

MÉTRICA, CLASIFICACIÓN Y TECNOLOGÍA EN LOS REPERTORIOS MÉTRICOS

METRICS, CLASSIFICATION AND TECHNOLOGY
IN THE POETIC REPERTOIRES

José Luis RODRÍGUEZ GÓMEZ

Real Biblioteca. Patrimonio Nacional

jluis.rodriguez@patrimonionacional.es

Resumen: En el presente artículo se lleva a cabo un análisis del tratamiento de la métrica, las clasificaciones formales y temáticas, y la tecnología presente en el desarrollo de los repertorios métricos. Para ello se han seleccionado tres obras en formato convencional de libro impreso, que explotan el corpus de la poesía cancioneril castellana de los siglos XIV y XV, y una muestra representativa de repertorios electrónicos, que en muchas ocasiones, se presentan como actualización o retroconversión de obras impresas.

Abstract: The main aim of this article is the analysis of the treatment of metrics, formal and thematic classifications, and technology in the development of poetic repertoires. We have selected three works in conventional printed book format that exploit the corpus of Castilian “cancioneril” poetry in the 14th and 15th centuries, and a representative sample of electronic repertoires, which often are conceived as an upgrade or former printed works.

Palabras clave: Poesía. Métrica. Repertorios métricos. Humanidades digitales.

Key Words: Poetry. Metrics. Poetic Repertoires. Digital Humanities.

1. INTRODUCCIÓN

En este artículo se realiza un análisis del tratamiento de la métrica, las clasificaciones formales y temáticas, así como la tecnología presentes en el desarrollo de los repertorios métricos. Las tres primeras obras que se analizan están en formato convencional de libro impreso, y se centran en el amplio corpus de la poesía cancioneril castellana de los siglos XIV y XV (1). A continuación se analizan los repertorios electrónicos, que en muchas ocasiones, se presentan como actualización o retroconversión de una obra impresa (2). Para la obtención de los datos técnicos de cada uno de los proyectos examinados se ha tenido como base los datos de la encuesta realizada por Elena González-Blanco, "Preparatory Survey for Madrid Meeting" (2014). También se ha considerado como punto de partida el trabajo de González-Blanco García y Sélaf (2013), en el que se abordan el estudio de los repertorios métricos en formato digital con la finalidad de elaborar una propuesta de repertorio colectivo que permita la interoperabilidad entre las distintas propuestas. Los datos tomados de estas dos referencias han sido completados para este artículo con otros extraídos de la propia observación e interacción con las aplicaciones Web que constituyen la interfaz pública de estos recursos electrónicos. Finalmente, se presentan las conclusiones sobre los tres aspectos que constituyen los objetivos de este trabajo: métrica, clasificaciones y tecnología (3).

2. REPERTORIOS CASTELLANOS EN PAPEL

2.1. BIBLIOGRAFÍA DE CANCIONEROS CASTELLANOS... (STENOU Y KNAPP, 1975)

En su amplia nota preliminar se detalla la estructura de la obra y se delimita su dominio: cancioneros colectivos de fines del siglo XIV hasta principios del XVI. El volumen I recoge el "inventario bibliográfico", es decir, la noticia de los cancioneros y de los poemas contenidos en cada uno de ellos. Los volúmenes 1-2 se centran propiamente en el repertorio de los géneros literarios de la poesía cancioneril del siglo XV, objeto de estudio. Las entidades de que se sirve para estructurar los datos son las siguientes:

- E1: Sigla alfabética: "ofrece determinadas informaciones sobre las poesías de los cancioneros" (Stenou y Knapp, 1975: I: 39), hasta un total de 23. Generalmente determinan relaciones formales, materiales o temáticas entre dos o más poesías y pormenores de la atribución. Por ejemplo, la sigla C indica que el poema al

que se refiere incorpora textos ajenos; la B señala que el autor se ha identificado recurriendo a otras fuentes que recogen ese mismo texto. También se utiliza esta sigla para indicar la presencia de prólogos o epílogos, la pertenencia a un determinado ciclo, e, incluso, informaciones más complejas como la que, en relación a cancioneros musicales, nos informa de que “una copla, independiente en apariencia, es considerada como la mudanza de una poesía de forma fija (villancico o canción)” (Stenou y Knapp, 1975: I: 43).

- E2: Número del manuscrito o edición: Este número se corresponde con el “Inventario de manuscritos y ediciones antiguas” (Stenou y Knapp, 1975: I:106-123). Para cada entrada de esta lista, que contiene 90 referencias, se ofrece la ciudad, biblioteca y signatura que localiza el manuscrito o impreso, el título, extensión, dimensiones, bibliografía, y, en algunos casos, notas aclaratorias.
- E3: Número de la poesía: Es el número de orden interno de la poesía dentro del cancionero.
- E4: Localización de los folios en los que se encuentra la poesía.
- E5: Versificación: Se indica el número de estrofas y de versos de la composición. La codificación de este campo utiliza nueve dígitos, que responden a cuatro informaciones diferentes: Los dos primeros y los dos últimos (posiciones 0-1 y 7-8) indican, respectivamente, el número de versos de una estrofa-*preludio* o *estribillo*, y el de una estrofa final. El segundo grupo (posiciones 2-4 corresponde al número de estrofas), y el tercero (posiciones 5-6) al número de versos de cada estrofa. De esta forma, la cadena 000080804 se refiere a una serie de 8 estrofas de 8 versos, con una finida de 4 versos.
- E6: Categoría silábica: Este elemento informático está relacionado con el un Sistema de categorías silábicas previamente establecido. La referencia se establece mediante un código de dos dígitos. Por ejemplo:
 - 81: estrofas isométricas: versos de 8 sílabas y un verso corto en cada estrofa (Stenou y Knapp, 1975: I:35)
- E7: Tipo formal: Toma valores de la tabla “Sistemas de los géneros según la forma”, que representa un gran esfuerzo clasificatorio (Stenou y Knapp, 1975: I:45-50). Cada epígrafe está identificado por dos dígitos, siguiendo una única numeración, de 01 a 99.
- E8: Género según el contenido. Al igual que el anterior está relacionado con una tabla, en este caso el “Sistema de los géneros según el contenido” (Stenou, y Knapp, 1975: I: 51-78). Puede tomar más de un valor, y para su referenciación se utiliza también un sistema numérico con tres dígitos.
- E9: Lengua. Identifica con un dígito que remite a la tabla de lenguas el idioma en el que está escrita la composición.

- E10: Autores. Se consignan con tres dígitos que remiten a la tabla “Índice de autores” (Stenou y Knapp, 1975: I: 775-789). El listado está normalizado y contiene referencias de las formas no válidas a las aceptadas.
- E11: Incipit. Recoge los dos primeros versos de cada composición.

Estos 11 elementos informativos recogen la totalidad de los datos con los que este repertorio pretende organizar y sistematizar la materia objeto de estudio. A partir de ellos, se obtienen varios listados con la finalidad de facilitar la consulta.

Cada uno de los campos señalados consta de uno o más atributos, tal como mostramos en el diagrama de entidad-relación al que hemos traducido la estructura subyacente del repertorio (Ilustración 1). La mayoría de las relaciones son del tipo 1...n, es decir cada registro de una tabla secundaria corresponde a múltiples registros de la tabla principal. Es lo que ocurre, por ejemplo, con la tabla E2_Ms_Ed, manuscritos y ediciones, que está en relación con la tabla principal. Como ejemplo de relación de n...m, es decir de varios a varios, tenemos la que se establece entre la tabla E4_versificacion y la tabla principal: Un determinado modelo de versificación puede corresponder a múltiples poemas, y un poema puede contener múltiples versificaciones. Tipográficamente, el repertorio de Steunou, marca con el salto de línea cada una de estas versificaciones.

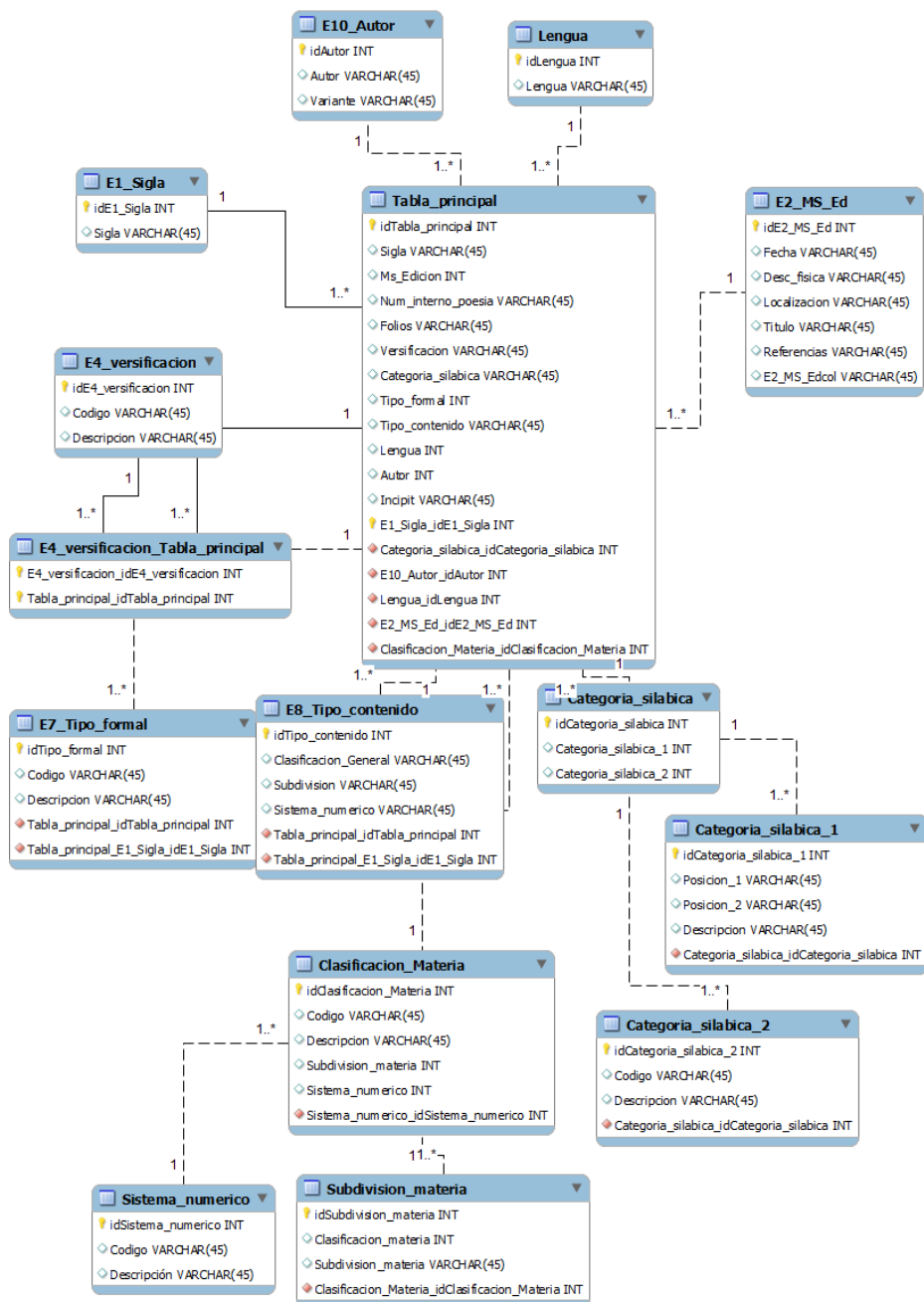


Ilustración 1: Esquema Entidad - Relación del Repertorio Steunou y Knnp

En la Nota Preliminar se atiende a la presentación de los datos. El volumen 1 recoge el contenido de cada uno de los cancioneros analizados, lo que denomina inventario

bibliográfico. Lo volúmenes 1-2 se centran propiamente en el repertorio de los géneros literarios de la poesía cancioneril del siglo XV.

El catálogo resultante es un listado impreso procedente probablemente del tratamiento de texto delimitado por etiquetas y tratado con programas realizados a medida. El Centro de Cálculo de la Maison de Sciences de l'homme es el responsable del tratamiento informático.

2.2. EL CATÁLOGO-ÍNDICE... (DUTTON, 1982)

La obra de Dutton ha tenido una enorme fortuna entre los estudiosos hasta convertirse en una referencia estándar para la identificación de la poesía cancioneril del siglo XV. A diferencia de la anterior, considera un corpus de testimonios mucho más amplio, beneficiándose en este aspecto de los trabajos de Joaquín González Cuenca (1978; 1980). Los 60 manuscritos y 30 impresos manejados por Steunou, pasan, respectivamente, a 199 y 221 en la obra de Dutton. Sin embargo, el trabajo clasificatorio de este corpus casi exhaustivo se limita a la consignación del género tal como aparece en la fuente, dejando al margen una clasificación genérica más estándar. La métrica, por su parte, queda reflejada con una fórmula numérica. En este sentido está más cerca del inventario que del repertorio métrico (Maguire, 2013)¹.

Hay una primera tabla que, encabezada por el epígrafe "Catálogo de fuentes manuscritas e impresas", constituye la relación exhaustiva de todas las composiciones de cada uno de los testimonios del corpus, siguiendo su orden de aparición en el manuscrito o impreso. Identifica en primer lugar el manuscrito o impreso con una sigla seguida de la ciudad y biblioteca en que se custodia, la signatura, una breve descripción física, la lengua y la datación. A continuación se ofrece el contenido, registrando un poema por entrada. Cada entrada contiene el identificativo, que remite al índice maestro, el número de orden interno dentro del cancionero, el autor, el folio, la fórmula métrica, los dos primeros versos, y, ocasionalmente, alguna nota aclaratoria. También se consigna el género o epígrafe de título, tal como aparece reflejado en el testimonio.

La tabla principal de este sistema de referencias es el denominado "Índice maestro" (Dutton, 1982: II, 1-140). Cada registro está encabezado por un número identificativo que remite unívocamente a un poema. A esa identificación sigue una fórmula numérica que declara el número de estrofas y versos, los datos de todas las versiones conocidas del poema, dejando constancia de los testimonios en los que aparece, el número de

¹ Como continuación, ampliación y retroconversión digital del proyecto de Dutton debe citarse el "Cancionero Manuscript Corpus" (<http://cancionerovirtual.liv.ac.uk>). Sobre los objetivos y aspectos técnicos de este proyecto digital puede consultarse (Maguire, 2013).

veces que es copiado o es objeto de glosa, lo que cuantifica su fortuna literaria, sus relaciones, sus atribuciones, etc...

Para facilitar el uso del repertorio, cuya función es localizar poemas en un corpus amplio previamente establecido, el investigador cuenta además de las dos tablas mencionadas, con los siguientes índices:

- Índice de primeros versos.
- Tabla de relaciones.
- Índice analítico de autores y poemas asociados a cada autor.
- Índice de destinatarios.
- Índice de títulos.
- Índice de relaciones.
- Índice de manuscritos y de fuentes impresas.

Para traducir el esquema subyacente a un modelo conceptual, se debe determinar una tabla principal y un conjunto de tablas auxiliares con las que se relaciona². En el caso que nos ocupa, el "Catálogo de fuentes manuscritas e impresas" realiza la función de tabla principal, ya que recoge toda la información que se extrae del dominio cancioneril. Como tablas relacionadas, se utilizan las siguientes:

- Tipos de relación establecida entre los poemas (Dutton, 1982: xi-xii) Se mencionan 14 relaciones.
- Índice de fuentes manuscritas (Dutton, 1982: 233-236).
- Índice de fuentes impresas (Dutton, 1982: 239-244).

Añadiendo los atributos correspondientes a cada tabla, obtenemos el siguiente esquema de tipo Entidad-Relación:

2 Este repertorio, al igual que el anterior, no fue realizado con un sistema de gestión de bases de datos relacionales. Nos servimos de un esquema relacional ya que nos permite homologar metodológicamente las distintas propuestas. Para una descripción técnica del proceso, véase "Apéndice I. Descripción técnica", por Francisco Santoyo Vázquez, en Dutton (1982: t. II, 269-274).

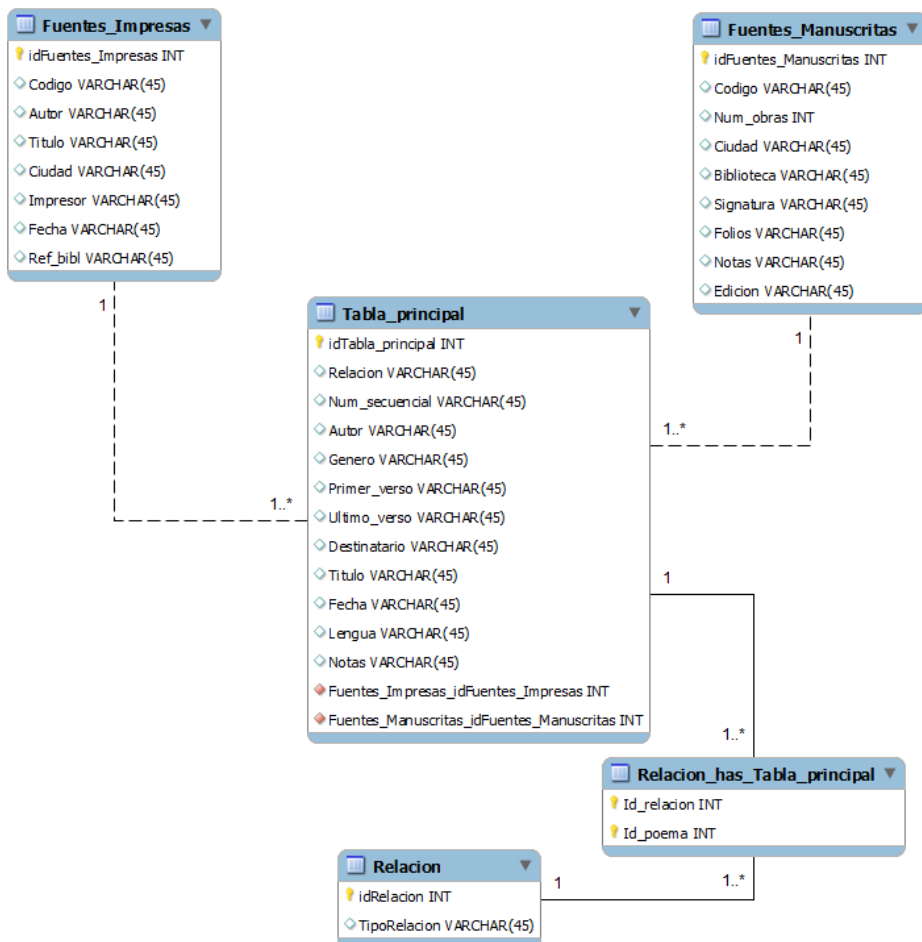


Ilustración 2: Esquema entidad-relación del Índice de Dutton

2.3. EL REPERTORIO MÉTRICO DE GÓMEZ-BRAVO (1998)

En Dialnet se presenta así el trabajo de Gómez-Bravo:

El presente “Repertorio” contiene la información métrica correspondiente al corpus cancioneril publicado por Brian Dutton. El “Repertorio” está inspirado en el modelo que István Frank realizó para la lírica provenzal y al que en la misma línea, siguieron el de Tavani y otros varios. Este trabajo toma los textos poéticos según su versión o versiones manuscritas³.

3 Dialnet: <http://dialnet.unirioja.es/servlet/libro?codigo=195165> (Consultado 15/09/2015).

En efecto, esta obra suple la carencia que hemos señalado en el Catálogo-Índice (1982), al añadir los campos de información métrica, y tuvieron a su disposición la transcripción electrónica del proyecto de Dutton. El responsable de la programación y del formato de salida de los datos fue Thorsten von Eicken (Cornell University). Se utilizó un ordenador con UNIX, y TROFF, que es el procesador de textos que AT&T desarrolló para ese sistema operativo.

En cuanto a la selección de los textos, con buen criterio no se recurre a ediciones críticas, sino que se tomarán directamente de los testimonios, dada la dificultad, a veces insalvable, que oponen los avatares de la transmisión manuscrita de la poesía a la labor ecdótica para la fijación de un texto crítico válido. Los distintos testimonios no solo presentan numerosas variables significativas, sino también variaciones en el patrón métrico.

El criterio de ordenación de las entradas del repertorio es jerárquico. El primer criterio es el número de versos de la estrofa, y es ascendente, al igual que el segundo criterio, que es el orden alfabético de los esquemas de la rima. El tercer criterio es de ordenación descendente y corresponde al esquema de sílabas. Dentro de cada uno de esos criterios, hay una serie de convenciones para resolver casos por coincidencia a causa de repetición de versos, por ejemplo. La jerarquización de esta organización alcanza hasta ocho criterios. Cada entrada, con todos los campos informativos, se estructura en tres líneas, dando legibilidad al catálogo, a pesar de su complejidad.

La recuperación de información, al margen de la métrica, que como hemos descrito está perfectamente organizada, se garantiza con los índices de primeros versos, autores, e identificativos que se corresponden con el *Catálogo-Índice...* de Dutton (1982). También es útil el índice de manuscritos e impresos que encabeza el repertorio y que identifica con precisión y concisión las fuentes de las que se toman los poemas.

Cuenta este repertorio, desde el presente año (2015), con una versión digital en línea (Escribano, Juanjo 2014). Para la creación de la aplicación Web la tecnología utilizada fue el RDBMS MySQL y PHP como lenguaje de servidor. Los campos que se han tenido en cuenta son:

- ID_Dutton: Numero identificativo de la composición en Dutton.
- Man: Identificación del cancionero. Se utilizan las siglas de Dutton.
- Poem_serial: Número de orden interno de la composición en el cancionero.
- Folio: Localización del folio en el que se encuentra el poema.
- Stanza: Número de versos de las estrofas.
- Syllabe Signature: Esquema métrico.
- First - second - last (line): Primero, segundo y último verso.
- Esquema rimático.
- Rimas.
- Número de sílabas del verso.

Esta primera versión permite filtrar la información por manuscrito y esquema métrico, siguiendo las convenciones notacionales del repertorio. Así, la fórmula "1×4@8", que sigue las convenciones notacionales del repertorio, en una búsqueda exacta, nos devuelve los poemas integrados por una única cuarteta octosilábica, tanto en la cabeza como en el cuerpo del poema.

3. REPERTORIOS DIGITALES

3.1. BASE DE DATOS DA LÍRICA PROFANA GALEGO-PORTUGUESA [MEDDB]

URL: <http://www.cirp.es/bdo/med/meddb.html> (ISSN 1989 4546)

Duración del proyecto: En curso. Con varias fases finalizadas.

Especificaciones técnicas: Partiendo de ediciones críticas, se codifican los datos en XML. Esta información etiquetada se transforma a un modelo entidad-relación para ser volcada en una base de datos relacional. A esta base de datos se añaden índices para optimizar las búsquedas y permitir lanzar consultas a los campos de texto. El usuario final se conecta con su navegador a la aplicación Web que accede a esta base de datos. Una serie de formularios e índices permiten, respectivamente, introducir los criterios de búsqueda, o seleccionarlos a partir de un listado. También están disponibles toda una serie de operadores booleanos para refinar la interrogación.

Contenido: El dominio de MedDB es la lírica profana gallego-portuguesa. El periodo cronológico va desde el último cuarto del siglo XII a la primera mitad del siglo XIV (ca. 1180-1350). Tiene como base el repertorio del corpus lírico profano gallego-portugués (Tavani, Giuseppe 1967).

Estado del proyecto: En este momento (2015) se están incorporando las imágenes y la edición paleográfica de las composiciones. La edición paleográfica considerará todos los elementos que rodean al texto: numeración en el códice, anotaciones, ángulos de inicio de cantiga, marcas de refrán, etc.

Se trata de un proyecto en curso, en actualización constante, y atento a las publicaciones especializadas -monografías, artículos, ediciones, etc.- de ámbito nacional e internacional. Las referencias a estas publicaciones, así como las modificaciones derivadas de ellas, se incorporan a MedDB.

Finalmente, en colaboración con equipos de las universidades de Roma, Cosenza y Siena, se está trabajando en la lematización. Como resultado final, se prevé la integración de MedDB con las bases de datos occitana, francesa e italiana.

Institución responsable: Centro Ramón Piñeiro y la Universidad de Santiago de Compostela. Está financiado por la Xunta de Galicia y la Consellería de Cultura y Educación.

Investigador responsable: Mercedes Brea y Pilar Lorenzo.

MedDB aborda no solo los datos, es decir la ficha técnica de las distintas entidades -versos, estrofas o poemas-, sino que también ofrece la transcripción de los textos y la edición facsímil, lo que supone un avance importante, incrementando su utilidad como herramienta de investigación⁴.

La base de datos relacional está formada por un total de cuatro tablas. La tabla "Cantigas" es la principal y consta de los siguientes campos o atributos:

I.- Tabla "Cantigas" (Songs).

- IDC: Identificador (uso interno de la BD).
- IDE: Código que identifica la cantiga.
- NUCANT: Número de cantiga.
- TC: Tipo.
- NE: Número de estrofas.
- SG: "de seguir", booleano SI / NO.
- XE: Género.
- OL: Otras lenguas que no sean el gallego.
- INCIPIT: Primer verso de la composición.
- DATOAU_NOME: Fecha de edición (1a parte).
- DATOAU_NUA: Fecha de edición (2a parte).
- TEXTO: Texto completo de la cantiga.
- BMED_IDT: Clave foránea que se relaciona tabla de trovadores.

La segunda tabla, Estrofas, está relacionada con la anterior por un campo IDC, identificativo de la cantiga. Se trata de una relación de composición

II.- Table "Estrofas"

- IDS: Identificador. Es la clave primaria, de uso interno para la BD.
- IDC: Identificador de la cantiga. Es clave foránea que toma valores del identificador de la tabla Cantigas.
- TIPO: Tipo de estrofa.
- NUM: Número de estrofa.
- NUTOTVER: Número total de versos.
- ESQRIM_TIPO: Tipo de esquema rítmico.
- NUMESQ: Número de esquema rítmico.
- ESQMET: Esquema métrico.

4 Para una aproximación a sus posibilidades de explotación es de utilidad el "Manual MedDB 2", disponible en el sitio Web del proyecto.

- TEXTO: Texto completo de la estrofa.

La tercera tabla registra la información de cada uno de los versos, y está relacionada con las dos anteriores

III.- Tabla "Versos".

- IDV: Identificador del verso. Es clave primaria, de uso interno para la BD.
- IDS: Clave foránea que relaciona el verso con la estrofa a la que pertenece.
- IDC: Clave foránea que relaciona el verso al poema al que pertenece.
- REFRÁN: Si el verso es o no parte de un refrán.
- N: Número de verso.
- NS: Número de sílabas.
- RI: Rima.
- ERI: Elemento de la rima.
- LL: Transcripción completa del verso.

La tabla cuarta se ocupa de los trovadores. De ella toma sus valores el campo BMED_IDT de la tabla "Cantigas".

IV.- Table "Trovadores".

- IDT: Identificador. Clave primaria de uso interno de la BD.
- NTROVADOR: Número de trovador.
- NOME: Nombre.
- ACTDESDE: Periodo de actividad. Año *a quo*...
- ACTATA: Periodo de actividad. Año *ante quem*.
- CONSOC: Status social.
- FIAORI: Grado de certidumbre del campo ORIXE.
- ORIXE: Lugar de nacimiento.
- NOMEABREV: Nombre abreviado del trovador.
- PERIODO: Periodo cronológico.
- BIOT: Reseña biográfica.
- CANTIGAS_IDC: Clave foránea que remite a la cantiga.

La quinta tabla da cuenta del manuscrito que copia la composición referenciada por la clave foránea IDC.

V.- Tabla "RefManus".

- IDC: Identificador de la Cantiga.
- NM: Número de manuscrito.
- F: Folio en el que se encuentra la cantiga.
- MU: Si contiene notación musical.
- EM: Si hay o no espacio reservado para la notación musical.
- Finalmente, la tabla VI recoge las materias de las composiciones.
- VI.- Tabla "Temas".

- IDC: Identificador de la cantiga.
- Tema: Tema principal de la cantiga.
- Subtema1: Campo temático más preciso.
- Subtema2: Otro campo temático más preciso.

3.2. CANTIGAS DE SANTA MARÍA DATABASE (CSM)

URL: www.csm.mml.ox.ac.uk/

Especificaciones técnicas: Sistema de gestión de bases de datos relacionales MySQL

Contenido: Recoge la poesía religiosa del Alfonso X el Sabio. Se trata de poemas de tipo narrativo, que concuerdan temáticamente con las recopilaciones de milagros de la Edad Media, y que constituyen en muchos casos la fuente de la que se sirven los textos del rey. El proyecto recoge también la descripción analítica de las principales colecciones de milagros marianos, generalmente en prosa, y la versión digital del *Index* de Poncelet (1902) El plan de CSM atiende a aspectos que van más allá de un repertorio métrico al considerar la diferentes versiones de un poema e incluir una extensa base de datos bibliográfica y un índice de milagros, tomado del mencionado Poncelet.

Institución responsable: Universidad de Oxford.

Investigador responsable: Stephen Parkinson.

CSM consta de once tablas relacionadas. La principal es Poems, que incluye un extenso conjunto de atributos, suficiente para un análisis riguroso de los poemas. Además de los campos puramente identificativos y de localización del texto, incluye información especializada sobre los aspectos métricos:

- Stanzanumb: número de estrofas.
- Zéjel: Booleano yes/no.
- Stanformula: Esquema métrico de las estrofas. Por ejemplo: 10 10 10; 15[7'7] 15'[7'7']; 6! 6!
- refformula: Esquema métrico para el refrán.
- Rhymecode: Esquema de la rima para todo el poema. Por ejemplo: AA|bbba.
- metformula: Esquema métrico para todo el poema, construido a partir de los otros campos.
- Altstform: Esquema métrico alternativo.
- Alrhyme: Esquema de la rima alternativo.
- Mettype: Código numérico para el esquema métrico completo (cod. 1-13), tomado de Mettmann 1986/89.
- Metnote: Comentario libre sobre la métrica.
- Rhymenote: Comentario libre sobre la rima.

Las tablas relacionadas atienden a atributos como palabras clave, tipo de rima, bibliografía o testimonio. Esta última, 'witness', refleja con exactitud la puesta en página del poema en el testimonio concreto del que se toma. Así, incluye información tan precisa como:

- stanzaform: Líneas de texto usadas en el MS.
- textunderlaid, runningtext: División del texto
- refslost: Omisión del refrán por razones de la puesta en página.

El identificativo del Index de Poncelet se recoge en la tabla principal, en el campo Poncelet. Este campo es una clave foránea relacionada con la tabla Miracles, que describe cada una de las entradas del Index. Por otra parte, Miracles se relaciona con Collection, que vincula cada milagro con la recopilación medieval que lo contiene.

Al igual que MedDB, la base de datos CSM ofrece el acceso a los textos, en este caso en formato pdf. Es la tabla poem_link_map la permite relacionar la tabla Poem, que describe el poema, con su versión en pdf. En futuros desarrollos se prevé la creación de una tabla tipo texto que recoja la edición codificada en XML.

3.3. CORPUS RHYTHMORUM MUSICUM [CRM]

URL: <http://www.corimu.unisi.it>

Especificaciones técnicas: Software desarrollado por Luigi Tessarolo. Cuenta con una versión en línea disponible a través de Internet, y de una versión más completa en CD-ROM.

Contenido: Repertorio poético de textos latinos con música de los siglos IV-IX. Contiene esquemas métricos, textos poéticos completos, concordancias, referencias bibliográficas. El *Corpus...* incorpora la edición crítica de todos los testimonios conocidos y la reproducción facsímil de cada manuscrito. De esta forma, el investigador podrá verificar las lecturas propuestas y disponer de los materiales necesarios para su estudio.

Estado del proyecto: En progreso.

Institución responsable: Università di Siena, Dipartimento di Teoria e documentazione delle tradizioni culturali, Società Internazionale per lo Studio del Medioevo Latino.

Investigador responsable: Francesco Stella

Otros datos: Está en relación con otras bases de datos como el *Archivio della Letteratura Italiana del Medioevo* (ALIM).

Se trata de uno de los proyectos más innovadores y ambiciosos. El filólogo especializado tiene la posibilidad de verificar las lecturas de los diversos testimonios a través del facsímil o de sus transcripciones. En cuanto a recuperación de la información, el usuario dispone de varios

paneles en la barra horizontal superior⁵. Si accedemos Al menú Texts y pinchamos en una entrada al azar visualizaremos su edición crítica. Al paso del ratón se nos abrirá en una ventana nueva información textual de interés filológico: aparato crítico, fuentes o lugares paralelos en textos coetáneos. En este mismo panel, el botón Synopsis nos traslada a una edición crítica, enriquecida, en algunos casos, con con diferentes versiones históricas. La pestaña 'Criterios generales' despliega una ficha descriptiva con los campos necesarios para la identificación del texto: Número, incipit, autor, título, fecha, procedencia geográfica, número ICL, número total de testimonios, testimonios no utilizados en versiones anteriores, número de versiones, y referencias filológicas y estudios. La pestaña Introduction nos lleva a un estudio filológico riguroso realizado por el editor crítico, profusamente anotado, con el análisis de todos los testimonios e incluyendo en varias ocasiones el *stemma codicum*. En estos casos, la sigla del manuscrito está enlazada con su transcripción. En Versification accedemos al registro técnico completo de los datos métricos: número de versos, patrón métrico, SAG1, SAM1, tipo estrofa, tipo verso, STM Verse, diéresis, hiatos, sístole, diástole o hipometría. De especial interés, es el panel Manuscript Tradition and Music, que despliega varios subpaneles. En el subpanel Transcripts accedemos a la identificación de los testimonios que recogen el texto en cuestión, a la transcripción, y a su reproducción digital. Si el testimonio contiene notación musical, la pestaña Notation nos permitirá leer la información sobre estas anotaciones e incluso su transcripción alfanumérica. También es posible acceder a las transcripciones históricas o creativas producidas por musicólogos de los siglos XIX y XX y a los correspondientes registros de audio. Finalmente, la tabla Language nos conduce a un análisis lingüístico, con el esquema sintáctico, las locuciones preposicionales o el vocalismo.(1998)

Frente a otros repertorios, el presente *Corpus...* añade un módulo estadístico, que recoge información sobre la distancia Sujeto / Verbo y Nombre / Adjetivo, la frecuencia y posición de las preposiciones, y las relaciones entre funciones sintácticas como Sujeto / objeto o grupos de complementos.

3.4. BIBLIOGRAPHIA ELETTRONICA DEI TROVATORI (BEDT)

URL: <http://www.bedt.it>

Duración del proyecto: Comienza en 2003. La versión 2.5 está disponible desde 2012.

Especificaciones técnicas: Aplicación propietaria desarrollada en ASPNET, entorno Windows, creada en 2003 y actualizada en 2008.

Contenido: Dedicado a la lírica los trovadores provenzales, integra y actualiza los repertorios bliográficos de Pillet y Carstens (1933) y el repertorio métrico de Frank (1953). Incluye:

5 Véase "Introductions" y "Help" en <http://www.corimu.unisi.it/> [Consultado 21/09/2015] para una descripción detallada del sitio y su funcionamiento.

- Descripción analítica del contenido de los cancioneros que incluyen textos trovadorescos.
- Relación de los cancioneros y manuscritos con los textos.
- Actualización de la bibliografía crítica.

Estado del proyecto: Actualizaciones periódicas. En 2012 aparecer la versión 2.5 (última versión consultada).

Vocabularios controlados y taxonomías: Se ha tenido como base las clasificaciones de repertorios clásicos como Pillet y Frank, y, en general, la riquísima tradición de los estudios sobre lírica provenzal.

Institución responsable: Università la Sapienza, Roma.

Investigador responsable: Stefano Asperti.

Las tablas principales de las base de datos son Autores, Textos, Fuentes y Bibliografía. La primera, mejorada desde la versión 2,5, recoge la información biográfica detallada de los autores: referencia a Pillet y Carstens (PC), nombre, lugar de origen y actividad, y otros detalles. La tabla de textos remite también a PC y a Frank, y recoge el incipit, la tradición e hipótesis nuevas en cuanto a atribuciones del texto, hipótesis de atribuciones, género de la composición de acuerdo con varias tablas clasificatorias (BedT, PC), argumento o la tipología siguiendo a Frank. Por lo que se refiere a la métrica, se consigna el esquema métrico y rítmico. Termina esta tabla con el campo "vidas e razos", que recoge la relación con posibilidad de detalle de la vida y "razos" de los trovadores.

La información sobre las fuentes es suficiente amplia para la identificación del cancionero (Sigla, Biblioteca). Contiene campos de relación que vinculan la pieza con el texto, autor, y la clasifican de acuerdo con varios esquemas de clasificación.

Este proyecto se propone como germen de un archivo en línea de toda la tradición lírica europea medieval, incluyendo autores, textos, métrica, manuscritos y bibliografía crítica actualizada.

Finalmente, la tabla Bibliografía contiene una ficha estándar con los datos bibliográficos básicos, a los que se añaden palabras clave, un resumen, y enlace, si está disponible en Internet, al contenido.

3.5. LE NOUVEAU NAETEBUS. POÈMES STROPHIQUES NON-LYRIQUES EN FRANÇAIS DES ORIGINES JUSQU'À 1400 [NAETEBUS]

URL: <http://nouveaunaetebus.elte.hu>

Duración del proyecto: 2004-2008 pero aún en desarrollo.

Especificaciones técnicas: Base de datos en SQL accesible vía Web. El RDBMS es PostgreSQL.

Contenido: Repertorio de la poesía narrativa medieval francesa no lírica producida entre los años 1100 y 1400. Tiene como punto de partida la retroconversión digital del clásico repertorio de Gotthold Naetebus (1891) El registro de los poemas va acompañado de sus esquemas estróficos, de una taxonomía temática y de funciones y de una base de datos bibliográfica. Además de las composiciones recogidas por Naetebus, considera también:

- Los poemas antiguos franceses que se corresponden formalmente con los recogidos por Naetebus pero ausentes en su repertorio
 - Otros poemas heterogéneos del francés antiguo, ausentes de Naetebus y de Molk y Wolfzettel (1972), independientemente de su carácter lírico o no-lírico
- Es especialmente útil la concordancia con los principales repertorios.

Estado del proyecto: Se han tratado ya alrededor de 400 composiciones, y se han incorporado datos del incipitario de Arthur Långfors (1918), e información nueva extraída directamente de los testimonios manuscritos.

Vocabularios controlados y taxonomías: Se proponen categorías genéricas modernas, además de anotar las denominaciones genéricas medievales. Algunos de los grandes grupos clasificatorios de esta propuesta taxonómica son: Función, Uso, Tema, Registro, Tono, Narratividad, Lengua.

Institución responsable: Budapest, Eötvös Loránd University of Sciences

Investigador responsable: Levente Seláf.

El formulario de búsqueda, perfectamente desarrollado, con campos para la introducción de criterios y para la selección de valores, permite recuperar información de forma precisa.

3.6. SKALDIC POETRY OF THE SCANDINAVIAN MIDDLE AGES [SKALDIC]

URL: <http://abdn.ac.uk/skaldic>

Especificaciones técnicas: Aplicación Web. S.O Linux, RDBMS MySQL, Programación Web PHP. Posibilidad de exportación a formato TEI XML para edición.

Duración del proyecto: En continua actualización, desde 2001.

Investigador responsable: Tarrin Wills.

Otros datos: Acceso mixto: libre y bajo licencia. El material publicado por Brepols solo es accesible a través de los volúmenes impresos. Otros materiales externos también están protegidos.

Vocabularios controlados y taxonomías: Existen tablas para controlar las categorías clasificatorias, palabras claves, esquemas métricos, tipos de dicción.

La BD consta 148 tablas con cerca de 2000 atributos. Las entidades principales en torno a las que se articula esta estructura son:

- Poetas/autores (Tabla 'skalds').
- Obras en verso y prosa (Tabla 'text').
- Manuscritos y otros objetos físicos soporte de textos (tabla 'mss').
- Estrofas (Tabla 'verses' table).
- Líneas (Tabla 'line').
- Palabras (Tabla 'word').
- Diccionario de lemas (Tabla 'lemma').
- Formas nativas de dicción poéticas (Tablas 'kenning', 'kenningpart', 'kenningref').
- Bibliografía (Tablas 'bibl_works', 'bibl_authors', 'bibl_journals').
- Categorías métricas del Norse medieval (Tabla 'metre').
- Categorías métricas de versos que usan Sievers (Tabla 'linetype').

En cuanto al sistema de recuperación de información, la aplicación dispone de un formulario que lanza los criterios de búsqueda introducidos por el usuario a las distintas tablas: biografías, textos introductorios, versos y notas. También permite el recurso a expresiones regulares POSIX. Por otra parte, un menú horizontal en la parte superior facilita el acceso a cada una de las tablas principales de la base de datos, brindando la posibilidad de navegación alfabética por índice o búsqueda por palabra clave.

3.7. RÉPERTOIRE DE LA POÉSIE HONGROISE ANCIENNE [RPHA]

URL: <http://rpha.elte.hu/>

Duración del proyecto: Desde 1979.

Especificaciones técnicas: Base de datos en SQL.

Contenido: Poemas húngaros desde finales del s. XIII hasta el s. XVI. Recoge variantes métricas a partir de todos sus testimonios manuscritos.

Estado del proyecto: Terminado.

Investigador responsable: Iván Horváth.

Institución responsable: El proyecto se desarrolló primero en la Universidad de Szeged (1979–1992) y luego pasó a Eötvös Loránd University (ELTE, Budapest). Está vinculado con el proyecto RPHA17, con nuevas perspectivas tecnológicas, en particular, etiquetado TEI XML.

El *Répertoire de la poésie hongroise ancienne* (RPHA) implementa una concepción actual de la crítica textual, al considerar el poema como la interscisión de objeto material que le sirve de soporte y texto abstracto. Así pues, se abandona la reconstrucción de un arquetipo que postula su acercamiento a la intención del autor. En su lugar, se tienen en cuenta todos los testimonios, manuscritos e impresos, ya que es toda la serie la que constituye la integridad del poema. El RPHA registra la descripción de todas y cada una de las variantes de un texto.

3.8. PEDE CERTO [PEDECERTO]

URL: <http://www.pedecerto.eu/home.jsp>

Contenido: Es parte del proyecto FIRB *Traditio patrum*. Incorpora también el Archivo digital Musique Deoque, que aborda textos poéticos latinos del período arcaico.

Institución responsable: Università degli Studi di Udine.

Responsables: Emanuela Colombi, Luca Mondin, Luigi Tessarolo, Andrea Bacianini.

El motor de búsqueda permite satisfacer las necesidades de la más exigente investigación en métrica, con múltiples aproximaciones al corpus métrico. Así, son posibles seis modos de interrogación: por palabra clave, por palabra-tipo (secuencia de largas y breves), esquema del verso (secuencia de metro y pausas), esquemas y particularidades, cadena concreta en un determinado esquema, o secuencia métrica en cualquier posición. Pero la gran aportación de este proyecto es el programa de análisis automático de la poesía latina: La opción "free scansion" nos permite comprobar el funcionamiento del programa de análisis e interpretación, introduciendo versos, siguiendo unas convenciones de transcripción, y obteniendo automáticamente su escansión.

3.9. DIÁLOGO MEDIEVAL. EDICIÓN DIGITAL [DMED]

URL: http://sade.textgrid.de/exist/apps/SADE/Dialogo_Medieval/index.html

Duración del proyecto: Desde 2015.

Especificaciones técnicas: Este proyecto es pionero en este campo en la implementación de TextGrid («The Project TextGrid - TextGrid» 2015), proyecto financiado por el German Federal Ministry of Education and Research. TextGrid da respuesta a la creciente demanda desde el mundo de las humanidades de herramientas digitales de investigación ofreciendo servicios y productos de distribución gratuitos. En concreto, el entorno de investigación virtual TextGrid facilita las tareas más demandadas desde

la investigación: la edición, el almacenamiento y la publicación de datos en formatos estándar⁶.

Contenido: Aborda la edición y publicación digital de 50 debates y poemas dialogados incluidos en el *Cancionero de Baena*⁷, con este objetivo:

it seeks to establish the basic characteristics of verse structure in dialogue poetry and give account of possible imitations or contrafactures, common verse, and metrical and rhyming patterns, in order to create a system able to answer types of research questions that are common in medieval versification studies and that involve regularities that show intertextual meanings conventionally associated with a given poetic meter (Río Riande, Gimena del 2015) .

Investigador responsable: Gimena del Río Riande.

Vocabularios controlados y taxonomías: El proyecto cuenta con un Tesauros para el control de vocabulario y categorías clasificatorias. Desarrollado en TemaTres está disponible a través de Internet⁸. Las grandes categorías en torno a las que se articula son las siguientes:

- Agentes de producción.
- Clasificación general.
- Epítetos poéticos.
- Estructuras poéticas.
- Prácticas compositivas.
- Tipo de poema según forma.
- Tipo de poema según tropo poético.

Como se desprende de la nota de contenido, se trata de un proyecto de investigación para esclarecer la estructura formal a través de patrones de un corpus poético muy concreto, trascendiendo así la función de repertorio métrico. En cualquier caso, es de especial interés para nuestro propósito por la utilización que se hace de TEI-Verse⁹ para el análisis métrico.

6 TextGrid utiliza la base de datos eXist, XML Nativa («eXistdb - The Open Source Native XML Database» 2015). Su sitio Web ofrece información detallada y enlaces a recursos sobre este entorno de edición y publicación digital: («SADE Publish Tool - TextGrid - Dariah wiki» 2015)

7 Madrid, BNE, MSS/5636; Philobiblon, texid BETA 1419

8 Véase (ReMetCa 2014). TemaTres es un sistema open source para la creación de Tesauros, taxonomías, ontologías o listas de valores, es decir, representaciones del conocimiento en varios formatos. Debido a su arquitectura orientada al servicio, permite que otras aplicaciones compartan estos datos (Véase («TemaTres Controlled Vocabulary server» 2015)).

9 Véase «Tei Verse» en («TEI: Text Encoding Initiative» 2014)

Se aporta también la edición en paralelo del facsímil y la transcripción, con la opción siempre visible de acceso a la ficha descriptiva de información de tipo métrico y al archivo fuente XML-TEI, lo que facilita la interoperabilidad con proyectos semejantes. En particular, DMED comparte datos con ReMetCa.

3.10. EL REPERTORIO MÉTRICO DIGITAL DE LA POESÍA MEDIEVAL CASTELLANA [REMETCA]¹⁰

URL: <http://www.remetca.uned.es>

Duración del proyecto: desde 2012.

Especificaciones técnicas: Base de datos en MySQL. Contiene un campo tipo XML con el etiquetado TEI, módulo TEI Verse, de los textos.

Contenido: Repertorio métrico en castellano desde sus primeras manifestaciones, a finales del siglo XII, hasta la riqueza y variedad de formas métricas de la poesía cancioneril de los siglos XV y principios del XVI, hasta 1511, fecha del *Cancionero General*, de Hernando del Castillo. El proyecto incluye el contenido del repertorio de Gómez Bravo (1998).

Estado del proyecto: En proceso de realización.

Institución responsable: UNED, Madrid.

Investigador responsable: Elena González-Blanco García.

Vocabularios controlados y taxonomías. En la tabla 'Obra' se incluye el campo 'claveHismetca', que realiza una clasificación de la poesía cancioneril siguiendo las categorías establecidas para el proyecto Hismetca, del Centro de Estudios Cervantinos (Gómez Redondo, Fernando, dir. 2015), un total 16:

- 01 Principios poéticos.
- 02 Épica y juglaresca.
- 03 Debate. • 04 Noticieros e historiográficos.
- 05 Poesía clerical.
- 06 Poesía aljamiada y clerecía rabinica
- 07 Poesía lírica castellana tradicional
- 08 Poesía lírica castellana cortés
- 09 Poesía hagiográfica y didáctica
- 10 Poesía cortesana y cancioneril
- 11 Oraciones y textos litúrgicos
- 12 Poemas historiográficos
- 13 Prosimetra: exégesis, alegoría y ficción sentimental
- 14 Romancero

¹⁰ ReMetCa es el repertorio que seguiremos más de cerca en nuestra propuesta. Para una aproximación más detallada que la que ofrecemos en este epígrafe, puede verse (González-Blanco y Rodríguez, 2014).

- 15 Poesía italianizante
- 16 Dramaturgia

Por su parte, la tabla 'Poema' contiene dos campos clasificatorios, 'tipo' y 'tema'. El primero se rige por criterios temáticos y formales. El segundo está integrado por un listado abierto de palabras claves que consignan la materia del poema.

Las tablas principales de la BD de ReMetCa son Obra y Poema (Ilustración 3). La primera consta de 22 atributos con los que se pretende dar cuenta de las propiedades que definen una obra que incluye varios poemas. Esta obra puede haber sido transmitida por más de un manuscrito o edición, es decir, puede tener múltiples testimonios, de ahí su relación con las tablas testimonio y manuscritos. Del mismo modo, se relaciona con la tabla 'bibliografía, que da cuenta de las referencias en las que ha sido tratada una determinada obra. El 'Autor' tiene también una tabla, con un número de atributos mínimo pero suficiente para su identificación unívoca y la consignación de alguna variante de nombre.

La tabla Poema consta de 18 atributos, y describe cada uno de los poemas, desde sus elementos identificativos (id, incipit) y aspectos puramente métricos (número de versos, estrofas, fórmula métrica, etc.), hasta su asignación a distintas categorías (palabras clave o tipo). Sus tablas relacionadas son el 'nombreMetrico', las clasificaciones por 'tipos' y 'palabrasClave', y por supuesto, la 'Obra' a la que pertenece el Poema. De especial interés es el campo 'codigo_tei', de tipo XML, y que se utiliza para codificar el poema en formato TEI¹¹, tal como se muestra en el siguiente ejemplo:

```
<| n="1">¡Oh viejo desventur<rhyme label="a">ado</rhyme></l>
<| n="2">Negra dicha fue la m<rhyme label="b">ía</rhyme></l>
<| n="3">en casarme con Mar<rhyme label="b">ía</rhyme></l>
<| n="4">por quien fuese deshonor<rhyme label="a">ado</rhyme>.</l>
<| n="5">Ya la veo bien preñ<rhyme label="c">ada</rhyme>,</l>
<| n="6">no sé de quién, nin de cu<rhyme label="d">ánto</rhyme>.</l>
<| n="7">Dizen que d'Espiritu S<rhyme label="d">anto</rhyme>,</l>
<| n="8">mas yo d'esto non sé n<rhyme label="c">ada</rhyme></l>
```

11 Por lo que se refiere al tipo de campo XML, en particular en MySQL, puede verse («MySQL 5.1 Reference Manual: 12.11 XML Functions» 2015).

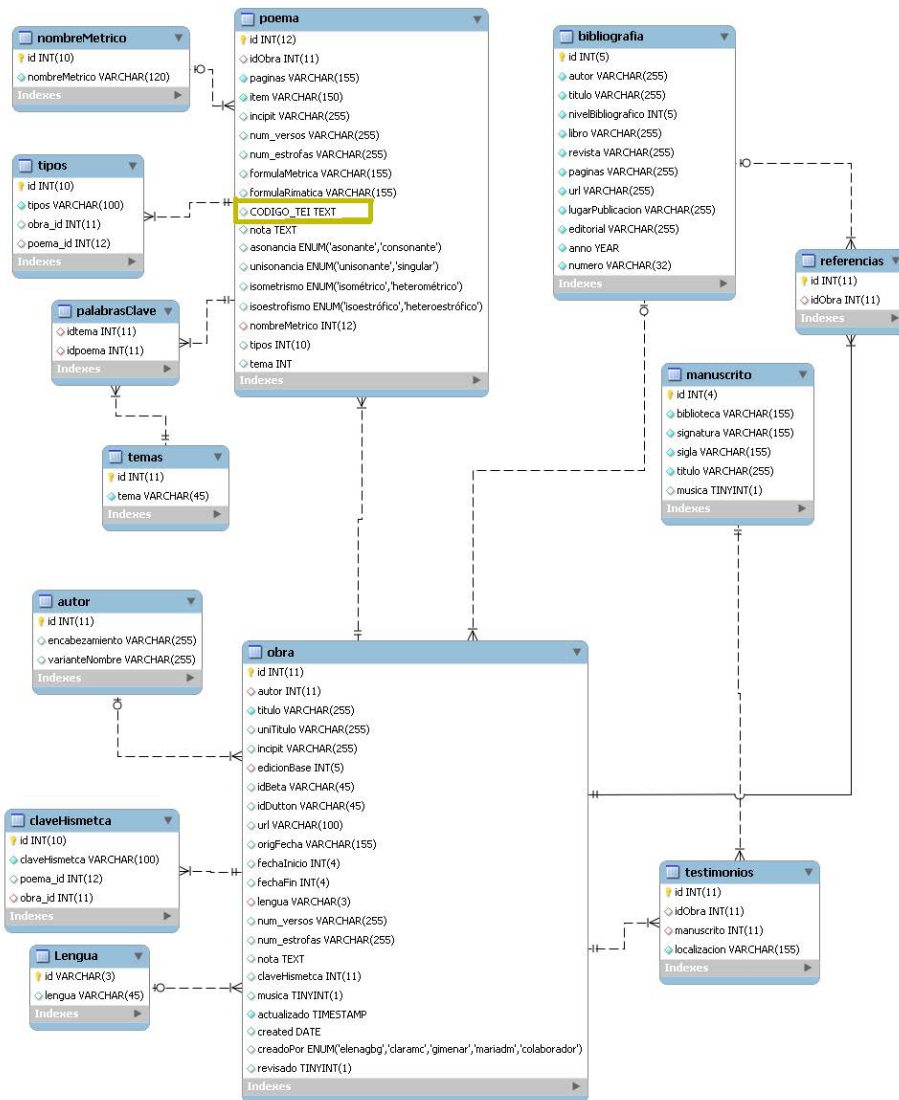


Ilustración 3. Esquema conceptual de ReMetCa

4. MÉTRICA, CLASIFICACIÓN Y TECNOLOGÍA

Dejaré al margen del análisis de los repertorios registrados en el apartado anterior los aspectos organizativos y de gestión de los proyectos, tales como el estado actual, la duración, los equipos que intervienen, y su titularidad y responsabilidad, corporativa o individual, para centrarme en los rasgos más técnicos.

Analizaré en primer lugar los aspectos métricos, y, en particular, los esquemas utilizados para mostrar formulariamente el número de sílabas, la rima y su distribución en la unidad considerada, poema o estrofa (3.1). Los vocabularios controlados, las clasificaciones y taxonomías se tratan en el punto 3.2. Finalmente, las conclusiones sobre tecnología se recogen en 3.3, poniendo de relieve los sistemas de gestión más utilizados y las propuestas más innovadoras.

4.1. ASPECTOS MÉTRICOS

El primero de los repertorios en papel analizados nos da ya las principales pautas para trazar una definición y un esquema métrico. Stenou y Knapp (1975) establece un conjunto de elementos informativos para el tratamiento específico de los aspectos métricos. Categorías silábicas y versificación son los dos apartados que, siguiendo una codificación precisa, recogen formulariamente toda esta información. La categoría silábica atiende al número de sílabas que componen los versos, y el campo más complejo de Versificación utiliza un total de nueve dígitos para informar del tipo de estrofa, rima, presencia de preludios o finidas, etc... A ello se añade el campo Tipo formal, que asigna cada composición a un determinado género formal, tomado de un listado previamente establecido y suficientemente consistente para dar cuenta de la enorme variedad formal del corpus.

Más parco es Dutton (1982), cuya finalidad es inventariar con la mayor exhaustividad, con una ampliación notable del corpus respecto a Stenou, y mostrar relaciones entre los textos, relegando la métrica a una escueta fórmula numérica, y a la transcripción literal del género o forma poética, sin afán de normalización.

En ambos casos la recuperación de esa información es complicada o nula. En Dutton la propuesta de recuperación alcanza tan solo al índice de los géneros tomados literalmente de la fuente.

Los repertorios digitales, por su parte, enriquecen, en general, el tratamiento métrico, con útiles refinamientos conceptuales. Orientados a *corpora* muy concretos, con particularidades muy específicas, descienden en muchos casos a los aspectos más especializados. Así CSM, al margen de campos métricos más genéricos, dedica un campo a la constatación de la presencia o ausencia de zéjel, y otro al refrán. También MedED utiliza un campo booleano para discriminar si una composición es "de seguir" o los versos que pertenecen al refrán. Del mismo modo BedT dedica campos particulares a 'strofe di invio', 'numero tornadas', 'lunghezza tornadas', 'tipo di tornada', 'tipo di refrain' o 'lunghezza refrain', es decir, aspectos especializados de su dominio.

En ocasiones, se complementa la información métrica con la referencia a categorías establecidas en repertorios que podríamos considerar canónicos. Es el

caso de CSM, que cumplimenta el campo 'Mettype' con un código numérico que corresponde unívocamente al esquema métrico completo tomado de (Mettmann, Walter 1986) También MedED recurre al repertorio de Tavani (1967 para identificar el número de esquema métrico, dejando ese campo en blanco, si no se registra en la obra del italiano.

Al margen de estas referencias, las fórmulas para los esquemas rimáticos y métricos suelen ser las habituales, utilizando convenciones para marcar determinados aspectos, como por ejemplo que el verso sea agudo o grave, o esté inserto en el refrán, tal como vemos en MedED:

Esquema métrico da estrofa con distinción de rimas masculinas ou agudas, e femininas ou graves (9' indica verso eneasílabo grave, fronte a 9 que sinala o eneasílabo agudo). Son exemplos: 10 10 10 10, 13' 13' 3', 7 6 7 7 7 8,....

Esquema rimático da estrofa... Os versos de refrán van indicados en maiúscula. Algúns exemplos son: abab, abbaa, abbcAA, ccc,....¹²

CSM sigue una formulación semejante, asignando campos distintos para el cuerpo de la composición y el refrán. Es frecuente también la combinación en la misma fórmula de medida y rima. Se completa la información indicando el número de estrofas total y, en varias ocasiones, un campo libre para comentar aspectos métricos; por ejemplo, "Nota a forma metrica" (BEdT)

ReMetca considera los campos habituales: número de versos y estrofas, fórmula métrica y rimática, pero indicando claramente sus convenciones notacionales, como la codificación de la rima, que resuelve la complicación que añade la presencia de estribillos o retronx:

Poemas con estribillo: el estribillo se codificará con las últimas letras del alfabeto (xyz...); cuando una rima del estribillo se use en el cuerpo de la parte glosadora, se usará la letra correspondiente a dicha rima en el estribillo; cuando haya retronx (sea total o parcial), la letra definitoria de la rima se marcará mediante asterisco precediendo a la letra¹³.

Además, añada campos derivados de las características métricas y rimáticas, como 'asonancia', 'unisonancia', 'isometrismo', e 'isoestrofismo', y el "nombre métrico", seleccionado de un listado previamente definido.

12 "Menú de ayuda contextual" en (Centro Ramón Piñeiro para a Investigación en Humanidades 2015)

13 Tomado de la página de "Instrucciones" (González-Blanco, dir. 2015).

4.2. CLASIFICACIONES Y TAXONOMÍAS

En la mayoría de los proyectos descritos se definen tablas clasificatorias y vocabularios controlados. En ocasiones están desarrollados específicamente para el propio proyecto por su equipo técnico, formado generalmente por especialistas en el corpus objeto de estudio. En otros casos se utilizan clasificaciones preexistentes, sancionadas por una larga tradición y generalmente aceptadas por los estudiosos del dominio.

Las clasificaciones pueden ser temáticas o formales. Los repertorios en papel hacen un buen uso de estos listados para indizar los materiales. (Stenou y Knapp, 1975) presenta una propuesta de sistemas clasificatorios que, por su amplitud, sigue siendo válida para los cancioneros castellanos del siglo XV¹⁴. El primer sistema establece los géneros según el contenido, y constituye un análisis en profundidad de la diversidad de materias que encierra el corpus cancioneril. Los bloques temáticos son:

- Poesía lírica.
- Poesía moral y satírica.
- Poesía lírico narrativa y lírico-dramática.
- Referencias generales.

Cada una de estas grandes divisiones se subdivide en multitud de aparatos hasta construir un sistema de clasificación exhaustivo y consistente.

Del mismo modo, el “Sistema de los géneros según la forma” refleja la diversidad formal de la poesía cancioneril. Los cinco grandes grupos en torno a los que se estructuran todas estas variedades compositivas son:

- Composiciones de forma fija.
- Composiciones estróficas con elementos de la poesía de forma fija.
- Composiciones con estrofas independientes (en cuanto a la rima).
- Composiciones de estrofa única.
- Composiciones no-estróficas.

Más especializada, limitada a su corpus temático, es la propuesta de CSM, que recurre al clásico listado de Poncelet (1902), volcado en una tabla de la BD, para identificar los milagros de la Virgen narrados en las cantigas. Otro listado de palabras claves permite la recuperación por temas muy concretos.

MedDB, que opera también sobre un corpus muy establecido, propone varios esquemas de clasificación: Tipo de Cantiga, que adscribe el corpus a cuatro apartados, incluido NA 'No aplicable'; clasificación genérica que determina los géneros y las modalidades genéricas de las cantigas, términos familiares para el investigador de este tipo de materiales; y un grupo más abierto que incluye temas y subtemas.

14 Dutton (1982, II: 223-226) se limita a recoger literalmente de la fuente los epígrafes que aluden al género.

Naetebus utiliza como tablas clasificatorias la de "Genres et classes" y "Caractère thématique". Por lo que se refiere a ReMetCa, añade a la clasificación temática y genérica el índice HisMetCa, que además de proponer una clasificación temática consistente, permite poner en relación los materiales analizados con los estudios recogidos en Gómez Redondo (2015)

En conclusión, diversidad de los materiales tratados por los distintos repertorios y los intereses del proyecto y de sus equipos investigadores tiene como consecuencia una disparidad clasificatoria, si bien hay al menos tres ejes que organizan estas clasificaciones: aspectos formales, aspectos de contenido, y vinculación con recursos externos y listas canónicas.

4.3. TECNOLOGÍA

Los repertorios en papel analizados son, en su mayor parte, salidas impresas de un ordenador con un sofisticado tratamiento. En concreto el desarrollo técnico llevado a cabo para la elaboración del *Catálogo-Índice...* de Dutton (1982) está perfectamente documentado en Anexo I, por su responsable de programación, Santoyo Vázquez (1982), del Computer Services Office (CSO) de la Universidad de Illinois. Describe los dos ordenadores utilizados: el primero, dotado con sistema operativo NOS, se ocuparía del análisis de los datos, y el segundo, equipado con UNIX, se destinaría a la fotoimpresión. Por lo que se refiere a software, para la edición de los textos se utilizó un procesador propio desarrollado por CSO. La indexación se ejecutó con el programa SORT/MERGE, y el análisis con SPITBOL 6000. Este software fue desarrollado en la misma Universidad, en concreto en el Illinois Institute of Technology, por Robert Dewar y Ken Belcher¹⁵.

En cuanto a la edición de los datos, para la identificación de los elementos informativos se utilizaron una serie de convenciones gráficas, que difieren poco de un sistema de marcación actual. Cada uno de los campos, hasta un total de 14, se identificó con un número entre corchetes triangulares, que precede a los datos, y cada registro está separado por el símbolo dólar "\$". Por motivos de disponibilidad de almacenamiento de los discos magnéticos, se crearon un total de siete archivos base para los manuscritos y cuatro para los impresos, a partir de los cuales se creó el Archivo maestro, en el que se reunificaron las entradas de las distintas versiones de un poema en un único registro. El Archivo maestro fue revisado manualmente por los investigadores, y, una vez depurado, fue la base para la extracción de los índices.

Pero nos interesa más el examen de las tecnologías que entran en juego en la programación y diseño de los repertorios métricos digitales. Dominan las bases de

15 A un proceso semejante, utilizando el texto digitalizado del proyecto anterior, fue sometido el proyecto de Gómez-Bravo (1998), realizado en Cornell University.

datos relacionales, que mayoritariamente están implementadas en MySQL, sistema de gestión de libre distribución. Pero también están presentes otros sistemas, como Oracle Database, que es el seleccionado por el proyecto MedDB, o PostgreSQL, opción de Naetebus, y Access, que es el RDBMS de BEdT. En cualquier caso, sean de distribución libre o propietarios, la opción mayoritaria por los RDBMS es la conclusión esperable

El recorrido del esquema relacional y sus robustos sistemas de gestión en todos los sectores y actividades de la economía es extenso. Desde los años 70, con una fundamentación teórica sólida y con posibilidades de implementación física altamente fiables, su predominio en el almacenamiento y recuperación de datos es incontestable. En los últimos años, sin embargo, está siendo objeto de una doble competencia. Por una parte las bases de datos NoSQL, que se proponen como la mejor opción para grandes volúmenes de datos no estructurados, y, por otra, el lenguaje de marcación XML, cuyos archivos son gestionados por bases de datos mixtas -con campos tipo XML-, o nativas.

En efecto, la opción XML está ganando terreno en el ámbito que nos ocupa, bien para dar persistencia a la totalidad de los datos del proyecto, o bien para codificar datos concretos. Se ha discutido mucho sobre los beneficios e inconvenientes de cada una de las propuestas de acuerdo con la estructura de los dominios de aplicación. En IBM (2005) se establecen "best practices" para seleccionar uno u otro modelo. En ellas se recomienda la elección de XML cuando la aplicación tiene multitud de atributos, y se dan de forma irregular en sus distintas instancias. Llevado a nuestro terreno, piénsese, por ejemplo, en la amplia casuística que ofrece la fijación de un texto según tenga uno o varios testimonios, y según sean más o menos laxos los criterios ecdóticos establecidos. Tendríamos textos codificados con un subconjunto de poco más de una docena de etiquetas, y otros que necesitarían cientos de etiquetas para dar cuenta de la huella que la transmisión de una obra deja en sus testimonios. En un esquema relacional, necesitaríamos una tabla con la totalidad de los atributos disponibles, y de muchos de estos atributos se derivarían tablas relacionadas. La recuperación de esta información a través de una aplicación web debería realizar múltiples operaciones de combinación de tablas, para recuperar esa información, y el coste en recursos sería muy alto, o, al menos sería más costoso que la alternativa XML.

Otra de las recomendaciones es utilizar XML "when data represents business in which the component part do not make sense removed from the context of that business object". En nuestro dominio los versos son parte componente directa o bien de las estrofas, o bien de los poemas, en el caso de que estos sean monoestróficos; y los poemas, a su vez, son parte componente de un cancionero. Un verso tiene sentido dentro de esa categoría superior con la que mantiene no solo una relación de pertenencia, sino también de rima con los otros versos o partes componentes. La utilización de un esquema relacional

normalizado (Silberschatz, Abraham y Korth, Henry F. 2002, 161-172)¹⁶ para mostrar esa estructura exigiría por lo menos el uso de una tabla por cada parte componente, y la búsqueda incurriría también en costosas operaciones de combinación.

Los documentos XML son además autodescriptivos, es decir, contienen los datos y los metadatos que su editor haya considerado necesarios de acuerdo con un esquema previamente establecido. También es jerárquico y sensible a la ordenación secuencial. Estas tres características se avienen perfectamente con el tratamiento deseable para nuestro dominio.

En general, la naturaleza de los datos a tratar es determinante a la hora de seleccionar uno u otro lenguaje. La primera oposición se establece entre condición tabular o jerárquica de los datos. En el primer caso, es aconsejable el modelo relacional y, en el segundo, el modelo jerárquico. Trasladado a nuestro dominio, nos encontramos con las dos estructuras. Hay, por un lado, una serie de metadatos representables fácilmente en tablas, como pueden ser la autoría y atribuciones, el título, o las entradas correspondientes a una clasificación. Por otro lado, observamos en el análisis del propio texto una estructura jerárquica clara: Un poema está compuesto de estrofas, y las estrofas están compuestas de versos. Jerarquizando y anidando estos componentes se muestra la estructura del poema, y estas posibilidades estructuración nos las ofrece XML, y, en particular, el modelo de de datos TEI. Sin ahondar mucho más, parece claro que estamos ante un modelo mixto.

En el proyecto CSM se apunta la solución mixta como *desideratum* para futuras versiones, anunciando que un campo de la base de datos, de tipo XML, incluirá la edición del texto marcado con TEI, en sustitución de los actuales archivos PDF vinculados al registro. La pretensión de CSM con esta mejora parece ser la de ofrecer un texto crítico con las posibilidades de anotación filológica que ofrece TEI («TEI: Text Encoding Initiative» 2014), frente a un texto plano sin información metatextual.

ReMetCa (González-Blanco y Rodríguez, 2014)¹⁷, por su parte, implementa este modelo introduciendo la codificación XML en un campo XML Type de MySQL. Finalmente, la opción de DMED, que utiliza el sistema de edición y publicación TextGrid, tiene detrás una base de datos nativa XML, en concreto ExistDB («eXistdb - The Open Source Native XML Database» 2015).

En conclusión, el empuje de TEI como estándar descriptivo en XML en el ámbito de las humanidades es probable que determine la balanza, en cuanto a opción tecnológica,

16 En la obra citada puede verse una caracterización de las distintas formas normales de tablas en un esquema relacional.

17 En este trabajo se analiza el modelo de datos mixto, y se muestran aspectos de su implementación en Oracle Database, si bien en el sistema en producción se usa MySQL, con menos posibilidades de explotación de archivos XML.

hacia las soluciones XML, bien de tipo mixto, bien en su totalidad, gracias a las bases de datos nativas XML y a la disponibilidad de sistemas, como TextGrid, que hacen transparente al usuario final la arquitectura tecnológica.

REFERENCIAS BIBLIOGRÁFICAS

- Dutton, Brian (1982). *Catálogo-Índice de la Poesía Cancioneril del Siglo XV*. Madison: The Hispanic Seminary of Medieval Studies, Ltd.
- Escribano, Juanjo (2014). "PoeMetCa | Recursos digitales para el estudio de la Poesía Medieval Castellana". <http://poemetca.lindh.es/> [Consultado el 03/09/2015].
- "eXistdb - The Open Source Native XML Database" (2015). <http://exist-db.org/exist/apps/homepage/index.html> [Consultado el 03/09/2015].
- Frank, István (1953). *Répertoire métrique de la poésie des troubadours*. Paris: Champion.
- Gómez-Bravo, Ana María (1998). *Repertorio métrico de la poesía cancioneril del siglo XV*. Alcalá de Henares: Universidad.
- González-Blanco, Elena y Rodríguez, José Luis. (2014). "ReMetCa: A Proposal for Integrating RDBMS and TEI-Verse". *Journal of the Text Encoding Initiative* 8. <https://jtei.revues.org/1274> [Consultado el 03/09/2015].
- González-Blanco, Elena, y Sélaf, Levente (2013). "Megarep: A comprehensive research tool in medieval and renaissance poetic and metrical repertoires". En *Humanitats a la xarxa: món medieval / Humanities on the web: the medieval world*. Oxford, Bern, Berlin, Bruxelles, Frankfurt am Main, New York, Wien: Peter Lang.
- González Cuenca, Joaquín (1978). "Cancioneros manuscritos del Pre-renacimiento". *Revista de Literatura* 40: 107-42.
- (1980). *Cancionero de la Catedral de Segovia. Textos poéticos castellanos. Edición y facsímil*. Ciudad Real: Museo de Ciudad Real.
- IBM (2005). "Comparing XML and relational storage: A best practices guide". IBM. <http://xml.coverpages.org/IBM-XML-GC34-2497.pdf> [Consultado el 03/09/2015]
- Långfors, Arthur (1918). *Les Incipit des poèmes français antérieurs au XVIe siècle*. Paris: Champion.
- Maguire, Fiona (2013). "Cancionero Virtual. An Electronic Library of 15th century Castilian Cancionero Manuscript Corpues". En *Actas del Seminario Internacional sobre Bibliotecas Digitales y Bases de Datos especializadas para la investigación de las Literaturas Hispánicas*, Mercedes Fernández Valladares (dir.), 156-195. Madrid: IUMP (UCM).
- Mölk, Ulrich & Wolfzettel, Friedrich (1972). *Répertoire métrique de la poésie lyrique française des origines à 1350*. München: W. Fink.
- Naetebus, Gotthold (1891). *Die nicht-lyrischen Strophenformen des Altfranzösischen*. Leipzig: Hirzel.

- Pillet, Alfred y Cartens, Henry (1933). *Bibliographie der Troubadours von Dr. A. P. ergänzt, weitergeführt und herausgegeben von Dr. H. C. Halle*: Niemeyer.
- Poncelet, Albert (1902). "Index miraculorum B.V. Mariae quae saec. VI-XV latine conscripta sunt". *Analecta Bollandiana* 21, 242-360.
- Silberschatz, Abraham y Korth, Henry F. (2002). *Fundamentos de bases de datos*. Madrid: McGraw-Hill.
- Stenou, Jacqueline y Knapp, Lothar (1975). *Bibliografía de los cancioneros castellanos del siglo XV y repertorio de sus géneros poéticos*, vol. I. Paris: Éditions du Centre National de la Recherche Scientifique.
- Tavani, Giuseppe (1967). *Repertorio metrico della lirica galego-portoghese*. Roma: Edizioni dell'Ateneo.
- "TEI: Text Encoding Initiative" (2014). <http://www.tei-c.org/index.xml>. [Consultado el 03/09/2015].
- "The Project TextGrid – TextGrid" (2015). <https://textgrid.de/en/projekt> [Consultado el 03/09/2015].

REPERTORIOS DIGITALES

- BEdT = Asperti, S. BedT. Bibliografía Elettronica dei Trovatori. <http://www.bedt.it>
- CRM = Stella, F. Corpus Rhythmorum Musicum. <http://www.corimu.unisi.it/>
- CSM = Parkinson, S. The Oxford Cantigas de Santa Maria Database. <http://csm.mml.ox.ac.uk/>
- DMED = Río Riande, Gimena del. Diálogo Medieval. Edición Digital. http://sade.textgrid.de/exist/apps/SADE/Dialogo_Medieval/index.html
- MedDB = Brea, M. MedDB: Base de datos da Lírica profana galego-portuguesa. <http://www.cirp.es/bdo/med/meddb.html>
- NAETEBUS = Seláf, L. Le Nouveau Naetebus. Répertoire des poèmes strophiques non-lyriques en langue française d'avant 1400. <http://www.nouveaunaetebus.elte.hu>
- PEDECERTO = Colombi, Emanuela (dir.). Pede Certo. Metrica latina digitale. <http://www.pedecerto.eu/>
- ReMetCa = González-Blanco, E. ReMetCa: Repertorio Digital de Poesía Medieval Castellana. <http://www.uned.es/remetca>
- RPHA = Horváth; I. Répertoire de la poésie hongroise ancienne. <http://rpha.elte.hu/>
- SKALDIC = Clunies Ross, M. SKALD. Skaldic Project. Medieval Norse-Icelandic skaldic poetry. <https://www.abdn.ac.uk/skaldic/m.php?p=skaldic>.

Recibido el 14 de octubre de 2015.

Aceptado el 30 de octubre de 2015.