SUPPORTING RESEARCH WITH POETIC METADATA: THE OXFORD CANTIGAS DE SANTA MARIA DATABASE

Stephen Parkinson
University of Oxford
stephen.parkinson@mod-langs.ox.ac.uk

1. The Cantigas de Santa Maria

Alfonso X’s Cantigas de Santa Maria is a multifaceted work, whose complexity and scale are the greatest challenges to its study. Composed and compiled in the late 13th century as a personal project of Alfonso X of Castile, it comprises 419 poems in medieval Galician-Portuguese, celebrating the Blessed Virgin Mary, to whom the King professed particular devotion. Most of the poems are set to music and many are richly illustrated. The entire collection is contained in four precious manuscripts (To, T, F, E) all produced in Castile before 1283, representing different stages in the conceptualization and execution of the King’s project. The smallest manuscript To incorporates an original nucleus of 100 poems, flanked by a prologue and an epilogue, to which various appendices are added. T and F together were intended to present 400 poems with the same prologue and epilogue, but remain unfinished.


2To (the Toledo MS) = Madrid, Biblioteca Nacional de España, MS 10069; T (the códice rico) = Real Monasterio de San Lorenzo de El Escorial, T.1.1; F (the Florence MS) = Florence, Biblioteca Nazionale, Banco Rari, 20; E (the códice de los músicos) = Real Monasterio de San Lorenzo de El Escorial, B.1.2.
E, the most complete and worst executed, achieves the numerical target of 400 items (with several internal repetitions), and an appended cluster of additional pieces.

The work itself defines the principal parameters for its study, which are textual, metrical, codicological, musical, hagiographical, historical, and iconographic.

At the textual level the CSM is a collection of troubadour poetry, with its own internal genre division and metrical conventions, and an order and structure defined by the patron/poet. The poems fall into three main classes: miracle narratives or *miragres*, lyrical praise poems or *cantigas de loor*, and para-liturgical celebrations of the Feast Days of the BVM, *cantigas de festas*.

RPM, 33 (2019), pp. 77-86, ISSN: 1137-8905

---


4 A small number of poems found in To are not repeated in the more complete MSS precisely because they do not fit this classification. They include five poems for non-Marian feast days (*Festas de Jesucristo* 1-5 in To, nos. 423-427 of the standard numbering), and one celebrating Maytime (the first poem of the appendix of additional *cantigas* in To, no. 406).
A bbba A pattern. This pattern, when applied to poems of up to 50 strophes, creates opportunities for virtuoso rhyming not available to the secular lyric\(^5\).

Codicologically, it is important to recognise that at least three of the four MSS were produced in the royal scriptorium, which was also responsible for most of the scientific and historical MSS of the Alfonsine cultural project\(^6\). They are nevertheless unique in their integration of text, paratext, music and illustration, on a large scale, which involves a high degree of scriptorium organisation. Detailed codicological study has revealed the importance of calculations of page extent in the planning of the manuscripts and the distribution of scribal tasks, and the pressure of page layout on the inclusion or omission of elements such as refrains\(^7\).

Musically, the CSM are one of the largest unified bodies of musically notated medieval song, with all but four poems accompanied by music in at least one manuscript. In most cases the music is applied to the opening refrain and strophe; occasionally more than one strophe is underlaid. As the musical notation is essentially syllabic, and the texts are precisely accentual-syllabic, the relations between music and text can be clearly established.

Hagiographically, the CSM are a late manifestation of the explosion of Marian miracle literature of the 12\(^{th}\) and 13\(^{th}\) centuries\(^8\). The collection of 356 separate miracle stories contained in the four MSS is one of the largest collections of miracles of the Blessed Virgin\(^9\). Indeed, the size and variety of Alfonso’s mariale magnum is one of its main design features, along with the richness of the two fully illustrated manuscripts; it was calculated to impress both the Blessed Virgin and all those who knew of it. It contains a number of clusters of poems relating to particular shrines of the BVM, some

---


\(^2\) MSS E T and F were produced in the Royal Scriptorium in Seville. See Laura Fernández Fernández, *Arte y Ciencia en el Scriptorium de Alfonso X el Sabio*, Puerto de Santa María, Cátedra Alfonso X el Sabio – Universidad de Sevilla, 2013; Fernández, «Los manuscritos...», *art. cit.*


\(^5\) The 356 miracle stories of the CSM can be compared with the 92 of the presumed Mariale magnum BN lat 3177, the 78 of Vincent de Beauvais’s *Speculum historiale*, the 61 of John of Garland’s *Stella maris*, and the 58 *Miracles de Nostre Dame* by Gautier de Coinci.
taken from extant local sources, some combining several, some unique to the
CSM, some piously fabricated from accounts associated with other shrines or
saints\textsuperscript{10}. An essential part of this study is to connect the narratives of the CSM
to the wider universe of Marian miracles.

Historical study is important as many of the poems refer to recent events
in the history of Iberia or in the personal life of Alfonso himself, and to con-
temporary personages. Historical and prosopographical data are thus an im-
portant element of metadata.

The iconographic dimension reflects the fact that each poem included in
the \textit{códices de las historias} was to be followed by a graphic narrative in the
form of one or more pages of illustration, in the form of a full page miniature
divided into six vignettes, each one supplemented by a caption. The visual
narratives at times diverge from the textual narrative partly in emphasis but
also in content\textsuperscript{11}.

The study of any part of this complex cultural product involves deploying
knowledge of its place in the complete structure. The narratives have to be
understood both as poems and as miracle stories, constrained both by their
genre and their metrical form. The illustrations constitute a parallel narrative
whose intersection with the poetic narrative and its sources is not straight-
forward. Adjacencies between poems have to be defined in terms of their
grouping in different manuscripts\textsuperscript{12}. There is a complex interaction between
text and paratext. An interactive relational database is the only way in which
all the information relative to textual study can be deployed and integrated
with the texts themselves.

\textsuperscript{10} Stephen Parkinson, «The Miracles Came in Two by Two: Paired Narratives in the
\textit{Cantigas de Santa Maria}», in \textit{Gaude Virgo Gloriosa: Marian Miracle Literature in the Iberian
Peninsula and France in the Middle Ages}, ed. by Juan-Carlos Conde and Emma Gatland, PMHRS 69,
Cantigas de Santa Maria, Códice Rico, Ms. T-I-1, Real Biblioteca del Monasterio de San Lorenzo de El
Escorial, El Códice Rico de las Cantigas de Santa Maria}, coord. Laura Fernández Fernández and

\textsuperscript{11} This results from a workshop system in which the artists were provided with a schematic narrative
before the texts were finalised. In some cases, they elaborated the illustrations with elements not men-
tioned in the poems, sometimes taken from alternative sources, sometimes from their own experience. See
Stephen Parkinson, «Alfonso X, Miracle Collector»; Stephen Parkinson and Deirdre Jackson, «Collection,
159-172; Francisco Prado-Villar, «\textit{Iudeus sacer}: Life, Law, and Identity in the “State of Exception” Called
“Marian Miracle”»), in \textit{Judaism and Christian Art: Aesthetic Anxieties from the Catacombs to Colonialism},

\textsuperscript{12} Many scholars who contrast the first hundred poems with the complete collection forget that \textit{Canti-
gas} 1-100 of the standard edition is the first 100 poems of MS E, which do not correspond exactly to the
100 poems contained in To.
2. The Database

The Oxford *Cantigas de Santa Maria* Database\(^{13}\) was launched in 2005 as the first project of the Centre for the Study of the *Cantigas de Santa Maria*\(^{14}\). The overall aim of the Centre is to promote and support work on the *Cantigas de Santa Maria*. Its two specific projects were the database and the new critical edition of the CSM. These two projects date back to 1994, when the PI, Stephen Parkinson, was engaged in a joint enterprise to transcribe the MSS, as well as other projects of digitisation of Old Portuguese texts\(^{15}\). The final act of this collaboration was the *Cobras e Son* conference of 1994\(^{16}\), where it became apparent that close study of MSS, texts and music were leading to a renewal of CSM study, based on consideration of all MSS sources and their distinctive constitutions, allied to metrics, sources and layout. The ongoing partnership with the Universidade Nova, Lisbon, and its project for the transcription and edition of the musical sources, dates from this period.

The Centre has been clear that the database is essentially a store of structured metadata rather than a textual edition supplemented by hyperlinked information. The long term aim has always been to incorporate the texts, initially as freestanding PDF files and ultimately as linkable XML files.

3. Structure and evolution of the database

The database is the third incarnation of a research database begun in Microsoft Access, ported to the proprietary relational database R:BASE and finally installed as an online MySQL database accessed by a website. Its guiding principles are:

– full relational organisation, allowing access to all relevant information from various entry points
– form-based presentation of complex information

---

\(^{13}\) Link: <csm.mml.ox.ac.uk> [accessed: 07/11/2019].

\(^{14}\) The main funding for this period has been from the Leverhulme Trust, under Research Grant F/08 736/B from 2003 to 2005, and an Emeritus Fellowship 2015-2017, and the British Academy Small Grant SG-46903 and Research Development Grant BR100062. Substantial grants were also received from the Research Development Fund of Oxford University, and the Modern Humanities Research Association. Much of the key work on the database was undertaken by Deirdre Jackson (Leverhulme Trust Research Assistant 2003-2005, MHRA Research Associate 2005-2006), Alison Campbell and Roberto Ceolin (Research Assistants 2005-2006). David Barnett (British Academy and Leverhulme Trust Research Assistant 2013-2017) was responsible for further work on sources and on the critical edition.

\(^{15}\) Funded by the Leverhulme Trust and the Hume University Fund.

– free user access
– interactive online editing through protected editing site, to allow remote updating by centre associates
– instant availability of updated information through dynamically generated display forms
– cross-field searching

The database is structured around the POEMS table, which contains all the essential information on the texts. Each distinct poem of the collection has a separate record, listing:

– numbers in editions and manuscripts
– summary of textual narrative, location, protagonists
– description of illustrations (graphic narrative), transcription of captions
– metrical analysis
– incipit, refrain

The table provides links to other tables:

– **MS/Layout** giving details of location, extent and page layout of each witness for each poem, including the assignment of space to the rubric, the staves, the running text, the decorative miniatures of T/F and the miniatures preceding every *cantiga de loor* in E, and the text of the rubric. Each witness record is linked to the poem it represents, and all records can be browsed in MS sequence.

– **MIRACLES** giving details of Latin and vernacular miracle narratives which form part of the network of sources. For this purpose, the database includes a digitised form of the authoritative Poncelet index of Latin miracles\(^{17}\). It supplements that source by providing full listings of each miracle collection indexed by Poncelet, and an increasing number of vernacular collections not yet indexed. The database also implements an internal cross reference system devised by Poncelet, who listed groups of miracles which are either substantially the same or demonstrably similar. In the database we create a Poncelet key number which is a single reference number applied to all narratives grouped together by Poncelet or shown to be so, and a separate *similes* key which groups similar narrative. The independence of the Poncelet and miracles tables from the Poems tables allows the digital Poncelet to be used by a wider

---

hagiographical community, and for new associations to be registered between individual cantigas and the hagiographical tradition.

– **BIBLIOGRAPHY**, listing all publications indexed to a particular poem. The Bibliography table is independently browsable and searchable, to allow the conclusion and deployment of bibliography not uniquely linked to individual poems. The Bibliography incorporates Joseph Snow’s Alfon-sine bibliographies. The database also provides external links to partner databases, notably the Philobiblon catalogue of texts, manuscripts and bibliography, and Pierre Roberge’s discography, using bookmarks inserted by Roberge to facilitate hyperlinks. It also has links to the Centre’s own and to archives of edited texts and musical transcriptions.

Entry to the database is through the homepage which offers access to the main tables through colour coded tabs, as well as other documentation provided by the Centre. Each table has a presentation screen with explanatory text in English and Galician, and offering searching and browsing of that table, with links available at various points in the records.

The main «Poems» view screen integrates a large amount of data drawn from and linked to the POEMS table, to give users a digest of relevant information. Users are given on-screen options to display content in different forms, to access linked databases and texts, and to view the more detailed information contained in specific tables. The different forms of the narratives (textual, graphic, captions) can be viewed individually or as aligned parallel text.

### 4. Research issues and publications

The database is currently accessed approximately 3000 times per month. It is also accessed internally by Oxford researchers, through the editing module, which makes available additional data not displayed on the public site. It is

---


22 The Archives of texts and music are independently accessed, with some 900 downloads of PDF files per week.
used to inform the ongoing work on the critical edition of the texts, which in turn feeds back corrections of bibliographical data, metrical schemes and summaries of text and miniatures. Researchers on narrative structures and thematic elements make extensive use of the summaries and keywords. The «Poncelet» and «Miracles» tables are the first and only digitization of the Poncelet index, and are regularly consulted by hagiographical researchers. The analyses of the miniatures are used for work on iconography and text-image relations.

5. Future perspectives

The text of the new critical edition, together with palaeographical transcriptions of all witnesses, will be integrated in the database as searchable and browsable XML files, beside the current provision for downloadable PDF files of the complete edition with apparatus. An upgraded search facility will allow textual features to be cross-searched with metadata elements. The texts will also be accessible to indexing and concordancing software, allowing searching both for lexical and phrasal elements, as well as for graphic variations.

It would be highly desirable to be able to link images of the manuscript pages, including the illustrated pages of T and F, to the appropriate records of the database. Access to images is currently an unresolved issue. While a variety of online image sources are available, the only colour images cur-


24 Santiago DiSalvo, Los monjes de la Virgen: representación y reelaboración de la cultura monacal en las Cantigas de Santa María de Alfonso X, Newark, Juan de la Cuesta, 2013; Parkinson «Alfonso X, Miracle Collector» was based on source information from the database, compiled by David Barnett.


rently available are those of To provided by the BNE website, but not in a form which would allow direct linkage from website records to individual page images. Images of T and F are copyright protected in published facsimiles, and cannot yet be deployed. It is hoped that there may soon be an official colour facsimile of E which could be accessed through the site.

The «Miracles» table needs to be supplemented by information relating to other possible sources and collaterals, such catalogues of ballads, folktales and exempla. Links will be provided to future projects mapping the location of Iberian cantigas.

The published metrical handbooks by Maria Pia Betti are both insufficient and need to be supplemented by more precise information from the database29. Recent study has revealed that the full complexity of CSM metrics cannot be captured by simple metrical schemes, or even by alternative schemes and the separation of refrain and strophe schemes. A freestanding Metrics table will be deployed to allow full metrical analyses for each poem, and the listing of poems of identical or similar metrical structure.

It is hoped to provide a fully bilingual interface to the database, with all titles and content in both Galician and English. A tentative start has been made with bilingual information pages and text and miniature summaries, provided by Déborah González.

Musical editions generated by the Universidade Nova de Lisboa editing project, as well as their completed database of notational forms, will be linked to the database or incorporated in it.

6. Evolving programming environment

The online implementation of the database was initially programmed as procedural PHP code, with free combination of XML presentation of tables governed by style sheet and MySQL database calls30. The data tables are held on secure server with the frontend and editing suite on dedicated webserver, which allows for offline development of code without risk to data. Management of the servers is by the staff of the Faculty of Medieval and Modern Languages of Oxford University. The site has survived one hacking attempt, and has proved a robust and effective sharing of research data. It has proved

unauthorised colour images of T are incorporated in video montages of images, text and music on sites such as YouTube.

29 Maria Pia Betti, Rimario e lessico in rima delle ‘Cantigas de Santa Maria’ di Alfonso X di Castiglia, Pisa, Pacini, 1997; Repertorio metrico delle ‘Cantigas de Santa Maria’ di Alfonso X di Castiglia, Pisa, Pacini, 2005.

30 By Cyan Collier.
increasingly more problematic to upgrade the PHP code in line with evolving standards, and is currently being reformulated so as to separate stable XML content and updatable PHP code. This will allow the database to be maintained and updated to meet evolving standards, while expanding content such as the Bibliography table can be continually updated with new entries.

Received: 25/04/2019
Accepted: 21/08/2019

SUPPORTING RESEARCH WITH POETIC METADATA:
THE OXFORD CANTIGAS DE SANTA MARIA DATABASE

ABSTRACT: Alfonso X’s Cantigas de Santa Maria is a multifaceted work, whose complexity and scale are the greatest challenges to its study. Thus, digital resources like the Oxford Cantigas de Santa Maria Database are required in order to comprehensively explore this complex work as a whole.


APoyo a la investigación con metadatos poéticos:
la base de datos de Oxford de las Cantigas de Santa María

RESUMEN: Las Cantigas de Santa María conforman una obra polifacética cuya complejidad y escala son los retos más importantes para su estudio. Es por eso por lo que recursos digitales como la Oxford Cantigas de Santa Maria Database son necesarios para poder estudiar de una manera exhaustiva esta obra en su conjunto.


31 By Richard Buckner, Oxford University IT Services, who was also responsible for upgrades and repairs in 2011 and 2015.