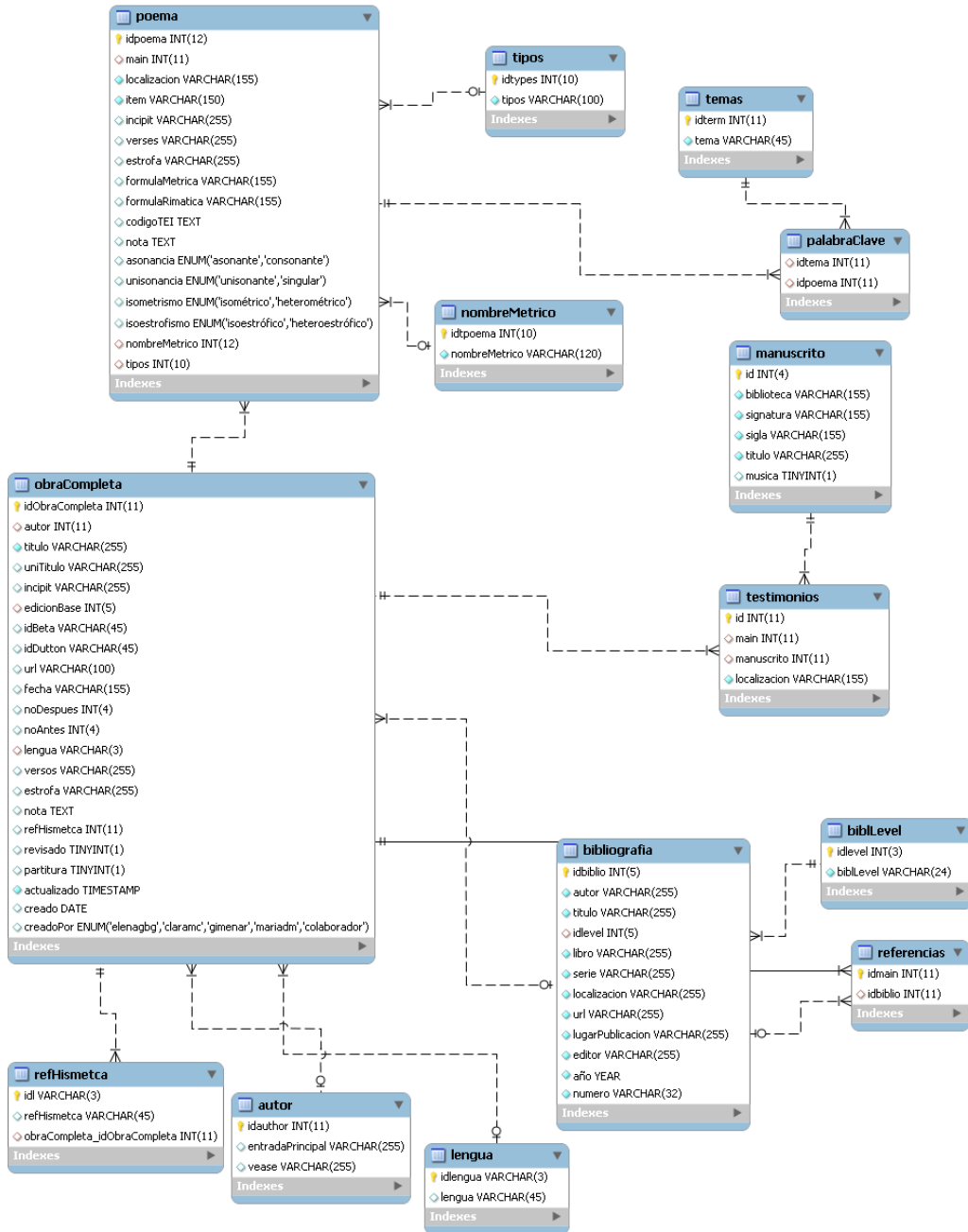
Figure 1. Conceptual model.



5. The Logical Schema: Modeling the Database E-R Model

- The conceptual model of entities, attributes, and relationships is transformed into an Entity-Relationship model after applying the transformation rules that convert it into a database.

Figure 2. Logical Model.



- 7 The relationships established among entities can be “one-to-many” or “many-to-many.” In the first category you can find the relationship between *obraCompleta* (literary work) and *refHismetca* (Classification System). As one *refHismetca* may be shown in many different literary works, the solution given in the ER model is to define a table for *refHismetca* with at least two attributes: one to identify the *refHismetca*, and the other to name it. For the entity *obraCompleta* (literary work), our database has a field with a foreign key which points to the *id* field of the *genre* table [*refHismetca*]. The second type of relationship, many-to-many, is larger. It is necessary to create a third table to contain each of them. This happens, for example, in order to express the relationship between topics and poems, whose third entity consists of as many tuples as there are topics assigned to each poem. Every tuple has at least two attributes to identify *poema* (poem) and *tema* (topic). A similar situation happens with the entity *bibliografía* (bibliography), whose third entity is called *referencias*.
- 8 The structure described above covers the majority of entities in our relational database model. Some problems arise when entering entities with a complex textual description, such as the terms used to describe parts of the literary work. *Poema* (poem), *estrofa* (stanza), and *verso* (line) are the components that define the hierarchical structure of each literary work analyzed. Applying the previously described E-R model would drive us into a complex model of relationships among those components which are very difficult to represent in a database. These relationships are created through composition, which means that the entity *poema* is composed of one or more entities *estrofa*, and the entity *estrofa* consists of one or more entities *verso*. The relations of multiplicity among the components vary from one work to another, depending on the number of stanzas and lines in each entity. For example, a sonnet shows a multiplicity of 1 for each stanza contained by it (i.e., each stanza is part of only one sonnet), and the stanza has a multiplicity of 4 in relation to the sonnet (the sonnet contains four stanzas). The problem of representing this composition in an E-R model is that the representation is data-centered, and it does not work for poem, line, and stanza, as these components need to be analyzed as textual items. It is necessary to insert metadata into these textual items to show their compositional relation and to add other relevant information to them. The E-R model is inappropriate for this purpose due to its center-based structure, with the entities of poem, line, and stanza in the middle of its referential domain of study.

- 9 The difficulties described above remain in each stage of the project and reappear in the programming phase and when designing the different web components, both in the administrative part of the project (for aspects related to the creation, modification, and deletion of registers) and in its querying module. Also, the structure of each of the three components mentioned (poem, stanza, and line) differs significantly depending on the literary work, as there are different kinds of poems, stanzas, and lines. For example, some of the poems are monostrophic, whereas others are composed of several different stanzas. Lines may be rhymed or unrhymed, may be inserted into a strophic structure or not, may form part of a series of compositions, or may constitute a literary composition themselves. From a technical point of view, all of them could be considered “semi-structured data.”

6. The Addition of XML Tagging: The Verse Module of TEI P5

- 10 This hierarchical and semi-structured pattern is quite ungainly for an E-R database model, which is not flexible enough to show hierarchical structures. For this reason, using an XML-based markup language, such as TEI, is the perfect solution to show all the complex relations and properties. As the ReMetCa model is built on a database which works very well in terms of structure to describe the main fields of our projects, TEI is just used for tagging entities and relationships that cannot be represented in the database structure because of the complexity and diversity of the contents of the texts included, in terms of metrical definitions and properties. We have selected only the tags of the Verse module of the TEI P5 Guidelines, as this module works perfectly to show the hierarchy and properties needed, which may vary depending on the type of poem. Here is an example:

Example 1. *Auto de la huida a Egipto*, escena III, vv. 36–44.

```
<lg type="estrofa" subtype="redondilla" asonancia="consonante" met="8,8,8,8"
rhyme="cddc">
  <l n="36">Guía al hijo y a la m<rhyme label="c">adre</rhyme>,</l>
  <l n="37">guía al viejo pecad<rhyme label="d">or</rhyme>,</l>
  <l n="38">que se parte sin tem<rhyme label="d">or</rhyme></l>
  <l n="39">a donde manda Dios p<rhyme label="c">adre</rhyme>;</l>
</lg>

<lg type="estrofa" asonancia="consonante" met="8,8" rhyme="+b+a">
  <note>estos dos versos son de vuelta</note>
  <l n="40">y pues al niño bend<rhyme label="b">ito</rhyme></l>
  <l n="41" rhyme="aste">y a nosotros tú sac<rhyme label="a">aste</rhyme>,</l>

<lg type="estribillo" asonancia="consonante" met="8,8,8" rhyme="*a*b*b">
  <l n="42">Ángel, tú que me mand<rhyme label="a">aste</rhyme></l>
  <l n="43">de Judea ir a Eg<rhyme label="b">ipto</rhyme>,</l>
  <l n="44">guíanos con el chiqu<rhyme label="b">ito</rhyme>.</l>
</lg>
<!-- lg are nested because "estribillo" belongs to the stanza but it needs to
be marked as an indepent structure -->
</lg>
```

- 11 We have designed a TEI customization based on the TEI Verse module and represented as an XML Schema (.xsd) which includes the following tags and attributes:

- The stanza, marked with the <lg> tag, is our highest tagging level and bears the following attributes:
 - @type describes the function of the stanza in the poem and may have one of three values, "estrofa", "estribillo", or "cabeza", depending on its function: a normal stanza, a chorus, or the head position of the poem (which is usually repeated with variations).

- @subtype contains the name of the stanza form (e.g., "soneto", "cuadernavia", "tirada", "romance", "zéjel"), provided it has a standardized name in the poetic tradition. The classification of the terms is gathered in a controlled vocabulary registered in the guidelines established by ReMetCa. If there is no known conventional name (which happens in many cases), the attribute may not appear.
- @n contains the number of the stanza as per the print edition on which our digital editions are based. Since the transcriptions of the texts we use are not complete, it is necessary to number the lines in order to make references to other editions.
- @met contains the metrical structure, based on the number of syllables in each line and on the number of lines, separated by commas. A stanza of four lines of octosyllables would be encoded as met="8,8,8,8".
- @rhyme contains the rhyming structure of each stanza, represented by letters, as for example in rhyme="abba", to show that the first and fourth lines have the same rhymes and so do the second and third.
- @asonancia is a new attribute which does not exist in the TEI Verse module and has been added to our schema. Although our philosophy is to respect and follow the TEI Guidelines (TEI Consortium 2013), we have felt obliged to introduce three new attributes. This one indicates the two possible values of the rhyming typology: "asonante" (which means that only vowel sounds are repeated at the end of two rhyming lines), and "consonante" (which means that every sound after the stressed syllable is repeated). The second added attribute is @unisonancia, which takes the values of either "unisonante" or "singular" and shows whether the same rhyme scheme is repeated in different stanzas (unisonante: abba, abba) or not (singular: abba, cddc). The third new attribute is @isometrismo, which takes the values of either "isométrico" or "heterométrico", which indicates whether all the descendant stanzas have the same number of syllables ("isométrico") or not ("heterométrico").
- The element "line," <l>, a child of <lg>, is a mixed element which contains the child element <rhyme> to mark the rhyming string of the line which is pronounced after the stressed vowel, and one attribute, @n, which marks the line number, following the same principles as those indicated by the same attribute described at <lg>.

7. Possibilities for Extending our TEI Schema

- 12 As the ReMetCa project was initially built as a relational model, the only TEI component that has been added to its data model is the Verse module. However, in the future there will be more TEI markup additions to reflect the need for gathering more information on the different manuscripts witnessing each poem. Our team is analyzing the possibility of replacing the manuscript table with the element `<msDesc>`. The reasons for this decision derive from the complexity of describing many different witnesses, each of them with complex relations to many collections, libraries, owners, and copyists. Following the E-R model would force us to multiply the number of tables in order to be able to reflect the complexity; therefore, using a model based on XML would ease our work by offering both exhaustiveness and flexibility, but this advantage is counterbalanced by the scale of changes that would be required in order to introduce XQuery functionality to the system, so each decision to change in this way has to be taken gradually and carefully.

8. The Physical Model

- 13 Once the logical model is established, the next task was to implement the selected schema into a database management system, what is known as the “physical model.” (Silberschatz 2002: 40) As far as the selection of the system is concerned, native XML databases, such as eXist or BaseX, were discarded since we wanted to preserve the original relational structure of our logical system, and also because we considered it interesting to use an RDBMS model, seeing that most of the online repertoires are built on similar systems. The problem lay in choosing a combined system to integrate both types of information: E-R relational structures and text marked up with XML. We compared two of the most popular E-R database systems: MySQL and Oracle9i XML. Both of them offered the possibility of adding an XMLType data column, thus offering the possibility of introducing hierarchical structures inside the relational model via the provision of XML nesting structures. The combination of both systems offers great advantages, such as the possibility of making combined queries with SQL and XPath languages.⁹
- 14 This first trial model of ReMetCa, combining the MySQL RDBMS and TEI XML fields, is already working online and can be visited at <http://ww.remetca.uned.es>. To access the database, it is necessary to log in to “Área de trabajo” and then into “base de datos,” using a username and

password that can be requested from the same website. Eventually the search screen will be online, free and open with no password, but not yet, as we are still working on populating the database and designing the search engine.

- 15 Oracle databases contain fields defined as XMLType, thanks to the XDB component. This component is usually installed by default, but its status can be checked by using the following SQL statement:

```
SQL> select comp_name, status from dba_registry where comp_name='Oracle XML
Database' ;
```

```
COMP_NAME
```

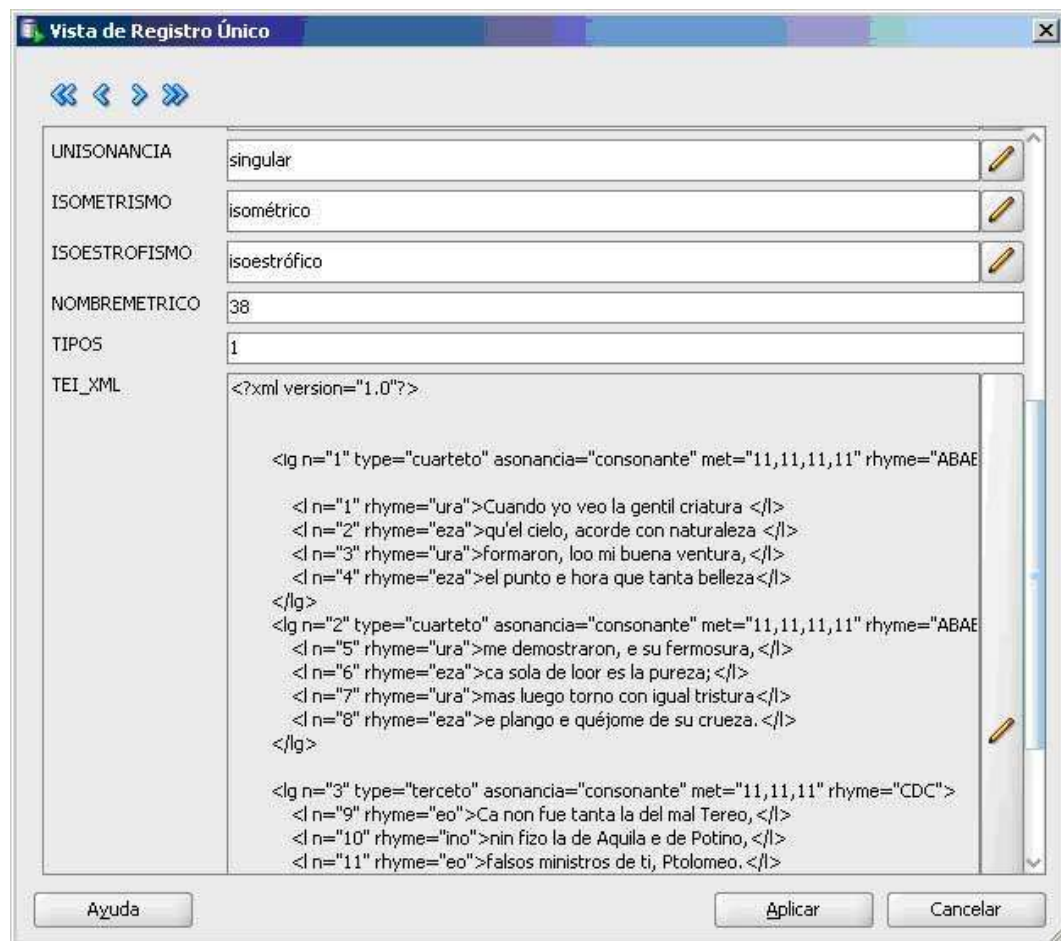
```
-----
STATUS
-----
```

```
Oracle XML Database
```

```
VALID
```

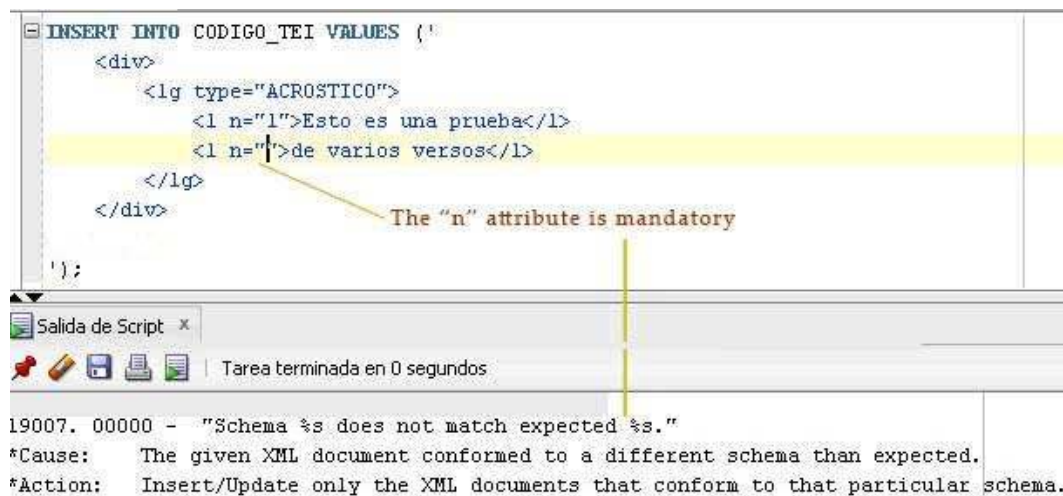
- 16 In the case of ReMetCa, this XMLType field needs to appear only once, in the table Poema, which contains <lg> elements and their content, as described above. Oracle SQL Developer also allows updates to contents with its unique record view.

Figure 3. Oracle SQL Developer.



- 17 Oracle stores XSD schemas into its system, so the XML introduced is validated. To test this property, we tried introducing an XSD schema for this <lg> element, adding attributes and constraints, and as figure 4 shows, it was perfectly validated. The system also accepts XSLT stylesheets. Oracle registers stylesheets and can apply them to any XMLType field.

Figure 4. Schema validation error.



- 18 Finally, it is important to create an index for XML elements to test the performance of the database.

Figure 5. Index of verses.

```

create index idx_verse on poema(tei_xml) INDEXTYPE IS XDB.xmlindex PARAMETERS
('PATHS (INCLUDE (/div/lg/l))');

```

9. Conclusion

- 19 As described above, the ReMetCa project is under constant development. As our textual corpus grows, the initial schema has to be adjusted to its needs. We are working to expand the technical model and to make it capable of describing the complex panorama of Medieval Spanish poetry. During the next stages of the project, more TEI modules and elements will be added in order to complete some of the sections that are not yet satisfactorily described by the E-R model in the database, and also to make it possible to export all the database contents to full TEI. A further stage will consist of working on a linked-data model to allow interoperability among the different European repertoires based on a common ontology.

- 20 Although we have adopted Oracle because of its power to work with XML technologies and TEI, there are other open-source systems like MySQL or PostgreSQL which offer enough functions to implement a mixed system like the one proposed in this paper. Actually, it is predictable that both systems will include similar tools in the future to work with XMLType fields, as XML is becoming more and more widely used.

BIBLIOGRAPHY

Bibliography on Metrical Studies

- Beltrán, Vicenç. 2007. "Bibliografía sobre Poesía medieval y Cancioneros." In *Biblioteca Virtual Joan Lluís Vives*. Accessed 5 February, 2014. <http://www.lluivives.com/FichaObra.html?Ref=25467&portal=1>.
- Domínguez Caparrós, José. 1985. *Diccionario de métrica española*. Madrid: Paraninfo.
- . 1988. *Contribución a la bibliografía de los últimos treinta años sobre la métrica española*. Madrid: C.S.I.C.
- Duffell, Martin J. 2007. *Syllable and Accent: Studies on Medieval Hispanic Metrics*. Papers of the Medieval Hispanic Research Seminar 56. London: Department of Hispanic Studies, Queen Mary and Westfield College, University of London.
- García Calvo, Agustín. 2006. *Tratado de Rítmica y Prosodia y de Métrica y Versificación*. Zamora: Lucina.
- Gómez Redondo, Fernando. 2001. *Artes Poéticas Medievales*. Madrid: Laberinto.
- González-Blanco García, Elena. 2010. *La cuaderna vía española en su marco panrománico*. Madrid: FUE.
- González-Blanco García, Elena and Levente Seláf. 2014. "Megarep: A Comprehensive Research Tool in Medieval and Renaissance Poetic and Metrical Repertories." In *Humanitats a la xarxa: món medieval / Humanities on the Web: The Medieval World*, edited by Lourdes Soriano, Helena Rovira, Marion Coderch, Glòria Sabaté, and Xavier Espluga, 321–32. Oxford: Peter Lang.
- Micó, José María. 2009. *Bibliografía para una historia de las formas poéticas en España*. Biblioteca Virtual Miguel de Cervantes: Alicante. Accessed February 5, 2014. <http://www.cervantesvirtual.com/nd/ark:/59851/bmc1g127>.
- Silberschatz, Abraham (et al.). 2002. *Fundamentos de Bases de datos*, Madrid: Mc Graw Hill.
- TEI Consortium. 2013. *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Version 2.5.0. Last updated July 26. N.p.: TEI Consortium. <http://www.tei-c.org/Vault/P5/2.5.0/doc/tei-p5-doc/en/html/>.

Metrical Repertoires Published in Print

- Antonelli, Roberto. 1984. *Repertorio metrico della scuola poetica siciliana*. Palermo: Centro di Studi Filologici e Linguistici Siciliani.
- Betti, Maria Pia. 2005. *Repertorio Metrico delle Cantigas de Santa Maria di Alfonso X di Castiglia*. Ospedaletto (Pisa): Pacini.
- Brunner, Horst, Burghart Wachinger, and Eva Klesatschke, eds. 1986–2007. *Repertorium der Sangsprüche und Meisterlieder des 12. bis 18. Jahrhunderts*. Tübingen: Niemeyer.
- Frank, István. 1953–57. *Répertoire métrique de la poésie des troubadours*. Paris: H. Champion.
- Gómez-Bravo, Ana María. 1999. *Repertorio métrico de la poesía cancioneril del siglo XV*. Alcalá de Henares: Universidad.
- Gorni, Guglielmo. 2008. *Repertorio metrico della canzone italiana dalle origini al Cinquecento (REMCI)*. Florence: Franco Cesati.
- Mölk, Ulrich, and Friedrich Wolfzettel. 1972. *Répertoire métrique de la poésie lyrique française des origines à 1350*. Munich: W. Fink Verlag.
- Naetebus, Gotthold. 1891. *Die Nicht-Lyrischen Strophenformen des Altfranzösischen. Ein Verzeichnis Zusammengestellt und Erläutert*. Leipzig: S. Hirzel.
- Pagnotta, Linda. 1995. *Repertorio metrico della ballata italiana: secoli XIII e XIV*. Milan: Ricciardi.
- Parramon i Blasco, Jordi. 1992. *Repertori mètric de la poesia catalana medieval*. Barcelona: Curial, Abadia de Montserrat.
- Pillet, Alfred, and Henry Carstens. 1933. *Bibliographie der Troubadours*. Halle: M. Niemeyer.
- Raynaud, Gaston. 1884. *Bibliographie des chansonniers français des xiii e [treizième] et xiv e [quatorzième] siècles: comprenant la description de tous les manuscrits, la table des chansons classées par ordre alphabétique de rimes et la liste des trouvères*. 2 vols. Paris: Vieweg.
- Solimena, Adriana. 1980. *Repertorio metrico dello Stil novo*. Rome: Presso la Società.
- . 2000. *Repertorio metrico dei poeti siculo-toscani*. Palermo: Centro di studi filologici e linguistici siciliani.
- Tavani, Giuseppe. 1967. *Repertorio metrico della lirica galego-portoghese*. Rome: Edizioni dell'Ateneo.
- Touber, Anton H. 1975. *Deutsche Strophenformen des Mittelalters*. Stuttgart: Metzler.
- Zenari, Massimo. 1999. *Repertorio metrico dei "Rerum vulgarium fragmenta" di Francesco Petrarca*. Padova: Antenore.

List of Digital Repertoires and Databases

- Asperti, Stefano, Fabio Zinelli, et al. *BEtI, Bibliografia Elettronica dei Trovatori*. Sapienza Università di Roma, Dipartimento Studi greco-latini, italiani, scenico-musicali. Accessed February 5, 2014. <http://www.bedt.it/>.
- Brea, Mercedes, et al. 1994-. *MedDB: Base de datos da Lírica Profana Galego-Portuguesa*. Centro Ramón Piñeiro para a Investigación en Humanidades. Accessed February 5, 2014. <http://www.cirp.es/bdo/med/mddb.html>.
- Dreves, Guido M., Clemens Blume, H.M. Bannister, González-Blanco, Elena, et al. 2013-. *ReMetCa: Repertorio métrico digital de la poesía medieval castellano [A Digital Repertory of Medieval Spanish Metrics]*. Last updated July 22, 2015. <http://www.remetca.uned.es>.
- Grijp, Louis, et al. *Dutch Song Database*. Meertens Institute, Royal Netherlands Academy of Arts and Sciences (KNAW). Accessed February 5, 2014. <http://www.liederenbank.nl/index.php?lan=en>.
- Horváth, Iván, et al. 1991-. *Répertoire de la poésie hongroise ancienne*. Eötvös Loránd University (ELTE). Accessed February 5, 2014. <http://rpha.elte.hu/>.
- Parkinson, Stephen, et al. *The Oxford Cantigas de Santa Maria Database*. Oxford University. Accessed February 5, 2014. <http://csm.mml.ox.ac.uk/>.
- Rauner, Erwin. *Analecta Hymnica Medii Aevi Digitalia*. Erwin Rauner Verlag. Accessed February 5, 2014. http://webserver.erwin-rauner.de/crophius/Analecta_conspectus.htm.
- Seláf, Levente. *Le Nouveau Naetebus: Poèmes strophiques non-lyriques en français des origines jusqu'à 1400*. Accessed February 5, 2014. <http://nouveaunaetebus.elte.hu/>.

NOTES

- 1 <http://www.remetca.uned.es/>.
- 2 http://www.remetca.uned.es/index.php?option=com_content&view=article&id=76:megarep-a-comprehensive-research-tool-in-medieval-and-renaissance-poetic-and-metrical-repertoires&catid=80&Itemid=514&lang=es.
- 3 With the works of Gaston Raynaud (1884), Gotthold Naetebus (1891), and Pillet and Carstens (1933), among others.
- 4 With the classic work of Frank on Provençal troubadours' poetry (Frank 1953–57); it has continued with the editions of printed metrical repertoires, in Old French lyrics by Mölk and Wolfzettel (1972), in Italian by Solimena (1980), Antonelli (1984), Pagnotta (1995), Zenari (1999),

Solimena (2000), and Gorni (2008); in the Hispanic Peninsula by Tavani (1967), Parramon i Blasco (1992), Gómez-Bravo (1999), and Betti (2005); in German by Touber (1975) and Brunner, Wachinger, and Klesatschke (1986–2007).

5 For the URLs and other details, see “Repertoires and Digital Databases” in the bibliography below.

6 Some examples are the publications of Domínguez Caparrós (1985 and 1988), Micó (2009), Gómez Redondo (2001), García Calvo (2006), Duffell (2007), Beltrán (2007) or González-Blanco (2010).

7 <http://e-spacio.uned.es/revistasuned/index.php/rhythmica>.

8 <http://www.arsmetrica.eu/>.

9 The disadvantage of MySQL compared to Oracle is that MySQL does not offer validation for XML fields, as it considers them just as text fields, and there are no added XSD schemas or any kind of transformational stylesheets, like XSLT. TEI code has to be composed and validated outside the database by using XML editors (we use Oxygen), since the MySQL RDBMS exploitation of XML technologies is limited to XPath and can only be recovered through the `extractValue` function of MySQL, which shows important limitations. Oracle Database 9i and 10g, however, are much more powerful in that sense, as they combine their E-R nature with XML technologies, resulting in a much more advanced tool to work with both systems, hierarchical and relational, at the same time.

ABSTRACT

This paper describes the technical structure of the project ReMetCa (Repertorio Digital de Métrica Medieval Castellana), the first online repertoire of Medieval Spanish metrics and poetry. ReMetCa is based on the combination of traditional metrical and poetic studies (rhythm and rhyme patterns) with digital humanities technology, TEI-XML integrated into a Relational Database Management System (RDBMS) through an XMLType field, thus opening up the possibility of launching simultaneous searches and queries by using a searchable, user-friendly interface. The use of the TEI Verse module to tag metrical and poetic structures is especially important for two reasons: first, because it lets us tag different kinds of poems with a variable metadata structure, making it possible to express a very high level of detail for poetry description that cannot be registered in a conventional database; and second, as it enables extensibility and adds more information to enrich the conceptual model, which is under constant development.

INDEX

Keywords: poetry, medieval metrics, repertoire, RDBMS, MySQL, verse, stanza

AUTHORS

ELENA GONZÁLEZ-BLANCO

Elena González-Blanco teaches in the Spanish Literature Department at Universidad Nacional de Educación a Distancia (Open University) of Spain in Madrid. Her main research and teaching areas are comparative medieval literature, metrics and poetry, and digital humanities. She is the Director of LINHD, the Digital Humanities Center at UNED, leads the project ReMetCa, and is also the coordinator of the linked-data Unedata project at UNED. She is a member of the Executive Committee of EADH, member of the first executive committee of GO::DH, member of the Clarin Scientific Advisory Board, member of Centernet Executive Committee, and the secretary of the Spanish association for digital humanities HDH.

JOSÉ LUIS RODRÍGUEZ

José Luis Rodríguez is a systems librarian at the Royal Library (Real Biblioteca) in Madrid. He participated in projects devoted to building the manuscript catalogs of the Salamanca University Library and the Royal Library. He has designed and developed several web applications to manage historical bookbinding collections, old inventories of books, exlibris, digital libraries, and other historical bibliographic materials. He is an expert in Integrated Library Systems and he is currently working with the Koha Open Source ILS. His publications include several articles on the history of Galician language, book history, and digital humanities. He is a member of the ReMetCa Project.