

PREFACIO

Helena Bermúdez Sabel y Elena González-Blanco García

Université de Lausanne

IE y Coverwallet

helena.bermudezsabel@unil.ch

egonzalezblancog@faculty.ie.edu

1. La razón de ser de este monográfico

La necesidad de intercambiar información en las diferentes lenguas y culturas ha hecho necesario crear estándares internacionales en la mayoría de campos de investigación. Este fenómeno ha sido especialmente importante en las disciplinas científicas, con estándares ampliamente adoptados e incluso regulaciones para el intercambio de datos. En este sentido, las humanidades han tenido una evolución más independiente, ya que hay factores importantes como la historia, la creatividad o la identidad propia, que influyen en cada tradición con diferentes resultados y sin contemplar esa necesidad de estandarización que hemos observado en otras disciplinas. El caso de la poesía es especialmente significativo, ya que cada país, tradición filológica y género literario ha seguido una evolución independiente y particular. Como resultado de ello, el acceso en línea a las colecciones de poesía está muy fragmentado.

El objetivo de este volumen es poner en relevancia este problema, recogiendo el caso de algunos de estos testimonios significativos a través de la creación de repertorios en diferentes lenguas. El hecho de que sea la *Revista de Poética Medieval* la que dedique un volumen monográfico a este tema resulta especialmente relevante, pues, como se verá a lo largo de estas páginas, la Edad Media es el período en el que se centran la mayor parte de los repertorios digitales, además de agrupar el origen de las problemáticas filológicas que se irán resolviendo en cada una de las diferentes tradiciones

de manera divergente. El hecho de abordar esta aproximación tecnológica y digital partiendo de este núcleo ayuda a entender la propia evolución de los estudios de métrica y prosodia en las diferentes tradiciones, y a su vez a entender la complejidad de un problema como la incomunicación interlingüística e intercultural que se produce en el mundo poético a lo largo de los años. Sin embargo, gracias a la tecnología y a la inclusión de los *big data* y de los algoritmos podemos hoy día enfrentarnos a esta particularidad y abordarla desde un punto de vista digital, haciendo que los propios datos en los que los ordenadores convierten a los textos y a los corpus, nos permitan volver a hacer estas tradiciones «interoperables». De este nuevo reto, que se aborda desde el marco de las Humanidades Digitales, trataremos en las próximas páginas, que recogen tan solo unas muestras digitales de la variedad, riqueza y singularidad de la poesía medieval europea a través de sus diferentes evoluciones. Esta es la perspectiva que se aborda en el proyecto europeo de investigación *Poetry Standardization and Linked Open Data* (POSTDATA)¹, en cuyo marco se desarrolla el contenido de este monográfico, y cuyo origen, como detallaremos más adelante, es justamente la idiosincrasia que reflejan los artículos recogidos en este volumen.

2. Introducción a los repertorios digitales de poesía medieval. Una perspectiva histórica

La literatura europea occidental es, desde los primeros testimonios existentes en lenguas romances, un conjunto de géneros y formas que adivinamos que circulaban con fluidez entre las diferentes lenguas. Este flujo se ve confirmado por la presencia de manifestaciones poéticas compartidas en las diferentes literaturas, acompañadas de géneros y melodías similares. Analizar todos estos fenómenos desde una perspectiva comparativa arroja importantes resultados para el estudio de los orígenes de la cultura europea y sus tradiciones literarias. Sin embargo, resulta muy complejo realizar un estudio interlingüístico a través de las diferentes tradiciones debido a la dispar organización de la investigación y a la diversidad bibliográfica, amén de las distancias geográficas, que han impedido a los investigadores la recolección de todos los datos, así como la investigación fuera de sus propias escuelas literarias. El esfuerzo por recopilar datos con un espíritu enciclopédico es el origen de los repertorios poéticos. Su historia puede dividirse en tres períodos significativos. La primera etapa coincide con el Positivismo (finales del siglo XIX), con

¹ERC Starting Grant. Grant Agreement nº: 679528. Enlace: <<http://www.postdata.linhd.es>> [consulta: 21/10/2019].

las obras de Gaston Raynaud², Gotthold Naetebus³ y Pillet y Carstens⁴, entre otras. La segunda etapa comienza después de la Segunda Guerra Mundial, con la obra de István Frank sobre la poesía de los trovadores provenzales⁵, y continúa durante mucho tiempo con las ediciones de repertorios métricos impresos en distintas lenguas: en francés antiguo (Mölk y Wolfzettel⁶), en italiano (Solimena⁷, Antonelli⁸, Zenari⁹, Pagnotta¹⁰ y Gorni¹¹), en las lenguas ibero-románicas (Tavani¹², Parramon i Blasco¹³, Gómez Bravo¹⁴ y Betti¹⁵) y en alemán¹⁶. Los avances tecnológicos promovidos por el desarrollo de la web permitieron crear una tercera generación de repertorios, en la que el tiempo de investigación se reduce considerablemente. El primer repertorio poético digital en línea fue el *RPHA: Répertoire de la Poésie Hongroise Ancienne jusqu'à 1600*¹⁷. Por otra parte, investigadores gallegos crearon *MedDB: Base de datos da Lírica profana galego-portuguesa*¹⁸; investigadores italianos también trabajaron la digitalización en *BEDT: Bibliografia Elettronica*

² Gaston Raynaud, *Bibliographie des chansonniers français des XIII^e et XIV^e siècles, comprenant la description de tous les manuscrits, la table des chansons classées par ordre alphabétique de rimes et la liste des trouvères*, Paris, F. Vieweg, 1884.

³ Gotthold Naetebus, *Die nicht-lyrischen Strophenformen des Altfranzösischen*, Leipzig, Hirzel, 1891.

⁴ Alfred Pillet y Henry Carstens, *Bibliographie der Troubadours*, Halle a.S., M. Niemeyer, 1933.

⁵ István Frank, *Répertoire métrique de la poésie des troubadours*, Paris, Edouard Champion, 1953.

⁶ Ulrich Mölk y Friedrich Wolfzettel, *Répertoire métrique de la poésie lyrique française des origines à 1350*, Munich, W. Fink, 1972.

⁷ Adriana Solimena, *Repertorio metrico dello Stil Novo*, Roma, Società filologica romana, 1980; Id., *Repertorio metrico dei poeti siculo-toscani*, Palermo, Centro di Studi Filologici e Linguistici Siciliani, 2000.

⁸ Roberto Antonelli, *Repertorio metrico della scuola poetica siciliana*, Palermo, Centro di studi filologici e linguistici siciliani, 1984.

⁹ Massimo Zenari, *Repertorio metrico dei Rerum vulgarium fragmenta di Francesco Petrarca*, Padua, Antenore, 1999.

¹⁰ Linda Pagnotta, *Repertorio metrico della ballata italiana: secoli XIII e XIV*, Milán, R. Ricciardi, 1995.

¹¹ Guglielmo Gorni, *Repertorio metrico della canzone italiana dalle origini al Cinquecento*, Florencia, F. Cesati, 2008.

¹² Giuseppe Tavani, *Repertorio Metrico della lirica galego-portoghese*, Roma, Edizioni dell'Ateneo, 1967.

¹³ Jordi Parramon i Blasco, *Repertori mètric de la poesia catalana medieval*, Barcelona, Curial Edicions Catalanes, 1992.

¹⁴ Ana María Gómez Bravo, *Repertorio métrico de la poesía cancioneril del siglo XV*, Madrid, Universidad de Alcalá de Henares, 1998.

¹⁵ María Pía Betti, *Repertorio metrico delle Cantigas de Santa María di Alfonso X di Castiglia*, Pisa, Ospedaletto Pacini, 2005.

¹⁶ Anton Heinrich Touber, *Deutsche Strophenformen des Mittelalters*, Stuttgart, J.B. Metzler, 1975; Brunner, Horst, Burghart Wachinger y Eva Klesatschke, *Repertorium der Sangsprüche und Meisterlieder des 12. bis 18. Jahrhunderts*, Tübingen, Halle a.S., M. Niemeyer, 1986-2007.

¹⁷ Iván Horváth *et al.*, *Répertoire de la poésie hongroise ancienne (1991-2014)*. Enlace: <<http://rpha.elte.hu/>> [consulta: 21/10/2019].

¹⁸ Mercedes Brea *et al.*, *MedDB: Base de datos da lírica profana galego-portuguesa (1994-2014)*. Enlace: <<http://www.cirp.es/bdo/med/meddb.html>> [consulta: 21/10/2019].

*dei Trovatori*¹⁹; más tarde aparecieron el *Nouveau Naetebus*²⁰, la base de datos *Oxford Cantigas de Santa María*²¹, la *Analecta Hymnica Digitalia*²², la base de datos de canciones holandesas²³, el *Corpus Rhythmorum Musicum*²⁴, el *Repertorio della tradizione italiana dai Siciliani a Petrarca*²⁵, el *Digital Index of Middle English Verse*²⁶, la balada inglesa *Broadside Ballad*²⁷, *The Last Song of the Troubadours*²⁸, el *Czech Versification Research Group*²⁹, y finalmente *ReMetCa: Repertorio Digital de Poesía Española Medieval*³⁰ que, bajo la dirección de Elena González-Blanco desde 2011, es la última pieza de este poético rompecabezas multilingüe de Babel.

La interoperabilidad entre repertorios poéticos no es sencilla, ya que no solo se trata de cuestiones técnicas, sino también de problemas conceptuales y terminológicos: cada repertorio pertenece a su propia tradición poética, y cada tradición ha desarrollado una terminología analítica idiosincrásica de manera independiente durante años. El resultado de esta descoordinación es un conjunto de terminologías variadas para explicar fenómenos métricos análogos en cada uno de los sistemas poéticos, cuyas correspondencias apenas se han estudiado.

Desde el punto de vista del análisis literario, los estudios sobre métricas y poesía se vincularon en un primer momento con la gramática y la retórica, y no se consideraron independientes como *ars metrica* o *ars poetica* hasta el siglo

¹⁹ Stefano Asperti et al., *BedT: Bibliografia Elettronica dei Trovatori*. Enlace: <www.bedt.it> [consulta: 21/10/2019].

²⁰ Levente Seláf, *Le Nouveau Naetebus. Répertoire des poèmes strophiques non-lyriques en langue française d'avant 1400*. Enlace: <<http://nouveaunaetebus.elte.hu/>> [consulta: 21/10/2019].

²¹ Stephen Parkinson et al., *The Oxford Cantigas de Santa Maria Database*. Enlace: <<http://csm.mml.ox.ac.uk/>> [consulta: 21/10/2019].

²² Erwin Rauner, *Analecta Hymnica Medii Aevi Digitalia*. Enlace: <http://webserver.erwin-rauner.de/crophius/Analecta_conspectus.htm> [consulta: 21/10/2019].

²³ Louis Grijp et al., *Dutch Song Database*. Enlace: <<http://www.liederenbank.nl/index.php?lan=en>> [consulta: 21/10/2019].

²⁴ Francesco Stella, *Corpus Rhythmorum Musicum*. Enlace: <<http://www.corimu.unisi.it>> [consulta: 21/10/2019].

²⁵ Lino Leonardi, *Repertorio della tradizione poetica italiana dai Siciliani a Petrarca*. Enlace: <<http://www.mirabileweb.it>> [consulta: 21/10/2019].

²⁶ Linne R. Mooney, *Digital Index of Middle English Verse*. Enlace: <<http://www.dimev.net/>> [consulta: 21/10/2019].

²⁷ Patricia Fumerton, *English Broadside Ballad*. Enlace: <<http://ebba.english.ucsb.edu/>> [consulta: 21/10/2019].

²⁸ Anna Alberni, *The Last Song of the Troubadours*. Enlace: <<http://www.lastsongtroubadours.eu/>> [consulta: 21/10/2019].

²⁹ Petr Plecháč, *Czech Versification Research Group*. Enlace: <<http://www.versologie.cz/en/index.html>> [consulta: 21/10/2019].

³⁰ Elena González-Blanco et al., *ReMetCa: A Digital Repertoire on Medieval Spanish Metrics* (2013). Enlace: <<http://poemetca.linhd.uned.es>> [consulta: 21/10/2019].

XIV, como señala Gómez Redondo³¹. Durante la Edad Media y el Renacimiento, la poderosa influencia del latín como lengua vehicular de la cultura, hizo que los estudiosos heredaran la terminología de los tratados clásicos y la aplicaran a las lenguas romances, independientemente de su forma de estructuración. Pero cuando empiezan a surgir teorías vernáculas, cada tradición literaria establece su propia terminología y sistema de clasificación. Esta multiplicidad conduce a situaciones paradójicas y complejas, como la creación de géneros conceptuales que solo existen en algunas escuelas literarias. Un claro ejemplo de ello es el llamado «Mester de clerecía», considerado exclusivamente por nuestra tradición académica como una creación española, pero, como se ha demostrado en los trabajos de González-Blanco³², es sólo una variación de la estrofa alejandrina francesa y latina del cuarto pliegue, heredada de la poesía goliárdica latina medieval. Incluso el término «alejandrino» sirve para denominar el verso de 14 sílabas en la métrica española, pero es un dodecasílabo en la tradición francesa, ya que esta lengua no cuenta las sílabas después del último acento. Esto podría ser lógico, pero aplicando esta misma teoría, el octosílabo francés sería igual al verso español de 9 sílabas (como sucede con algunos de los primeros poemas octosílabos en español), por lo que nos enfrentamos una y otra vez a incoherencias similares al intentar realizar estudios comparativos. Las escuelas lingüísticas del siglo xx –especialmente el Estructuralismo y el Generativismo– también han tratado de instaurar sus propios principios al análisis métrico. Por tanto, el trabajo académico actual de análisis de esquemas rítmicos se basa en modelos abstractos de estas dos escuelas, lo que dificulta el intercambio de información entre ambas líneas de investigación.

3. El proyecto POSTDATA como marco de acceso digital a las colecciones de poesía y a los repertorios medievales

El proyecto POSTDATA es, a día de hoy, uno de los proyectos que ha contado con mayor financiación de la Comisión Europea para el análisis poético dentro de su programa de excelencia ERC. POSTDATA nace con el objetivo

³¹ Fernando Gómez Redondo, *Artes Poéticas Medievales*, Madrid, Laberinto, 2001.

³² Elena González-Blanco, «Las raíces del *Mester de Clerecía*», en *Revista de Filología Española*, 88 (2008), pp. 195-207; Id., «El exordio de los poemas romances en *cuaderna vía*. Nuevas claves para contextualizar la segunda estrofa del *Alexandre*», en *Revista de Poética Medieval*, 22 (2009), pp. 23-84; Id., *La cuaderna vía española en su marco panrománico*, Madrid, FUE, 2010; Id., «Heterodoxia en la *cuaderna vía*: Nueva revisión del concepto de las “sílabas contadas” a la luz de los poemas franceses e italianos», en *eHumanista*, 18 (2011), pp. 66-93; Id., «Los orígenes de la *cuaderna vía*: entre los himnos mozárabes y la poesía goliárdica», en *Estudios de Latín Medieval Hispánico*, ed. de J. Martínez *et al.*, Florencia, Sismel-Edizioni del Galluzzo, 2011, pp. 755-764.

de acortar la brecha digital existente entre la poesía y la tecnología, desarrollando soluciones de interoperabilidad y fomentando un trabajo interdisciplinar que proporciona resultados que están a la vanguardia de lo realizado hasta el momento. El proyecto se basa en la aplicación de tecnologías de web semántica para enlazar y publicar conjuntos de datos poéticos de forma estructurada, en una nube de datos enlazada. Esta labor de convertir la poesía en datos legibles estructurados permitirá la posibilidad de vincular, indexar y extraer nueva información a través de la combinación de los diferentes conjuntos de datos. La adición de una capa «semántica» de datos a las diferentes bases de datos y recursos digitales existentes dedicados a la poesía potenciará la interoperabilidad entre ellos y permitirá a los investigadores desarrollar estudios comparativos innovadores, que antes no era posible llevar a cabo.

La estandarización es la clave para lograr este objetivo, tanto en el ámbito filológico como en el tecnológico. La simbiosis de estos dos mundos permitirá a los usuarios analizar, clasificar y publicar ediciones digitales de poemas, primero para corpus en español, francés e italiano y más adelante para corpus en lenguas adicionales. El desarrollo tecnológico refleja los cambios experimentados por las humanidades digitales en los últimos años. Dependiendo del enfoque y la naturaleza de los proyectos, se han utilizado diferentes tecnologías. Las más populares son las bases de datos SQL, los lenguajes de marcado TEI y XML, los estándares de tecnología de web semántica (RDF, OWL, SKOS), los sistemas de procesamiento de lenguaje natural (NLP) y las herramientas de visualización.

4. Modelo de datos

4.1 Metodología

Como paso previo a la creación de una red de ontologías de la poesía europea, el equipo de POSTDATA ha creado un modelo de datos a partir del análisis de diferentes repertorios poéticos. El objetivo era extraer los diferentes conceptos que los proyectos examinados utilizan (y necesitan) para describir su objeto de estudio.

En un primer momento, POSTDATA contactó con representantes de diferentes proyectos a los que se les solicitó la estructura interna de sus bases de datos, así como cualquier otra documentación adicional que facilitase el análisis de su proyecto. Aplicando un método de ingeniería inversa, se utilizaron técnicas de ingeniería de software para reconstruir los modelos conceptuales de estas bases de datos³³.

³³ La metodología se describe en detalle en Mariana Curado Malta, Paloma Centenera y Elena

En total se analizaron veinticuatro repertorios, veintitrés de los cuales estaban disponibles en la web. Estos repertorios se han utilizado en momentos diferentes del proceso de definición (y de validación) del modelo de datos común. Para extraer y comparar todos los conceptos de cada uno de los modelos de datos de dichos proyectos, el equipo de POSTDATA elaboró modelos conceptuales de once de los repertorios:

- *Corpus of Spanish Golden-Age Sonnets*³⁴
- *Corpus Rhythmorum Musicum*³⁵
- *Digital Edition of the Index of Middle English Verses*³⁶
- *Kalevala poetry*³⁷
- *Lyrik des hohen Mittelalters*³⁸
- *MedDB - Base de Datos da Lírica profana galego-portuguesa*³⁹
- *Métrique en Ligne*⁴⁰
- *The Oxford Cantigas de Santa Maria Database*⁴¹
- *Répertoire de la poésie hongroise ancienne*⁴²
- *Repertorio Métrico Digital de la Poesía Castellana*⁴³
- *Corpus of Czech Verse (Versologie)*⁴⁴

De este estudio inicial se elaboró una primera versión del modelo de datos común de la poesía europea, que fue presentando en un taller en Madrid en

González-Blanco «Using Reverse Engineering to Define a Domain Model: The Case of the Development of a Metadata Application Profile for European Poetry», en *Developing Metadata Application Profiles*, ed. de M. C. Malta, A. A. Baptista y P. Walk, Hershey (Pennsylvania), IGI Global, 2017, pp. 146–180. Véase también Helena Bermúdez Sabel, Mariana Curado Malta y Elena González-Blanco García, «Towards Interoperability in the European Poetry Community: The Standardization of Philological Concepts», en *Language, Data, and Knowledge: First International Conference, LDK 2017, Galway, Ireland, June 19-20, 2017, Proceedings*, ed. de J. Gracia, F. Bond, J. P. McCrae, P. Buitelaar, C. Chiarcos y S. Hellmann, Cham, Springer International Publishing, 2017, pp. 156-165.

³⁴ Borja Navarro Colorado, María Ribes Lafoz y Noelia Sánchez, «Metrical annotation of a large corpus of Spanish sonnets: representation, scansion and evaluation», en *Proceedings of the 10th edition of the Language Resources and Evaluation Conference*, Portorož, 2016. Enlace: <<http://lrec2016.lrec-conf.org/en/>> [consulta: 21/10/2019].

³⁵ Stella, *ob. cit.*

³⁶ Mooney, *ob. cit.*

³⁷ AA.VV., *SKVR-tietokanta – kalevalaisten runojen verkkopalvelu*. Enlace: <<https://skvr.fi/>> [consulta: 21/10/2019].

³⁸ Manuel Braun, Sonja Glauch y Florian Kragl, *Lyrik des hohen Mittelalters* (2012-2019). Enlace: <<http://www.ldm-digital.de>> [consulta: 21/10/2019].

³⁹ Brea *et al.*, *ob. cit.*

⁴⁰ AA.VV., *Métrique en Ligne*. Enlace: <<https://crisco2.unicaen.fr/verlaine/index.php>> [consulta: 21/10/2019].

⁴¹ Parkinson, *ob. cit.*

⁴² Horváth *et al.*, *ob. cit.*

⁴³ González-Blanco *et al.*, *ReMetCa*, *ob. cit.*

⁴⁴ Plecháč, *ob. cit.*

2017. Asistieron representantes de diez de los once proyectos analizados junto con un delegado de la base de datos *Corpus des Troubadours*. Una de las actividades de dicho taller consistió en un proceso de validación de la propuesta de modelo de datos. Los representantes de los repertorios tenían que rellenar unas hojas de cálculo en las cuales se recogía el modelo de datos. Así, tenían que poblar estas hojas con la información de un recurso real presente en sus respectivas bases de datos⁴⁵. Después de este proceso de validación, se creó una segunda versión del modelo de datos. El siguiente paso consistió en añadir más elementos al análisis:

1. Estudio de seis modelos de datos adicionales:

- *Corpus des Troubadours*⁴⁶
- *Dainu Skapis*⁴⁷
- *Eighteenth-Century Poetry Archive (ECPA)*⁴⁸
- *Nouveau Naetebus*⁴⁹
- *Pedecerto*⁵⁰
- *Répertoire métrique de la poésie néotroubadouresque*⁵¹
- *Sound and Metre in Italian Narrative Verse*⁵²

2. Detección de las necesidades informativas de cinco proyectos de poesía a partir del estudio de sus sitios web:

- *Eighteenth-Century Poetry Archive (ECPA)*
- *Estonian runic songs database*⁵³
- *Nederlandse Liederenbank*⁵⁴
- *Nouveau Naetebus*
- *Skaldic Poetry of the Scandinavian Middle Ages*⁵⁵

⁴⁵ Para una explicación pormenorizada de los procesos de validación, véase Mariana Curado Malta, Helena Bermúdez-Sabel, Ana Alice Baptista y Elena González-Blanco, «Validation of a metadata application profile domain model», en *International Conference on Dublin Core and Metadata Applications* (2018), pp. 65-75.

⁴⁶ Enlace: <<https://troubadors.iec.cat/>> [consulta: 21/10/2019].

⁴⁷ AA.VV., *Dainu Skapis*, Enlace: <<http://en.lfk.lv/collection/folksong-cabinet/>> [consulta: 21/10/2019].

⁴⁸ Huber, Alexander, *Eighteenth-Century Poetry Archive* (2019). Enlace: <<https://www.eighteenth-centurypoetry.org/>> [consulta: 21/10/2019].

⁴⁹ Sélaf, *Le Nouveau Naetebus*, *ob. cit.*

⁵⁰ Emanuela Colombi, *Pedecerto*. Enlace: <www.pedecerto.eu> [consulta: 21/10/2019].

⁵¹ Proyecto de Dominique Billy quien amablemente nos cedió su dataset y análisis.

⁵² AA.VV., *Sound and Metre in Italian Narrative Verse*, 2017. Enlace: <<http://www.italianverse.reading.ac.uk/>> [consulta: 21/10/2019].

⁵³ AA.VV., *Estonian runic songs database*. Enlace: <<http://www.folklore.ee/regilaul/andmebaas/?ln=ee>> [consulta: 21/10/2019].

⁵⁴ Grijp *et al.*, *ob. cit.*

⁵⁵ AA.VV., *Skaldic Poetry of the Scandinavian Middle Ages*. Enlace: <<https://skaldic.abdn.ac.uk/db.php>> [consulta: 21/10/2019].

Este proceso tuvo como resultado la tercera versión del modelo de datos de POSTDATA. Fue validada escogiendo recursos de proyectos no utilizados durante la etapa de desarrollo:

- *Bardic Poetry Database*⁵⁶
- *Cantigas de Santa Maria for Singers*⁵⁷
- *The computationally assisted analysis of formal features in Russian poetry*⁵⁸

Estos repertorios representan diferentes tradiciones literarias, lenguas y culturas. Con el fin de priorizar la diversidad y garantizar la representatividad de la muestra, proyectos como *The Last Song of the Troubadours*⁵⁹, *Bibliografia Elettronica dei Trovatori*⁶⁰ o *Cantigas Medievais Galego-Portuguesas*⁶¹, que habían sido seleccionados cuando se hizo el estado de la cuestión inicial, no fueron analizados porque modelan tradiciones ya incluidas.

Uno de los criterios determinantes a la hora de seleccionar los recursos analizados fue su accesibilidad, tanto a la hora de poder acceder a su estructura interna como a la hora de poder comprender y estudiar sus contenidos. Incluso así, se han llevado a cabo grandes esfuerzos para conseguir una muestra representativa en lo que se refiere a la lengua, período de composición y sistema prosódico.

4.2. Modelado

La diversidad de los marcos teóricos, así como la heterogeneidad a la hora de formalizar las diferentes tradiciones poéticas, conlleva una serie de dificultades a la hora de definir un modelo común. Por ejemplo, existen diferentes características métrico-rimáticas a través de las cuales se caracterizan las diferentes tradiciones prosódicas. La ontología de POSTDATA tiene que dar cuenta de todas ellas, pero al mismo tiempo necesita implementarse como un modelo flexible y lo suficientemente abstracto para garantizar la

⁵⁶ Mícheál Hoynes *et al.*, *Bardic Poetry Database*. Enlace: <<https://bardic.celt.dias.ie/>> [consulta: 21/10/2019].

⁵⁷ Andrew D. Casson, *Cantigas de Santa Maria for Singers*, 2019. Enlace: <<http://www.cantigas-desantamaria.com/>> [consulta: 21/10/2019].

⁵⁸ David Birnbaum y Elise Thorsen, *The computationally assisted analysis of formal features in Russian poetry*. Enlace: <<http://poetry.obdurodon.org/>> [consulta: 21/10/2019].

⁵⁹ Alberni, *ob. cit.*

⁶⁰ Asperti *et al.*, *ob. cit.*

⁶¹ Graça Videira Lopes, Manuel Pedro Ferreira *et al.*, *Cantigas Medievais Galego Portuguesas [base de dados online]*. Lisboa, Instituto de Estudos Medievais-FCSH/NOVA, 2011-. Enlace: <<https://cantigas.fcsh.unl.pt/>> [consulta: 21/10/2019].

interoperabilidad. Si no, solo se podrían establecer comparaciones entre las tradiciones más próximas.

Dependiendo del tipo de versificación, podemos clasificar los sistemas métricos según diferentes elementos definitorios: hay tradiciones que cuentan sílabas, otras se caracterizan por la distribución de la longitud de sus vocales y otras buscan la presencia de sílabas acentuadas en intervalos regulares. El equipo de trabajo tiene que tener un profundo conocimiento de la prosodia para poder buscar los elementos en común entre diferentes tradiciones y, al mismo tiempo, poder modelar las características únicas de cada una de ellas. Es habitual clasificar la poesía europea/occidental en términos de verso cualitativo o cuantitativo, es decir, categorizar una composición atendiendo a si es la cantidad silábica o el acento la característica que define el metro⁶². Sin embargo, hemos decidido combinar muchos atributos de estos dos tipos de metro, ya que la distribución de acentos y la distribución de sílabas largas son equivalentes si dejamos a un lado la caracterización lingüística y hablamos en su lugar de sílabas en posiciones fuertes y débiles. Por lo tanto, nos hemos centrado en crear propiedades que permitan diferenciar, dentro del metro cualitativo, si los acentos del verso necesitan ser regulares o si solamente la posición de una sílaba determinada (por ejemplo, la última) tiene que ser fija. Así, hablamos de esquemas de patrones (*patterning scheme*) para definir la distribución de sílabas fuertes y débiles y esquemas métricos de cómputo (*counting scheme*) para aquellos que cuentan en qué posición se encuentra la última sílaba acentuada del hemistiquio o verso. Este sistema permite que un verso trocaico cuantitativo sea interoperable con uno acento-silábico ya que siguen el mismo tipo de esquema.

Algunas de las dificultades de modelado no son específicas del género poético, sino que vienen dadas por las peculiaridades de la transmisión de textos literarios, véase por ejemplo toda la complejidad que puede surgir del modelado de la autoría.

En la mayor parte de los casos, nos encontraremos con atribuciones autoriales sencillas en las que hay uno o varios autores que crean una obra. También hay que tener en cuenta, además, la posibilidad de asignar una autoría específica a diferentes partes de una composición (pensemos, por ejemplo, en géneros dialogados como la *tençó* en la que dos trovadores alternan la autoría de las estrofas).

Otro caso especial en lo que se refiere a la autoría es el de los acrósticos como fuente de una atribución autorial. El modelo de datos no solo tiene que

⁶² José Domínguez Caparrós, *Métrica española*, Madrid, UNED, 2014, pp. 41-43.

incluir las propiedades necesarias para definir los contenidos de un acróstico, sino que tiene que formalizar que esta es la fuente de una atribución de autoría concreta. Además, también tiene que haber cabida para otros fenómenos específicos de la composición literaria: pseudónimos, heterónimos, ortónimos y alónimos.

Un modelo que se marca como objetivo la capacidad de formalizar cualquier composición poética europea tiene que ofrecer la posibilidad de codificar conceptos como la «certeza» a la hora de asignar una autoría. Al mismo tiempo, tenemos que tener presente las especificidades de determinados procesos de transmisión: como sabemos, atribuciones erróneas son relativamente frecuentes cuando se trabaja con textos enviados a través de manuscritos. Generalmente, esta atribución es realizada en un paratexto entendido como cualquier texto diferente del principal, es decir, del poema. Sin embargo, esta información puede estar presente en diferentes partes de los materiales textuales: la rúbrica de un testimonio, el índice de un códice entre otras posibilidades. Todos estos conceptos deben existir dentro del modelo de datos y deben estar correctamente formalizados para que la comunidad científica pueda describir con detalle las particularidades de su objeto de estudio.

El modelo de datos que será la base de la ontología de POSTDATA ha sido construido siguiendo un proceso metodológico riguroso, utilizando métodos y técnicas propias de la ingeniería de software con el fin de analizar las estructuras de datos y los sitios web a partir de los cuales extraer la información necesaria para la conceptualización del modelo. Los cimientos del análisis fueron veinticuatro repertorios poéticos que son una muestra que responde a la diversidad de lenguas, períodos de composición y sistemas prosódicos. Además, se han incorporado procedimientos de validación que garantizan la calidad del modelo. El siguiente paso es el de conseguir la aceptación de la comunidad científica y que esta haga uso de una ontología que permitirá responder a preguntas de investigación que antes no podían ser formuladas. Este monográfico es una vía más de difusión del contenido de esta ambiciosa propuesta con la que se espera dar visibilidad al esfuerzo coordinador que se está desempeñando. De manera especial, este número quiere dar cuenta de los recursos existentes en el ámbito digital, frecuentemente desconectados, y servir de altavoz de estos proyectos innovadores en el campo literario en general, y en el de la poética medieval en particular.

5. El contenido de este monográfico

Este volumen viene a cubrir un hueco en las publicaciones de literatura medieval comparada y también en las de Humanidades Digitales. Su propósito es recopilar algunos testimonios significativos de estos repertorios digitales que han formado parte del modelo de POSTDATA y, a su vez, dar visibilidad a los problemas que plantea a los investigadores la desconexión del mundo digital en literatura comparada, al que se suman la construcción literaria divergente en el ámbito de las diferentes métricas y tradiciones, unidas a una diferente evolución tecnológica condicionada por las tecnologías de distintas épocas y la carencia de estándares.

Dado que la mayor parte de los repertorios digitales pioneros surgieron en torno a los textos literarios y poéticos de la Edad Media, nos pareció muy pertinente utilizar el espacio de esta revista para dar un paso más hacia los estudios de literatura medieval comparada, esta vez desde la perspectiva de las Humanidades Digitales y del ámbito de la estandarización y la interoperabilidad o comunicación entre los diferentes recursos.

Así, este número recoge, por una parte, una reflexión teórico-metodológica sobre los corpus poéticos digitales de la mano de Levente Seláf con el artículo «Les voies de la création d'une base de données poétique générale». Por otra, se hace una exposición individualizada de diferentes repertorios digitales que representan empíricamente la heterogeneidad de la poesía europea.

Abre la sección sobre recursos digitales un artículo de Irene Volpi sobre *Corpus Rhythmorum Musicum*, un sitio web que ofrece una base de datos filológica que contiene tanto las ediciones críticas como las reproducciones musicales de las más tempranas composiciones medievales en lengua latina (siglos IV-IX).

Pasamos del latín a la tradición romance con dos contribuciones que versan sobre uno de los movimientos literarios más importantes en la península ibérica durante la Plena y Baja Edad Media: la escuela gallego-portuguesa. Concretamente, Pilar Lorenzo Gradín y Carmen de Santiago Gómez presentan un nuevo recurso para el estudio de la lírica profana, *Base de datos paleográfica da lírica galega medieval* (PalMedDB) que complementa a la pionera *MedDB: Base de datos da lírica profana galego-portuguesa*, disponible en red desde 1998. La exploración de la vertiente religiosa de esta corriente lírica es desarrollada por Stephen Parkinson que nos presenta cómo un recurso electrónico como la *Oxford Cantigas de Santa Maria Database* se convierte en una herramienta indispensable para el estudio de las *Cantigas de Santa María* debido a lo compleja y multifacética que es esta monumental obra.

Pasamos a continuación a otra tradición medieval completamente diferente, la poesía escáldica compuesta por poetas guerreros vikingos entre los siglos IX y XIV. Tarrin Wills expone en su artículo sobre el *Skaldic Project* una serie de recursos integrados que permiten un análisis holístico de esta tradición poética en nórdico antiguo.

Finalmente, cierra este número una contribución de Borja Navarro-Colorado, precursor de la aplicación de métodos computacionales al estudio de la poesía española. En este artículo, Navarro-Colorado nos presenta un corpus lírico castellano que está anotado tanto desde el punto de vista lingüístico como métrico y expone las funcionalidades de investigación que un recurso de este tipo posibilita.

Como puede deducirse de estas breves descripciones, lo que caracteriza a este volumen es la variedad y riqueza, tanto desde un punto de vista literario como tecnológico y metodológico. Se abordan desde los principios más teóricos de la construcción de los repertorios métricos hasta las últimas tendencias de la computación que van desde los lenguajes de marcado y anotación a los algoritmos de aprendizaje profundo de la inteligencia artificial. Se recoge una importante variedad geográfica, tanto de los autores, como de los propios textos y corpus analizados (desde Galicia hasta Escandinavia, pasando por Hungría e Italia), y se manifiesta en la propia lengua de redacción de los trabajos, que incluye cuatro idiomas. Si bien algunos trabajos extrapolan ligeramente los límites cronológicos del período medieval, su inclusión se debe a su relevancia como contribución metodológica dentro del panorama de los repertorios poéticos. Toda la riqueza que las próximas páginas contienen viene unida por un hilo conductor común: el esfuerzo de los investigadores a lo largo ya de varias décadas para hacer disponible a la comunidad científica un conjunto de corpus –con sus fuentes, manuscritos y recursos complementarios–, herramientas de análisis y acceso a diferentes tipos de contenidos relacionados con la poesía (la música, la hagiografía, la propia historia), desde una interfaz digital. Nuestro propósito en el número que nos ocupa es agrupar y dar visibilidad a estos importantes esfuerzos.