# Scire vias

## Humanidades digitales y conocimiento

Fátima Díez Platas
César González-Pérez
(coords.)

digitalia humanistica

# Scire vias

## Humanidades digitales y conocimiento

Fátima Díez Platas,
César González-Pérez (coords.)

UNIVERSIDADE DA CORUÑA

*SCIRE VIAS*
HUMANIDADES DIGITALES Y CONOCIMIENTO

Fátima Díez Platas y César González-Pérez (coords.)
A Coruña, 2024
Universidade da Coruña, Servizo de Publicacións

COLECCIÓN
*Digitalia Humanistica*, nº. 2

DIRECCIÓN DA COLECCIÓN
Carlota Fernández Travieso (UDC)

COMITÉ EDITORIAL
Mar Campos Souto (Universidade de Santiago de Compostela)
Fátima Díez Platas (Universidade de Santiago de Compostela)
Ana García Serrano (Universidad Nacional de Educación a Distancia)
Sagrario López Poza (Universidade da Coruña)
Stefano Neri (Università di Verona - Italia)
Nuria Rodríguez Ortega (Universidad de Málaga)
Amelia Sanz Cabrerizo (Universidad Complutense de Madrid)
Juan Luis Suárez (Western University - Canada)

Esta obra foi revisada e avaliada por pares cegos, externos á UDC.

368 pp.
Índice: p. 4

DOI: https://doi.org/10.17979/spudc.9788497498852
ISBN: 978-84-9749-885-2

Thema: UGC AMG 1DSE-ES-LBA 3MNQV
IBIC: UGC AMG 1DSEL 3JH

EDICIÓN
Servizo de Publicacións <www.udc.gal/publicacions>

© da edición: Universidade da Coruña
© dos textos: os autores

ASESORAMENTO TÉCNICO DE DESEÑO
Cecilia Labella Lozano



Esta obra publícase baixo unha licenza Creative Commons Atribución-NonComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0)

Reservados todos os dereitos. Nin a totalidade nin parte deste libro pode reproducirse ou transmitirse por ningún procedemento electrónico ou mecánico, incluíndo fotocopia, gravación magnética ou calquera almacenamento de información e sistema de recuperación, sen o permiso previo e por escrito das persoas titulares do copyright.

# Scire vias

## Humanidades digitales y conocimiento

Fátima Díez Platas,
César González-Pérez (coords.)

# ÍNDICE

## PRÓLOGO

[7] *Scire vias*: Humanidades Digitales y Conocimiento
Fátima Diez Platas, César González-Pérez

## I. LAS VÍAS DE LA VISIÓN, LA RECONSTRUCCIÓN Y EL SENTIDO DEL PATRIMONIO

[13] Más allá de la representación tridimensional:
la arquitectura (des)aparecida del Asilo de las Mercedes
en el espacio digital
Estefanía López Salas

[45] Humanidades digitales y Arqueología:
experimentando con modelos de reconocimiento automático
de texto manuscrito en español sobre patrimonio cultural de
inicios del siglo XX
Raquel Liceras-Garrido, Alba Comino

[79] Estudio interdisciplinar de una relación festiva:
de la descripción textual a la representación visual de la
arquitectura efímera y su iconografía
Estefanía López Salas, José Julio García Arranz, Manuel Alberto Santos
Miñambres, Sagrario López Poza

[113] 3D más allá de la reconstrucción virtual:
una herramienta al servicio de la gestión integral del
patrimonio en la ermita de Sant Llorenç de la Senabra
Pablo del Fresno Bernal, Esther Travé Allepuz, Isidre Ot Padilla, José Manuel Martínez
Torrecilla, Josep Socorregut Domenech

[145] El trabajo-proyecto *Patrimonio Herido.*
Una propuesta de innovación educativa transdisciplinar
María Marcos Cobaleda, Julia García González, Antonio Jesús Santana
Guzmán, Antonio Cruces Rodríguez, Nuria Rodríguez-Ortega

## II. LAS VÍAS DEL LENGUAJE, DEL TEXTO Y DE LA LITERATURA

[181] *Mapping Chivalry:*
La evolución de los libros de caballerías castellanos en el tiempo y el espacio
Giulia Tomasi

[207] Entre la Historia y la Filología.
El proyecto CORDICan y el valor de la interdisciplinariedad en las HD a través de la documentación relativa a *dotes y arras*
Ana Viña Brito, Dolores Corbella, Roberto J. González Zalacain

[235] Las grandes bibliotecas de textos griegos y latinos ante los retos de la nueva era de edición digital. Luces y sombras
Cecilia Criado

[263] Proyecto HDATEATROUNIR.
Realidades y perspectivas para el estudio del teatro desde las HD
Teresa Santa María Fernández, Monika Dąbrowska

## III. LAS VÍAS DEL MODELADO, DE LOS DATOS Y DE LA INFORMACIÓN

[295] Identificando el sujeto del paisaje:
modelado e integración de la dimensión social para una (re) lectura de la geografía altomedieval
Sonia Medina Gordo, Esther Travé Allepuz

[325] HILAME.
De la persona al dato, del dato a la persona
Agurtzane Paz Moro, Teresa Jular Pérez-Alfaro

[347] Estrategias de recuperación de información mediante *pooling* para la construcción de colecciones de referencia desde redes sociales.
Caso de estudio durante las tensiones raciales de 2020
Patricia Martín-Rodilla, David Otero

PRÓLOGO

# *SCIRE VIAS*: HUMANIDADES DIGITALES Y CONOCIMIENTO

Fátima Díez Platas
Universidade de Santiago de Compostela
César González-Pérez
Instituto de Ciencias del Patrimonio. Consejo Superior de Investigaciones Científicas

> "Se debe por tanto aspirar al conocimiento,
> adquirirlo y usarlo convenientemente,
> pues por medio de él tendremos todo en buen orden"
> Aristóteles, *Protréptico*, 8.

Es tarea difícil mostrar la cohesión y el sentido de una monografía sobre Humanidades Digitales con una serie variada de contribuciones no solo de temas diferentes, sino con distinto uso de las herramientas digitales y con distintos objetivos metodológicos. Sin embargo, el hilo conductor de este volumen, que procede en última instancia de la reunión de una selección de contribuciones al V Congreso de la Sociedad de Humanidades Digitales Hispánicas celebrado en el año 2021 en modalidad *online*, se encuentra en la búsqueda de diferentes vías para alcanzar conocimientos nuevos. Como sucede con el resto de las disciplinas, las Humanidades Digitales, un entorno científico que ya no puede considerarse emergente de una manera simple, persiguen la creación de conocimiento a través de la constatación de que el encuentro de los dominios humanísticos con las herramientas y las tecnologías digitales suscita unas preguntas de investigación que no son necesariamente nuevas,

pero que sí producen resultados novedosos que no se alcanzarían partiendo de las disciplinas cultivadas en soledad. Estos resultados son conocimientos realmente nuevos. Y estos conocimientos alcanzan a todos los ámbitos del saber humanístico y científico: de la lengua y el lenguaje a la literatura, de los textos a los libros, de la arqueología a la historia del arte y sus objetos presentes o perdidos, de los datos y la información y su tratamiento al patrimonio en todas sus realidades. Y muestran el avance imparable del desarrollo tecnológico que da la réplica a la investigación y se funde con ella.

De este modo, este volumen, que se presenta bajo el lema del conocimiento y de la búsqueda de los caminos para conseguirlo en el entorno de este encuentro entre el campo privilegiado del estudio científico del pensamiento humano que son las humanidades y las tecnologías digitales y el mundo virtual, se estructura en tres secciones que, a través de doce contribuciones, exploran tres territorios de las Humanidades Digitales que a la vez engloban diversos aspectos relacionados entre sí.

En la primera sección se contemplan desde distintos puntos de vista y con objetivos y objetos diferentes cuestiones relacionadas con el patrimonio cultural y su valor ("Humanidades Digitales y Arqueología: experimentando con modelos de reconocimiento automático de texto manuscrito en español sobre patrimonio cultural de inicios del siglo XX ", "El trabajo-proyecto Patrimonio Herido: Una propuesta de innovación educativa transdisciplinar") y con el patrimonio construido, reconstruido y visualizado ("Más allá de la representación tridimensional: la arquitectura (des)aparecida del Asilo de las Mercedes en el espacio digital", "Estudio interdisciplinar de una relación festiva: de la descripción textual a la representación visual de la arquitectura efímera y su iconografía" y "3D más allá de la reconstrucción virtual: Una herramienta al servicio de la gestión integral del patrimonio en la ermita de Sant Llorenç de la Senabra").

La segunda sección se ocupa del lenguaje, el texto y la literatura, ofreciendo análisis de corpus ("Entre la Historia y la

Filología: El proyecto CORDICan y el valor de la interdisciplinariedad en las HD a través de la documentación relativa a dotes y arras"), estudios de literatura ("La evolución de los libros de caballerías castellanos en el tiempo y el espacio"), una reflexión sobre las bibliotecas digitales con ediciones críticas ("Luces y sombras: Las grandes bibliotecas de textos griegos y latinos ante los retos de la nueva era de edición digital") y los estudios de teatro ("Proyecto HDATEATROUNIR: Realidades y perspectivas para el estudio del teatro desde las HD").

Por último, en la tercera sección se recogen tres trabajos que se ocupan del modelado ("Identificando el sujeto del paisaje: Modelado e integración de la dimensión social para una (re) lectura de la geografía altomedieval"), de los datos ("HILAME: De la persona al dato, del dato a la persona") y de la información obtenida de las redes sociales ("Estrategias de recuperación de información mediante *pooling* para la construcción de colecciones de referencia desde redes sociales: Caso de estudio durante las tensiones raciales de 2020").

Creemos que estas tres secciones, con los artículos que contienen, son buen ejemplo del diverso panorama de la práctica de las Humanidades Digitales hoy, y esperamos que inspiren nuevos trabajos y retos de investigación.

# I

# LAS VÍAS DE LA VISIÓN, LA RECONSTRUCCIÓN Y EL SENTIDO DEL PATRIMONIO

# MÁS ALLÁ DE LA REPRESENTACIÓN TRIDIMENSIONAL:
## LA ARQUITECTURA (DES)APARECIDA DEL ASILO DE LAS MERCEDES EN EL ESPACIO DIGITAL

Estefanía López Salas

**Resumen**

La reconstrucción virtual de arquitecturas completamente desaparecidas se realiza principalmente por medio de sistemas de dibujo y modelado computarizado 2D y 3D, que requieren la búsqueda e interpretación humana de fuentes de diferente naturaleza, con información sobre la apariencia física de la entidad histórica. La producción de un modelo digital riguroso a partir de estas, es decir, no una creación artística, sino una representación lo más exacta posible a nivel geométrico e histórico del edificio, conduce a un proceso complejo y diferente del de la investigación exclusivamente realizada a través de textos y/o imágenes para producir resultados eminentemente textuales.

Cuando intentamos reconstruir de modo gráfico un espacio perdido nos enfrentamos a múltiples cuestiones sobre su forma, dimensiones y materialidad, entre otras, claves para su definición en el mundo real, pero igualmente en el virtual. Al intentar darles respuesta de modo visual adquirimos un mejor conocimiento del edificio histórico, que se produce durante el proceso de recuperar el plano y hacer el modelado. El dibujo y modelado digital se convierten así en un método de investigación en sí mismos, pues revelan numerosas cuestiones espaciales que pasan desapercibidas con el uso de otros procedimientos tradicionales.

Estefanía López Salas, Universidade da Coruña
https://orcid.org/0000-0002-2889-6465
estefania.lsalas@udc.es

En el presente artículo mostramos el proceso de investigación, adquisición de conocimiento y recuperación de la memoria de una arquitectura desaparecida que se produce durante la creación de su modelo digital. En concreto, presentamos los resultados del trabajo en curso en el marco del proyecto I+D+i «Memoria del patrimonio arquitectónico desaparecido en Galicia. El siglo XX» (MEPAGAL) sobre el demolido Asilo de las Mercedes de Lugo, proyectado por el arquitecto Nemesio Cobreros Cuevillas en 1876 que, aunque no llegó a utilizarse con su función original, fue pionero en su tipología en la Galicia del cambio del siglo XIX al XX.

**Palabras clave:** Historia de la arquitectura; reconstrucción virtual; diseño asistido por ordenador; modelado histórico 3D; arquitectura hospitalaria decimonónica

**Abstract:**
Virtual reconstructions of lost architecture are now mainly accomplished with 2D computer-aided design software and 3D computer graphics. The drawing and modelling processes usually involve the search and human interpretation of a variety of historical sources that gathered data about the non-extant entity. Based on them, it is possible to create a rigorous 3D model that aims to not be an artistic creation, but an accurate geometrical description of the objects. In this case, the processes of drawing and modelling themselves lead the scholar through a new and complex research scenario in which a different approach to the lost entity is possible from previous methods based on textual and graphic historical data to also produce textual research results.

When we try to reconstruct a lost architecture in a graphic way, we are forced to deal with a variety of questions regarding geometry, measurement, and materiality, to cite a few, that are key to define a physical entity not only in the real world, but also in the virtual one. When we try to investigate and give a visual answer to each of them, we gain a better understanding of the historical artifact. This new knowledge is acquired in the process of drawing and modelling. That way, the process of making the final digital plans and model turns into a research process as it makes visible a number of spatial issues that will remain hidden with the use of other procedures.

This article presents the research process, gained knowledge and memory revival of a lost building that derives from the creation of its 2D plan and 3D digital reconstruction. In particular, we show the results of an in-process work within the research project titled «Memory of the Lost Architectural Heritage in Galicia. The 20th Century» (MEPAGAL) and the building called Hospice and Maternity Hospital of the Mercedes at Lugo. It was designed by the architect Nemesio Cobreros Cuevillas in 1876. Although it was not finally used for the original purpose, it is considered as a pioneer proposal in Galicia, in what refers to the field of hospital architecture in the transition from the 19th to the 20th century.

**Keywords:** History of architecture; virtual reconstruction; computer-aided design; historical 3D modelling; 19th-century hospital architecture

## 1. INTRODUCCIÓN

El primer propósito de la reconstrucción digital aplicada al campo de la investigación y divulgación de la arquitectura histórica es el de crear un modelo virtual espacial, temporal o semántico que responda a la pregunta de cómo era una entidad patrimonial en un determinado momento de su pasado. Tener acceso, desde la contemporaneidad, a la percepción de un espacio histórico, total o parcialmente no conservado, abre las puertas a la recuperación de su memoria y significado, y a la posibilidad de llevar a cabo nuevos análisis de esa arquitectura histórica. La cuestión del cómo es así pareja a la segunda pregunta más importante del por qué abordar su representación con herramientas digitales.

En el presente artículo proponemos mostrar el conocimiento adquirido sobre una arquitectura histórica concreta durante las primeras fases del flujo de trabajo llevado a cabo para su reconstrucción virtual. Nuestro caso de estudio es el edificio proyectado y construido como Asilo de las Mercedes en Lugo, a finales del siglo XIX. Pionero en su tipología en Galicia, no llegó nunca a utilizarse con su función original y sufrió múltiples transformaciones

en el tiempo, principalmente para su uso como colegio militar y cuartel, hasta su completa desaparición en 1964 (Castro Freire, 1951: 179-150; Abel Vilela, 1996: 181; Prado Gómez y Reboredo Pazos, 2001: 57; Fernández Fernández y Sánchez García, 2006: 192). De sus diferentes fases evolutivas, la que abordamos en este trabajo es la primera y más antigua, la del asilo, por ser esta la que definió su fábrica original y condicionó las posteriores obras de adaptación a otros programas funcionales. Con la reconstrucción digital de este espacio histórico no sólo buscamos recuperar la imagen perdida del complejo asistencial, sino también sentar las bases para, en una fase futura, analizar las intenciones del arquitecto en la creación del espacio y la importancia de la propuesta en su época.

A partir de una recopilación de planos históricos y fotografías del proyecto original y de otros de reformas sucesivas localizados en archivos y prensa histórica, elaboramos las trazas principales del complejo, primeramente como dibujos digitales bidimensionales con medidas reales para cada espacio componente en programas de diseño asistido por ordenador (CAD), extraídas de datos contenidos en fuentes textuales o del escalado de la planimetría histórica. Las trazas resultantes sirvieron de base para el modelado tridimensional posterior, asegurando una total consistencia entre los dos productos virtuales. En esta fase del flujo de trabajo se produjo un constante proceso comparativo entre las diferentes fuentes disponibles, para poder afrontar la representación del espacio no conservado. Este proceso sacó a la luz cuáles son los datos más relevantes para la comprensión del espacio histórico desaparecido, cuáles están ausentes en las fuentes, qué información espacial presenta incongruencias entre unas evidencias documentales y otras, y en qué partes son necesarias, posibles o limitadas, las hipótesis para producir un modelo final completo, coherente y riguroso desde un punto de vista histórico-artístico (Münster, Friedrichs y Hegel, 2018).

Documentar y reflexionar sobre esta parte del flujo de trabajo nos permitirá mostrar, por un lado, el cambio de lente

que los métodos de reconstrucción virtual conllevan con respecto a prácticas previas en el estudio del espacio de la arquitectura histórica y, por otro, los problemas asociados a la investigación en la que los métodos y los resultados son eminentemente visuales, frente a la textual, todavía dominante en el mundo académico.

## 2. EL PROYECTO MEPAGAL Y EL ASILO DE LAS MERCEDES COMO CASO DE ESTUDIO

El presente trabajo se enmarca dentro del proyecto I+D+i «Memoria del patrimonio arquitectónico desaparecido en Galicia. El siglo XX», concedido por el Ministerio de Ciencia e Innovación para el periodo 2020-2023. El objetivo de este proyecto, liderado por Jesús Ángel Sánchez García y Alfredo Vigo Trasancos, es documentar de forma exhaustiva el patrimonio gallego desaparecido en Galicia a lo largo de un periodo temporal concreto, de 1936 al 2001, con el fin de recuperar su conocimiento y reactivar su memoria.

Cada miembro del equipo se ocupa del estudio de entre tres y cinco monumentos, edificios o conjuntos urbanos de interés, desaparecidos o muy alterados, y ubicados en alguna de las capitales de provincia o principales ciudades de la comunidad autónoma de Galicia. La metodología que se aplica para su estudio se basa, por un lado, en la recopilación y análisis de documentos, planos, fotografías, noticias de época o testimonios, entre otros, que nos permiten documentar: su historia constructiva, su proceso de alteraciones y destrucción final, así como los agentes y dinámicas sociales, económicas o políticas implicadas en ella, sus características y los valores de los elementos perdidos. Por otro lado, este estudio se complementa con la elaboración de un modelo digital, que visualiza la arquitectura perdida y nos permite responder a la pregunta de cómo era esta entidad en un determinado momento de su pasado.

Dentro del proyecto MEPAGAL, uno de los edificios seleccionados en la ciudad de Lugo, es el desaparecido Asilo de

las Mercedes, que luego fue colegio preparatorio y cuartel militar (Abel Vilela, 2003: 320-321). Proyectado como hospicio y casa de maternidad en diciembre de 1876, se ubicaba en un amplio terreno en las afueras de la ciudad, al pie de la antigua carretera de Castilla, actual Avenida de Madrid[1]. El solar fue cedido a la Diputación Provincial de Lugo en febrero del mismo año (Abel Vilela 1996, 177)[2]. En él existía un edificio que se empezara a construir con destino a hospicio a finales del siglo anterior, pero sin llegar a buen término. La única condición de la cesión del terreno era que las obras de una construcción destinada a fines benéficos se iniciasen de forma inmediata, para poner fin a los diversos problemas que la ciudad tenía en la atención a los pobres, enfermos y necesitados (Castro Freire, 1951: 122-123).

El proyecto fue costeado por la Diputación Provincial y por el Ayuntamiento de Lugo que nombraron una comisión mixta para disponer todos los trabajos preparatorios, entre los que se incluía la elaboración de la memoria, planos y presupuestos para la construcción, que se encargó al arquitecto provincial, Nemesio Cobreros Cuevillas[3]. Este está considerado como el más importante de la historia contemporánea de Lugo, no sólo por el desarrollo que la ciudad experimentó en el periodo en el que ejerció como arquitecto provincial, finales del siglo XIX, sino también por su implicación en numerosas e importantes obras públicas, de las cuales era ejemplo nuestro caso de estudio (Abel Vilela, 2003: 320).

El presupuesto total del nuevo edificio «destinado a casa provincial de Maternidad y Expósitos y a Hospicio o casa municipal

1. Proyecto de Hospicio y casa de maternidad para la ciudad de Lugo, 1876-1888. Localizado en: Archivo Histórico Provincial de Lugo (en adelante, AHPL), Fondo Concello, signatura: 0442-03.
2. Proyecto de Hospicio y casa de maternidad para la ciudad de Lugo, 1876-1888. Localizado en: AHPL, Fondo Concello, signatura: 0442-03, f. 6r.
3. Proyecto de Hospicio y casa de maternidad para la ciudad de Lugo, 1876-1888. Localizado en: AHPL, Fondo Concello, signatura: 0442-03, f. 6r.

de Beneficencia» ascendía a 264.386,44 pesetas[4], un primer dato que revela la magnitud del proyecto planteado. Una vez construido, el conjunto lo formaban: un cuerpo central de planta baja y primera para la casa de maternidad y diversos servicios generales del asilo; dos pabellones laterales sobresalientes, con diferentes estancias distribuidas en planta baja y una altura, comunicadas por galerías exteriores cubiertas, que definían dos patios abiertos, uno de mujeres y otro de hombres; y, finalmente, un tercer cuerpo en la parte posterior de planta baja, para almacenes[5]. Con una superficie construida final de alrededor de cinco mil metros cuadrados, el Asilo de las Mercedes estaba preparado para acoger un total de doscientos cincuenta asilados en sus diferentes estancias, rodeadas de un amplio terreno para su uso y disfrute.

A nivel tipológico, la solución propuesta por Cobreros Cuevillas era deudora de los modelos de hospitales franceses e ingleses de la segunda mitad del siglo XIX que, abandonando las organizaciones claustrales de los hospitales históricos, apostaban por una distribución en pabellones, que priorizaba la organización, control y aislamiento de cada tipo de usuario, a la par que garantizaba su relación con el entorno, clave para su recuperación o crecimiento (Zayas Fernández, 2012: 108). Aunque la función y programa de este asilo, representa la continuidad de la concepción de las casas de beneficencia, hospicios y hospitales existentes en los siglos anteriores, su propuesta tipológica ha sido considerada como pionera dentro del contexto gallego de cambio del siglo XIX al XX y ahí, en parte, radicaba el valor de esta arquitectura desaparecida (Fernández Fernández y Sánchez García, 2006: 192-193). Esta es la razón principal que determina la decisión de abordar la

4. Proyecto de Hospicio y casa de maternidad para la ciudad de Lugo, 1876-1888. Localizado en: AHPL, Fondo Concello, signatura: 0442-03, f. 7r.

5. Proyecto de Hospicio y casa de maternidad para la ciudad de Lugo, 1876-1888. Localizado en: AHPL, Fondo Concello, signatura: 0442-03, ff. 83r-85r.

reconstrucción digital de la primera y más antigua de las diferentes fases evolutivas que experimentó este edificio a lo largo de su trayectoria vital, entre la recepción provisional de la obra en 1885 y su demolición de 1964[6].

## 3. AUSENCIAS, EVIDENCIAS E INCONSISTENCIAS DE LAS FUENTES HISTÓRICAS

Para alcanzar nuestro objetivo, el primer paso fue la localización de documentación histórica con información relativa a cómo era la arquitectura del primitivo asilo. El reto planteado es reconstruir el edificio histórico en un ecosistema digital y, para ello, es necesario reunir datos de diverso tipo que, de hecho, son similares a los que necesitaríamos para hacer el mismo ejercicio en el mundo real. En este sentido, el instrumento fundamental de la arquitectura para definir un espacio es el plano, como condensador de las trazas del espacio proyectado. El plano revela cuestiones formales,compositivas y dimensionales, entre otras, que son clave para la comunicación y estudio de la arquitectura, y también para la recuperación de una obra desaparecida. Por tanto, la localización de planos históricos constituye el primer paso del proceso de trabajo desarrollado de cara a la creación del modelo tridimensional.

En la memoria presentada por el arquitecto provincial se hace referencia a que se elaboraron un total de nueve planos para definir el edificio[7]. El primero de ellos era un plano de situación del solar, en el que se representaba la edificación empezada a construir

6. Demolición del viejo cuartel de Las Mercedes, 8 diciembre 1964. *El Pueblo Gallego*, p. 12. Localizada en: Biblioteca Pública de Pontevedra Antonio Odriozola. Copia digitalizada disponible en: http://biblioteca.galiciana.gal/gl/catalogo_imagenes/grupo.do?path=1116090&posicion=12&presentacion=pagina [último acceso, 31 diciembre 2021].

7. Proyecto de Hospicio y casa de maternidad para la ciudad de Lugo, 1876-1888. Localizado en: AHPL, Fondo Concello, signatura: 0442-03, f. 7r.

en el siglo XVII, el terreno anexo a ella y las curvas de nivel. Perfiles y rasantes del terreno estaban incluidos en el plano número dos. El tercero y cuarto de los documentos gráficos contenían la planta baja y principal del nuevo asilo, respectivamente. Las fachadas frontal, lateral y posterior se presentaban en los planos quinto, sexto y séptimo. La definición del edificio se completaba con una sección transversal en el plano número ocho y, finalmente, con la planta de cubiertas en el noveno documento. Por tanto, el proyecto contenía un conjunto de planos, que podríamos llamar técnicos o de ejecución, cuyo propósito era no sólo mostrar la obra al promotor, sino también servir de instrumento para su perfecta construcción.

El proyecto original del asilo se conserva en el Archivo Histórico Provincial de Lugo, sin embargo, no contiene ninguno de los planos citados. La explicación se encuentra dentro del propio documento, en una nota manuscrita añadida con fecha de enero de 1886, en la que se indica que: «Se han desglosado los planos del edificio que expresa este expediente y se remitieron con conocimiento del Sr. Alcalde a Madrid, bajo sobre certificado dirigidos a Don Teolindo Soto Barro», diputado por la provincia de Lugo[8].

Sabemos que, en agosto de 1885, se había producido la entrega provisional del edificio a las corporaciones copropietarias, es decir, estaba terminado en su mayor parte[9]. No obstante, la recepción definitiva de la obra se retrasó hasta abril de 1887 y[10],

8. Proyecto de Hospicio y casa de maternidad para la ciudad de Lugo, 1876-1888. Localizado en: AHPL, Fondo Concello, signatura: 0442-03, f. 86r.

9. Locales, 17 agosto 1886. *El Eco de Galicia*, p. 3. Localizada en: Instituto Padre Sarmiento de Estudos Galegos. Copia digitalizada disponible en: http://biblioteca.galiciana.gal/gl/catalogo_imagenes/grupo.do?path=1037379&posicion=3&presentacion=pagina [último acceso, 31 diciembre 2021].

10. La ciudad, 20 abril 1887. *El Regional*, p. 2. Localizada en: Biblioteca Dixital de Galicia (en adelante, Galiciana). Copia digitalizada disponible en: http://biblioteca.galiciana.gal/gl/catalogo_imagenes/grupo.do?path=1051254&posicion=2&presentacion=pagina [último acceso, 31 diciembre 2021].

en ese corto periodo temporal, se sucedieron una serie de noticias en la prensa de la época que informaban de varias gestiones en curso con el Ministerio de Hacienda para cambiar la funcionalidad del edificio recién construido. En ellas encontramos la referencia a la petición de envío de los planos del proyecto del asilo al senador Manuel Vázquez de Parga Somoza, Conde de Pallares, en junio de 1886, que era una de las personas que estaba intentando negociar en la capital la conveniencia o no sobre la instalación de otros usos en el edificio[11]. La confirmación de la llegada de los planos a Madrid, o de una copia de ellos, figura en otra noticia de la prensa local de julio del mismo año y allí seguían a finales de marzo de 1888[12],es decir, poco tiempo antes de que, desechada su apertura como asilo, se acordase la instalación en el nuevo edificio de un colegio preparatorio militar, que estuvo en funcionamiento desde noviembre de 1889 hasta 1893[13].

Una parte de los planos mencionados, en concreto, las plantas baja y principal, aunque no se conservan entre la documentación del proyecto original del archivo, nos son conocidas, porque se utilizaron para ilustrar un amplio reportaje sobre el edificio, publicado por el diario local *El Regional*, en los números

11. Locales, 25 junio 1886. *El Eco de Galicia*, p. 3. Localizada en: Instituto Padre Sarmiento de Estudos Galegos, Real Academia Galega. Copia digitalizada disponible en: http://biblioteca.galiciana.gal/gl/catalogo_imagenes/grupo.do?path=1037336&posicion=3&presentacion=pagina [último acceso, 31 diciembre 2021].

12. Locales, 20 julio 1886. *El Eco de Galicia*, p. 3. Localizada en: Instituto Padre Sarmiento de Estudos Galegos. Copia digitalizada disponible en: http://biblioteca.galiciana.gal/gl/catalogo_imagenes/grupo.do?path=1037356&posicion=3&presentacion=pagina [último acceso, 31 diciembre 2021]; La orfandad de Lugo, 31 marzo 1888. El Eco de Galicia, p. 1. Localizada en: Instituto Padre Sarmiento de Estudos Galegos. Copia digitalizada disponible en: http://biblioteca.galiciana.gal/gl/catalogo_imagenes/grupo.do?path=1037863&posicion=1&presentacion=pagina [último acceso, 31 diciembre 2021].

13. Las Migas, 12 noviembre 1889. *El Eco de Galicia*, p. 2. Localizada en: Galiciana. Copia digitalizada disponible en: http://biblioteca.galiciana.gal/gl/catalogo_imagenes/grupo.do?path=1038334&posicion=2&presentacion=pagina [último acceso, 31 diciembre 2021].

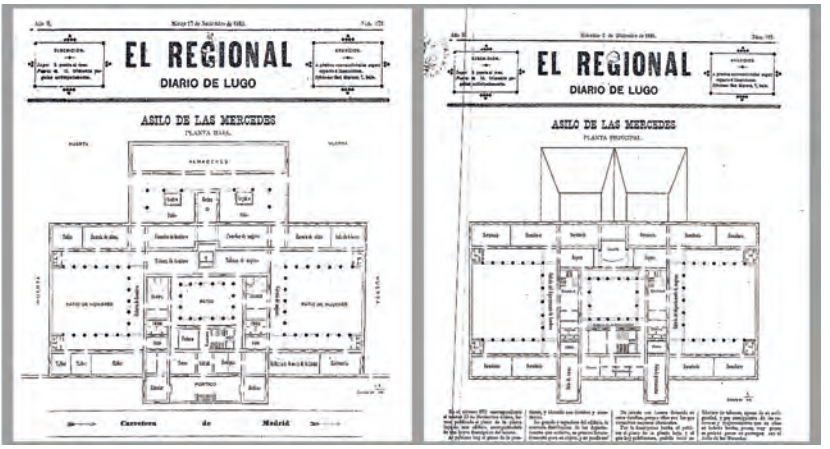



Figura 1. Planta baja y principal del Asilo de las Mercedes publicadas a finales de 1885. Fuente: Asilo de las Mercedes, 17 noviembre y 2 diciembre 1885. *El Regional*, p. 1.

373 y 388 de noviembre y diciembre de 1885, respectivamente[14]. Estos dos planos son los documentos gráficos más tempranos que hemos podido localizar del primitivo asilo y un punto de partida para el estudio de esta arquitectura desaparecida (Figura 1). No obstante, carecen de escala gráfica, elemento que también es necesario para dar los primeros pasos de una reconstrucción rigurosa desde un punto de vista dimensional.

Para completar la información ausente en esos dos primeros testimonios gráficos, es necesaria la revisión de la memoria del proyecto y de otras fuentes de esa época o posterior, que puedan contener datos escritos o dibujados relativos a las medidas del edificio o de alguno de sus espacios, y no sólo en planta, sino también en la tercera dimensión o altura. En concreto,

14. Asilo de las Mercedes, 17 noviembre y 2 diciembre 1885. *El Regional*, pp. 1-2; 1. Localizada en: Galiciana. Copia digitalizada disponible en: http://biblioteca.galiciana.gal/gl/catalogo_imagenes/grupo.do?path=1050759&presentacion=pagina&posicion=1&registrardownload=0 y http://biblioteca.galiciana.gal/gl/catalogo_imagenes/grupo.do?path=1050773&presentacion=pagina&posicion=1&registrardownload=0 [último acceso, 1 enero 2022].

en la memoria del proyecto de 1876 se señala que el nuevo asilo ocuparía una superficie de cuatro mil quinientos metros cuadrados y que, la planta baja de la esquina sur de la fachada principal, se separaría del nivel del terreno tres metros[15].

El reportaje de 1885 es algo más preciso a este respecto, pues enfatizaba que el edificio medía «90 metros de fachada por 56 de fondo, haciendo un total de 5.000 metros de edificación»[16]. Aunque este último es un dato diferente al indicado en el documento anterior, que clasificaríamos dentro de la no poco habitual inconsistencia de las fuentes históricas, los otros dos sí son relevantes, pues nos permiten un primer ejercicio de escalado de los planos históricos con el uso de herramientas digitales. En otras palabras, podemos aumentar o reducir el tamaño del plano con un programa de diseño asistido por ordenador, en la parte correspondiente a su fachada o fondo, y de acuerdo a cada uno de los valores dimensionales indicados en la fuente textual. Con este ejercicio se abre una vía para recuperar la escala de la que carece el documento original.

Si asignamos 90 m de fachada al plano de la planta baja, lo primero que podemos comprobar es que el fondo total del conjunto alcanzaría los 73 m y los 56 que se indican en el reportaje serían tal, sin contar el pórtico sobresaliente en el frente, ni los almacenes de la parte posterior (Figura 2). Asimismo, con este proceso de escalado digital se pone de manifiesto que la superficie construida del edificio, con esas dimensiones de referente, era más próxima a los cinco mil metros cuadrados señalados en el reportaje, que a los cuatro mil quinientos de la memoria del proyecto, quizás como consecuencia de que el edificio ya se había

15. Proyecto de Hospicio y casa de maternidad para la ciudad de Lugo, 1876-1888. Localizado en: AHPL, Fondo Concello, signatura: 0442-03, f. 83v-84r.

16. Asilo de las Mercedes, 17 noviembre 1885. *El Regional*, p. 2. Localizada en: Galiciana. Copia digitalizada disponible en: http://biblioteca.galiciana.gal/gl/catalogo_imagenes/grupo.do?path=1050759&presentacion=pagina&posicion=2&registrardownload=0 [último acceso, 1 enero 2022].

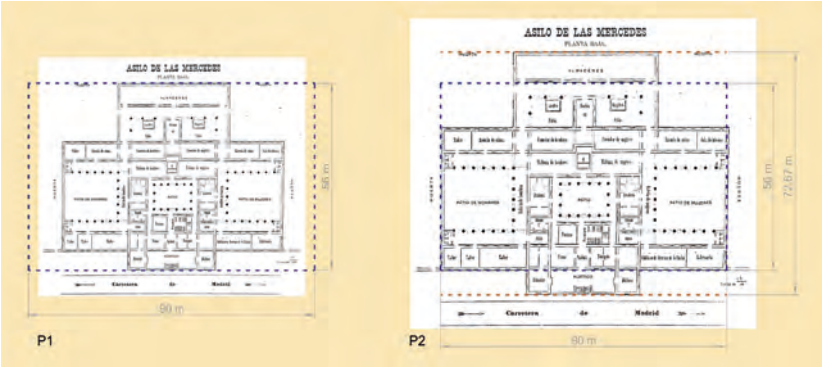



Figura 2. Pruebas de escalado de la planta baja del asilo publicada en 1885 y las cotas de los resultados. Fuente: Elaboración propia.

materializado y, por tanto, en 1885 se describía una realidad. El ejercicio de escalar un plano histórico sin escala gráfica a partir de un dato dimensional que procede de una fuente textual debe, por tanto, basarse en una interpretación cuidadosa de los datos combinados y, sobre todo, contrastarlos con otras fuentes que permitan su validación o no. En cualquier caso, en este punto del proceso de trabajo tenemos ya una primera aproximación a las dimensiones reales del edificio en planta, que la revisión de otros documentos nos permite completar.

En mayo de 1888, *El Eco de Galicia* publicaba la copia de una misiva enviada por el arquitecto provincial al Ministerio de Guerra, en nombre del Ayuntamiento de Lugo, para explicar las buenas condiciones del edificio en un momento en el que se estaba todavía decidiendo qué uso darle[17]. Entre los datos recogidos en la carta se indicaba que la altura de los pisos era de cuatro metros en la planta baja y de cinco en la principal. Estos nos permiten empezar a conocer las dimensiones en altura del edificio, claves

17. La Academia militar preparatoria, 8 mayo 1888. *El Eco de Galicia*, p. 1. Localizada en: Galiciana. Copia digitalizada disponible en: http://biblioteca.galiciana.gal/gl/catalogo_imagenes/grupo.do?path=1037895&posicion=1&presentacion=pagina [último acceso, 1 enero 2022].

en un proceso de reconstrucción del espacio histórico en el que el modelo final es tridimensional.

En una Real Orden de marzo de 1889, emitida para valorar el proyecto de reforma del asilo para la instalación de un colegio militar, encontramos otra referencia dimensional. Aparece cuando al determinar la necesidad de construir de nueva planta un ala de edificio que cerrase el patio meridional, se indicaba que los pisos bajo y principal debían formar una crujía de seis metros de ancho, con galería como las que había en el edificio existente[18].

Para localizar datos más precisos sobre las dimensiones del espacio desaparecido, hemos recurrido también a algunos de los proyectos que se conservan de las diversas reformas y transformaciones que experimentó el edificio para ser adaptado a cuartel militar tras el fin de su uso como colegio preparatorio en 1893. Muchos de ellos son planos parciales en los que tan sólo se representaron los añadidos, pero alguno de los expedientes sí incluye planos completos del estado del edificio, dibujados con más o menos precisión[19].

En concreto, destacan los tres planos asociados a un proyecto de reforma y ampliación del edificio de 1923, que hoy se conservan en el Instituto de Historia y Cultura Militar dependiente de la Subdirección General de Patrimonio del Ministerio de Defensa de

18. Crónica parlamentaria municipal, 8 marzo 1889. *El Eco de Galicia*, pp. 2-3. Localizada en: Galiciana. Copia digitalizada disponible en: http://biblioteca.galiciana.gal/gl/catalogo_imagenes/grupo.do?path=1038134&posicion=2&presentacion=pagina [último acceso, 1 enero 2022].

19. En el AHPL, además del proyecto de 1876, se conservan varios expedientes relativos a este edificio en el Fondo Concello, de los años 1893 a 1914, con signaturas 0443-1 a 0443-25. Sólo los de signaturas 0443-1 y 0443-5 tienen un plano total y otro parcial del edificio, respectivamente. Los demás expedientes contienen planos de los nuevos edificios que se proyectaron en ese periodo, como almacenes, cuadras, cobertizos,... Asimismo, nos falta por consultar los expedientes que sobre este edificio se conservan en el Archivo General Militar de Segovia, Sección 3ª, Legajo 556, en cuya descripción solo figura la existencia de dos planos, uno de situación, en diversos expedientes de entre 1893 a 1925.

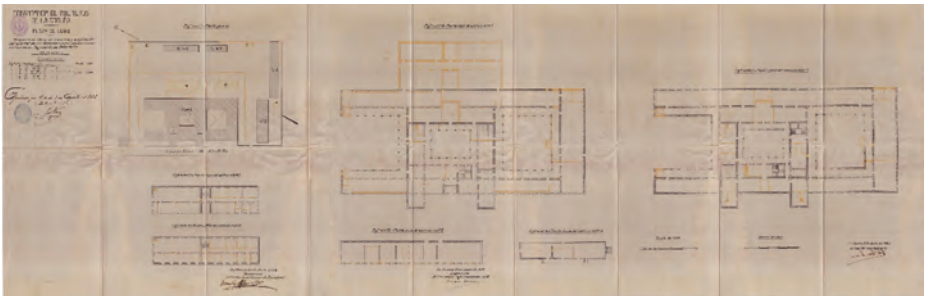


Figura 3. Plano nº 1 del proyecto de obras de reforma y ampliación en el cuartel de las Mercedes para alojamiento definitivo de un regimiento de infantería, 1923. Fuente: IHCM, signatura: SH AFI-353/2.

España, en Madrid[20]. El primero de ellos, titulado «Demoliciones», incluye una representación de la planta general de la parcela con un esquema de los varios edificios que componían el cuartel militar, seguida por dos planos de mayor detalle de las plantas baja y primera del edificio principal, y de los demás independientes que se situaban en sus inmediaciones (Figura 3). Cada plano utiliza el trazo negro para representar las partes de los edificios que se propone mantener y, en amarillo, aquellas otras a demoler. En realidad, es un conjunto de planos que podemos llamar del «estado actual» del cuartel, antes de acometer su reforma y ampliación. El plano número dos contiene el proyecto de reforma en sí, es decir, la planta general, baja y principal del cuartel y sus edificios anexos con la representación en color rojo de todos los añadidos y nuevos edificios que los autores del proyecto proponían. En último lugar, el plano número tres, titulado «Alzados y cortes», se ocupaba de mostrar los trabajos de reforma y ampliación a ejecutar por medio del grafiado de las fachadas parciales y de algunas secciones principales de las construcciones existentes y de otras de nueva planta.

20. Proyecto de obras de reforma y ampliación en el Cuartel de las Mercedes para alojamiento definitivo de un regimiento de infantería, 1923. Localizado en: Instituto de Historia y Cultura Militar (en adelante, IHCM), signatura: 8182.3. Los planos tienen por signaturas: SH AFI-353/2, SH AFI-353/3 y SH AFI-353/4.

A diferencia de los planos de 1885, estos de 1923 sí incluyen una escala gráfica que, sin necesidad de recurrir a otras fuentes, permite el conocimiento de las dimensiones de cada elemento representado. No obstante, es importante tener en cuenta que el conjunto edificado de 1923 que se visualiza en el primero de los planos no es el asilo terminado en 1887, sino el cuartel militar que fue definido a partir de aquel tras muy diversos y complejos procesos de adaptación y reforma.

## 4. HACIA EL MODELADO DIGITAL DE LA ARQUITECTURA DESAPARECIDA

Para poder reconstruir el Asilo de las Mercedes, el edificio originalmente proyectado por Nemesio Cobreros Cuevillas, es necesario combinar la información que podemos obtener de las diversas fuentes presentadas previamente, sin olvidar el uso de las fotografías históricas que, aunque no son abundantes en este caso, sí nos pueden ayudar a completar algunas cuestiones formales o compositivas, y sobre todo relativas a la materialidad, que no reflejen las otras vías de aproximación al espacio histórico. Asimismo, nos permiten testear la validez de aquellas, pues las fotografías sí muestran lo realmente construido, algo que no siempre sucede en los planos de proyecto o en las memorias que los acompañan, que pudieron experimentar cambios durante la ejecución de las obras, que no quedaron plasmados en los documentos gráficos.

Si comparamos los planos de la planta baja del edificio de 1885 con la del estado actual de 1923 podemos reconocer que el nivel de detalle de la primera —en cuanto a la representación de muros, divisiones, escaleras, carpinterías, o incluso elementos en proyección—, es inferior a la segunda (Figura 4). Esto puede deberse a que su calidad original fue reducida para su inclusión como ilustración de un reportaje de prensa, o bien a que no se trata del plano de Cobreros, sino de una copia simplificada. De hecho, los nombres que identifican en este plano cada una de las

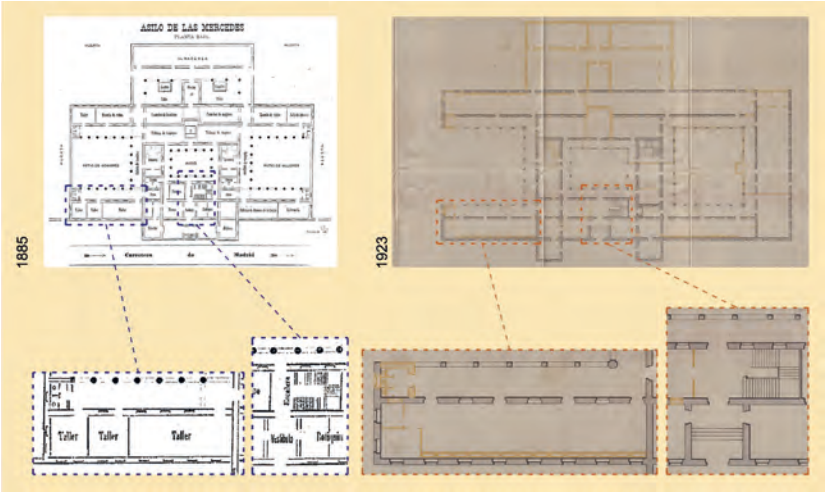



Figura 4. Análisis comparativo del nivel de definición gráfica del plano de la planta baja de 1885 y de 1923 para dos mismos espacios. Fuente: Elaboración propia.

estancias que componían el asilo están escritos con tipografías de imprenta, a pesar de que en los planos de época, el texto, si existía, era manuscrito por el propio técnico que elaboraba el proyecto.

Sea como fuere, esa mayor precisión de la representación gráfica determinó el uso de la serie de planos de 1923 como base sobre la que restituir el edificio de 1885. El documento histórico digitalizado fue escalado de acuerdo a su propia escala gráfica y, tras ello, es posible comprobar que, tal y como indicaba el reportaje de 1885, el asilo de las Mercedes tenía un frente de 90 m, pero su fondo no era de 56 m, sino que alcanzaba alrededor de los 70 (Figura 5). Para seguir testeando la validez de la fuente, escalamos también los planos correspondientes a los alzados y secciones de acuerdo a la escala gráfica del documento. Una vez realizado este ejercicio comprobamos que la altura de la planta baja es de cuatro metros y la del piso principal de cinco, tal y como indicaba la carta de Nemesio Cobreros que fue publicada por *El Eco de Galicia* en 1888. En cuanto a los seis metros de ancho que se decía que tenían los pabellones del asilo en la Real Orden de 1889, en el plano de

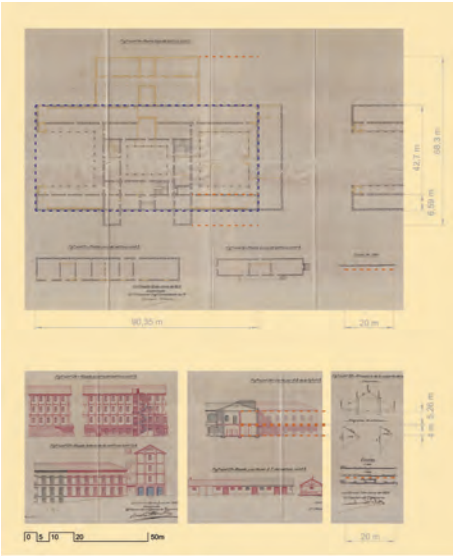


Figura 5. Escalado de la planta baja, alzado y sección de los planos del proyecto de reforma de 1923 a partir de un programa de diseño asistido por ordenador. Fuente: Elaboración propia.

1923 escalado podemos comprobar que solo los dormitorios, sin las galerías a las que se abrían, superaban esa dimensión. Lo mismo ocurre con el plano histórico de 1885 que, además, presenta una cierta distorsión, en forma de estiramiento de las dimensiones, en la dirección de la profundidad del edificio.

Durante el dibujo de las trazas del asilo desaparecido, se llevó a cabo un constante proceso comparativo entre las diferentes fuentes disponibles, para poder afrontar la representación del espacio no conservado con rigor. Este proceso hizo visibles otra serie de ambigüedades de los planos históricos, en cuanto a cómo el espacio arquitectónico fue representado que, en uno de los casos, reduce su grado de confiabilidad como fuente. Para ilustrar esta cuestión sirvan los tres ejemplos que se exponen a continuación.

Cuando nos enfrentamos a la representación de la planta baja del cuerpo central del edificio, si atendemos al plano de 1923, vemos que la simetría en la apertura de huecos a ambos lados del eje central del asilo era total. No obstante, si observamos los planos de 1885, esta simetría no existe ni en planta baja ni en la principal. Esta

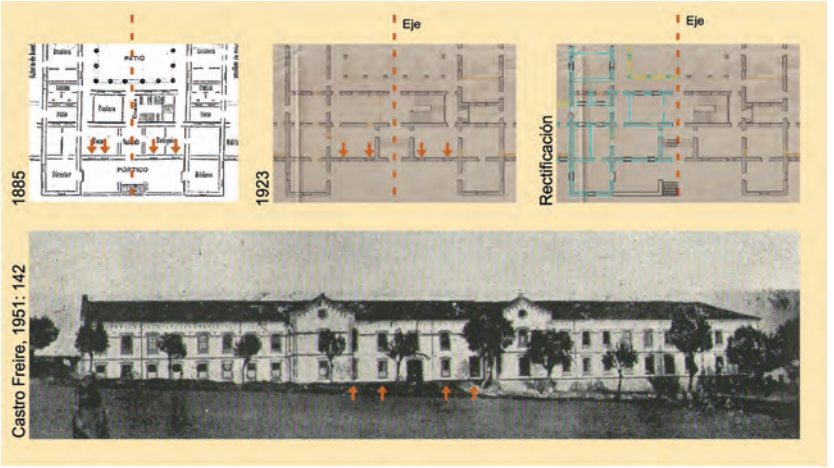



Figura 6. Primer ejemplo del proceso de recuperación de la traza en planta de parte del cuerpo central con indicación de las ambigüedades detectadas en la representación de la parte central de la fachada. Fuente: Elaboración propia.

ambigüedad que surge de la comparación de ambos documentos, se puede resolver acudiendo a las fotografías del edificio, que muestran una completa simetría en la disposición de ventanales y puertas a ambos lados de la entrada principal (Figura 6).

Algo similar ocurre cuando nos enfrentamos a la representación de las galerías de los dos patios abiertos que definían los pabellones sobresalientes hacia el norte y el sur del cuerpo central del edificio. La lectura comparativa de ambos documentos gráficos revela claustros distintos en cuanto al número de sus arcos. En las plantas publicadas en 1885 se representaron dos patios sensiblemente cuadrados, definidos por dos pandas paralelas de seis arcadas y un tramo ciego y otra panda transversal de seis arcos ligeramente más anchos que todos los demás. Sin embargo, si atendemos al plano de 1923, son siete tramos en total los que definen tanto las galerías paralelas de los pabellones, como la fachada interior transversal del cuerpo central que los une (Figura 7). La inconsistencia de las fuentes, en este caso, resulta complicada de resolver acudiendo a fotografías históricas pues,

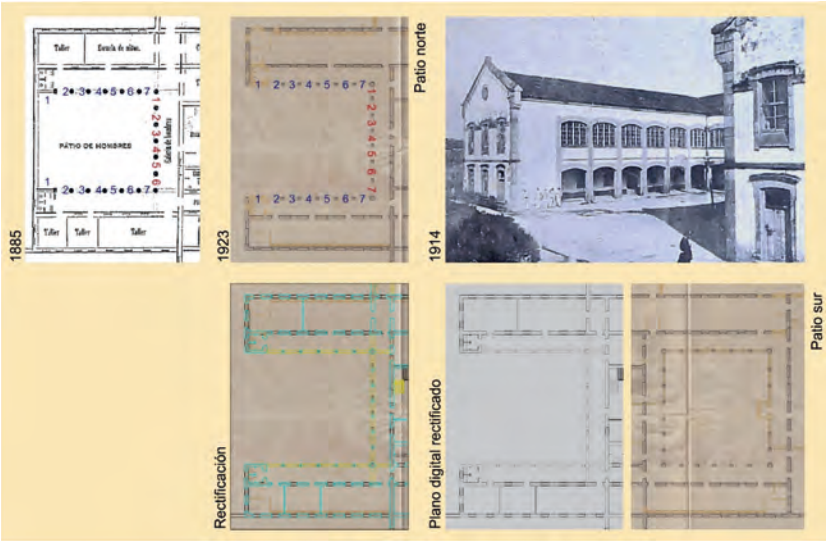



Figura 7. Ejemplo de inconsistencia entre las fuentes históricas en la representación de los patios. Fuente: Elaboración propia. De la fotografía de 1914: Gráficos de actualidad, 30 noviembre 1914. *Vida gallega*, p. 17.

hasta la fecha, no hemos localizado ninguna instantánea que capturara la dimensión completa de esas fachadas interiores que, además, fueron muy modificadas en sucesivas reformas. Por tanto, la reconstrucción que se plantea de esa zona es hipotética, a falta de localizar otros documentos históricos gráficos que conviertan en evidencia una u otra opción. Asimismo, aunque optamos por la existencia de siete arcadas en cada frente, al realizar el plano digital corregimos la falta de simetría que muestra la planta baja de 1923 en la representación de pilastras de la galería transversal del patio norte, pero no en la del sur y que, muy probablemente, son resultado de errores de grafiado en el documento original.

El tercer y último ejemplo concentra su atención en las fachadas laterales de las alas del edificio que eran ligeramente sobresalientes con respecto al plano de la fachada principal, y flanqueaban lo que en la planta de 1885 se identifica como pórtico de entrada. En este documento de finales del XIX observamos que

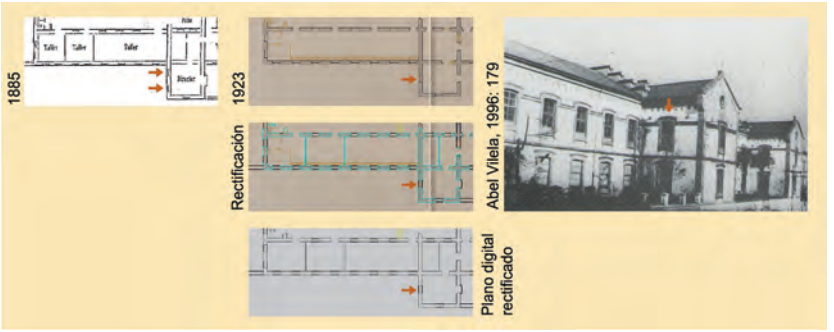



Figura 8. Tercer ejemplo del proceso de dibujo asistido por ordenador de los cuerpos sobresalientes en la fachada principal. Fuente: Elaboración propia.

se representaban dos ventanales para iluminar tanto la sala del director del centro y del médico ubicadas en planta baja, como las salas de cunas y parturientas del cuerpo superior. Sin embargo, el plano de 1923 y una fotografía de 1914 reflejan una composición de esos tramos de fachada que se apoya en la apertura de un único hueco en ambos niveles (Figura 8). No parece que los diferentes datos que aportan esas dos fuentes planimétricas se deban a una reforma del edificio original, sino más bien a un error en la representación del más antiguo de los dos documentos.

Con el desarrollo de este tipo de dibujo intelectual, es posible obtener como resultado las plantas, alzados y secciones rectificadas del desaparecido Asilo de las Mercedes (Figura 9). Estas trazas digitales bidimensionales contienen la información necesaria sobre su forma, dimensiones, organización, estancias, elementos arquitectónicos y ornamentales,... en el estado actual de conocimiento que tenemos sobre la obra, derivada de la comparación de las fuentes disponibles, más no de su digitalización directa. Asimismo, son la base para el inicio de la elaboración del modelo tridimensional que, en este trabajo, se lleva a cabo en el programa multiplataforma de Blender.

El primer paso a realizar dentro del ecosistema digital de Blender es importar los planos vectoriales creados con CAD y, a partir

de ellos, levantar el modelo digital. Una cuestión que a veces surge en las discusiones sobre cómo abordar un modelado histórico es por qué no elaborar directamente la representación tridimensional de la obra con el programa de computación gráfica seleccionado, sin el trabajo previo de recuperación de las plantas, alzados y secciones en un software de diseño asistido por ordenador. En otras palabras, un segundo camino posible y utilizado en el campo de la recreación virtual histórica, es integrar como imagen los documentos planimétricos de archivo en el programa de modelado, aquellos que creemos que mejor representan el edificio y, a partir de ellos, afrontar la recreación virtual del mismo. Si así se hace, el proceso iterativo de trabajo, basado en la comparación de las diversas fuentes disponibles, se trasladaría al entorno tridimensional. No obstante, en nuestra experiencia, este camino presenta una serie de limitaciones y dificultades cuando se trabaja con arquitecturas que no se conservan y de las que no hay planimetría disponible o esta es escasa.

La principal es que la evaluación de la confiabilidad de las fuentes planimétricas, si se aborda, se traslada a un entorno en el que la libertad para el testeo de los datos dimensionales y formales —en cuyo proceso de representación se forma una mejor comprensión del tema—, es menor que en programas de diseño asistido por ordenador. Estos últimos, al estar orientados a la creación, modificación, análisis u optimización de un diseño, permiten un dibujo más técnico y, por tanto, más próximo al de definición de los planos de ejecución de un edificio. Y, en el caso de una reconstrucción virtual, lo que realmente queremos hacer es volver a construir el edificio, no en el mundo físico real, sino en el virtual.

No obstante, los programas de computación gráfica como Blender, son la mejor opción para a partir de las trazas rectificadas, construir el modelo y prepararlo para renderizaciones fotorrealistas o cuando se requieren entornos de realidad virtual o realidad aumentada. Asimismo, cada vez son más importantes incluso para la creación de recursos atractivos a través de los cuales difundir o transferir los resultados de una investigación

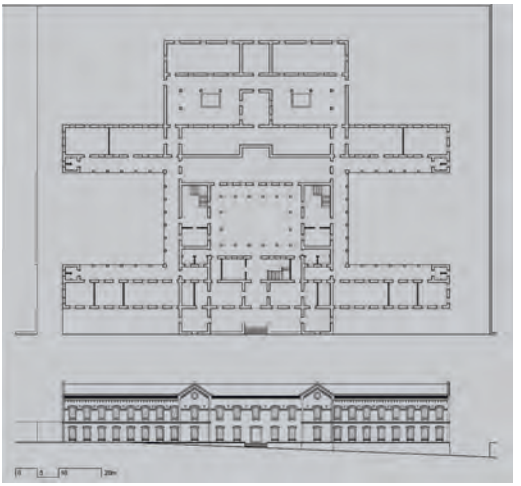


Figura 9. Planta baja y alzado principal del Asilo de las Mercedes en 1885 resultado del proceso de rectificación con CAD. Fuente: Elaboración propia.




Figura 10. Del plano restituido al modelado 3D de la planta baja del asilo de 1885. Fuente: Elaboración propia.

(De Francesco y D'Andrea, 2008). En resumen, las dos vías son válidas, combinadas permiten aprovechar las mejores capacidades de ambas, si bien convendría tratar de mejorar y estandarizar las formas en las que interactúan entre ellas.

Al integrar los planos vectoriales rectificados del edificio histórico en Blender y hacer el modelado en base a estos, se asegura una total consistencia entre los dos productos digitales si, al iniciar el proceso, se activan las herramientas de encajado de objetos a determinados elementos de la escena, los vértices. Asimismo, con este método, se consigue trasladar al modelo toda la precisión que, en la fase previa de representación bidimensional, fue encapsulada en los planos rectificados (Figura 10).

El proceso de creación del modelo conduce también a nuevos interrogantes sobre el edificio histórico, que ahora tienen un mayor foco de atención en la materialidad, especialmente una vez que, creada la volumetría, se acomete su texturizado. En este punto se inicia otro proceso de consulta de las fuentes históricas disponibles para poder representar, con el mayor rigor posible, el material de cada uno de los elementos componentes del edificio que, a su vez, conduce a una mejor comprensión de la arquitectura histórica desaparecida, que se adquiere en el propio proceso de hacer el modelo.

En nuestro caso de estudio, el material con el que fue hecha cada parte de la obra está descrito, en algunas ocasiones, en las fuentes textuales, principalmente en las memorias de proyectos iniciales o de reforma, presupuestos de obra, pero también en alguna noticia de prensa. Como ejemplo del tipo de información que podemos obtener en las fuentes textuales, sirva de ejemplo, en primer lugar, la breve descripción que Nemesio Cobreros hizo de los aspectos constructivos del asilo en la memoria del proyecto de 1876:

> Los muros los suponemos construidos de mampostería de piedra pizarra, la sillería de granito, invirtiéndose este material en los ángulos del edificio, cornisas y marcaciones de los huecos de la fachada primera,

> pilares y arcos de las galerías, y en el pavimento de los portales y galerías de la planta baja. Las vigas de la armadura las suponemos de madera de castaño, las del piso principal tablones de pino tea de los Estados Unidos y toda la carpintería de taller, y tabla de los pisos de este último material. Nos hemos sujetado también a las formas y dimensiones más usuales en las construcciones de esta Capital.[21]

Para este mismo fin de recuperar la materialidad, las memorias de los proyectos de reforma son igualmente muy útiles ya que en ellas es habitual la descripción del estado previo y de la propuesta de cambio, especialmente de aquellas que concentran la atención en el mantenimiento de ciertos elementos del conjunto. En un proyecto de renovación de las cubiertas del edificio de 1933 se indicaba que tenía «su cubierta en la actualidad constituida por placas de pizarra natural de 0,40 x 0,40 m sobre enlatado correas y armaduras de madera», pero que, de todos esos elementos, «el forro exterior, formado por la pizarra dicha, está verdaderamente en muy mal estado»[22]. El acabado de pizarra era posiblemente el del edificio original. Aunque los problemas que tenía la cubierta se habían intentado subsanar con reparaciones parciales, el proyecto de 1933 proponía la sustitución de la pizarra por «placas onduladas de cemento y amianto comprimido» de gran onda, en color gris[23].

En esta fase del trabajo, las fotografías históricas son también clave, pues facilitan mucho la identificación de materiales y, sobre todo, de texturas, es decir, de su apariencia externa y superficial, o bien la realidad de otros edificios sí conservados del mismo arquitecto en la ciudad u otros de la época que, en la

21. Proyecto de Hospicio y casa de maternidad para la ciudad de Lugo, 1876-1888. Localizado en: AHPL, Fondo Concello, signatura: 0442-03, ff. 84r-84v.

22. Proyecto de renovación de cubiertas en el cuartel de las Mercedes, Memoria descriptiva, 1933. Localizado en: IHCM, signatura: 8182.2, f. 1r.

23. Proyecto de renovación de cubiertas en el cuartel de las Mercedes, Memoria descriptiva, 1933. Localizado en: IHCM, signatura: 8182.2, f. 2r.

cuestión de la materialidad, son hoy huella de unos modos de hacer que caracterizaron la arquitectura de aquellos años.

Si volvemos a la cuestión de los acabados de las cubiertas, la consulta de fotografías históricas nos permite añadir alguna información más a la que obtuvimos previamente de las fuentes textuales. Las instantáneas conservadas de finales de siglo dejan ver una cubierta de pizarra, de corte irregular, pero las placas de amianto son las que caracterizan la imagen de la cubierta del edificio en las fotografías conservadas de los años sesenta, antes de su demolición (Figura 11). Hay otro grupo de fotografías que, por su baja resolución o por ser tomadas con cierta distancia, no permiten la identificación del material, lo cual supone una limitación a su uso como fuentes del texturizado histórico.

Los ejemplos presentados creemos que son ilustrativos del trabajo desarrollado para poder recuperar no sólo el tipo de material de cada elemento del edificio, sino también su apariencia externa y superficial. Revelan el profundo estudio y reposado razonamiento que se puede esconder tras la imagen final de una reconstrucción virtual histórica, si se hace con rigor académico (Figura 12).

La mayor dificultad en este punto del trabajo se produjo en el texturizado de los espacios interiores ya que no sólo las fuentes textuales son más parcas en describir sus características, sino que también las fotografías históricas que se conservan son pocas, parciales y posteriores al periodo objeto de estudio. Las más tempranas que hemos localizado datan de 1889, cuando el edificio ya había sufrido varias reformas para su utilización como colegio preparatorio. Como consecuencia, es sobre todo en el texturizado del espacio interior donde aumenta el número de hipótesis de diferente grado. En ciertas ocasiones, conocemos el tipo de material y su textura y estaríamos ante una evidencia, pero en muchas otras aunque el primer dato es conocido, no tenemos información sobre su apariencia externa y la que se plantea en el modelo es, en consecuencia, hipotética. En el peor de los casos las fuentes no revelan ni cuál era el material, ni cuál era la textura y, por tanto,

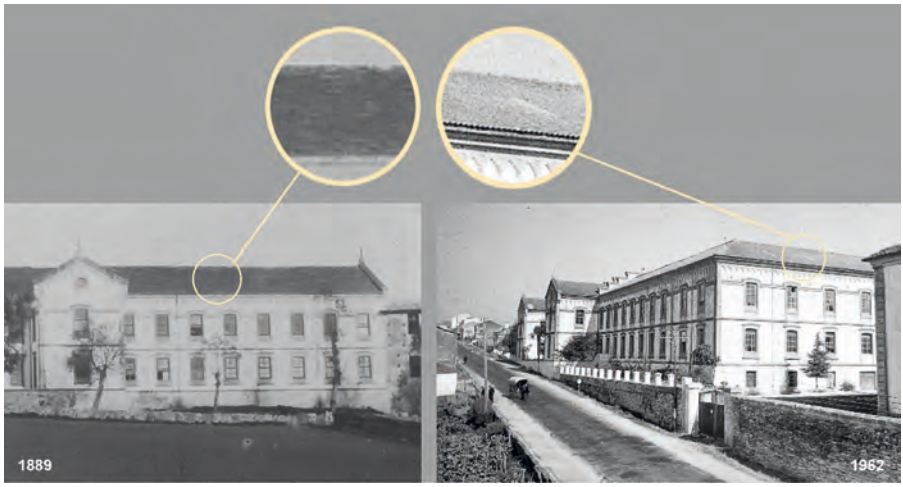



Figura 11. Identificación del material y textura de las cubiertas a partir de fotografías históricas. Fuente: De la fotografía de 1889, IHCM, signatura: F.15310. De la fotografía de 1962, AHPL, Fondo Vega, signatura: 5243-4.




Figura 12. Modelado 3D de los ventanales del Asilo de las Mercedes en 1885. Fuente: Elaboración propia.

no es posible su texturizado. Se pueden plantear materialidades posibles o supuestas, pero en cualquier caso, completamente hipotéticas. Por tanto, al igual que con las dimensiones, formas y elementos compositivos, una reconstrucción virtual también debe incluir una diferenciación entre las evidencias y las hipótesis y sus diferentes niveles de certeza en la representación de las texturas para que el resultado visual haga posible la comprensión del grado de conocimiento que sobre ellas representan (London Chapter, 2009; ICOMOS, 2017).

Tanto los planos rectificados creados como el modelo virtual son editables al estar definidos, en el primer caso, por puntos, líneas y polígonos, y en el segundo, por una gran variedad de primitivas geométricas, entre las que se incluyen mallas poligonales. Al ser editables permiten la integración futura de cambios o la elaboración de nuevas versiones de los planos y modelos históricos, derivados de la localización de nuevas fuentes documentales, así como la creación desde el principio de capas y escalas de color para la diferenciación de evidencias e hipótesis en diferentes grados. A lo anterior hay que añadir que todas sus partes se pueden medir y analizar desde un punto de vista compositivo, para estudiar, por ejemplo, la relación de esta propuesta con otras de su época o incluso integrar el edificio en representaciones pasadas o contemporáneas de la ciudad y estudiar su papel o huella en la definición de la trama urbana actual.

## 5. CONCLUSIONES

Recuperar el conocimiento del patrimonio gallego desaparecido y reactivar su memoria es el objetivo principal del proyecto MEPAGAL. El presente artículo ha intentado explicar el papel que la reconstrucción histórica tridimensional puede jugar en el proceso de investigación de la arquitectura histórica a partir de un caso de estudio concreto.

Una reflexión de conjunto sobre todos los aspectos tratados revela que la producción de un modelo digital, a partir

de fuentes históricas de diferente índole, conduce a un proceso complejo y diferente del de la investigación exclusivamente realizada a través de textos y/o imágenes para producir resultados eminentemente textuales. Cuando intentamos reconstruir de modo visual un espacio perdido nos enfrentamos a múltiples cuestiones sobre su forma, dimensiones y materialidad, entre otras, claves para su materialización física en el mundo real, pero igualmente en el virtual. Al intentar darles respuesta de modo visual adquirimos un mejor conocimiento del edificio histórico, que se produce durante el proceso de recuperar su plano y hacer el modelado.

El constante trabajo comparativo entre las diferentes fuentes disponibles al que conduce la creación de un modelo tridimensional riguroso de un espacio no conservado hizo patente, en primer lugar, cuáles son los datos más relevantes para una verdadera comprensión del espacio histórico desaparecido, entre los que podemos distinguir: los formales, compositivos, dimensionales, funcionales, estructurales, materiales y paisajísticos. En segundo lugar, durante el proceso de hacer el modelo se hicieron visibles tanto las ausencias de ciertos datos en las fuentes disponibles, como las ambigüedades o inconsistencias de estas, así como la importancia de su contraste para validar el grado de confiabilidad de los documentos históricos. En último lugar, el proceso también dejó patente en qué partes son necesarias las hipótesis para producir un modelo final completo, coherente y exacto desde un punto de vista histórico-artístico y, por tanto, en qué aspectos nuestra comprensión del edificio histórico es parcial y como tal debe ser representada.

Cuando trasladamos el rigor académico que caracteriza a los métodos de investigación tradicionales a la creación de un modelo digital tridimensional, el proceso de recuperar el plano y el modelado se convierten en un método de investigación en sí mismos que, por tanto, va más allá del producto visual final. Los planos y el modelo resultantes no sólo son técnicamente más rigurosos, sino mucho más sólidos desde un punto de vista metodológico y,

como consecuencia, creemos que deben ser considerados como una nueva forma de producción de conocimiento que condensa, en sus múltiples elementos gráficos, tanta o más información que la que, sobre aspectos espaciales, se puede explicar con palabras. El trabajo de reconstrucción virtual rigurosa nos permite no sólo hacer aparecer el espacio desaparecido, sino también aumentar el conocimiento que sobre él tenemos antes de iniciar el proceso. Como académicos, en el ejercicio de dibujar y modelar nos sumergimos en un proceso complejo e iterativo de investigación, que permite nuevos descubrimientos en el propio proceso de un hacer, paulatino, lento, hacia un todo cohesivo y riguroso, más allá de la representación digital del espacio histórico.

## REFERENCIAS BIBLIOGRÁFICAS


Abel Vilela, Adolfo de (1996). *Urbanismo y arquitectura en Lugo. Arquitectura isabelina y de la restauración*. A Coruña: Edicións do Castro.

Abel Vilela, Adolfo de (2003). Nemesio Cobreros y Cuevillas. En: Carlos del Pulgar Sabín (ed.). *Artistas galegos arquitectos. Da ilustración ó eclecticismo*. Vigo: Nova Galicia Edicións, pp. 320-344.

Abel Vilela, Adolfo de. 12 de agosto de 2007. Orígenes del asilo de As Mercedes. *La Voz de Galicia*, p. 6. Disponible en: https://www.lavozdegalicia.es/noticia/lugo/2007/08/12/origenes-asilo-as-mercedes/0003_6053904.htm [actualizada 17 agosto 2007; citada 31 diciembre 2021].

Castro Freire, Salvador (1951). *Lugo y sus hombres: ensayo de síntesis histórica*. Lugo: Ediciones Celta Lugo.

De Francesco, Giuliana y D'Andrea, Andrea (2008). Standards and guidelines for quality digital cultural three-dimensional content creation. En: Marinos Ioannides et al. (ed.). *Digital heritage: proceedings of the 14th international conference on virtual systems and multimedia*. Budapest: Archaeolingua, pp. 229-233.

Fernández Fernández, Carlos M. y Sánchez García, Jesús Ángel (2006). Espacios para el refugio. La asistencia a la vejez y los asilos en Galicia. *Sémata. Ciencias sociais e humanidades*, 18: 167-203.

ICOMOS (2017). *Los principios de Sevilla. Principios internacionales de la arqueología virtual*. Disponible en: https://icomos.es/wp-content/uploads/2020/06/Seville-Principles-IN-ES-FR.pdf [actualizada 6 junio 2020; citada 27 diciembre 2021].

London Chapter (2009). La carta de Londres para la visualización computarizada del patrimonio cultural. Disponible en: https://www.londoncharter.org/downloads.html [actualizada 7 febrero 2009; citada 27 diciembre 2021].

Prado Gómez, Antonio y Reboredo Pazos, Julio (2001). El desarrollo urbano de Lugo en el siglo XIX. *CROA: boletín da asociación de amigos do museo do Castro de Viladonga*, 11: 45-62.

Münster, Sander; Friedrichs, Kristina y Hegel, Wolfgang (2018). 3D reconstruction techniques as a cultural shift in art history? *International Journal for Digital Art History. Digital Space and Architecture*, 3: 38-60, https://doi.org/10.11588/dah.2018.3.32473

# HUMANIDADES DIGITALES Y ARQUEOLOGÍA:
## EXPERIMENTANDO CON MODELOS DE RECONOCIMIENTO AUTOMÁTICO DE TEXTO MANUSCRITO EN ESPAÑOL SOBRE PATRIMONIO CULTURAL DE INICIOS DEL SIGLO XX

Raquel Liceras-Garrido, Alba Comino

**Resumen**

Los documentos en formato legible por ordenador son cruciales para explorar textos sobre Arqueología desde enfoques de análisis de datos. En la última década, han proliferado los proyectos de digitalización masiva de documentos arqueológicos e históricos, produciendo grandes colecciones. Hasta el momento, estas colecciones se encontraban principalmente accesibles en formato de imagen, disponibles solo para llevar a cabo análisis manuales que requerían enormes inversiones de tiempo e impedían enfoques computacionales. Recientemente, el uso de algoritmos de aprendizaje automático ha logrado una alta precisión en la producción de transcripciones automáticas a partir de archivos de imagen. Tecnologías como el reconocimiento automático de texto manuscrito permiten la creación de modelos precisos para producir transcripciones automáticas a partir de imágenes utilizando arquitecturas de redes neuronales. Este estudio ha seleccionado tres volúmenes manuscritos

Raquel Liceras-Garrido, Departamento de Prehistoria, Historia Antigua y Arqueología, Universidad Complutense de Madrid
https://orcid.org/0000-0002-5552-9273
rliceras@ucm.es

Alba Comino, Institute of Contemporary History, NOVA School of Social Sciences and Humanities / IN2PAST — MSCA Nº 101063220
albacomino@fcsh.unl.pt

del *Catálogo Monumental de España*, elaborados a principios del siglo XX, para probar el potencial de estas técnicas de Inteligencia Artificial en documentos arqueológicos. Se ha utilizado el software Traskribus obteniendo una precisión entre el 91% y 95% en el reconocimiento automático de caracteres.

**Palabras clave:** Aprendizaje Automático, Inteligencia Artificial, Reconocimiento de texto manuscrito, Arqueología, Prehistoria

**Abstract**
Documents in a machine-readable format are crucial for exploring texts on Archaeology from big data approaches. In the last decade, massive digitisation projects of archaeological and historical documents have proliferated, producing large collections. Until now, these collections were mainly accessible in image format, available only for manual analyses that required huge time investments and precluded computational approaches. Recently, the use of Machine Learning algorithms has achieved high accuracy in the production of automatic transcriptions from image files. Using neural network architectures, handwritten text recognition allows the creation of accurate models to produce automatic transcriptions from input images. This study has selected three handwritten volumes of the Monumental Catalogue of Spain, produced in the early twentieth century, to test the potential of these Artificial Intelligence techniques on archaeological documents. We have used the software Transkribus, obtaining an accuracy between 91% and 95% in the automatic recognition of characters.

**Keywords:** Machine Learning, Artificial Intelligence, Handwritten text recognition, Archaeology, Prehistory

## 1. INTRODUCCIÓN

Hace más de un siglo que la Arqueología se convirtió en una disciplina científica. Desde sus inicios, ha liderado la colaboración entre las Humanidades y las Ciencias de la Computación. Sin embargo, la aplicación potencial de técnicas computacionales a recursos textuales sobre Arqueología durante este tiempo permanece, en gran parte, inexplorada. Asimismo, la utilización de métodos y análisis de la Lingüística Computacional, el Procesamiento del Lenguaje Natural o la Minería de Textos es aún extremadamente escasa (Amrani *et al.*, 2008; Murrieta-Flores y Gregory, 2015; Richards, Tudhope y Vlachidis, 2015; Liceras-Garrido *et al.*, 2019; Liceras-Garrido, Comino y Murrieta-Flores 2022). Para implementar estos análisis, es imprescindible disponer de texto en formato legible por ordenador. En las últimas décadas, se han llevado a cabo grandes proyectos de digitalización por parte de bibliotecas, archivos y museos, escaneando en formato de imagen y compartiendo abiertamente una gran cantidad de datos como, por ejemplo, publicaciones arqueológicas, informes o correspondencia postal, que posibilitan enfoques de análisis de *big data*.

Dentro de la Inteligencia Artificial, el Aprendizaje Profundo ha evolucionado como una categoría avanzada del Aprendizaje Automático, impulsando mejoras sustanciales en el campo de la Visión Artificial y desarrollando métodos innovadores para la identificación y clasificación de imágenes, desde artefactos hasta paisajes (Van der Maaten *et al.*, 2006; Hörr, Lindinger y Brunnett, 2014; Verschoof-van der Vaart y Lambers, 2019; Domínguez-Rodrigo *et al.*, 2020). Las redes neuronales son un componente fundamental de cualquier arquitectura de Aprendizaje Profundo. Aunque existe una amplia gama de redes neuronales, para este trabajo son relevantes dos: 1) redes neuronales recurrentes, que usan capas apiladas de redes neuronales para aprender y procesar secuencias de entrada de longitud variable (Tealab, 2018), implementadas en gran parte para reconocimiento de voz y reconocimiento de texto manuscrito (HTR a partir de ahora por las siglas en inglés: *Handwritten Text*

*Recognition*); y 2) redes neuronales convolucionales, que procesan grandes matrices estructuradas de datos, ampliamente utilizadas para el reconocimiento y clasificación de imágenes y, también, el procesamiento del lenguaje natural (Krizhevsky, Sutskever y Hinton, 2013). Sobre la base de estas arquitecturas, el HTR es una tecnología que dota al ordenador de la capacidad de interpretar datos de entrada que contienen texto manuscrito en formato de imágenes y producir texto codificado de alta calidad inteligible para computadoras (Govindaraju, Srihari y Srihari, 1995; Romero, Toselli y Vidal, 2012).

La técnica de reconocimiento óptico de caracteres (OCR a partir de ahora por su acrónimo en lengua inglesa *Optical Character Recognition*) es un caso similar, pero enfocado a imágenes de texto impreso. Esta compara los caracteres de la imagen de entrada con plantillas existentes utilizando una amplia gama de diccionarios de idiomas (Govindan y Shivaprasad, 1990; Kichuk, 2015). Cuando los resultados obtenidos por OCR son de baja calidad debido a la transcripción automática de imágenes de baja resolución como, por ejemplo, escaneos de periódicos antiguos, es posible utilizar técnicas de HTR en textos impresos para generar versiones legibles por ordenador de mejor calidad (Ströbel, Clematide y Volk, 2020).

En los últimos años han proliferado numerosos softwares de HTR, tanto de acceso abierto como propietario. Entre ellos destaca Transkribus, una herramienta en línea ampliamente utilizada para la investigación. Fue desarrollada por los proyectos europeos "*tranScriptorium*" (2013-2015) y "*Recognition and Enrichment of Archival Documents* (READ)" (2016-2019) mediante los que se creó y perfeccionó una plataforma de acceso abierto, aunque recientemente se ha comercializado para garantizar su continuidad. Transkribus utiliza una combinación de redes neuronales recurrentes y redes neuronales convolucionales para reconocer la escritura manuscrita utilizando una cantidad limitada de datos de entrenamiento (Kahle *et al.*, 2017; Weidemann *et al.,* 2017; Michael, Weidemann y Labahn, 2018; Müehlberger *et al.*, 2019).

Este trabajo utiliza la tecnología HTR integrada en Transkribus[1] para explorar las posibilidades de producir transcripciones automatizadas de tres volúmenes manuscritos del *Catálogo Monumental de España* (CME en adelante), creado por las Comisiones Provinciales de Monumentos a principios del siglo XX. Este catálogo es una fuente fundamental en español para disciplinas como la Arqueología, la Historia, el Patrimonio, las Artes y la Arquitectura, entre otras, porque enumera y describe los principales edificios, monumentos y antigüedades conocidos hasta entonces, incluyendo datos únicos para la investigación arqueológica. Así, se ha experimentado con los volúmenes escritos por Manuel Gómez-Moreno Martínez para la provincia de Ávila, Juan Cabré Aguiló para Soria y Narciso Sentenach y Cabañas para Burgos (Figura 1). El corpus inicial es una versión en imagen digital de los manuscritos, accesible a través de la Biblioteca Virtual del Consejo Superior de Investigaciones Científicas (CSIC), libre de derechos de autor. El corpus comprende tres volúmenes y cada volumen contiene varios libros:

— Manuel Gómez-Moreno (1900-1901). *Catálogo monumental de España: provincia de Ávila*: tres libros.
— Juan Cabré (1916-1917). *Catálogo monumental de la provincia de Soria*: ocho libros.
— Narciso Sentenach (1924). *Catálogo monumental y artístico de la provincia de Burgos*: siete libros.

La elección de estos tres catálogos está relacionada con nuestra línea de investigación sobre la Edad del Hierro, pero también con la importante implicación de estos tres autores en el desarrollo de la Arqueología como disciplina científica en España. Eran humanistas, interesados por la Historia, la Literatura, las Artes y la Arqueología,

1. Transkribus: https://readcoop.eu/transkribus/ [último acceso, 29 de agosto 2022].

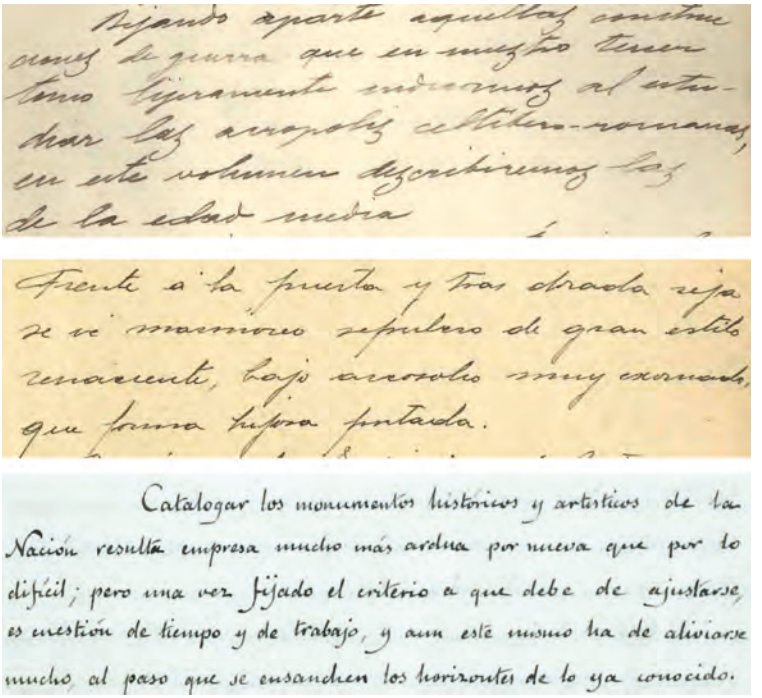


Figura 1. Muestras de escritura manuscrita de los Catálogos Monumentales. De arriba a abajo: Cabré, Sentenach y Gómez-Moreno.

que compartían vínculos con la Real Academia de Bellas Artes de San Fernando (RABASF). La RABASF fue creada en 1752 y sigue siendo una institución que busca promover la creación artística, así como el estudio, difusión y protección de las artes y el patrimonio cultural. Las trayectorias profesionales de Narciso Sentenach (1853-1925) (Díaz-Andreu, 2009; López-Yarto Elizalde, 2012), Manuel Gómez-Moreno Martínez (1870-1970) (Pasamar Alzuria, 2009a; López-Yarto Elizalde, 2012) y Juan Cabré (1882-1947) (Pasamar Alzuria, 2009b; López-Yarto Elizalde, 2012) son de gran interés, ya que fueron figuras de transición entre la práctica *amateur* y la Arqueología científica, surgida en España debido al proceso de institucionalización y profesionalización de la disciplina durante el primer tercio del siglo XX.

Este artículo está organizado en cuatro apartados. En el apartado 2, se abordan los antecedentes históricos del CME, desde su creación hasta la digitalización del corpus. El apartado 3 proporciona una breve introducción a la tecnología HTR, incluidas las pautas diseñadas específicamente para este proyecto. Esta sección consta de tres subapartados. En el primero se detalla el experimento realizado para transcribir las letras manuscritas de Cabré y Sentenach calculando seis modelos HTR. En el segundo se expone la utilización de los modelos creados en la sección anterior para transcribir una escritura diferente. En el tercero, se hace una evaluación general del número de errores entre los caracteres producidos por el ordenador y las revisiones manuales de las transcripciones. En el apartado 4 se analizan los desafíos a los que se ha enfrentado este trabajo. Finalmente, el apartado 5 presenta unas conclusiones.

## 2. EL CATÁLOGO MONUMENTAL DE ESPAÑA

A principio del siglo XX, siguiendo las tendencias europeas de creación de catálogos científicos, el Real Decreto de 1 de junio de 1900 proyectó la elaboración de los CME para las cuarenta y siete provincias españolas, aunque solamente se llegaron a publicar diecisiete. El CME tuvo como objetivo: a) recopilar información patrimonial relevante para las disciplinas histórico-artísticas; b) catalogar sistemáticamente cada monumento, obra de arte y yacimiento arqueológico; c) proteger el patrimonio; y d) crear una publicación que recogiera todo el patrimonio español (Muñoz Cosme, 2012).

El director de la RABASF, Juan Facundo Riaño, fue el encargado de diseñar la estrategia e impulsar la ejecución de los trabajos de recopilación e investigación del patrimonio histórico-artístico de España. Inicialmente, encargó a Manuel Gómez-Moreno la realización de la colección completa. Sin embargo, dada la gran envergadura del proyecto y las críticas por parte de la autoría exclusiva, se dictó un segundo Real Decreto en 1902, que modificaba el anterior para ampliar el número de estudiosos necesario para su

compilación. De esta forma, lo que se podía perder en rigor se ganó en velocidad (Morales, 1996; López-Yarto Elizalde, 2012; López-Ocón Cabrera, 2012). La RABASF fue la encargada de seleccionar a los profesionales que elaboraron el CME. Sin embargo, esta no fue una tarea fácil debido principalmente a dos razones. Por un lado, en la España de principios del siglo XX no había suficientes profesionales disponibles para trabajar simultáneamente sobre el tema. Por otro, la retribución económica y el prestigio de la obra atrajeron a personas sin formación suficiente, pero con amigos influyentes (Hidalgo Brinquis, 2012). Además, como el CME pretendía tratar un tema tan amplio como el patrimonio en todas sus formas, los autores enfatizaron sus propias áreas de conocimiento, dejando de lado aquellos aspectos en los que estaban poco o nada versados.

Si bien en muchas ocasiones las mujeres se involucraron en la edición final de algunos volúmenes del CME como, por ejemplo, Elena Rodríguez-Bolívar, esposa de Gómez-Moreno (Hidalgo Brinquis, 2012; Lorenz Arribas y Pérez Martín, 2017), no fue hasta 1930 cuando las mujeres comenzaron a integrarse en las instituciones patrimoniales y arqueológicas (Díaz- Andreu y Sanz-Gallego, 1994). Así, veintitrés hombres fueron los encargados de completar el CME. El catálogo resultante se configuró entre 1901 (Ávila, Gómez-Moreno) y 1924 (Burgos, Sentenach) (Hidalgo Brinquis, 2012). No obstante, aunque uno de los objetivos era establecer un registro estandarizado, la gran cantidad de autores involucrados, las diferencias de conocimiento y la falta de criterios comunes, dieron como resultado una recopilación desigual de información que dependía de los intereses y la formación de cada autor. Esto es evidente en los catálogos analizados en este artículo.

Los volúmenes del CME ofrecen descripciones detalladas e información sobre la historia del patrimonio arqueológico e histórico español conservado a comienzos del siglo XX, además de una extensa colección de fotografías, dibujos y bocetos de objetos y edificios desaparecidos o transformados. Esta colección reúne aproximaciones teóricas, interpretaciones sobre la cultura material descrita y

extractos sobre las experiencias de los autores durante la elaboración del catálogo. Esta variedad de datos convierte al corpus en una fuente valiosa no solo para evaluar el estado del patrimonio en ese momento, sino también para acercarse a la historia de la Arqueología.

Gracias al convenio firmado en 2008 entre el Consejo Superior de Investigaciones Científicas (CSIC) y el Instituto Español de Patrimonio Cultural (IPCE), el proyecto "El Catálogo Monumental de España (1900-1961)" digitalizó y puso a disposición de investigadores, profesionales y público más amplio esta importante información (Albert de León, 2012). El proceso de digitalización fue realizado por el IPCE, y el almacenamiento y mantenimiento de datos por el CSIC. La realización de este proyecto permitió no sólo la preservación de esta fuente de conocimiento, sino también cumplir con el objetivo original del CME de difundir a la sociedad una publicación que recopilara todo el patrimonio histórico-artístico de España.

## 3. CREACIÓN DE LOS DATOS DE ENTRENAMIENTO

Como se ha mencionado anteriormente, este trabajo se realizó utilizando la plataforma Transkribus para producir texto legible por computadora de los catálogos de Ávila, Soria y Burgos, configurando así un corpus de dieciocho libros. La elección de dicha herramienta se debió principalmente a que era de acceso abierto y permitía a diferentes perfiles profesionales y académicos crear y compartir datos en abierto.

Se diseñaron dos estrategias para optimizar el proceso de transcripción. En primer lugar, se crearon dos modelos HTR personalizados para los textos manuscritos de Cabré y Sentenach, ya que eran los conjuntos de datos *a priori* más amplios. En segundo lugar, se desarrolló una propuesta que buscaba combinar los modelos creados en la fase anterior y transcribir automáticamente los libros producidos por Gómez-Moreno. El objetivo de esta segunda estrategia era probar la eficacia de modelos creados con datos de entrenamiento

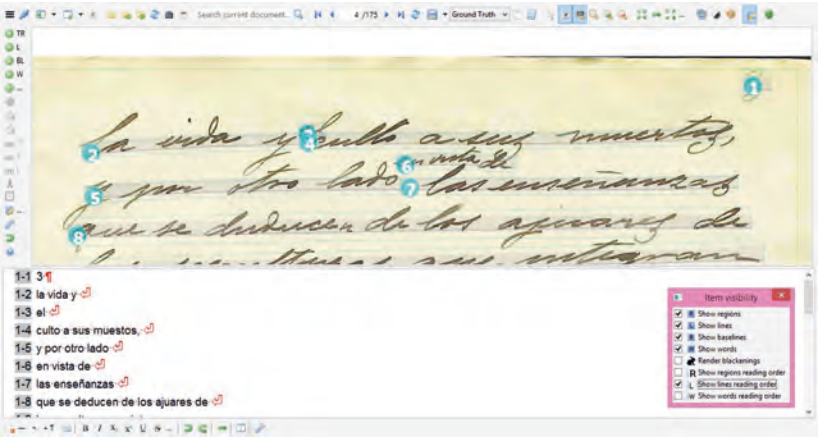



Figura 2. Ver la corrección que hizo el autor introduciendo una línea (6) sobre otra (7) (Cabré, libro 3).

de diferentes escrituras y evaluar su precisión a la hora de producir transcripciones automáticas de caligrafías "no vistas" previamente.

Como nuestra investigación se centraba en la Arqueología Prehistórica, especialmente en la Edad del Hierro (I milenio a.C.), comenzamos generando el conjunto de datos de entrenamiento con las transcripciones manuales del libro 3 de Cabré y el libro 1 de Sentenach. Para ello, el primer paso fue identificar las regiones con texto dentro de cada página y segmentarlas en líneas de base, usando una arquitectura CITlab (redes neuronales recurrentes) (Strauß et al., 2016; Michael *et al.*, 2018). El resultado fue relativamente preciso, identificando generalmente un cuerpo de texto continuo (región) por página en los documentos de Cabré y dos en los manuscritos de Sentenach, ya que marcaba como región independiente el número de página situado en la esquina superior. El mayor desafío fue mantener el orden de lectura de las líneas porque los autores introdujeron constantes correcciones, añadidos y modificaciones entre las líneas de texto (ver Figura 2).

Los algoritmos de aprendizaje profundo requieren grandes cantidades de datos para aprender y generalizar. Por tanto, una segunda etapa consistió en transcribir manualmente las 15.000

palabras recomendadas para entrenar un modelo HTR (Proyecto READ, 2019). Dado que la consistencia y la estandarización eran claves en el proceso de creación de datos de entrenamiento, fue necesario establecer algunas pautas o instrucciones que recogiesen la complejidad de esta fuente:

1. Transcribir con la mayor precisión posible lo que escribe el autor, incluyendo las faltas de ortografía, acentos, errores gramaticales, etc.
2. Agregar un espacio al final de cada línea, teniendo en cuenta las oraciones terminadas en guion cuando la última palabra está dividida en dos líneas.
3. Cuando el texto esté subrayado, mantener las palabras, letras y oraciones con la mayor precisión posible al original. Sin embargo, debido a las limitaciones del software, no diferenciamos si el subrayado era simple, doble u ondulado.
4. Marcar el texto tachado con la opción de tachado. En muchas ocasiones, transcribir el texto detrás de los tachones fue un desafío. Para minimizar el impacto de esto, se creó un registro con los fragmentos de texto problemáticos que luego fueron revisados por una segunda transcriptora.
5. Transcribir las leyendas de las imágenes en su posición original cuando dividen el cuerpo del texto principal de la página. Si no hay saltos en el texto, las leyendas se transcriben al final.
6. Los pies de página se clasifican como tales añadiéndose a los metadatos.
7. Las fotografías se marcan como imágenes.
8. Los mapas, planos, bocetos, etc. se marcan como dibujos de línea/bocetos.

Seguir estas directrices incrementó el tiempo necesario para generar los datos de entrenamiento, aunque mejoró la siempre necesaria estandarización. Observamos, por ejemplo, cómo la presencia continua de imágenes, dibujos y texto hacía el proceso de transcripción



Figura 3. Ejemplo de texto intercalado con fotografías y dibujos (Sentenach, libro 4).

manual más lento (Figura 3). No obstante, las principales dificultades que encontramos en este proceso fueron la interpretación de palabras ilegibles, abreviaturas, tachaduras y faltas de ortografía.

Una vez terminadas las transcripciones manuales del libro 1 de la provincia de Burgos de Sentenach el resultado era de 24.424 palabras (206 páginas), rebasando así la recomendación de 15.000 palabras de una escritura nítida para entrenar un modelo HTR en Transkribus (Proyecto READ, 2019). Sin embargo, la transcripción del libro 3 de Soria de Cabré solo constaba de 13.941 palabras (175 páginas), por lo que fue necesario incluir la transcripción de 58 páginas adicionales del libro 4, llegando a las 18.271 palabras (233 páginas) para este segundo autor.

### 3.1. EXPERIMENTO 1: ENTRENAMIENTO HTR PARA LA LETRA MANUSCRITA DE CABRÉ Y SENTENACH

El modelo de aprendizaje profundo que utiliza Transkribus en la creación del modelo HTR son redes neuronales convolucionales anidadas y memoria bidireccional a largo plazo para simular las variaciones de la escritura manuscrita. Para ello, usa como datos

de entrada la imagen de la línea de texto extraída por el algoritmo de detección de la línea de base mencionado en el apartado anterior (Michael *et al.*, 2018). El entrenamiento de un modelo HTR requiere introducir los datos transcritos manualmente en dos conjuntos de datos principales, uno para entrenar el modelo de reconocimiento de texto y otro como un conjunto de validación para evaluar su rendimiento. Una vez introducidos, en nuestro caso, el tiempo de cálculo para cada modelo fue de unas siete horas.

La precisión del modelo resultante se mide mediante la ratio de caracteres erróneos (CER, acrónimo en inglés de *Character Error Rate*) que calcula el porcentaje de caracteres que el modelo ha transcrito incorrectamente. Esta ratio se calculó en los dos conjuntos de datos: entrenamiento y validación. El conjunto de validación se considera más relevante, ya que muestra cómo funciona el nuevo modelo en texto que no se ha utilizado en el entrenamiento y compara los resultados con las transcripciones manuales proporcionadas (Proyecto READ, 2019). De este modo, el porcentaje máximo de la CER para que el modelo HTR sea aceptable para realizar búsquedas de palabras clave es del 20%-30%, mientras que una ratio inferior al 10% permite realizar transcripciones automáticas eficientes para análisis de Lingüística de Corpus (Proyecto READ, 2019). En nuestro proyecto, los resultados de esta primera iteración se encuentran en la Tabla 1, mostrando valores ligeramente superiores al 5% en el conjunto de validación. Esto significó que casi un 95% de los caracteres fueron transcritos automáticamente de modo correcto (Cabré: 94,59%, Sentenach: 94,88%).

**Tabla 1. Resultados de la primera iteración.**

| MODELO HTR | CER EN LOS DATOS DE ENTRENAMIENTO | CER EN LOS DATOS DE VALIDACIÓN | Nº DE PALABRAS | DATOS USADOS PARA CREAR EL MODELO HTR |
|---|---|---|---|---|
| Cabré_v1 | 2,21% | 5,41% | 18.271 | – Transcripción manual del libro 3 y parte del libro 4 |
| Sentenach_v1 | 2,20% | 5,12% | 24.424 | – Transcripción manual del libro 1 |

El proceso de aprendizaje de los modelos HTR fue inicialmente muy intenso, aunque descendió drásticamente tras una serie de convoluciones, tal y como se puede ver en el gráfico de la Figura 4. El eje vertical representa el porcentaje de precisión por cada valor de la CER, mientras que el eje horizontal muestra el rendimiento del modelo a lo largo de las diferentes iteraciones en las que se evalúan los datos de entrenamiento (*epoch*). De este modo, la línea azul representa los valores de la CER en el conjunto de datos de entrenamiento y la línea roja se refiere a los datos de validación. Finalmente, la línea verde marca el punto en el que el modelo HTR presenta el funcionamiento óptimo y se guarda.

Los modelos HTR resultantes fueron utilizados para transcribir automáticamente el libro 5 de Cabré y el libro 2 de Sentenach. A pesar de los prometedores valores de la CER, los documentos resultantes nos ayudaron a identificar la necesidad de combinar estos primeros modelos HTR con un diccionario. Los diccionarios ofrecen una normalización ligera que ayuda a la máquina a reconocer la ortografía de ciertas palabras. Esto fue significativo debido a la confusión entre letras manuscritas como 'u' y 'n' (con errores tipográficos como 'gnerra' por 'guerra'). Para la elaboración del diccionario, se utilizó una publicación contemporánea sobre las excavaciones del yacimiento arqueológico de Numancia (Mélida, 1912) que recopilaba vocabulario especializado sobre Arqueología de la época, compuesto por un total de 33.363 palabras. Dicho diccionario fue implementado en la plataforma bajo el nombre "Spanish_20th_Century_v1.dict", disponible para todos los usuarios interesados en transcribir textos en español de principios del siglo XX relacionados con Arqueología o Patrimonio.

Tras la transcripción automática y revisión manual de los dos libros anteriores, obtuvimos un total de 27.483 palabras para Cabré y 42.873 para Sentenach en formato de texto legible por ordenador. Una vez llevado a cabo el entrenamiento y validación del modelo, realizamos un test independiente del proceso de construcción y ajuste de los modelos para evaluar el error real cometido por los

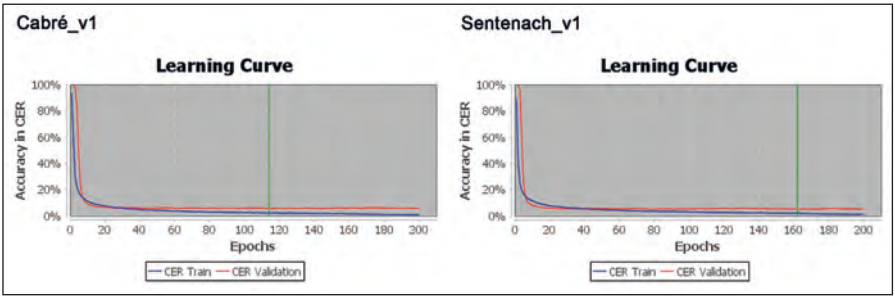



Figura 4. Curvas de aprendizaje en la primera iteración.

mismos. Para ello, se calculó manualmente el número total de modificaciones realizadas por documento (revisiones de caracteres), utilizando la opción "comparar" incluida en Microsoft Word 2021, con el fin de cotejar la diferencia de porcentajes de cambios entre las transcripciones automáticas y las versiones revisadas manualmente. Esta herramienta contabiliza y destaca cuándo se ha eliminado, modificado o insertado un carácter o palabra, considerando cada transcripción automática como un documento original y la versión revisada como un documento actualizado (Tabla 2). De este modo, se puede observar que la precisión de los modelos creados en la primera iteración disminuye ligeramente en la evaluación externa respecto a la CER en los datos de validación, con una eficiencia de en torno al 92% (Cabré: 92,65%, Sentenach: 92,85%).

**Tabla 2. Test de la primera iteración.**

| DOCUMENTO | NÚMERO DE CARACTERES | | CARACTERES REVISADOS | DIFERENCIA DE CARACTERES (%) | MODELO DE ENTRENAMIENTO HTR | CER EN CONJUNTO DE VALIDACIÓN (%) | PÁGS. |
|---|---|---|---|---|---|---|---|
| | TRANSCRIPCIÓN AUTOMÁTICA | TRANSCRIPCIÓN REVISADA | | | | | |
| Cabré libro 5 | 52501 | 53214 | 3858 | 7.35 | Cabré_v1 | 5.41 | 121 |
| Sentenach libro 2 | 96242 | 96747 | 6882 | 7.15 | Sentenach_v1 | 5.12 | 151 |

En este punto, quisimos testar dos elementos principales: 1) si era posible mejorar la precisión del modelo HTR creando desde cero un nuevo modelo a partir de un conjunto mayor de palabras

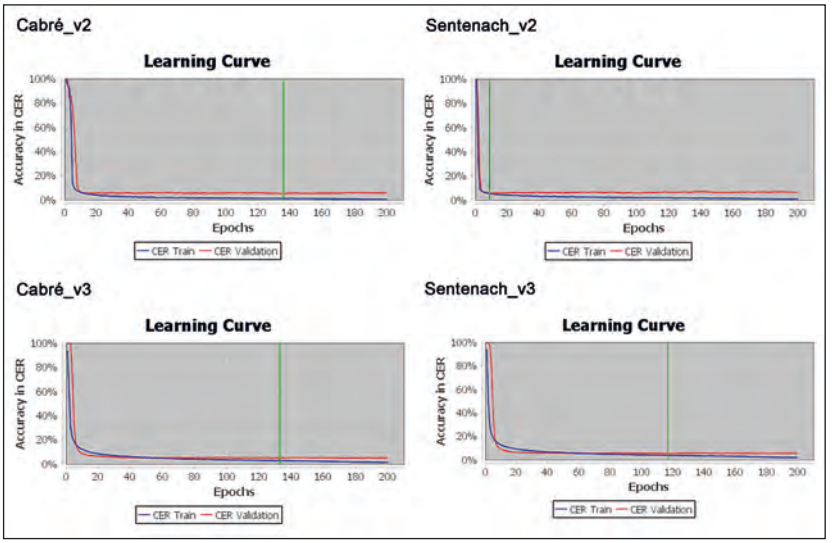



Figura 5. Curvas de aprendizaje.

y 2) probar el rendimiento de los modelos base que ofrecía esta plataforma. Esta segunda propuesta nos permitía utilizar los modelos de reconocimiento de texto Cabré_v1 y Sentenach_v1 como modelos base para el entrenamiento e integrar datos nuevos para crear otro modelo. Esta opción, además, permitía reducir el tiempo de computación en el cálculo de los modelos y requería una muestra de entrenamiento menor.

Durante esta segunda iteración, no observamos cambios significativos en los porcentajes de precisión más allá de un número reducido de fluctuaciones, particularmente en la CER de los datos de entrenamiento (Tabla 3). Las curvas de aprendizaje presentaron resultados similares y la misma tendencia que las anteriores (Figura 5). Frente a estos datos, surgieron dos conclusiones principales: a) no existen diferencias sustanciales en la CER utilizando un modelo base o creando un modelo HTR desde cero; y b) aumentar el número de palabras por encima del umbral de 15.000 palabras no ofrece mejoras significativas en la CER.

**Tabla 3. Resumen de los modelos HTR creados hasta el momento, incluyendo la CER en el conjunto de datos de entrenamiento, los datos de validación, número de palabras y datos de entrada.**

| MODELO HTR | CER EN LOS DATOS DE ENTRENAMIENTO | CER EN LOS DATOS DE VALIDACIÓN | Nº DE PALABRAS | DATOS USADOS PARA CREAR EL MODELO HTR |
|---|---|---|---|---|
| Cabré_v1 | 2.21% | 5.41% | 18.271 | – Transcripción manual del libro 3 y parte del libro 4 |
| Cabré_v2 | 0.98% | 5.03% | 9.212 | – Modelo base: Cabré_v1<br>– Transcripción automática + correcciones manuales del libro 5 |
| Cabré_v3 | 2.68% | 4.75% | 27.483 | – Transcripción manual del libro 3 y parte del libro 4<br>– Transcripción automática + correcciones manuales del libro 5 |
| Sentenach_v1 | 2.20% | 5.12% | 24.424 | – Transcripción manual del libro 1 |
| Sentenach_v2 | 5.32% | 5.50% | 18.460 | – Modelo base: Sentenach_v1<br>– Transcripción automática + correcciones manuales del libro 2 |
| Sentenach_v3 | 3.64% | 5.15% | 42.873 | – Transcripción manual del libro 1<br>– Transcripción automática + correcciones manuales del libro 2 |

Los volúmenes restantes fueron transcritos con los modelos HTR creados a partir de un mayor número de palabras, ya que presentaban la transcripción automática más óptima. Así, los libros de Juan Cabré fueron transcritos con el modelo Cabré_v3 que coincide con la CER más baja en los datos de validación de este autor. Sin embargo, los libros restantes de Narciso Sentenach fueron transcritos con el modelo Sentenach_v3 que, si bien no es el modelo

con la CER más baja, es el modelo con mayor número de palabras. Además, la diferencia entre este y el que tiene una CER menor para este autor no es significativa (Tabla 3). En este punto, asumimos que aquellos modelos entrenados con un mayor número de palabras serían capaces de realizar transcripciones más eficientes, ya que recogían una mayor variabilidad en los caracteres manuscritos menos frecuentes como eran las letras mayúsculas o los números.

Nuevamente, testamos la eficacia de los modelos con el procedimiento de evaluación externa mencionado anteriormente. Los resultados ofrecieron un resultado dispar (Tabla 4). El modelo para Sentenach_v3 arrojó resultados muy similares a Sentenach_v1 con una precisión estable entre el 92,25% y 92,77%. Por el contrario, el modelo Cabré_v3 mostró una variación mayor en cuanto a la precisión del mismo, oscilando entre los 91,21% y los 95,39% de caracteres transcritos correctamente. En la revisión manual del texto, no pudimos percibir ningún cambio significativo que justificase la ejecución diferencial del algoritmo.

**Tabla 4. Test de la segunda iteración.**

| DOCUMENTO | NÚMERO DE CARACTERES | | CARACTERES REVISADOS | DIFERENCIA DE CARACTERES (%) | MODELO DE ENTRENAMIENTO HTR | CER EN CONJUNTO DE VALIDACIÓN (%) | PÁGS. |
|---|---|---|---|---|---|---|---|
| | TRANSCRIPCIÓN AUTOMÁTICA | TRANSCRIPCIÓN REVISADA | | | | | |
| Cabré libro 1 | 27966 | 27868 | 2437 | 8.71 | Cabré_v3 | 4.75 | 100 |
| Cabré libro 2 | 83147 | 77176 | 6041 | 7.27 | Cabré_v3 | 4.75 | 201 |
| Cabré libro 4 | 162579 | 157056 | 9808 | 6.03 | Cabré_v3 | 4.75 | 217 |
| Cabré libro 6 | 95792 | 95379 | 4417 | 4.61 | Cabré_v3 | 4.75 | 220 |
| Cabré libro 7 | 83066 | 82424 | 5609 | 6.75 | Cabré_v3 | 4.75 | 226 |
| Cabré libro 8 | 69225 | 68678 | 6084 | 8.79 | Cabré_v3 | 4.75 | 182 |
| Sentenach libro 3 | 142442 | 141143 | 10465 | 7.35 | Sentenach_v3 | 5.15 | 237 |
| Sentenach libro 4 | 153910 | 152316 | 11132 | 7.23 | Sentenach_v3 | 5.15 | 253 |
| Sentenach libro 5 | 136143 | 133114 | 10547 | 7.75 | Sentenach_v3 | 5.15 | 208 |
| Sentenach libro 6 | 120150 | 117124 | 8812 | 7.33 | Sentenach_v3 | 5.15 | 195 |
| Sentenach libro 7 | 150371 | 149459 | 11378 | 7.57 | Sentenach_v3 | 5.15 | 260 |

## 3.2. EXPERIMENTO 2: TRANSCRIPCIÓN AUTOMÁTICA DE UNA ESCRITURA DIFERENTE (GÓMEZ-MORENO)

Los tres libros escritos por Manuel Gómez-Moreno para la provincia de Ávila fueron los elegidos para comprobar si era posible transcribir con precisión la caligrafía de un tercer autor no incluido en los datos de entrenamiento. Para ello, se creó un modelo HTR que combinaba las caligrafías de Juan Cabré y Narciso Sentenach, en concreto los mismos datos de entrenamiento y validación usados para calcular los modelos Cabré_v3 y Sentenach_v3. El nuevo modelo fue denominado "*Early20thcentury*" y constaba de 70.356 palabras en total. La precisión de dicho modelo fue de un 95,3 % (Tabla 5), cuya curva de aprendizaje mostraba la misma tendencia que los modelos anteriores (Figura 6).

**Tabla 5. Datos del modelo HTR de principios del siglo XX (*Early20thcentury*).**

| MODELO HTR | CER EN LOS DATOS DE ENTRENAMIENTO | CER EN LOS DATOS DE VALIDACIÓN | Nº DE PALABRAS | DATOS USADOS PARA CREAR EL MODELO HTR |
|---|---|---|---|---|
| Early20thcentury | 4,12% | 4,97% | 70.356 | Cabré:<br>– Transcripción manual del libro 3 y parte del libro 4<br>– Transcripción automática + correcciones manuales del libro 5<br><br>Sentenach:<br>– Transcripción manual del libro 1<br>– Transcripción automática + correcciones manuales del libro 2 |

Posteriormente, se utilizó dicho modelo para transcribir el primero de los libros de Gómez-Moreno, utilizando el diccionario Spanish_20th_Century_v1 con el fin de mejorar el reconocimiento de caracteres. Sin embargo, a pesar de que los resultados de la CER fueron prometedores, las caligrafías de Cabré y Sentenach que entrenaron

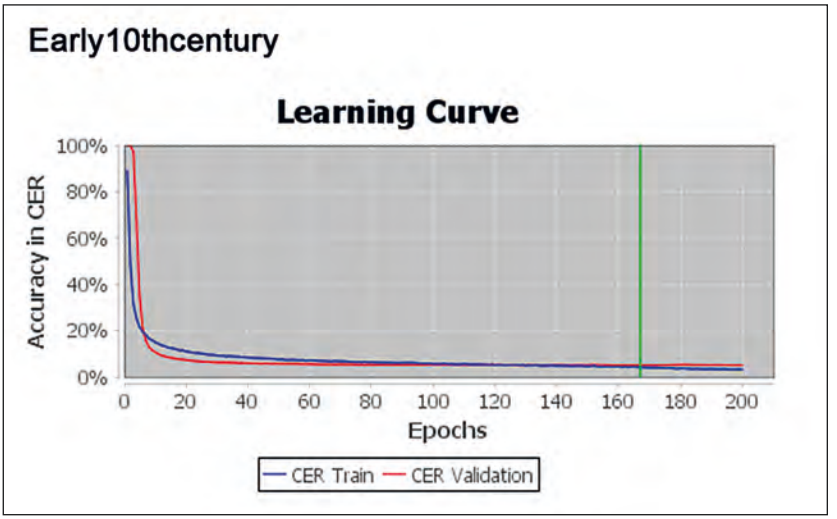



Figura 6: Curva de aprendizaje del modelo de principios del siglo XX.

el modelo diferían considerablemente (ver la Figura 1 para una evaluación visual) y eran particularmente problemáticas en el trazo de las letras mayúsculas y los números. Ante la problemática expuesta, y con el objetivo de mejorar el rendimiento en la transcripción de la escritura de Gómez-Moreno, se revisaron y corrigieron manualmente las transcripciones automáticas de las primeras 50 páginas del libro 1.

Posteriormente, se calculó un modelo HTR personalizado para la escritura del catálogo de Gómez-Moreno usando "*Early20thcentury*" como modelo base. El modelo resultante (Gomez-Moreno_v1) presentó una tasa de precisión del 98,57% significativamente superior a los ejemplos anteriores (Tabla 6). De esta forma, se demostraba que incluir modelos HTR relacionados con el contexto como base permitía dos cosas principales: a) generar una muestra de entrenamiento mucho más rápidamente, ya que solo requería correcciones en las transcripciones producidas automáticamente, en lugar de realizar una transcripción manual desde cero y b) el número de palabras para crear un conjunto de datos de entrenamiento exitoso es menor.

**Tabla 6. Datos del modelo Gomez-Moreno_v1.**

| MODELO HTR | CER EN LOS DATOS DE ENTRENAMIENTO | CER EN LOS DATOS DE VALIDACIÓN | Nº DE PALABRAS | DATOS USADOS PARA CREAR EL MODELO HTR |
|---|---|---|---|---|
| Gomez-Moreno_v1 | 0,88% | 1,43% | 11.938 | – Modelo base: Early20thcentury<br>–Transcripción automática + correcciones de 50 páginas. |

Usando el modelo Gomez-Moreno_v1 y el diccionario anterior transcribimos el resto del texto de este autor. Finalmente, se evaluó la eficiencia de los modelos arrojando resultados similares a los anteriores. A pesar de que el modelo específico para la letra del catálogo de Gómez-Moreno presentaba una CER muy baja, el test externo demostró que el porcentaje de caracteres transcritos correctamente se encontraba entre el 92,19% y el 95,87% (Tabla 7).

**Tabla 7. Test del segundo experimento.**

| DOCUMENTO | NÚMERO DE CARACTERES | | CARACTERES REVISADOS | DIFERENCIA DE CARACTERES (%) | MODELO DE ENTRENAMIENTO HTR | CER EN CONJUNTO DE VALIDACIÓN (%) | PÁGS. |
|---|---|---|---|---|---|---|---|
| | TRANSCRIPCIÓN AUTOMÁTICA | TRANSCRIPCIÓN REVISADA | | | | | |
| Gómez-Moreno libro 1 | 71875 | 71782 | 4376 | **6.09** | Early20thCentury | 4.97 | 50 |
| Gómez-Moreno libro 1 | 284619 | 272664 | 22224 | **7.81** | Gomez-Moreno_v1 | 1.43 | 208 |
| Gómez-Moreno libro 2 | 309372 | 306921 | 12792 | **4.13** | Gomez-Moreno_v1 | 1.43 | 254 |

## 3.3. EFICIENCIA DE LOS MODELOS HTR

Cada transcripción automática fue revisada manualmente para verificar el orden de lectura de las regiones y líneas, incorporar subrayados, incluir saltos de línea, agregar o quitar acentos, corregir signos de puntuación, etiquetar imágenes y dibujos, etc. El resultado final fue un corpus revisado en formato legible por computadora (Liceras-Garrido, Comino y Murrieta-Flores, 2020 a, b, c).

La comparación entre las trascripciones automática y las versiones editadas mediante una evaluación externa a la formación de los modelos no parece altamente significativa en términos de esfuerzo para las revisoras, ya que ofrece un promedio de modificaciones del 6,99% para todos los textos analizados (Tabla 8). Así, la modificación de caracteres del corpus para obtener transcripciones exactas de las versiones automatizadas estuvo entre el 4,13% y el 8,79%, no muy lejos del rango de 5,50% y 1,43% calculado por el CER en el conjunto de validación de los modelos HTR. El bajo número de modificaciones muestra el buen rendimiento de la arquitectura HTR de Transkribus.

**Tabla 8. Comparación entre las transcripciones automáticas y las versiones editadas**

| DOCUMENTO | NÚMERO DE CARACTERES | | CARACTERES REVISADOS | DIFERENCIA DE CARACTERES (%) | MODELO DE ENTRENAMIENTO HTR | CER EN CONJUNTO DE VALIDACIÓN (%) | PÁGS. |
|---|---|---|---|---|---|---|---|
| | TRANSCRIPCIÓN AUTOMÁTICA | TRANSCRIPCIÓN REVISADA | | | | | |
| Cabré libro 1 | 27966 | 27868 | 2437 | 8.71 | Cabré_v3 | 4.75 | 100 |
| Cabré libro 2 | 83147 | 77176 | 6041 | 7.27 | Cabré_v3 | 4.75 | 201 |
| Cabré libro 3 | Transcripción manual | | | | | | 175 |
| Cabré libro 4 | Transcripción manual | | | | | | 58 |
| Cabré libro 4 | 162579 | 157056 | 9808 | 6.03 | Cabré_v3 | 4.75 | 217 |
| Cabré libro 5 | 52501 | 53214 | 3858 | 7.35 | Cabré_v1 | 5.41 | 121 |
| Cabré libro 6 | 95792 | 95379 | 4417 | 4.61 | Cabré_v3 | 4.75 | 220 |
| Cabré libro 7 | 83066 | 82424 | 5609 | 6.75 | Cabré_v3 | 4.75 | 226 |
| Cabré libro 8 | 69225 | 68678 | 6084 | 8.79 | Cabré_v3 | 4.75 | 182 |
| Sentenach libro 1 | Transcripción manual | | | | | | 206 |
| Sentenach libro 2 | 96242 | 96747 | 6882 | 7.15 | Sentenach_v1 | 5.12 | 151 |
| Sentenach libro 3 | 142442 | 141143 | 10465 | 7.35 | Sentenach_v3 | 5.15 | 237 |
| Sentenach libro 4 | 153910 | 152316 | 11132 | 7.23 | Sentenach_v3 | 5.15 | 253 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sentenach libro 5 | 136143 | 133114 | 10547 | 7.75 | Sentenach_v3 | 5.15 | 208 |
| Sentenach libro 6 | 120150 | 117124 | 8812 | 7.33 | Sentenach_v3 | 5.15 | 195 |
| Sentenach libro 7 | 150371 | 149459 | 11378 | 7.57 | Sentenach_v3 | 5.15 | 260 |
| Gómez-Moreno libro 1 | 71875 | 71782 | 4376 | 6.09 | Early20thCentury | 4.97 | 50 |
| Gómez-Moreno libro 1 | 284619 | 272664 | 22224 | 7.81 | Gomez-Moreno_v1 | 1.43 | 208 |
| Gómez-Moreno libro 2 | 309372 | 306921 | 12792 | 4.13 | Gomez-Moreno_v1 | 1.43 | 254 |

## 4. DISCUSIÓN

El proyecto proponía utilizar Transkribus para crear una edición diplomática del CME, en la medida en que los caracteres tipográficos existentes lo permitieran, por lo que se han mantenido determinadas particularidades y errores ortotipográficos. Transkribus es una plataforma ideal para transcriptores no expertos en técnicas computacionales, ya que es muy intuitiva y existen varios manuales accesibles que guían paso a paso los procesos de transcripción manual, entrenamiento y edición de resultados. Sin embargo, para paleógrafas y paleógrafos o expertas y expertos en computación, ha sido señalada la existencia de *black boxes*, respecto a la opacidad y falta de transparencia en los procesos de aprendizaje y automatización, privando al usuario del control sobre los diferentes procesos de la transcripción automática. En relación con ciertos aspectos técnicos, comprobamos que el uso de más de 15.000 palabras en el entrenamiento del modelo no afectaba al proceso de transcripción automática, no siendo significativas las diferencias entre los porcentajes de correcciones manuales, como se demostró al comparar los porcentajes de corrección entre las transcripciones realizadas con Cabré_v1-Sentench_v1 y Cabré_v3-Sentench_v3 (Tabla 6).

Uno de los objetivos de CME era recopilar información estandarizada. Sin embargo, la heterogeneidad es la característica

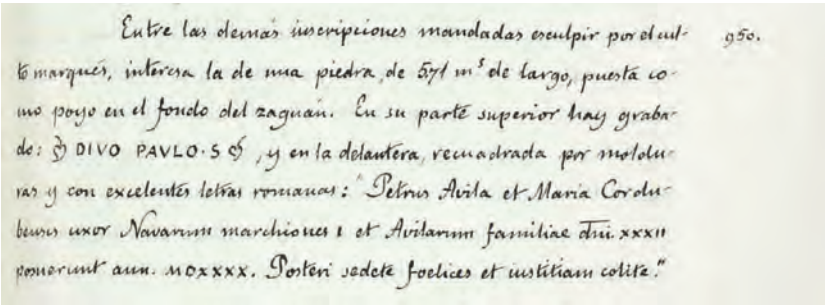

Entre las demás inscripciones mandadas esculpir por el mis- 950.
mo marqués, interesa la de una piedra, de 5y1 m.s de largo, puesta co-
mo poyo en el fondo del zaguán. En su parte superior hay grabar-
do: D DIVO PAVLO·S D, y en la delantera, recuadrada por mololu-
ras y con excelentes letras romanas: "Petrus Avila et Maria Corolu-
bensis uxor Navarum marchiones i et Avilarum familiae dni. xxxii
posuerunt ann. mDxxxx. Posteri sedete foelices et iustitiam colite."

Figura 7. Fragmentos de texto en latín y castellano e inscripción con los signos de *hedera distinguens y vinculum* (Gómez-Moreno, libro 2).

que define esta recopilación. A lo largo de estos volúmenes se pueden encontrar textos manuscritos, mapas, bocetos, fotografías, recortes de publicaciones impresas, etc. Del mismo modo, aunque la lengua predominante de este corpus es el español, está salpicado de fragmentos en latín clásico, vulgar y eclesiástico, francés, inglés, alemán, árabe y hebreo (Figura 7). Además de eso, hay una variedad de caracteres especiales, como flechas, cruces, triángulos, entre otros, que se han incluido utilizando el formato UTF-8. Sin embargo, estos manuscritos, como otros documentos históricos, presentan una serie de desafíos para las tecnologías de codificación y HTR, dado que resulta complejo transcribir palabras sobrescritas, tachones, subrayados de abreviaturas o signos de corrección ortográfica, por citar solamente unos ejemplo.

Por otra parte, estos manuscritos presentan una serie de retos para las tecnologías de codificación y transcripción automática. Tras la inclusión del diccionario, las transcripciones mejoraron notablemente y se solucionaron algunos problemas relacionados con la separación silábica de las palabras o la acentuación en español. No obstante, es preciso señalar que menos del 10% de los caracteres necesitaron modificaciones para lograr una transcripción precisa. Así, durante la revisión manual, las modificaciones más comunes consistieron en la corrección

de errores ortográficos, la adición o eliminación de acentos y la modificación de caracteres especiales como ñ, ç o ü. En este sentido, es importante tener en cuenta que la caligrafía de cada autor es diferente y, por tanto, los errores no son consistentes. Por este motivo, fue necesario realizar una revisión manual, ya que las confusiones de letras podían cambiar por completo el significado de las palabras. De este modo, si se modifica la primera vocal de la palabra masa, se obtiene mesa, misa o musa. Siguiendo el mismo ejemplo, si lo que cambia es la primera consonante, se transformaría en basa, casa o tasa. A estos errores se añaden las peculiaridades de las palabras tritónicas en castellano, aquellas que pueden acentuarse en sus tres sílabas, cambiando así su significado como, por ejemplo, término, termino y terminó. En la mayoría de los casos, estas confusiones tienen poca importancia, ya que se sobreentienden por el contexto. Sin embargo, a veces se necesitan conocimientos especializados sobre el tema. Por ejemplo, en uno de los libros de Sentenach se hacía referencia a la advocación mariana de una iglesia, pero no quedaba claro si se trataba de la virgen de la Ascensión o de la Asunción, lo que solamente el conocimiento histórico puede aclarar.

El tiempo y el esfuerzo de las revisiones también se vieron influidos por la eliminación del ruido de las transcripciones automáticas y los cambios en el orden de lectura de las líneas de texto. Por ejemplo, los documentos escritos por Cabré requirieron de un sinfín de modificaciones del orden de lectura, ya que el autor realizaba constantes rectificaciones y añadidos de información entre las líneas ya escritas. Los documentos escritos por Sentenach presentan numerosas manchas de tinta que el algoritmo interpretaba como texto y trataba de transcribir. Por último, en los libros de Gómez-Moreno, el principal problema estuvo determinado por la imagen escaneada, ya que a veces se incluían fragmentos de otras páginas en los márgenes, lo que generaba nuevas regiones de texto y transcripciones que se eliminaron durante la revisión manual (Figura 8).

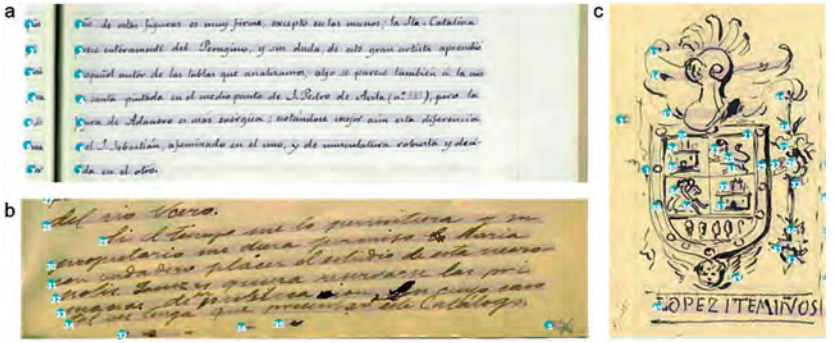



Figura 8. Identificación incorrecta de regiones de texto: (a) texto de otra página (Gómez-Moreno, libro 1), (b) manchas de tinta identificadas como texto (Cabré, libro 4) y (c) dibujo identificado como texto (Sentenach, libro 4).

Finalmente, la revisión de cada página llevó un tiempo medio de 10 minutos para lograr una transcripción precisa, en comparación con los 15 minutos necesarios para transcribir el mismo texto desde cero, lo que puede variar dependiendo de las habilidades paleográficas y conocimientos de la persona que transcriba. Teniendo en cuenta las diferencias de tiempo y esfuerzo entre la transcripción manual de todo el corpus y la mera revisión de las transcripciones automáticas, se observa que el uso de un modelo automático de reconocimiento de texto afecta sustancialmente el tiempo dedicado a esta tarea. En los casos de estudio expuestos, menos del 10% de los caracteres necesitaron modificaciones para lograr una transcripción completamente precisa. Por lo tanto, dependiendo del uso del corpus, se podría decidir si la transcripción automática es lo suficientemente precisa para los propósitos de la investigación y omitir la revisión manual. En el caso de algunas aplicaciones de minería de textos, análisis lingüísticos de corpus o enfoques de *big data*, este margen de error puede ser aceptable dependiendo del tamaño del corpus.

## 5. CONCLUSIONES

El reconocimiento automático de textos manuscritos ha demostrado ser una tecnología eficiente para producir transcripciones

automatizadas en poco tiempo. Una vez entrenados los modelos, la posibilidad de trabajar con diferentes idiomas (incluso al mismo tiempo) y mejorar la detección de los caracteres a través de diccionarios hace que estas técnicas sean extremadamente versátiles. Los principales desafíos a los que nos hemos enfrentado en este artículo son universales y se aplican a la mayoría de los documentos históricos escritos a mano. Además, la codificación UTF-8 está limitada en el contexto de documentos históricos, incapaz de representar símbolos como marcas de cantero o abreviaturas de escriba, subrayados diversos o signos de corrector.

El texto en formato legible por ordenador permite realizar análisis más allá de la lectura individual de los textos, como los análisis de contenido o discurso (Lingüística de Corpus); la identificación, clasificación y recuperación de conceptos clave (Procesamiento del Lenguaje Natural); o cuestionar la autoría de los manuscritos (Estilometría). Este último fue especialmente relevante para nuestra investigación porque al realizar una investigación paralela sobre la autoría del catálogo de Ávila de Manuel Gómez-Moreno, comprobamos que aunque él fue el autor intelectual, su hermana Sacramento fue la amanuense encargada de pasar a limpio sus investigaciones (Liceras-Garrido, Comino y Murrieta-Flores, 2022).

El resultado final de este proyecto son ocho modelos HTR entrenados para transcribir la letra de tres académicos imprescindibles para la historia de la Arqueología en España, así como un diccionario con terminología arqueológica de principios del siglo XX en español. Estos están disponibles en Transkribus para crear transcripciones automáticas o nuevos modelos HTR utilizándolos como modelos base. Además, se ha generado un corpus de 386.444 palabras que contiene dieciocho libros del CME en texto codificado por computadora disponible en Figshare y GitHub bajo una licencia *Creative Commons* (Liceras-Garrido, Comino y Murrieta-Flores, 2020 a, b, c).

**Agradecimientos**

Esta investigación forma parte del proyecto "*Goodbye reading glasses: a Machine Learning experiment on handwritten documents*" (SZA 1410) financiado por la *Faculty of Arts and Social Sciences* y el *Digital Humanities Hub* de la Universidad de Lancaster. También ha contado con el apoyo del proyecto "*Digging into Early Colonial Mexico: A large scale computational analysis of 16th century historical sources*", financiado por T-AP Digging into Data (ES/R003890/1).

Por otra parte, nos gustaría dar las gracias a Patricia Murrieta Flores, Katherine Bellamy y James Taylor por su apoyo y consejos. Finalmente, también queremos expresar nuestro agradecimiento al equipo de Transkribus por su asesoramiento durante el proceso de creación de los modelos.

## REFERENCIAS BIBLIOGRÁFICAS

Albert de León, Ángeles (2012). Introducción. En: Domingo, María. y Cendón, Óscar. (eds.). *El Catálogo Monumental de España (1900-1961). Investigación, restauración y difusión.* Madrid: Ministerio de Cultura - Secretaría General Técnica, 7.

Amrani, Ahmed, Abajian, Vicken, Kodratoff, Yves y Matte-Tailliez, Oriane (2008). A Chain of Text-mining to Extract Information in Archaeology. En: *3rd International Conference on Information and Communication Technologies: From Theory to Applications, Institute of Electrical and Electronics Engineers Computer Society*. Piscataway, 1-5 https://doi.org/10.1109/ICTTA.2008.4529905

Díaz-Andreu, Margarita (2009). Sentenach y Cabañas, Narciso. En: M. Díaz-Andreu, G. Mora, G. y J. Cortadella (eds.). *Diccionario histórico de la arqueología en España (siglos XV-XX)*. Madrid: Marcial Pons, 585-586.

Díaz-Andreu, Margarita y Sanz-Gallego, Nuria (1994). Women in Spanish Archaeology. En: M. Nelson, S. Nelson y A. Wylie (eds.). *Equality Issues for Women in Archaeology*. Arlington: American Anthropological Association, 121-130.

Domínguez-Rodrigo, Manuel, Cifuentes-Alcobendas, Gabriel, Jiménez-García, Blanca, Abellán, Natalia, Pizarro-Mozo, Marcos, Organista,

Elia y Baquedano, Enrique (2020). Artificial intelligence provides greater accuracy in the classification of modern and ancient bone surface modification. *Scientific Reports*, 10, 18862, https://doi.org/10.1038/s41598-020-75994-7

Govindan, V. K. y Shivaprasad, A. P. (1990). Character recognition - A review. *Pattern Recognition*, 23 (7), 671-683. https://doi.org/10.1016/0031-3203(90)90091-X

Govindaraju, Venu, Srihari, R. y Srihari, S. (1995). Handwritten text recognition. En: Spitz, A. Lawrence., Dengel, Andreas (eds.). *Document Analysis Systems.* Singapore: World Scientific Publishing, 288-304.

Hidalgo Brinquis, María del Carmen (2012). Interpretación material de los Catálogos Monumentales de España. En: María Domingo y Óscar Cendón (eds.). *El Catálogo Monumental de España (1900-1961). Investigación, restauración y difusión*. Madrid: Ministerio de Cultura - Secretaría General Técnica, 77-107.

Hörr, Christian, Lindinger, Elisabeth y Brunnett, Guido (2014). Machine learning based typology development in archaeology. *Journal on Computing and Cultural Heritage*, 7 (1), 1-23. https://doi.org/10.1145/2533988

Kahle, Philip, Colutto, Sebastian, Hackl, Günter y Mühlberger, Günter (2017). Transkribus - A Service Platform for Transcription, Recognition and Retrieval of Historical Documents. En: *14th IAPR International Conference on Document Analysis and Recognition 4.* Los Alamitos: Institute of Electrical and Electronics Engineers Computer Society, 19-24. https://doi.org/10.1109/ICDAR.2017.307.

Kichuk, Diana (2015). Loose, Falling Characters and Sentences: The Persistence of the OCR Problem in Digital Repositories E-Books. *Libraries and the Academy*, 15(1), 59-91. https://doi.org/10.1353/pla.2015.0005.

Krizhevsky, Alex, Sutskever, Ilya y Hinton, Geoffrey E. (2013). ImageNet Classification with Deep Convolutional Neural Networks. En: P.F. Bartlett, F. Pereira, C.J.C. Burges, L. Bottou, y K.Q. Weinberger (eds.). *Advances in Neural Information Processing Systems* 25. Red Hook: Curran Associates, 1097-1105. https://papers.nips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

[dataset] Liceras-Garrido, Raquel, Comino, Alba y Murrieta-Flores, Patricia (2020a). *Transcripción del Catálogo Monumental de España: Provincia de Ávila por Manuel Gómez Moreno (1900-1901).* Figshare Data. https://doi.org/10.6084/m9.figshare.12006318.v1.

[dataset] Liceras-Garrido, Raquel, Comino, Alba y Murrieta-Flores, Patricia (2020b). *Transcripción del Catálogo Monumental de la Provincia de Soria por Juan Cabré (1916-1917).* Figshare Data. https://doi.org/10.6084/m9.figshare.12006273.v1.

[dataset] Liceras-Garrido, Raquel, Comino, Alba y Murrieta-Flores, Patricia (2020c). *Transcripción del Catálogo Monumental y Artístico de la Provincia de Burgos por Narciso Sentenach (1925).* Figshare Data. https://doi.org/10.6084/m9.figshare.12006327.v1.

Liceras-Garrido, Raquel, Comino, Alba y Murrieta-Flores, Patricia (2022). Mujeres en el Catálogo Monumental de España: discursos arqueológicos sobre Prehistoria y Edad del Hierro en las provincias de Ávila, Soria y Burgos. *Complutum*, 33(1): 269-288. https://doi.org/10.5209/cmpl.80895

Liceras-Garrido, Raquel, Favila-Vázquez, Mariana, Bellamy, Katherine, Murrieta-Flores, Patricia, Jiménez Badillo, Diego y Martins, Bruno (2019). Digital Approaches to Historical Archaeology: Exploring the Geographies of 16th Century New Spain. *Open Access Journal of Archaeology and Anthropology*, 2(1), http://dx.doi.org/10.33552/OAJAA.2019.02.000526.

López-Ocón Cabrera, Leoncio (2012). El papel de Juan Facundo Riaño como inductor del proyecto cultural del Catálogo Monumental de España. En: María Domingo y Óscar Cendón (eds.). *El Catálogo Monumental de España (1900-1961). Investigación, restauración y difusión.* Madrid: Ministerio de Cultura - Secretaría General Técnica, 51-73.

López-Yarto Elizalde, Amelia (2012). Los autores del Catálogo Monumental de España. En: María Domingo y Óscar Cendón (eds.). *El Catálogo Monumental de España (1900-1961). Investigación, restauración y difusión.* Madrid: Ministerio de Cultura - Secretaría General Técnica, 39-49.

Lorenzo Arribas, Josemi y Pérez Martín, Sergio (2017). *Excursiones Zamoranas 1903-1904. Epistolario de Manuel Gómez-Moreno y Elena Rodríguez-Bolívar*. Zamora: Editorial Semuret.

Mélida, José Ramón (1912). *Excavaciones de Numancia. Memoria presentada al Ministerio de Instrucción Pública y Bellas Artes por la Comisión Ejecutiva*. Madrid: Imprenta Artística de José Blas y Cía.

Michael, Johannes, Weidemann, Max y Labahn, Roger (2018). *HTR Engine Based on NNs P3 Optimizing speed and performance - HTR+*. Technical report, D7.9. https://readcoop.eu/wp-content/uploads/2018/12/Del_D7_9.pdf.

Morales, Alfredo J. 1996. El Catálogo de Bienes Muebles. Registro, conocimiento y tutela. En *Catalogación del Patrimonio Histórico.* Sevilla: Instituto Andaluz del Patrimonio Histórico - Junta de Andalucía, pp. 41-48.

Müehlberger, Guenter; Seaward, Louise; Terras, Melissa; Ares Oliveira, Sofia; Bosch, Vicente; Bryan, Maximilian; Colutto, Sebastian; Déjean, Hervé; Diem, Markus; Fiel, Stefan; Gatos, Basilis; Greinoecker, Albert; Grüning, Tobias; Hackl, Guenter; Haukkovaara, Vili; Heyer, Gerhard; Hirvonen, Lauri; Hodel, Tobias; Jokinen, Matti; Kahle, Philip; Kallio, Mario; Kaplan, Frederic; Kleber, Florian; Labahn, Roger; Lang, Eva Maria; Laube, Sören; Leifert, Gundram; Louloudis, Georgios; McNicholl, Rory; Meunier, Jean-Luc; Michael, Johannes; Mühlbauer, Elena; Philipp, Nathanael; Pratikakis, Ioannis; Puigcerver Pérez, Joan; Putz, Hannelore; Retsinas, George; Romero, Verónica; Sablatnig, Robert; Sánchez, Joan Andreu; Schofield, Philip; Sfikas, Giorgos; Sieber, Christian; Stamatopoulos, Nikolaos; Strauss, Tobias; Terbul, Tamara; Toselli, Alejandro Héctor; Ulreich, Berthold; Villegas, Mauricio; Vidal, Enrique; Walcher, Johanna; Weidemann, Max; Wurster, Herbert y Zagoris, Konstantinos (2019). Transforming scholarship in the archives through handwritten text recognition: Transkribus as a case study. Journal of Documentation; 75(5); 954-976. https://doi.org/10.1108/JD-07-2018-0114

Muñoz Cosme, Aalfonso (2012). Catálogo e inventarios del Patrimonio en España. En: María Domingo y Óscar Cendón (eds.). *El Catálogo Monumental de España (1900-1961). Investigación, restauración y difusión.* Madrid: Ministerio de Cultura - Secretaría General Técnica, 15-37.

Murrieta-Flores, Patricia y Gregory, Ian (2015). Further frontiers in GIS: Extending Spatial Analysis to Textual Sources in Archaeology. *Open Archaeology,* 1, 166-175. https://doi.org/10.1515/opar-2015-0010

Pasamar Alzuria, Gonzalo Vicente (2009a). Gómez-Moreno Martínez, Manuel. En: M. Díaz-Andreu, G. Mora, G. y J. Cortadella (eds.). *Diccionario histórico de la arqueología en España (siglos XV-XX)*. Madrid: Marcial Pons, 305-307.

Pasamar Alzuria, Gonzalo Vicente (2009b). Cabré Aguiló, Juan En: M. Díaz-Andreu, G. Mora, G. y J. Cortadella (eds.). *Diccionario histórico de la arqueología en España (siglos XV-XX)*. Madrid: Marcial Pons, 161-162.

Richards, Julian D., Tudhope, Douglas y Vlachidis, Aandreas (2015). Text Mining in Archaeology: Extracting Information from Archaeological Reports. En Juan Barceló y Igor Bogdanovic (eds.). *Mathematics and Archaeology*. Florida: CRC Press-Taylor & Francis Group, 240-254.

Romero, Verónica, Toselli, Alejandro Héctor y Vidal, Enrique (2012). *Multimodal Interactive Handwritten Text Transcription*. Singapore: World Scientific Publishing.

Strauss, Tobias, Grüning, Tobias, Leifert, Gundram y Labahn, Roger (2016). CITlab ARGUS for Keyword Search in Historical Handwritten Documents - Description of CITlab's System for the ImageCLEF 2016 Handwritten Scanned Document Retrieval Task. En: *Conference and Labs of Evaluation Forum Working Notes*, 399-412. http://ceur-ws.org/Vol-1609/16090399.pdf

Ströbel, Phillip Benjamin; Clematide, Simon y Volk, Martin (2020). How much data do you need? About the creation of a ground truth for black letter and effectiveness of neural OCR. En: *12th Conference on Language Resources and Evaluation,* 3551-3559. https://www.aclweb.org/anthology/2020.lrec-1.436.pdf

Tealab, Ahmed (2018). Time series forecasting using artificial neural networks methodologies: A systematic review. *Future Computing and Informatics Journal,* 3 (2), 334–340. https://doi.org/10.1016/j.fcij.2018.10.003

Van der Maaten, Laurens, Boon, Paul, Lange, Guus, Paijmans, Hands y Postma, Eric (2006). Computer Vision and Machine Learning for Archaeology. En: J.T. Clark y E.M. Hagemeister (eds.). *Digital Discovery. Exploring New Frontiers in Human Heritage*. Budapest: Archaeolingua, 112-130. https://publikationen.uni-tuebingen.

de/xmlui/bitstream/handle/10900/61550/CD49_Maaten_et_al_CAA2006.pdf.

Verschoof-van der Vaart, Wouter Baernd y Lambers, Karsten (2019). Learning to Look at LiDAR: The Use of R-CNN in the Automated Detection of Archaeological Objects in LiDAR Data from the Netherlands. *Journal of Computer Applications in Archaeology*, 2 (1), 31-40. http://doi.org/10.5334/jcaa.32

Weidemann, Max, Michael, Johannes Michael, Grüning, Tobias y Labahn, Roger (2017). HTR Engine Based on NNs P2. Building deep architectures with TensorFlow. *Technical report*, D7.8. https://readcoop.eu/wp-content/uploads/2017/12/Del_D7_8.pdf

READ project 2019. *How to train a handwritten text recognition model in Transkribus*. Project report. https readcoop.eu/transkribus/howto/how-to-train-a-handwritten-text-recognition-model-in-transkribus/

# ESTUDIO INTERDISCIPLINAR DE UNA RELACIÓN FESTIVA:
## DE LA DESCRIPCIÓN TEXTUAL A LA REPRESENTACIÓN VISUAL DE LA ARQUITECTURA EFÍMERA Y SU ICONOGRAFÍA

Estefanía López Salas, José Julio García Arranz,
Manuel Alberto Santos Miñambres, Sagrario López Poza

**Resumen**

El presente texto muestra el trabajo en curso dentro del proyecto I+D+i *Biblioteca Digital Siglo de Oro 6* (BIDISO 6) para la recuperación de cinco arquitecturas efímeras erigidas con ocasión de la entrada de la reina Isabel de Valois en Toledo en el año 1560 y su rico programa iconográfico. El artículo se organiza en cinco partes principales. En la primera de ellas presentamos el proyecto y sus objetivos en torno a la recreación visual de una relación festiva. La segunda parte la destinamos a introducir la entrada real y la fuente literaria que utilizamos para su estudio, así como el contexto artístico que rodeó a este acontecimiento. En tercer lugar, explicamos el flujo de trabajo desarrollado para representar las trazas principales de los aparatos efímeros con herramientas de diseño asistido

Estefanía López Salas, Universidade da Coruña
https://orcid.org/0000-0002-2889-6465
estefania.lsalas@udc.es

José Julio García Arranz, Universidad de Extremadura
https://orcid.org/0000-0002-7052-8754
turko@unex.es

Manuel Alberto Santos Miñambres, Universidad de Extremadura
https://orcid.org/0000-0002-9698-714X
msantosjx@alumnos.unex.es

Sagrario López Poza, Universidade da Coruña
https://orcid.org/0000-0003-1852-7938
sagrario.lopez.poza@udc.es

por ordenador. A continuación, el apartado cuatro se focaliza en exponer el método que elaboramos para la localización de elementos iconográficos que, en mayor o menor grado, se correspondan con las descripciones que de ellos contiene nuestra relación de partida. En último lugar, hacemos una reflexión conjunta acerca de los resultados alcanzados y las dificultades que surgieron en el proceso de trabajo, así como de los caminos para su desarrollo futuro, de cara a una aproximación integral a la arquitectura e iconografía de la fiesta pública en el Siglo de Oro.

**Palabras clave:** Arquitectura efímera, iconografía, crónica festiva, fiestas Edad Moderna, Siglo de Oro

**Abstract**
This paper shows the work in progress within the research project titled «Biblioteca Digital Siglo de Oro 6» (BIDISO 6) to recover five ephemeral triumphal arches built for a royal entry of Isabel de Valois in Toledo (1560) and their rich iconography. The article is organized into five main parts. It first presents the project aim around the graphic recreation of a festival news pamphlet. Second part is devoted to introduce the royal entry and the literary source we used in the study as well as the artistic context that surrounded the historical event. In the third section we explain the workflow to draw the ephemeral architecture with computer-aided design tools. After that, fourth part is focused on showing the method developed to search and identify the iconographic materials that match the descriptions of the literary source. Finally, we reflect on the achieved outcomes and the challenges we faced as well as the future lines of work in a way to a comprehensive approach to the architecture and iconography for a royal entry in the Spanish Golden Age.

**Keywords:** Ephemeral architecture, iconography, festival news pamphlet, Early Modern festivals, Spanish Golden Age

## 1. INTRODUCCIÓN

El estudio de los acontecimientos festivos históricos a partir de fuentes de diverso tipo viene constituyendo la base documental fundamental para la creación de espectáculos contemporáneos que se proponen recuperar la memoria de las tradiciones del pasado. Generalmente se configuran como eventos multisensoriales que, destinados a la participación activa de población local y turistas, acaban convirtiéndose en recursos de interés en la difusión del patrimonio cultural inmaterial entre un público experto o no. Además, suelen reportar importantes beneficios económicos a las ciudades que los organizan.

En el campo específico de la fiesta pública de los siglos XVI y XVII, este potencial patrimonial y turístico puede ser considerado como una primera razón que subyace en el notable desarrollo que la investigación humanística sobre este tipo de estudios ha experimentado en los últimos años. Asimismo, los avances en el conocimiento en este campo se han visto favorecidos por la creación de grupos de investigación interdisciplinares, que permiten enfoques diversos sobre la riqueza de las expresiones artísticas involucradas en tales acontecimientos festivos.

A las dos cuestiones anteriores debemos añadir el avance que la aplicación de las últimas tecnologías digitales ha supuesto tanto para el estudio como para la difusión de las investigaciones desarrolladas en torno a los espectáculos festivos del Siglo de Oro, sus escenarios y sus elementos artísticos componentes. Durante los últimos años se han presentado diversos trabajos en los que se ha recurrido al modelado tridimensional para la representación de arquitecturas efímeras (Antolini, 2020; Fernández González, 2017 y 2019; Marín Tovar y Román Aliste, 2020; Varela Flor, 2020). En otros casos, las herramientas digitales están siendo empleadas para recuperar el color de los aparatos festivos (Soto Caba y Solís Alcudia, 2019a y 2019b) o para integrar efectos sonoros o visuales a los escasos materiales gráficos conservados de este tipo de eventos (González Román, 2019: 114-115 y 2020: 174-176;

González Román y Risatti, 2021). La mayoría de las aproximaciones que han visto la luz en la última década se plantean como objetivo esencial caminar hacia la recreación de un ambiente escenográfico histórico, fundamentado en la autenticidad y transparencia de las fuentes utilizadas con el apoyo de tecnologías digitales, que permitan plantear nuevos interrogantes de carácter histórico-artístico en el estudio de la fiesta pública de la Edad Moderna y que, a su vez, faciliten la difusión de los resultados entre el público de forma atractiva.

Una categoría muy especial de festividad pública durante el Siglo de Oro estuvo constituida por las entradas solemnes de reinas que, casadas por poderes en su país, llegaban a España para ratificar el matrimonio y ser presentadas y recibidas por el pueblo. Las ciudades competían para poder organizar la entrada, de la que esperaban beneficios de diverso tipo; sin embargo, el monarca solo concedía tal privilegio a algunas de las villas solicitantes. Con la finalidad de dejar constancia del gasto y esfuerzos realizados, los ayuntamientos solían encargar a un cronista competente la exposición por escrito de todos los detalles del acontecimiento. Esas relaciones del suceso festivo constituyen fuentes históricas y literarias de capital importancia, si bien en raras ocasiones incluyen dibujos, grabados o pinturas de lo descrito (López Poza, 1998: 16). El lector del pasado, así como el contemporáneo, ha de imaginarse, gracias a la detallada descripción verbal, los arcos triunfales, soportales, escenarios, estatuas, fuentes y carros procesionales, así como la composición de las comitivas y atavíos de los asistentes.

El objetivo del presente artículo es mostrar el trabajo en desarrollo emprendido por un grupo interdisciplinar de investigadores e investigadoras de las áreas de Literatura Española, Historia del Arte y Composición Arquitectónica, para llevar a cabo la recreación virtual de una relación festiva: la redactada por el humanista Alvar Gómez de Castro por encargo del Ayuntamiento de Toledo relativa a la entrada de la reina Isabel

de Valois en la «ciudad imperial» que en aquel momento —1560— era sede de la corte de Felipe II[1]. Esta relación es la más importante y completa que se conserva del acontecimiento y, aunque contiene descripciones minuciosas de cada uno de los aparatos efímeros construidos para la entrada real, carece de cualquier tipo de representación gráfica (Gómez de Castro, 1561a y 1561b)[2]. A través de un determinado flujo de trabajo y método de visualización digital computarizada es posible recuperar, con rigor académico en el proceso y en los resultados, la imagen perdida de cinco de los diez aparatos festivos, su arquitectura y programa iconográfico. Se trata de un primer paso en el camino de descifrar el mensaje, eminentemente visual, de la compleja red de demostraciones y expresiones de la voluntad popular de la que tales estructuras formaron parte, proceso que, al mismo tiempo, abre la puerta hacia nuevas interpretaciones y modos de presentación y difusión de la fiesta pública del Renacimiento como acontecimiento histórico, artístico y cultural.

## 2. LA ENTRADA REAL DE 1560: ITINERARIO Y APARATOS EFÍMEROS

En 1559, Felipe II se casaba con su tercera esposa, Isabel de Valois, hija del rey Enrique II de Francia y Catalina de Médicis. La boda se había realizado por poderes en París, el 22 de junio, y se ratificó en Guadalajara, el 31 de enero de 1560. Se celebraron entradas festivas de la nueva reina en varias ciudades de Castilla, pero la más solemne tuvo lugar en Toledo, a donde Isabel de Valois llegó

1. Esta publicación es parte del proyecto de I+D+i *Biblioteca Digital Siglo de Oro 6* (BiDISO 6), referencia: PID2019-105673GB-I00 financiado por MCIN/ AEI/10.13039/501100011033/

2. Utilizamos la edición crítica que de esta relación de sucesos ha publicado Carlota Fernández Travieso en 2007 (Gómez de Castro, 1561a). También existe una edición facsimilar posterior, realizada por Julia de la Torre Fazio en 2012 (Gómez de Castro, 1561b).

el 12 de febrero de 1560. La recibió una gran fiesta pública que se extendía por toda la ciudad.

Uno de los recursos más utilizados en las entradas reales del siglo XVI fue la construcción de artefactos efímeros que, a modo de esculturas, jardines o arquitecturas temporales, modificaban el aspecto habitual de la ciudad y creaban una imagen ideal de esta. Las arquitecturas efímeras adquirían normalmente la forma de arcos de triunfo dispuestos en determinados puntos de un recorrido urbano previamente planificado. Los monarcas avanzaban por la ciudad, a la manera de emperadores romanos, y su solemne desplazamiento y paradas establecidas en los espacios distinguidos por las arquitecturas efímeras simbolizaban, entre otras cuestiones, la renovación de la lealtad a la Corona de la ciudad que los acogía, sus instituciones y habitantes (Ferrer Valls, 2003: 31).

Aunque los arcos de triunfo eran estructuras temporales de madera, lienzo y yeso, simulaban el mármol, el oro y la piedra de los monumentos permanentes. Asimismo, estaban decorados con una profusión de pinturas, inscripciones, estatuas y emblemas, conforme a un programa conceptual e iconográfico que, diseñado por un entendido en Letras Humanas, transmitía un mensaje cargado de significaciones políticas y propagandísticas que, por regla general, estaba destinado a glorificar a los monarcas y su estirpe. Su contenido iconográfico procedía de lecturas de obras que la imprenta había conseguido divulgar y también de la observación *in situ* de solemnidades celebradas en Italia y Flandes de tantos caballeros y miembros del séquito de Carlos V o del príncipe Felipe, su hijo, en el largo viaje que este realizó entre 1548 y 1552 por los estados de su padre (Marsden, 1960, 402-403). Calvete de Estrella (1552) dejó extraordinario testimonio escrito en *El felicíssimo viaje…* y su libro, con detalladas descripciones, fue inspiración posterior para quienes tenían que idear programas iconográficos.

En el caso de la entrada de la reina Isabel de Valois en Toledo, se construyeron un total de diez aparatos efímeros. El




Figura 1. Extracto de la vista de Toledo del pintor Anton van den Wyngaerde, 1563. Fuente: Biblioteca Nacional de Austria en Viena. https://onb.digital/result/10EDFF57

primero de ellos se levantó en las inmediaciones de la llamada puerta de la Bisagra, una entrada principal de la primera línea de murallas que cerraba el Toledo de finales del XVI por su cara norte. Esta puerta había sido reconstruida entre 1540 y 1548, tal y como podemos apreciar en la hermosa vista de Toledo que dibujó el flamenco Anton van den Wyngaerde en 1563 (Figura 1).

En este primer punto del itinerario real, la arquitectura efímera la constituía un arco de triunfo de tres calles y dos cuerpos. Era el más monumental de los aparatos desplegados, con 84 pies de ancho por 100 de alto y resuelto en orden dórico (Gómez de Castro, 1561a: 98). Tras cruzar la muralla y atravesar el llamado arrabal de Santiago, se dispuso un segundo arco en la segunda línea de murallas, junto a la llamada Herrería. De menores dimensiones, tenía una única calle y era de orden corintio compuesto (Gómez de Castro, 1561a: 122). Similar composición presentaba el tercer arco de triunfo, si bien con un solo cuerpo y decoración rústica. Se localizaba a pocos metros de distancia del anterior, en una zona ya por encima de la segunda muralla de la ciudad (Figura 2).

Las tres paradas siguientes de la reina tuvieron lugar en la plaza de San Vicente, en la que se había levantado un altar «a la antigua»; en la plaza situada delante de las llamadas casas del Conde

| Localización | Nombre del aparato efímero |
|---|---|
| 1 | Arco a la puerta de la Bisagra |
| 2 | Arco a la puerta del segundo muro, junto a la Herrería |
| 3 | Arco por encima de la segunda muralla de la ciudad, desde donde se ve el palacio de Galiana y la huerta del rey |
| 4 | Un altar a la antigua en la plaza de San Vicente |
| 5 | El huerto de las Hespérides, un bosque en la plaza de delante de las casas del Conde de Orgaz |
| 6 | Un gran coloso en la plaza de San Salvador |
| 7 | Arco en la puerta del Perdón de la catedral |
| 8 | Una figura de Lucrecia |
| 9 | Arco a la entrada de la plaza Zocodover |
| 10 | Una mesa con tres figuras de gigantes |

Figura 2. Relación de los diez aparatos efímeros de la entrada de Isabel de Valois en Toledo de 1560. Fuente: elaboración propia.




Figura 3. Localización de los aparatos efímeros de la entrada de Isabel de Valois de 1560, relacionados en la Figura 2. Fuente: elaboración propia sobre la ortofotografía actual de Toledo, Instituto Geográfico Nacional, PNOA, 2021.

de Orgaz, que fue convertida en un jardín de nombre «Huerto de las Hespérides», y en la plaza de San Salvador, donde se colocó la escultura de un gran coloso (Gómez de Castro, 1561a: 125, 127, 134). Desde aquí la reina fue conducida a la catedral de Toledo, en la que se había construido un cuarto arco de triunfo, de orden corintio, y el segundo de mayores dimensiones del conjunto, con 40 pies de ancho y 85 de alto (Gómez de Castro, 1561a: 136). Se dispuso pegado a la portada central de entrada al templo, llamada Puerta del Perdón, y transformaba temporalmente su aspecto medieval en una nueva fachada renacentista. Al salir de la catedral, la plaza de las Cuatro Calles se distinguía con la presencia de una figura de Lucrecia fingida de mármol (Gómez de Castro, 1561a: 143). Un poco más adelante, Isabel de Valois debía atravesar un quinto arco de triunfo para entrar en la plaza de Zocodover. Compartía con los tres anteriores el tener una única puerta de paso, pero se diferenciaba en el orden arquitectónico empleado, jónico. Su anchura estaba directamente determinada por la dimensión máxima que permitía la calle, 32 pies, pero en altura alcanzaba los 64 (Gómez de Castro, 1561a: 143-144). El décimo y último aparato efímero, antes de llegar al Alcázar de Toledo, punto final del recorrido, era una mesa de cinco pies de alto sobre la que se disponían tres figuras de gigantes de 22 pies de altura cada una (Gómez de Castro, 1561a: 150). Por tanto, en total fueron construidos cinco arcos de triunfo, un jardín y cuatro conjuntos escultóricos efímeros para esta entrada real (Figura 3).

Está documentado que el trazador y arquitecto del arco a la puerta de la Bisagra y de todos los demás espectáculos, con excepción del situado a la Puerta del Perdón de la catedral, fue Nicolás de Vergara (Gómez de Castro, 1561a: 26; 1561b: 51). Fernando Marías (1983: 371) y González Ramos (2018) atribuyen los arcos triunfales a Nicolás de Vergara, el Viejo (c. 1517-1574) que, aunque trabajó principalmente en obras de rejería, escultura y vidriería, también tuvo una pequeña producción arquitectónica, en buena parte en colaboración con Alonso de Covarrubias. El arco a la Puerta del Perdón de la catedral no fue de la misma autoría, pues se reaprovechó de otra celebración

anterior, la que tuvo lugar con motivo de la imposición del capelo cardenalicio al arzobispo Silíceo en 1556 (Marías, 1983: 141).

El cronista de la fiesta para Isabel de Valois nos dice que, en el diseño de los arcos, se siguieron las reglas del libro cuarto del *De Architectura* de Vitruvio, y que los vocablos utilizados para describirlos también proceden de la obra de ese tratadista romano (Gómez de Castro, 1561a: 98-99). Esta alusión condujo a Marsden (1960: 403) a concluir que los arcos de triunfo que se construyeron en España poco después de mediar el siglo seguían fielmente las formas prescritas por los entonces considerados maestros, es decir, los grandes tratadistas italianos, a diferencia de la sencillez que presentaban los aparatos de principios del XVI y la complejidad basada en las mezclas que muestran las descripciones de los levantados hacia 1586.

Los tratados de arquitectura renacentista habían empezado a traducirse y publicarse en castellano precisamente en la primera mitad del quinientos. La primera aportación española al estudio de la arquitectura «a la romana», obra de Diego de Sagredo, vio la luz en Toledo en 1526 (Paniagua Soto, 1997: 242). Lo siguieron las traducciones del italiano al castellano del *Tercero y Cuarto Libro de Arquitectura* de Sebastiano Serlio de 1552, en Toledo, y del tratado de Vitruvio en un manuscrito que suele fecharse entre 1554 y 1564 (Paniagua Soto, 1997: 236, 241). La primera edición en castellano de este último fue hecha por Urrea en 1582 (Paniagua Soto, 1997: 235).

Esta literatura sobre cómo hacer la arquitectura permitió la difusión del arte renacentista en la España del siglo XVI, junto al trabajo de artistas que previamente se habían formado en Italia. De hecho, el propio Nicolás de Vergara, el Viejo, se sabe que poseía una amplia biblioteca en su estudio de Toledo, formada por «cien libros impresos y manuscritos», tal y como figura en el inventario redactado por su hijo, Nicolás de Vergara, el Mozo, tras su fallecimiento (Marías, 1983: 364). Asimismo, tenía un buen número de modelos de yeso, betún y cera de esculturas italianas y clásicas que, según Marías (1983: 365-366), puede hacernos suponer que las había adquirido directamente en Italia o bien

proceder del mercado de importación. Todo este conocimiento es fundamental para afrontar el estudio de los cinco arcos de triunfo proyectados para la entrada de Isabel de Valois en Toledo en 1560.

## 3. REPRESENTACIÓN DE LA ARQUITECTURA EFÍMERA: FLUJO DE TRABAJO Y RESULTADOS

El primer paso del método que planteamos para recuperar las trazas de los cinco arcos de triunfo es una lectura analítica de la mencionada relación del acontecimiento compuesta por Gómez de Castro. Esta se basa en la extracción de información de carácter espacial de las descripciones de los aparatos efímeros, es decir, datos textuales referentes a: entidades arquitectónicas, escultóricas o pictóricas —columnas, arcos, machones, basas, capiteles, cuadros, figuras de bulto...—; dimensiones —de altura, anchura y profundidad de cada entidad—; cuestiones compositivas —relaciones entre las entidades o entre cada entidad y el todo—, y urbanas o paisajísticas —sobre la localización y entorno del arco triunfal— (Figura 4).

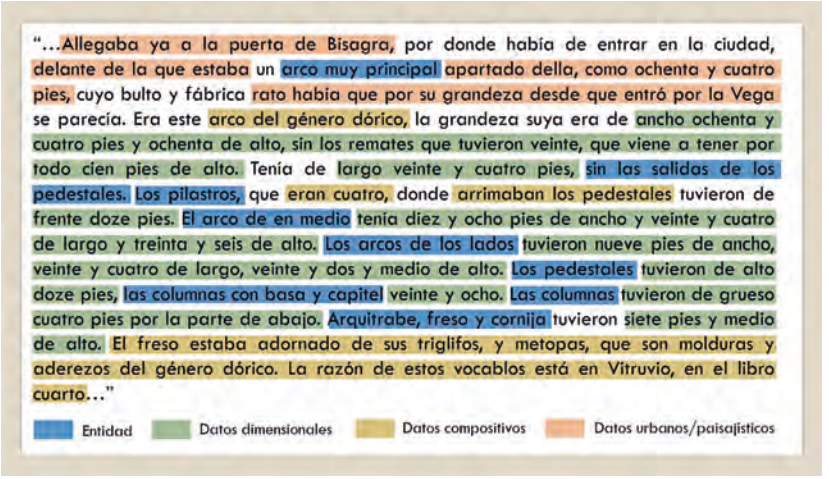



Figura 4. Ejemplo de lectura analítica para la extracción de datos espaciales de la descripción del arco a la puerta de la Bisagra de Gómez de Castro (1561a: 98). Fuente: elaboración propia.

De todos ellos, los datos más importantes para iniciar el trabajo de recuperación de las trazas arquitectónicas son los dimensionales que definían el volumen total de los arcos, pero también los correspondientes a cada una de las entidades que los componían. Esta información nos permite afrontar con rigor la representación bidimensional de unas arquitecturas que, por su naturaleza efímera, hoy no se conservan. Si nuestra relación de partida no hubiera recogido tales datos, como sucede en el caso de otras crónicas elaboradas sobre esta misma entrada real, el proceso que aquí planteamos no podría llevarse a cabo. En ausencia de las dimensiones obtenidas a partir de una fuente escrita, sólo la localización de pinturas, grabados o dibujos a escala podría permitir el alcance del objetivo propuesto.

No obstante, los datos dimensionales aportados por Gómez de Castro, aunque abundantes, no son completos. Como consecuencia de ello, uno de los primeros retos que hay que superar en el proceso de trabajo radica en la interpretación de las referencias que se encuentran ausentes en la descripción. Por ejemplo, en el caso del arco de la puerta de la Bisagra, Gómez de Castro especificó sus medidas generales, la anchura de cada machón, el alto y ancho de cada arco de paso, la altura de los pedestales, la de las columnas y su diámetro en la base y la altura del arquitrabe, friso y cornisa, así como la medida de algún elemento escultórico (Figura 5). Con toda esta información resulta posible comenzar el proceso de dibujo de las trazas del aparato efímero con un programa de diseño asistido por ordenador (CAD). En la Figura 6 se muestra hasta dónde llegamos utilizando exclusivamente los datos dimensionales de la relación de sucesos: una primera representación proporcionada del volumen general del arco, de sus tres puertas de paso, de cada uno de sus cuerpos en sentido vertical y de la altura de las partes componentes del primero de ellos, además del diámetro de la base de sus columnas.

Para continuar el proceso se hace necesario combinar estos datos con la aplicación de las leyes propias de la gramática clásica que se utilizaban en la época en la que fueron proyectados.

| Entidad | Alto | Ancho | Profundidad |
|---|---|---|---|
| Volumen | 100 | 84 | 24 |
| Remates | 20 | | |
| Pilastros | | 12 | |
| Arco central | 36 | 18 | 24 |
| Arcos laterales | 22,5 | 9 | 24 |
| Pedestales | 12 | | |
| Columnas | 28 | 4 | |
| Entablamento | 7,5 | | |
| Diosas Vesta y Palas | 18 | | |
| Tribuna | | 18 | |
| Suela tribuna | 8 | 18 | |
| Genio de la ciudad | | 14 | |
| Resalto a la ciudad | | 34 | |
| … | | | |

Figura 5. Datos dimensionales recogidos por Gómez de Castro sobre el arco a la puerta de la Bisagra. Fuente: elaboración propia.

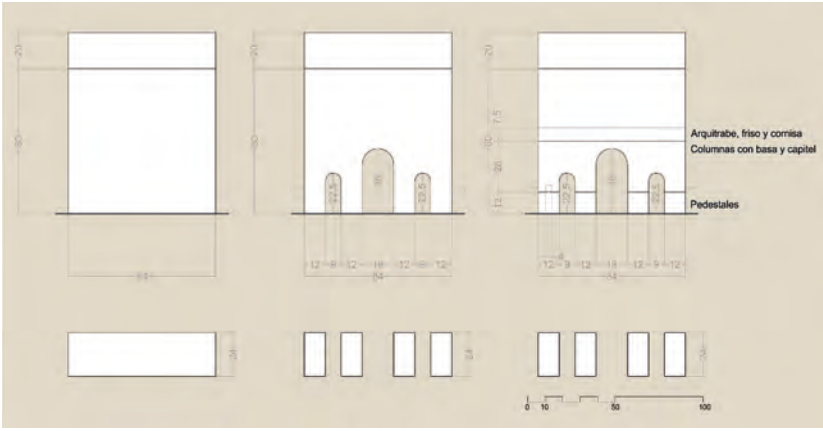



Figura 6. Tres primeros pasos del proceso de recuperación de las trazas del alzado principal del arco a la puerta de la Bisagra. Cotas y escala en pies. Fuente: elaborados por Estefanía López Salas.

Esta gramática se basa en un sistema de proporciones llamado orden arquitectónico, que sirve para componer un edificio a partir de la combinación de unos determinados elementos que se relacionan entre sí de un modo coherente por medio de unas normas prefijadas. Entre ellas figura el uso de un módulo o dimensión que se toma como unidad para referir las medidas principales. El sistema de proporciones varía de acuerdo con el orden del arco, un dato que también señala Gómez de Castro.

Volviendo al caso del arco a la puerta de la Bisagra, se dice que era de orden dórico. En el tratado de Vitruvio se establece que, para este orden, el diámetro de la columna tendrá dos módulos y la altura, incluyendo el capitel, catorce módulos (Oliver Domingo, 1997: 103). Por tanto, el módulo o unidad de medida en este primer orden es la mitad del diámetro inferior de la columna, medido encima de la basa. Tomando este sistema de modulación como referencia, tanto este teórico del arte clásico como los tratadistas del Renacimiento que en él se basaron describían en sus obras, con mayor o menor detalle, la altura de las columnas de cada orden, su separación, la dimensión de basas y capiteles, pedestales, entablamento...

En el flujo de trabajo que desarrollamos, hemos acudido a la aplicación de las reglas compositivas que, para cada orden arquitectónico, se recogen tanto en el libro cuarto de Vitruvio, como en el tratado de Serlio, publicado apenas ocho años antes de que tuviera lugar la entrada real que aquí nos ocupa. De hecho, la aplicación de estas reglas proporcionales es una vía a la que nos invita el propio Gómez de Castro en su crónica oficial.

Con el fin de ilustrar lo que acabamos de explicar, vamos a poner un ejemplo relativo de nuevo al arco a la puerta de la Bisagra. Gómez de Castro señala que la anchura de cada columna del arco a la altura de la base era de cuatro pies castellanos. Teniendo en cuenta el orden del arco, dórico, y aplicando la relación proporcional descrita por Vitruvio, es posible deducir que el módulo del arco era de dos pies. Asimismo, la altura de las columnas dóricas dice Vitruvio que debía ser de catorce módulos. Si unimos esta norma con el dato

espacial obtenido de la crónica oficial, concluimos que cada columna alcanzaba una altura de 28 pies castellanos. Este ejercicio, en este segundo caso, lo hemos aplicado suponiendo que la altura de las columnas fuera un dato ausente de la relación del acontecimiento, pero no es así. Gómez de Castro sí señala que las columnas, con su basa y capitel, tenían 28 pies. Por tanto, esta metodología de trabajo nos permite obtener datos dimensionales desconocidos, pero supone al mismo tiempo una forma de contrastar la validez de la fuente.

El siguiente paso en la reconstitución del arco a la puerta de la Bisagra habría de ser, por ejemplo, enfrentarse a la recreación del capitel que coronaba cada columna, pero cuyas dimensiones no figuran en la relación festiva. Vitruvio indica que la altura del capitel de la columna dórica debía ser de un módulo y su anchura de dos módulos más una sexta parte (Oliver Domingo, 1997: 103).

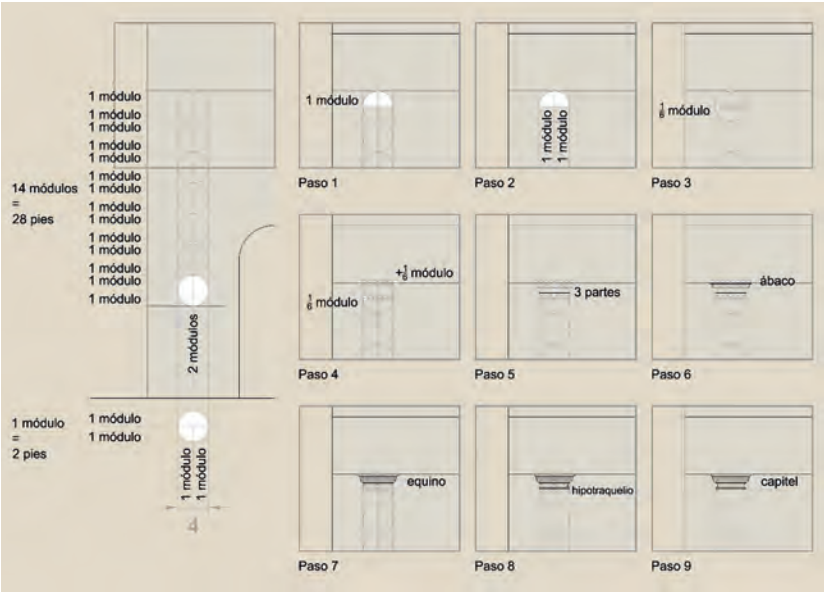



Figura 7. Representación de un capitel dórico paso a paso según las reglas proporcionales establecidas por Vitruvio. Fuente: elaborada por Estefanía López Salas.




Figura 8. Pasos 4 al 17 en la recuperación de las trazas del alzado al campo del arco a la puerta de la Bisagra y resultado final. Cotas y escala en pies. Fuente: elaborados por Estefanía López Salas.

Una vez representada la altura total, el tratadista romano señala que era necesario dividirla en tres partes de las cuales una sería para el ábaco junto con su cimacio, otra para el equino junto con los anillos y la tercera parte para el hipotraquelio. Con el dibujo de estas relaciones proporcionales, paso a paso, obtenemos una representación rigurosa de los capiteles dóricos del primer arco triunfal (Figura 7).

Los anteriores son ejemplos concretos de cómo aplicar la gramática clásica que repetimos sucesivamente para cada elemento del arco: basa, pedestal, arquitrabe, friso, cornisa..., de modo que, poco a poco, vamos recuperando su traza. Como resultado de este proceso se crearon no menos de 70 planos parciales en los que, como si se tratara de los fotogramas de una película, cada arco se visualiza paulatinamente hasta definir sus trazas completas en planta, alzados y secciones (Figura 8).

En este proceso fue también muy importante la consulta de modelos de arcos triunfales próximos a nivel tipológico o temporal. En tal sentido, contamos con los propios arcos de triunfo romanos ilustrados en los tratados de arquitectura, así como con los dibujos, grabados o pinturas conservados de los arcos triunfales realizados para eventos festivos previos o inmediatamente posteriores al recibimiento de Isabel de Valois en Toledo. Estos referentes nos ayudan a disponer de información muy útil sobre los modos de hacer este tipo de aparatos.

En el caso del arco a la puerta de la Bisagra, en la relación se indica que, sobre cada una de las columnas, por encima del nivel de la segunda cornisa, se colocó «por remate, un vaso antiguo con llamas de fuego que salían de él» (Gómez de Castro, 1561: 121). Este es un elemento decorativo que, como tal, carece de unas relaciones proporcionales sistematizadas por los tratadistas. Para reconstruirlo, se tomaron como referentes diversos grabados de arcos triunfales con este tipo de remate. Asimismo, dado que la forma así obtenida es hipotética, estos elementos se distinguieron con un color diferente en la traza digital. Así lo hacemos para que,

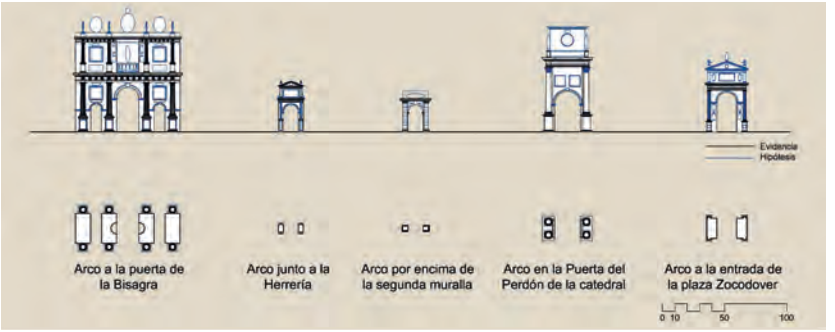



Figura 9. Alzados principales y plantas de los cinco arcos de triunfo de la entrada de la reina Isabel de Valois en Toledo en 1560: primera versión de las trazas recuperadas a partir de la relación de sucesos de Gómez de Castro. Fuente: elaborados por Estefanía López Salas.

tanto el proceso de visualización como el producto resultante, expresen de forma clara el grado de conocimiento que representan (London Chapter, 2009; ICOMOS, 2017).

La combinación de las tres vías de aproximación descritas tanto al conjunto como a cada elemento de esta estructura nos permite recuperar la imagen perdida de la arquitectura de los cinco arcos triunfales erigidos para la entrada de Isabel de Valois en Toledo. En la Figura 9 mostramos las trazas recuperadas correspondientes al alzado principal y planta de dichos arcos efímeros objeto de estudio en el proyecto BIDISO 6. Representan una primera versión que podemos calificar de «desnuda», pues carecen aún de todos los elementos iconográficos que revestían aquellas máquinas con el fin de comunicar un determinado mensaje.

## 4. RECUPERACIÓN DE IMAGINARIOS FESTIVOS: MÉTODO Y RESULTADOS

En este apartado vamos a mostrar la labor llevada a cabo en el ámbito del mencionado proyecto en relación con el rastreo y localización de los elementos iconográficos que ofrezcan una correspondencia temática y formal lo más estrecha posible con los componentes icónicos de aquellas construcciones efímeras. El objetivo fundamental es la integración de las imágenes aportadas, no sólo en las trazas previamente recuperadas, sino también en un futuro modelo digital tridimensional, obteniendo de ese modo un conocimiento más completo del programa diseñado para la decoración de la entrada real que aquí nos ocupa. Nuestro punto de partida es, de nuevo, la detallada crónica del acontecimiento realizada por Gómez de Castro, que nos permite disponer de una descripción literaria somera, pero suficiente a la hora de abordar el propósito que nos hemos planteado. En esta fase inicial del trabajo, hemos desarrollado una suerte de protocolo metodológico en paralelo aplicado, hasta la fecha, a la búsqueda y localización de las diversas figuras y composiciones incluidas en el primer arco del recorrido —el situado junto a la puerta de la Bisagra—, con el fin de que esta experiencia pueda servir para abordar posteriormente con mayor eficiencia la recuperación de la iconografía de las cuatro estructuras restantes.

A partir de las referencias textuales contenidas en la crónica, y teniendo muy en cuenta como referente comparativo otras construcciones coetáneas de similares características, cuyo aspecto sí nos es conocido a través de estampas o dibujos, se ha llevado a cabo una labor de rastreo de aquellas efigies o escenas que mejor se pudieran adaptar a los detalles ornamentales y simbólicos de las máquinas temporales. Nuestro criterio prioritario ha sido la selección de las imágenes que respondieran en el mayor grado posible a la convergencia entre dos variables principales:

1. La primera, de carácter formal en el sentido de tratar de obtener imágenes que presenten la mayor similitud posible

—o, cuando menos, un parecido razonable— con respecto a lo que se detalla en la mencionada relación, atendiendo tanto a los rasgos de índole tipológica (si se trata de representaciones pintadas, o bien de figuras escultóricas en relieve o de bulto) o temática (máxima semejanza del personaje/figura o episodio que protagonizan con lo apuntado en la crónica), como compositiva (esto es, que los personajes y/o elementos se agrupen o interactúen en escenas similares a las apuntadas en el texto de referencia). Además, ha de tenerse en cuenta el formato/disposición de las imágenes en el sentido de su ajuste a las características del marco o espacio al que estaban destinadas en las estructuras arquitectónicas (intercolumnios, pedestales, enjutas, remates...). Finalmente, deben ser de igual modo tomadas en consideración las indicaciones —cuando las hay— acerca del material con que estaba realizada la obra o, en determinados casos, aquel que se intentaba imitar, como sucedía con los relieves de piedra o bronce simulados.

2. La segunda variable fundamental es de índole cronológica: se busca encontrar imágenes datadas con la mayor proximidad posible al acontecimiento, ya sean inmediatamente anteriores, ante la posibilidad de que hayan podido funcionar como fuentes documentales del imaginario de los arcos efímeros, o bien posteriores. En algunos casos, cuando no hemos conseguido respetar esa coincidencia temporal en imágenes afines tipológica o temáticamente, tuvimos que ir retrasando la cronología, habiendo dispuesto el siglo XVIII como fecha límite, con el fin de obtener el mayor rigor posible en los resultados.

La exploración de las imágenes se ha llevado a cabo mediante una doble vía. Por un lado, hemos realizado una consulta sistemática de bibliografía ilustrada, tanto impresa como manuscrita, coetánea

en la medida de lo posible al momento de la celebración: manuales mitográficos, repertorios de personificaciones alegóricas, libros de emblemas, series específicas de estampas, catálogos heráldicos, series de retratos... Por otro, a través de los habituales motores de búsqueda en la *web*, hemos realizado rastreos introduciendo las palabras clave referidas tanto a los personajes y composiciones como a sus atributos, y ensayando con distintas combinaciones e idiomas. En los resultados icónicos obtenidos con este segundo procedimiento, más aleatorio, siempre se ha corroborado *a posteriori* la autoría y, sobre todo, la cronología y procedencia de estos con el fin de sopesar convenientemente su idoneidad.

De este modo, se ha elaborado un corpus de imágenes susceptibles de ser usadas en la posterior labor de revestimiento icónico de las trazas previamente recuperadas de las arquitecturas efímeras, así como en su futuro modelado tridimensional (Figura 10). Para cada elemento iconográfico descrito por Gómez de Castro, al que asignamos un código de identificación, proponemos una o más imágenes como resultado del proceso de localización. Se establece así una clasificación en tres grupos en función de su grado de correspondencia con la fuente textual de partida:

1. Propuestas con un elevado grado de correspondencia en relación con el descriptivo de referencia (color verde).
2. Propuestas en las que existen similitudes, pero también diferencias significativas con respecto al relato de la crónica, por ejemplo, en los atributos de los personajes o en las composiciones de las escenas (color amarillo).
3. Ausencia de propuestas, en aquellos casos en los que no nos ha sido posible hasta el momento encontrar imágenes adecuadas, ni siquiera aproximadas a las descripciones de la relación de sucesos (color rojo).

Asimismo, si bien la tipología/temática de los ítems visuales aportados para la recreación del arco levantado frente a la puerta

| Cód. | Elemento iconográfico | Propuestas | GC | Fuente |
|---|---|---|---|---|
| N1 | Neptuno |  | ● | Cartari, 1608: 4 |
| | |  | ● | Cartari, 1608: 229 |
| | |  | ● | Cartari, 1608: 234 |
| | |  | ● | Du Choul, 1559: 88 |
| N2 | Tritón |  | ● | Alciato, 1549: 63 |
| | |  | ● | Cartari, 1608: 226 |
| N3 | Proteo en forma de viejo |  | ● | Alciato, 1549: 6 |
| | |  | ● | Alciato, 1549: 150 |
| N4 | Aeolo | | ● | |
| N5 | Viejo | | ● | |
| N6 | Cirene | | ● | |
| N7 | Las Alciones |  | ● | Alciato, 1549: 40 |
| N8 | Aretusa y Parténope |  | ● | Cartari, 1608: 241 |
| N9 | Galatea | | ● | |
| N10 | Viejo Nero |  | ● | Galle, 1586: ? |
| ... | ... | ... | ... | ... |

Figura 10. Extracto del corpus de elementos iconográficos localizados para el arco a la puerta de la Bisagra. Fuente: elaborado por José Julio García Arranz y Manuel Alberto Santos Miñambres.

de la Bisagra es en apariencia diversa, hemos podido establecer tres categorías principales: personajes y escenas de temática mitológica, personificaciones y composiciones de carácter alegórico, y «otras temáticas», en la que se incluyen escudos heráldicos o retratos de reyes. Con la finalidad de ilustrar mejor el proceso y los resultados alcanzados en la creación de este corpus, vamos a mostrar a continuación, para cada tipología/temática, las fuentes principales consultadas y algunos ejemplos concretos de propuestas de imágenes con diferente grado de correspondencia.

Comenzando por los abundantes elementos visuales de carácter mitológico que integraban este conjunto ornamental, en el proceso de búsqueda de estos se ha de señalar la importancia de dos tratados impresos de referencia que, además de haber sido editados por primera vez en fechas ligeramente anteriores al año 1560 y de haber experimentado una notable difusión en lo sucesivo, nos han servido de herramienta para identificar y proponer un buen número de imágenes con un elevado grado de correspondencia. De hecho, estas conexiones visuales entre lo escrito por Gómez de Castro y lo representado gráficamente en esos dos títulos creemos que hacen más que evidente que los que se responsabilizaron de la decoración del arco a la puerta de la Bisagra conocían tales fuentes iconográficas y las utilizaron como modelo.

La primera de las obras a las que nos referimos fue editada en 1549 en la imprenta lionesa de Guillaume Rouillé. Se trata de una traducción al castellano por parte de Bernardino Daza Pinciano de los célebres *Emblemata* de Andrea Alciato, titulada *Los emblemas de Alciato traducidos en rimas españolas*, y en la que identificamos una serie de figuras que se ajustan bien a determinadas descripciones de la referida crónica. Este es el caso de las *picturae* pertenecientes a los emblemas de las páginas 63, 150 o 159[3], en las que, respectivamente,

3. Correspondientes con los lemas «Que d'el estudio de las letras nasçe la inmortalidad», «Que qualquiera invençion es muy antigua» y «Las Graçias», respectivamente.




Figura 11. Dos ejemplos de elementos iconográficos con alto grado de correspondencia. Fuente: elaboración propia.

aparecen representadas las figuras de Tritón, Proteo y las Gracias dotadas de una disposición y de unos atributos en perfecta consonancia con los pormenores ofrecidos por el cronista para referirse a la ornamentación de dos de los netos del arco (Gómez de Castro, 1561a: 100-101) y de uno de los cuadros ubicados en los intercolumnios de su primer cuerpo (Gómez de Castro, 1561a: 111) (Figura 11).

Sucede algo similar con el segundo de los títulos, las *Immagini degli dei degli antichi* de Vincenzo Cartari, una obra publicada en Venecia en 1556, con un amplio repertorio de xilografías de temática mitológica que nos ha permitido plantear algunas propuestas iconográficas íntegramente fidedignas para con la crónica: así sucede, por ejemplo, con la imagen de Hércules que figura en la página 314 de la edición impresa en 1608 por Pietro Paolo Tozzi, que es claramente equiparable a las palabras dedicadas por el cronista a la decoración de uno de los intercolumnios del segundo cuerpo del arco (Gómez de Castro, 1561a: 115); en otros casos, nos ha facilitado una serie de ilustraciones que, combinadas mediante herramientas de edición digital de imágenes, han posibilitado la recreación de escenas

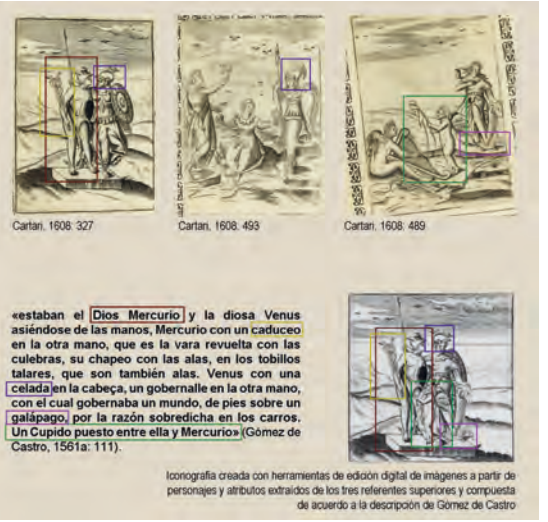



Figura 12. Historia de bronce del dios Mercurio y la diosa Venus, que ocupaba el intercolumnio inferior izquierdo de la cara frontal del arco a la puerta de la Bisagra, creada a partir de varios referentes localizados en la obra de Cartari. Fuente: elaboración propia.

de mayor complejidad compositiva, como la del cuadro que ocupaba el intercolumnio derecho del alzado principal del primer cuerpo del arco, en la que interactúan Mercurio, Venus y Cupido (Figura 12).

Por último, conviene apostillar que otros libros ilustrados, como el *Discours de la religion des anciens Romains* de Guillaume Du Choul (Lyon, 1556), también nos han brindado algunas soluciones iconográficas de temática mitológica; sin embargo, aún persisten algunos casos, como sucede con las representaciones de Cirene o de Eolo, para los que no hemos logrado aportar propuestas mínimamente convincentes por el momento.

El segundo grupo de imágenes se encuadra dentro de la categoría de personificaciones y composiciones de carácter alegórico, también muy presentes en la decoración de la estructura. Para su localización hemos rastreado, en primer lugar, varios tipos de repertorios. Hemos recurrido de manera preferente a aquellas referencias que, por afinidad temática y cronológica, pudieron servir de fuente visual directa. En el caso del imaginario del arco a la puerta de la Bisagra contamos con el manuscrito titulado *Relación de la christiana rogativa que con Christianissimo coraçón la*

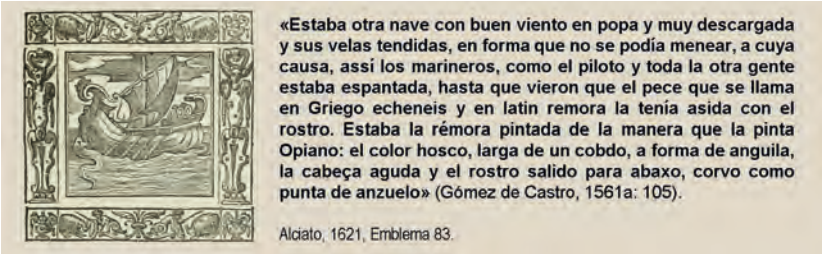



Figura 13. Referente visual para la duodécima iconografía de los netos del arco a la puerta de la Bisagra. Fuente: elaboración propia.

*imperial cibdad de Toledo hizo por la salud del muy alto y muy poderoso Príncipe don Carlos nuestro señor* (Alcalá de Henares, 1562). Una de sus ilustraciones nos proporcionó el precedente de la «Pintura de la Enciclopedia en forma de mujer», que nos permite recuperar la imagen del cuadro que se situaba en el lateral este del segundo cuerpo del primer aparato festivo (López Poza, 1996: 144-145; Fernández Travieso, 2011: 304).

Para otras composiciones, como ya se indicó más arriba, también la literatura de emblemas ha constituido un útil campo de análisis en el que nos hemos encontrado ante una doble situación. En algunos casos, hemos hallado *picturae* emblemáticas coincidentes con representaciones figuradas del arco ya existentes con anterioridad a la crónica y que, por tal razón, podrían considerarse fuentes directas de inspiración. De esta primera situación es ejemplo la imagen descrita por Gómez de Castro de la duodécima composición de los netos del arco, que se adecua perfectamente al emblema de Andrea Alciato con el lema *IN FACILE A VIRTUTE DESCISCENTES*, presente en todas las ediciones desde la *princeps* de Augsburgo de 1531 (Figura 13).

Sin embargo, para otras descripciones tan solo hemos hallado imágenes correspondientes a las *picturae* de emblemas posteriores a la entrada. Así ocurre en el caso de otra representación de las situadas al nivel de los netos, en la que Gómez de Castro describe unos personajes y una acción que podemos relacionar

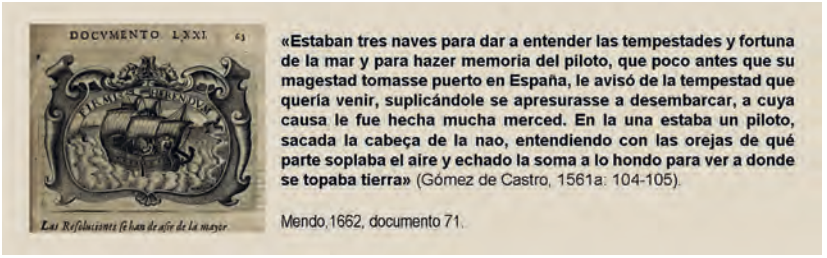



Figura 14. Referente visual para la undécima iconografía de los netos del arco a la puerta de la Bisagra. Fuente: elaboración propia.

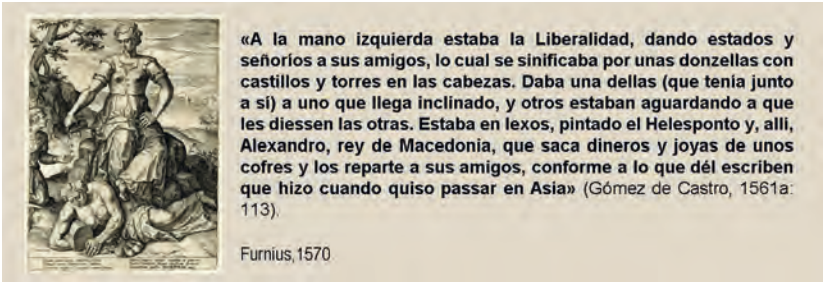



Figura 15. Referente visual para la historia de la Liberalidad. Fuente: elaboración propia.

con la *pictura* de un emblema de Juan de Solórzano (1653: emblema 46) o Andrés Mendo (1662, documento LXXI), ambos con el lema *FIRMIS HAERENDUM* (Figura 14).

No obstante, dado que nos encontramos ante figuras de carácter alegórico, nuestro trabajo se ha orientado fundamentalmente al seguimiento de series y colecciones de personificaciones del siglo XVI, esencialmente flamencas e italianas, pertenecientes sobre todo a libros impresos o series grabadas, incluyendo las ediciones de la *Iconología* de Cesare Ripa, aunque estas sean cronológicamente posteriores al evento, pues la primera versión ilustrada de esta obra vería la luz en Roma en 1603. Esta labor de rastreo, en ocasiones ardua, nos ha permitido descubrir, por ejemplo, una estampa de Pieter Furnius con el título *Liberalitas et Avaritia* (ca. 1570) bastante afín al correspondiente detalle de la

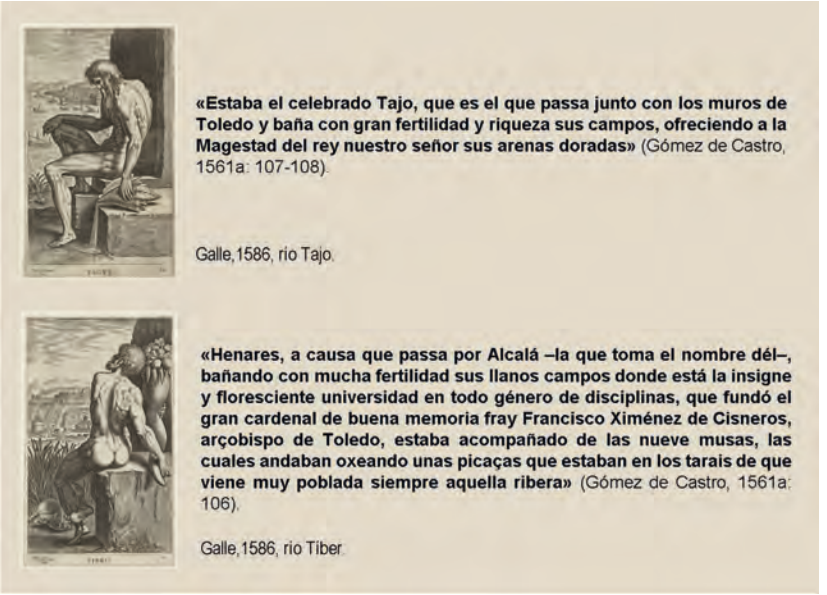



Figura 16. Referentes visuales para las personificaciones de los ríos Tajo y Henares. Fuente: elaboración propia.

crónica para la historia en bronce de la *Liberalidad*, que ocupaba el intercolumnio izquierdo inferior de la cara posterior del arco (Figura 15).

Finalmente, en cuanto a otras figuraciones de carácter geográfico, nos hemos servido de la serie *Personificaciones de ríos* impresa por Philip Galle en Amberes (1586). Además de la propia representación de *Tagus*, que se ajusta a la imagen del río Tajo descrita por Gómez de Castro (1561a: 107-108), otras figuras, como la del *Tibris* (Tíber), se pueden asimilar sin dificultad a otras descripciones de la relación de sucesos, como la correspondiente al río Henares (Figura 16).

En último lugar, las imágenes que clasificamos en el grupo de «otras temáticas», además de diversos escudos heráldicos, incluyen varios de los motivos que se correspondían con retratos de reyes godos, especialmente vinculados a la ciudad de Toledo.

Aunque bastante tardíos, de estos hemos localizado ejemplos idóneos en las estampas que Arnold van Westerhout realizó para la edición de la *Historia general de España* de Juan de Mariana (Madrid, 1797) en las figuras de Recaredo I, Sisebuto, Sisenando o Wamba.

## 5. UN PROYECTO INTEGRAL: ALGUNAS CONCLUSIONES

Una vez presentados los métodos y principales resultados provisionales del presente estudio interdisciplinar de una relación festiva, nos gustaría, a modo de conclusión, señalar en primer lugar que el trabajo realizado constituye una aportación novedosa en cuanto a que supone la recuperación gráfica integral tanto del soporte arquitectónico como del imaginario icónico de los aparatos efímeros descritos en la crónica de Gómez de Castro. Hemos recuperado las trazas de planta, alzado y sección para los cinco arcos del recorrido, obteniendo una primera versión basada exclusivamente en los contenidos de la mencionada relación. En el caso de la iconografía, nos hemos centrado hasta la fecha en el análisis de una de las máquinas festivas, si bien la más compleja desde el punto de vista del número y variedad de sus componentes icónicos. El análisis conjunto de los resultados obtenidos en este proceso muestra que, de 51 elementos rastreados, hemos aportado ejemplos con alto grado de correspondencia en 32 casos (62,7 %), imágenes con necesidad de cambios o ajustes en quince (29,4 %) y, en tan solo cuatro figuras, no nos ha sido posible, hasta la fecha, localizar referentes apropiados (7,8 %). Por tanto, entendemos que los resultados provisionales son satisfactorios, de modo que el método desarrollado ya se está aplicando a los cuatro aparatos temporales restantes.

Las dificultades a las que nos hemos enfrentando en la recuperación de las trazas han surgido, principalmente, a raíz de la ausencia de ciertos datos dimensionales o compositivos en la crónica de Gómez de Castro, que será necesario completar a la

luz de los contenidos en otras relaciones disponibles sobre este o similares acontecimientos. En cuanto al proceso de búsqueda de los elementos iconográficos, el mayor inconveniente al que nos hemos enfrentado posiblemente sea el derivado de la necesidad de proporcionar imágenes en las que el grado de correspondencia con la descripción de la crónica fuese total, sobre todo en el caso de personificaciones o escenas de cierta complejidad. En este sentido, el porcentaje de ejemplos identificados que pueden ser considerados como fuente visual directa, y que por tanto no necesitan de modificación en alguno de sus elementos para responder plenamente a la descripción de la relación de sucesos, ha resultado bastante bajo.

Como consecuencia de lo anterior, una vez creado el corpus de imágenes, se hace necesario un trabajo paciente de edición digital destinado a la composición de aquellas pinturas o esculturas que, por ejemplo, no presentan los mismos atributos que los descritos por Gómez de Castro, o bien incluyen sólo algunos de los personajes de la escena completa descrita en la relación. Son únicamente dos ejemplos de la múltiple casuística que ha sido posible detectar hasta la fecha, y de la que hemos presentado algunas muestras en este trabajo; pero la realidad es que el abordaje de esta fase del proceso, previa a la integración de los componentes icónicos en las trazas recuperadas, requiere de un estudio específico que detallaremos en futuras publicaciones. En cualquier caso, los resultados alcanzados que aquí presentamos ya permiten obtener una nueva percepción y una comprensión mejorada del acontecimiento festivo toledano al exhibir en su dimensión visual lo que solo estaba descrito, lo cual ha sido posible con el uso de herramientas informáticas avanzadas para el proceso de diseño y, sobre todo, con la fusión de estrategias de análisis y saberes múltiples.

## REFERENCIAS BIBLIOGRÁFICAS


Alciato, Andrea (1549). *Los emblemas de Alciato traducidos en rimas españolas*. Lyon: Guillaume Rouille.

Alciato, Andrea (1621). *Andreae Alciati Emblemata cum commentariis Claudii Minois I. C. Francisci Sanctii Brocensis & notis Laurentii Pignorii...* Padua: Pietro Paolo Tozzi.

Antolini, Margherita (2020). Operational methodology for a historical, critical and virtual reconstruction of Baroque ephemeral apparatuses. En: *2020 IMEKO TC-4 International conference on metrology for archaeology and cultural heritage proceedings*. Trento (Italia): IMEKO, pp. 617-622. https://www.imeko.org/publications/tc4-Archaeo-2020/IMEKO-TC4-MetroArchaeo2020-117.pdf

Calvete de Estrella, Juan Cristóbal (1552). *El felicísimo viaje del muy alto y muy poderoso príncipe don Felipe*. Anuers: Casa de Martin Nucio. https://bvpb.mcu.es/es/consulta/registro.do?id=439026

Cartari, Vincenzo (1608). *Immagini degli dei degli antichi*. Padua: Pietro Paolo Tozzi.

Du Choul, Guillaume (1559). *Discorso della religione antica de Romani*. Lyon: Guillaume Rouille.

Fernández González, Laura (2017). Virtual worlds: visualizing early modern festivals in the Iberian world. *Newsletter of the association of Spanish and Portuguese historical studies*, 7: 6-13. https://eprints.lincoln.ac.uk/id/eprint/25234/

Fernández González, Laura (2019). Imagining the past? Architecture and public rituals in Early Modern Lisbon, Madrid and Valladolid. *Bulletin of Spanish Visual Studies*, 3(2): 1-16. https://doi.org/10.1080/24741604.2019.1621091

Fernández Travieso, Carlota (2006). La cultura emblemática en la entrada en Toledo de Isabel de Valois en 1560. En: Anthony J. Close y Sandra María Fernández Vales (eds.). *Edad de oro cantabrigense: Actas del VII Congreso de la Asociación Internacional de Hispanistas del Siglo de Oro*. Cambridge: Asociación Internacional del Siglo de Oro (AISO), pp. 245-254.

Fernández Travieso, Carlota (2011). Reutilización de elementos del programa iconográfico creado para el recibimiento a Isabel de Valois en Toledo en 1560. En: Rafel Zafra Molina, José Javier Azanza López (coords.). *Emblemática trascendente: hermenéutica*

*de la imagen, iconología del texto*. Pamplona: Universidad de Navarra, pp. 301-310.

Ferrer Valls, Teresa (2003). La fiesta en el Siglo de Oro: en los márgenes de la ilusión teatral. En: *Teatro y fiesta del Siglo de Oro en tierras europeas de los Austrias*. Madrid, SEACEX, pp. 27-37. https://www.uv.es/entresiglos/teresa/pdfs/FiestaCatalogo.pdf

Gómez de Castro, Álvar (1561a). *Recebimiento que la imperial ciudad de Toledo hizo a la magestad de la reina nuestra señora doña Isabel…*, edición y estudio de Carlota Fernández Travieso, 2007. A Coruña: Sociedad de Cultura Valle Inclán.

Gómez de Castro, Álvar (1561b). *Recebimiento que la imperial ciudad de Toledo hizo a la magestad de la reina nuestra señora doña Isabel…*, Estudio introductorio de Julia de la Torre Fazio, 2012. Málaga: Servicio de Publicaciones e Intercambio Científico de la Universidad de Málaga.

González Ramos, Roberto (2018). Vergara, Nicolás de. El Viejo. En: Real Academia de la Historia. *Diccionario Biográfico electrónico (DB~e)*. Disponible en: https://dbe.rah.es/biografias/5294/nicolas-de-vergara [actualizada 3 mayo 2018; citada 26 diciembre 2021].

González Román, Carmen (2019). Proyectos digitales para la re-creación/re-construcción de la cultura escenográfica y performativa en la Edad Moderna. Un estado de la cuestión. En: Rosa Romojaro (ed.). *Las Humanidades en el mundo digital, el mundo digital en las Humanidades*. Valencia: Tirant lo Blanch, pp. 103-118.

González Román, Carmen (2021). Scenographic festival books. Towards a multisensory archive. En: A. von Rosen & V. Kjellmer (eds.). *Scenography and art history: performance design and visual culture*. London: Bloomsbury Visual Arts, pp. 165-182.

González Román, Carmen y Risatti, Rudi (2021). Exposición virtual. De la vida real al mundo del arte. Sobre el encuentro entre las artes visuales y escénicas en la cultura festiva europea de los siglos XVI al XVIII. Disponible en: https://www.artes-exhibition.digital/es/exposicion/ [actualizada 10 junio 2021; citada 26 diciembre 2021].

ICOMOS (2017). Los principios de Sevilla. Principios internacionales de la arqueología virtual. Disponible en: https://icomos.es/wp-content/uploads/2020/06/Seville-Principles-IN-ES-FR.pdf [actualizada 6 junio 2020; citada 26 diciembre 2021].

London Chapter (2009). La carta de Londres para la visualización computarizada del patrimonio cultural. Disponible en: https://www.londoncharter.org/downloads.html [actualizada 7 febrero 2009; citada 26 diciembre 2021].

López Poza, Sagrario (1996). Emblemas españoles manuscritos en Toledo en 1562. En: Sagrario López Poza (ed.). *Literatura emblemática hispánica: actas del I Simposio Internacional*. A Coruña: Universidade da Coruña, pp. 129-174.

López Poza, Sagrario (1998). Introducción. En: Sagrario López Poza y Begoña Canosa Hermida (eds). *Jorge Báez de Sepúlveda. Relación verdadera del recibimiento que hizo la ciudad de Segouia a la magestad de la reyna nuestra señora doña Anna de Austria, en su felicissimo casamiento que en la dicha ciudad se celebró*. Segovia: Fundación Don Juan de Borbón: 7-39.

Marías, Fernando (1983). *La arquitectura del Renacimiento en Toledo (1541-1631), I*. Madrid: CSIC/Instituto Provincial de Investigaciones y Estudios Toledanos.

Marín Tovar, Cristóbal y Román Aliste, Sergio (2020). La entrada en Madrid de la Reina Margarita de Austria: reconstrucción digital de las arquitecturas efímeras. En: *Workshop: fiesta, arte efímero y humanidades digitales: virtualización y digitalización como instrumentos de rescate. Reflexiones, problemáticas y estudios de caso*. Madrid: UNED. Disponible en: https://canal.uned.es/video/5f5b73d35578f253117f5bc3 [actualizada 6 octubre 2020; citada 26 diciembre 2021].

Marsden, C. A. (1960). Entrées et fêtes espagnoles au XVI[e] siècle. En: Jean Jacquot (ed.). *Les fêtes de la Reinassance, II. Fêtes et cérémonies au temps de Charles Quint*. París: Éditions du Centre National de la Recherche Scientifique, pp. 389-411.

Mendo, Andrés (1662). *Príncipe perfecto y ministros aiustados, documentos políticos, y morales*. Lyon: Horacio Boissat y George Remeus.

Oliver Domingo, José Luis (1997). *Vitruvio. Los diez libros de arquitectura*. Madrid: Alianza Forma.

Paniagua Soto, José Ramón (1997). La teoría de la arquitectura en España en el siglo XVI. Algunas consideraciones sobre las fuentes de la literatura. *Anales de Historia del Arte*, 7: 231-244. https://revistas.ucm.es/index.php/ANHA/article/view/ANHA9797110231A

Redondo, Augustin (1997). Fiesta, realeza y ciudad: Las relaciones de las fiestas toledanas de 1559-1560 vinculadas al casamiento de Felipe II con Isabel de Valois. En: Sagrario López Poza y Nieves Pena Sueiro (eds.). *La fiesta. Actas del II seminario de relaciones de sucesos.* Ferrol: Sociedad de Cultura Valle Inclán, 303-313.

Ripa, Cesare (1603). *Iconologia*. Roma: Lepido Facii.

Solórzano Pereyra, Juan (1653). *Emblemata Regio Politica in centuriam unam redacta*. Madrid: Domingo García Morrás.

Soto Caba, Victoria y Solís Alcudia, Isabel (2019a). «Adereçados y pintados de pinçel». Una recreación virtual: policromía efímera en la Lisboa de 1619. *Revista de humanidades digitales*, 3. https://doi.org/10.5944/rhd.vol.3.2019.23172

Soto Caba, Victoria y Solís Alcudia, Isabel (2019b). De la policromía efímera. Metodología e informática para una recreación virtual del color. Los arcos triunfales lisboetas del desembarco de Felipe III (1619). En: Inmaculada Rodríguez Moya (coord.). *El rey festivo: palacios, jardines, mares y ríos como escenarios cortesanos (siglos XVI-XIX)*. Valencia: Universitat de Valéncia, pp. 67-80.

Varela Flor, Susana (2020). Os arcos das festas de casamento de D. Afonso VI e D. Maria Francisca de Sabóia (1666): uma reconstituição possível entre 1662 e 1687. En: *Workshop: fiesta, arte efímero y humanidades digitales: virtualización y digitalización como instrumentos de rescate. Reflexiones, problemáticas y estudios de caso*. Madrid: UNED. Disponible en: https://canal.uned.es/video/5f7ed5285578f2675f49d06c [actualizada 7 octubre 2020; citada 26 diciembre 2021).

# 3D MÁS ALLÁ DE LA RECONSTRUCCIÓN VIRTUAL:
## UNA HERRAMIENTA AL SERVICIO DE LA GESTIÓN INTEGRAL DEL PATRIMONIO EN LA ERMITA DE SANT LLORENÇ DE LA SENABRA

Pablo del Fresno Bernal, Esther Travé Allepuz, Isidre Ot Padilla, José Manuel Martínez Torrecilla, Josep Socorregut Domenech

**Resumen**

Los trabajos realizados en la ermita de Sant Llorenç de la Senabra (Alt Penedès, Barcelona) en la última década nos han llevado a plantear la necesidad de acudir a la reconstrucción virtual como elemento polivalente, tanto para tareas de investigación de base, como de difusión y trabajo colaborativo. Esta ermita románica posteriormente transformada en masía, ha sido objeto de actuaciones arqueológicas y de conservación-restauración realizados intermitentemente desde 2011. En el marco de la redacción de un Plan Director orientado a la patrimonialización del lugar, la reconstrucción

Pablo del Fresno Bernal, Sistemes de Gestió de Patrimoni SCCL
0000-0001-8775-9113
pdfsgp@gmail.com

Esther Travé Allepuz, Universitat de Barcelona – Institut d'Arqueologia
0000-0002-6769-4487
esther.trave@ub.edu

Isidre Ot Padilla, Virtual Past
0000-0002-2043-0767
i.otpadilla@hotmail.com

José Manuel Martínez Torrecilla, Qark Arqueología SL
0000-0002-0688-2505
torre@qark.es

Josep Socorregut Domenech1, Sistemes de Gestió de Patrimoni SCCL
0000-0003-1234-722
josepsoco@gmail.com

virtual nos ha permitido generar distintos modelos para la interpretación y difusión de los resultados y propuestas para la posible creación de un equipamiento cultural construido sobre las ruinas del edificio.

En el proceso de reconstrucción virtual surgen necesidades, problemas metodológicos y reflexiones que traemos a colación en este trabajo, en el que presentamos los resultados obtenidos hasta el momento y la metodología empleada para ello. Hacemos hincapié en la necesidad de ahondar en los procesos de normalización del registro y en las posibilidades de integración de un registro tridimensional, construido sobre unas bases científicas sólidas, en la investigación acerca de la realidad estructural de este patrimonio edificado, sus características y las posibilidades viables de conservación y revalorización del mismo.

**Palabras clave:** Reconstrucción virtual 3D, Gestión Integral, Patrimonio construido, Documentación virtual, Sant Llorenç de la Senabra

**Abstract**

In the last decade, research carried out in the hermitage of Sant Llorenç de la Senabra (Alt Penedès, Barcelona) has led us to consider the need to develop virtual reconstruction as a multipurpose element, both for basic research tasks and for dissemination and collaborative purposes. This Romanesque hermitage was transformed into a farmhouse, and has been the object of discontinuous archaeological and conservation-restoration work since 2011. Within the propitious framework of a Master Project aimed at the heritage preservation at this place, we created virtual models for the interpretation and dissemination of the results, together with some proposals for the further creation of a cultural equipment built upon the remains of the building.

The virtual reconstruction of this site was a thought-provoking process that led us to face methodological problems leading to the reflective outcome introduced in this paper. We aim at offering the results obtained and the methodology used. We will emphasize the need to delve into the normalization processes of the register and the possibilities of integrating a three-dimensional register, based upon solid research, within the structural reality of built heritage, assessing the building features and the possibilities of feasible conservation and dissemination as a cultural value.

**Key words:** 3D Virtual Reconstruction, Integrated Management, Built Heritage, Virtual Documentation, Sant Llorenç de la Senabra

## 1. INTRODUCCIÓN: NUEVOS ESCENARIOS Y NUEVOS RETOS EN LA GESTIÓN DEL PATRIMONIO

La reconstrucción virtual y las representaciones tridimensionales constituyen en la actualidad un elemento cada vez más presente en la arqueología, especialmente como herramienta de difusión. Las denominadas Humanidades Digitales se han interesado por este campo y, a lo largo de los últimos años, el recurso a la tridimensionalidad, de uno u otro modo, se ha convertido en un elemento cotidiano de los trabajos arqueológicos. En cualquier caso, resulta imprescindible a nuestro juicio una reflexión serena acerca de las posibilidades que este nuevo escenario nos ofrece. Nuestra contribución pretende dar a conocer los resultados que en este sentido hemos obtenido en Sant Llorenç de la Senabra (Santa Margarida i els Monjos, Alt Penedès, Barcelona) desde el inicio de los trabajos de recuperación del lugar hasta hoy, en el marco de un proyecto de gestión integral del Patrimonio.

Estos trabajos nos han situado en la tesitura de reflexionar acerca de los procesos de restitución virtual, de la metodología empleada para ello y, en especial, de las finalidades y objetivos de la misma. Es indiscutible, al margen del indudable valor difusivo, que el uso de las reconstrucciones virtuales tiene un valor nada menospreciable en la investigación arqueológica de base, pero –en contra de una creencia que presumimos excesivamente extendida– dicho valor no es intrínseco. En el ámbito profesional, la reconstrucción virtual adquiere su verdadero sentido cuando se convierte en una herramienta útil para resolver problemas y responder a preguntas de calado. De lo contrario, las más de las veces acaba por convertirse en un ente finalista, un producto destinado más a impresionar y atraer a un público no especializado (que en muchas ocasiones carece incluso de los elementos necesarios para valorar la calidad y veracidad de la restitución final) que en un recurso destinado a contribuir realmente a la ciencia arqueológica.

Por todo ello, nuestro trabajo pretende ser una propuesta para el diálogo y la reflexión colectiva acerca de los procesos, la

metodología de la restitución, la trazabilidad de la información que la posibilita y el potencial de dichas restituciones con un doble objetivo: por un lado, la investigación arqueológica de base desarrollada en múltiples escenarios; y, por otro, las estrategias de difusión patrimonial como un elemento cultural de primer orden en la necesaria socialización del conocimiento. Presentamos a lo largo de las siguientes páginas las bases y líneas maestras de un proyecto de gestión integral del patrimonio que nos ha permitido con los años poner en valor un yacimiento singular haciendo uso de la restitución tridimensional y, que, a su vez, ha generado el marco propicio para una reflexión acerca de las razones por las que trabajamos en entornos virtuales en contextos de gestión patrimonial y de las herramientas con que contamos para ello.

Sin olvidar el ámbito de las humanidades digitales, y haciendo especial hincapié en la idea de *humanidades*, este trabajo se ha gestado en un marco colaborativo e interdisciplinar que ha favorecido sinergias entre los sectores público y privado y entre las distintas disciplinas dedicadas a la preservación, análisis y difusión del patrimonio y en concreto del patrimonio construido. Sistemes de Gestió de Patrimoni SCCL ha sido la encargada de coordinar la redacción de un Plan Director, y para ello ha configurado un equipo de trabajo integrado por empresas especializadas en la conservación y restauración (ÀBAC SL), en la investigación y gestión del patrimonio construido (QARK Arqueología SL), en la reconstrucción virtual (Virtual Past) y en la ilustración de evocaciones históricas (Dbòlit).

El proyecto ha sido financiado por la Diputación de Barcelona y el Consorcio del Parc del Foix quienes han mostrado su interés por la revalorización patrimonial de un enclave presidido por una ermita que en 2011 se hallaba en un estado de completo abandono. Las administraciones públicas propiciaron los trabajos arqueológicos iniciales y, ante los resultados obtenidos, promovieron la redacción de un proyecto que definiera las necesidades principales y que estableciera las actuaciones

necesarias de gestión integral del patrimonio. Diez años después, el trabajo al servicio del patrimonio edificado empieza a dar sus frutos y ha permitido que, en un marco de trabajo colaborativo, podamos reflexionar sobre las necesidades del mismo, e integrar el elemento 3D como una de las soluciones para algunas de estas necesidades. Nos hallamos por tanto en un momento propicio para dar a conocer y discutir en el ámbito académico algunas de estas inquietudes a la luz de los resultados obtenidos.

## 2. EL YACIMIENTO: LA ERMITA DE SANT LLORENÇ DE LA SENABRA Y SU GESTIÓN INTEGRAL

La ermita de Sant Llorenç es un bien cultural de interés local (BCIL) incluido en el Catálogo de Patrimonio Cultural catalán y se sitúa en un pequeño promontorio, a una altura de 344 m, localizada a unos 3 km hacia el este de la iglesia de Santa María de la Senabra y algunos espacios de hábitat anexos, a la derecha del torrente homónimo, y se accede a ella a través de un sendero que nace al fondo del torrente y al pie del promontorio. La zona se integra en el Parc del Foix, un parque natural protegido que forma parte de la red de parques naturales de la Diputación de Barcelona. Con anterioridad al inicio de los trabajos, esta pequeña iglesia románica presentaba un estado de degradación y ruina notables (Mauri, Travé y Del Fresno, 2013). Una primera prospección inicial de la zona, permitió realizar un reconocimiento del terreno, elaborar la topografía inicial del enclave, una documentación de los restos constructivos y el estudio de algunos materiales de superficie, con cronologías posteriores al siglo XV (Mauri, Travé y Del Fresno, 2011).

Pese que no tener referencias escritas para la iglesia anteriores a 1305, el lugar aparece ya en la documentación en 992, con el topónimo *Sinapsis*, y en 993 (*Senabre*) (Pladevall 1992: 141). La ermita, que por sus características estructurales y arquitectónicas data del siglo XII, habría organizado tal vez un

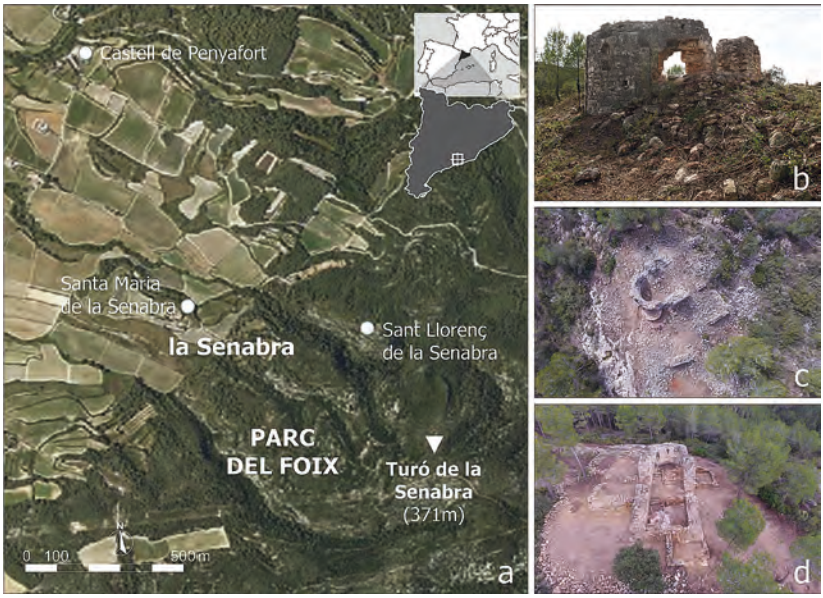



Figura 1. Localización del yacimiento (a) y vistas generales del mismo: Ábside de la ermita con los restos del horno cubiertos por el derrumbe durante los trabajos de prospección en 2011 (b), vista aérea del conjunto tras la excavación y consolidación del horno (c) y vista aérea del conjunto tras la excavación de la ermita en 2018 (d).

pequeño núcleo de poblamiento (Díaz 2020: 23-25) –del cual no tenemos constancia material a tenor de las excavaciones realizadas, aunque aparece frecuentemente mencionado en las fuentes escritas– que posteriormente se habría trasladado al llano alrededor de la iglesia de Santa María, de construcción algo más tardía. Entre ambos edificios de culto es posible el contacto visual y ambos parecen experimentar una historia compartida, con transformaciones similares que tendremos ocasión de comentar. La ermita de Sant Llorenç es un edificio románico, del cual sólo quedan en pie parte de los muros laterales y un ábside decorado con bandas lombardas.

Los trabajos iniciales, ya permitieron detectar la existencia de al menos dos fases diferenciadas, que fueron corroboradas

por las excavaciones posteriores. Las sucesivas campañas arqueológicas han puesto de manifiesto una ocupación del lugar que podemos sintetizar en dos grandes fases históricas (Del Fresno y Socorregut, 2018). Un primer momento de construcción y uso de la iglesia románica, probablemente en funcionamiento desde el siglo XII y hasta el XIV; y un segundo momento de época moderna, en el cual el edificio fue reutilizado como centro neurálgico de una pequeña explotación agrícola y ganadera que transformó la iglesia ya en ruinas en un espacio de hábitat a lo largo de los siglos XVIII y XIX (Fig. 1).

Así, a través del ensanchamiento de una de las ventanas a la derecha del ábside, se construyó un horno relacionado con la vivienda instalada en el antiguo edificio de culto, que transformó el ábside en cocina como ya atestiguaba una de las imágenes más recientes del edificio tomada en 1912, antes de la degradación creciente del mismo. Dicha evolución no parece anómala en este territorio, puesto que una transformación similar fue acaecida en la vecina iglesia de Santa María. En el momento inicial inmediatamente anterior a la intervención, los trabajos de prospección habían permitido delimitar la extensión general de los niveles de derrumbe, identificar los restos del horno y definir una primera sectorización que –sin correspondencia con ningún criterio interpretativo de la historia del lugar– pretendía simplemente definir la estrategia organizativa del trabajo por realizar.

La complejidad y el estado de conservación precario de los restos aconsejaron plantear una intervención por fases. Durante la primera fase se llevó a cabo la excavación y consolidación del horno (Del Fresno, 2012), su consolidación se realizó en 2016, momento que supuso una parada temporal de los trabajos a la espera de comprobar la efectividad de la restauración y el efecto del paso del tiempo. En 2018 se retomaron las excavaciones que dieron a conocer las características de todo el conjunto y su evolución, seguidas por una intervención de conservación preventiva de




Figura 2. Vistas de la ermita tras su excavación, en la que se documentaron las transformaciones y cambios de uso que transformaron el ábside en una cocina (a) con un horno anexo (b) y distintas dependencias adosadas al antiguo templo (c). La excavación permitió también definir la planta completa del edificio de culto, parcialmente oculta por los derrumbes (d) y fue objeto de una conservación preventiva (e-f).

los abundantes y deteriorados restos materiales documentados correspondientes a ambas fases constructivas (Fig. 2).

A partir de aquel momento, la ermita presentó la problemática común a muchos otros elementos patrimoniales, es decir, la ausencia de estrategias de gestión integral del sitio que permitan su revalorización patrimonial y, en muchas ocasiones, la propia integridad estructural. Por este motivo, en el año 2020, la dirección del parque encargó la redacción de un Plan Director que estructurase y coordinase la secuencia de actuaciones que permitiesen tanto la salvaguarda del conjunto, como su socialización y gestión como equipamiento cultural integrado dentro de la estructura del parque y la red patrimonial municipal. En todo este proceso, tanto en los trabajos arqueológicos iniciales de carácter prospectivo como la excavación posterior, la documentación fotogramétrica y la restitución tridimensional han constituido herramientas valiosas para la interpretación del conjunto, por lo que deberían seguir siéndolo en las estrategias futuras de preservación, puesta en valor y difusión.

## 3. RETOS METODOLÓGICOS: HERRAMIENTAS DE REVALORIZACIÓN EN UN ENTORNO NORMALIZADO

Los objetivos del Plan Director se organizaron según la Cadena de Valor del Patrimonio Cultural (Barreiro, 2013; Criado-Boado, 1996), que entendemos como la secuencia de instancias valorativas que intervienen en el proceso de estudio y gestión del patrimonio cultural con el fin de que un objeto o edificio pueda ser considerado bien patrimonial. Es decir, el carácter de bien patrimonial de un elemento determinado es el resultado final de la implementación de los distintos eslabones de esta cadena, de tal modo que se reivindica el carácter *construido/otorgado* del valor patrimonial que la sociedad confiere a un bien determinado, que no necesariamente tiene valor por sí mismo.

En nuestro caso, estas fases de puesta en valor parten necesariamente de la significación del monumento, que pasa por interpretar, entender y contextualizar cada una de las fases de ocupación y abandono del conjunto. El estado de los restos descubiertos debe ser evaluado a fin de valorar el estado de conservación y los riesgos estructurales, ambientales y sociales que podrían comprometerlo. La finalidad de ello es garantizar la conservación del mismo a partir de la redacción de proyectos de conservación estructurados en fases de actuación, es decir, un proyecto de conservación preventiva que ponga freno a la degradación del patrimonio, un proyecto curativo que permita sanear la estructura y, finalmente, un proyecto de restauración que contemple la integración de elementos nuevos que restituyan la estructura. Este valor patrimonial que adquiere el monumento tras su puesta en valor implica también la definición de dinámicas de circulación y difusión que permitan una socialización del conocimiento generado y que debe conseguirse a través de la redacción y ejecución de proyectos de adecuación y narrativa (Fig. 3). Las tareas encaminadas a significar, evaluar, conservar y difundir, por tanto, constituyen las líneas maestras de este proyecto de puesta en valor, en línea con las propuestas actuales del sector (López y Lorés, 2021).

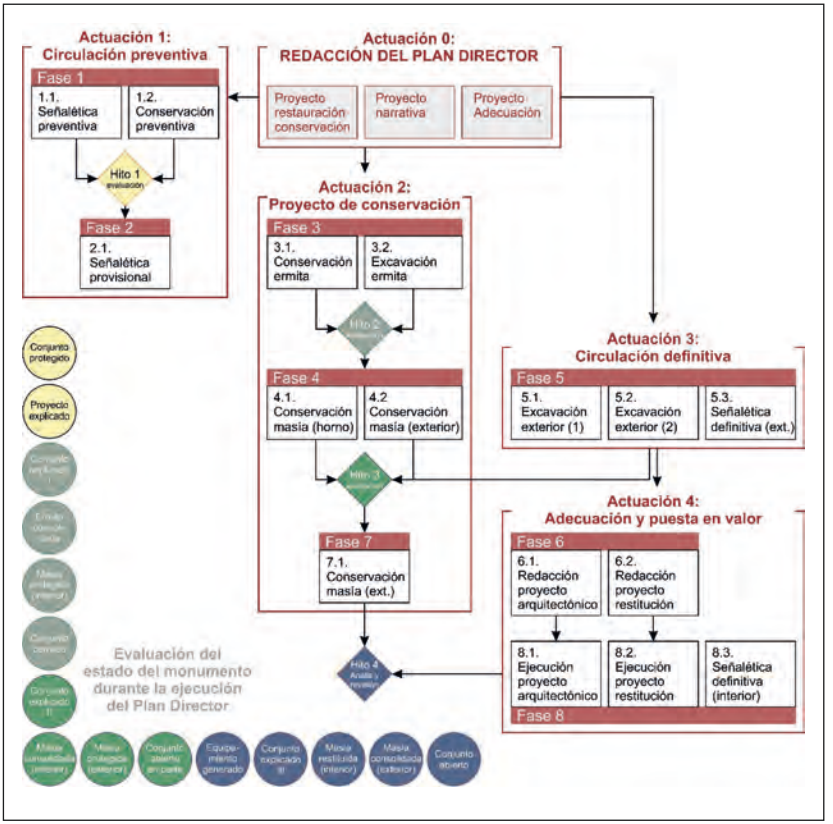



Figura 3. Tabla de Pert (Olañeta, 2021) del desarrollo general del Plan Director, en la que se consideran las distintas actuaciones a realizar, las tareas que conllevan cada una de ellas y los hitos de evaluación de la marcha del proyecto.

La información compilada, analizada y estructurada en el Plan Director debe constituir la base, entre otras cuestiones, para la redacción de un proyecto arquitectónico que permita la transformación del edificio en un equipamiento cultural, tal como han planteado las administraciones competentes. Es de este aspecto, en concreto, de donde surge la necesidad de modelar en entornos virtuales en dos sentidos o con dos objetivos diferenciados: por un lado desde la perspectiva de la interpretación arqueológica del conjunto, es decir, planteando cuál habría sido el aspecto inicial

de la ermita durante su uso como edificio de culto –la altura, los volúmenes, la cubierta, etc.– y también cómo podemos interpretar los hallazgos que atestiguan la transformación de la ermita en una vivienda o masía; y por otro lado aventurar el aspecto final que podría tener el equipamiento cultural, valorar su posible impacto visual en la zona, la adecuación de los materiales constructivos, o el encaje entre las evidencias arqueológicas y los elementos de reconstrucción, entre otros aspectos. Para trabajar en esta línea, nos resulta imprescindible adoptar sistemas de normalización de la información que permitan el trabajo colaborativo e interdisciplinar entre todos los agentes implicados en la redacción y ejecución de los distintos proyectos que conforman el plan director y, a su vez, establecer los momentos de evaluación durante el transcurso de su desarrollo.

En este sentido, hemos acudido a metodologías y desarrollos ya existentes que dan respuesta a los problemas planteados que encajan con nuestro entorno conceptual e instrumental actual. Esta cuestión se materializa en la plataforma de gestión integral Horai, cuyas bases parten de la creación de una propuesta de normalización aplicada a la gestión del registro arqueológico (Del Fresno, 2016). Nos referimos de manera específica a un modelo de datos basado en los conceptos de Unidad Topográfica, Unidad Estratigráfica y Actor. Estos conceptos, de los que se nutre el modelo de datos subyacente en Horai, permiten integrar el uso de distintos lenguajes normalizados para el registro de procesos en la reconstrucción virtual sobre los que discutiremos a lo largo de nuestro trabajo. En definitiva, se trata de gestionar la información generada en cada fase de desarrollo del Plan Director en base a una metodología de trabajo colaborativa destinada a centralizar la información del proyecto en un modelo de información digital creado por todos sus agentes.

### 3.1. UNIDADES ESTRATIGRÁFICAS Y UNIDADES TOPOGRÁFICAS

Dado que nuestro trabajo parte necesariamente de un planteamiento arqueológico, la Unidad Estratigráfica (UE), según la define E.C. Harris (1979) como evidencia material de una acción sucedida en el pasado, es el concepto base que opera en nuestra gestión de la información. Sin embargo, la vinculación intrínseca de la definición de UE con la materialidad limita en cierto modo con la integración en el registro arqueológico de aquellos elementos carentes de dicha materialidad, como la información procedente de las fuentes gráficas, escritas, orales o de cualquier otro origen o naturaleza. Por ello, nuestro sistema de gestión de la información Horai, utiliza también el concepto de Unidad Topográfica (UT) como indicador de la evidencia de una acción sucedida en el pasado (Mauri, 2006), con independencia de la materialidad de la misma o de la naturaleza de la fuente que proporciona dicha información.

El concepto de UT está directamente vinculado al de Actor, como partícipe de dicha acción. No abundaremos en este trabajo sobre las fuentes informativas acerca de los actores, muy habitualmente de carácter escrito ya sea bibliográfico o documental y cuya relevancia es limitada en el proceso de reconstrucción virtual. En todo caso, nos remitimos a la literatura existente sobre el tema (Travé *et al.*, 2020; 2021) y a otros trabajos que versan de manera más específica sobre el concepto de Actor y su normalización (Medina y Travé, 2021). El trabajo realizado en Sant Llorenç de la Senabra, sin embargo, se ha servido del concepto de UT para identificar y expresar distintos elementos informativos que constituyen la base de las reconstrucciones virtuales que presentamos en este trabajo, definiendo una dialéctica entre UT y UE en función de la materialidad del dato sobre la que ya nos hemos referido en otros trabajos al respecto (Travé *et al.*, 2021: 49).

La tabla 1 que presentamos a continuación ofrece algunos ejemplos de la atribución de distintos elementos informativos relacionados con el yacimiento en cuestión que contribuyen a su restitución virtual con independencia de las fuentes que informan sobre ellos.

**Tabla 1: Extracto del registro de UT y UE, en donde se consignan sus atributos principales y las fuentes de obtención. A fin de evitar redundancias, no consignamos en la tabla el atributo 'localización' que es para todos los casos la ermita de Sant Llorenç de la Senabra.**

| UT | DEFINICIÓN | BIÓTICO | ABIÓTICO | ANTRÓPICO | CRONOLOGÍA | FUENTE(S) |
|---|---|---|---|---|---|---|
| 01 | Tejado de la masía | | | Construcción | 1912 | Gráfica: fotografía |
| 02 | Chimenea | | | Construcción | 1912 | Gráfica: fotografía |
| 03 | Horno | | | Construcción | 1912 | Gráfica: fotografía |
| 04 | Visita pastoral | | | Visita, registro | 1303 | Documental |
| 05 | Espadaña (existencia) | | | Construcción, Testimonio | 1303 | Documental |
| 06 | Altar (existencia) | | | Construcción, Testimonio | 1303 | Documental |
| 07 | Estado de abandono | | | Situación, Testimonio | 1303 | Documental |
| 08 | Fisuras del edificio | | Meteorización | | post 1912 | Reconocimiento visual |
| 09 | Desprendimientos | | Meteorización | Vandalismo | post 1912 | Reconocimiento visual |
| 10 | Tejado de la ermita | | | Elemento virtual | 2021 | Reconstrucción virtual |
| 11 | Espadaña | | | Elemento virtual | 2021 | Reconstrucción virtual |
| 12 | Puertas | | | Elemento virtual | 2021 | Reconstrucción virtual |
| 01 | Muro | | Piedra | Elemento constructivo | s. XII | Registro arqueológico |
| 02 | Pavimento | | Cal, arena | Elemento constructivo | s. XII | Registro arqueológico |
| 03 | Jamba | | Piedra | Elemento constructivo | s. XII | Registro arqueológico |
| 04 | Saqueo | | | Interficie | s. XVIII | Registro arqueológico |
| 05 | Limpiezas | | | Interficie | 2016 | Trabajos de restauración |
| 06 | Consolidaciones | | Material constructivo | Conservación curativa | 2016 | Trabajos de restauración |

### 3.2. REPRESENTACIONES, NORMALIZACIÓN Y PROCESOS DE RECONSTRUCCIÓN VIRTUAL.

Recientemente, se han llevado a cabo algunos esfuerzos en vistas a una normalización de los procesos de reconstrucción virtual, con el fin de definir lenguajes formales mediante los cuales realizar un seguimiento de dichos procesos para documentar de manera rigurosa la actividad científica que subyace en una propuesta de restitución determinada y validar, a través de la trazabilidad del proceso, los distintos elementos restituidos. Una de las propuestas más sólidas en este sentido es la definición de la denominada *Extended Matrix* (EM) o matriz extendida (Demetrescu, 2015), que aúna en una representación bidimensional –a la manera de las matrices tradicionalmente usadas en arqueología– tanto los datos arqueológicos procedentes del registro de excavación como los elementos virtuales añadidos durante la reconstrucción y documentados como *unidades estratigráficas virtuales* (UEV; USV son las siglas de uso habitual en inglés) para las cuales se consignan también las fuentes de información que posibilitan su existencia y el grado de fiabilidad de las mismas.

Tendremos ocasión de reflexionar más adelante sobre el potencial destacable de estas aplicaciones, que por ahora sólo hemos explorado de manera incipiente en el desarrollo del proceso de reconstrucción virtual de Sant Llorenç de la Senabra. En cualquier caso, cabe destacar la versatilidad del concepto de Unidad Topográfica para englobar en su definición también los procesos de reconstrucción virtual que en otros casos se han propuesto como UEV i USV, dado que derivan también del concepto de Unidad Estratigráfica.

Al margen de la matriz extendida, los esfuerzos por validar la calidad de las reconstrucciones virtuales a partir de la transparencia en la determinación de la fiabilidad de los distintos elementos reconstruidos, que tienen grados distintos, ha sido objeto también de otras propuestas de normalización y representación (Aparicio y Figueiredo, 2016). El carácter visual

sobre la reconstrucción virtual de este tipo de propuestas supone un elemento elocuente que transmite al propio espectador la escala de fiabilidad de los elementos reconstruidos de manera llana y sencilla y que, por lo tanto, constituye un ejercicio de transparencia, especialmente hacia el gran público, a la hora de interpretar el resultado. Nuestras propuestas de reconstrucción tridimensional hacen uso habitual de este tipo de representaciones.

Dicha codificación se incorporó al resultado final de la reconstrucción, elaborada tomando como punto de partida el modelo fotogramétrico, debidamente escalado y georreferenciado. La fotogrametría sobre la que se han trabajado las distintas reconstrucciones que presentamos ha sido realizada mediante la toma de fotografías bidimensionales a partir de las cuales se ha generado el modelo 3D. Existen, ciertamente, otras técnicas de captura tridimensional, como el escáner laser. No obstante, dada la alta calidad -tanto geométrica como visual- que ofrecen en la actualidad los resultados fotogramétricos, así como el menor coste de este proceso frente a otras técnicas de digitalización activa, se ha considerado la fotogrametría la mejor técnica para nuestros propósitos. La propuesta de reconstrucción virtual se realizó utilizando como programa principal Blender, software de libre disposición (*open source*) cada vez más empleado en el ámbito de la arqueología virtual. Tras la validación de la hipótesis virtual, que hubo de ser modificada distintas veces a lo largo del proceso, se llevó a cabo la obtención de distintos pases de render que, posteriormente, fueron postprocesados en otros softwares de edición gráfica. Finalmente, en el caso de las recreaciones virtuales se utilizaron técnicas de matte painting para recrear la posible ambientación.

## 4. RESULTADOS: PROPUESTAS DE ESTUDIO, RESTITUCIÓN VIRTUAL E INTERPRETACIÓN DEL YACIMIENTO

El recurso a la tridimensionalidad mediante la utilización, de las técnicas y herramientas de registro y documentación digitales como elemento para la investigación arqueológica de base ha sido una práctica habitual en la ermita de Sant Llorenç desde que se iniciaron los trabajos en 2011. De hecho, deberíamos referirnos propiamente a la *cuatridimensionalidad*, puesto que la variable temporal supone un elemento indisociable de cualquier análisis arqueológico. Los trabajos de prospección inicial nos llevaron a realizar un modelo de terreno en base a la toma sistemática de puntos y la generación de curvas de nivel. Dicha reconstrucción permitió valorar de forma fiable, el alcance de los volúmenes de derrumbe que colmataban el yacimiento a la vez que lo protegían

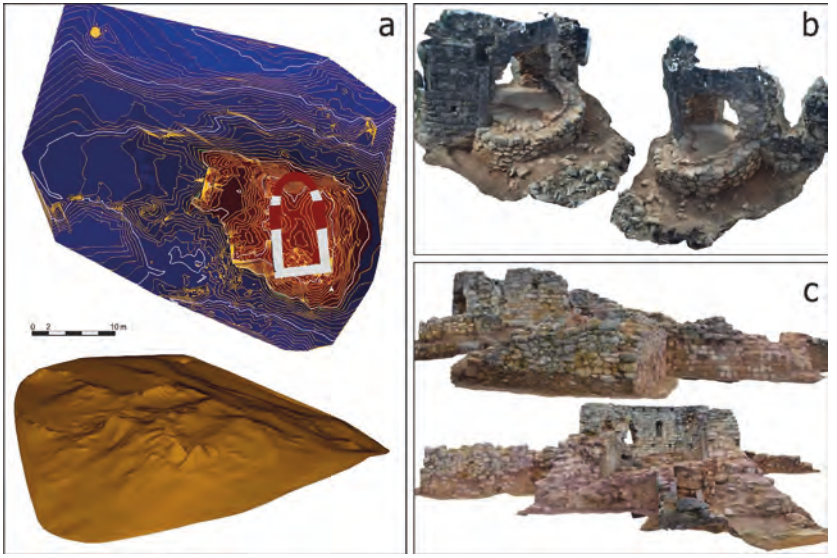



Figura 4. Modelo de terreno tridimensional del yacimiento previo a su excavación realizado en función de las curvas de nivel (a) y alzados fotogramétricos del horno (b) y del conjunto de la ermita y edificios anexos (c).

de su degradación (Fig. 4a). Como resultado de las intervenciones posteriores realizadas para llevar a cabo la excavación puntual del horno en 2012 (Fig. 4b) y del conjunto de la ermita en 2018 (Fig. 4c) se realizaron modelos fotogramétricos de los conjuntos intervenidos. A partir de esta práctica, habitual en numerosos yacimientos hoy en día, se inició un proceso de bifurcación de vías para los trabajos de investigación y la difusión que es lo que nos ha llevado a reflexionar sobre los problemas que aún existen en los procesos de reconstrucción tridimensional y virtualización.

Utilizando el modelo fotogramétrico como base, se procedió a realizar las distintas reconstrucciones virtuales del edificio en sus distintas fases como un elemento tanto de investigación arqueológica como de difusión. Una de las claves para este tipo de reconstrucciones es valorar de forma correcta la información de la que se dispone para reconstruir. Es decir, partir del análisis del registro arqueológico y documental (planimétrico, bibliográfico y gráfico) del propio yacimiento para las estructuras que todavía se conservan, y contar con paralelos fiables para las partes que no se conservan es el procedimiento recomendable, incluso cuando se cuenta con documentación textual explícita sobre de la existencia de cualquier elemento arquitectónico ausente).

En el caso de la ermita –nos referimos específicamente al primer uso del edificio, antes de ser transformado en un espacio de hábitat y producción–, y pese al pobre estado de conservación del edificio, contábamos con una exhaustiva documentación planimétrica y tridimensional de los restos fruto de las intervenciones realizadas entre los años 2011 y 2018. Además, la existencia de algunas fotografías tomadas a principios del siglo XX (Vidal, 1912) nos ayudaron a definir las características del ábside (como su altura) y su decoración de bandas verticales y arcuaciones o arquillos característicos del románico lombardo. La pervivencia de los muros laterales del edificio y parte del ábside proporcionaban, asimismo, información suficiente para proyectar su continuidad.

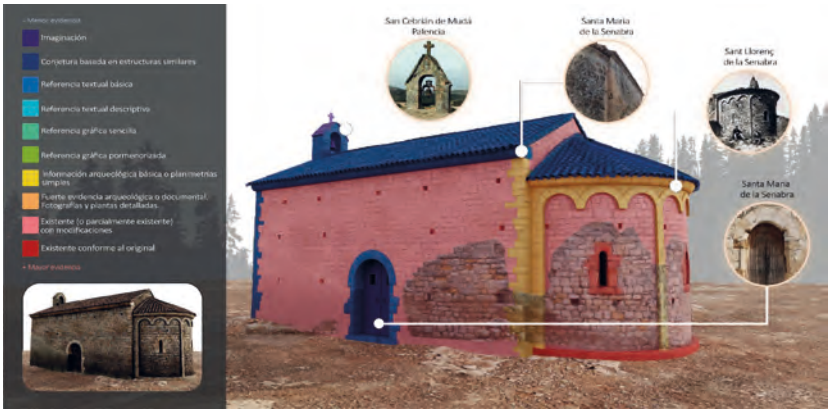



Figura 5. Restitución virtual de la ermita de Sant Llorenç de la Senabra, codificada a través de la escala de evidencia histórico-arqueológica (Aparicio y Figueiredo, 2016) para ilustrar la fiabilidad de la existencia y características de los elementos no conservados.




Figura 6. Diversos paralelos de masías para la representación de las distintas estancias, especialmente de las áreas de cocina (a-c) o de acceso y distribución de dependencias (d). Fuente: Arxiu Fotogràfic de Catalunya. Colección Estudi de la Masia Catalana.

Por otro lado, para los elementos no conservados, tuvimos la suerte de contar con un paralelo evidente en la coetánea iglesia de Santa María, situada a escasa distancia de nuestra ermita. La comparación con la misma nos permitió aventurar la posible apariencia del acceso lateral, la cornisa y la cubierta (Fig. 5). Otros paralelos, como el caso castellano de San Cebrián de Mudá (Palencia), ofrecen algunas informaciones acerca del posible aspecto de la espadaña.

Algo más compleja es la reconstrucción del edificio durante la fase de reutilización como masía, que nuevamente requiere de una detenida comparación de los paralelos documentados (Fig. 6). En este caso, la reconstrucción pretendía comprender las diversas soluciones constructivas desaparecidas y adoptadas en diferentes épocas, pero superpuestas en el mismo espacio, por lo que fue necesario reconstruir las diferentes arquitecturas para comprobar la viabilidad de las teorías adoptadas. Es decir, reconstruimos virtualmente para entender la construcción física ya desaparecida, y su relación con aquello que aún hemos podido documentar in situ.

Los trabajos que hemos realizado en los últimos años a partir de volúmenes reales (virtuales, pero reales, al fin y al cabo) nos han permitido contrastar y corregir algunas de las ideas iniciales respecto a la apariencia de la construcción en sus diferentes fases. Así, por ejemplo, el trabajo de modelado virtual ha permitido descartar la existencia de algunos elementos estructurales arquitectónicos propios de la ermita románica, pero ausentes ya en la masía, como la cubierta de la nave central o la misma cúpula del ábside. Conclusiones, estas, a las que sólo hemos podido llegar tras una restitución virtual aproximada de los volúmenes de ambas construcciones (Fig. 7) en las que se han tenido como referencia, especialmente en el caso de la masía, los escasos recursos gráficos que han llegado hasta nuestros días, en los cuales podía apreciarse la altura original de las diversas naves.

Por otro lado, el uso de las herramientas digitales no sólo ha permitido la reconstrucción del modelo virtual en sus fases

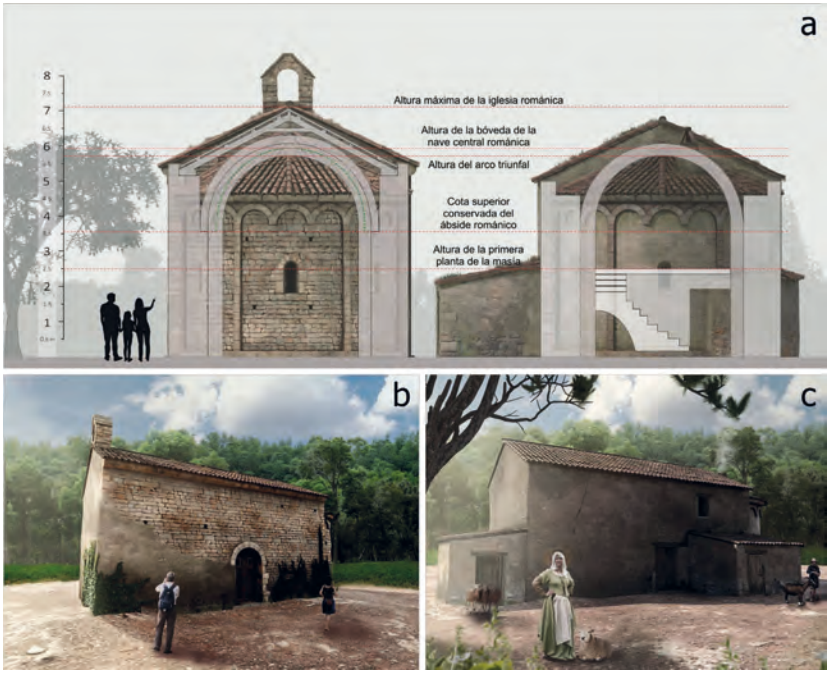



Figura 7. Reconstrucciones virtuales de la ermita de Sant Llorenç a partir de las estimaciones de las dimensiones a partir del registro arqueológico y los paralelos (a), representada como edificio de culto en un escenario contemporáneo posterior a su utilización (b), y como masía en un momento coetáneo a su uso (c).

pretéritas, si no también reflejar de manera visual y detallada una posible propuesta de restitución para este monumento como equipamiento cultural (Fig. 8) en el futuro. La propuesta museográfica, de hecho, contempla la preservación de los elementos propios de la masía a partir de la reconstrucción del piso superior; y es en este sentido que el recurso a la visualización tridimensional ofrece elementos de análisis y valoración del conjunto en las claves propuestas por los BIM, es decir, en un escenario de trabajo en equipo que garantice la interoperabilidad (Eastman et al., 2011: 100).

En el marco del proyecto que ha articulado los trabajos tanto de excavación como de conservación y explicación de los restos, hemos necesitado generar reproducciones 3D que nos




Figura 8. Reconstrucción virtual del lugar como equipamiento cultural museizado y visitable. Alzados tridimensionales del prototipo en una vista elevada de la fachada sur (a) y del ábside (b); desde el interior del equipamiento, con acceso al piso superior de la masía (c) y representada como visualización virtual del equipamiento en uso con elementos explicativos y de difusión para los visitantes (d).

permitieran documentar, entender, interpretar y por último representar aquello que está siendo objeto de estudio y gestión. La necesidad de sistematizar estos trabajos a fin de posibilitar la trazabilidad de los procesos, nos ha llevado a identificar otras necesidades de carácter metodológico, conceptual e instrumental que de momento no hemos podido integrar en nuestra dinámica de trabajo, pero que implican la necesidad de desarrollar líneas de investigación inmediatas sobre las que discutiremos a continuación.

## 5. DISCUSIÓN: MÁS ALLÁ DE LA RECONSTRUCCIÓN VIRTUAL

En la actualidad, son infinitos los ejemplos del uso reconstrucciones o recreaciones virtuales asociados a proyectos de carácter patrimonial. Es evidente que, con la democratización de la tecnología necesaria, la obtención de estas restituciones virtuales

pretéritas está mucho más al alcance de lo que estaban solo hace unos años. Pero, la pregunta de calado que debería subyacer es, ¿Por qué lo hacemos? ¿Con qué finalidad? ¿Hay siempre una justificación detrás de cada reconstrucción virtual? ¿Llegamos a lo virtual por qué podemos, o por qué debemos? No sería descabellado plantear que, en muchas ocasiones, detrás de magníficas propuestas virtuales y puestas en escena prevalecen las motivaciones económicas frente a la mera generación de conocimiento.

La ejecución de este trabajo en Sant Llorenç nos ha llevado a reflexionar acerca del planteamiento de acercamiento a lo virtual. Como arqueólogos e historiadores, creemos que las infinitas posibilidades del mundo virtual deben ir más allá de la simple evocación de una construcción en un momento determinado, sin quitarle un ápice de valor al inmenso poder divulgativo que esta posibilidad representa. Así pues, consideramos que, especialmente en el ámbito profesional, deben ayudar a entender y modelar tanto el propio elemento objeto de estudio, como el contexto histórico en el que esas construcciones se han erigido. En este sentido, nos alineamos con los planteamientos de la plataforma Virtual Multimodal Museum Plus (VIMM+), que fue creada como Acción de Coordinación y Apoyo financiada por el programa H2020 y que vincula a las distintas organizaciones del sector de la Museología Virtual y el Patrimonio Digital con el fin de desarrollar políticas de calidad que promuevan estas líneas de actuación.

Tal como expresan sus principios acerca de reconstrucción virtual, no puede darse por sentado que la reconstrucción virtual es siempre y a priori la mejor de las soluciones para comunicar o investigar acerca del patrimonio cultural. Lo será precisamente en función de la documentación que se genere para cada caso a fin de justificar la elección del método más apropiado y las posibilidades de evaluar su adecuación y trazabilidad (Denard, 2009). Esta necesidad de evaluación del elemento patrimonial y de las estrategias de difusión refuerza la convicción de que teniendo en cuenta la dimensión de realidad cambiante atribuible al Patrimonio Cultural

y en extensión al arqueológico, el registro –cualquier registro, pero más el que se puede obtener a partir del método arqueológico– nos permite entender y explicar las diferentes realidades documentadas, las aún existentes y las ya desaparecidas de un objeto patrimonial. Por ende, nos hacemos eco del planteamiento de A. Rodríguez en relación al carácter patrimonial del propio registro, según el cual "es precisamente esta evolución constante la que origina que la información herede el carácter patrimonial de los elementos que describe dado que, desde el momento en que éstos ya no existen tal y como se documentaron, la información capturada queda como única fuente para su estudio" (Rodríguez 2014: 1). De ahí infiere el autor –y nos sumamos a dicha interpretación– que la información que generamos sobre los elementos patrimoniales es, en sí misma, patrimonio, por lo que debe ser conservada. Parece obvio que la reconstrucción virtual se acoge plenamente a estos principios y que, como tal, debe ser estandarizada y normalizada.

En este sentido, nos resultan especialmente interesantes las propuestas que abogan por la elaboración de matrices –como la matriz extendida (EM) (Demetrescu, 2015) – que permitan documentar los procesos de reconstrucción y la información empleada para ello. El enfoque, en este caso, es el de la lectura estratigráfica, común también al ámbito de la arqueología de la arquitectura, y tiene como objetivo crear un marco común que conecte la documentación arqueológica y la reconstrucción virtual en las primeras etapas de la excavación de un sitio o un estudio 4D de un monumento (Demetrescu y Fanini, 2017), entendiendo que la temporalidad es una variable intrínseca del estudio arqueológico en cualquiera de sus facetas. En todo caso, el planteamiento parte de la necesidad de otorgar valor científico a las restituciones virtuales y de validarlas como parte integrante de un proceso de investigación sólido (Demetrescu, 2018) a través de la creación de un lenguaje normalizado (Demetrescu y Ferdani, 2021).

Los trabajos derivados de esta propuesta se han incrementado en los últimos años y la EM parece consolidarse

como un elemento válido de representación del conocimiento relacionado con la reconstrucción virtual, a pesar del carácter marcadamente arqueológico (Ferdani *et al.*, 2019) que subyace en todas las propuestas. La construcción de la EM para un proceso de reconstrucción virtual determinado implica una codificación específica para cada tipo de fuente que permite la creación de una unidad estratigráfica virtual, mientras que nuestro enfoque, engloba todas estas fuentes bajo el concepto de Unidad Topográfica. Nos remitimos a la tabla 1, mostrada anteriormente, como ejemplo de ello. De todos modos, resulta innegable el valor de la EM como propuesta de normalización de la disparidad de fuentes y procesos que intervienen en la reconstrucción virtual y que, como tales, es necesario integrar. En este sentido, el potencial de la propuesta se ha revelado ya muy prometedor no sólo en la construcción de un registro arqueológico que incluye elementos virtuales, sino también en la posibilidad de integración de los elementos puramente interpretativos o explicativos del yacimiento en su totalidad y en las posibilidades de estas reconstrucciones rigurosas en el campo de la difusión (Demetrescu *et al.*, 2020).

Otro elemento para la reflexión y el debate lo constituyen las propuestas colaborativas en el marco de los Modelos de Información para la Construcción (*Building Information Modeling [BIM])* para la creación y gestión de este tipo de proyectos (Azhar, Khalfan y Maqsood, 2012). Esta necesidad surge, en nuestro caso, de los trabajos de proyección virtual de un equipamiento museográfico para la preservación y difusión del yacimiento y su conversión en un centro de interpretación patrimonial. Desde el momento en que se espera que el destino final de la ermita de Sant Llorenç sea su transformación en un equipamiento cultural accesible para el visitante, la integración de los trabajos realizados en este tipo de propuestas parece que cobra todo su sentido y puede ser un elemento de especial utilidad para el desarrollo de trabajos futuros.

En este sentido, el BIM es un proceso colaborativo desde el inicio de su gestación que aúna bajo un mismo concepto tanto el

proceso como la tecnología que lo posibilita. Facilita a los usuarios (y financieros) potenciales la visualización final de la propuesta y permite cuantificar de manera adecuada los volúmenes constructivos a fin de afinar, entre otros cálculos, los presupuestos de trabajo y material. No exento de debate acerca de la generalización de su uso (Borrman, 2018), el sistema BIM se ha presentado también como un recurso útil para los problemas específicos de los trabajos de construcción directamente relacionados con el patrimonio a partir de la especificidad de un modelo de carácter patrimonial *(Heritage Building Information Modeling [hBIM])*, sobre el cual existen ya algunas propuestas recientes que ilustran aplicaciones prácticas (Godinho *et al.*, 2020) o el potencial de la aplicación para evaluar propuestas de intervención patrimonial, sus efectos o posibles afectaciones negativas sobre un elemento construido con alto valor patrimonial, previo a su ejecución (Lee *et al.*; 2019). En el caso que nos ocupa, la aplicación de las propuestas en el entorno del BIM tiene todavía mucho recorrido por delante y se encuentra simplemente en un estadio inicial. Nos parece necesario, aun así, discutir sobre el potencial, las necesidades de uso, o el recorrido que pueden tener aplicaciones de este tipo.

En cualquier caso, ¿dónde están los límites entre las necesidades de estandarización en el desarrollo de las distintas propuestas? Las propuestas de representación como la matriz extendida o –en nuestro caso– el uso de Unidades Topográficas para identificar ausencias y elementos reconstruidos a partir de la información contenida en otras fuentes complementarias deberían poder utilizarse también para evocaciones interpretativas de otro tipo como la ilustración, por ejemplo (Fig. 9). En este campo, para el que contamos con numerosos ejemplos realizados desde hace décadas (Ollich y Rocafiguera, 2001; Padilla, 2003; López, Álvaro, Travé, 2016) el proceso interpretativo es similar a la reconstrucción virtual; no en vano, también podríamos considerar estas representaciones como reconstrucciones virtuales, si bien en un soporte distinto y mediante una metodología ligeramente

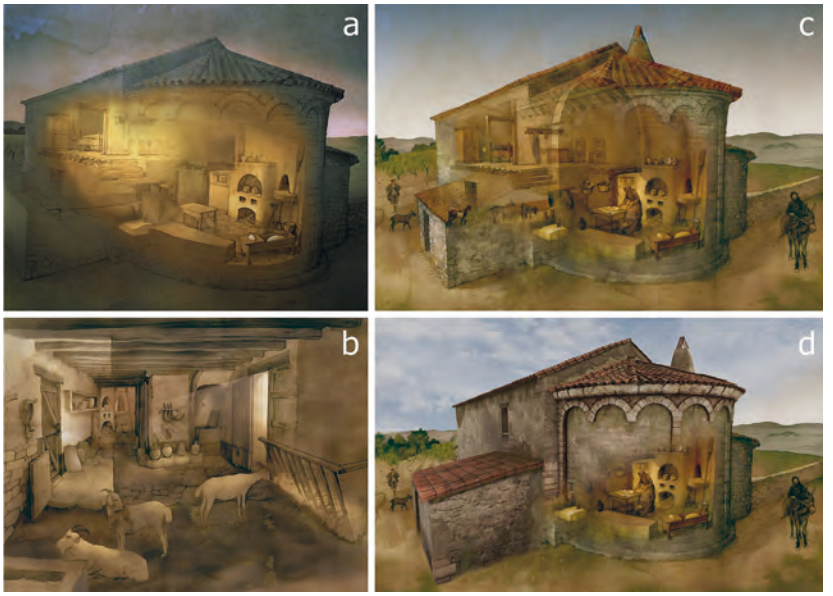



Figura 9. Evocación en un mismo plano ilustrativo de la secuencia evolutiva del yacimiento, que transformó una ermita románica en una explotación agropecuaria. El proceso de creación de la ilustración, implica la generación de bocetos en distintos momentos (a) y escenarios (b), definiendo propuestas interpretativas (c) hasta llegar a la representación definitiva utilizada para la difusión, que en este caso integra la evocación con la reconstrucción tridimensional (d). Ilustraciones elaboradas por Dbòlit.

diferenciada. Parece posible entender que los procesos de registro y normalización de los procesos de reconstrucción virtual podrían aplicarse también a estrategias de este tipo.

En suma, podríamos afirmar que el Patrimonio Arqueológico, aquel que se configura a través de la mirada arqueológica de las fuentes históricas, no sólo está formado por los elementos materiales existentes, sino también por las diferentes evidencias identificadas, por los diferentes indicadores informativos, y por todas aquellas recreaciones documentadas, virtuales o no, que podamos construir a partir del Registro Arqueológico recogido de manera normalizada y sistemática (Del Fresno 2016). Y, a su vez, la sistematización con la que abordemos la construcción de

reproducciones virtuales, es decir, la capacidad que tengamos de transformar los diferentes elementos reproducidos, en unidades informativas que podamos sintetizar junto con aquellas evidencias procedentes de la lectura arqueológica de la ruina material, determinará si estamos ante un ejercicio meramente evocador del pasado, o construyendo herramientas de investigación histórica. En definitiva, cuando reproducimos virtualmente, también debemos documentar históricamente y hacia esta finalidad apuntan la mayoría de herramientas desarrolladas al respecto (Denard, 2009; Ferdani *et al.*, 2019), por lo que parece obvio que no nos encontramos ante una preocupación particular, sino ante un tema de reflexión colectiva. Como casi todas las cuestiones de nuestro campo de procedencia, la Arqueología, las problemáticas son comunes, pero las soluciones no lo son tanto y nunca están exentas de un cierto debate. En nuestro caso, la normalización del registro tanto material como virtual a través de las unidades estratigráficas y topográficas, empleando la codificación como UE para los elementos materiales y la UT para el resto, nos parece una solución adecuada que se ajusta a las herramientas actualmente en desarrollo.

## CONCLUSIÓN

A lo largo de esta contribución hemos presentado los trabajos de reconstrucción virtual realizados en nuestra ermita con el fin de explicar una evolución compleja del edificio marcada por un profundo cambio de uso a lo largo de su historia. Estas virtualizaciones se han desarrollado en el marco de un Plan Director diseñado para establecer la estrategia de gestión patrimonial del yacimiento en pro de su preservación, conservación y difusión en un ejercicio de difusión cultural de unos resultados de tipo arqueológico. En el contexto del hBIM, hemos desarrollado una propuesta de adecuación del espacio a fin de transformarlo en un equipamiento cultural museizado y visitable a través de la reconstrucción virtual.

Este es, precisamente, uno de los puntos fuertes del proyecto: que utilizamos las potencialidades y posibilidades de las herramientas digitales para conjugar, a partir de un modelo 3D, pasado, presente y futuro de Sant Llorenç, articulando a través de él todo el proceso de gestión integral que incorpora las conclusiones aportadas desde los distintos campos profesionales que han intervenido en el proyecto. Y todo ello sin causar un impacto real e irreversible sobre el propio monumento, pues, el trabajo en entornos virtuales nos permite determinar e integrar algunas de las soluciones constructivas ya empleadas en el pasado, valorando la idoneidad de las mismas y, de cara a posibles actuaciones futuras, reducir el impacto visual y estructural de las intervenciones, manteniendo una coherencia arquitectónica e histórica que ayude a la comprensión de la diacronía existente.

Nuestro trabajo es, sobre todo, un punto de partida y de manera deliberada lo hemos presentado como un elemento para la reflexión acerca de la orientación y finalidad de nuestras virtualizaciones y, especialmente, de los procesos de normalización y estandarización que podemos adoptar para garantizar su fiabilidad y trazabilidad. El contexto actual, en entornos interdisciplinares que conforman verdaderos equipos de humanistas digitales, parece el escenario propicio para abundar en dichos debates y propuestas a fin de buscar soluciones adecuadas a unas necesidades nuevas.

Agradecimientos

Este trabajo ha sido posible gracias al apoyo del Consorci del Parc del Foix y la Xarxa de Parcs Naturals de la Diputació de Barcelona, y del Ayuntamiento de Santa Margarida i Els Monjos (Alt Penedès, Barcelona). Nuestra actividad se incluye en la línea de investigación de Arqueología del Paisaje del Grupo de Investigación en Arqueología Medieval y Postmedieval de la Universidad de Barcelona (2021 SGR 00236) del cual son miembros algunos de los autores.

## REFERENCIAS BIBLIOGRÁFICAS


Aparicio, Pablo y Figueiredo, César (2016). El grado de evidencia histórico-arqueológica de las reconstrucciones virtuales. Hacia una escala de representación gráfica. *Otarq. Otras arqueologías*, 1: 235-247.

Azhar, Salman; Khalfan, Malik; Maqsood, Tayyab (2012). Building Information Modeling (BIM): Now and Beyond. Australasian Journal of Construction Economics and Building, DOI: 10.51307AJCEB. V12I4.3032.

Barreiro, David. (2013). *Arqueológicas. Hacia una arqueología aplicada*. Barcelona (España): Bellaterra Arqueología.

Criado-Boado, Felipe (1996). Hacia un modelo integrado de Investigación y Gestión del Patrimonio Histórico: la cadena interpretativa como propuesta. *PH Boletin Patrimonio Histórico,* 16: 73-78.

Del Fresno, Pablo (2012). *Memòria. Fase 1. Excavació del forn annex a l'ermita de Sant Llorenç de la Sanabra. Santa Margarida i els Monjos (Alt Penedès)*. Barcelona (España): Sistemes de Gestió de Patrimoni SCP. Informe inédito. Disponible en: http://calaix.gencat.cat/handle/10687/425334#page=1.

Del Fresno, P. (2016): *Sistema de Información Arqueológica: Propuesta de Normalización, Desarrollo Conceptual e Informático.* [Tesis doctoral inédita], Vitoria-Gasteiz: Universidad del País Vasco. Disponible en: https://addi.ehu.es/handle/10810/18925.

Del Fresno, Pablo y Socorregut, Josep (2018). *Memòria de la intervenció arqueològica preventiva a Sant Llorenç de la Senabre d'acord amb la resolució d'autorització 470 K121 N-661 2018-1/22745 SantaMargarida i els Monjos (Alt Penedès)*. Barcelona (España): Sistemes de Gestió de Patrimoni SCCL. Informe inédito. Disponible en: https://www.academia.edu/40432667/.

Demetrescu, Emanuel (2015). Archaeological stratigraphy as a formal language for virtual reconstruction. Theory and practice. *Journal of Archaeological Science,* 57: 42-55.

Demetrescu, Emanuel (2018). Virtual Reconstruction as a Scientific Tool: The Extended Matrix and source-based Modelling Approach. En: *Digital research and education in Architectural Heritage*, Cham (suiza): Springer International, pp. 102-116.

Demetrescu, Emanuel; D'Annibale, Enzo; Ferdani, Daniele y Fanini, Bruno (2020): Digital replica of cultural landscapes: An experimental

reality-based workflow to create realistic, interactive, open world experiences. *Journal of Cultural Heritage*, 41: 125-141.

Demetrescu, Emanuel y Fanini, Bruno (2017). A White-Box Framework to Oversee Archaeological Virtual Reconstructions in Space and Time: Methods and Tools. *Journal of Archaeological Science: Reports,* 14: 500–514.

Demetrescu, Emanuel y Ferdani, Daniele (2021). From field archaeology to Vrtual Reconstruction: A five steps method using the Extended Matrix. *Applied Sciences*, 11: 5206.

Denard, Hugh (2009). *The London Charter for the Computer-based Visualisation of Cultural Heritage*. Londres (Reino Unido): King's College London.

Díaz Ros, Marçal (2020). *Estudi històric a l'entorn dels paratges de Sant Llorenç de la Senabra (Santa Margarida i els Monjos, Alt Penedès)*. Barcelona (España): Diputació de Barcelona, SPAL (informe técnico inédito).

Eastman, Chuck; Teicholz, Paul; Sacks, Rafael y Liston, Kathleen (2011). *BIM Handbook. A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers, and Contractors*. Hoboken, New Jersey (Estados Unidos): John Wiley & Sons.

Ferdani, Daniele; Demetrescu, Emanuel; Cavalieri, Marco; Pace, Gloriana y Lenzi, Sara (2019). 3D Modelling and Visualization in Field Archaeology. From survey to interpretation of the past using digital tecnologies. *GROMA. Documenting archaeology*, 4: 1-21.

Godinho, Márcia; Machete, Rita; Ponte, Magdalena; Falcão, Ana Paula; Gonçalves, Alexandre B. y Bento, Rita (2020). BIM as a resource in heritage management: An application for the National Palace of Sintra, Portugal. *Journal of Cultural Heritage*, 43: 153-162.

Harris, Edward C. (1979). *Principles of archaeological sratigraphy*, Londres (Reino Unido): Academic Press.

Lee, Jongwook; Kim, Junki: Ahn, Jaehong; Woontack, Woo (2019). Context-aware risk management for architectural heritage using historic building information modeling and virtual reality. *Journal of Cultural Heritage*, 38: 242-252.

López, Olga y Lorés, Immaculada (coords.) (2021). *Planificació estratègica de museus i centres patrimonials*. Manuals de Museologia, 4. Barcelona (España): Associació de Museòlegs de Catalunya.

López, Mª Dolores; Álvaro, Karen y Travé, Esther (2016). *Yacimiento arqueológico de Revenga. Iglesia, necròpolis, aldea. Comunero de Revenga (Burgos)*. Burgos (España): Diputación de Burgos.

Mauri, Alfred (2006). *La Configuració del Paisatge Medieval: El Comtat de Barcelona Fins al Segle XI*. [Tesis doctoral inédita]. Barcelona: Universidad de Barcelona. Disponible en: http://hdl.handle.net/10803/2065.

Mauri, Alfred; Travé, Esther y Del Fresno, Pablo (2011). *Memòria de la prospecció arqueológica a l'ermita de Sant Llorenç de la Senabra. Santa Margarida i els Monjos (Alt Penedès)*. Gelida (España): Sistemes de Gestió de Patrimoni SCP. Informe inédito. Disponible en: https://parcs.diba.cat/documents/178697/410051/p12d104.pdf.

Mauri, Alfred; Travé, Esther y Del Fresno, Pablo (2013). Estudi preliminar de l'ermita de Sant Llorenç de la Senabre. Santa Margarida i els Monjos (Alt Penedès). En: *III Monografies del Foix*, Barcelona (España): Diputació de Barcelona, pp. 150-157.

Medina, Sonia y Travé, Esther (2021): Identificando el sujeto del paisaje: Modelado e integración de la dimensión social para una (re)lectura de la geografía altomedieval. *Scire Vias. Congreso de Humanidades Digitales Hispánicas (HDH2021)*. Santiago de Compostela (España): Comunicación oral (actas en preparación).

Olañeta, Juan Antonio (2021). Disseny i gestió de projectes en museus i equipaments patrimonials. En: *Planificació estratègica de museus i centres patrimonials*. Barcelona (España): Associació de Museòlegs de Catalunya, pp. 97-121 [texto en castellano: pp. 321-335].

Ollich, Imma y Rocafiguera, Montserrat de (2001). *L'Esquerda. 2500 anys d'Història. 25 anys de recerca*. Roda de Ter (España): Fundació Privada l'Esquerda.

Padilla, José Ignacio (2003). *Yacimiento arqueológico de Cuyacabras. Despoblado, iglesia y necròpolis. Eremitorio de Cueva Andrés. Quintanar de la Sierra (Burgos)*. Barcelona (España): Publicacions de la Universitat de Barcelona.

Pladevall, Antoni (1992). *Catalunya Romànica: El Penedès, Anoia.* Barcelona (España): Enciclopèdia Catalana.

Rodríguez Miranda, Álvaro (2014): *Documentación espacial del patrimonio: preservación de la información. Necesidades, posibilidades, estrategias y estánd*ares. [Tesis doctoral inédita], Vitoria-Gasteiz:

Universidad del País Vasco. Disponible en: https://addi.ehu.es/handle/10810/13617.

Travé, Esther; Del Fresno, Pablo; Mauri, Alfred (2020). Ontology-mediated Historical Data Modelling: Theoretical and Practical Tools for an Integrated Construction of the Past, *Digital Humanities Special Issue. Information*, 11: 182.

Travé, Esther; Del Fresno, Pablo; Mauri, Alfred; Medina, Sonia (2021). The Semantics of History. Interdisciplinary Categories and Methods for Digital Historical Research, *International Journal of Interactive Multimedia and Artificial Intelligence*, 6 (5): 47-56.

Vidal, Enric (1912). La conca del riu Foix i sos principals afluents. *Butlletí del Centre Excursionista de Catalunya*, 206: 65-79.

# EL TRABAJO-PROYECTO *PATRIMONIO HERIDO.* UNA PROPUESTA DE INNOVACIÓN EDUCATIVA TRANSDISCIPLINAR

María Marcos Cobaleda, Julia García González,
Antonio Jesús Santana Guzmán, Antonio Cruces Rodríguez,
Nuria Rodríguez-Ortega

**Resumen:**
El objetivo principal de esta contribución es presentar una línea de indagación docente basada en la hibridación entre metodologías de prototipado, estrategias ABP (Aprendizaje Basado en Problemas) e implementaciones tecnológicas, incluida en las acciones de innovación educativa desarrolladas por TransUMA (Laboratorio de Competencias Transdisciplinares de la Universidad de Málaga). Esta publicación se centra específicamente en el trabajo-proyecto Patrimonio Herido, en el que ha participado el alumnado del grado en Historia del Arte y del grado en Turismo de la Universidad de Málaga.

María Marcos Cobaleda, Universidad de Málaga
ORCID: 0000-0002-3742-5913
mmcobaleda@uma.es

Julia García González, Universidad de Granada
ORCID: 0000-0002-7092-7691
juliagargon@ugr.es

Antonio Jesús Santana Guzmán, Universidad de Málaga
ORCID: 0000-0001-8132-192X
asantana@uma.es

Antonio Cruces Rodríguez, Universidad de Málaga
ORCID: 0000-0001-6376-9608
antonio.cruces@uma.es

Nuria Rodríguez-Ortega, Universidad de Málaga
ORCID: 0000-0003-3460-397X
nro@uma.es

El objetivo central de este proyecto consiste en crear mecanismos a través de los cuales los estudiantes (y la ciudadanía en general) puedan contribuir de manera efectiva y directa al cuidado y preservación del patrimonio cultural próximo. Para alcanzar este fin, se ha partido de la selección de bienes patrimoniales en riesgo para los que, en primer lugar, los estudiantes han elaborado una ficha básica de diagnóstico, en la que se incluyen los daños ocasionados sobre el bien. En una segunda fase se han propuesto acciones de conservación, consolidación y de mejora para paliar esos daños. Para recoger la información recopilada, se ha creado la plataforma tecnológica Patrimonio Herido, accesible online, en la que se permite también la participación ciudadana.

Además de la plataforma mencionada, como resultados de aprendizaje destacan el acercamiento directo al patrimonio por parte de los estudiantes, una mayor sensibilización hacia las problemáticas actuales que afectan a los bienes patrimoniales y el desarrollo por parte del alumnado de herramientas para la resolución de problemas reales que se encontrará en su actividad profesional. Asimismo, la plataforma Patrimonio Herido se convierte en un foro de discusión entre los diferentes agentes relacionados con el patrimonio, con un destacado papel de la ciudadanía, redundando en una mayor visibilidad de su estado de conservación, y poniendo de manifiesto el valor del patrimonio como elemento identitario de la sociedad.

**Palabras clave:** Patrimonio Herido, peligro, protección, Proyecto de Innovación Educativa (PIE), ciudadanía.

**Abstract:**
The aim of this contribution is to present a line of teaching research based on the hybridisation between prototyping methodologies, PBL (Problem Based Learning) strategies and technological implementations, included in the educational innovation actions developed by TransUMA (Laboratory of Transdisciplinary Competences of the University of Malaga). This publication focuses specifically on the work-project Wounded Heritage, in which the students of the degree in History of Art and the degree in Tourism of the University of Malaga have participated.

The main objective of this project is to create mechanisms through which students (and citizens in general) can contribute effectively and directly to the care and preservation of the local cultural heritage. In order to achieve this goal, the starting point was the selection of heritage

assets at risk. Firstly, the students drew up a basic diagnostic form, which included the damage caused to the asset. In a second phase, conservation, consolidation and improvement actions were proposed to mitigate the damage. In order to collect the information gathered, the technological platform Wounded Heritage has been created. This platform is accessible online, which also allows citizen participation.

In addition to the aforementioned platform, the learning results include the direct approach to heritage by students, a greater awareness of the current problems affecting heritage assets and the development by students of tools for solving real problems that they will face in their professional activity. Likewise, the Wounded Heritage platform becomes a forum for discussion between the different agents related to heritage, with a prominent role for citizens, resulting in greater visibility of its state of conservation, and highlighting the value of heritage as an identifying element of society.

**Keywords:** Wounded Heritage, Hazard, Protection, Education Innovation Project, Citizens.

## 1. INTRODUCCIÓN

El objetivo principal de esta contribución es presentar una línea de indagación docente basada en la hibridación entre metodologías de prototipado, estrategias ABP (Aprendizaje Basado en Problemas) (Restepo Gómez, 2005; Gutiérrez Carrillo y Romero Sánchez, 2019) e implementaciones tecnológicas que forma parte de las acciones de innovación educativa llevadas a cabo por TransUMA (Laboratorio de Competencias Transdisciplinares de la Universidad de Málaga). TransUMA es un ecosistema de aprendizaje orientado a facilitar una formación transdisciplinar, entendida como integración de saberes heterogéneos donde se combinan acción, pensamiento divergente y discurso crítico, trascendiendo los límites de las disciplinas involucradas y proporcionando un enfoque holístico[1].

En concreto, esta publicación centra su atención en un estudio de caso específico: el trabajo-proyecto *Patrimonio Herido*, en el que ha participado el alumnado del grado en Historia del Arte y del grado en Turismo de la Universidad de Málaga. El objetivo central de este proyecto consiste en crear mecanismos a través de los cuales los estudiantes (y la ciudadanía en general) puedan contribuir de manera efectiva y directa al cuidado y preservación del patrimonio cultural próximo.

Gran parte de este patrimonio se encuentra en un estado de degradación importante debido a diferentes factores, tanto de origen natural como antrópico (Camuffo, 1998). En el primero de los casos, juegan un importante papel los factores ambientales, mientras que en el segundo, las causas más destacadas son los actos vandálicos, las restauraciones inapropiadas o la falta de planes urbanísticos que tengan en cuenta la protección del

1. TransUMA es resultado del proyecto de innovación educativa de la Universidad de Málaga PIE17-123 que continúa en el marco del PIE19-178 *Imaginando futuros posibles*, para el que sha sido concedida su continuidad con el reconocimiento de Grupo Permanente de Innovación Educativa y el proyecto Estrategias STEAM-H aplicadas al ámbito patrimonial: el proyecto Patrimonio Herido (PIE22-165). Véase en http://transuma.iarthislab.eu/ [consultado 21 marzo 2022].

patrimonio. A esto hay que sumar, en muchas ocasiones, la falta de un mantenimiento adecuado de estas estructuras (AA. VV., 2019), lo que conlleva el daño material de los bienes, así como la pérdida de sus valores patrimoniales, comprometiendo, incluso, la seguridad de las personas (Gutiérrez-Carrillo *et al.*, 2020).

Considerando esta realidad, la elaboración de estrategias de conservación preventiva y de mantenimiento resulta esencial para garantizar una protección y conservación más sostenibles (Garrido y Gutiérrez-Carrillo, 2015). La *Carta de Cracovia* (UNESCO, 2020) supuso una importante actualización de los criterios de conservación y restauración, destacando la relación de los bienes patrimoniales con su entorno y el paisaje. En ella se promueve un modelo de conservación basado en la mínima intervención y evitando la reconstrucción, incidiendo en la compatibilidad con los materiales y técnicas empleados en la construcción del bien, así como en la reversibilidad de las acciones. Asimismo, se recomienda el uso de las nuevas tecnologías en la gestión pública del patrimonio, dando prioridad al uso de bases de datos, sistemas computacionales y modelos virtuales de difusión (Gutiérrez-Carrillo *et al.*, 2020).

Estas recomendaciones han fomentado en los últimos años la conformación de equipos transdisciplinares para abordar de forma holística el estudio y conservación del patrimonio en toda su complejidad (Marcos Cobaleda y Gutiérrez-Carrillo, 2021), integrando a los diferentes agentes que intervienen en su conservación. De esta realidad surge el proyecto *Patrimonio Herido* como propuesta de innovación educativa orientada a fomentar el conocimiento del patrimonio cercano y su conservación, poniendo de manifiesto el valor del patrimonio como elemento identitario de la sociedad (Cuenca-López, Molina-Puche y Martín-Cáceres, 2018), a partir del desarrollo de una plataforma digital accesible en abierto a través de la web, cuyos objetivos, metodología y resultados obtenidos presentamos a continuación.

La creación de esta plataforma digital ha supuesto un gran reto, puesto que trasciende los límites de las aulas, estando a

disposición de la ciudadanía, gracias a su publicación en abierto, de tal forma que cualquier ciudadano puede ser partícipe del proyecto y enviar a través de su propio teléfono móvil sus denuncias sobre los daños causados sobre el patrimonio cercano, acompañadas de fotografías. El principal beneficio de su implementación es que, al ser una plataforma de amplio alcance, el volumen de datos que pueden incluirse en ella no está limitado a lo que consigan introducir en la base de datos los miembros del proyecto, sino que esta información puede ampliarse sustancialmente, ayudando así a una mayor visibilidad del proyecto y a convertirlo en una verdadera herramienta para la conservación del patrimonio, fin con el que fue concebido. No obstante, el mayor de los retos al que nos enfrentamos es lograr que esta participación ciudadana sea realmente efectiva, por lo que una de las líneas estratégicas del proyecto propuestas para el curso 2022-2023 se basa precisamente en las labores de difusión de esta plataforma digital[2].

## 2. OBJETIVOS Y METODOLOGÍA

Teniendo en cuenta la variedad de agentes que intervienen en la conservación del patrimonio y su carácter transdisciplinar, durante el curso académico 2020-2021, se escogió una metodología de actividad formativa caracterizada por la transdisciplinariedad para la realización de un trabajo-proyecto colaborativo entre el alumnado de varias asignaturas de distintos grados de la Universidad de Málaga. Por un lado, participaron

2. En este sentido, para lograr una mayor involucración de la ciudadanía en el proyecto se están llevando a cabo diferentes acciones de difusión al gran público, como la participación en la Noche Europea de los Investigadores que tuvo lugar en otoño de 2022, con la actividad "En busca del Patrimonio Herido". Si bien esta filosofía de participación ciudadana se encuentra en la idea fundacional del proyecto, todavía no ha podido implementarse completamente como parte del proceso de enseñanza-aprendizaje. Esta dimensión será, sin embargo, central en el marco del nuevo GpIE iniciado en el curso 2022-2023.

estudiantes de la asignatura Gestión del Patrimonio Cultural del grado en Turismo (primer curso), con un total de 240 estudiantes. Por otro lado, se implicaron en el proyecto estudiantes del grado en Historia del Arte matriculados en las asignaturas Arte Islámico, Andalusí y Mudéjar (primer curso), con un total de 101 estudiantes; Historia y Conceptos Fundamentales del Patrimonio Cultural (segundo curso), con un total de 84 estudiantes; y de la asignatura Arquitectura Contemporánea (cuarto curso), con un total de 32 estudiantes.

### 2.1. OBJETIVOS ESPECÍFICOS

En esta misma línea, para la realización de la actividad se plantearon los siguientes objetivos específicos:

1. Propiciar la transdisciplinariedad entre el grado en Historia del Arte y el grado en Turismo de la Universidad de Málaga.
2. Involucrar al estudiantado en el contacto directo con el patrimonio cercano y fomentar su conservación.
3. Promover acciones que permitan sanar el patrimonio cultural herido.
4. Concienciar de la importancia del turismo cultural como factor de preservación del patrimonio.
5. Familiarizar al alumnado con herramientas digitales para la difusión, protección y conservación del patrimonio, y proveerles de la capacitación técnica necesaria a través de talleres de formación específicos.

### 2.2. FASES METODOLÓGICAS

En virtud de los objetivos anteriores, metodológicamente el trabajo-proyecto *Patrimonio Herido* se ha estructurado en tres fases. La primera fase ha consistido en la selección por parte del alumnado matriculado en las asignaturas Gestión del Patrimonio Cultural; Arte Islámico, Andalusí y Mudéjar y Arquitectura Contemporánea de un bien patrimonial (preferiblemente localizado en la provincia




Figura 1. Plantilla de ficha de diagnóstico facilitada a los estudiantes para su cumplimentación. En azul aparecen indicados los diferentes apartados de la ficha; en gris aparece la información que los estudiantes deben completar; en rojo se incluyen las indicaciones dadas por el profesorado de las diferentes asignaturas para la cumplimentación de la ficha (2021). Fuente: elaboración propia.

de Málaga, ya que se trata del territorio en el que se estaban cursando los estudios implicados) que, por alguna razón, se encontrara en riesgo de conservación o que hubiera sufrido algún deterioro o destrucción. Una vez elegido el bien, los estudiantes elaboraron una ficha básica de diagnóstico, en la que se incluyeron los datos para su identificación, su cronología, su ubicación exacta, el tipo de patrimonio en el que se inscribe y una breve descripción histórico-artística. Junto a ello, en la ficha de diagnóstico, los estudiantes debían señalar los daños sufridos por el bien seleccionado. En este caso, debían incluir la ubicación concreta del daño, su descripción, la fecha en la que se produjo (siempre que se dispusiera de esta información) y las causas por las que el daño se había producido, conteniendo además un apartado de indicaciones en el que debían incluir una reflexión personal sobre el estado del bien patrimonial escogido. Asimismo, se solicitó al alumnado la inclusión de una bibliografía y documentación de referencia para familiarizarse con los diferentes sistemas de citación; una serie de imágenes que permitiera la identificación del bien y la valoración de los daños sufridos; un mapa básico de localización; y una reflexión sobre las posibles medidas que posibilitarían la sanación del bien (Figura 1).

Una vez confeccionadas las fichas de diagnóstico, el profesorado de las asignaturas procedió a una selección de los bienes más representativos, atendiendo a varios criterios:

1. Calidad de los trabajos.
2. Viabilidad de la realización de un proyecto de conservación sobre el bien.
3. Necesidades sociales en el entorno donde se emplaza el bien y las posibilidades de regeneración que su recuperación ofrece.
4. Inclusión de ejemplos de otros lugares de Andalucía.
5. Grado del daño.
6. Viabilidad de acciones de mejora.
7. Amplitud tipológica.

Así, el número total de fichas de diagnóstico seleccionadas fue de sesenta y siete.

La segunda fase del proyecto fue realizada por los estudiantes matriculados en la asignatura Historia y Conceptos Fundamentales del Patrimonio Cultural del grado en Historia del Arte, quienes trabajaron sobre las fichas de diagnóstico seleccionadas por el equipo docente, recopilando los datos esenciales del inmueble, entre ellos los relativos a su catalogación, así como realizando un análisis de la situación en la que se encuentra el bien, en la que se incluye una propuesta de uso, ideas para frenar el daño mediante unos "primeros auxilios", y una intervención más específica y definitiva para recuperar el elemento. El alumnado tuvo que acompañar esta documentación con imágenes históricas, actuales y de la futura propuesta, a la vez que ampliar el número de referencias y bibliografía utilizada, haciendo una búsqueda de documentos más específicos de su disciplina.

Junto al alumnado de la asignatura Historia y Conceptos Fundamentales del Patrimonio Cultural, los estudiantes de Arquitectura Contemporánea, además de haber participado en la primera fase del trabajo-proyecto con la selección del patrimonio en peligro, realizaron igualmente la ampliación de la ficha descrita en el párrafo anterior. Esta decisión vino planificada debido a diversas cuestiones. Por un lado, a que dicha asignatura se imparte en el cuarto curso, año del grado en el cual la formación del alumnado es más amplia, al haber cursado ya varias asignaturas relativas al patrimonio cultural y su conservación. Por otra parte, durante el curso 2018-2019, estos estudiantes ya habían participado en una edición anterior del proyecto al cursar la asignatura Historia y Conceptos Fundamentales del Patrimonio Cultural, por lo que se encontraban familiarizados con el sistema. Junto a todo ello, ante la posibilidad de que no fuesen seleccionados bienes inmuebles de este periodo cronológico por parte del alumnado del grado en Turismo, se permitió la elección

directa a los estudiantes de Arquitectura Contemporánea, puesto que la sociedad no suele tener un vínculo tan directo con estos bienes, principalmente debido a su construcción más reciente y al uso de un lenguaje no tan aceptado por el público general, lo que conlleva que sean más frágiles en el ámbito expuesto en este proyecto. Su inclusión supone, así, la ampliación del abanico de bienes patrimoniales en peligro.

Como se detallará en el apartado 2.3 de esta contribución, para albergar los trabajos llevados a cabo por los estudiantes, en el marco de este proyecto y en la línea de las propuestas del Plan Nacional de Educación y Patrimonio coordinado por el Instituto de Patrimonio Cultural en España (IPCE) (Ibáñez-Etxeberria, Fontal Merillas y Rivero Gracia, 2018), en una tercera fase se ha creado la infraestructura tecnológica *Patrimonio Herido*[3], desarrollada *ad hoc*, que por su especificidad, se describirá a continuación en un apartado independiente. El uso de esta plataforma ha permitido a los estudiantes adquirir competencias muy diversas, entre las que se incluyen las propiamente tecnológicas.

Cabe destacar en el apartado metodológico la realización el pasado 28 de mayo de 2021 de un seminario bajo el título Ética del cuidado y patrimonio cultural. Estrategias y participación ciudadana. En este marco, se reunió al alumnado adscrito al proyecto de los grados en Turismo e Historia del Arte con profesionales en la materia y la ciudadanía para dar a conocer el proyecto. En él participaron el Presidente de ICOMOS España, Jordi Tresserras; el director de Participación e Innovación Social de la Universidad de Granada y director de Medialab UGR, Esteban Romero Frías; la arqueóloga Carmen Íñiguez; la directora de la Cátedra de Patrimonio en la Universidad de Granada, María Luisa Bellido; el comunicador cultural Miguel Cajigal (conocido como El Barroquista); y el delegado de la asociación Hispania Nostra para Andalucía Oriental, Luis Machuca Santa-Cruz. Todos ellos, desde

3. https://patrimonioherido.iarthislab.eu [21 de marzo de 2022].




Figura 2. Informe DAFO (2021). Fuente: elaboración propia.

diferentes perspectivas, ofrecieron a los asistentes[4] un panorama del estado del patrimonio cultural en España y diferentes acciones que se están llevando a cabo desde las diferentes instituciones para salvaguardarlo. Asimismo, el profesorado implicado en las asignaturas circunscritas al proyecto *Patrimonio Herido* realizó una presentación de esta iniciativa, como una estructura donde la creación de sinergias entre Universidad, alumnado y sociedad civil se presenta como pilar fundamental sobre el que se construye este proyecto. El objetivo principal de este seminario consistió en presentar públicamente el proyecto *Patrimonio Herido*, posibilitando que éste trascendiera los límites del aula y adquiriese así una función social que permita la retroalimentación

4. Debido a las restricciones de aforo en el auditorio del Museo Thyssen por la pandemia de COVID-19, el seminario fue retransmitido en *streaming*, con una enorme aceptación, puesto que en muchos momentos se alcanzó el número máximo de usuarios conectados permitido por el servidor.

de todos los agentes implicados en la conservación patrimonial, teniendo en cuenta que la participación ciudadana ha sido uno de los elementos fundamentales a tener en cuenta a la hora de configurar el proyecto.

A partir del análisis y evaluación del proyecto, el grupo de docentes planteó, siguiendo la línea de otras investigaciones (García González y Ríos Moyano, 2021), acciones de mejora de cara a su implementación durante el curso 2021-2022 utilizando como herramienta la elaboración de un informe DAFO (Figura 2), en el que advertimos las debilidades y fortalezas del proyecto en el curso 2020-2021 y propusimos, tras el estudio de las posibles amenazas, las nuevas oportunidades para el curso 2021-2022.

### 2.3. LA PLATAFORMA PATRIMONIO HERIDO: UNA HERRAMIENTA CIUDADANA EN DEFENSA DEL PATRIMONIO CULTURAL

Una vez culminados los trabajos por parte del alumnado, en la tercera fase del proyecto, toda la información recopilada fue incorporada por parte de los estudiantes y el equipo docente a la plataforma *Patrimonio Herido* (https://patrimonioherido.iarthislab.eu), desarrollada por Antonio Cruces Rodríguez. Esta plataforma ha sido creada con la finalidad de hacer visibles los daños en el patrimonio por parte de cualquier ciudadano. Para facilitar la supervisión del profesorado sobre las contribuciones al trabajo-proyecto realizadas por el alumnado, y, al mismo tiempo, permitir una publicación rápida, de alcance global, con un método de trabajo cuya curva de aprendizaje fuese muy leve, sin dejar por ello de ofrecer todas las garantías de seguridad y privacidad, *Patrimonio Herido* (PH) se ha implementado mediante un mecanismo internacionalmente reconocido y utilizado con éxito en diversos proyectos del grupo de investigación iArtHis_Lab y su rama tecnológica iASD (iArtHis_Lab Software Development), cuyo despliegue ha dado lugar a una tesis doctoral (Cruces Rodríguez, 2021), y que ha servido como herramienta clave en dos más (Ortiz Tello, 2020; Romero Ferón, en curso).

Se trata de una combinación de temas y complementos –*plugins*– desarrollados sobre la aplicación de fuente abierta WordPress (WP), cimentada en un conjunto de código PHP y JavaScript, más MySQL/MariaDB como base de datos. Haciendo uso de su versatilidad, los estudiantes, previamente identificados y una vez aceptados los términos legales de privacidad y las restricciones derivadas del uso de imágenes, envían a la plataforma sus hallazgos y colaboraciones[5]. Tales envíos son inmediatamente notificados al profesorado destinatario, que, tras realizar la correspondiente evaluación, puede publicar lo recibido, o bien reclamar ulteriores revisiones a los remitentes.

El sistema, que puede consultarse en español, y en un futuro también en inglés e italiano, es tan adaptable y configurable que permite, en realidad, la participación de cualquier persona que desee aportar su ayuda en la difícil pero necesaria tarea de la defensa del patrimonio cultural, bajo la supervisión de investigadores y docentes pertenecientes durante el curso 2020-2021 al Departamento de Historia del Arte de la Universidad de Málaga. Adicionalmente, PH ofrece no sólo una interfaz de publicación electrónica con acceso directo desde Internet para cualquier internauta, sino como repositorio OAI2-PMH, estándar abierto que posibilita su incorporación directa a depósitos de información de alcance internacional tales como *Hispana*, dependiente del Ministerio de Cultura, y *Europeana*, auspiciado por la Unión Europea, ampliando así de forma muy notable el ámbito de difusión de sus contenidos (Figura 3).

Para la localización de los bienes seleccionados, en el caso concreto de la plataforma *Patrimonio Herido*, se ha utilizado

5. Para que los estudiantes tuvieran los conocimientos técnicos necesarios para poder darse de alta y subir la información a la plataforma, se desarrollaron varios seminarios formativos bajo el título "Formación para las herramientas del portal de Patrimonio Herido" impartidos por Antonio Cruces Rodríguez, tanto para el alumnado del grado en Turismo como para el del grado en Historia del Arte.

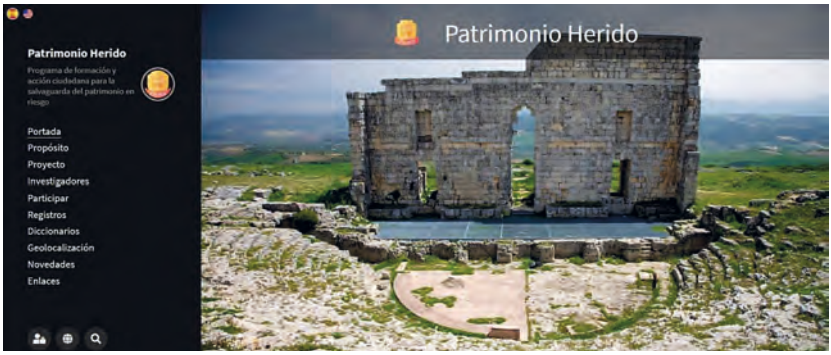



Figura 3. Vista de la portada de la página de inicio del portal *Patrimonio Herido* (2021). Fuente: ©patrimonioherido.iarthislab.eu.




Figura 4. Vista de la portada de la sección de Geolocalización del portal *Patrimonio Herido* (2021). Fuente: ©patrimonioherido.iarthislab.eu.

un sistema cartográfico *open source* (fundamentalmente OpenStreetMap) y un mecanismo de gestión Leaflet, basado en JavaScript. Gracias a ello, la plataforma puede mostrar mapas globales en los que se visualiza la localización de todos los bienes culturales incluidos en el sistema (Figura 4). Asimismo, gracias a la georreferenciación de los bienes, PH puede aportar operaciones SIG, como la generación de *heatmaps*, clusterización, operaciones geométricas, interpolación, etc.

Además de los registros con los bienes del patrimonio herido subidos por los estudiantes y revisados por el profesorado,

los incluidos directamente por los docentes en relación a bienes en los que han trabajado de manera directa, y de su geolocalización, la plataforma PH ofrece información complementaria sobre el propio proyecto y el equipo que lo conforma, así como un apartado de noticias bajo el epígrafe "Novedades", en el que se publican regularmente noticias de prensa y de diversas instituciones de carácter patrimonial como ICOMOS o Hispania Nostra relacionadas con el ámbito del proyecto, para hacer visible cualquier actuación, daño o peligro que afecte al patrimonio cultural. Asimismo, se ha creado un apartado de "Participación" destinado a la intervención ciudadana, en el que se explica cómo darse de alta en la plataforma y cómo enviar la información para que sea revisada por los miembros del proyecto. Este apartado cuenta con una herramienta de fotodenuncia en el que cualquier ciudadano puede enviar desde su móvil de forma sencilla y rápida una imagen de un bien en peligro que haya encontrado, con un breve comentario.

## 3. PRESENTACIÓN DE LOS RESULTADOS

Como muestra de los resultados obtenidos en el trabajo-proyecto, incluimos a continuación varios ejemplos de los bienes patrimoniales en riesgo en la provincia de Málaga que han sido seleccionados para el proyecto. En aras de alcanzar el segundo de los objetivos propuestos ("Involucrar al estudiantado en el contacto directo con el patrimonio cercano y fomentar su conservación"), se limitó el marco geográfico del trabajo-proyecto a la mencionada provincia, para que los estudiantes pudieran acceder físicamente a los bienes patrimoniales escogidos y pudieran así comprobar *in situ* los daños experimentados y su estado general de conservación. El curso 2018-2019 fue el primero en el que se decidió implicar al alumnado de ambos grados; considerándose ya desde entonces la necesidad de que el estudiantado visitase, de manera presencial,




Figura 5. Muralla islámica de Málaga, tramo de la Plazuela Virgen de las Penas (2021). Fuente: ©Google Maps.

el bien sobre el que iba a desarrollar su investigación (Santana Guzmán, 2020)[6].

En primer lugar, podemos mencionar el caso de la muralla islámica de la ciudad de Málaga. Ésta comenzó a erigirse en el siglo XI bajo el gobierno de la dinastía Ḥammūdī (Íñiguez Sánchez, 1995; Íñiguez Sánchez, Rambla Torralbo y Mayorga Mayorga, 2003), siendo fruto de diversas ampliaciones en época nazarí y tras la conquista de los Reyes Católicos (Ruiz Povedano, 2000). A pesar de las destrucciones llevadas a cabo en los siglos XVIII y XIX, algunos de los tramos de esta cerca se conservan todavía entre el caserío urbano del centro histórico de Málaga. Entre ellos, sus

6. Aunque el marco propuesto para el trabajo con los estudiantes fue la provincia de Málaga, debido a que en el comienzo de su elaboración la Universidad de Málaga aún se encontraba en un contexto de enseñanza telemática a causa de la situación de alerta sanitaria provocada por la pandemia de COVID-19, algunos de los trabajos fueron realizados sobre bienes patrimoniales en riesgo de otras provincias andaluzas, como Córdoba o Almería, lugar de residencia de algunos de los estudiantes. Actualmente, el marco territorial del proyecto *Patrimonio Herido* ha sido ampliado a la totalidad de la Comunidad Autónoma de Andalucía, participando asimismo en el proyecto estudiantes de las Universidades de Granada y Sevilla.

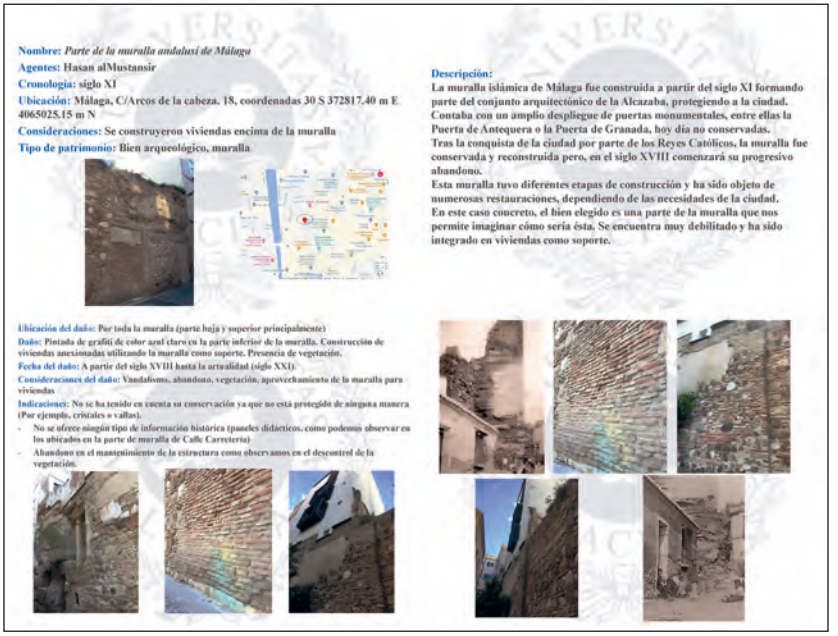



Figura 6. Detalle de ficha de diagnóstico de la muralla islámica de Málaga (2021). Autoras: Carmen Garrocho Portillo, Gloria Hidalgo Serrano y Laura Muñoz Aguilar.

lienzos visibles en la Plazuela Virgen de las Penas y en la calle Arcos de la Cabeza han sufrido varios daños, causados sobre todo por el vandalismo y por la presión urbanística (Figura 5).

En este contexto, los estudiantes que trabajaron sobre el bien patrimonial detectaron que la muralla había sido integrada en varias viviendas, pudiendo verse claramente cómo la cerca medieval actúa directamente como soporte de éstas (en la mayoría de los casos, conformadas por varios pisos en altura). Por otro lado, también incidieron en cómo había sido víctima de actos vandálicos, incluyendo la presencia de *graffiti* en numerosos puntos, hecho que afecta igualmente al propio entorno del bien patrimonial, que se caracteriza por el depósito de restos de basura y otros residuos, los cuales suponen una importante contaminación visual de la muralla.

A esto se suma la presencia de un naranjo que crece junto a la cerca y que entra en contacto con ella en diversos puntos, provocando nuevamente un importante impacto visual sobre el bien patrimonial, además del evidente deterioro del lienzo conservado.

Asimismo, el alumnado constató cómo el abandono tras la última restauración llevada a cabo en el año 2009 por la empresa *Qibla Restaura* (Quibla Restaura, 2011) ha provocado la aparición de vegetación en los muros. Por último, los estudiantes remarcaron cómo este tramo conservado de la muralla de Málaga carece de cualquier señalización que la identifique con su pasado islámico y remarque sus valores patrimoniales, al contrario de lo que sucede en otros tramos conservados, como el de la calle Carretería, nº 64, quedando sorprendidos por esta disparidad en el tratamiento del mismo bien patrimonial. Toda esta información fue recogida en las fichas de diagnóstico elaboradas por los diferentes grupos que trabajaron sobre la muralla, como puede apreciarse en el detalle de la ficha de diagnóstico (Figura 6).

Otro caso a reseñar es el del *Molino de la Bóveda*, en Torremolinos, cuyos orígenes se remontan al siglo XV, si bien, incluso, pudieran ser anteriores (Palop, 2006). Esta arquitectura vernácula ha sobrevivido gracias a su adaptación a distintas funciones en su devenir histórico, siendo los momentos más críticos el desplome de su actividad primaria durante los siglos XIX y XX, así como la radical transformación de la localidad en la segunda mitad de la última centuria. En esos momentos su uso se adaptó al monopolio del sector turístico, convirtiéndose en restaurante y sala de espectáculos. Actualmente, este edificio, que destaca por sus valores etnológicos, está completamente abandonado; su desuso se convierte en su principal verdugo ante una situación de peligro de desaparición inminente (Figura 7).

Entre los factores que producen el riesgo de pérdida de este inmueble cabe añadir: la entrada de personas ajenas a la propiedad, el crecimiento descontrolado de vegetación, y la pérdida o desprendimientos de parte de su material constructivo.

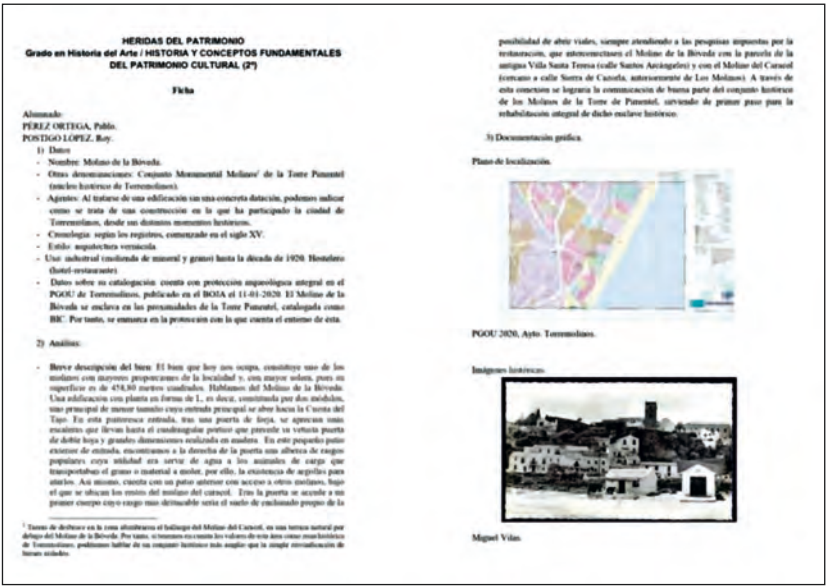


Figura 7. Antiguo *Molino de la Bóveda,* en la Cuesta del Tajo, Torremolinos (Málaga, 2016). Fuente: Miguel Vila.

Los estudiantes proponen su reconversión en espacio cultural multidisciplinar, enfocado principalmente al papel artístico y vinculado a la ciudadanía. Del mismo modo, se pretende ponerlo en valor como elemento destacado del entramado urbano recuperando su presencia en el paisaje (Figura 8).

Cabe destacar que esta investigación es fruto del interés personal de dos estudiantes que conocen a la perfección Torremolinos y, por lo tanto, se decidió que desarrollaran un trabajo *ex profeso* incluyendo un bien que no había sido elegido en la primera fase; esta opción se ha planteado en el marco del proyecto de manera excepcional al considerar la importancia de dar difusión a bienes en peligro con los que el alumnado mantuviese un vínculo directo.

La preocupación de los estudiantes implicados no sólo derivó en un buen trabajo de investigación, sino también en la difusión de este estudio a través de una actividad llevada a cabo en colaboración con el Ayuntamiento de Torremolinos. Así, el 18 de junio de 2021, en el Centro Cultural Pablo Ruiz Picasso, el alumnado, a través de sendas ponencias, presentó los dos bienes



**HERIDAS DEL PATRIMONIO**
**Grado en Historia del Arte / HISTORIA Y CONCEPTOS FUNDAMENTALES DEL PATRIMONIO CULTURAL (2º)**

**Ficha**

Alumnado:
PÉREZ ORTEGA, Pablo.
POSTIGO LÓPEZ, Roy.

1) Datos
- Nombre: Molino de la Bóveda.
- Otras denominaciones: Conjunto Monumental Molinos' de la Torre Pimentel (núcleo histórico de Torremolinos).
- Agentes: Al tratarse de una edificación tan una concreta datación, podemos indicar como se trata de una construcción en la que ha participado la ciudad de Torremolinos, desde sus distintos momentos históricos.
- Cronología: según los registros, comenzado en el siglo XV.
- Estilo: arquitectura vernácula
- Uso: industrial (molienda de manomil y grano) hasta la década de 1920. Hostelero (hotel-restaurante)
- Datos sobre su catalogación: cuenta con protección arqueológica integral en el PGOU de Torremolinos, publicado en el BOJA el 11-01-2020. El Molino de la Bóveda se enclava en las proximidades de la Torre Pimentel, catalogada como BIC. Por tanto, se entiende en la protección con la que cuenta el entorno de ésta.

2) Análisis:
- Breve descripción del bien: El bien que hoy nos ocupa, constituye uno de los molinos con mayores proporciones de la localidad y, con mayor solera, pues su superficie es de 458,80 metros cuadrados. Hablamos del Molino de la Bóveda. Una edificación con planta en forma de L, es decir, constituida por dos módulos, uno principal de menor tamaño cuya entrada principal se abre hacia la Cuesta del Tajo. En esta pintoresca entrada, tras una puerta de forja, se aprecian unas escaleras que llevan hasta el cuadrangular portico que precede la vetusta puerta de doble hoja y grandes dimensiones realizada en madera. En este pequeño patio existen dos estancias, encontramos a la derecha de la puerta una alberca de origen popular, cuya utilidad era servir de agua a los animales de carga que transportaban el grano o material a moler, por ello, la existencia de agrolitos para atarlos. Asi mismo, cuenta con un patio anterior con acceso a otros enclaves, bajo el que se ubican los restos del molino del cañaveral. Tras la puerta se accede a un primer cuerpo cuyo rasgo mas destacable seria el suelo de enclinado propio de la

[1] Tuvo de destacar en la zona alrededores el balneario del Molino del Carasol, en una terraza natural por debajo del Molino de la Bóveda. Por tanto, si tuviéramos en cuenta los valores de este área como enclave histórico de Torremolinos, podríamos hablar de un conjunto histórico más amplio que la simple enumeración de bienes aislados.

posibilidad de abrir viales, siempre atendiendo a las pequeñas superficies por la restauración, que interconectasen el Molino de la Bóveda con la parcela de la antigua Villa Santa Teresa (calle Santos Arcángeles) y con el Molino del Carasol (cercano a calle Sierra de Cazorla, anteriormente de Los Molinos). A través de esta conexión se lograría la comunicación de buena parte del conjunto histórico de los Molinos de la Torre de Pimentel, sirviendo de primer paso para la rehabilitación integral de dicho enclave histórico.

3) Documentación gráfica.

Plano de localización.

PGOU 2020, Ayto. Torremolinos.

Imágenes históricas.

Miguel Vélez.

Figura 8. Detalle de ficha de diagnóstico del antiguo *Molino de la Bóveda* de Torremolinos (Málaga, 2021). Autores: Pablo Pérez Ortega y Roy Postigo López.

de la citada localidad que ese año habían sido estudiados: el *Molino de la Bóveda y la Cuesta del Tajo* (Pérez Ortega y Postigo López, 2021) y el *Conjunto los Manantiales* (Antón Valencia *et al.*, 2021). Este evento ha ampliado aún más el sector de público al que el proyecto quiere involucrar, acercando a la ciudadanía de cada territorio el patrimonio herido existente en él. Por ello, la presentación del acto estuvo precedida por la ponencia titulada *La web Patrimonio Herido y la concienciación ciudadana*, que fue impartida por el docente Antonio Jesús Santana Guzmán (coordinador de las asignaturas implicadas en ambas investigaciones). Esta actividad servirá de modelo a exportar a otras localidades en un futuro.

Otra de las investigaciones ha versado sobre el antiguo *Cine Lope de Vega*, edificio de 1962, obra del arquitecto Andrés Escassi Corbacho, que tuvo diversos usos hasta su clausura en la década de los 80, entre ellos como *Discoteca Bobby Logan*



Figura 9. Antiguo *Cine Lope de Vega,* en la Avenida Juan Sebastián Elcano (Málaga, 2021). Fuente: Cayetano Villodres.

(Asociación de vecinos y vecinas de El Palo, 2016). Está emplazado en un entorno donde mayoritariamente se localizan edificios residenciales de principios y mediados del siglo XX, entre los que destaca por su volumetría. Actualmente, no cuenta con ningún tipo de protección, sin embargo, las posibilidades que ofrece para incluir usos carentes en este sector urbano hacen que pueda ser recuperado (Figura 9).

El equipo de estudiantes encargado de este trabajo destacó su vinculación con la arquitectura del relax, así como la situación de peligro en la que se encuentra, propiciada principalmente por su abandono. Pero tal y como se propone en su ficha, podría reconvertirse en centro cultural, recuperándose, en la medida de lo posible, su aspecto original, tanto en el exterior como en el interior (incluyéndose imágenes virtuales con el aspecto que podrían tener), a la par que se solucionan



Figura 10. Detalle de ficha de diagnóstico del antiguo *Cine Lope de Vega* (Málaga, 2021). Autores: Rafael Córdoba Prats, Yaiza María Ortiz Mota, María Sanzo González, Noelia Vigo Martínez y Cayetano Villodres Fernández.

los probables problemas constructivos derivados de los años de abandono. También se incluyen en estos datos las propuestas de recuperación que ha tenido el inmueble por parte de las administraciones: por un lado, la Junta de Andalucía y, por otro, el Ayuntamiento. Si bien ninguna de ellas fue llevada a término, con su inclusión se exponen las posibilidades de esta arquitectura, así como el interés mostrado por parte de estas instituciones públicas (Figura 10).

Por último, de época romana, destacamos el caso de la *Villa romana de los Molinillos* en Benalmádena (Málaga) que, a pesar del buen estado de conservación en el que se encontraba durante su excavación, dando lugar al conocimiento de la *pars urbana* (concretamente lo que se ha identificado como el espacio termal) y de la *pars fructuaria* de la villa, en la actualidad adolece de medidas de conservación y adecuación que permitan la comprensión y el disfrute por parte de la sociedad siendo objeto de actos vandálicos, tales como pintadas en las estructuras (Figura 11).



Figura 11. Vista del estado actual del yacimiento *Villa romana de Los Molinillos*, Benalmádena (Málaga, 2021). Fuente: ©Google Maps.

Este bien fue trabajado por dos estudiantes del municipio que conocían la presencia de un solar abandonado en Benalmádena, junto al paseo marítimo, pero no el valor arqueológico del mismo. Concretamente se acercaron a la *pars fructuaria* de una villa datada entre el siglo I d. C y el siglo V d. C. dedicada, en una primera fase, a la producción de aceite, y posteriormente a la industria de salazón. La misma fue hallada en 2002 mientras se cimentaba un edificio que se pretendía de uso turístico (Pineda de las Infantas Beato, 2003). En primer lugar se centraron en contextualizar el bien, conocer su adscripción tipológica, aprender por qué se levantan villas en la costa malacitana en época romana y cómo se configuran (Rodríguez Oliva y Beltrán Fortes, 2008). Tras ello dieron un salto al presente para comprender el proceso de tutela administrativa y la complejidad de la compra de un espacio urbano, de propiedad privada que quiere adecuarse por parte del Ayuntamiento para ser devuelto a la ciudadanía. Así, el alumnado se centró en el proceso de compra del terreno y las problemáticas derivadas de ello (Figura 12).

Los daños advertidos por los estudiantes son principalmente el estado de ruina en el que se encontraba, la falta de




Figura 12. Detalle de ficha de diagnóstico del yacimiento *Villa romana de Los Molinillos*, Benalmádena (Málaga, 2021). Autores: Alejandro Castellanos Álvarez e Inmaculada González García.

protección física del espacio que permitía el acceso a la zona y la realización de *graffiti* sobre las estructuras, la contaminación visual o la nula señalización del bien, a pesar de encontrarse en una zona muy transitada y que desgraciadamente se utiliza para verter desechos por parte de los viandantes.

## 5. RESULTADOS DE APRENDIZAJE

Entre los resultados de aprendizaje obtenidos con la realización del trabajo-proyecto *Patrimonio Herido*, hemos observado que este tipo de actividades formativas contribuyen a familiarizar al alumnado, en especial al de primer curso, con las problemáticas actuales relativas al patrimonio cultural, haciéndolo a su vez más próximo a su realidad cotidiana (Marcos Cobaleda y Calderón Roca, 2018). Por su carácter transversal, así como por partir de la metodología ABP, este trabajo-proyecto ha acercado además al

estudiante a la resolución de problemas reales que se encontrará en su actividad profesional, fomentando la originalidad en los planteamientos resolutivos y en el proceso de la investigación (Cambil Hernández y Romero Sánchez, 2017), que le serán de gran utilidad para afrontar desafíos posteriores en su formación, como el TFG, el TFM o una Tesis Doctoral.

Asimismo, este trabajo-proyecto ha contribuido al desarrollo de una metodología de trabajo basada en fichas de diagnóstico sobre bienes patrimoniales, que le será de gran utilidad para su futura práctica profesional. Igualmente, al incorporar en ellas la descripción histórico-artística del bien, ésta le ayuda, por un lado, a manejar la terminología específica de esta disciplina y a su mejor asimilación, y, por otro, a estructurar correctamente y de forma sintética los comentarios de los bienes culturales, tarea a la que tanto el alumnado de Turismo como el de Historia del Arte deberán enfrentarse a lo largo de su formación académica y en el desarrollo posterior de su actividad profesional.

El hecho de permitir que sea el propio alumnado quien seleccione los bienes patrimoniales sobre los que trabajar ha redundado en una mayor motivación, basada en la autonomía a la hora de realizar esta selección, así como en la realización de sus propios materiales didácticos gracias al uso de las TICs. Esto puede verse, por ejemplo, en el caso concreto de la elaboración de los mapas de localización de los bienes patrimoniales seleccionados como complemento a las fichas de diagnóstico, para los que los estudiantes han utilizado las herramientas de Google, en concreto Google Earth y Google Maps. El uso de estos programas ha supuesto una primera aproximación del alumnado al uso de herramientas de carácter geográfico, que permiten posicionar un determinado objeto patrimonial en el espacio y tener en cuenta herramientas clave en el mismo como son la inserción de la escala y el Norte.

En este contexto, en la asignatura Historia y Conceptos Fundamentales del Patrimonio Cultural, se fue un paso más allá, aproximando al alumnado al uso de los Sistemas de Información

Geográfica (SIG). Para ello, se llevó a cabo el *I seminario de Patrimonio Cultural. El patrimonio arquitectónico: la importancia de su estudio y puesta en valor*, en el que tuvo lugar la actividad impartida por el consultor territorial Anton Ozomek Fernández, bajo el título "Territorio y Patrimonio: GIS online". A través de la misma, el alumnado se familiarizó con portales como: Google Maps, la Oficina Virtual del Catastro, la Fototeca Digital del Instituto Geográfico Nacional, la Guía Digital del Patrimonio Cultural de Andalucía, y el portal de la Gerencia Municipal de Urbanismo, Obras e Infraestructuras de Málaga. El conocimiento del manejo de estas herramientas, así como de estas fuentes de bases de datos relacionadas con el patrimonio cultural, ha supuesto un aumento del interés del alumnado en el uso de herramientas digitales aplicadas al estudio patrimonial, adquiriendo competencias ampliamente demandadas por el actual mercado laboral, donde las Humanidades Digitales han adquirido en los últimos años un gran protagonismo.

## 6. CONSIDERACIONES FINALES Y PROPUESTAS DE FUTURO

La propuesta de innovación educativa basada en el trabajo-proyecto *Patrimonio Herido* que aquí presentamos tiene un carácter abierto en el sentido más amplio. En primer lugar, no está limitada a un marco temporal, puesto que el hecho de haber implementado la plataforma PH permite que el estudiante pueda seguir colaborando con el proyecto una vez culminada la asignatura implicada, destacando así su vocación de futuro.

En segundo lugar, este proyecto trasciende el marco del aula, implicando de forma directa a la ciudadanía a través de instrumentos como la fotodenuncia. Esta herramienta es fundamental para conseguir una implicación ciudadana efectiva y fomentar su acercamiento al patrimonio cercano y su interés por su cuidado y protección, lo que redundará en una mejor conservación de éste (Marcos Cobaleda y Calderón Roca, 2018).

En tercer lugar, la propuesta es abierta dada la amplia variedad de patrimonio cultural que es susceptible de ser incluido en la plataforma, destacando especialmente el que se encuentra en riesgo o amenazado, convirtiéndose así en un instrumento de denuncia de esta situación.

En cuarto lugar, la publicación en *Open Access* de la plataforma PH permite que cualquier persona con conexión a Internet pueda visualizarla y colaborar con el proyecto. Un análisis de los datos de visitas a la misma durante los últimos nueve meses (entre el 24 de junio de 2021 y el 24 de marzo de 2022) posibilita medir el alcance del proyecto en estos primeros meses de publicación. Las cifras de visitas ascienden a 5.482, con un promedio de 22,56 visitas por día, destacando una mayor afluencia a partir del mes de febrero de 2022, como puede verse en el siguiente gráfico (Figura 13).

Esta mayor afluencia se debe, sobre todo, a la difusión de la plataforma que se está llevando a cabo en el marco de diferentes congresos nacionales e internacionales, publicaciones científicas, organización de seminarios y encuentros[7], páginas webs (principalmente de TransUMA y de iArtHis_Lab), redes sociales y la difusión entre el alumnado de los grados implicados en el proyecto.

Por último, aspiramos a que la plataforma *Patrimonio Herido* se convierta en un foro de discusión entre los diferentes agentes relacionados con el patrimonio, donde tenga voz la

7. Durante el mes de mayo de 2022 se llevó a cabo un encuentro en Antequera en el que participaron los estudiantes de las Universidades de Málaga, Granada y Sevilla implicados en el proyecto. Aunque inicialmente el proyecto de *Patrimonio Herido* estaba destinado a los estudiantes de la Universidad de Málaga, durante el curso 2021-2022 se han incorporado estudiantes de los grados en Turismo y en Historia del Arte de la Universidad de Granada, y del grado en Historia del Arte de la Universidad de Sevilla. Para el curso 2022-2023 se incorporarán además estudiantes del grado en Edificación de la Universidad de Granada, que contribuirán a remarcar la transdisciplinariedad del proyecto.

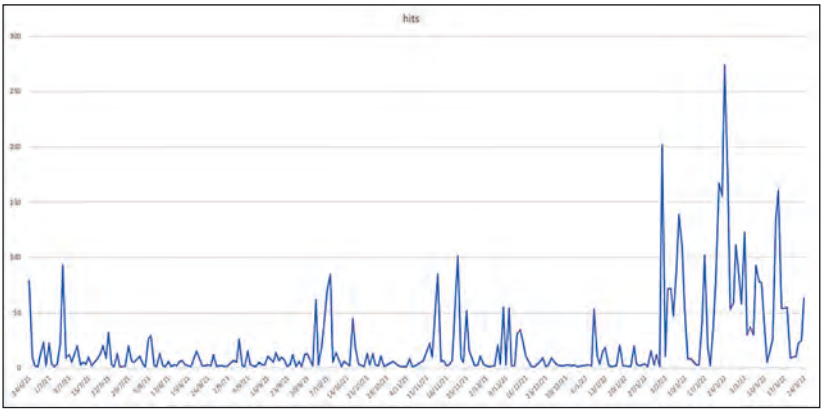



Figura 13. Visitas realizadas a la plataforma PH entre el 24 de junio de 2021 y el 24 de marzo de 2022. Autores: elaboración propia.

ciudadanía, y en el que colaboren instituciones culturales, administraciones públicas, asociaciones y otros organismos relacionados con el patrimonio cultural. El alcance del proyecto, con un total de 457 estudiantes implicados en el curso académico 2020-2021, el aumento durante el curso 2021-2022 con los estudiantes de la Universidad de Granada y de la Universidad de Sevilla, y el previsto para el curso 2022-2023 con la incorporación de los estudiantes del grado en Edificación de la Universidad de Granada; la presencia de tres universidades andaluzas; las evaluaciones que realiza el equipo docente al finalizar cada curso para promover las acciones de mejora; y el crecimiento progresivo del número de visitas a la plataforma PH son indicativos del creciente interés por el patrimonio cultural, por lo que consideramos que esta propuesta de innovación educativa de amplio alcance puede convertirse en herramienta que contribuya a garantizar su conservación.

## REFERENCIAS BIBLIOGRÁFICAS

AA. VV. (2019). *HeritageCARE. Guide Méthodologique de Conservation Préventive*. Sevilla: UMinho/DRCN/CCG/USAL/IAPH/FSMLR/UCA/ULIM.

Antón Valencia, María; Gaitán Nekkebroek, Raquel; Gallardo Martín, Laura; Granados Hita, Alberto y Villaverde Calvo, Jesús Miguel (2021/06/18). "Conjunto los Manantiales" [ponencia]. *Investigación [Patrimonio Herido en Torremolinos]*. Torremolinos: Ayuntamiento y Departamento de Historia del Arte de la Universidad de Málaga.

Asociación de vecinos y vecinas de El Palo (01/05/2016). *Los vecinos de Pedregalejo reiteran a Urbanismo que permute Bobby Logan*. Disponible en: https://www.elpalo.org/los-vecinos-de-pedregalejo-reiteran-a-urbanismo-que-permute-bobby-logan/ [18 de marzo de 2021].

Cambil Hernández, María Encarnación y Romero Sánchez, Guadalupe (2017). "Metodología por proyectos: un modelo innovador para la enseñanza y el aprendizaje del patrimonio cultural". En: María Encarnación Cambil Hernández, Antonio Tudela Sancho (coords.). *Educación y patrimonio cultural. Fundamentos, contextos y estrategias didácticas*. Madrid: Pirámide, pp. 61-80.

Camuffo, D. (1998). *Microclimate for Cultural Heritage*. San Diego: Elsevier.

Cruces Rodríguez, Antonio (2021). *ExpoFinder y Pathfinder. Infraestructura tecnológica para la generación de conocimiento en el ámbito de la Historia del Arte* [Tesis doctoral]. Universidad de Málaga: Málaga. Disponible en: https://riuma.uma.es/xmlui/bitstream/handle/10630/23224/TD_CRUCES_RODRIGUEZ_Antonio_Jesus.pdf?sequence=1 [1 de abril de 2022].

Cuenca-López, José María; Molina-Puche, Sebastián y Martín-Cáceres, Myriam J. (2018). Identidad, ciudadanía y patrimonio. Análisis comparativo de su tratamiento didáctico en Museos de Estados Unidos y España. *Arbor*, 194 (788): a447.

García González, Julia y Ríos Moyano, Sonia (2021). Historiando para diseñar. Lab_DisMuseos: un laboratorio para el conocimiento de la historia del diseño en tiempos de covid-19. *Artseduca*, 30: 111-130. Doi: https://doi.org/10.6035/artseduca.5809. Disponible en: https://www.e-revistes.uji.es/index.php/artseduca/article/view/5809/6651 [30 de marzo de 2022].

Garrido, J. y Gutiérrez-Carrillo, Mª Lourdes (2015). Andalusian cultural heritage and natural hazards prevention. En: G. Lollino, D. Giordan, C. Marunteanu, B. Christaras, I. Yoshinori y C. Margottini (eds.). *Engineering Geology for Society and Territory*. Cham: Springer, vol. 8, pp. 437-440.

Gutiérrez Carrillo, Mª Lourdes y Romero Sánchez, Guadalupe (2019). Arquitectura defensiva de Granada. Propuesta metodológica fundamentada en itinerarios didácticos y ABP para su conocimiento y conservación". En: Mª Paz Bermúdez Sánchez (coord.). *Actas del congreso: Evaluación de la calidad de la investigación y de la Educación Superior. XVI FECIES*. Granada: Asociación Española de Psicología Conductual (AEPC), p. 527.

Gutiérrez-Carrillo, Mª Lourdes; Bestué Cardiel, Isabel; Molero Melgarejo, Emilio y Marcos Cobaleda, María (2020). Pathologic and Risk Analysis of the Lojuela Castle (Granada-Spain): Methodology and Preventive Conservation for Medieval Earthen Fortifications. *Applied Sciences*, 10: 6491. Doi: https://doi.org/10.3390/app10186491. Disponible en: https://www.mdpi.com/2076-3417/10/18/6491 [31 de marzo de 2022].

Ibáñez-Etxeberria, Alex; Fontal Merillas, Olaia y Rivero Gracia, Pilar (2018). Educación patrimonial y TIC en España: marco normativo, variantes estructurantes y programas referentes. *Arbor*, 194 (788): a448. Doi: https://doi.org/10.3989/arbor.2018.788n2008. Disponible en: https://arbor.revistas.csic.es/index.php/arbor/article/view/2261/3184 [1 de abril de 2022].

Íñiguez Sánchez, Mª del Carmen (1995). La muralla islámica de Málaga: referencias textuales y constatación arqueológica. *Arqueología y Territorio Medieval*, 2: 129-138. Disponible en: http://www.ujaen.es/revista/arqytm/PDF/R2/R2_7_Iniguez.pdf [5 de marzo de 2021].

Íñiguez Sánchez, Mª del Carmen; Rambla Torralbo, José Antonio y Mayorga Mayorga, José Francisco (2003). La construcción de la Muralla Musulmana de Málaga, un hito en la historia de la ciudad. *Revista Mainake*, 25: 133-176.

Marcos Cobaleda, María y Calderón Roca, Belén (2018). Los trabajos-proyecto en Educación Infantil: propuesta metodológica para

la enseñanza del Patrimonio cultural andaluz". En: David Caldevilla Domínguez, Elena Alarcón Orozco, Virginia Alarcón Martínez (coords.). *Reformulando la docencia actual*. Barcelona: Gedisa, pp. 203-2018.

Marcos Cobaleda, María y Gutiérrez-Carrillo, Mª Lourdes (2021). Almoravid Works on Defensive Architecture in Southeast Al-Andalus: Analysis of Their Remains and Proposal for Preventive Conservation. *Sustainability*, 13 (24): 13597. Doi: https://doi.org/10.3390/su132413597. Disponible en: https://www.mdpi.com/2071-1050/13/24/13597 [31 de marzo de 2022].

Ortiz Tello, María (2021). *Picasso en el circuito expositivo: coleccionistas, préstamos y circulación artística* [Tesis doctoral]. Universidad de Málaga: Málaga. Disponible en: https://hdl.handle.net/10630/19864 [1 de abril de 2022].

Palop, J. J. (2006). *Los molinos de Torremolinos*. Torremolinos: Ayuntamiento.

Pérez Ortega, Pablo y Postigo López, Roy (2021/06/18). "Molino de la Bóveda y Cuesta del Tajo" [ponencia]. *Investigación [Patrimonio Herido en Torremolinos]*. Torremolinos: Ayuntamiento y Departamento de Historia del Arte de la Universidad de Málaga.

Pineda de las Infantas Beato, Gonzalo (2003). Los Molinillos, un complejo industrial de época romana en Benalmádena-Costa. *Cilniana: Revista de la Asociación Cilniana para la Defensa y Difusión del Patrimonio Cultural*, 26: 65-74.

Quibla Restaura (2011). *Informe de intervención de conservación y restauración sobre los restos emergentes de la muralla medieval de Málaga en Calle Arco de la Cabeza*.

Restepo Gómez, Bernardo (2005). Aprendizaje Basado en Problemas (ABP). Una innovación didáctica para la docencia universitaria. *Educación y educadores*, 8: 9-20.

Rodríguez Oliva, Pedro y Beltrán Fortes, José (2021). Arqueología de las *Villae* romanas de la costa malacitana. *Habis,* 39: 223-243. Doi: http://dx.doi.org/10.12795/Habis.2008.i39.17 [30 de marzo de 2021].

Romero Ferón, Bárbara (en curso). *Discusión crítica del concepto de arte español a través de sus exposiciones y catálogos (Europa y EE.UU., 1850-1939)* [Tesis doctoral]. Universidad de Western Ontario: Ontario (Canadá).

Ruiz Povedano, José María (2000). *Málaga, de musulmana a cristiana: la transformación de la ciudad a finales de la Edad Media*. Granada: Universidad de Granada.

Santana Guzmán, Antonio Jesús (2020). Experiencias sobre el acercamiento del patrimonio malagueño a la ciudadanía. *Universitas*, 33: 103-124. Doi: https://doi.org/10.17163/uni.n33.2020.05 Disponible en: https://dialnet.unirioja.es/descarga/articulo/7774711.pdf [30 de marzo de 2022].

UNESCO (2020). *Carta de Cracovia*. Disponible en: https://ipce.culturaydeporte.gob.es/dam/jcr:b3b6503d-cf75-4cb0-adaf-226740ebd654/2000-carta-cracovia.pdf [31 de marzo de 2022].

# II

# LAS VÍAS DEL LENGUAJE, DEL TEXTO Y DE LA LITERATURA

# *MAPPING CHIVALRY:* LA EVOLUCIÓN DE LOS LIBROS DE CABALLERÍAS CASTELLANOS EN EL TIEMPO Y EL ESPACIO[1]

Giulia Tomasi

## 1. *MAPPING CHIVALRY*: PREMISAS Y OBJETIVOS.

Los libros de caballerías castellanos de los siglos XVI y XVII tuvieron una notable incidencia en el contexto literario de su época y su influencia se trasladó a los siglos sucesivos a través de la reelaboración de sus formas y contenidos. Para dar relieve a tal propagación, estamos desarrollando un proyecto que se centra en el estudio de la literatura caballeresca hispánica desde distintas perspectivas, siendo apoyados por una financiación trienal por parte del Ministero dell'Università e la Ricerca en el ámbito de los proyectos de interés nacional italiano. *Mapping Chivalry: Spanish Romances of Chivalry from Renaissance to 21st Century, a Digital Approach* tiene por objeto la exploración de los ricos contenidos de los libros de caballerías castellanos y de su difusión fuera de España,

1 Este trabajo se inserta en el macroproyecto italiano PRIN 2017 Mapping Chivalry. Spanish Romances of Chivalry from Renaissance to 21st Century: a Digital Approach (2017JA5XAR), concedido por el MIUR. La investigadora principal es Anna Bognolo (Università di Verona). Más detalles acerca del proyecto se encuentran en Tomasi (2022).

Giulia Tomasi
Università di Trento

más allá de la novela y hasta la época contemporánea[2]. Por primera vez los especialistas de distintas Universidades italianas, quienes desde hace muchos años se dedican al estudio de la literatura caballeresca española, se han reunido trabajando de forma sinérgica entre sí y con otros centros de investigación que se ocupan del mismo tema. La investigadora principal del proyecto es Anna Bognolo de la Università di Verona. Las otras unidades que conforman el proyecto son: la de la Università di Trento, cuya responsable es Claudia Demattè; la Università di Roma, La Sapienza, encabezada por Elisabetta Sarmati; y la unidad de la Università di Salerno, dirigida por Daniele Crivellari. Cada equipo explora un aspecto distinto y a la vez decisivo del éxito y la difusión de la literatura caballeresca a partir de un común núcleo de obras, precisamente los libros de caballerías castellanos de los siglos XVI y XVII.

El trabajo sinérgico entre distintos equipos que el proyecto *Mapping Chivalry* supone, permite explorar los diversos ámbitos en los que las obras caballerescas ejercieron su influencia. Pues bien, gracias al enfoque digital, nos proponemos ahondar en campos aún desconocidos de estos textos; sistematizar los datos sacados a luz a través de métodos computacionales, combinándolos con el conocimiento del género por parte de estudiosos expertos; y, finalmente, representar el complejo fenómeno de la difusión de estas obras a través de mapas conceptuales y bases de datos (Bognolo y Bazzaco, 2019; Tomasi, 2020 y 2022).

En la Università di Trento, bajo la dirección de Claudia Demattè, tenemos el objetivo de realizar el banco de datos MeMoRam[3],

2 Como señala Lucía Megías, se cuentan «decenas de traducciones a todas las lenguas de cultura de su momento [...] sin olvidar las reescrituras teatrales, los juegos escénicos, los motivos para la celebración de fiestas y de actos solemnes» (Sales Dasí, 2004: 11).

3 La presentación de las distintas unidades que confrman el proyecto sigue un orden cronológico que parte de las fuentes literarias compartidas por todo el equipo, es decir, los libros de caballerías castellanos, hasta llegar a sus reescrituras de época contemporánea pasando por las traducciones de los siglos XVI y XVII y el teatro del Barroco.

donde se coleccionan los principales motivos y personajes detectados en los textos que constituyen las fuentes del entero proyecto, es decir, el corpus completo de los libros de caballerías castellanos de los siglos XVI y XVII. Dicho corpus se ha definido a partir de varios catálogos a nuestro alcance, entre ellos destacan la *Bibliografía de libros de caballerías castellanos* de Eisenberg y Marín Pina (2000) y los trabajos de Lucía Megías (2002 y 2019). La lista cuenta, finalmente, con ochenta y dos obras. Estas se han sistematizado en la plataforma *Muruca Bibliography*[4], donde se recogen unas fichas con información bibliográfica sobre los textos. Junto a la realización de las fichas, empezamos a organizar el material de la base de datos MeMoRam, que aloja los motivos literarios que aparecen en las obras del corpus. Para corroborar el planteamiento según el que los libros de caballerías no son «todos una mesma cosa» (Cervantes [1977] 2015, I, XLVII: 618) como, en cambio, opinaba el canónigo en el *Quijote*, es fundamental reconocer las unidades narrativas mínimas y recurrentes que estructuran las tramas, ya que, para hacer tan solo un ejemplo emblemático, en el caso de los combates las secuencias «se articulan de acuerdo con unos estereotipos tradicionales, ya sea el combate a caballo con la lanza, ya sea el combate a pie con la espada, pero ninguno resulta idéntico» (Cacho Blecua, 2002: 32)[5]. En la base de datos MeMoRam tendrán especial relieve los fragmentos concretos en los que los motivos se desarrollan, para que puedan apreciarse los posibles desvíos a los que se someten y delinearse, de esta manera, las metamorfosis de los motivos, que en ocasiones siguen logrando éxito a través de sus divergencias,

4 La herramienta fue elaborada por Net7, una empresa italiana líder en las Humanidades Digitales, con la que colaboramos para la realización de nuestro proyecto. Se trata de un framework que presenta una gran cantidad de funcionalidades distintas y permite crear bibliotecas digitales. Más abajo ahondamos en los detalles acerca del diseño de las fichas que integran la plataforma.

5 Véanse también las definiciones que Cacho Blecua ofrece del motivo en 2012 y 2020.

o bien fracasan, sin que vuelvan a repetirse en textos sucesivos. Gracias a un convenio con la Universidad de Alcalá de Henares, contamos con las ediciones de los libros de caballerías que se han publicado hasta la fecha en la colección Libros de Rocinante[6]. Nos fue posible, pues, empezar a poner a prueba unas herramientas de Humanidades Digitales con el fin de detectar los motivos recurrentes de los libros de caballerías y trazar su uso y transformaciones en el corpus[7]. Entre los recursos que desde hace ya varios años se emplean en la investigación literaria y el estudio del lenguaje natural, la minería de textos se ha demostrado especialmente eficaz[8]. Sin embargo, la mera aplicación de una herramienta, por útil que sea, no es suficiente para obtener datos fiables. Por eso el producto digital que queremos proporcionar a los usuarios es el fruto de la combinación de diversos instrumentos: por un lado nos apoyamos en los índices de motivos (folclóricos y caballerescos) que ya existen para elaborar una lista representativa de las unidades narrativas que componen nuestras obras[9]; por otro lado, utilizamos programas y plataformas digitales tanto para procesar los textos, como para extraer de los mismos sus datos. Así pues, el enfoque tradicional, basado en los estudios que facilitan un hondo conocimiento del contenido de las obras, se combina con el uso de medios computacionales para ofrecer una interpretación eficaz

6 La editorial cuenta con cuarenta y dos obras editadas, de las que manejamos cerca de treinta textos. Agradecemos a Carlos Alvar y José Manuel Lucía Megías su apoyo al presente proyecto y la posibilidad de utilizar estas ediciones que se han convertido en una herramienta imprescindible para nuestro trabajo. Para una lista actualizada de sus publicaciones consúltese: http://www3.uah.es/iimigueldecervantes/es/publicaciones/libros-de-rocinante.

7 Más abajo se encuentran detalles sobre los campos que componen la ficha de cada motivo y un ejemplo práctico de cómo utilizamos las herramientas digitales.

8 Véanse los trabajos de Blevins (2010), Jockers y Mimno (2012), Karsdorp (2013) y Voigt, et al., (1999), entre otros.

9 Se trata especialmente de Thompson (1975); Bueno Serrano (2007) y Luna Mariscal (2013; 2017; 2020).

de los datos que dichos medios nos brindan y acrecentar, de esta manera, el conocimiento acerca de los libros de caballerías.

En el ámbito de la difusión de la materia caballeresca fuera de España, el grupo de la Università di Verona desarrolla su proyecto en el cauce del *Progetto Mambrino*, en el que lleva varios años estudiando la recepción de los libros de caballerías en Italia en lengua original, o mediante traducciones y ahondando en el fenómeno de las obras que Mambrino Roseo da Fabriano agrega a los ya nutridos ciclos españoles (las llamadas *aggiunte*). En el marco del presente proyecto, la unidad tiene por objetivo el de ofrecer ediciones científicas digitales del ciclo del *Amadis di Gaula*, coleccionándolas en una biblioteca digital. La elaboración de dichas ediciones permitirá a los usuarios «la navigazione dei loro contenuti in modo dinamico, accompagnando la lettura delle trascrizioni con immagini in alta risoluzione delle fonti e collegamenti ipertestuali ad altre risorse, inedite o già presenti in rete (cataloghi bibliografici; indici dei nomi, luoghi, motivi; immagini o altri oggetti multimediali)» (Bognolo y Bazzaco, 2019: 25). Esta labor se apoya en la experiencia de decenios de trabajo llevado a cabo por el equipo del *Progetto Mambrino*, gracias al que ya contamos con la digitalización de los ejemplares que se conservan en las bibliotecas locales, como puede apreciarse en su página web[10]. Más allá de estos recursos que tienen un alcance ya notable en el campo de las Humanidades, la unidad de Verona está utilizando herramientas informáticas especializadas para lograr su objetivo de creación de una Biblioteca de Ediciones Científicas Digitales. Entre ellas encontramos los métodos de reconocimiento óptico de los caracteres (OCR) para facilitar las transcripciones (Bognolo y Bazzaco, 2019: 27; Bazzaco, 2018). Los estudiosos que integran la unidad se dedicaron a entrenar el *software* puesto a disposición

10 Las bibliotecas que colaboraron con el equipo son la Biblioteca Civica di Verona, la Biblioteca Civica Bertoliana di Vicenza y la Biblioteca Internazionale La Vigna (Vicenza). Consúltese: https://www.mambrino.it/it, en la sección *Collezioni Digitali.*

por Transkribus[11] a la letra y las características particulares de las obras con las que trabajan, a partir de imágenes en alta resolución, para llegar a obtener resultados rigurosos desde un punto de vista filológico (Bazzaco, 2018: 269-271; Mancinelli, 2016: 256). Como afirma Bazzaco acerca del uso de Transkribus «la variedad de las imágenes cargadas a la plataforma ha permitido ampliar el espectro de aplicación del software, que está acostumbrado ahora a realizaciones gráficas y tipográficas muy distintas» (2018: 270). En esta rama tenemos, pues, un claro testimonio de cómo las Humanidades Digitales, a la vez que sirven de apoyo a nuestro trabajo, avanzan gracias a las aportaciones de los especialistas en las materias de estudio, en una constante interacción entre la interpretación de los datos y los instrumentos utilizados para alcanzarlos[12]. La transcripción de las obras no constituye el punto de llegada del trabajo, sino que de ahí parte su pasaje a en un lenguaje de marcado XML, que facilita la extracción de datos como los nombres de los personajes, o de los lugares. Además, los usuarios contarán con los resúmenes de cada capítulo de las obras que forman parte del corpus (Bazzaco, 2018: 270-71), gracias a las aportaciones del *Progetto Mambrino*, que en este campo elaboró,

11 Se trata de una plataforma puesta a disposición por Readcoop, que utiliza instrumentos de inteligencia artificial para la digitalización, la transcripción y la búsqueda de documentos históricos. Se utiliza especialmente para los manuscritos, pero su empleo se ha demostrado eficaz también para otros tipos de caracteres. El desarrollo de la herramienta se debe a la Universidad de Innsbruck en colaboración con distintos grupos de investigación europeos y se enmarca en el proyecto UE Horizon 2020 READ. Véase https://readcoop.eu/transkribus/.

12 Entre los más relevantes pasos adelante de la investigación a través de herramientas digitales, se encuentra el hecho de que los instrumentos para el análisis textual se nutren de las obras con que los estudiosos decidan alimentarlas y esto apunta a la creación de un «historical electronic text corpus» (Springmann y Anke 2017 cit. indirecta, véase Bazzaco, 2018: 270). Véase también lo que señala Mancinelli sobre el amplio alcance del tipo llamado *Aldino* (*italic*, o cursiva) que caracteriza las ediciones que maneja la unidad de Verona: «training a software on this particular font would be a very useful project for a huge corpora of books» (2016: 256).

además, repertorios de motivos y personajes de los ciclos italianos de *Amadís* (Bognolo, Cara, Neri, 2013) y de *Palmerín* (Bellomi, *et al.*, en prensa). Todos estos resultados irán volcándose en una base de datos puesta a disposición de los usuarios en acceso libre.

Como se ha apuntado arriba, el material narrativo de los libros de caballerías castellanos sobrevivió durante tanto tiempo también gracias a su capacidad de amoldarse a nuevas formas y la unidad de la Università di Salerno, dirigida por Daniele Crivellari, se dedica a la colección de los datos relativos a las obras dramáticas del siglo XVII centradas en episodios, personajes, o motivos de los libros de caballerías castellanos, con el fin de realizar el banco de datos TeatroCaballeresco. La unidad tiene dos objetivos fundamentales: el estudio de las obras y su edición digital. En lo que atañe al primer aspecto, los investigadores se plantean ofrecer el acceso a los textos, sus datos bibliográficos y su relación con las fuentes más directas, con las que trabaja la unidad de Trento; a nivel intratextual se presta atención a temas y motivos detectados en las obras y, a nivel transtextual, se pondrán en relación las obras dramáticas con los libros de caballerías de los que proceden. Este último aspecto es clave en el estudio del teatro caballeresco y la clasificación de las distintas reelaboraciones del material narrativo se remonta al *Repertorio bibliografico e studio interpretativo del teatro cavalleresco spagnolo del sec. XVII* de Demattè (2005), punto de partida del trabajo de la unidad. En el repertorio se identifican las técnicas transtextuales que los dramaturgos normalmente adoptaban al escribir sus piezas: la dramatización de episodios de libros de caballerías que se podían identificar de manera clara; el uso libre de elementos caballerescos; la dramatización en segundo grado de la materia caballeresca; la reescritura a lo divino y la paródica y/o burlesca (Demattè 2005: 37-46). Por lo que se refiere al segundo objetivo de la unidad de Salerno, el grupo tiene prevista la publicación de ediciones críticas fiables de los textos dramáticos áureos relacionados, a varios niveles, con libros de caballerías; las ediciones, que serán preparadas en colaboración con la editorial Reichenberger, se alojarán en la Biblioteca Virtual Miguel de Cervantes.

Lejos de sucumbir en la hoguera a la que se destinaron algunos de los libros que se encontraban en la biblioteca de don Quijote (Cervantes [1977] 2015, I, VI), las aventuras caballerescas dieron prueba de gran vitalidad, hasta volver a florecer en épocas distintas y lejanas respecto a la que las vio nacer. Elisabetta Sarmati, encabezando la unidad de la Università di Roma La Sapienza, explora los frutos contemporáneos que proceden de las fuentes caballerescas hispánicas. No obstante la importante revitalización que la caballería tuvo en el siglo XX, especialmente en la época después de la guerra civil, y en la contemporaneidad, el fenómeno queda aún por la mayoría inexplorado. Es por esto por lo que el objetivo principal de la unidad es elaborar un banco de datos de las obras de narrativa caballeresca contemporánea llamado AmadíssigloXX, que aloja unas fichas bibliográficas articuladas en trece campos, donde se recogen los metadatos acerca de los textos (Sarmati, 2021: 250-252). Entre estas obras destacan, por ejemplo, *Las aventuras del caballero Florestán del Palier*, de Wenceslao Fernández Flórez (1959)[13]; *El caballero, la muerte y el diablo*, *La historia del caballero Rafael* (1956) y *Merlín y familia* (1955) de Álvaro Cunqueiro; o *Libro de caballerías* (1957) y *Las aventuras del caballero Kosmas* (1981) de Joan Perucho. En ocasiones, los autores, a través de la representación de la antigua caballería, reaccionan al realismo que caracteriza la narrativa de mediados del siglo XX. También las reelaboraciones contemporáneas del *Quijote* de Cervantes tendrán un lugar de relieve en la base de datos[14]. De todas las obras incluidas en el corpus se estudian las variaciones respecto a los textos originales en relación a distintas categorías: las estructuras narrativas, los motivos, los personajes, el espacio y el tiempo de la acción. Se plantean problemas de crítica textual, puesto que las obras del corpus pueden configurarse como imitaciones de los textos originales, o bien como novelas

13 Un estudio pormenorizado de la obra se encuentra en Sarmati (2020).

14 Para el corpus completo véase Sarmati (2020 y 2021).

caballerescas amoldadas a situaciones y escenarios típicos de la era contemporánea, o bien en ellas pueden coexistir los dos mundos. Hay que determinar, pues, si el tema caballeresco en los textos es predominante, parcial o tan solo aludido. Entre los objetivos primarios de la unidad de investigación se encuentra también el análisis de la revitalización del género caballeresco desde un punto de vista socio-cultural.

## 2. LOS LIBROS DE CABALLERÍAS CASTELLANOS Y SU ORGANIZACIÓN EN FICHAS.

Merece la pena detenerse en la presentación de los libros de caballerías, ya que desempeñaron un papel fundamental en el desarrollo de la novela moderna, como han demostrado los estudios más recientes acerca de estas obras. Afirma Lucía Megías: «El *Quijote*, como repite hasta la indecencia la crítica, es la primera novela moderna...pero lo es por su calidad literaria y por algunas de sus propuestas (...), no por su capacidad de crear una ficción de entretenimiento: este había sido mérito de los libros de caballerías» (Sales Dasí, 2004: 10). En efecto, es destacable el influjo que estos libros tuvieron en la realización de la obra maestra cervantina. El autor logró escribir el más revolucionario de los libros de caballerías (Cacho Blecua, 2002: 28), sacando partido de una materia bien conocida tanto por él como por su público, reelaborando el material narrativo hasta crear una obra en la que, según señala Lucía Megías, los aficionados podían encontrar: «algo conocido (las aventuras -en este caso, humorísticas- de un caballero, las aventuras típicas y tópicas de un caballero andante junto a su escudero), pero a su vez, (...) algo distinto: ni el caballero era tal ni las aventuras terminaban como se esperaba. El *Quijote, como libro de caballerías,* puede calificarse como una "feliz síntesis" de la tradición caballeresca del siglo XVI» (2002: 33). Así pues, a pesar de las pasadas lecturas del libro de Cervantes como causa de la decadencia del género, hoy en día a la pregunta sobre si

«hubiese tenido alguna excusa o pretexto Cervantes para escribir su inmortal relato de no haber existido los libros de caballerías» (Sales Dasí, 2008: 200) podemos contestar de forma negativa. Tampoco puede pasar desapercibido el enorme éxito que estos textos tuvieron en la escena editorial renacentista, al lado de las más estudiadas manifestaciones de las letras áureas hispánicas[15]. Los libros «grandes, muy bien encuadernados» (Cervantes [1977] 2015, I, VI: 150) que le secaron el cerebro al famoso hidalgo manchego, ocupan los talleres de imprenta a partir de la primera edición conocida del *Amadís de Gaula* de Montalvo, a finales del siglo XV[16], y permanecen en auge durante más de cien años (Lucía Megías, 2002: 9) con obras, como precisamente la refundición de Montalvo, que cuentan con hasta veinte reediciones, volviendo a ser los *best-sellers* de la época (Bognolo y Bazzaco, 2019: 23; Sales Dasí, 2008: 199)[17].

Según la metáfora de Gayangos, la literatura caballeresca «arranca de las profundidades de la Edad Media y esparce sus ramas y derrama su influencia casi hasta nuestros días» (1857: LXI). Los libros de caballerías que aparecen en el Renacimiento son el fruto de la elaboración de distintas fuentes que, en los siglos anteriores, se habían difundido por toda Europa. Entre ellas encontramos la llamada Materia de Bretaña (Alvar, 2008), con los relatos del rey Arturo y los caballeros de la Mesa Redonda, a los que se añaden los amores y las aventuras de Tristán e Iseo, como transfondo literario a partir del que florecen las versiones medievales del *Amadís de*

15 Entre ellas encontramos la ficción sentimental, el género celestinesco, la poesía de cancionero, el petrarquismo, la novela picaresca y los libros de pastores (Lucía Megías y Sales Dasí, 2008: 19).

16 La primera edición del *Amadís de Gaula* de Montalvo que ha llegado hasta nuestros días es la impresa en Zaragoza en 1508 en el taller de Jorge Coci.

17 Para un catálogo completo de los títulos de libros de caballerías y datos sobre reediciones, secuelas y continuaciones, véase Lucía Megías (2002: 46-60). Para una nueva propuesta de clasificación de los ciclos caballerescos véase Demattè (en prensa).

*Gaula* (Gómez Redondo, 2008[18], Rodríguez de Montalvo, [1988] 2008: 19-37). Así pues, los descendientes renacentistas de Amadís[19] empiezan a dar sus primeros pasos en una época de grandes cambios, cuando se abre el camino a la Edad Moderna, durante el reinado de los Reyes Católicos[20]. Se trata de «un período durante el que conviven miradas que amalgaman las sólidas herencias del pasado y los albores de unos tiempos paulatinamente diferentes» (Mérida Jiménez, 2007: s.p). Colocándose en el centro de tal transformación, los libros de caballerías del siglo XVI se vuelven de alguna manera los testimonios de su momento histórico, caracterizado por descubrimientos y conquistas[21], a la vez que son entre los productos más prolíficos de un nuevo y potente medio

18 El estudioso hace hincapié también en el papel crucial que desempeña el *Libro del cavallero Zifar* en la conformación de la literatura caballeresca hispánica y en la recepción de la materia artúrica. Véase Gómez Redondo (2008: 56-59).

19 Entre los títulos de las series más famosas, además de los libros de Amadís, se encuentran por ejemplo: *Palmerín de Olivia* (1511), con *Primaleón* (1512) y *Platir* (1533); *Clarián de Landanís* (1518) de Gabriel Velázquez de Castillo, del que Jerónimo López escribe la segunda parte titulada *Floramante de Colonia* (1550), el libro tercero (1524) y *Lidamán de Ganaíl* (1528), a los que se añade el libro segundo de la primera parte (1522) por Álvaro de Castro; en la segunda mitad del siglo XVI empiezan a salir los volúmenes de la serie de entretenimiento *Espejo de príncipes y caballeros* (1555, Diego Ortúñez de Calahorra), de la que en 1623 se vuelven a publicar la tercera y cuarta partes por Marcos Martínez y que cuenta con un quinto libro manuscrito (Lucía Megías 2002: 46-59).

20 Los libros de caballerías se vuelven, en ocasiones, instrumentos de propaganda y promotores de comportamientos modélicos para la sociedad de la época (Lucía Megías y Sales Dasí, 2008: 20-22). Véase también Marín Pina (1995).

21 Mérida Jiménez subraya la fundamental diferencia entre los textos caballerescos más antiguos y los libros de caballerías que se publican a partir de la primera edición conservada del *Amadís* de Montalvo y afirma, además, que «sólo a partir de la ascensión al trono del emperador Carlos los libros de caballerías iniciaron su rápida adaptación a las nuevas geografías. Entre sus lectores, además del propio Carlos V, cronistas como Gonzalo Fernández de Oviedo o soldados como Hernán Cortés, quienes pudieron vivir sus conquistas, por primera vez, como si fueran amadises y como recreadores de las valientes hazañas de su hijo Esplandián» (2007: s.p.).

de difusión: la imprenta, llegando a ser «uno de los pilares más estables del comercio editorial castellano» (Lucía Megías y Sales Dasí, 2008: 21).

Con el proyecto *Mapping Chivalry* seguimos en el cauce de la innovación, abordando los contenidos de los libros de caballerías desde la perspectiva de las Humanidades Digitales, que constituyen una parte fundamental de nuestras investigaciones. En este contexto, el diseño conceptual de los bancos de datos de cada equipo es el fruto de la colaboración con la empresa informática Net7[22], que cuenta con años de experiencia en este ámbito, al realizar muchos portales sobre literatura tanto italiana como internacional. Como afirma Antonucci, para la creación de productos digitales los estudiosos de literatura no podemos limitarnos a «decirle al informático qué es lo que se quiere crear y dejarle luego que trabaje por su cuenta, recibiendo un producto final sin más compromiso de nuestra parte, sino de seguir paso a paso el trabajo del informático» (2018: 82). En efecto, la elaboración de las fichas que recogen los metadatos de los libros de caballerías en *Muruca Bibliography* nos ocupó durante varios meses de intenso trabajo, cooperación e intercambio con la empresa y los otros miembros de nuestro equipo. Lo que nos pareció oportuno poner de relieve en las fichas, además de la información bibliográfica sobre las obras, fueron los vínculos con los corpora que se estudian en las otras unidades que conforman el proyecto. Dichos vínculos están garantizados por unos códigos diferenciadores, que identifican a cada obra del corpus caballeresco compartido por todas las unidades como fuente y se conectan con las posibles traducciones (unidad de Verona) y reelaboraciones tanto dramáticas (unidad de Salerno) como de narrativa contemporánea (unidad de Roma). Cada ficha se completa con los datos sobre el autor, de no tratarse de textos anónimos, el lugar y la fecha de la primera edición. Aparece también la posible colocación de la obra

22 https://www.netseven.it. Agradecemos a Chiara Aiola su colaboración en todas las fases del proyecto.

en un ciclo, como es una de las características sobresalientes del género; su naturaleza impresa o bien manuscrita (Lucía Megías, 2002: 31-32) y, en fin, a través de un hipervínculo puede accederse a la base de datos *Clarisel* de la Universidad de Zaragoza que, en la sección *Amadís*, colecciona las fichas detalladas de los estudios críticos sobre los libros de caballerías[23].

Al lado de la selección y sistematización de las obras del corpus, empezamos a manejar los textos mismos, puesto que, como apuntamos arriba, contamos con las ediciones de los libros de caballerías en edición moderna (Libros de Rocinante). La base de datos MeMoRam recoge noticias acerca de las primeras ediciones conservadas de las obras y las modernas cuando las hay. Su implementación supuso una intervención directa en los textos, que se transcribieron a través del programa *Oxygen XML Editor* para su inserción en un campo específico de la base y su estudio mediante herramientas de minería de textos (Tomasi, 2022). Las distintas porciones que marcamos son el título, que permite la identificación de los libros, junto con el prólogo, las posibles partes y los capítulos, por ser las secciones donde se localizan los motivos, objeto principal de nuestros estudios. Pues bien, el proceso de etiquetado de los textos nos llevó a tomar decisiones sobre su representación, ya que, en ocasiones, las normas editoriales de la colección dan lugar al uso de símbolos que obstaculizan la lectura en un lenguaje de marcado. Se trata, por ejemplo, de las erratas, señaladas en las obras editadas entre ángulos. En virtud de la redacción en *Oxygen*, que utiliza simismo los ángulos para las etiquetas, decidimos recopilar cada texto limpio de erratas para facilitar al programa su lectura. Para permitir el reconocimiento de las palabras por parte de *softwares* de procesamiento del lenguaje, sustituimos también otros símbolos que no son compatibles con las normas XML-TEI, como

23 Respecto a la elaboración de las fichas presenté unos detalles también en Tomasi (2022: 273).

la & y juntamos palabras que, en las ediciones modernas, quedan separadas por números o guiones[24].

## 2. MEMORAM. IDENTIFICACIÓN Y SISTEMATIZACIÓN DE LOS MOTIVOS CABALLERESCOS: UN CASO PRÁCTICO.

Tras esta fase preliminar[25], el flujo de trabajo que nos planteamos en la Università di Trento, nos lleva a la parte más experimental del proyecto, es decir, el empleo de herramientas digitales para abordar los contenidos de las obras que manejamos. El centro de nuestras investigaciones es el motivo literario, que se ha vuelto «uno de los ejes de reflexión más importantes en el desarrollo de la crítica literaria moderna» (Luna Mariscal, 2018: 79) y un elemento imprescindible para entender la poética medieval de la que los libros de caballerías son herederos. Este enfoque hacia la literatura de la Edad Media, especialmente la artúrica, llevó a la realización de distintos índices de motivos de raigambre folclórica, que, siguiendo la estela del pionero y monumental *Motif Index* de Thompson (1975), acrecentaron el conocimiento del contenido de distintos corpora. Entre ellos podemos citar los trabajos de Guerreau-Jalabert sobre los *romans* artúricos en francés de los siglos XII y XIII (1992)[26], el índice de Ruck que abarca asimismo temas y motivos de los poemas

24 Otra cuestión relevante a la hora de procesar los textos a través de herramientas de Humanidades Digitales es el carácter conservativo de las normas editoriales, que implica el uso de una grafía antigua difícil de descifrar por dichas herramientas. Este aspecto abre la senda a reflexiones sobre la necesidad de que los instrumentos reconozcan la lengua diacrónica.

25 Dicha fase no ha terminado, ya que el contenido de la base de datos es susceptible de incrementarse a medida que aumenta el número de obras editadas.

26 La estudiosa vuelve a reflexionar sobre la utilidad de los índices de motivos y nos otorga valiosas aportaciones teóricas también en un artículo sucesivo, donde, acerca de los diversos límites que la crítica reconoce en la aplicación del índice de Thompson, apunta: «il vaut mieux un outil imparfait, mais disponible que pas d'outil du tout» (Guerreau-Jalabert, 2005: 3).

artúricos del siglo XII (1991), o las aportaciones de Bordman en su índice de motivos de los *romances* ingleses en verso (1963). Estos valiosos instrumentos de investigación son el testimonio de cómo evolucionan unas mismas unidades narrativas según el diferente contexto histórico, literario y cultural al que se amolden.

Así pues, la carga simbólica que caracteriza a los motivos les permite ser reconocidos más allá de sus múltiples metamorfosis en textos diferentes y a lo largo del tiempo. Además, ellos son los elementos de los libros de caballerías que Román Ramírez[27] supo aprovechar para repetir «la sustancia de las aventuras» cambiando «las razones como le parece» (Cacho Blecua, 2002: 33). En el ámbito de la materia caballeresca hispánica existen varios acercamientos que privilegian el motivo literario como clave de lectura de sus contenidos. Se trata de los ya mencionados estudios de Cacho Blecua (2002; 2012 y 2020), en los que el autor plantea problemas de definición del motivo como unidad. Luna Mariscal elaboró distintos índices de motivos centrados en el corpus de las historias caballerescas breves (2013), en el *Baladro del Sabio Merlín* (2017) y en la *Demanda del santo Grial* (2020). La estudiosa sigue el modelo utilizado por Thompson en su índice, a cuya lista añade diversas entradas a partir de los contenidos específicos de los corpora que maneja. En lo que atañe a los libros de caballerías, Neumayer analizó las tramas de unas obras del ciclo amadisiano (precisamente de *Amadís de Gaula*, *Florisando* y *Lisuarte de Grecia* de Feliciano de Silva) basándose en una concepción del motivo acorde a las fuciones proppianas (2008). Bueno Serrano, en su tesis doctoral (2007) dirigida por Cacho Blecua, nos ofrece un índice

27 La última parte del nombre de nuestra base de datos (-Ram) es un homenaje a Román Ramírez, condenado en auto público de fe después de su muerte y cuya excepcional memoria se consideraba el fruto de un pacto con el diablo (Tomasi 2020: 146, nota 43). Es por esto por lo que en el año 1600 se prendió fuego a una figura que le representaba para que no quedara «memoria del dicho Román sobre la faz de la Tierra» (Archivo Diocesano de Cuenca, Inquisición, legajo 343, expediente 4876, f. 243v, cit. indirecta, véase Díaz Migoyo, 2004: 42).

de motivos de los primeros siete libros de caballerías a partir del *Amadís de Gaula* de Montalvo publicado en 1508. A la obra fundacional se añaden las *Sergas de Esplandián* (1510), *Florisando* (1510), *Palmerín de Olivia* (1511), *Primaleón* (1512), *Lisuarte de Grecia* (de Feliciano de Silva, 1514) y *Floriseo* (1516). Tal índice constituye un punto de partida fundamental para nuestras investigaciones, tanto desde el lado teórico, como práctico. A nivel teórico la estudiosa se plantea varios problemas de definición e identificación de los motivos en los libros de caballerías castellanos y establece que se trata de unidades narrativas (o meta-narrativas) recurrentes[28] que pueden distinguirse a través de marbetes que describen su contenido[29]. Desde el punto de vista práctico, Bueno Serrano reconoce diversas limitaciones en la aplicación del *Motif-Index* de Thompson al grupo de obras seleccionado y para superarlas elabora varias listas de motivos más representativos de su propio material narrativo[30]. Contamos, entonces, con la localización puntual de un abundante número de secuencias, a partir de la que podemos trazar un nuevo conjunto de motivos que se adecúe al corpus completo de los libros de caballerías castellanos, ya que nuestro objetivo es incluir en el estudio todas las obras a nuestro alcance.

En la fase de experimentación sobre la búsqueda automatizada de los motivos en los textos, utilizamos las obras que manejamos en formato digital y el índice de motivos paradigmáticos simples de Bueno Serrano como punto de partida (2007: 1935-2130)[31].

28 O unidad del motivo: la manera en la que este se realiza en cada contexto.

29 Se trata del enunciado del motivo que se elabora a partir de sustantivos deverbales. Para profundizar en este asunto véase Bueno Serrano (2007: 153).

30 La autora realiza también un índice de motivos folclóricos para los libros de caballerías (Bueno Serrano, 2007: 641-938).

31 Agradecemos a Alessandro Bertozzi (Università di Bologna) su ayuda en la elaboración de los datos comentados, que constituyen un paso adelante respecto a los resultados presentados en Tomasi (2022: 278-281).

A través de un *script* del programa Python, *xml_parser*, pasamos los documentos en XML al formato *json*. Dividimos el material obtenido de esta manera en dos carpetas distintas: una dedicada a las obras que se consideran en el índice de Bueno Serrano (*json-books-for_index*)[32] y la otra a las demás obras de nuestro corpus, de las que tenemos marcado el texto (*json-books*).

Escogimos luego una de las entradas del índice, es decir, el motivo que se expresa con el marbete *Curación de las heridas* y, mediante el comando *index_extractor* de Python, convertimos el índice en un documento *json* con una estructura que se pudiese analizar en los pasajes sucesivos de la investigación. A través de otro *script* (*chapters_extractor.py*), que toma en consideración la carpeta *json-books-for_index* y el documento *json* con las referencias puntuales al motivo en los mismos, se coleccionan, en un file TXT (*chapters_extracted.txt*), todos los capítulos identificados en el índice de Bueno Serrano. Para el motivo *Curación de las heridas* los capítulos detectados son 56 en total.

Después de coleccionar estos datos, nos ocupamos de seleccionar los fragmentos específicos en los que el motivo *Curación de las heridas* puede detectarse de forma concreta y para esta operación empleamos la lectura cercana de los capítulos. A partir de esta validación obtuvimos un file TXT como subgrupo de *chapters_extracted.txt*. Las secciones extraídas (*section_extracted*) se sometieron luego a preprocesamiento empleando un *script* (*section_preprocessing.py*) de Python, con el fin de mantener en el texto tan solo sustantivos, verbos y adjetivos, es decir, las partes del discurso que conllevan una carga semántica significativa para nuestros objetivos (POS, *parts of speech*), señalando al mismo tiempo su reiteración. Este enfoque nos proporcionó una lista de palabras (*ranked list words*) que analizamos detenidamente para verificar su validez.

32 Allí se encuentran las obras que manejamos en edición moderna y ya se han pasado a XML: *Lisuarte de Grecia* (de Feliciano de Silva), *Palmerín de Olivia, Primaleón* y *Floriseo.*

Desambiguamos y agrupamos, pues, los términos con la misma raíz para comprobar su incidencia en el texto y obtuvimos un modelo de nueve palabras representativas del motivo *Curación de las heridas* (*cleanded ranked list words*). En los libros de caballerías que se han considerado este motivo suele desarrollarse según dos secuencias principales: el aposentamiento de los caballeros heridos en ricas habitaciones, expresándose la escena con *echar* en *ricos lechos*, o *aposentar*; y el acto práctico de la curación de las heridas gracias a la intervención de personajes específicos[33], que normalmente se explicita con *curar* de las *heridas* (o *llagas*), *apretar* y en unos casos también *atar* las *heridas* (o *llagas*)[34]. A menudo el motivo se presenta también en otro momento, es decir, durante o inmediatamente después de la lucha, al considerar el narrador, o algún personaje, que sería necesario sanar las heridas de los caballeros. A las palabras que pusimos en cursiva el programa atribuyó el mayor peso en los textos procesados[35] y, gracias a la conexión que los términos detectados establecen entre sí, el motivo resulta siempre reconocible, a pesar de que cambie la forma con la que las palabras aparecen, debido a la conjugación de verbos, la declinación de adjetivos y sustantivos o al uso de grafías antiguas.

La elaboración de una lista de palabras vinculadas con el motivo seleccionado, nos permitió hacer luego inferencias en otras obras del corpus de las que tenemos las versiones en XML y *json* (*json*-

33 Entre los aspectos más interesantes del motivo encontramos justamente las figuras a las que está confiada la curación y se trata, por ejemplo, de maestros (médicos), monjas, magos, magas, doncellas o ermitaños.

34 Bien diferente es el caso, analizado por Campos García Rojas en la leyenda tristaniana (2010), de búsqueda de salud tras una herida que causa enfermedad por envenenamiento, donde este último aspecto cobra más importancia y constituye un motivo en si mismo (Bueno Serrano 2007: 1972).

35 Para crear nuestra lista de palabras clave consideramos las que logran un peso de tres o cuatro (en referencia al número de textos en los que aparecen), añadiendo unos sinónimos que nos parecen significativos y las variantes u oscilaciones de una misma palabra.

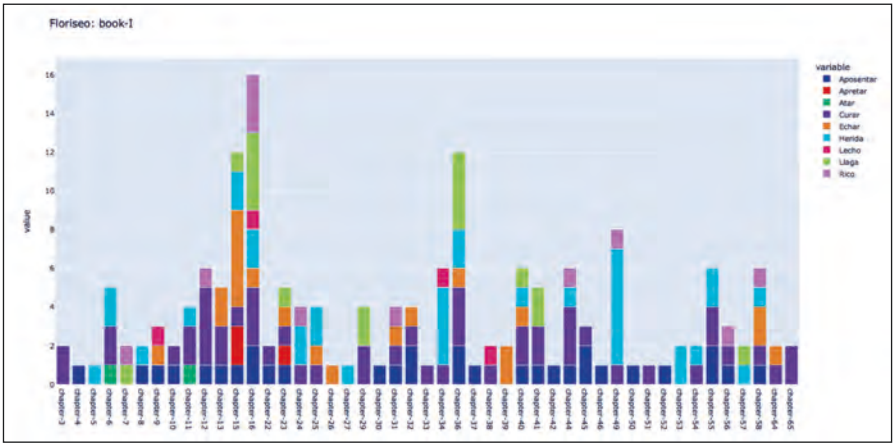



Figura 1. Distribución de las palabras del modelo en el libro I de *Floriseo*. Producido con Plotly.js (v2.9.0).

*books*). Para esta operación utilizamos la función *make_inference.py*, que nos devolvió los capítulos de las obras en los que aparecen las palabras que conectamos con nuestro motivo (*chapters_hot_spots.txt*). Tras obtener estos datos, pudimos visualizarlos de forma eficaz a través de unos gráficos para identificar los *clusters* de palabras que pueden asociarse al motivo buscado. Volviendo a acercarnos a los textos pudimos averiguar la real presencia del motivo en las secciones evidenciadas por el programa y analizar su manera de expresarse en obras diferentes. Este pasaje permite ampliar el modelo de palabras significativas, otorgándonos una identificación más precisa del motivo en el corpus. Además, cruzando los resultados que restituyen estas búsquedas con las referencias indicadas en el índice de Bueno Serrano, es posible establecer una vez más su eficacia y añadir otros episodios en los que tiene lugar la *Curación de las heridas* y se escaparon a la selección manual. En el caso del primer libro de *Floriseo*, por ejemplo, encontramos el motivo también en los capítulos 15, 24, 34, 40 y 49. En el gráfico arriba se aprecia la distribución de las palabras de nuestro modelo en esta obra (figura 1).

En esta fase de la experimentación a través de instrumentos de Humanidades Digitales la labor requiere aún nuestra supervisión hasta que se convierta en un proceso más rápido y ágil. Sin embargo, mediante este trabajo nos pudimos plantear cuestiones sobre la definición del motivo en sentido abstracto, al lado de su detección concreta en los textos. En lo que atañe al primer aspecto fue fundamental revisar las entradas coleccionadas por Bueno Serrano en su índice, eligiendo las etiquetas que no conllevan detalles ni matices y confiando el reconocimiento de posibles divergencias y peculiaridades a la aportación de los estudiosos, que en ningún momento podrá ser reemplazada por los programas informáticos. Estos constituyen, por su parte, un eficaz apoyo en la detección de fragmentos concretos en los textos, como se desprende del caso presentado. A partir de los resultados obtenidos, nuestro trabajo se centrará en el perfeccionamiento de la búsqueda, seleccionando líneas, frases o párrafos en los que las palabras de nuestro modelo tienen que aparecer juntas para que en el estudio no entren fragmentos de texto demasiado amplios, que invaliden el sentido de la investigación.

El largo viaje de los libros de caballerías castellanos, que mezclan a las atmósferas medievales los primeros atisbos hacia la Edad Moderna, pasan por la pluma de un apasionado lector y traductor italiano del siglo XVI, se transforman en las enredadas piezas del teatro barroco e insuflan su vitalidad hasta en la novela contemporánea, aún no ha terminado. Al contrario, los instrumentos de las Humanidades Digitales brindan la ocasión a los estudiosos de acercarse a estas obras con una lupa innovadora y seguir el camino de sus exóticos escenarios descubriendo puntos de conexión entre obras de diferentes épocas y contextos, sin renunciar al rigor de los estudios filológicos que siguen siendo imprescindibles para respaldar y consolidar nuestro conocimiento.

## REFERENCIAS BIBLIOGRÁFICAS


Alvar, Carlos (2008). La Materia de Bretaña. En: José Manuel Lucía Megías (ed.) «Amadís de Gaula»: *1508, quinientos años de libros de caballerías*. Madrid: Biblioteca Nacional de España: Sociedad de Conmemoraciones Culturales: 19-46. Disponible en la Biblioteca Virtual Miguel de Cervantes: https://www.cervantesvirtual.com/obra/la-materia-de-bretana-2/ [9/08/2022].

Antonucci, Fausta (2018). Una nueva herramienta para el estudio del teatro clásico español: Calderón Digital. Base de datos, argumentos y motivos del teatro de Calderón. *Bulletin of the Comediantes*, 70, 1: 79-95.

Bazzaco, Stefano (2018). El Progetto Mambrino y las tecnologías OCR: Estado de la cuestión. *Historias Fingidas*, 6: 257-272 https://historiasfingidas.dlls.univr.it/article/view/89/135 [9/08/2022].

Bellomi, Paola, Bognolo, Anna, Neri, Stefano, Zoppi, Federica (en prensa). *Repertorio delle continuazioni italiane ai romanzi cavallereschi spagnoli. Ciclo di «Palmerino di Oliva»*. Roma: Bulzoni.

Bernal, Fernando, (2003). *Floriseo* (Valencia, Diego de Gumiel, 1516). Ed. Javier Guijarro Ceballos, Alcalá de Henares: Centro de Estudios Cervantinos.

Blevins, Cameron (2010). Topic Modeling Martha Ballard's Diary. *Cameron Blevins.org* https://www.cameronblevins.org/posts/topic-modeling-martha-ballards-diary/. [publicado 1/04/2010] [9/08/2022].

Bognolo, Anna y Bazzaco, Stefano (2019). Tra Spagna e Italia: per l'edizione digitale del Progetto Mambrino. *eHumanista/IVITRA*, 16: 20-36 https://www.ehumanista.ucsb.edu/sites/default/files/sitefiles/ivitra/volume16/3.%20Bognolo.pdf [9/08/2022].

Bognolo, Anna, Cara, Giovanni, Neri, Stefano (2013). *Repertorio delle continuazioni italiane ai romanzi cavallereschi spagnoli. Ciclo di «Amadis di Gaula»*. Roma: Bulzoni.

Bordman, Gerald (1963). *Motif-Index of the English metrical romances*. Helsinki: Suomalainen Tiedeakatemia, Academia Scientiarum Fennica.

Bueno Serrano, Ana Carmen (2007). *Índice y Estudio de Motivos en los Libros de Caballerías Castellanos (1508-1516)*. [Tesis doctoral dir. por Juan Manuel Cacho Blecua]. Universidad de Zaragoza: Zaragoza.

Cacho Blecua, Juan Manuel (2002). Introducción al estudio de los motivos en los libros de caballerías: la memoria de Román Ramírez, En: Eva Belén Carro Carbajal, Laura Puerto Moro, María Sánchez Pérez (eds.). *Libros de caballerías, (de «Amadís» al «Quijote»). Poética, lectura e identidad*. Salamanca: Seminario de Estudios Medievales y Renacentistas, Sociedad de Estudios Medievales y Renacentistas: 27-57.

Cacho Blecua, Juan Manuel (ed.) (2012). *Revista de Poética Medieval. El motivo en la literatura sapiencial*, 26.

Cacho Blecua, Juan Manuel (2020). El *Motif-Index* de S. Thompson y sus aplicaciones en la literatura caballeresca, *Historias Fingidas*, 8: 5-54 https://historiasfingidas.dlls.univr.it/article/view/729/pdf [9/08/2022].

Campos García Rojas, Axayácatl (2010). Heridas, veneno y búsqueda de salud: apuntes comparativos para la leyenda de Tristán e Iseo. *Destiempos. Revista de curiosidad cultural*: 257-278.

Cervantes, Miguel de [1977] (2015). *Don Quijote de la Mancha*. Vol. I, ed., J. J. Allen, Madrid: Cátedra.

Clarisel: Amadís. Base de datos de literatura caballeresca (Juan Manuel Cacho Blecua y María Jesús Lacarra Ducay). https://clarisel.unizar.es/paginas/index.php?base=amadis&opcion=presentacion [9/08/2022].

Dematтè, Claudia (2005). *Repertorio bibliografico e studio interpretativo del teatro cavalleresco spagnolo del sec. XVII*. Trento: Editrice Università degli Studi di Trento.

Dematтè, Claudia (en prensa). Ciclos de caballerías hispánicos *versus* libros de caballerías 'únicos' y 'sueltos': una propuesta de nueva clasificación. En: Daniel Gutiérrez Trápaga y María Gutiérrez Padilla (eds.). *Libros de caballerías: aproximaciones a la poética de un género literario*. Ciudad de México: Facultad de Filosofía y Letras, UNAM.

Díaz Migoyo, Gonzalo (2004). Memoria y fama de Román Ramírez. En: Francisco Domínguez Matito y María Luisa Lobato López (eds.) *Memoria de la palabra: actas del VI Congreso de la Asociación Internacional Siglo de Oro. Burgos-La Rioja 15-19 de Julio 2002*. Vol 1: 39-54.

Eisenberg, Daniel y Marín Pina, María Carmen (2000). *Bibliografía de los libros de caballerías castellanos*. Zaragoza: Prensas Universitarias.

Silva, Feliciano de, (2002). *Lisuarte de Grecia* (1514). Ed. Emilio Sales Dasí, Alcalá de Henares: Centro de Estudios Cervantinos.

Gayangos, Pascual de (1857), *Libros de caballerías. Con un discurso preliminar y un catálogo razonado por don Pascual de Gayangos*. BAE, 40 Madrid: Atlas.

Gómez Redondo, Fernando (2008). La literatura caballeresca castellana medieval. En: José Manuel Lucía Megías (ed.). «Amadís de Gaula»: *1508, quinientos años de libros de caballerías.* Madrid: Biblioteca Nacional de España: Sociedad de Conmemoraciones Culturales: 53-79. Disponible en la BIblioteca Virtual MIguel de Cervantes: https://www.cervantesvirtual.com/obra/la-literatura-caballeresca-castellana-medieval/ [9/08/2022].

Guerreau-Jalabert, Anita (1992). *Index des motif narratifs dans le romans arthuriens français en vers (XII-XIII siècles)*. Geneve: Droz.

Guerreau-Jalabert, Anita (2005). Une tentative d'indexation. *L'Index des motifs narratifs dans les romans arthuriens français en vers*, *Les Cahiers du Centre de recherches Historiques*, 35: 39-45. Disponible en: https://journals.openedition.org/ccrh/3014 [9/08/2022].

Jockers, Matthew L., Mimno, David (2012). Significant Themes in 19th-Century Literature. *Faculty Publications*, *Department of English*, 105, s.p. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1105&context=englishfacpubs [9/08/2022].

Karsdorp, Folgert, Van der Bosch, Antal (2013). Identifying Motifs in Folktales using Topic Models. In: *Proceedings of the 22 Annual Belgian-Dutch Conference on Machine Learning*: 41-49.

Libros de Rocinante: http://www3.uah.es/iimigueldecervantes/es/publicaciones/libros-de-rocinante. [9/08/2022]

Lucía Megías, José Manuel (2002). Libros de caballerías castellanos: textos y contextos. *Edad de Oro*, XXI: 9-61. Disponible en la Biblioteca Virtual Miguel de Cervantes: https://www.cervantesvirtual.com/obra/libros-de-caballerias-castellanos-textos-y-contextos/ [9/08/2022].

Lucía Megías, José Manuel (2019). Género literario, corpus y difusión de los libros de caballerías castellanos. En: Marta Haro Cortés (dir.). *Libros de caballerías castellanos. 2. Género literario, corpus y difusión*. Monografías *Aula Medieval*, 9 Valencia: Proyecto Parnaseo Universitat de València: 5-47. Disponible en: https://parnaseo.uv.es/

AulaMedieval/AulaMedieval.php?id=monografias#monografias [9/08/2022].

Lucía Megías, José Manuel y Sales Dasí, Emilio José (2008). *Libros de caballerías castellanos (siglos XVI-XVII)*. Madrid: Laberinto.

Luna Mariscal, Karla Xiomara (2013). *Índice de motivos de las historias caballerescas breves*. Vigo: Editorial Academia del Hispanismo.

Luna Mariscal, Karla Xiomara (2017). *El motivo literario en «El Baladro del Sabio Merlín» (1498 y 1535)*. México: El Colegio de México.

Luna Mariscal, Karla Xiomara (2020). De Stith Thompson a las plataformas digitales: algunas reflexiones (con un Índice de motivos de la *Demanda del Santo Grial*, Toledo, 1515). *Historias Fingidas*, 8: 55-128 https://historiasfingidas.dlls.univr.it/article/view/164/pdf [9/08/2022].

Mancinelli, Tiziana (2016). Early printed edition and OCR techniques: what is the state-of-art? Strategies to be developed from the working-progress Mambrino project work. *Historias Fingidas*, 4: 255-260 https://historiasfingidas.dlls.univr.it/article/view/65/104 [9/08/2022].

Marín Pina, María Carmen (1995). La historia y los primeros libros de caballerías españoles. En: Juan Paredes (ed.) *Medioevo y Literatura Actas del V congreso de la Asociación Hispánica de Literatura Medieval (Granada, 27 septiembre – 1 octubre 1993)*. III, Granada: Universidad de Granada, pp. 183-192.

Marín Pina, María Carmen (1998). Motivos y tópicos caballerescos. En: Francisco Rico (ed.) *Don Quijote de La Mancha*. Barcelona: Crítica: 857-902. Disponible en: https://cvc.cervantes.es/literatura/clasicos/quijote/introduccion/apendice/marin.htm [9/08/2022].

Mérida Jiménez, Rafael M. (2007), Los libros de caballerías en América: huellas culturales y cultura impresa (1492-1516). Tirant: Bulletí informatiu i bibliogràfic, 10.

Net7: https://www.netseven.it [9/08/2022].

Neumayer, Kristin (2008). *Index and Study of Plot Motifs in some Spanish 'Libros de Caballerías'*. [Tesis doctoral]. University of Winsconsin: Madison.

*Palmerín de Olivia*: (2004).*«Palmerín de Olivia»* (Salamanca, [Juan de Porras], 1511). Introd. M.ª Carmen Marín Pina. Ed. y apénd.

Giuseppe di Stefano. Rev. Daniela Pierucci, Alcalá de Henares: Centro de Estudios Cervantinos.

*Primaleón*: ¿Vázquez, Francisco?, (1998). *Primaleón.* Salamanca, 1512. Ed. M.ª Carmen Marín Pina, Alcalá de Henares: Centro de Estudios Cervantinos.

Progetto Mambrino: https://www.mambrino.it/it/ (Anna Bognolo y Stefano Neri) [9/08/2022].

Rodríguez de Montalvo, Garci [1988] (2008). *Amadís de Gaula*. Ed. Juan Manuel Cacho Blecua, Madrid: Cátedra.

Ruck, Elaine Heather (1991). *An index of themes and motifs in 12th century French Arthurian poetry*. Arthurian studies, 25, Cambridge: D. S. Brewer.

Sales Dasí, Emilio, José (2004). *La aventura caballeresca: epopeya y maravillas*. Pr. José Manuel Lucía Megías, Alcalá de Henares: Centro de Estudios Cervantinos.

Sales Dasí, Emilio, José (2008). Los libros de caballerías por dentro. En José Manuel Lucía Megías (ed.). «Amadís de Gaula»: 1508, *quinientos años de libros de caballerías*. Madrid: Biblioteca Nacional de España: Sociedad de Conmemoraciones Culturales: 197-242. Disponible en la Biblioteca Virtual Miguel de Cervantes: https://www.cervantesvirtual.com/obra/los-libros-de-caballerias-por-dentro/ [9/08/2022].

Sarmati, Elisabetta (2020). Motivos caballerescos y reescritura paródica: el *Florestán del Palier* (1959) de Wenceslao Fernández Florez. *Historias Fingidas*, 8: 257-273 https://historiasfingidas.dlls.univr.it/article/view/154/pdf [9/08/2022].

Sarmati, Elisabetta (2021). Para una aproximación a las reescrituras modernas (siglos XIX-XX) de los libros de caballerías castellanos y del *Quijote*. El portal AmadíssigloXX. *Historias Fingidas*, 9: 231-258 https://historiasfingidas.dlls.univr.it/article/view/1094/pdf [9/08/2022].

Thompson, Stith (1975). *Motif-Index of Folk Literature*. Bloomington and London: Indiana University Press.

Tomasi, Giulia (2020). Las Humanidades Digitales y la base de datos MeMoRam: para un enfoque sistemático hacia los motivos en los libros de caballerías. *Historias Fingidas*, 8: 129-156. https://historiasfingidas.dlls.univr.it/article/view/155/pdf [9/08/2022].

Tomasi, Giulia (2022). Realización de una base de datos de los motivos caballerescos: presentación y avances de MeMoRam. *Historias Fingidas*, Número Especial 1 Humanidades Digitales y estudios literarios hispánicos: 271-288. https://historiasfingidas.dlls.univr.it/article/view/1098/1233 [9/08/2022].

Transkribus: https://readcoop.eu/transkribus/ [9/08/2022].

Voigt, Vilmos, Preminger, Michael, Ládi, László, Darányi, Sándor (1999). Autmated Motif Identification in Folklore Text Corpora. *Folklore (Tartu)*, 12: 126-141. https://www.folklore.ee/folklore/vol12/motif.htm [9/08/2022].

# ENTRE LA HISTORIA Y LA FILOLOGÍA

## EL PROYECTO *CORDICan* Y EL VALOR DE LA INTERDISCIPLINARIEDAD EN LAS HD A TRAVÉS DE LA DOCUMENTACIÓN RELATIVA A *DOTES* Y *ARRAS*

Ana Viña Brito, Dolores Corbella, Roberto J. González Zalacain

**Resumen:**
En este artículo se da a conocer el proyecto *Corpus Digital de las Islas Canarias* (*CORDICan*), cuyo objetivo es el de ofrecer una edición digital de diversas tipologías de documentación manuscrita, procedente de archivos de las islas Canarias, correspondientes a la Baja Edad Media y a la Edad Moderna. Además de una presentación general del corpus, se muestran las posibilidades analíticas que ofrece a partir del estudio de caso de documentación relacionada con dotes y arras.

**Palabras clave:** Corpus, Historia, español de Canarias, dotes, documentación notarial.

**Abstract:**
This article presents *Corpus Digital de las Islas Canarias* (CORDICan) project. It wants to offer a digital edition of a variety of manuscript records from Canary Islands archives. The chronology of the projects covers the late

Ana Viña Brito
Instituto de Estudios Medievales y Renacentistas – Universidad de La Laguna
Dolores Corbella
Instituto de Estudios Medievales y Renacentistas – Universidad de La Laguna
Roberto J. González Zalacain
Instituto de Estudios Medievales y Renacentistas – Universidad de La Laguna

Middle Ages and the Modern Age. In this paper a general presentation of the corpus is made, and also the analytical possibilities offered by the case study of documentation related to wedding dowries are shown.

**Keywords:** Corpus, History, Canarian Spanish, Dowries, Notarial Records.

## 1. INTRODUCCIÓN

La investigación en Humanidades se enfrenta en la actualidad a múltiples retos: los propios derivados del objeto de estudio y los inherentes a la difusión y divulgación de sus resultados y, por tanto, a la transferencia del conocimiento y del patrimonio. La gestión de una cada vez más notable cantidad y variedad de datos y su puesta a disposición del público, tanto especializado como simplemente interesado en estas temáticas, constituye un desafío constante para quieres afrontan tareas de investigación. La integración de la tecnología en todos estos procesos relacionados con el conocimiento científico en los distintos campos humanísticos ha abierto múltiples caminos para optimizar los resultados y diseminarlos de manera más eficaz. En las páginas que siguen se va a detallar el desarrollo de un proyecto específico, el *Corpus Documental de las Islas Canarias* (*CORDICan*), cuya trayectoria se puede insertar en esa progresiva incorporación de las nuevas tecnologías al análisis multidimensional de la sociedad, la lengua y la cultura en el pasado.

Desde hace algo más de un lustro un equipo de investigación interdisciplinar formado por historiadores, filólogos y especialistas de archivo han puesto en marcha, en diversas fases, un proyecto destinado a desarrollar *CORDICan.* Los primeros pasos se dedicaron a comprender desde un punto de vista histórico y lingüístico la importancia de la agroindustria azucarera en el proceso de colonización de las islas Canarias. Se trataba de desentrañar el papel que la terminología utilizada en la puesta en marcha de esa industria tuvo en la configuración del español atlántico y en la expansión transoceánica. La explicación tradicional, emanada de los repertorios lexicográficos americanos, adscribían a este ámbito, y más concretamente en muchos casos al área de Brasil, la formación de un léxico azucarero específico. Esta primera investigación, liderada por Cristóbal Corrales, Dolores Corbella y Ana Viña, requirió la consulta archivística y la transcripción de un número significativo de documentos manuscritos de diversa tipología,

pero todos relacionados con la implantación de la agroindustria y la instalación de ingenios, en un marco cronológico amplio que iba desde la fase final de la conquista de las islas Canarias, a fines del siglo XV, hasta comienzos del siglo XVIII.

Fruto de estos primeros trabajos fueron tres monografías. Una de ellas se dedicó al análisis del proceso de difusión del léxico azucarero específico por el mundo atlántico (Corrales, Corbella y Viña Brito, 2014) mientras que las otras recogían un amplio conjunto de documentación conservada en los archivos de Canarias (Viña Brito, Corrales y Corbella, 2014 y 2016), primer eslabón de la selección documental que se integra en el corpus. A su vez, esa labor conjunta desarrollada por filólogos e historiadores dio lugar a la reflexión sobre el valor que la puesta a disposición de los especialistas de estos campos del saber científico tendría no solo la colección de transcripciones documentales realizada, sino también la infraestructura informática desarrollada al efecto. Esto permitiría superar las limitaciones de la edición en papel al tiempo que aprovechar al máximo el conjunto de metadatos utilizados con el fin de facilitar el análisis propuesto. Se encuentra ahí, precisamente, la base de este *Corpus Documental de las islas Canarias* (*CORDICan*), concebido como un proyecto *in fieri*, que recopila, transcribe y analiza registros manuscritos redactados en las islas o sobre las islas en el marco cronológico ya señalado. Para esta labor se ha contado con el apoyo financiero obtenido mediante convocatorias públicas procedente de diversas fuentes[1].

El diseño del proyecto se inscribe en la convergencia de una doble tradición de investigación. Por un lado, se enmarca en la consolidada trayectoria de edición de fuentes históricas que se lleva a cabo en las islas Canarias desde hace casi un siglo, en la que se viene reflexionando sobre la necesidad de respetar al máximo

1. FFI2016-76154-P (2016-2020, Ministerio de Economía y Competitividad, IP Dolores Corbella, y Programa Mª Rosa Alonso (2018-2019, Cabildo Insular de Tenerife), IP Dolores Corbella y Ana Viña Brito.

la representación original del texto manuscrito (Viña Brito y Corbella, 2018: 3-4). A su vez, desde la perspectiva filológica, se fundamenta en todo el trabajo sobre la conformación de corpus históricos, en cuyo desarrollo se ha producido el impacto positivo de la revolución tecnológica aplicada a ellos, con unos resultados significativos en diferentes contextos académicos en los últimos años (Claridge, 2008).

Para su desarrollo se optó por la utilización de TEITOK. Se trata de un aplicativo web destinado específicamente a la creación de corpus (Janssen, 2016). La elección vino motivada por la experiencia previa de proyectos con ambiciones similares, como *P.S. Post Scriptum*[2], que tuvo como finalidad la publicación digital de varios miles de cartas con el doble objetivo de diseminación de información y de constitución de dos corpus anotados, uno en español y otro en portugués, de un millón de palabras cada uno (Vaamonde, 2015). La elección de este sistema viene motivada porque permite combinar tanto la anotación textual como lingüística en un único documento XML, marcado a partir del esquema TEI. Además, TEITOK incorpora varias herramientas de programación neurolingüística (PNL) que facilitan el procesamiento de los textos. Desde el punto de vista de la experiencia de usuario, ofrece múltiples posibilidades de visualización de la transcripción en función del formato de salida elegido, y a su vez integra un visor de imágenes que permite acompasar la lectura con la reproducción digital de los documentos, propiciando una mayor transparencia en los procesos de transcripción y ofreciendo a la comunidad científica todos los elementos necesarios para un aprovechamiento máximo de la documentación publicada.

La puesta en marcha definitiva de la plataforma, que se encuentra en la última fase de desarrollo, se ha llevado a cabo merced a la obtención de un proyecto de investigación actualmente

2. http://teitok.clul.ul.pt/postscriptum/. Fecha de consulta 10-04-2022.




Figura 1. Esquema de metadatos de *de CORDICan*.

en vigor, y que ha permitido la implementación de la aplicación y su perfeccionamiento desde un punto de vista técnico a partir de las posibilidades ofrecidas por TEITOK. Para la gestión de la información contenida, se ha diseñado un esquema de metadatos que incorpora los créditos de identificación del documento junto con otro conjunto de elementos básicos para la gestión y cruce de información. En cuanto a la caracterización de las personas intervinientes, ya como productoras, ya como destinatarias, se incorpora asimismo una serie de campos específicos (ocupación, género, clase social) que permiten mejorar las búsquedas, tal y como veremos en el ejemplo

descrito en las páginas siguientes. Además, los metadatos incluyen datos cronológicos y geográficos. Estos parámetros permiten, en el conjunto del corpus, poder refinar y acotar las búsquedas, mientras que la herramienta de geolocalización y visualización de datos a partir de coordenadas facilita la identificación de los lugares de producción de los documentos. Veremos también un ejemplo concreto en las páginas que siguen.

En paralelo a este trabajo, en la plataforma se ha ido ampliando la base documental del corpus con el objetivo de enriquecer su contenido en la publicación definitiva. A la documentación vinculada con la agroindustria azucarera que, como ya se ha indicado, ha constituido el punto de partida, se ha ido sumando una amplia selección de textos manuscritos pertenecientes a diversas tipologías: notarial, judicial y epistolar. Además, se ha incorporado la transcripción de numerosos testamentos e inventarios, entre los que las dotes y arras constituyen un apartado destacado. La estimación de documentos publicados cuando finalice esta fase del proyecto se situará en torno al medio millar de testimonios de diversas tipologías, con aproximadamente un millón de palabras.

## 2. DOTES Y ARRAS EN *CORDIcan*

Los estudios sobre dotes y arras son relativamente abundantes desde hace años, tanto desde la óptica jurídica como histórica. No obstante, se ha solido abordar el análisis de las elites en el ámbito urbano, con investigaciones señeras sobre el ajuar doméstico centrado en este sector (Ronquillo Rubio, 1992), mientras que se ha tratado en menor medida el otro contingente importante de población radicada en el mundo rural y no siempre tan favorecida (Borrero Fernández, 1984). Este desfase se debe, en parte, a la escasez y al carácter reiterativo de los bienes dotales que contienen los documentos ligados a los ambientes con menos recursos, pero también a la dificultad que entraña el acceso a las fuentes documentales originales y mucho más cuando estas son periféricas.

De ahí el interés que presenta el corpus *CORDICan* para el análisis de esta temática, al ofrecer una serie relativamente importante de manuscritos centrados en las aportaciones recibidas por los contrayentes o sus familiares al concertar el vínculo matrimonial.

El análisis de este subcorpus dentro de la base de datos documental permite rastrear el significado de los conceptos 'dote' y 'arras' a través de cartas, recibos, memoriales, testamentos, litigios, etc. de todos los grupos que conformaron la sociedad canaria de los siglos XVI y XVII, contextualizándolos con la legislación coetánea y con las características propias de una sociedad de frontera. Este último factor facilita el estudio de los aportes de la población foránea y los distintos mecanismos de integración, ya que los lazos establecidos por matrimonio se convirtieron en una de las vías principales de adhesión y cohesión de la nueva sociedad insular (González Zalacain, 2005; Viña Brito 2012 y 2017). El repositorio documental se convierte así en una fuente primaria para establecer la prosopografía de los distintos personajes y para dilucidar las líneas de transmisión de sus bienes, al tiempo que permitirá contrastar la especificidad diatópica en relación con otras áreas.

### 2.1. DATOS TÉCNICOS DEL SUBCORPUS

La muestra seleccionada para este artículo contiene tres docenas de documentos de distintas tipologías relacionadas con la transmisión de bienes con motivo del matrimonio. Entre todos ellos suman 45 000 tokens, una muestra significativa que proporciona, como veremos, un conjunto de análisis de sumo interés.

El esquema de metadatos de CORDICan permite una selección asequible de la documentación a partir del *query builder* de TEITOK. En el campo de "Tema" se han identificado, a partir de un listado controlado de términos, aquellos descriptores significativos en relación con la temática del documento. A partir de una búsqueda sencilla se puede, por tanto, recuperar la documentación relacionada con ese tema.




Figura 2. Distribución geográfica de la documentación en la isla de Tenerife

Además, la geolocalización que incorpora TEITOK permite visualizar la distribución geográfica de esos documentos. En ella se observa cómo la mayor parte de estos registros proceden de la isla de Tenerife, y más concretamente de su antigua capital, San Cristóbal de La Laguna.

## 2.2. CARTAS DE DOTE Y ARRAS EN LA DOCUMENTACIÓN CANARIA

Aunque genéricamente la *carta de dote* o *escritura de dote* es la denominación habitual de esta tipología documental, debe distinguirse, junto a la escritura de dote o escritura de capitulaciones matrimoniales, ya sea paterna, materna o mixta, la *carta de pago de dote* o *finiquito de dote* que, a diferencia de las *cartas de dote*, se realizaba una vez consumado el matrimonio. En este

tipo generalmente se incluyen no solo todos los bienes dotales sino la tasación individualizada de cada uno de ellos. Son sin duda estas dos categorías (las *cartas de pago de dotes* y los *finiquitos*) las que pueden proporcionar datos específicos que desvelan la realidad cotidiana y el valor añadido que supuso este tipo de patrimonio aportado por la mujer en la constitución de los nuevos núcleos familiares, con datos singulares relativos a toda la escala social presente en el archipiélago en comparación con los bienes dotales recibidos en otros espacios del dominio hispánico.

Los *inventarios* o *relaciones de bienes* que acompañan a esta modalidad textual ayudan a contextualizar este análisis ya que, por lo común, tienen como finalidad establecer el listado previo de posesiones de las que disponía el contrayente y que aportaba al matrimonio, lo que ofrece un léxico muy rico en detalles, además de informaciones de gran interés para los estudios históricos. Por lo general, es el varón el que suele disponer de bienes patrimoniales anteriores a la firma del contrato matrimonial, pero también se encuentra documentación en la que es el mismo cónyuge quien decide detallar la relación de bienes que traía consigo la esposa al nuevo vínculo establecido, como acaeció con el alcalde Alonso Suárez Carreño al contraer matrimonio con doña Leonor Peraza de Ayala, procedente de uno de los linajes más importantes y antiguos del archipiélago[3]. La documentación muestra, además, algunos *memoriales de bienes* que se asemejan a un *inventario de dote*, como el llevado a cabo por María de Maya a favor de Martín Pérez por el enlace con su hija Jacomyna Martín[4]. Otra tipología que complementa el análisis es la de los *testamentos*, en los que suelen aparecer de manera detallada las dotes y arras

3. Es el caso del inventario del mercader Pedro Guillet. 1610, abril, 25. San Cristóbal de La Laguna. AHPTF, PN 689, ff. 181r-184v. Salvo que se especifique otra fuente bibliográfica, toda la documentación manuscrita citada procede de *CORDICan.*

4. 1585, abril, 22. Garachico. AHPTF, PN 1434, ff. 329r-331v. Este inventario incluía "casas terreras" y tributos por tierras y "arvejas".

satisfechas por los contrayentes, como en el testamento de Catalina de las Cuevas o en el de Lázaro Hernández[5], en este caso referente a la dote aportada por su mujer Juana Joven. En otra escala están los testamentos de la población indígena, como el otorgado por Diego Díaz, guanche, marido de Luisa de Vera, hija de Francisco de Tacoronte, en el que se explicita: "mando se cumpla y pague a mi muger Luisa de Vera todas las cabrillas y tierras y otras cosas que reçibi con ella en dote y casamiento"[6].

En cuanto a la tipología específica que nos interesa en este análisis, hay que indicar que las cartas de dote presentes en los protocolos notariales canarios son relativamente abundantes en el período escogido. Limitando el análisis a las que se han transcrito en *CORDICan*, se puede afirmar que, en líneas generales, todas siguen un esquema similar: presencia del progenitor y del cónyuge ante el escribano para fijar la cuantía de la dote y ausencia significativa de la comparecencia de la mujer. Se trata de una situación generalizada en todos los grupos sociales. Así se advierte en el caso de Juan Martín Gallegos, vecino de la isla de La Palma, que ofrece en dote a Gómez Suárez de Toledo, natural de Sevilla y estante en la isla, por el matrimonio que "espera hazer celebrar por palabras de presente" con su hija María de la Roche, la cantidad de mil trescientas doblas de oro, en diferentes partidas, entre las que se encuentran tierras, cuevas y un "tanque" en Aguatavar por valor de quinientas doblas; doscientas doblas en joyas de oro y plata, vestidos y ajuar y seiscientas doblas en dineros de contado; además, se compromete a abonar todos estos bienes en dos años, con la posibilidad de permutar alguna de las cantidades por un tributo sobre una casa en la capital insular, sin olvidar que la alimentación y residencia de los nuevos

5. 1626, octubre, 28. La Orotava. AMC/ACFA, 157015, leg. 31.

6. 1518, agosto, 29. San Cristóbal de La Laguna. AHPTF, PN 9, ff. 860r-861v (*vid.* Tabares de Nava y Marín y Santana Rodríguez, 2014: 67-69).

contrayentes correría a su cargo hasta que se hiciera efectivo el pago de la dote[7].

Al tratarse, en esencia, de un contrato entre familias, la cuantía de estas donaciones siempre estaba en estrecha relación con la categoría social de los contrayentes, aspecto registrado también en los metadatos de *CORDICan*. El adelantado de Canarias, por ejemplo, dada su posición social, cuando contrajo matrimonio con Juana de Masieres recibió una dote que ascendía a dos mil doscientos ducados de oro (825 000 maravedís) en dinero, joyas y preseas, y aportó en arras la misma cantidad[8]. Más reducida fue, sin embargo, la dote de Isabel Suárez, que incluía algunos enseres de casa como una "talla" y que se valoró en torno a sesenta y nueve doblas, unos 34 600 maravedís[9].

Si bien todas las tipologías documentales que se han registrado siguen una estructura predeterminada, característica de la documentación notarial (López Mora y García Aguiar, 2014), un análisis minucioso de los textos transcritos desvela algunas variantes que inciden en que, junto al cumplimiento estricto del poder contractual, se intuye a veces algún aspecto personal o una diferencia social que trata de justificarse cuando se acude ante el escribano. También puede ocurrir que sea el mismo amanuense el que incluya en el documento una digresión más amplia como alegato. Este hecho puede observarse, por ejemplo, en la escritura de arras que Martín López, "calderero" y vecino de Tenerife, otorga a su esposa Catalina Fernández[10] por cuantía de 50 000 maravedís

7. 1553, octubre, 21. Santa Cruz de La Palma. Archivo General de La Palma. Protocolos de Domingo Pérez, C. 2, ff. 572v-573r (*vid.* Hernández Martín, 1999: 204-205).

8. 1519, septiembre, 7. Barcelona. AHPTF, PN 387, f. 155r. La escritura se encuentra en el referido Protocolo en el que el adelantado dice lo que recibió en dote de su matrimonio con Juana Masieres y lo hizo ante el escribano Antón de Vallejo el 27 de junio del año 1514 (*vid.* Lobo Cabrera, 1979: 373-374).

9. 1544, septiembre, 24. San Cristóbal de La Laguna. AHPTF, PN 766, minuta 11.

10. 1510, mayo, 29. La Laguna. AHPTF, PN 374, d. 209.

de la moneda de estas islas. La diferencia de estamento social entre los contrayentes llevó al escribano a anotar y justificar este hecho del que era consciente el contrayente: "la dicha mi muger soys noble e de buen linaje". Se trata de algo poco habitual, pero constituye una muestra fehaciente de las singularidades de las sociedades de frontera en lo relativo a los contratos matrimoniales entre personas procedentes de grupos sociales dispares.

En la documentación se advierte asimismo la diferencia estamental entre la población indígena y el resto de los avecindados durante los primeros años de colonización, fruto del proceso de aculturación posterior a la conquista (Baucells Mesa, 2013). Aunque no se han conservado cartas de dote consignadas por indígenas, a través de los testamentos se deduce que la cuantía de los bienes dotales se circunscribía en la mayoría de los casos a la entrega de ganados, tal como se muestra en el testamento de Alonso de Salazar, guanche, marido de Sancha Verde: "item confieso que al tiempo que case con Sancha Verde [...] mi mujer trajo una burra preñada que pario"[11]. Si la aportación era monetaria, la cuantía no solía superar las cuarenta doblas, algunas tierras y ajuar. A veces, incluso se explicita que la mujer no contribuía con dote alguna, como en el testamento de Francisca González, guancha, esposa de Gonzalo de Ibaute: "por quanto yo me case y contraje matrimonio con el dicho Gonzalo de Ibaute puede haber tres años, poco mas o menos, y a el tiempo de dicho casamiento yo no traje ni lleve conmigo ningunos bienes ni los tengo[12]. A medida que avanzaba la centuria los bienes entre la población indígena se diversificaron y, si bien la entrega de ganadería continuaba siendo prioritaria, ya se advierte la aportación de solares o la tasación de ajuar y enseres o dinero en efectivo, como se aprecia

11. 1529, agosto, 21. San Cristóbal de La Laguna. AHPTF, PN 369, ff. 869r-871r (*vid.* Tabares de Nava y Marín y Santana Rodríguez, 2014: 117-118).

12. 1529, enero, 22. San Cristóbal de La Laguna. AHPTF, PN 197, d. 347 (*Ibidem,* 2014: 116-117).

en el testamento de Francisca de Júsar, esposa de Miguel de las Casas, que llevó en dote "100 doblas en dinero y en ropas de vestir y en ajuar"[13], o en la descripción detallada de la dote aportada por Ana Vizcaína en su matrimonio con Juan Prieto: tierras, casas, mantas, arcas, almohadas, chapines, etc.[14]

### 2.2.1. Cartas de dote otorgadas por mujeres

El empleo de metadatos con la marcación de género permite afinar la búsqueda en *CORDICan* y observar que, aunque las *cartas de dote* generalmente eran otorgadas por varones, en ocasiones y de manera excepcional se encuentran referencias de mujeres como otorgantes. En la mayoría de estos casos la presencia femenina ante el escribano estaba justificada al tratarse de viudas o huérfanas, por lo que eran ellas mismas las que tenían que acudir al notario para declarar y hacer relación de los bienes que llevaron al matrimonio. Si se comparan dos de estas escrituras de dote realizadas por mujeres a mediados del siglo XVI y mediados del XVII, utilizando el criterio de búsqueda múltiple en la base de datos, se puede advertir que no difieren sustancialmente en sus características de las otorgadas por los hombres, salvo en la cuantía de los bienes que, como se ha indicado, no estaba en relación con el género del firmante sino con el estatus de las otorgantes. Son los casos, a modo de ejemplo, de la escritura de dote otorgada por Juana Suárez, viuda de Antón González, a favor de su hija María Suárez, ante las futuras nupcias que contraería con el licenciado Cristóbal Suárez[15]: en este documento se obliga a dar en dote unas tierras atributadas, la sexta parte de unas casas en La Laguna y ciento veinte doblas en ajuar y preseas de casa, junto a un novillo y varias fanegas de trigo. En el testimonio llama la atención

13. 1552, agosto, 17. San Cristóbal de La Laguna. AHPTF, PN 3374, s/f. (*Ibidem*, 2014: 141-142).

14. 1550, junio, 27. San Cristóbal de La Laguna. AHPTF, PN 634, ff. 397r-400r (*Ibidem*, 2014: 134-137).

15. 1665, octubre, 25. San Cristóbal de La Laguna. AHPTF, PN 711, ff. 181v-183v.

que la propia escritura indica expresamente que esta donación se efectuará siempre que su hijo apruebe la demasía concedida, lo que demuestra que, aunque la mujer, viuda en este caso, pudiera disponer de los bienes, la figura masculina continuaba estando presente directa o indirectamente. Bien es verdad que en esta escritura la anuencia explícita del hijo probablemente se hacía constar para evitar litigios posteriores por reclamaciones de herencia.

Si bien da la impresión de que en las cartas de dote analizadas priman las tierras junto al obligatorio "ajuar de casa y preseas", y las cantidades donadas no parecen relativamente importantes, lo que demuestra que una parte de la sociedad canaria del XVI estaba constituida por labradores o pequeños propietarios, también es posible encontrar otras cartas de dote otorgadas por mujeres a través de las cuales queda de manifiesto el poder y autoridad que ostentaban porque pertenecían al sector privilegiado de la sociedad. En esta situación se encontraba Águeda de Monteverde (Viña Brito, 2002), que otorgó una escritura por el matrimonio de su hija Ana de Monteverde con don Diego de Ayala, segundo conde de La Gomera, en el año 1557[16]. La dote que incluía, entre otros, el "yngenio de Tesacorte", "tierraz de sequero" y "medio tanque de agua", ascendió, en su conjunto, a unas dieciséis mil doblas, una suma relativamente elevada que pone de manifiesto que los contratos matrimoniales constituían la principal partida de gasto familiar, además de conformar, como bien ha señalado la historiografía desde hace décadas, el principal escenario en la transmisión de bienes en el Antiguo Régimen (Gacto Fernández, 1987). Si bien la cuantía principal de la dote provenía del fondo patrimonial, la propia Águeda de Monteverde añadió dos mil doblas y la abuela de la contrayente, Águeda de Socarrás, otras mil doblas más, que le había dejado por manda testamentaria. Este matrimonio se concertó en parte aprovechando las dificultades financieras del

16. 1557, julio, 13. Santa Cruz de La Palma. BMSC. Fondo Casa Fuerte, sign. 1226, leg. 3, nº 6, ff. 1r-22v.

conde de La Gomera y, aunque pudiera parecer que el beneficio era para los Ayala, a la larga fue clave para la intervención directa de los Monteverde en las islas de señorío, con lo que se demuestra una vez más los vínculos que se configuraban a partir de los esponsales. Entre los bienes dotales de doña Ana de Monteverde se encontraba una décima parte del ingenio de Tazacorte, así como viñas en La Breña (también en la isla de La Palma) y tierras de cereal. En la escritura se fijaba, incluso, que el matrimonio se realizaría en la Navidad de aquel mismo año de 1557. Por su parte, Diego de Ayala "de vos la dicha Águeda de Monteverde el dicho prometimiento de dote e reçibo e entrega" la cantidad reseñada de las partidas más cuatrocientas doblas para el cumplimiento total de las dieciséis mil doblas en que fue apreciada la dote. Una situación similar se intuye en las cartas de dote de las otras dos hijas de doña Águeda de Monteverde, la de doña Margarita con el hermano del conde de La Gomera, don Melchor, tasada en torno a las quince mil doblas en el año 1562, y la de su hija menor doña Águeda con Pedro de Liaño en 1583. Se puede trazar de esta manera la historia documental de estos linajes que, con sus matrimonios, aseguraron su permanencia en la elite isleña.

Un caso algo infrecuente, y por ello difícil de localizar en la documentación notarial, es la escritura de dote otorgada por una huérfana. Fue lo que le sucedió a Magdalena Francisco que, tras demostrar ante el escribano que sus progenitores habían fallecido, realizó personalmente la carta de dote "por quanto al servicio de dios nuestro señor y para honra y gloria suya esta tratado y consertado qasar lexitimamente en has de la santa yglesia de Roma con Blas de la Carrera"[17]. Llama la atención que uno de los testigos presentes, sin especificar si le une algún tipo de parentesco con la otorgante, dona asimismo cuatro pipas de vino de malvasía de la viña que posee en La Orotava y, "por haserle buena

17. 1666, marzo, 28. San Cristóbal de La Laguna. AHPTF, PN 711, ff. 13r-15v.

hobra y caridad", se compromete a abonar el alquiler de la casa en que viva por tiempo de seis años. Por su parte, Blas de la Carrera, tras la aceptación de la escritura, "promete y se obliga a casarse lexitimamente despues de haver presedido las tres amonestaciones que el Concilio de Trento manda", aportando en arras al matrimonio cien ducados de once reales cada uno. En este caso, y como solía ser habitual, los otorgantes no firmaron por no saber escribir y lo hizo en su nombre un testigo, lo que constituye una evidencia fehaciente del analfabetismo del conjunto de la sociedad.

No tenían un estatus especial las dotes establecidas para el ingreso de las monjas en los conventos. En la documentación aparecen algunas licencias que permitían a las novicias realizar este tipo de escrituras, como la concedida por el provincial de San Francisco y la abadesa del convento de Santa Clara para que se dictara escritura de dote, según lo estipulado por el convento, de Anna María de San Joaquín Suárez, hermana de la citada comunidad[18]. El provincial concede autorización a la abadesa para que "pueda dar y de la profession para religiosa de velo negro según forma canónica y regular". De esta manera entraría a profesar con sus bienes dotales correspondientes.

### 2.2.2. Cartas de dote otorgadas por parientes

En otras ocasiones, las cartas de dote no fueron estipuladas por los progenitores, padre o madre, sino que también resulta frecuente encontrar algunas donaciones de parientes. Juan de Tres Puentes, por ejemplo, dotó a su hermana María Vélez[19] por encontrarse sus padres ausentes como vecinos de la villa de Medinaceli. La aportación patrimonial de Juan de Tres Puentes al matrimonio de su hermana consistía en unas casas en la ciudad de Las Palmas cuyo valor no superaba los quinientos sueldos, que era lo establecido por ley. Se

18. 1641, agosto, 8. San Cristóbal de La Laguna. AHPTF, PN 733, f. 50r.
19. 1522, febrero, 28. Las Palmas de Gran Canaria. AHPLP, PN 735, ff. 257r-258r.

comprometía asimismo a declarar este hecho en su testamento con el fin de evitar cualquier reclamación al respecto.

Cuando la donación la realizaban algunos parientes cercanos a la futura contrayente, aparecen justificados los motivos que llevaron a contribuir con la dote. En la carta de finiquito de Andrés Pérez, labrador, tras su matrimonio con María González, se especifica que su esposa "a servido a Gonsalo Peres y a su ermana... y por el servicio y deudo y amor y buena voluntad" este le ha concedido en dote, en varias partidas, ajuar, mobiliario y ganado valorado en 36 126 maravedís[20].

Otras veces fueron familiares colaterales, generalmente tíos, los que contribuyeron con algunos bienes a la mejora de la aportación llevada por la mujer al matrimonio. El racionero de la catedral de Las Palmas, Bartolomé López, legó parte de su patrimonio a su sobrina Juana Alemán al contraer matrimonio con Joan Narciso de Lissaga, catalán[21]. La donación del racionero fue conjunta con otros parientes de Telde, su tío Sebastián Marques y la esposa de este, Melchora de los Reyes. Cada uno detalló las cantidades aportadas que, por parte del racionero, consistió en trescientos ducados de once reales cada uno, que haría efectivos en dos plazos, por Navidad y por el mes de agosto; por su parte, los otros tíos, avecindados en Telde, coadyuvaron con unas casas bajas tejadas tasadas en ciento cincuenta doblas, más unas tierras en su ciudad apreciadas en cincuenta doblas y el ajuar de casa. Afortunadamente, en este caso es posible seguir la trayectoria de la dote de Juana Alemán a través de varios recibos otorgados por su marido Joan de Lissaga pocos días después de consumado el matrimonio, como el efectuado el 15 de octubre de 1610[22], a favor

20. 1589, marzo, 5. San Cristóbal de La Laguna. AHPTF, PN 101, ff. 85v-87r.

21. 1610, septiembre, 25. Las Palmas de Gran Canaria. AHPLP, Fondo Sall y Tascón, ff. 68r-71v.

22. 1610, octubre, 15. Las Palmas de Gran Canaria. AHPLP, Fondo Sall y Tascón, ff. 66r-68v.

del canónigo don Bartolomé López, del ajuar con que había dotado a su sobrina.

Pocos años más tarde, en 1637, se tiene constancia de otro recibo de la dote prometida al alférez don Francisco Betancor Cabrera por su matrimonio con doña Catalina de Alemán y Lissaga, "con quien es casado legítimamente". Este nuevo finiquito corresponde al aprecio de los bienes muebles, en el que se incluyen una casa, tierras ("suerte y media de tierra calma... en el tabaybal desta ciudad"), un esclavo negro, dinero de contado, ajuar, mobiliario y otros utensilios como "un harnero en dos reales una bazinilla en ocho reales un espejo en nueve reales...", por importe de 11 176 reales. A todo ello se añadía la cantidad que el canónigo Bartolomé López le entregó en presencia del escribano, esto es, "quatro mil y quatrocientos reales en reales de a ocho y de a quatro en tostones y reales sencillos"[23].

### 2.2.3. Otra documentación relacionada con las dotes

En *CORDICan* se localizan otros conjuntos de tipologías documentales vinculados con esta temática, aunque no específicos, pero que ofrecen también importante información para los estudios interdisciplinares, como los *inventarios* y *reclamaciones*. No obstante, algunos de los manuscritos ya citados resultan relativamente abundantes e interesantes ya que, otorgados habitualmente por el marido una vez consumado el matrimonio, en ellos se detalla la tasación individualizada de los distintos bienes. Así, en los *recibos de dote* disponibles en el corpus, el elemento primordial es el ajuar de casa, al que se unen pequeñas cantidades monetarias de contado, junto a algún ganado. Además, en aquellos casos en los que el documento se elabora en un entorno rural, es frecuente que las cartas de dote añadan la entrega de tierras. Así ocurrió, por ejemplo, con el recibo de dote de Francisco Pérez a sus suegros Luis Hernández y

23. 1637, agosto, 1. Las Palmas de Gran Canaria. AHPLP, Fondo Sall y Tascón, ff. 67r-70v.

Juana Magdalena, tras su matrimonio con Isabel Rodríguez, en el que se da cuenta de que su esposa recibió setenta doblas en ajuar (que equivalían a 50 476 maravedís), en el que se incluía "una casa grande con su serradura y su baja de vinatico", aparte de setenta doblas en dineros de contado y una dotación de tierras[24].

Si las escrituras y recibos de dote son tipos documentales de primer nivel por la detallada información que proporcionan, también los *inventarios* constituyen una fuente de extraordinario interés. Generalmente estas relaciones se realizan a petición del futuro contrayente para que quede constancia de los bienes que aporta al matrimonio: "porque en todo tiempo esten y parescan los vienes e caudal queste otorgante lleva por su capital y propios suyos", señalaba el mercader Pedro Guillet[25] al desposarse con Catalina Vázquez de Moya. En esa relación de posesiones quedaban reflejadas, entre otros, las casas de su morada valoradas en ocho mil reales, así como una ingente cantidad de textiles, dada su condición de mercader floreciente que importaba este tipo de mercancías desde varios puntos de Europa. La relación permite conocer en detalle los diferentes tipos de telas importadas: "dies y ocho varas de fustan negro en treinta y seis reales çiete varas de anascote a ocho reales la vara quatro piesas de bocaran a treinta reales veinte y çiete varas de toquilla negra en çien reales, etc."[26], así como el ajuar de casa, mobiliario, esclavos y un caballo con su silla y freno, indispensable para los desplazamientos por la isla y poder ejercer su oficio de comerciante. El inventario, según señala el escribano Juan de Pineda, se realizó y se redactó en la tienda y casa del mercader, que le mostró un libro donde se detallaban las cuentas de sus acreedores, así como las minutas que tenía pendientes de pago. Como se puede comprobar, tanto el léxico como la propia

24. 1591, abril, 10. San Cristóbal de La Laguna. AHPTF, PN 1226, ff. 274r-277r.

25. 1610, abril, 25. San Cristóbal de La Laguna. AHPTF, PN 689, ff. 181r-184v, ya citado.

26. *Ibidem*, f. 181v.

información contenida aportan datos de extraordinario valor para los estudios comparados históricos, sociales y lingüísticos.

No siempre los inventarios que se realizaban a petición del futuro marido eran redactados para dejar constancia de las posesiones previas, sino que a veces el objetivo era consignar los fondos vinculados a la futura esposa. En el inventario de la dote de doña Leonor Peraza de Ayala, por ejemplo, se señala que "trae cantidad de bienes de los quales quiere para su seguridad y lo que mas le conviene que coste en todo tiempo", frente al patrimonio de su futuro marido el alcalde mayor de Daute, Alonso Suárez Carrillo[27]. En este caso, la categoría de la contrayente requería el inventario previo en el que se explicitaba el traspaso de tierras de viña y de sembradura, solares, casas, tributos, vacas, becerros, asnos, etc., aparte de un oficio de regimiento, cinco esclavos negros, así como el ajuar de casa en que se incluían textiles, menaje, algunos elementos decorativos y varias piezas de oro y plata.

El estado civil de la contrayente determinaba asimismo la realización de un inventario preliminar a las nupcias. Es el caso de las viudas con hijos de un primer matrimonio que, a fin de salvaguardar el patrimonio de estos y evitar litigios posteriores, dejaban constancia ante el escribano de la relación de bienes procedentes de su anterior enlace. En este supuesto se encontraba Catalina Fernández, viuda, cuando concertó su nuevo matrimonio con Afonso, portugués, por lo que declaró por escrito que poseía tierras en el "malpays de Ycode", casas, ganados y una esclava de su propiedad[28].

Ya se ha indicado que, además de las tipologías ante-dichas, algunos testamentos añadían determinadas cláusulas referentes a la dote y arras alegadas. En el testamento de Lázaro Hernández[29], mercader, se incluye una detallada relación de lo aportado por su mujer Juana Joven, aunque en este caso el objetivo primordial de tal

27. 1585, abril, 22. Garachico. AHPTF, PN 1434, ff. 329r-331v, ya citado.
28. 1535, octubre, 8. Malpaís de Icod. AHPTF, PN 3386, s/f.
29. 1608, marzo, 16. San Cristóbal de La Laguna. AHPTF, PN 1176, ff. 414r-418v.

especificación era la reclamación de una parte de la dote. En concreto, se trataba de las mejoras realizadas por dos de sus cuñados, que eran frailes de la orden de San Agustín, a los que reclamaba una viña en Icod de los Vinos (al norte de Tenerife), el valor de un esclavo y un tributo de diez fanegas de trigo. Al mismo tiempo, solicitaba la parte de la herencia que le correspondía del padre de su esposa, Juan de Carminatis, argumentando que el licenciado Francisco de Carminatis y Juan de Carminatis, su hermano difunto, fueron fiadores de la dote que se prometió a Juana Joven. Según declaraba, aún no habían hecho efectiva esta carta dotal que ascendía a cuatro mil doblas de quinientos maravedís cada una de la moneda de estas islas. El citado mercader había aportado en arras al matrimonio quinientas doblas, sin duda una cantidad también bastante relevante. En el testamento declaraba que, por cédula real, fue reconocido como hijo legítimo de maese Domingos Hernández y los bienes que poseía eran fruto de las actividades comerciales con la península ibérica e Indias.

En este sentido, resulta asimismo de interés otra dote inserta en el testamento de doña Catalina de las Cuevas[30], donde deja constancia del conjunto de bienes que llevó al matrimonio por no haberse levantado escritura de dote cuando contrajo matrimonio con Francisco Suárez de Lugo. Las cláusulas detallaban los fondos aportados por este concepto, como seis esclavos y un inmenso ajuar compuesto por bienes muebles, menaje de casa, ropa de vestir y algunas joyas, junto a las tercias de los frutos de Adeje y Garachico procedentes de los mayorazgos de sus hermanos Nicoloso y Alonso de Ponte. Además, se señalaba que, según el testamento de su padre, hasta que no recibiera los doce mil ducados prometidos en dote, podría gozar de una serie de rentas para su manutención. En otra de las cláusulas del testamento reclamaba la "herensia que me dexo mi hermana doña Ygnes Benitez marquesa de Lanzarote, assi su dote, harras e multiplicados", pertenencias que sumaban

30. 1626, octubre, 28. La Orotava. AMC/ACFA, 157015, leg. 31, ya citado. La copia del testamento ante el escribano Nicolás de Curras se realizó el 7 de diciembre de 1761.

dieciséis mil ducados, aparte de los acrecentamientos y mejoras. En este caso, al fallecer su hermana sin descendencia, la dote revertía de nuevo en la unidad familiar, por lo que reclamaba la parte correspondiente por consanguinidad.

Es precisamente a través de la documentación generada por estas grandes familias donde es posible rastrear el conjunto de dotes otorgadas a sus herederos. En la cláusula diecisiete del testamento de Pedro Ponte[31] se puede apreciar la dote que legó a sus hijos: a Inés Benítez de las Cuevas por su matrimonio con el conde de Lanzarote, ocho mil ducados; a María de las Cuevas por su matrimonio con Andrés Suárez Gallinato de Lugo, seis mil trescientos ducados; a Francisca de Ponte por su matrimonio con Bartolomé de Ponte, su sobrino, seis mil ducados y otros bienes. Fundó, además, mayorazgo para sus dos hijos Nicoloso de Ponte y Alonso de Ponte, mientras que dejó sin especificar las cantidades fijadas para las hijas que aún eran doncellas, Isabel, Ana y Catalina.

Finalmente, una última tipología de documentos relacionados con la dote disponible en *CORDICan* es la compuesta por las reclamaciones de impagos. Eran también asientos frecuentes en el contexto de la conflictividad familiar bajomedieval y moderna (González Zalacain, 2013: 235-253), como se aprecia, entre otros, en el testamento de Pedro Santana, guanche, marido de Juana Martín, que reclamaba a Rodrigo Alonso, vecino de La Palma, veinte doblas de resto de cien doblas "que fueron adjudicadas en mi mujer en dote"[32]. Un caso singular aparece en el testamento de Juan de Simancas, herreño y vecino de Acentejo, marido de María Hernández, quien declaró que "los bienes que tiene son todos de su mujer que los trajo en dote, parte de una viña y unas tierras en Sentejo"[33].

31. 1568, septiembre, 1. Heredamiento de Adeje. AMA/ACFA, 157003, nº 23.

32. 1535, febrero, 22. AHPTF, PN 621, ff. 543r-544r (*vid.* Tabares de Nava y Marín y Santana Rodríguez, 2014: 144-145).

33. 1550, febrero, 7. San Cristóbal de La Laguna. AHPTF, PN 2041, ff. 266r-267r (*Ibidem*, 2014: 330-331).

El montante de las dotes establecidas en los testamentos no difería de las demás tipologías, pues todo dependía de la posición que ocupaban los beneficiarios en la escala social y, entre los miembros de la misma familia, las prebendas que recibían las mujeres solían ser de una cuantía similar. Isabel González, viuda de Juan Bello a principios del siglo XVII[34], por ejemplo, legaba a su hija María Bello trescientas cincuenta doblas, y para su otra hija Ana Bello estipulaba una cantidad de cuatrocientas doblas.

En cuanto a las reclamaciones, algunas mujeres acudían al escribano para solicitar el reintegro de su dote. Los motivos que las forzaron a ello eran muy variados. Así, Catalina Sánchez, mujer de Bartolomé González, se presentó ante el gobernador y justicia mayor de Tenerife para pedir licencia a la autoridad con la finalidad de poder disponer de sus bienes dotales "para su reparo y el de sus hijos", ya que su marido estaba desterrado. En este caso confirmó su reclamación con dos testigos y, pocos días después, el gobernador accedió a la solicitud para que "la dicha Cathalina Sanches pueda pareser en juizio e pedir e demandar y cobrar qualesquier bienes que pertenescan al dicho Batholome"[35].

En otras ocasiones, la mujer acudía al notario para hacer valer sus derechos ante una situación de violencia doméstica. Clara de Vitoria, por ejemplo, en 1612, se presentó ante el escribano Marcos Gutiérrez porque temía la reacción de su marido, Juan Gutiérrez de Albornoz, al denegarle la autorización para poder vender unas tierras en Tijoco que pertenecían a sus bienes dotales. Argumentaba que "la a de forsar y obligar con justos temores y fuersas y malos tratami*ent*os y amenasas obradas en su persona", añadiendo que recibía "malos tratos que le hase que no tienen parada"[36]. En esta

34. 1600, junio, 21. Granadilla de Vilaflor. AHPTF, PN 1443, ff. 470r-473v. Poseía "treinta colmenas e nueve cofres de drago", así como una "casa pajiza".

35. 1567, abril, 7. La Laguna. AHPTF, PN 650, ff. 176r-176v. La solicitud se presentó el 7 de abril de 1567 y la licencia le fue concedida el 9 de abril del mismo año.

36. 1612, noviembre, 5. La Laguna. AHPTF, PN 263, ff. 473r-473v.

ocasión concreta, el problema se planteaba tanto por la violencia de género recibida como por el incumplimiento ante la venta de esas tierras por parte de su marido, lo que supondría una considerable merma para sus descendientes.

En conjunto, la documentación disponible acredita el interés de la creación de subcorpus dentro de *CORDICan*. La selección temática de un conjunto de documentos del tipo de los aquí comentados ofrece amplias posibilidades de estudio en el campo de la historia social, a la vez que aporta material de primer nivel para los estudios del léxico específico acopiado en los textos, en muchas ocasiones de marcado carácter geolectal.

## 3. CONCLUSIONES

*CORDICan* se plantea como objetivo la puesta en marcha de un recurso digital para el estudio y la difusión de documentación canaria durante el Antiguo Régimen (siglos XV-XVIII) y la Edad Moderna, que aporte información de interés para diferentes disciplinas y subdisciplinas del conocimiento histórico, filológico y cultural. Al conformarse como un recurso de libre acceso va a permitir ofrecer a los usuarios datos textuales y extratextuales relevantes para este tipo de análisis.

El análisis de caso de las posibilidades derivadas del estudio de un subcorpus, el de *dotes y arras*, constata todos estos objetivos planteados. El vocabulario reflejado, abundante en detalles relacionados con la vida cotidiana, permitirá llevar a cabo investigaciones léxicas y semánticas, con el aporte de primeras dataciones y testimonios inéditos para la lexicografía histórica. Además, la posibilidad que ofrece TEITOK de realizar de manera relativamente sencilla la marcación morfosintáctica de los textos, no prevista inicialmente en esta etapa de desarrollo de *CORDICan*, abre la puerta a su desarrollo posterior por parte de historiadores de la lengua interesados en el estudio del llamado español meridional a través de una documentación transcrita literalmente y con variados formatos de presentación.

A partir del análisis de los metadatos y del contenido de la treintena de ejemplos disponibles hemos podido comprobar, a su vez, una serie de cuestiones destacables. Por lo general, esta documentación fue otorgada mayoritariamente por hombres, aunque hay casos excepcionales en los que es una mujer la que acude al notario para dejar constancia de sus bienes dotales. En el corpus, solo uno de cada seis documentos fue redactado por iniciativa femenina; en la mayoría de las ocasiones, fueron los maridos o futuros cónyuges o parientes los que otorgaron las escrituras.

Un aspecto singular en las dotes redactadas en el archipiélago es la relevancia de la población foránea, con unas políticas matrimoniales perfectamente definidas, tanto en el caso de los flamencos como de los portugueses. Canarias era en aquellos momentos una sociedad de frontera y las alianzas por matrimonio aseguraban la consolidación de determinadas redes familiares (Viña Brito, 2020). Esos lazos permiten trazar la transmisión de los bienes a partir de unas alianzas y unas estrategias planificadas por esas familias, de tal manera que apellidos como Vandevalle, Monteverde, Ponte o Lugo aparecen con relativa frecuencia en la documentación.

Las cartas de dote analizadas presentan una estructura muy similar, solo varían las aportaciones en función de la escala social de cada otorgante y del grado de parentesco que poseen con la dotada. Ello es debido al carácter universal del ajuar en la dote y al papel que desempeñó el matrimonio en la etapa inicial de conformación de las sociedades isleñas. En este período, las teorizaciones, la legislación y la realidad cotidiana constituían tres aspectos diferentes, aunque complementarios e imprescindibles (Monzón Perdomo y Viña Brito, 2006). Este conjunto de dotes ofrece, además, un acercamiento a la mujer como sujeto histórico y social (Segura Graíño, 2013: 40), una línea en continuo proceso de revisión que, en aquellos momentos, apenas quedaba visible en el resto de la documentación manuscrita.

## REFERENCIAS BIBLIOGRÁFICAS


Baucells Mesa, Sergio (2013). *Aculturación y etnicidad. El proceso de interacción entre guanches y europeos (siglo XIV-XVI)*. San Cristóbal de La Laguna: Instituto de Estudios Canarios.

Borrero Fernández, Mercedes (1984). El ajuar de la casa campesina sevillana a fines de la Edad Media. En *La sociedad medieval andaluza, grupos no privilegiados: Actas del III Coloquio de Historia de Andalucía*. Jaén: Diputación de Jaén, 211-226.

Claridge, Claudia (2008). Historical corpora. En Anke Ludeling, Merja Kyto (eds.), *Corpus Linguistics: An International Handbook*. Berlín / NuevaYork: Walter de Gruyter, 242-259.

Corrales, Cristóbal; Corbella, Dolores y Viña Brito, Ana (2014). *Léxico azucarero atlántico*. San Millán de la Cogolla (Logroño): Cilengua.

Gacto Fernández, Enrique (1987). El grupo familiar de la Edad Moderna en los territorios del Mediterráneo hispánico: una visión jurídica. En *La familia en la España mediterránea*. Barcelona: Crítica, 36-64.

González Zalacain, Roberto J. (2005). *Familia y sociedad en Tenerife a raíz de la conquista*. San Cristóbal de La Laguna: Instituto de Estudios Canarios.

González Zalacain, Roberto J. (2013). *La familia en Castilla en la Baja Edad Media: violencia y conflicto*. Madrid: Congreso de los Diputados.

Hernández Martín, Luis Agustín (1999). *Protocolos de Domingo Pérez, escribano público de La Palma (1546-1553)*. Santa Cruz de La Palma: Caja General de Ahorros de Canarias.

Janssen, Maarten (2016). TEITOK text-faithful annotated corpora. En Nicoletta Calzolari; Khalid Choukri, Thierry Declerck y Asunción Moreno (eds.). *10th conference on International Language Resources and Evaluation (LREC'16)*. Portorož (Slovenia), 4037-4043.

Janssen, Maarten y Vaamonde, Gael (2020). Da edición dixital á análise lingüística. A creación de corpus históricos na plataforma TEITOK. En Rosario Álvarez y Ernesto González Seoane (eds.): *Calen barbas, falen cartas. A escrita en galego na Idade Moderna*. Santiago de Compostela: Consello da Cultura Galega (Ensaio & Investigación), 271-292.

Lobo Cabrera, Manuel (1979). *Protocolos de Alonso Gutiérrez (1520-1521)*. San Cristóbal de La Laguna: Instituto de Estudios Canarios y Aula de Cultura de Tenerife.

López Mora, Pilar y García Aguiar, Livia Cristina (2014). Pragmática del documento notarial. Mímesis e impostura en la tradición diplomática, *Anuario de estudios filológicos*, 37, 139-157.

Monzón Perdomo, María Eugenia y Viña Brito, Ana (2006). Ellos opinan de ellas. Las mujeres en la historiografía canaria, desde la conquista al siglo XVIII. En *Mujeres en movimiento. Historia y Literatura*, coord. Rosa Sierra del Molino. Las Palmas de Gran Canaria: ULPGC-Instituto Canario de la Mujer.

Ronquillo Rubio, Manuela (1992). El ajuar de la casa urbana en las Islas Canarias a fines de la Edad Media. *Vegueta,* 0, 37-42.

Segura Graíño, Cristina (2013). Las mujeres medievales: perspectivas historiográficas. En *Las mujeres en la Edad Media*, coord. por Isabel Val Valdivieso y Juan Francisco Jiménez Alcázar. Lorca (Murcia): SEEM-Editum, 33-54.

Tabares de Nava y Marín, Leopoldo y Santana Rodríguez, Lorenzo (2014). *Testamentos de guanches (1505-1550)*. San Cristóbal de La Laguna: Instituto de Estudios Canarios.

Vaamonde, Gael (2015). P. S. Post Scriptum: Dos corpus diacrónicos de escritura cotidiana. *Procesamiento del Lenguaje Natural*, 55, 57-64.

Viña Brito, Ana (2002). Águeda de Monteverde y la administración de un patrimonio familiar. *Revista de Historia Canaria*, 184, 341-360.

Viña Brito, Ana (2012). Los flamencos en Canarias en el siglo XVI ¿Una comunidad extranjera? Especificidades en la isla de La Palma. *Revista de Historia Canaria*, 194, 161-191.

Viña Brito, Ana (2017). Solidaridad y endogamia de los portugueses en Canarias durante el siglo XVI. *Portuguese Studies Review*, 24 (1), 3-29.

Viña Brito, A. (2020). Canarias, sociedad de frontera. En *Las fronteras en la Edad Media hispánica (siglos XIII-XVI)*, coord. por Manuel García Fernández. Sevilla: SEEM, 475-492.

Viña Brito, Ana y Corbella, Dolores (2018). Corpus documental de las islas Canarias: un nuevo reto en Humanidades Digitales. *Anuario de Estudios Atlánticos*, 65, 1-17.

Viña Brito, Ana; Corrales, Cristóbal y Corbella, Dolores (2014 y 2016). *Islas y voces del azúcar I (Tenerife, La Gomera y La Palma)* y *II (Gran Canaria)*. San Cristóbal de La Laguna - Las Palmas de Gran Canaria: Gobierno de Canarias - Archivos Históricos Provinciales de Tenerife y de Las Palmas de Gran Canaria.

# LAS GRANDES BIBLIOTECAS DE TEXTOS GRIEGOS Y LATINOS ANTE LOS RETOS DE LA NUEVA ERA DE EDICIÓN DIGITAL. LUCES Y SOMBRAS

Cecilia Criado

**Resumen:**
A comienzos de este milenio, la Filología Clásica fue consciente de que había que someter a revisión los criterios editoriales de las colecciones que ofrecían los textos griegos y latinos canónicos en acceso libre. Éstas comenzaban a acusar obsolescencia tecnológica y obvias deficiencias filológicas, tocantes sobre todo a la calidad de las ediciones. En consecuencia, se establecieron directrices consensuadas por la comunidad investigadora internacional para establecer cómo habría de ser la presencia *online* de estos textos en el futuro. Se propugnó la adopción del estándar TEI XML, el uso de los CTS URNs y, ya estrictamente en el terreno filológico, se recomendó que las grandes plataformas y proyectos abordasen el reto de crear ediciones críticas basadas en ediciones de referencia preexistentes. El objeto de este artículo es, primero, describir las líneas de actuación propuestas y, segundo, dar cuenta de cómo las sustanciaron las tres grandes bibliotecas de acceso libre, esto es, *Open Greek and Latin* (OGL), que es la última y ambiciosa versión de la *Perseus Digital Library*, la *Library of Digital Latin Texts* (LDLT) y *Musisque Deoque* (MQDQ).

**Palabras clave:** bibliotecas digitales, textos griegos y latinos, TEI XML, LOD, edición crítica

Cecilia Criado, Universidad de Santiago de Compostela
ORCID 0000-0003-2056-1086
cecilia.criado@usc.es

**Abstract**
At the beginning of this millennium, Classical Philology became aware that the editorial criteria of the collections offering canonical Greek and Latin texts in open access had to be revised. They were beginning to suffer from technological obsolescence and obvious philological deficiencies, especially concerning the quality of the editions. Consequently, consensus guidelines were established by the international research community to establish how the online presence of these texts should be in the future. The adoption of the TEI XML standard and the use of CTS URNs were advocated, and, strictly in the philological field, it was recommended that the platforms and projects address the challenge of creating critical editions based on pre-existing standard editions. The purpose of this article is, first, to describe the proposed lines of action and, second, to give an account of how the three largest open access libraries, namely *Open Greek and Latin* (OGL), which is the latest and the most ambition version of the *Perseus Digital Library*, *The Library of Digital Latin Texts* (LDLT) and *Musisque Deoque* (MQDQ), have implemented them.

**KEYWORDS:** digital libraries, Greek and Latin texts, TEI XML, LOD, critical edition

## 1. INTRODUCCIÓN

Las bibliotecas de textos griegos y latinos antiguos creadas en el último cuarto del s. XX por el *Thesaurus Linguae Graecae* (TLG, University of California, Irvine)[1], el *Packard Humanities Institute* (PHI)[2] y la *Perseus Digital Library* (PDL, Tufts University)[3] evidencian el papel pionero de la Filología Clásica en el panorama internacional de las Humanidades Digitales. Desde esas fechas hasta la actualidad, son innumerables las colecciones que se han creado. No es mi intención, sin embargo, ofrecer un panorama de la presencia *online* de los textos clásicos[4]. Mi interés se centrará en dos puntos. Primero, analizaré los efectos menos benévolos de la temprana aptitud digital de la Filología Clásica, causados, aunque no exclusivamente, por el rápido avance que experimentó el desarrollo web en los últimos años del siglo pasado. La oferta de los textos clásicos tuvo que ser sometida a un drástico replanteamiento con el fin de actualizar datos e infraestructuras. Ello se concretó en propuestas que propugnaron la revisión de las directrices tecnológicas y filológicas asumidas hasta el momento.

En segundo lugar, someteré a examen la oferta actual de bibliotecas que tienen las siguientes características: los textos están en acceso libre; tienen un tratamiento informático desarrollado, esto es, no son webs estáticas[5], sino que ofertan valores añadidos (traducciones, diccionarios, aparatos críticos,

1. http://stephanus.tlg.uci.edu/ [actualizada: 29 septiembre 2021; citada: 1 febrero 2022].
2. http://latin.packhum.org [actualizada: s.d.; citada: 1 febrero 2022].
3. http://www.perseus.tufts.edu [actualizada: 2014; citada: 1 febrero 2022].
4. Un excelente estado de la cuestión se encuentra en Dániel Kiss, 2019: 221-237.
5. Es el caso, por ejemplo, de *The Latin Library* (dir. William L. Carey. Disponible en https://www.thelatinlibrary.com/ [actualizada: s.d.; citada: 10 febrero 2022]), una precoz biblioteca puesta en funcionamiento en 1994. En su momento, cubrió de forma excepcional las necesidades de los usuarios de un Internet que todavía no poseía un ancho de banda que permitiese mayores desarrollos.

etc.); están en constante desarrollo; y, por último, son plataformas de carácter comprenhensivo, entendiendo por tales aquéllas concebidas para contener fondos ingentes de, al menos, las obras canónicas antiguas en lengua griega, latina o ambas. No serán, por tanto, consideradas las colecciones cuyo acceso está sujeto a suscripción (como sucede con TLG)[6], las que no están en desarrollo hoy día (por ejemplo, PHI)[7], las que se alimentan de las obras de una temática o de un periodo cronológico concretos (textos tardíos, textos medievales, humanistas, etc.) o los proyectos que elaboran *born-digital critical editions* de un único autor u obra[8].

Las bibliotecas que se adecuan a mi objeto de estudio son *Open Greek and Latin* (OGL), que es la última versión de la *Perseus Digital Library* (PDL), la *Library of Digital Latin Texts* (LDLT) y *Musisque Deoque* (MQDQ). A través del análisis crítico de su metodología editorial, intentaré establecer el grado de éxito alcanzado por cada una de ellas en la consecución de las metas que la comunidad internacional de Filología Clásica estableció para

6. TLG, no obstante, ofrece una *Abridged version* de acceso gratuito que es de gran utilidad para la enseñanza de las lenguas clásicas. Disponible en: http://stephanus.tlg.uci.edu/abridged.php [actualizada: 2014; citada: 1 febrero 2022]

7. Desde 2015, el *Packard Humanities Institute* permite acceso libre a las obras contenidas originariamente en las bases de datos elaboradas entre 1987 y 1997. Desde entonces, el corpus no ha sido ampliado.

8. Un avanzado a su tiempo fue Michael Hendry quien, entre 2000 y 2008, publicó en su web ***Curculio*** ediciones, anotadas con variantes textuales básicas, de las obras de Propercio, Ovidio, Juvenal, Marcial y Claudiano (Disponible en: http://curculio.org/ [actualizada: 31 julio 2022; citada: 10 septiembre 2022]. Aun a riesgo de incurrir en imperdonables omisiones y sin afán de exhaustividad, citaré como proyectos señeros de edición crítica digital de una única obra *Callimachus' Aetia* (ed. Susan Stephens, Stanford University. Disponible en: https://dcc.dickinson.edu/callimachus-aetia/the-aetia [actualizada: 20 mayo 2022; citada: 15 septiembre 2022]), *Catullus Online* (ed. Dániel Kiss, Universitat de Barcelona. Disponible en: http://www.catullusonline.org/ [actualizada: s.d.; citada: 15 septiembre 2022]) y *Digital Statius: the Achilleid* (dir. Damien Nelis, Université de Genève. Disponible en: https://achilleid.unige.ch/welcome [actualizada: s.d.; citada: 15 septiembre 2022]).

afrontar la nueva era de edición digital a comienzos del s. XXI. Como se comprobará, las revisiones metodológicas propuestas se están materializando de forma desigual en unas y otras bibliotecas.

## 2. REPLANTEAMIENTO TECNOLÓGICO Y FILOLÓGICO DEL 'COSMOS' DEL *PERSEUS PROJECT*

La primera versión electrónica de la *Perseus Digital Library* (PDL), lanzada por la Tufts University en 1998 bajo la dirección de Gregory R. Crane, supuso un hito sin precedentes en el campo de las bibliotecas digitales de cualquier ámbito. La plataforma fue concebida para posibilitar un entorno de lectura lineal y radial del vasto número de textos contenidos en el CD-Rom publicado en 1998. Fue diseñado «as a hypertext, with many cross-references complementing a clear hierarchical structure» (Crane, 2002: 8) que ofrecía innumerables valores añadidos. En esta primera versión, cada palabra ya contenía un enlace que dirigía a los analizadores morfológicos que vinculaban las formas flexionadas con las entradas de los diccionarios. Asimismo, podía visualizarse la traducción inglesa en paralelo (aunque no alineada) de las fuentes primarias y permitía enlazar con multitud de gramáticas, enciclopedias, mapas, imágenes, artículos científicos relacionados y un largo etcétera[9]. En definitiva, el objetivo era la creación de una biblioteca que ampliara la infraestructura digital con una gran masa de materiales no textuales (Lang, 2018a: 2). PDL continúa siendo la mayor y la más poderosa herramienta sobre el mundo antiguo en acceso libre, lo cual es coherente con su aspiración, ya explicitada en las fases iniciales del proyecto, de contribuir a la democratización de la herencia cultural de la Humanidad, defendiendo que el acceso a tal legado había de ser «a right, not a privilege» restringido al mundo académico (Crane, 2002: 5).

9. Para la descripción pormenorizada de las utilidades de lectura radial ofrecidas por la plataforma, *cf.* Crane (2002: 8-13) y Lang (2018: 7-8).

Con todo, tras el lanzamiento de las tres primeras versiones de la PDL, a principios de este milenio hubo que rendirse a la evidencia de que la precocidad digital del ámbito de la Filología Clásica tenía un precio. Frente a lo que sucedía en el caso de bibliotecas que contenían corpus de otras lenguas y que, por ser de factura más reciente, diseñaron sus sistemas de forma acorde con los desarrollos web del momento, la plataforma de la PDL dependía de una serie de servicios de enlace automático, de extracción de información y de visualización que los sistemas de las primeras versiones del *Perseus Project*, más orientados al catálogo que a la biblioteca propiamente dicha, estaban siendo incapaces de soportar o de garantizar su sostenibilidad a largo plazo[10]. Por ello, se optó por abandonar la codificación Perl[11], al tiempo que se adaptó el sistema de la PDL y de su creciente número de objetos con el fin de poder utilizar el repositorio FEDORA (*Flexible Extensible Digital Object and Repository Architecture*)[12] (Chen, 2016: 136 y Babeu, 2019: 61). La migración de datos desde el *Tufts University Repository* a FEDORA comenzó en 2006, año en que se concluyó el diseño de la intezfaz de esta nueva versión 4.0 de *Perseus*, denominada *Perseus Hopper*, cuyo código fuente fue puesto en abierto de forma casi inmediata[13]. No obstante, la necesidad de desarrollar FEDORA para que pudiese gestionar la inmensidad de datos contenidos en la PDL y la toma de conciencia de que era necesario construir un nuevo sistema de biblioteca digital provocaron que el proceso se dilatase en el tiempo y se produjesen importantes pérdidas y duplicación de datos (Babeu, 2019: 62).

Existían, además, obvias deficiencias tocantes exclusivamente al planteamiento filológico del proyecto. Con el fin de

10. http://www.perseus.tufts.edu/hopper/help/archived/perseus4.0.ann.full.html [actualizada: 26 mayo 2005; citada: 15 septiembre 2022].
11. https://www.perl.org/ [actualizada: 10 septiembre 2022; citada: 8 septiembre 2022].
12. https://duraspace.org/fedora/about/ [actualizada: 30 agosto 2022; citada: 8 septiembre 2022].
13. https://sourceforge.net/projects/perseus-hopper/ [actualizada: 26 enero 2016; citada: 8 septiembre 2022].

no vulnerar los derechos de autor, los textos literarios y las traducciones que nutrían la *Perseus Digital Library* eran (y siguen siendo) digitalizaciones de viejas ediciones, muchas de ellas de la Loeb Classical Library, que datan de final del s. XIX y de principios del s. XX. Aparte de las múltiples erratas derivadas del escaneo y del OCR al que fueron sometidos, no son ediciones que puedan hoy día considerarse estándares, pues no dan cuenta de la ingente labor ecdótica realizada en la fijación de los textos desde esas fechas hasta la actualidad. En lo que respecta a las traducciones, el inglés adolece de arcaísmo ortográfico y léxico, siendo, en muchas ocasiones, su comprensión más ardua que la del texto original. Y, por último, los textos no contienen los aparatos críticos de las ediciones originales. Todas estas circunstancias llevaron a cuestionar la idoneidad del uso de los materiales textuales de la PDL con fines investigadores.

La comunidad internacional de Filología Clásica, fuertemente implicada en el *Perseus Project*, fue consciente de que, una vez conseguido el objetivo de que gran parte de los textos antiguos canónicos estaban en acceso libre en la red, urgía someter a reconsideración la manera en que las colecciones textuales digitales habían de ser publicadas. La necesidad de una depuración fue explícitamente formulada en septiembre de 2006 en la reunión de investigadores *Open Source Critical Editions* (OSCE), celebrada en el Center for Computing in the Humanities (King's College, London)[14]. Este *workshop* constituyó el acta fundacional de una nueva era no sólo del 'cosmos' *Perseus*, sino del concepto de edición digital de los textos clásicos. Se abogó por que la presencia en Internet de los textos antiguos fuese más científica, menos divulgativa y dirigida, también, a un público altamente especializado. Ello implicaba que habría que velar por la calidad de los textos y enfrentar el desafío de realizar ediciones críticas que incluyesen las variantes textuales e incorporasen las lecturas de los manuscritos que o no habían sido

14. Véanse las problemáticas, y soluciones, planteadas en OSCE (2006) en Garcés y Bodard (2006) y en Bodard y Garcés (2009).

tenidos en cuenta por los mencionados viejos editores, o fueron hallados posteriormente. La existencia de ediciones críticas en las bibliotecas digitales es siempre una necesidad imperiosa, pero más si cabe, cuando se trata de obras antiguas en las que, inevitablemente y sin que existan excepciones, el texto del autor nos llega filtrado por las lecturas ofrecidas durante siglos por sucesivos editores (Boschetti, 2007: 1).

En 2006, Gabriel Bodard y Juan Garcés[15] defendieron una *praxis* que, si bien comenzaba a ser frecuente en el ámbito de creación de *software*, no tenía precedentes en el mundo de la ecdótica. Propusieron la puesta en abierto de las ediciones críticas que se elaborasen, para que pudieran ser descargadas por el usuario y, por tanto, reutilizada la totalidad de sus datos textuales, codificación, fuentes y *software*. De este modo, futuros investigadores podrían servirse del trabajo preexistente y mejorar, completar o modificar los aparatos críticos con nuevos hallazgos o conjeturas. Podrían, asimismo, incrementar el número de instancias textuales o metatextuales introduciendo notas, imágenes, acceso a comentarios, visualización de manuscritos, topónimos, nombres propios, etc.

Para que el conjunto de estos objetivos pudiera verse cumplido en un futuro más o menos cercano, se sugirió la conveniencia de que los proyectos que planeaban crear grandes bibliotecas digitales adoptasen un lenguaje estándar de codificación para, así, garantizar que los archivos de los textos pudiesen ser compartidos y utilizados con distintos *softwares* o en distintas plataformas. Se propuso también la adopción de un protocolo común que permitiese la interoperabilidad entre las bases de datos de las distintas bibliotecas. Sobre la primera cuestión, se optó por el uso del lenguaje de marcado extendido XML[16] (o su versión simplificada, SGML) según las directrices del consorcio TEI (*Text Encoding Initiative*)[17],

15. Garcés y Bodard (2006), y Bodard y Garcés (2009).

16. *Extensible Markup Language.* Disponible en: https://www.w3.org/XML/ [actualizada: 11 octubre 2016; citada: 4 febrero 2022].

recomendado por *The World Wide Web Consortium* (W3C)[18]. Respecto al segundo punto, es decir, conseguir enlazar los contenidos de las distintas bases de datos, proporcionar a los textos un mecanismo de referencia común y, en definitiva, lograr una interoperabilidad plena entre los repositorios textuales *online* presentes o futuros, pareció idóneo adoptar los identificadores estables CTS URNs (*Canonical Text Services Universal Resource Names*). La razón fue que el protocolo CTS[19] había sido desarrollado en fechas recientes por Christopher Blackwell y Neel Smith[20], con el fin de que las ediciones digitales del *Homer Multitext Project*[21] (dir.s Casey Dué y Mary Ebbott, Center for Hellenic Studies de la Universidad de Harvard) pudiesen ser publicados en *Linked Open Data* (LOD). Los URNs (*Universal Resource Names*), por su parte, son los identificadores únicos utilizados por los dos corpus canónicos de textos griegos y latinos, el TLG (*Thesaurus Linguae Graeca*) y el PHI 5.3. (*Packard Humanities Institute)*. Ellos son los que hacen posible que se pueda recuperar un pasaje concreto de una obra contenida en una u otra de estas bases de datos[22].

No obstante, el paso del tiempo ha demostrado que el uso de los CTS URNs no está exento de problemas. Señalaré los tres

17. https://tei-c.org/ [actualizada: 15 octubre 2021; citada: 4 febrero 2022].

18. La comunidad internacional W3C, dirigida por Tim Berners-Lee y Jeffrey Jaffe, tiene desde 1994 como cometido el desarrollo de estándares abiertos que garanticen el crecimiento y sostenibilidad del World Wide Web a largo plazo. Disponible en: https://www.w3.org/ [actualizada: 14 marzo 2022; citada: 14 marzo 2022].

19. http://cite-architecture.github.io/cts_spec/ [actualizada: abril 2014; citada: 4 febrero 2022].

20. Blackwell y Smith (2019: 73). Veánse también, Smith (2009) y Berti et al. (2014). Un exhaustivo análisis del protocolo CTS en la práctica editorial de la Humanidades Digitales Clásicas se encuentra en Jochen Tiepmar (2017).

21. https://www.homermultitext.org/ [actualizada: 20 mayo 2020 s.d.; citada: 24 febrero 2022].

22. Me limito a poner un ejemplo de la aplicación del protocolo CTS al canon TLG. En éste, el identificador universal de Eurípides es 0006 y el de su tragedia Medea, 003. Con el protocolo CTS, la URN de los 29 primeros versos de dicha tragedia tiene esta codificación: urn:cts:greekLit:tlg0006.tlg003:0-29.

más obvios. Existen textos (conservados en papiros, o tardíos, o medievales o modernos) que muchas bibliotecas tienen proyectado publicar, pero que no aparecen en los cánones de TLG y PHI 5.3. Por otro lado, investigaciones futuras pueden poner en duda la identificación de un autor o la autoría tradicionalmente admitida para una obra, con lo cual el identificador universal de los corpus canónicos quedaría sin efecto (Monella, 2017: 153-154). Y, por último, la adopción de los CTS URNs varía de unas comunidades académicas a otras. Por ello, para construir la arquitectura de los *Linked Open Data* de la *Digital Latin Library* (DLL), su director, Samuel J. Huskey (*per litt.*), ha optado por incluir no sólo los CTS URNs, sino también otros identificadores como los del *Virtual International Authority File*, de la *Library of Congress* e, incluso, de colecciones individuales.

En todo caso, transcurridos 16 años y dado el estado actual de la cuestión, no cabe negar que las propuestas de OSCE resultaron tener un carácter visionario. Aunque no de forma inmediata ni generalizada aún hoy día, se produjo la adopción del lenguaje de marcado TEI XML por parte de las bibliotecas comprenhensivas de textos antiguos y de los emergentes repositorios de textos medievales y neolatinos. También, y a pesar de sus reticencias iniciales, en 2006 Gregory R. Crane tomó la decisión de poner en abierto, con licencia *Creative Commons*, los textos TEI XML de la *Perseus Digital Library*[23], cuyos textos y diccionarios pasaron a engrosar con rapidez las bases de datos de numerosas plataformas, bibliotecas y portales concebidos para la investigación[24], o para

23. Crane (2014 y 2015) explica el papel esencial jugado por los argumentos a favor de los datos abiertos esgrimidos por especialistas de la talla de Chris Blackwell, Gabby Bodard, Helma Dik, Tom Elliott, Sebastian Heath, Ross Scaife y Neel Smith.

24. Entre los muchos existentes, digno de especial mención es el proyecto de la University of Chicago *Perseus under PhiloLogic* (dir.a Helma Dik. Disponible en: https://perseus.uchicago.edu/ [actualizada: 21 mayo 2021; citada 4 febrero 2022]). Contiene, aproximadamente, una tercera parte de los textos de la *Perseus Digital Library.* A diferencia de ésta, su finalidad no es tanto la lectura enriquecida o radial, sino servir de herramienta para la investigación en lingüística de corpus,

la enseñanza del latín, griego y cultura clásica[25]. Y, finalmente, en 2013, comenzaron a darse los primeros pasos para la puesta en abierto de la totalidad de los datos y objetos de *Perseus Project* (Almas, Babeu y Krohn, 2014). De forma progresiva, el mundo de la Filología Clásica fue tomando conciencia de que era imprescindible estimular la colaboración tanto entre bibliotecas digitales (para lo cual era condición indispensable la mencionada sindicación), como entre investigadores (filólogos, informáticos o humanistas informáticos) separados por barreras de espacio, idioma y cultura.

## 3. EL PRESENTE DE LAS BIBLIOTECAS COMPREHENSIVAS DE TEXTOS GRIEGOS Y LATINOS

Desde 2005 hasta la actualidad, surgieron las otras tres grandes plataformas existentes en acceso libre, todas ellas en desarrollo: *Open Greek and Latin* (OGL), *Library of Digital Latin Texts* (LDLT) y *Musisque Deoque* (MQDQ). Ha de advertirse que la consecución de las prácticas editoriales recomendadas en OSCE tiene diverso alcance en unas y otras. OGL, innegablemente, se está centrando en aspectos que tocan a la implementación tecnológica, mientras que LDLT y MQDQ dan amplia cabida también a la vertiente netamente filológica, siendo su meta última la elaboración de ediciones críticas.

pues la codificación practicada permite búsquedas por ***lemmata*** y por combinación de características morfológicas. Ahora bien, la sintaxis de búsqueda, en su desarrollo actual, es un tanto «overwhelming» (Dik, 2018), inconveniente de cierta entidad en cuya resolución el equipo del proyecto está trabajando.

25. Dado el número ingente de estos últimos, sería imposible dar cuenta de ellos. Por razones de espacio, aduciré un único caso que pone en evidencia la celeridad con que distintos proyectos se apropiaron de los materiales de la PDL. Tan sólo unos pocos meses después de que Perseus ofreciese accesibilidad a sus colecciones, Jeremy March (Queens College, New York) creó *philolog.us* (Disponible en: https://philolog.us/ [actualizada: junio 2005; citada: 4 febrero 2022]), la primera de las plataformas concebidas para un manejo más amable de los diccionarios del Perseus Project que el que permite la interfaz de éste.

### 3.1. OPEN GREEK AND LATIN (OGL)[26]

OGL, versión 5.0 de la *Perseus Digital Library*, es el fruto de la expansión del *Perseus Project* desde su base inicial en la Tufts University a las principales redes académicas de todo el mundo. Este proceso culminó en 2014 con el establecimiento de una segunda base del *Perseus Project* en el *Humboldt Chair of Digital Humanities* de la Leipzig Universität (Lang, 2018: 2 y 9). Es un proyecto colaborativo monumental, acorde con su también monumental objetivo. No sólo da accesibilidad libre a los textos antiguos canónicos contenidos en las versiones previas de *Perseus*, sino que ofrecerá el mayor número de ediciones posibles de la totalidad de textos griegos y latinos, de procedencia diversa (tradición manuscrita, inscripciones, papiros, etc.), escritos desde la antigüedad hasta la actualidad[27]. Con este fin, están siendo escaneadas miles de fuentes primarias y sus traducciones, en todos los idiomas en que éstas existen. El texto de las ediciones de las obras canónicas antiguas es, por tanto, el mismo que el de las viejas ediciones impresas, libres de derechos de autor, que alimentaban *Perseus Tufts* y *Perseus Hopper*. No hay indicios de que forme parte de la agenda inmediata de OGL la elaboración de ediciones críticas que, sin embargo, es la meta prioritaria de las dos bibliotecas de las que trataré en los epígrafes siguientes.

En la primera fase de OGL, también dirigida por Gregory R. Crane, éste estableció las tareas prioritarias (Crane, 2014): primero, someter a corrección las numerosas erratas existentes en los textos, que aún hoy día conforman el corpus canónico de la *Perseus Digital Library*, que iban a ser integrados en OGL; segundo, convertir a Unicode el estándar Betacode utilizado para codificar el

26. https://opengreekandlatin.org/ [actualizada: 5 enero 2021; citada: 15 febrero 2022].

27. En palabras de Gregory R. Crane, se dará accesibilidad «to a global community not just a canon of secular Greek and Latin sources but thousands of years of human critical experience that survives in these languages» (Crane, 2020).

griego politónico en los textos de *Perseus*; y, por último, dotar a las ediciones ya existentes de un sucinto pero homogéneo etiquetado inicial TEI XML. Para ello, OGL utilizó el subconjunto de etiquetas que había sido elaborado para que sirviese de Guía de Etiquetado en la edición de los textos de *Epigraphic Documents in TEI XML* (EpiDoc)[28]. La finalidad era que futuros editores pudiesen elaborar con más facilidad sus ediciones, así como añadir la codificación que sus fines investigadores requiriesen.

Según se ha dicho, en 2013 *Perseus* había implementado el protocolo CTS para las URNs de los cánones de TLG y PHI. Ahora era necesario crear un *software* que garantizase el mantenimiento de datos y la interoperabilidad del gigante *Perseus* 5.0 o, si se prefiere, OGL. El resultado fue *CapiTainS*[29], creado por Thibault Clérice (École Nationale des Chartes) y Bridget Almas (Tufts University). Sus componentes son interoperativos con cualquier *software* compatible con CTS (y, por tanto, también con EpiDoc XML, utilizado por OGL) y permite mejorar las funcionalidades del motor de búsqueda, la actualización de los datos y la forma de citación, puesto que mantiene los metadatos generales separados de los metadatos textuales[30]. En el conjunto de herramientas de CapiTainS se basa el *Scaife Viewer*[31], el nuevo entorno de lectura creado para OGL y para los corpus o fondos de cualquier biblioteca que utilice los *CapiTainS CTS-services* (Crane, 2018a). Su primera versión, lanzada en 2018, mejoró sustancialmente la experiencia de lectura que aún ofrece la interfaz de la *Perseus Digital Library*. No obstante, en la actualidad, presenta numerosos *bugs* e

28. https://epidoc.stoa.org/ [actualizada: noviembre 2021; citada: 25 febrero 2022]

29. https://github.com/Capitains [actualizada: 27 septiembre 2021; citada: 25 febrero 2022].

30. *Cf.* al respecto, *las Guidelines* de *CapiTainS*. Disponibles en: http://capitains.org/pages/guidelines#foreword [actualizada: 8 junio 2018; citada: 25 febrero 2022].

31. https://scaife.perseus.org/ [actualizada: 14 marzo 2022; citada: 16 marzo 2022].

inconsistencias (Sansom, 2019), que encuentran difícil explicación si se tiene en cuenta el número de ingenieros y humanistas informáticos que forman parte del equipo, la copiosa financiación de que dispone el proyecto y el tiempo ya empleado en su desarrollo (más de seis años). También resulta extraño el silencio del equipo editorial respecto a la existencia de estos fallos y a las estrategias previstas para subsanarlos.

Inevitablemente, el usuario constata que el alineamiento de texto primario y traducción no siempre funciona, que persisten algunas (o muchas) de las erratas de los textos de *Perseus Tufts* y de su siguiente versión, *Perseus Hopper*, y que la navegación por autores produce más errores de los deseables. Cabe, asimismo, deducir que está habiendo problemas con la integración en el *Scaife Viewer* de los diccionarios de *Perseus Tufts*, puesto que el acceso al significado de las palabras sólo es posible en los textos griegos y, aun así, de forma muy parca. En lo que respecta al motor de búsqueda, personalmente, lo encuentro poco intuitivo[32] y, al menos a día de hoy, poco fiable. Éste es un hecho fácilmente comprobable si se comparan los resultados obtenidos en una búsqueda en OGL, con los resultados arrojados por la versión 4 de *Diogenes*[33], la aplicación creada por P. Heslin (University of Durham) para la búsqueda y navegación de los textos contenidos en los CD-Roms de TLG y PHI. Se da la circunstancia, además, de que las búsquedas por forma, en ocasiones, producen más resultados que la búsqueda por *lemmata*.

Una de las últimas funcionalidades incorporadas es la interoperabilidad con parte de los contenidos del Center for Hellenic Studies (Harvard University). El usuario puede leer los

32. The Center for Hellenic Studies ofrece un tutorial para adiestrar al usuario en los criterios de búsqueda. Disponible https://www.youtube.com/watch?v=ZvLx0UxrNVs [actualizada: 10 junio 2020; citada: 25 febrero 2022].

33. https://d.iogen.es/d/index.html [actualizada: 14 diciembre 2021; citada 4 marzo 2022].

textos de OGL y, en paralelo y sin cambiar de página, consultar los comentarios que se están realizando en la plataforma *New Alexandria*[34], creada por el citado centro (Crane, 2018b). Así, mientras se procede a la lectura de la *Ilíada*, es posible consultar los comentarios de Leonard Muellner y Gregory Nagy a los libros 1 y 2 del poema homérico. No obstante, salvo error por mi parte, tal consulta sólo funciona en el caso de la *Ilíada*, no así en el resto de comentarios (concretamente, de la obra de Píndaro, Pausanias y Herodoto) que se están elaborando en *New Alexandria*. En todo caso, aunque mucho de su potencial está aún en un desarrollo menor de lo que el usuario desearía, con seguridad OGL será (de hecho, ya lo es) una herramienta magnífica. Es un proyecto de una ambición sin precedentes, y con una vocación y solvencia tecnológica sin parangón en el mundo de la Filología Clásica.

### 3.2. *LIBRARY OF DIGITAL LATIN TEXTS* (LDLT)[35]

La LDLT es parte de la *Digital Latin Library* (DLL)[36], dirigida por Samuel J. Huskey (University of Oklahoma). La DLL surgió, precisamente, para dar respuesta a la necesidad de adecuar la calidad filológica de los textos clásicos a los requerimientos de la nueva era de edición digital. Es, como OGL, un proyecto de gran envergadura cuya creación fue promovida en 2014 por la Society for Classical Studies, la *Medieval Academy of America* y la *Renaissance Society of America*. Tiene dos secciones: a)el *DLL Catalog*[37], que permite un rápido acceso a más de cinco mil textos latinos, de calidad diversa, disponibles en distintos sitios web, y b)

34. https://oc.newalexandria.info/ [actualizada: 22 junio 2022; citada 7 septiembre 2022].
35. https://digitallatin.org/library-digital-latin-texts [actualizada 25 febrero 2022; citada: 25 febrero 2022].
36. https://digitallatin.org/ [actualizada 25 febrero 2022; citada: 25 febrero 2022]. Es posible que algunas de las URL de las páginas de la *DLL* den error pues se está procediendo a la migración del servidor.
37. https://catalog.digitallatin.org/ [actualizada 25 febrero 2022; citada: 25 febrero 2022].

la *Library of Digital Latin Texts* (LDLT), plataforma que impulsa la creación de *born-digital critical editions* de textos latinos de todos los tiempos (Huskey, 2019: 19-34). Las ediciones, codificadas en XML según el estándar TEI, no sólo son de acceso libre sino que se publican en abierto. Es posible la descarga en formato XSLT, LaTeX y PDF y, por tanto, la reutilización de sus datos. Samuel J. Huskey y Hugh Cayless (Duke University) elaboraron una detalladísima *Guideline*[38] de etiquetado TEI XML que contempla la práctica totalidad de instancias textuales susceptibles de ser codificadas en los textos antiguos. Es una contribución inapreciable, pues evita a los editores el arduo trabajo de estudiar las, en ocasiones, inabordables *Guidelines* de TEI con el fin de seleccionar qué etiquetas se adecuan a la idiosincrasia de sus textos y sus objetivos editoriales.

La LDLT ofrece un entorno de lectura muy grato que, por su calidad, es parangonable con el de OGL[39], proyecto con el que comparte muchos recursos[40]. Si la visualización interactiva de los datos es fundamental en cualquier plataforma que posee elementos subyacentes complejos, lo es más en el caso de una biblioteca en la que los textos tienen aparato crítico. Por ello, LDLT confirió una prioridad alta a la implementación de técnicas de visualización[41]. A tal efecto, Chris Weaver (School of Computer

38. https://digitallatin.github.io/guidelines/LDLT-Guidelines.html. [actualizada: 4 marzo 2020; citada: 4 marzo 2022].

39. Tal como se puede comprobar en https://ldlt.digitallatin.org/editions/calpurnius_siculus/poem1.html [actualizada: 2017; citada: 19 enero 2022].

40. De hecho, ambas plataformas participaron en el desarrollo de las *Alpheios Reading Tools.* Este *software* permite enriquecer los textos griegos y latinos de OGL y de la LDLT, de tal modo que, picando sobre cada palabra, se genera su análisis léxico y morfológico. Disponible en: http://alpheios.net/ [actualizada: s.d.; citada: 25 febrero 2022]. Sin embargo, en lo que respecta a los *Linked Open Data*, como se ha dicho, LDLT, al contrario que OGL, ha optado por no limitarse a los CTS URNs para construir su arquitectura (cf. p. 8). p. 292

41. Cf. al respecto, https://digitallatin.org/library-digital-latin-texts/data-visualization [actualizada: s.d.; citada: 19 enero 2022].

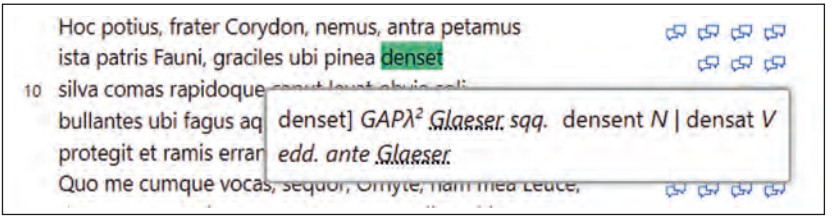



Figura 1. Fuente: Captura de imagen de la edición crítica de Calpurnius Siculus de la LDLT

Science de la University of Oklahoma) creó *Improvise*[42], que ofrecía una visualización de las variantes textuales que fue muy innovadora en 2017, fecha en que el análisis visual del texto era un territorio relativamente poco explorado (Huskey, 2019: 31).

Este *software* recibió subsiguientes desarrollos que requieren un procesamiento de la información cuya realización *online*, desafortunadamente, no es posible. Para que el usuario pueda beneficiarse de ellos, es preciso descargar una aplicación de escritorio para utilizarla con los archivos de la LDLT. En cambio, si se practica la lectura de los textos *online*, el usuario, como si de una edición crítica en soporte papel se tratase, ha de cambiar de página para saber a qué manuscritos se refieren las siglas G, A, P y λ que aparecen en la figura 1. Esto, según veremos, no sucede en el caso de la visualización de la información de los aparatos críticos de la biblioteca *Musisque Deoque*, que será objeto de mi atención en el siguiente epígrafe.

En lo que se refiere a la elaboración de las ediciones críticas, cualquier especialista puede postularse como editor de la *Library of Digital Latin Texts*, si su propuesta supera la evaluación de pares ciegos. A diferencia de as ediciones libres de *copyright*, utilizadas por OGL, los editores de LDLT trabajan sobre varias ediciones preexistentes en soporte papel que, por su calidad,

42. https://www.cs.ou.edu/~weaver/improvise/ [actualizada: 21 septiembre 2014; citada: 25 febrero 2022].

son consideradas hoy día como estándares. Ahora bien, no se contempla la posibilidad de que los editores de la LDLT realicen una colación autóptica de los manuscritos. Tal *modus operandi* suscita recelos en los sectores partidarios de una *práxis* ecdótica más tradicional. Sin embargo, existe consenso al respecto entre los especialistas en edición digital: la Filología Clásica digital ha de hacer el esfuerzo de superar prejuicios y aprovechar el gran trabajo editorial realizado durante siglos. Es necesario, defienden, dejar a un lado las definiciones rígidas de lo que es o debe ser una edición crítica digital[43], sobre todo cuando se da la circunstancia de que, como todos sabemos, muchas de las ediciones críticas en soporte papel (recientes y no tan recientes) de textos clásicos se basan, precisamente, en ediciones anteriores. Volver a hacer una colación de todos los manuscritos que contienen cada obra griega y latina no es necesario y, además, sería impensable si el objetivo es hacer una biblioteca que contenga una cantidad ingente de obras. Ello no implica que los editores no tengan que realizar una gran labor filológica, incorporando la información textual aportada por estudios recientes, incrementando el número de testimonios utilizados por los editores-base o detectando los numerosos errores, frecuentes en las ediciones impresas, de atribución de lecturas falsas a manuscritos.

43. Veánse las declaraciones de Samuel J. Huskey y de Monica Berti, dir.ª de *Leipzig Open Fragmentary Texts*, en *Summer Sunoikisis Digital Classics July 1, 2021, chaired by Monica Berti and Gabriel Bodard.* Disponible en: https://www.youtube.com/watch?v=CL-J-QSj4uA&t=4550s [actualizada: 1 julio 2021; citada: 4 abril 2022]. Cf. también Boschetti (2007).

### 3.3. MUSISQUE DEOQUE. A DIGITAL ARCHIVE OF LATIN POETRY, FROM ITS ORIGINS TO THE ITALIAN RENAISSANCE (MQDQ)[44]

Es un hecho que la Tufts University (PDL), la Leipzig Universität (OGL) y la *Society for Classical Studies*, principal promotora de la LDLT, ostentan un liderazgo indiscutible en el campo de las Humanidades Digitales Clásicas. No obstante, fue la Università Ca'Foscari Venezia la pionera en promover la creación de una biblioteca digital con ediciones críticas, empresa que ningún proyecto había acometido. La finalidad de este proyecto colaborativo era, y sigue siendo, crear un archivo digital de poesía latina y neolatina, a cuyas obras les están siendo incorporados los aparatos críticos. El ritmo de edición de MQDQ (dir. Paolo Mastandrea) ha sido extraordinario en la última década y se ha enriquecido con la integración de otros proyectos interdisciplinares. Algunos de ellos están aún en las primeras fases de desarrollo, pero dotan a la plataforma de muchas y muy interesantes funcionalidades en cuya descripción no voy a entrar: *Epigraphica*[45], *Metrical typology*[46] y la emergente serie *Hellenica*[47], que consta de 10 obras griegas en verso, todavía sin aparato crítico.

MQDQ surgió en una fecha muy temprana, 2005, por tanto, un año antes de la celebración del *workshop* OSCE. Esto puede explicar que, cuando las nuevas prácticas de edición digital se fueron sustanciando entre los filólogos clásicos, cosa que no se produjo de manera inmediata, MQDQ tenía un recorrido lo suficientemente largo como para que pudiera permitirse volver sobre sus pasos y replantearse su metodología, tanto en lo tocante a tecnología como a su enfoque editorial, que resulta ser un

44. http://mizar.unive.it/mqdq/public/index [actualizada: 24 julio 2022; citada: 21 septiembre 2022].

45. https://mizar.unive.it/mqdq/public/ce/presentazione [actualizada: 24 julio 2022; citada: 21 septiembre 2022].

46. https://mizar.unive.it/mqdq/public/indici/metri [actualizada: 24 julio 2222; citada: 21 septiembre 2022].

47. https://mizar.unive.it/hellenica/public/indici/autori [actualizada: 13 febrero 2021; citada: 21 septiembre 2022].

tanto errático (Gelderblom, 2008: 224). Sobre ninguno de ambos aspectos, se encuentra información consistente en su sitio web o en los estudios que describen el proyecto[48], que se limitan a señalar el hecho de que practican el etiquetado de los textos con el lenguaje XML y de que sus aparatos críticos no tienen la pretensión de ser exhaustivos. Al estar el corpus concebido para la lectura *online* y al no ser descargables los archivos XML, el usuario no puede reutilizar el material. Por la misma razón, tampoco es posible conocer el set de etiquetas XML que emplean, aunque cabe deducir que no siguen las directrices de TEI. De hecho, Daniele Fusi (Venice Centre for Digital and Public Humanities), al anunciar en 2020 la transición tecnológica que se está llevando a cabo en MQDQ, afirmó que las ediciones «will be reconstituted as a database with an abstract, highly atomized data model»[49], lo cual difiere de la estructura conceptual de TEI, aunque no es incompatible con el hecho de que TEI pueda ser uno de los *outputs* que dicho modelo genere.

Al menos en fechas previas a la anunciada remodelación, MQDQ trabajaba con un «cross-plataform software», denominada, MQDQ2, escrita en Java. Esta plataforma ponía a disposición del editor el texto, previamente escaneado y practicada la OCR. A través de sucesivos cuadros de dialogo, la aplicación iba guiando al autor de la edición en la introducción de los datos y metadatos: la edición-base, el nombre del autor de la edición digital, los *sigla* de los manuscritos y su descripción, y el *conspectus uirorum doctissimorum* que propusieron conjeturas. Ya por último, se procedía a la elaboración del aparato crítico, que incluía la lectura aceptada por el editor de MQDQ, las variantes más relevantes y, en caso de existir, las conjeturas (Manca, 2014: 131-137). Por tanto, era la plataforma MQDQ2 la que practicaba la codificación XML. El editor, al que no era visible el documento codificado, se limitaba a

48. Por ejemplo, Zurli y Mastandrea (2009), Manca (2009) y Mastandrea y Spinazzè (2011).

49. *Encoding a critical apparatus.* Disponible en: https://youtu.be/27_7iBQVjg0. [actualizada: diciembre 2020; citada: 21 septiembre 2022].

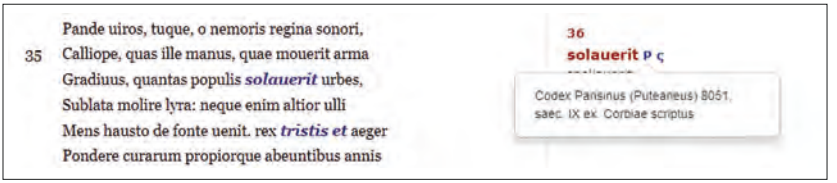



Figura 2. Fuente: Aparato crítico de Stat. Theb. 4.36 en MQDQ

rellenar los cuadros de diálogo con los datos que la aplicación le iba pidiendo. En ningún momento, se esperaba de él que fuese conocedor del lenguaje XML (Manca, 2014: 132). Es de suponer que, aunque la anunciada remodelación tenga lugar, la forma de proceder seguirá siendo muy similar en lo que a los editores respecta.

Tal como se ha mencionado, al menos en el momento actual, la visualización de los datos de los manuscritos y de las ediciones que aparecen en los aparatos mejora a la de LDLT (aunque la interfaz de ésta última es de mucha más calidad), pues esta información es accesible sin necesidad de cambiar de página (para la LDLT, véase Figura 1).

No obstante, en lo que se refiere tanto a interfaz como a información aportada, el proyecto *Thebarum fabula. Biblioteca Grecorromana sobre el mito tebano* de la Universidad de Santiago de Compostela, (ThDig)[50], en mi opinión, es superior. Gracias a la visualización de los datos en una doble ventana emergente, ofrece una mayor y mejor interactividad.

Prescindiendo ahora de la visualización propiamente dicha, las Figuras 2 y 3 ponen en evidencia la que considero la

50. El proyecto colaborativo ThDig (http://thebarumfabula.usc.es/proyecto/), dirigido por mí, escapa a los límites temáticos del presente estudio, pues es una biblioteca que sólo contiene las obras que documentan la historia del ciclo tebano. Su número es considerable, pero en forma alguna es parangonable con el de las obras que OGL, LDLT y MQDQ proyectan publicar. De todas formas, sobre la oferta existente, ThDig tiene la innegable ventaja de ofrecer traducciones originales a lenguas modernas (español, inglés, italiano, griego moderno, etc.), alineadas en paralelo al texto griego y latino con las etiquetas de alineamiento TEI XML.

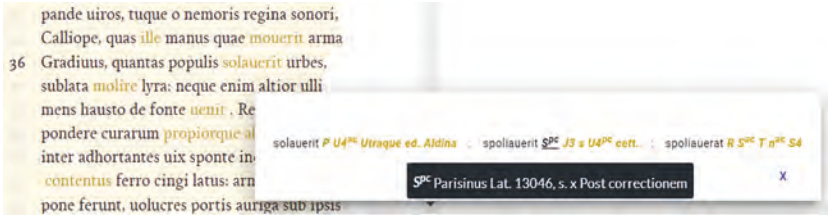



Figura 3. Fuente: Aparato crítico de Stat. *Theb.* 4.36 en ThDig

mayor objeción a este proyecto y que guarda relación con el enfoque filológico adoptado. Con mucho acierto, los textos y los aparatos críticos de MQDQ, como es también el caso de la LDLT, se basan en ediciones preexistentes en soporte papel pero que, al contrario de lo que sucede en el *Perseus Digital Library* y, consiguientemente, en OGL, son ediciones actualmente consideradas estándares. Así, por citar un ejemplo, el texto y aparato crítico de la *Eneida* de Virgilio toma como edición de referencia la de Marius Geymonat (1973/2008), frente al texto de la edición de James B. Greenough (1900) que ofrecen la *Perseus Digital Library*, la OGL y la práctica totalidad de bibliotecas digitales de acceso libre. Según he dicho, el equipo editorial de MQDQ explicita que el objetivo del proyecto es ofrecer para todo el corpus poético de la literatura latina, un sencillo aparato crítico que dé cuenta de las variantes y conjeturas más relevantes[51]. Tal declaración de principios denota una *praxis*, cuando menos, arriesgada, pues ¿cómo decidir qué variantes incluir y cuáles excluir?

Como se puede apreciar, el aparato crítico de ThDig del verbo '*solauerit*' del verso 36 del libro cuarto de la *Tebaida* de Estacio (Figura 3) es mucho más exhaustivo que el de la edición de MQDQ (Figura 2). ¿Es necesario? Sin duda lo es, si MQDQ pretende servir de instrumento de investigación también a especialistas en crítica textual. La causa de la divergencia en la información textual aportada por ambas bibliotecas es que el texto de la edición digital de MQDQ, realizada por Giuseppe C. Musa, tomó como edición-base

51. Boschetti (2007: 3) y Manca (2014: 138).

la de Donald E. Hill (1983/1996) cuando, sin duda, habría sido mejor opción haber utilizado la de John B. Hall (2008-2009). Mientras Donal E. Hill emplea alrededor de 31 testimonios, John B. Hall usa aproximadamente 92, y ello sin contar testimonios ignotos de los que sólo se tiene noticia por aparecer citados por algún estudioso, ni los siete incunables y las dos ediciones Aldinas cuyo valor, al no conocerse de qué manuscrito derivan, no hay que descartar (Martínez Zepeda, *per litt.*). Al haber utilizado un menor número de manuscritos, faltan en la edición de Donal E. Hill (que el editor de MQDQ simplifica al máximo) lecturas importantes que, sin embargo, ayudan a comprender la historia del texto. Además Donal E. Hill incurre en un error frecuente, admitiendo como conjeturas lecturas que ya aparecían en manuscritos que ni el editor de MQDQ ni el autor de la conjetura consultaron.

En definitiva, y para no extenderme en las objeciones estrictamente filológicas, es indudable el papel pionero jugado por MQDQ en la presentación digital de ediciones críticas de los textos griegos y latinos. No obstante, los aparatos críticos digitales, que ciertamente han de ser lo más sencillos y limpios posible, tienen que ofrecer una información textual exhaustiva, cosa que no es compatible con el apremiante ritmo editorial de la plataforma.

## 4. CONCLUSIONES

Como se ha expuesto, desde principios del presente siglo, el mundo de la Filología Clásica fue consciente de que había llegado el momento de conformar una segunda generación de bibliotecas digitales, que permitiese la adaptación de los datos existentes a los avances del desarrollo web experimentado desde finales del siglo XX y que garantizase, asimismo, el cumplimiento de los estándares de calidad filológica, más exigentes que en el pasado. Se planteó como indispensable que existiera consenso en puntos cruciales.

Hubo unanimidad en lo que respecta a la adopción del lenguaje XML con el fin de permitir la compartición de archivos.

Ahora bien, según se ha señalado, OGL ofrece una codificación de sus obras que he descrito como «sucinto etiquetado inicial». MQDQ, por su parte, al menos a día de hoy, no contempla la codificación de las innumerables instancias textuales que LDLT tiene previstas etiquetar. En contrapartida, a esta última le está resultando difícil encontrar filólogos clásicos dispuestos a aprender el lenguaje XML según las directrices del estándar TEI[52].

Respecto a la deseada interoperabilidad entre plataformas, OGL y LDLT han adoptado los protocolos CTS URNs, no así MQDQ. Ésta, con la asistencia del muy solvente Venice Centre for Digital and Public Humanities (dir. Franz Fischer), parece estar más interesada en buscar la interoperabilidad con las dos grandes bibliotecas temáticas italianas de textos latinos, la *Biblioteca digitale di testi latini tardoantichi*[53] y el *Archivio della Latinitá Italiana del Medioevo*[54].

Y, por último, en lo que se refiere a la proclamada necesidad de que las grandes bibliotecas elaboren ediciones críticas, los hechos demuestran que hará falta tiempo para que sea una realidad. Al mencionado temor que suscita la codificación TEI XML, hay que añadir las dificultades de encontrar financiación y el hecho de que la creación de una edición digital no computa como investigación (Kiss, 2019: 233).

En definitiva, hay mucho camino andado pero queda mucho, y muy apasionante, por hacer. Se hará, pero será preciso que los investigadores hagan gala de generosidad y permitan la reutilización de sus ediciones. Si no, no se podrán alcanzar tan ambiciosos objetivos.

52. En lo que se refiere al reconocimiento de la falta de competencias digitales por parte de los propios filólogos clásicos, cf. Monella (2018: 144 y Fischer, 2019: 203 y 206-207).

53. Università degli studi del Piemonte Orientale-Vercelli. Disponible en: https://digiliblt.uniupo.it/index.php [actualizada 23 febrero 2022; citada: 25 febrero 2022].

54. Universidades de Verona, Napoli, Palermo, Ca' Foscari Venezia, Siena-Arezzo y Basilicata. Disponible en: http://alim.unisi.it/ [actualizada: s.d.; citada: 25 febrero 2022]

## REFERENCIAS BIBLIOGRÁFICAS


Almas, Bridget, Babeu, Alison y Krohn, Anna (2014). Linked Data in the Perseus Digital Library. *ISAW Papers 7: Current Practice in Linked Open Data for the Ancient World*. New York: New York University. Disponible en: http://dlib.nyu.edu/awdl/isaw/isaw-papers/7/almas-babeu-krohn/ [actualizada: s.d.; citada: 15 febrero 2022].

Babeu, Alison (2019). The Perseus Catalog: of FRBR, Finding Aids, Linked Data, and Open Greek and Latin. En: Monica Berti (ed.ª). *Digital Classical Philology: Ancient Greek and Latin in the Digital Revolution*. Berlin y Boston: De Gruyter Saur, pp. 53-72.

Berti, Monica *et al.* (2014). The Linked Fragment: TEI and the Encoding of Text Reuses of Lost Authors. *Journal of the Text Encoding Initiative*, 8. *Selected Papers from the 2013 TEI Conference*. https://doi.org/10.4000/jtei.1218.

Berti, Monica (2021): *Digital Editions of Historical Fragmentary Texts*. Heidelberg: Propylaeum.

Blackwell, Christopher W. y Smith, Neel (2019). The CITE Architecture: a Conceptual and Practical Overview. En: Monica Berti (ed.ª). *Digital Classical Philology. Ancient Greek and Latin in the Digital Revolution*. Berlin y Boston: De Gruyter Saur, pp. 73-94.

Bodard, Gabriel y Garcés, Juan (2009). Open Source Critical Editions: a Rationale. En: Marilyn Deegan y Kathryn Sutherland (eds.). *Text Editing, Print, and the Digital World*. London: Ashgate Press, pp. 84-98.

Boschetti, Federico (2007). Methods to Extend Greek and Latin Corpora with Variants and Conjectures: Mapping Critical Apparatuses onto Reference Text. En: *Proceedings of the Corpus Linguistics Conference (CL 2007)*. Birmingham: University of Birmingham. Disponible en: http://ucrel.lancs.ac.uk/publications/CL2007/paper/150_Paper.pdf [actualizada: enero 2007; citada: 15 febrero 2022].

Chen, Jiangping (2016). *Multilingual Access and Services for Digital Collections*. Santa Barbara: CA Libraries Unlimited.

Crane, Gregory (2002). Cultural Heritage Digital Libraries: Needs and Components. En: *Proceedings of the 6th European Conference on Research and Advanced Technology for Digital Libraries*. London: Springer Verlag, pp. 626-637.

Crane, Gregory R. (2014). Open the Perseus Corpus and the Leipzig Corpus of Open Greek and Latin (Draft for a blog post). No disponible [actualizada: 8 febrero 2014; citada: 2016].

Crane, Gregory R. (2015). Getting to Open Data for Classical Greek and Latin: Breaking Old Habits and Undoing the Damage. A Call for Comment! *Perseus Digital Library Updates*. Disponible en: https://sites.tufts.edu/perseusupdates/2015/03/04/getting-to-open-data-for-classical-greek-and-latin-breaking-old-habits-and-undoing-the-damage-a-call-for-comment/ [publicada: 4 marzo 2015; citada: 21 marzo 2022].

Crane, Gregory R. (2018a). First Version of the Scaife Digital Library Viewer Goes Live: Building the Future while Remembering a Friend. Disponible en: https://sites.tufts.edu/perseusupdates/2018/03/15/first-version-of-the-scaife-digital-library-viewer-goes-live-building-the-future-while-remembering-a-friend/ [actualizada: 15 marzo 2018; citada: 14 febrero 2022].

Crane, Gregory R.(2018b). Individual Developments and Systemic Change in Philology. Disponible en: https://chs.harvard.edu/gregory-crane-individual-developments-and-systemic-change-in-philology/ [actualizada: 1 mayo 2018; citada: 4 febrero 2022].

Crane, Gregory R. (2020). Greek, Latin, and a Global Dialogue among Civilizations. En: Elmer, David *et al.* (ed.). *Donum natalicium digitaliter confectum Gregorio Nagy septuagenario a discipulis collegis familiaribus oblatum: A Virtual Birthday Gift Presented to Gregory Nagy on Turning Seventy by his Students, Colleagues, and Friends*. Washington, D.C. Center for Hellenic Studies. Disponible en: https://chs.harvard.edu/donum-natalicium-digitaliter-confectum-gregorio-nagy-septuagenario-a-discipulis-collegis-familiaribus-oblatum-toc/ [actualizada: 2 noviembre 2020; citada: 15 marzo 2021].

Dik, Helma (2018). Perseus Project Texts Loaded under PhiloLogic: Final Season for Philologic 3, Summer 2018. Disponible en: https://perseus.uchicago.edu [actualizada: 21 mayo 2021; citada 4 febrero 2022].

Fischer, Franz (2019). Digital Classical Philology and the Critical Apparatus. En: Monica Berti (ed.ª). *Digital Classical Philology:*

*Ancient Greek and Latin in the Digital Revolution*. Berlin y Boston: De Gruyter Saur, pp. 203-219.

Garcés, Juan y Bodard, Gabriel (2006). Open Source Critical Editions Workshop Report. Disponible en: http://www.methodsnetwork.ac.uk/activities/act9report.html [actualizada: 2008; citada: 4 febrero 2022].

Gelderblom, Werner (2008). Musisque Deoque: un archivio digitale di poesia Latina. *Variants*, 7: 223-225.

Geymonat, Marius (1973/2008). *P. Vergili Maronis opera*. Roma: Edizioni di Storia e Letteratura.

Giarratano, Cesare (ed.) (1910). *Calpurnii et Nemesiani Bucolica*. Napoli: Detken and Rocholl. Edición digital de Samuel J. Huskey y Hugh Cayless. Disponible en: https://ldlt.digitallatin.org/editions/calpurnius_siculus/ [publicada: 2017; citada: 19 enero 2022].

Greenough, James B. (1900). *Bucolics, Aeneid, and Georgics of Vergil*. Boston: Ginn & Co.

Hall, John B., Ritchie, Annabel y Mike J. Edwards (2007-2008). *P. Papinius Statius. Thebaid and Achilleid*, 3 vols. Newcastle (UK): Cambridge Scholars Publishing.

Hill, Donald E. (1983/1996). *Papini Statii Thebaidos Libri XII*. Leiden: Publischer Brill.

Huskey, Samuel J. (2019). The Digital Latin Library: Cataloging and Publishing Critical Editions of Latin Texts. En: Monica Berti (ed.ª). *Digital Classical Philology: Ancient Greek and Latin in the Digital Revolution*. Berlin y Boston: De Gruyter Saur, pp. 19-34.

Huskey, Samuel J. y Cayless, Hugh (2022). Guidelines for Encoding Critical Editions for the Library of Digital Latin Texts. Disponible en: https://digitallatin.github.io/guidelines/LDLT-Guidelines.html. [actualizada: 4 marzo 2020; citada: 4 marzo 2022].

Kiss, Dániel (2019). Classics into Code: Latin Texts in the Digital Space. *Storie e linguaggi*, 5 (1): 221-237.

Lang, Sarah (2018). [Review of *Perseus Digital Library 1985-2017*, Gregory Crane (ed.)]. *RIDE. A Review Journal for Digital Editions and Resources*. Disponible en: https://ride.i-d-e.de/issues/issue-8/perseus/ [actualizada: 10 diciembre 2017; citada: 14 marzo 2022].

Manca, Massimo (2009). Database and Corpora of Ancient Texts towards the 'Second Dimension': Theory and Practice of

Musisque Deoque Project. En: Anreiter, Peter y Kienpointner, Manfred (eds.). *Computational Linguistics and Latin Philology. 15th Colloquium on Latin Linguistics*. Innsbruck: Universität Innsbruck, pp. 697-702.

Manca, Massimo *et al.* (2014). Musisque Deoque: Text Retrieval on Critical Editions. *Journal for Language Technology and Computational Linguistics*, 26: 129-140.

Mastandrea, Paolo y Spinazzè, Linda (eds.) (2011). *Nuovi archivi e mezzi d'analisi per i testi poetici: i lavori del progetto Musisque Deoque, Venezia 21-23 giugno 2010*. Amsterdam: Hakkert.

Monella, Paolo (2017). Forme del testo digitale. En: Raul Mordenti (ed.). *In filologia digitale: problemi e prospettive*. Roma: Bardi Edizioni, pp. 145-148.

Monella, Paolo (2018). Why Are There no Comprehensively Digital Scholarly Editions of Classical Texts? En: Adele Cipolla (ed.ª). *Digital Philology: New Thoughts on Old Questions*. Padova: libreriauniversitaria.it edizioni, pp. 141-159.

Sansom, Stephen A. (2019). [Review of *Perseus Digital Library Scaife Viewer*]. *Society for Classical Studies Blog*. Disponible en: https://classicalstudies.org/scs-blog/stephensansom/review-perseus-digital-library-scaife-viewer [publicada: 1 marzo 2019; citada: 16 febrero 2022].

Smith, Neel (2009). Citation in Classical Studies. *Digital Humanities Quarterly* 3 (1). *Changing the Center of Gravity: Transforming Classical Studies through Cyberinfrastructure*. Disponible en: http://www.digitalhumanities.org/dhq/vol/3/1/000028/000028.html.

Tiepmar, Jochen (2017). *Implementation and Evaluation of the Canonical Text Service Protocol as Part of a Research Infrastructure in the Digital Humanities* [Tesis doctoral inédita]. Universität Leipzig: Leipzig. Disponible en: http://asv.informatik.uni-leipzig.de/publication/file/406/Volltext___PDF_.pdf.

Zurli, Loriano y Mastandrea, Paolo (eds.) (2009). *Poesia latina, nuova E-filologia. Opportunità per l'editore e per l'interprete. Atti del convegno internazionale. Perugia 13-15 settembre 2007.* Roma: Hedder Editrice e Libreria.

# PROYECTO HDATEATROUNIR.
## REALIDADES Y PERSPECTIVAS PARA EL ESTUDIO DEL TEATRO DESDE LAS HD

Teresa Santa María Fernández, Monika Dąbrowska

**Resumen:**
Se recorre la trayectoria del grupo GHEDI – HDAUNIR, desde su origen hasta la actualidad -, reseñando los diferentes proyectos y resultados obtenidos en estos años. Para ello, se describen, en primer lugar, los diferentes proyectos realizados, como el de la *Biblioteca Electrónica Textual de Teatro Español (1968-1939)* (BETTE); el proyecto de *Visualización de redes de sociabilidad mediante grafos en el teatro español: del Siglo de Oro a la Edad de Plata* y el actual, *HDATEATROUNIR,* junto con los investigadores, procesos de trabajo y resultados de los tres proyectos. En segundo lugar, se indica las tareas que han llevado a cabo dentro del actual proyecto, que gira en dos vertientes: seguir nutriendo BETTE de textos etiquetados en TEI, con un objetivo claro de dar mayor visibilidad a las dramaturgas de la Edad de Plata; y continuar con la investigación a partir de redes sociales de los diferentes corpus que se integran en DraCor. Por último, se apuesta por una tercera vía en el futuro, que va dirigida al estudio y puesta en práctica de las posibilidades didácticas de los textos dramáticos en TEI en estudiantes y profesores de niveles de Secundaria y Bachillerato. Con

Teresa Santa María Fernández, Universidad Internacional de La Rioja (UNIR)
Orcid ID 0000-0003-0281-7754
teresa.santamaria@unir.net

Monika Dąbrowska, Universidad Internacional de La Rioja (UNIR)
Orcid ID 0000-0001-7421-0164
monika.dabrowska@unir.net

todas ellas se pretende describir la investigación llevada hasta la fecha, los problemas encontrados y resueltos, las aportaciones, así como los desafíos y retos que suponen incluir las metodologías de las Humanidades Digitales aplicadas en el estudio del teatro español de la Edad de Plata.

**Palabras clave:** Teatro de la Edad de Plata, Biblioteca Electrónica Textual, Textos en TEI, redes de sociabilidad, Transferencia en el aula preuniversitaria

**Abstract**
It is an overview of the trajectory of the group GHEDI - HDAUNIR, from its origins to present day -, reviewing its different projects and the results obtained through these years. For that, on the first place, there is a description of the different projects that have been made, such as the "Electronic Textual Library of Spanish Theatre (1968-1939)" (Biblioteca Electrónica Textual del Teatro Español, BETTE); the project "Visualization of the sociability networks in Spanish theatre through graphs: from the Golden Age to the Silver Age", and the current, "HDATEATROUNIR", with the researchers, work processes and results of all three. On the second place, it indicates the tasks that have been developed in the current project, that revolves around two aspects: in keeping providing BETTE of TEI labelled texts, with an objective of giving more visibility to Silver Age female playwrights; and in continuing with the investigation based on the social networks from the different corpus that make up DraCor. Finally, a third way on the future is envisaged, destined to the study and implementation of the didactic possibilities of dramatic pieces on TEI on students and teachers on High School and bachelor's degree levels. Within all of them the purpose is of describing the investigation made up to this date, the problems that have been found and resolved, the contributions, as well as the challenges that supposes including the methodologies of Digital Humanities to the study of Spanish theatre on the Silver Age.

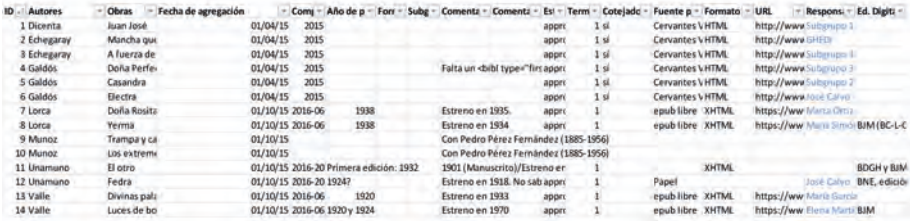
**Key words:** Silver Age Theatre, Textual Electronic Library, Texts in TEI, sociability Networks, Transfer in the pre-university classroom

Resulta interesante y oportuno explicar en las próximas páginas cómo nace, se desarrolla, continúa y hasta dónde aspira seguir uno de los proyectos sobre Humanidades Digitales que, como muchos otros, nacieron un poco por azar y un mucho por el tesón y esfuerzo de varios investigadores implicados en sacarlos adelante. Como en cualquier disciplina naciente no siempre el camino es recto ni libre de obstáculos, pero se espera que la descripción de los itinerarios iniciados sirva a otros proyectos y grupos para organizar su trabajo, pero, sobre todo, para perder el miedo a adentrarse en una vía que continúa con muchos tramos ignotos y llenos de interesantes descubrimientos dentro de las HD.

## 1. ENTRE GHEDI Y HDATEATROUNIR

A finales de 2014 se empezaron a dar los primeros pasos organizativos para que varios investigadores del Grupo de Humanidades y Edición Digital (GHEDI) se empezaran a formar en lenguaje de marcado XML-TEI (*eXtensible Markup Language-Textual Encoding Initiative*), de la mano de Juana María González, como investigadora principal y de José Calvo Tello, como formador y facilitador de todo el proceso. Nace así, un poco más tarde, en julio de 2015, el proyecto la creación de una *Biblioteca Electrónica Textual de Teatro Español (1868-1939)*, BETTE, cuya finalidad es editar y publicar en lenguaje de codificación XML-TEI y analizar con métodos digitales las obras dramáticas más relevantes de ese periodo (Martínez y Santa María, 2019; Jiménez y Calvo, 2020). El grupo continúa como GHEDI y consigue una financiación como grupo y proyecto propio de la Universidad Internacional de La Rioja (UNIR), ahora con la doctora Concepción Jiménez Fernández como su investigadora principal y continuando José Calvo Tello como apoyo formativo y organizador.

En estos primeros años son varios los investigadores que trabajaron en la edición de las veinticinco obras que se integran en el proyecto y que, por tanto, pasaron a lenguaje de marcado TEI las piezas dramáticas que se han incorporado en el corpus BETTE (https://github.com/HDAUNIR/BETTE). Aunque se especifica en



| ID | Autores | Obras | Fecha de agregación | Com | Año de p | For | Subg | Comentar | Comentar | Es | Term | Cotejado | Fuente p | Formato | URL | Responsa | Ed. Digiti |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Dicenta | Juan José | 01/04/15 | 2015 | | | | | | appr | 1 | sí | Cervantes V | HTML | http://www | Subgrupo 1 | |
| 2 | Echegaray | Mancha qu | 01/04/15 | 2015 | | | | | | appr | 1 | sí | Cervantes V | HTML | http://www | GHEDI | |
| 3 | Echegaray | A fuerza de | 01/04/15 | 2015 | | | | | | appr | 1 | sí | Cervantes V | HTML | http://www | Subgrupo 1 | |
| 4 | Galdós | Doña Perfe | 01/04/15 | 2015 | | | | Falta un <bibl type="firs | | appr | 1 | sí | Cervantes V | HTML | http://www | Subgrupo 3 | |
| 5 | Galdós | Casandra | 01/04/15 | 2015 | | | | | | appr | 1 | sí | Cervantes V | HTML | http://www | Subgrupo 2 | |
| 6 | Galdós | Electra | 01/04/15 | 2015 | | | | | | appr | 1 | sí | Cervantes V | HTML | http://www | José Calvo | |
| 7 | Lorca | Doña Rosita | 01/10/15 | 2016-06 | 1938 | | | Estreno en 1935. | | appr | 1 | | epub libre | XHTML | https://ww | Marta Ortiz | |
| 8 | Lorca | Yerma | 01/10/15 | 2016-06 | 1938 | | | Estreno en 1934 | | appr | 1 | | epub libre | XHTML | https://ww | María Simón | BJM (BC-L-C |
| 9 | Munoz | Trampa y ca | 01/10/15 | | | | | Con Pedro Pérez Fernández (1885-1956) | | | | | | | | | |
| 10 | Munoz | Los extrem | 01/10/15 | | | | | Con Pedro Pérez Fernández (1885-1956) | | | | | | | | | |
| 11 | Unamuno | El otro | 01/10/15 | 2016-20 | Primera edición: 1932 | | | 1901 (Manuscrito)/Estreno en | | | 1 | | | XHTML | | | BDGH y BJM |
| 12 | Unamuno | Fedra | 01/10/15 | 2016-20 | 1924? | | | Estreno en 1918. No sab | | appr | 1 | | Papel | | | José Calvo | BNE, edició |
| 13 | Valle | Divinas pal | 01/10/15 | 2016-06 | 1920 | | | Estreno en 1933 | | appr | 1 | | epub libre | XHTML | https://ww | María García | |
| 14 | Valle | Luces de bo | 01/10/15 | 2016-06 | 1920 y 1924 | | | Estreno en 1970 | | appr | 1 | | epub libre | XHTML | https://ww | Elena Martí | BJM |

Figura 1: Captura de parte del Excel con algunos datos que se van incluyendo sobre el corpus de BETTE

cada una de las piezas quienes realizaron dicha labor de marcado de texto, no está de más poder citarlos a continuación, por orden alfabético de apellido: María Aboal López, José Calvo Tello, Mireya Fernández Merino, María García Sánchez, Sila María Gómez, Juana María González, Concepción Jiménez Fernández, Elena Martínez Carro, Roxana Beatriz Martínez Nieto, Marta Ortiz, Julio Peiró, Ignacio Roldán, María Teresa Santa María, María Simón, María Ángel Somalo y Rocío Vilches. A este grupo inicial se unió, a finales de 2016 Guadalupe Frutos.

Para la realización del trabajo se efectuaron varias tareas previas. En primer lugar, resultaba necesario la elección de las obras que se iban a digitalizar en TEI. Se optó por incluir las de aquellos dramaturgos que se incorporan en monografías como las de Ruiz Ramón (1986; 1996) y Huerta (2003) y cuyos derechos de autor no estuvieran vigentes, porque ya habían transcurrido ochenta años o más desde su muerte. También se delimitó a que tuvieran piezas dramáticas publicadas en ese periodo que va desde 1868 a 1936 y, que se ha ampliado, como veremos más adelante, hasta 1939 a lo largo del proyecto.

Con todos estos datos se elaboró un Excel (Figura 1) que recoge todos los datos relevantes, donde se incluye: un número identificador (ID) de cada pieza, el autor, la obra, la fecha en que se incorporó en este listado y en qué fase del proyecto está previsto su digitalización. Además, se incorpora información sobre cada obra para facilitar su inclusión en el *TeiHeader* de las obras editadas: año

| AUTORES | OBRAS |
|---|---|
| GARCÍA LORCA | Doña Rosita<br>Yerma<br>La zapatera prodigiosa<br>Bodas de sangre<br>La casa de Bernarda Alba |
| VALLE-INCLÁN | Divinas palabras<br>Luces de bohemia<br>Águila de Blasón<br>Romance de Lobos<br>Cara de Plata |
| MUÑOZ SECA | Los pergaminos<br>El refugio<br>La conferencia de Algeciras<br>Usted es Ortiz |
| VALERA | Asclepigenia<br>La venganza de Atahualpa |
| UNAMUNO | El otro<br>Fedra |
| ECHEGARAY | Mancha<br>A fuerza de arrastrarse |
| CLARÍN | Teresa |
| DICENTA | Juan José |
| GALDÓS | Electra<br>Casandra<br>Doña Perfecta |

Figura 2. Especificación de obras en XML-TEI y sus autores, editadas por BETTE
Fuente: (Martínez y Santa María, 2019)

de publicación, de estreno, cuándo se ha finalizado su digitalización, procedencia del texto que va a ser ser procesado o la edición digital – en HTML o epub – que ha servido como base para la edición en TEI. También se aportan datos sobre quién es el responsable de dicha digitalización y dónde se encuentra – Biblioteca de la Fundación Juan March o Biblioteca Digital Hispánica de la Biblioteca Nacional de España – la primera edición tras el estreno con la que se realizó la revisión filológica antes de editar cada obra en formato digital y en las que se actualiza únicamente la ortografía según la edición de 2010 de la Real Academia Española.

De esta forma, quedaron seleccionadas, para la primera fase BETTE y para su marcado en TEI veinticinco obras de nueve autores diferentes y que cultivaron géneros dramáticos distintos: Leopoldo Alas Clarín, Joaquín Dicenta, José Echegaray, Benito Pérez Galdós, Federico García Lorca, Pedro Muñoz Seca, Miguel de Unamuno, Juan Valera y Ramón del Valle-Inclán. En total, los investigadores editaron con TEI veinticinco obras (Figura 2).

Una segunda tarea inicial relevante fue la formación de los investigadores en el lenguaje de marcado en TEI, a través de

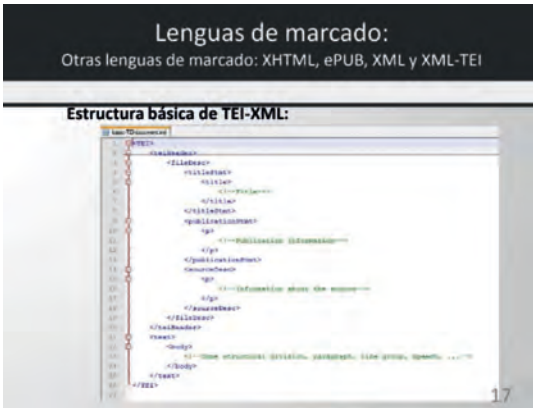



Figura 3: Pantallazo de una de las formaciones impartidas por José Calvo Tello para el proyecto BETTE

una serie de sesiones y material realizados por José Calvo Tello (Figura 3). De esta manera, se impartió un primer seminario sobre lenguajes de marcado, entre noviembre y diciembre de 2014 dividido en cuatro sesiones diferentes: la primera sobre "Introducción al HTML", la segunda sobre "Elementos de HTML", la tercera se refería a la "Estructura de un archivo HTML y CSS: creando nuestra primera web"; y la cuarta sobre "Otras lenguas de marcado"[1].

En las cuatro sesiones de un segundo seminario, que se llevó a cabo en julio de 2015, José Calvo Tello se centró en la formación más específica sobre la aplicación de TEI a los textos dramáticos que se van a digitalizar y versaron sobre: "TEI para teatro", "Estructura general y validación", "Aspectos prácticos al editar TEI" y "Expresiones regulares para TEI". Por último, un tercer seminario llevaba el título general de "¿Qué hacemos con los textos?" y fue impartido también por Calvo Tello, en mayo de 2016. En esta ocasión el seminario tenía como objeto la formación en: "CSS", "xPath", "Grafos" y "GitHub". Un poco antes, en febrero

1. El material y las grabaciones pueden solicitarse contactando con el proyecto HDATEATROUNIR a través de la web del Grupo de investigación Humanidades Digitales Aplicadas (HDAUNIR) http://gruposinvestigacion.unir.net/hdaunir/

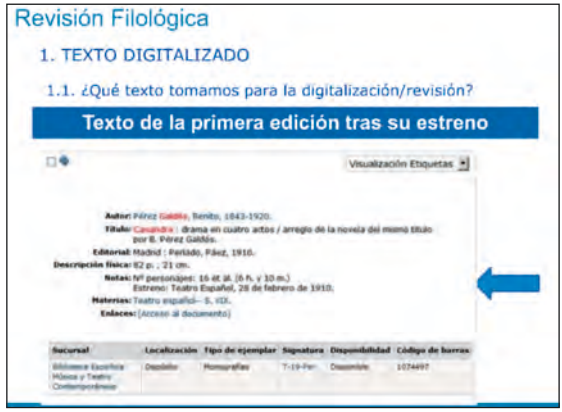



Figura 4. Pantallazo de la formación sobre "Revisión filológica"

de 2016, Teresa Santa María dio una sesión formativa sobre "Revisión filológica" (Figura 4).

En tercer lugar y en paralelo con estas formaciones los investigadores iniciaron la edición en lenguaje de marcado TEI de las veinticinco obras aludidas antes. El proceso de trabajo respondió al siguiente flujo de tareas:

— Primera fase: todos realizan una escena de la misma obra, en este caso, *Mancha que limpia* de José Echegaray.
— Segunda fase: divididos en grupos de tres investigadores, se realiza la digitalización de dicha obra para su marcado entre los diferentes grupos.
— Tercera fase: cada investigador lleva a cabo la digitalización de una obra diferente.

Esta distribución permitió trabajar desde el principio con dicha edición en TEI, con la ventaja de que los textos venían en una versión más "limpia" que era preparada previamente por José Calvo Tello a partir de un texto en HTML o ePub. También este investigador se responsabilizó de la tarea de aclarar dudas y revisar la versión final, que ya se consideraba la idónea. Con todo

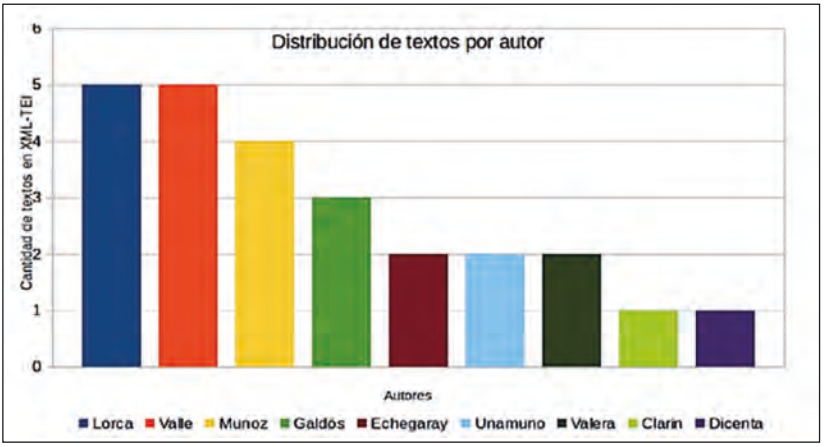



Figura 5: Distribución de textos teatrales según autores. Fuente: (Martínez y Santa María, 2019)

este proceso de trabajo se disponían en septiembre de 2017 de las veinticinco obras antes aludidas (Figura 2). El reparto de las obras por dramaturgo no es homogéneo, puesto se puede observar que algunos autores solo tienen una obra en TEI dentro de BETTE – como Clarín y Dicenta -, tres dos piezas – Echegaray, Unamuno y Valera -; uno tres piezas – Galdós – o cuatro – Muñoz Seca – y dos autores cinco piezas – García Lorca y Valle-Inclán - (Figura 5).

Además de la formación, los textos previamente preprocesados por Calvo Tello y el flujo de fases para preparar edición en lenguaje de marcado, se cuenta con un esquema propio de TEI, específico para las obras teatrales del corpus BETTE y un modelo de obra comentado, *Electra* de Galdós, donde se explica cada elemento. Ambos han sido realizados por José Calvo Tello.

Junto con todo este material de formación, esquema TEI y texto modelo de *Electra*, se ha incluido en una carpeta Dropbox todos los textos necesarios para que cada investigador pueda marcar sus obras en TEI. De este modo, existen cinco carpetas dentro del apartado "Corpus" (Figura 6) para las diferentes fases de los textos que se van a editar. En la carpeta "0- digitalizaciones"




Figura 6: División del corpus de BETTE en las diferentes fases de su digitalización

están los textos en PDF o Word y que constituyen la primera publicación tras el estreno, que han proporcionado las bibliotecas de la Fundación Juan March o de la Biblioteca Digital Hispánico; en la "1-descargados" están las obras de teatro en HTML o ePub extraídos, por ejemplo, de la Biblioteca Virtual Cervantes, en "2-preprocesados" se encuentran los textos en una primera versión XML, limpiados por Calvo Tello; en "3-procesados" aparecen las piezas dramáticas en TEI, pero sin haberse revisado filológicamente con la primera publicación de la obra tras el estrenos. Los textos, ya revisados y listas para su publicación en abierto, se incluyen en "4-finalizados".

Las veinticinco obras procesadas están libres de derechos de autor decantándonos por la variedad de temas, subgéneros teatrales o estilos de la dramaturgia española y sus cambios durante los casi setenta años que abarca la llamada Edad de Plata del teatro español (1868-1936), tomando el sentido amplio de dicho periodo, como han hecho otras estudiosas al vincularlo con el término de "Modernidad" (Romero López, 2014 y 2015; Thion Soriano-Mollá, 2021). Este corpus de teatro ha sido revisado filológicamente para su edición en TEI con miras a dar homogeneidad a todos los textos según unas directrices comunes, partiendo, como se ha comentado antes, de la primera edición

tras el estreno y solo actualizando el texto según los criterios de la *Ortografía* de la RAE de 2010. Por otro lado, entre los dramaturgos seleccionados están autores del Realismo y Naturalismo: Galdós, Valera y Clarín y los renovadores de la escena española a principios del siglo XX: Valle-Inclán, García Lorca y Unamuno; autores de comedias: Muñoz Seca o los de teatro de la Restauración: Echegaray y Dicenta (Arana, Santa María y Simón, 2015).

Otra cuestión relevante a tener en cuenta estriba en las etiquetas y categorías que se incluyeron en el marcado de obras en TEI dentro de BETTE. En una primera fase se incorporaron las más estructurales – división de escenas y actos, acotaciones e intervenciones de los personajes, *dramatis personae* u otros datos previos -; mientras que, en una segunda fase, tras la elaboración de los resúmenes de cada pieza a cargo del responsable o de los responsables, se enriquecieron dichos datos, aportando información relativa al sexo, papel, naturaleza o importancia de los personajes. Todas estas obras se pueden consultar y descargar, porque están en abierto, desde el repositorio de GitHub del grupo (https://github.com/GHEDI/BETTE). Para ello, es necesario entrar en la carpeta "Corpus" y se pueden abrir y descargar, a continuación, bien en formato TEI; bien en TXT cada una de las obras.

Todos estos procesos y flujos de trabajo se han mantenido en el segundo y tercer proyectos financiados por UNIR para poder seguir incrementado con nuevas obras el corpus BETTE, y con Teresa Santa María como investigadora principal. En efecto, a partir de septiembre de 2018, el grupo GHEDI queda rebautizado como HDAUNIR y dentro de él se incluyen propuestas no solo relativas a BETTE y sus proyectos de digitalización teatrales siguientes, sino otras relativas a análisis en redes sociales como medio de comunicación social – liderada por Julio Montero - o el uso de dichas redes sociales en el ámbito educativo, con un proyecto dirigido por Raquel Gil y Diego Calderón. Pero volviendo al proyecto BETTE, en 2018 se reestructura el grupo y se mantiene el objetivo de seguir digitalizando dentro de BETTE – que queda

como nombre de la biblioteca de textos en TEI – como uno de los que integra el segundo proyecto de investigación. Este proyecto, *Visualización de redes de redes de sociabilidad mediante grafos en el teatro español: del Sigo de Oro a la Edad de Plata*, consigue financiación de la UNIR para el bienio 2018-2020 y añade la dimensión investigativa de las redes sociales de personajes de las obras de la Edad de Plata y de autores dramáticos del Siglo de Oro como tema de estudio. En este proyecto participan, siguiendo el orden alfabético del primer apellido: Monika Dabrowska, Guadalupe Frutos, Concepción Jiménez, Elena Martínez Carro, Teresa Santa María, María Ángel Somalo y Alejandra Ulla. Ya en el segundo proyecto, HDATEATROUNIR, financiado para 2020-2022, se incorporan Javier Jacobo González y David Merino.

Asimismo, las veinticinco piezas aparecen bajo el nombre de *Spanish Drama Corpus* (https://dracor.org/span) en el *Drama Corpora Project* (*DraCo*r). Este proyecto está liderado por el profesor Frank Fischer, director de DARIAH (Digital Research Infraestructure for the Arts and Humanities). Dentro de DraCor (Fischer et al, 2019a), proyecto vinculado al Centre for Digital Humanities at HSE (Moscow) y a la Universität Potsdam (Alemania), se encuentras otras colecciones de teatro europeo, bien divididas por lenguas como ruso, latín, griego clásico, sueco, alemán, italiano, tártaro y alsaciano; bien vinculadas a algunos dramaturgos en concreto, como es el caso de los corpus de William Shakespeare y de Calderón de la Barca. Además de proporcionar los textos en TEI, DraCor enlaza las obras teatrales y sus autores con Wikidata y permite una investigación en abierto, así como la colaboración entre diferentes interesados y proyectos. De este modo, DraCor ofrece al proyecto BETTE mayor visibilidad y difusión a nivel internacional y hace posible la comparación de datos entre corpus teatrales de distintos países. Del mismo modo, para la enseñanza en HD, DraCor aporta herramientas interesantes, novedosas e incluso divertidas con las que se puede trabajar fácilmente (Fischer, Santa María, Calvo-Tello y Milling, 2021)) y cuenta con *Shiny DraCor* (https://shiny.dracor.org/) para generar los

grafos de dichas obras dramáticas sin tener que conocer y manejar otros programas informáticos para crearlos.

BETTE cuenta en la actualidad con estos textos en abierto, dentro del GitHub propio y de la base de datos de DraCor y, además, otras siete obras en TEI, que no pueden ponerse a la disposición de todos los investigadores, porque aún tienen vigentes sus derechos de autor. En concreto, están ya preparadas las ediciones de: *Espejo de avaricia,* de Max Aub, *Rosas de otoño,* de Jacinto Benavente, *Usted tiene ojos de mujer fatal*, *Cuatro corazones con freno y marcha atrás* y *Un marido de ida y vuelta,* de Enrique Jardiel Poncela, *Tres sombreros de copa,* de Miguel Mihura, y *La esfinge* de Unamuno[2]. El criterio de selección de estas obras y dramaturgos viene establecido por la inclusión en manuales de referencia de esta área (Figura 10) y ese mismo criterio se ha seguido para elegir las piezas que editarán en lenguaje de marcado TEI dentro de esta tercera fase o proyecto – HDATEATROUNIR. En efecto, en esta etapa, hemos primado la necesidad de poder contribuir a la estructura de textos para el estudio y trabajo de otros investigadores (Fischer et al. 2019b), con textos que se pudieran dejar en en abierto, a la vez que pretendíamos dar una mayor visibilidad a dramaturgas e incluirlas, por mérito propio en BETTE y en *Spanish Drama Corpus* de DraCor.

## 2. CONSOLIDACIÓN DEL CORPUS

En cuanto a la alimentación del corpus con nuevas obras, en la etapa actual del proyecto cobran protagonismo las dramaturgas españolas de este periodo. Puesto que no aparecía en esta biblioteca electrónica ninguna obra de autoría femenina, dentro de la Edad de Plata, se ha querido subsanar esta carencia y completar BETTE con el teatro de las creadoras de dicha época. Considerando el papel de la

2. Estos textos en TEI se pueden solicitar contactando con cualquiera de los investigadores del proyecto HDATEATROUNIR que se incluyen dentro del grupo HDAUNIR: https://gruposinvestigacion.unir.net/hdaunir/quienes-somos/

mujer en la escena española y vida teatral en sentido amplio, unido a la reconfiguración del su rol social y cultural en la España del inicio del siglo XX, parecía más que justificado rescatar sus aportes a la renovación escénica y darles más presencia y divulgación a las autoras de este momento histórico, por lo general ausentes en los libros de texto correspondientes. Asimismo, parece relevante facilitar el acceso libre a estas obras desde *Spanish Drama Corpus*, puesto que en el corpus internacional DraCor, de momento, hay únicamente tres colecciones, la alemana, la alsaciana y la sueca, que contienen piezas de escritoras dramáticas.

Tras cribar en la lista de autores que pasaban a dominio público se decidió ampliar el alcance del corpus hasta 1939 y centrarnos en los próximos años en incluir dentro de BETTE algunas de las doce dramaturgas cuyas obras pueden ponerse en abierto:

| | | | |
|---|---|---|---|
| Acuña, Rosario (1850-1923) | Enlace al catálogo | Autor en BDH | Autor en Datos.BNE.es |
| Calé y Torres de Quintero, Emilia (1837-1908) | Enlace al catálogo | Autor en BDH | Autor en Datos.BNE.es |
| Contreras de Rodríguez, María del Pilar (1861-1930) | Enlace al catálogo | Autor en BDH | Autor en Datos.BNE.es |
| Coronado, Carolina (1820-1911) | Enlace al catálogo | Autor en BDH | Autor en Datos.BNE.es |
| Díaz de Mendoza y Aguado, Carmen, Condesa de San Luis (1864-1929) | Enlace al catálogo | Autor en BDH | Autor en Datos.BNE.es |
| Freixas, Narcisa (1859-1926) | Enlace al catálogo | Autor en BDH | Autor en Datos.BNE.es |
| García Balmaseda de González, Joaquina (1837-1911) | Enlace al catálogo | Autor en BDH | Autor en Datos.BNE.es |
| Monserdà de Macià, Dolors (1845-1919) | Enlace al catálogo | Autor en BDH | Autor en Datos.BNE.es |
| Muñiz y Más, Adelaida m. 1906 | Enlace al catálogo | Autor en BDH | Autor en Datos.BNE.es |
| Pardo Bazán, Emilia Condesa (1851-1921) | Enlace al catálogo | Autor en BDH | Autor en Datos.BNE.es |
| Sáenz de Tejada, Victorina (1841-1909) | Enlace al catálogo | Autor en BDH | Autor en Datos.BNE.es |
| Velilla y Rodriguez, Mercedes (1852-1918) | Enlace al catálogo | Autor en BDH | Autor en Datos.BNE.es |

Figura 7. Listado de las dramaturgas en españolas de la Edad de Plata que pasaron a dominio público en 1938. Elaboración propia, a partir del catálogo de la BNE: https://datos.gob.es/es/catalogo/e00123904-autores-espanoles-en-dominio-publico-en-2019-fallecidos-en-1938

De entre todas ellas, se han elegido, para el bienio 2020-2022, a dos autoras, Emilia Pardo Bazán y Rosario Acuña ya que en ellas confluían dos características que las hacían idóneas. En primer lugar, porque se cumplía una efeméride para ambas, que podría ayudar a su mayor difusión. En efecto, en 2021 y 2023 se conmemora el centenario del fallecimiento de una y el nacimiento de otra, respectivamente. En segundo, lugar, porque nos permitirá comprobar la influencia en el futuro de dos autoras que han recibido, dentro de los manuales literarios y de teatro, un trato desigual. Mientras Emilia Pardo Bazán sí aparece en uno de ellos como dramaturga (Pedraza Jiménez y Rodríguez Cáceres, 1983: 783-784), Rosario Acuña no es referenciada en ninguno de los volúmenes de dicho manual ni en los dos más específicos sobre teatro de Ruiz Ramón (1986 y 1996). Incidimos, de este modo, en las posibilidades que ofrecen las HD de dar visibilidad y alcance a autores no tan valorados hasta la fecha, como se constata en proyectos *Mnemosine*, vinculado por el periodo de la Edad de Plata al que se circunscribe; y, dentro del ámbito de la mujer, *Bieses* (https://www.bieses.net/) y la iniciativa en Wikipedia de *Editatón* (Editatón. https://es.wikipedia.org/wiki/Editat%C3%B3n), por poner solo algunos ejemplos.

En concreto, las obras de dramaturgas femeninas en las que están trabajando Monika Dabrowska, Guadalupe Frutos, David Merino, Teresa Santa María y María Ángel Somalo, con el asesoramiento de Paloma Centenera, son las siguientes:

| | |
|---|---|
| **Emilia Pardo Bazán** | *El vestido de boda* |
| | *La verdad* |
| | *La suerte* |
| | *Cuesta abajo* |
| | *Las raíces* |
| | *El becerro de metal* |
| | *Juventud* |
| **Rosario Acuña** | *El padre* |
| | *Amor a la patria* |
| | *Tribunales de venganza* |
| | *Rienzi el tribuno* |

Figura 8: Listado de obras de dramaturgas que se incluirán en BETE (Fase 2020-2022)

A ellas hay que añadir otros textos dramáticos de autores que ya están en BETTE, pero que sirven para completar el corpus específico de estos dramaturgos, entre los que están Federico García Loca, José de Echegaray y Pedro Muñoz Seca:

| | |
|---|---|
| **Federico García Lorca** | *Los títeres de la cachiporra* |
| | *El público* |
| | *Mariana Pineda* |
| **José de Echegaray** | *El gran galeote* |
| **Pedro Muñoz Seca** | *La venganza de don Mendo* |

Figura 9: Listado de obras de dramaturgos ya presentes en BETE (Fase 2020-2022)

En el caso de estos autores se ha primado su presencia en los citados manuales de literatura y teatro, para intentar que el corpus recogiera los más representativos y referenciados en

estos y mejorar la correlación que Calvo Tello realizó teniendo en cuenta solo los veinticinco textos que están actualmente en abierto y donde solo cuatro autores superarían el 0,0500:

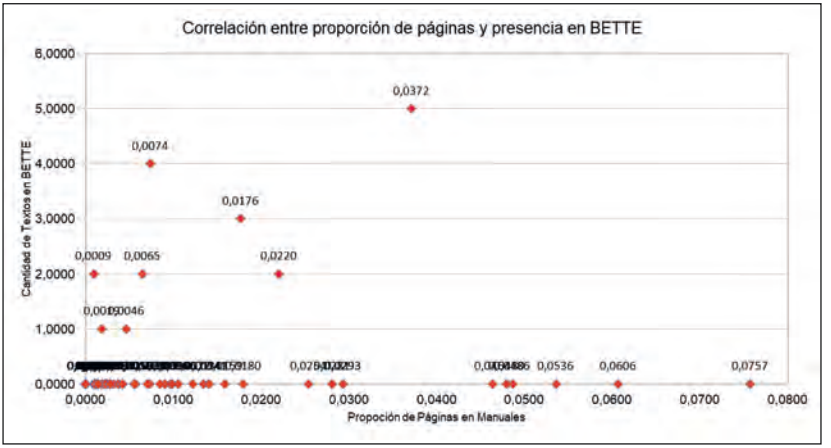



Figura 10. Fuente: Calvo-Tello

Nos parece que incluir autores más o menos canónicos contribuye a dar mayor visibilidad a algunos de ellos, que solo cuentan con ediciones en PDF en la Biblioteca Digital Hispánica y que, de esta manera pueden ser objeto de estudios futuros; a la vez que se proporcionan nuevas herramientas digitales para aquellos que se han considerado más canónicos y han sido objeto de mayor interés por parte de los investigadores.

Por último, conviene reseñar que, durante esta fase, la colaboración Paloma Centenera, que ha venido a sustituir a José Calvo Tello como asesora técnica del grupo, ha contribuido a continuar con las labores de marcado en TEI iniciadas, a la vez que se ha seguido con unas sesiones de formación que, al igual que las anteriores, constan de grabaciones y material que se puede solicitar contactando con el proyecto HDATEATROUNIR. En esta ocasión, Centenera incidió durante las dos sesiones que impartió en mostrarnos cómo covertir de texto en PDF a un texto XML, que se podía iadaptar a nuestro esquema TEI. De este

modo, todo el proceso previo – el paso de PDF a texto TEI – y la posterior revisión filológica, comparando el texto resultante con la primera edición tras el estreno de cada obra, queda resumido en este esquema:




Figura 11. Flujo de trabajo de BETTE-HDATEATROUNIR. Fuente: Elaboración propia

Un punto aparte en nuestro trabajo y formación son los Seminarios de Humanidades Digitales que organizamos en colaboración con el programa de Doctorado y el Máster en Humanidades Digitales de la UNIR. Además de instrucción y actualización para los miembros del grupo de investigación, es una forma de estrechar las relaciones con los académicos de otros grupos e intercambiar las experiencias. Durante el curso 2019-2020 han tenido lugar cuatro sesiones:

— Presentación de la *Biblioteca Electrónica Textual de Teatro en Español 1868-1939* (BETTE), por Concepción Jiménez (22 de octubre de 2019)
— Humanidades Digitales: entre Investigación e Infraestructura, por Frank Fischer, DARIAH (22 de octubre de 2019)
— Humanidades Digitales. Conceptos, logros consolidados y perspectivas de futuro, por Sagrario López Poza (12 de noviembre de 2019)

El curso siguiente, 2020-2021 otras dos sesiones:

— Una visión panorámica de la Tecnología aplicada a la Música (MIR), por Francisco Gómez, UCM (2 de abril de 2020)
— Hegel y Nietzsche desde el análisis computacional NLP, por Manuel Cebral, Tecnológico de Monterrey (24 de noviembre de 2020)

En el año académico en curso, 2021 -2022 se ha iniciado con la sesión "Nuevas luces entre candilejas: Emilia Pardo Bazán, una dramaturga innovadora" impartida por Dolores Thion Soriano-Mollá, de la Universidad de Pau et des Pays de l'Adour. Gabriela Baeza Ventura y Carolina A. Villarroel, del Centro de Humanidades Digitales (USLDH) de la University of Houston, han compartido su experiencia sobre la "Recuperación del legado escrito de latinas y latinos en los Estados Unidos".

## 3. NUEVAS LÍNEAS DE TRABAJO: USOS DIDÁCTICOS DEL PROYECTO BETTE/HDAUNIR

Una vez descritos los inicios del proyecto y su estado presente, llega el momento de plantearse su recorrido futuro, a corto y largo plazo. En este horizonte destacan tres vías de trabajo principales. Por un lado, la vía de digitalización, es decir, la necesidad de consolidar el corpus, completándolo con más obras de autores ya incluidos y con dramaturgos nuevos. En segundo lugar, la vía de la investigación para ahondar en conocimiento de su producción teatral con los métodos de estudio propios de las Humanidades Digitales. Por último, la vía de la educación, para dar a esos avances también el uso didáctico, de modo que sirvan no solo para académicos e investigadores, sino también para profesores y estudiantes, incluso de las etapas de educación preuniversitaria. De hecho, las tres líneas se retroalimentan, porque un mayor número de obras en la plataforma aumenta las posibilidades de su estudio

y de una exploración didáctica de los textos literarios y, viceversa, un mayor número de usuarios con diversos intereses científicos y didácticos permite enriquecer el corpus con funcionalidades adaptadas a estas necesidades lectoras e investigativas. Los resultados obtenidos de la investigación, a su vez, redundarán en variedad de aplicaciones didácticas, tanto en campo literario, como en otras disciplinas humanísticas.

Ante la creciente digitalización del mundo y de los propios estudios humanísticos, en nuestro trabajo no queremos perder de vista los retos educativos que se plantean para la enseñanza de la literatura en las aulas escolares y, por supuesto, a nivel académico. En la era de las pantallas, es necesario preguntarse por los nuevos modelos de enseñanza y por las habilidades y destrezas que necesitarán los alumnos, habituados al uso de herramientas computacionales. ¿Cómo formar a los futuros humanistas digitales y aplicar los medios computacionales a la educación de filólogos, maestros, espectadores teatrales, actores y profesionales de escena y jóvenes en general? ¿Cómo responder a las nuevas exigencias educativas y pedagógicas de la sociedad digital? ¿Cómo tener en cuenta los aspectos educativos a la hora de desarrollar archivos o bibliotecas digitales? Ante estas preguntas, que en el fondo miran hacia el futuro de la disciplina, nos hemos propuesto explorar la vertiente educativa de nuestro proyecto de Humanidades Digitales, y en concreto las potencialidades de nuestro corpus a la hora de enseñar el teatro español del fin de siglo XIX y comienzos del XXI.

Convencidos de la importancia de los métodos e instrumentos propios de las Humanidades Digitales para el futuro de las disciplinas humanísticas y la innovación educativa, ha surgido de forma natural la reflexión sobre nuestra propia docencia, dado que varios miembros del grupo la ejercen a nivel de educación superior y preuniversitaria. De allí surgieron propuestas de los posibles usos didácticos de nuestra biblioteca digital y ejemplos de su aplicación al aula para acercar a los estudiantes al teatro de la Edad de Plata. Así hemos iniciado esta nueva vía de trabajo didáctico.

En primer término, nos hemos acercado a los trabajos sobre la pedagogía digital, buscando bases teóricas, pautas y estrategias para potenciar la competencia digital y crítica, así como las experiencias de aula relacionadas con la lectura computacional y uso didáctico de archivos digitales. Como era de esperar, el avance de las herramientas tecnológicas al servicio de las Humanidades Digitales propicia una forma distinta de acercarse a las obras teatrales en todas sus dimensiones: la textual, sociohistórica, escénica, actoral, escenográfica y, por supuesto, abre nuevas posibilidades de tratarlas dentro de una clase de literatura.

Las oportunidades que ofrece la visualización de datos dinámica, la realidad aumentada y entornos 3D, los contextos de inmersión y simulación sin duda facilitan el diseño de nuevas experiencias educativas y enriquecerán los aprendizajes. Los ejemplos de actividades didácticas diseñadas por los grupos de investigación y cuerpos académicos del ámbito anglosajón dan muestra de las nuevas tendencias en la didáctica digital aplicada a las áreas humanísticas y resultan relevantes para abrir nuevos horizontes en innovación educativa. Las abundantes metodologías surgidas recientemente en este campo, como Object Based Learning (OBL), Objects Interactive Learning, Virtual Reality Simulation-Based Learning, Open Educational Resources (OER) u OER Enabled Pedagogy, etc. dan fe de los multiformes esfuerzos por llevar estos adelantos a las aulas e involucrar a los estudiantes en el uso de las aplicaciones interactivas para generar nuevos conocimientos. El panorama de las posibilidades es francamente enorme y, a la vez, muy ligado a las potencialidades de las propias plataformas e instrumentos integrados. Para empezar nuestra andadura, nos hemos querido centrar en los procedimientos que posibilitan la visualización de datos mediante grafos y el procesamiento y análisis digital del texto (*close reading*, *distant Reading*, *text analysis*).

De esta forma, una de las primeras líneas de investigación y de aplicación docente fue la visualización mediante grafos para el estudio de las redes sociales de personajes de las obras teatrales.

El sistema de grafos es una representación gráfica de los datos numéricos extraídos, en este caso de la obra literaria, mediante una red de nodos – personajes (visualizados como puntos) y aristas (líneas de relaciones entre estos nodos) que en su conjunto representan la trama de una obra teatral. Tras los acercamientos más académicos al tema, en forma de comunicación de Santa María y Somalo al III Congreso Internacional de Lingüística Computacional y de Corpus (Bogotá, 2020) y de Santa María en el 2nd International Conference of the European Association for Digital Humanities (EADH, Krasnoyarsk, 2021), así como los artículos de Martínez Carro, (2019) o Jiménez-Fernández y Calvo Tello (2020), nos hemos dado cuenta que esta forma gráfica e interactiva de sintetizar la estructura de una obra es un formidable recurso didáctico para explicar la renovación de la escena española a principios del siglo XX y mostrar de forma visual dichas innovaciones teatrales a los alumnos adolescentes que estudian en clase el movimiento modernista.

El primer intento de una propuesta estrictamente didáctica fue presentado en el III Congreso Internacional Educación, Investigación y Tecnología "Innovación educativa y uso de las TIC" (Bogotá, 2021), con el trabajo "Los grafos y la edición digital. Propuesta para enseñar Literatura en Secundaria y Bachillerato". La aportación ejemplifica las posibilidades de trabajo con grafos para comprender la configuración de la red de personajes de una obra teatral y las relaciones entre ellos y brinda una serie de actividades didácticas para comprender la estructura de un corpus de obras digitales, como es *Spanish Drama Corpus* dentro de DraCor (https://www.dracor.org/span). La siguieron otras comunicaciones encaminadas a familiarizar a los jóvenes con el estudio de obras teatrales a través de método de grafos, "Herramientas digitales de visualización de redes de personajes como recurso didáctico para analizar las obras del teatro en el aula de literatura", presentada en el XXII Congreso Internacional de la Sociedad Española de Didáctica de la Lengua y la Literatura (SEDLL, Cuenca, 2021), "Los personajes de *Bodas de sangre* de García Lorca desde la perspectiva digital" y

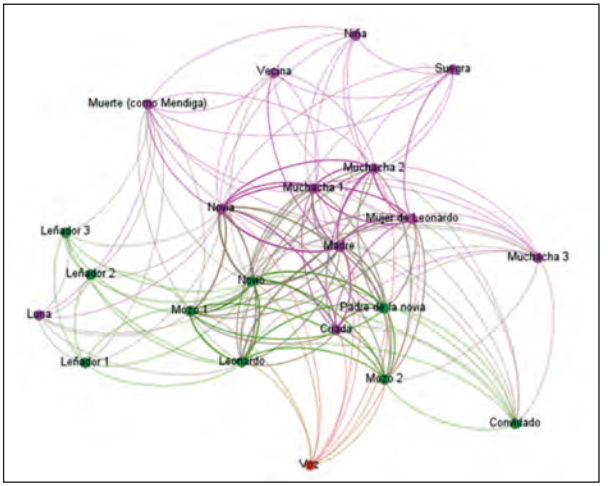



Figura 12. Grafo de *Bodas de sangre* de García Lorca creado con Gephi 0.9.2

"De la percepción del alumno a la percepción del profesorado: sobre la didáctica de *La casa de Bernarda Alba* en las aulas de 4 º de la ESO", expuestas en el XV Simposio Internacional SEDLL (Braga, 2022).

Por otro lado, nos parecía provechoso acercar a los estudiantes de Secundaria y Bachillerato al análisis computacional del texto, como una manera diferente de trabajar el análisis literario, que puede resultar más atractiva al alumnado joven, acostumbrado a la utilización de las TIC. Para ello se han esbozado algunas actividades y ejercicios prácticos para que los dicentes puedan llevar a cabo un análisis semántico y cuantitativo del léxico de las obras que están en nuestro corpus, pero también cualquier otra lectura escolar o texto de interés, con la herramienta *Voyant Tools*, una plataforma gratuita y que los estudiantes preuniversitarios pueden manejar de forma intuitiva. Quizá un ejercicio de análisis de una lectura a través de un software informático, atendiendo a los datos cuantitativos, pueda parecer un acto mecánico, sin embargo, se trata de mostrar su potencialidad para lograr descubrir los significados de la obra y el rol activo (e interactivo) del intérprete. Las prácticas con *Voyant Tools* suponen una ocasión para enriquecer los resultados interpretativos tradicionales.




Figura 13. Visualización de estadísticas textuales de *La casa de Bernarda Alba* de Federico García Lorca con *Voyant Tools*

Este programa incorpora, además, varias posibilidades de visualización de los datos textuales, generando gráficos que ayudan a explorar e interpretar el texto y su significado (grafos, nubes de palabras, dendrogramas, histogramas, etc.). Las representaciones visuales de información son de gran relevancia en la era del tratamiento de datos, por lo que conviene acercarse de forma reflexiva, propiciando el pensamiento crítico acerca de los resultados obtenidos, su interpretación y posibles sesgos. La visualización aplicada al proceso analítico e interpretativo de un texto literario estimula la implicación de los alumnos en la generación de preguntas y formulación de hipótesis de investigación, llegando a constituir una hermenéutica digital, complementaria a la tradicional, y novedosa en la forma de realizar los ejercicios de comentario de texto. En esta línea, hemos presentado una propuesta al mencionado III Congreso Internacional Educación, Investigación y Tecnología en Bogotá, que puede servir de guía para abordar en el aula en análisis digital de *La casa de Bernarda Alba* con *Voyant Tools* (Dabrowska, 2021), u otro diseño didáctico para explorar en clase otra obra lorquiana, titulado "Posibilidades de visualización digital de *Yerma*. Una propuesta de aula" (Dabrowska, 2022).

El tercer tipo de diseño pedagógico que estamos iniciando tiene que ver con la edición académica en el lenguaje de

marcado XML-TEI y está destinado a capacitar a los estudiantes universitarios en la codificación de textos teatrales según los estándares de *Textual Encoding Initiative* (TEI). Este lenguaje se considera un estándar para el marcado de textos de Humanidades y Ciencias Sociales. El consorcio TEI se encarga de su normalización mediante Directrices (*Guidelines for Electronic Text Encoding and Interchange*), actualmente en su versión P5 (2021), accesibles en inglés. La consistencia de TEI como lenguaje de marcado radica en la posibilidad de etiquetar otras capas de significado como metadatos, por ejemplo, las características semánticas o estructurales del texto, que posteriormente posibilitan una exploración informática. En otros términos, en el caso de una obra teatral, permite al editor marcar, aparte del contenido de la obra, su división interna en actos o escenas, detalles de los personajes, lugares, las ediciones o escenificaciones y muchas otras.

Esta cualidad precisamente hace a los autores como Allés y del Río (2018) y Allés (2019) abogar por las posibilidades educativas de la edición en TEI, puesto que introduce una perspectiva distinta de lectura y estudio de texto literario, ya no centrada únicamente en el desarrollo de la acción dramática (en el caso de obras teatrales), sino en sus características estructurales, en los conceptos propios del género, o en la conveniencia - o no - de marcar un tipo de atributos de los protagonistas, de los espacios, etc. Los estudiantes, a la hora de editar un texto en TEI, necesitan reflexionar previamente sobre la vinculación entre el contenido y el aspecto formal del texto y reparar en el proceso de etiquetado que van a realizar. Deben decidir y ponerse de acuerdo sobre qué conviene anotar y mediante qué atributos, estableciendo un conjunto de etiquetas que posibiliten futuros estudios sobre este texto. Todo ello tendrá implicaciones no solo a la hora de etiquetar, sino para las ulteriores investigaciones de la crítica e interpretación textual que posibilite, tal vez, esta codificación.

Así se explica que las actividades basadas en codificación TEI requieren una lectura detallada de la obra, una conciencia



Figura 14. Codificación en TEI del primer acto de *La casa de Bernarda Alba*, de García Lorca con Notepad++

metatextual, además de un conocimiento del proceso de marcación en dicho lenguaje. Como esto supone un proceso lento y laborioso, proponemos que los alumnos practiquen con la tipología de drama y anoten una escena de una obra teatral de la Edad de Plata, incorporando las etiquetas específicas. Previamente, se les puede explicar en qué consiste la codificación en TEI y los conceptos básicos que incorpora, como elementos de una etiqueta y la estructura de un documento (encabezado, cuerpo, divisiones). A continuación, se les facilita un esquema y tienen que realizar las pruebas usando primero el editor de texto Notepad++, y después Oxygen. Un siguiente paso es la visualización de la codificación en un interfaz visible con el navegador u otro *software*. Para esta instrucción en el marcado estructural básico se están creando ejercicios y hojas de trabajo con ejemplos y esquemas para guiar a los estudiantes a la hora de realizar sus digitalizaciones. Aunque es una línea apenas comenzada, estamos convencidos de que este tipo de actividades de aprendizaje constituyen una forma eficaz para acercarse a la textualidad digital y comprenderla.

## 4. CONCLUSIONES

Con todo lo expuesto esperamos haber dado a conocer mejor un proyecto de marcación y edición en TEI de textos dramáticos que se ha concebido desde su origen con el objetivo de aportar una infraestructura y herramientas, en abierto, que faciliten y hagan posible la tarea de otros investigadores. La colaboración y la transferencia nos parecen dos aspectos primordiales en el área de las Humanidades Digitales. De este modo, se han descrito los procedimientos de trabajo llevados a cabo por los investigadores, así como el acierto de algunas decisiones tomadas – incluir el corpus en DraCor – y el menor éxito de otras, como el de escoger para una de las fases autores que, por razones de derechos de autor, no pueden ser incluidos en una base de datos en abierto. En todo nuestro recorrido se confirma la necesidad y la conveniencia de un trabajo de carácter colaborativo.

Además, resulta fundamental ir aunando la construcción de infraestructura, en nuestro caso textos en TEI, con una formación en herramientas de las Humanidades Digitales y una labor de difusión e investigación. En este sentido los artículos, sesiones de formación interna y seminarios abiertos a otros investigadores resultan, en nuestra opinión, fundamentales, a la vez que presentar las líneas y adelantos en la investigación en congresos especializados en las Humanidades Digitales, tanto del ámbito hispano (Valladolid 2018; Ciudad de México 2018; Toledo 2019; III Congreso Internacional de Lingüística Computacional y de Corpus: Una mirada desde las tecnologías del lenguaje y las Humanidades Digitales, Bogotá 2020, Santiago de Compostela 2021), como en los promovidos por la Alianza de Organizaciones de Humanidades Digitales (ADHO) en Ottawa 2020, Krasnoyarsk 2021 y, Tokyo 2022, así como los organizados por The Association for Computers and the Humanities (ACH) en 2021 (Texas) y 2022 (DH Unbound).

Finalmente, el punto de encuentro entre las Humanidades Digitales y las prácticas escolares nos parece inspirador y fructífero, por ser un aporte a la formación de nuevas generaciones, y por

promover el uso de las tecnologías digitales con mayor conciencia y capacidad crítica. Aparte de que el enfoque pedagógico permite "aprender haciendo", resulta interesante impulsar la capacidad de razonamiento acerca del procesamiento electrónico de un texto literario, sus metodologías e implicaciones. En definitiva, queremos orientar nuestra labor de digitalización, investigación y docencia a los aspectos humanísticos, no solo tecnológicos, teniendo presente el postulado de Nuria Rodríguez-Ortega: "una agudización de nuestra conciencia crítica sobre el carácter del conocimiento que producimos y sobre las infraestructuras que lo hacen posible, para tomar decisiones responsables y éticas" (2018: 5).

## REFERENCIAS BIBLIOGRÁFICAS


Allés Torrent, Susana y del Rio Riande, Gimena (2018). Enseñar edición digital con TEI en español. *Aprendizaje situado y transculturación. Humanidades Digitales: Construcciones locales en contextos globales. Actas del I Congreso Internacional de la Asociación Argentina de Humanidades Digitales – AAHD*. Buenos Aires.

Allés Torrent, Susanna (2019). Introducción a la Text Encoding Initiative. Definición, aplicaciones prácticas y recursos (v.2). *TTHUB. Text Technologies Hub: Recursos sobre tecnologías del texto y edición digital*. https://tthub.io/aprende/l1-intro-a-tei/

Arana, Rocío; Santa María, Teresa y Simón, María (2015). El teatro de la España contemporánea (1868-1936). *Texto Digital,* 11(2) (julio-diciembre), 111-127 http://dx.doi.org/10.5007/1807-9288.2015v11n2p111

Huerta Calvo, Javier (2003). *Historia del teatro español*. Barcelona: Gredos

Calvo Tello, José; Jiménez, Concepción y Santa María, Teresa (2017). *Biblioteca Electrónica Textual del Teatro en Español (BETTE)*. Universidad Internacional de La Rioja. https://github.com/GHEDI/BETTE

Dabrowska, Monika (2021). Análisis semántico y cuantitativo de "La casa de Bernarda Alba" en el aula con Voyant Tools. Una

propuesta de aplicación didáctica. *Revista Internacional de Pedagogía e Innovación Educativa*, 2(2), 57–74. https://doi.org/10.51660/ripie.v2i2.82

Dabrowska, Monika (2022). Posibilidades de digitalización de "Yerma". Una propuesta de aula. En Juan Manuel Rodríguez Álvarez y Sandra Sánchez García (coord.). *Didáctica de la lengua y la literatura y nuevas tecnologías* (pp. 131-142). Madrid: Dykinson.

Fischer, Frank, et al. (2019a). Programmable Corpora: Introducing DraCor, an Infrastructure for the Research on European Drama. *In Proceedings of DH2019: "Complexities"*, Utrecht University, doi:10.5281/zenodo.4284002.

Fischer, Frank; Börner, Ingo; Göbel, Mathias; Hechtl, Angelika; Kittel, Christopher; Milling, Carsten y Trilcke, Peer (2019b). Programmable Corpora: Introducing DraCor, an Infrastructure for the Research on European Drama. In *Proceedings of DH2019: "Complexities"*, Utrecht University, doi:10.5281/zenodo.4284002.

Fischer, Frank; Santa María Fernández, Teresa; Calvo Tello, José y Milling, Carsten. (2021) Una perspectiva comparativa del teatro español: integración de la "Biblioteca Electrónica Textual del Teatro en Español" en la plataforma "DraCor" en M.ª Ángela Celís Sánchez (coord.), *Las humanidades digitales como expresión y estudio del patrimonio digital* (pp. 273-282). Universidad de Castilla La Mancha: Ediciones de la Universidad de Castilla-La Mancha.

Jiménez-Fernández, Concepción-María y Calvo Tello, José (2020). Grafos de escenas y estudios literarios digitales: una propuesta computacional crítica. *452ºF Revista de Teoría de la Literatura y Literatura Comparada*, 23: 78-101.

Jiménez, Concepción-María; Santa María, Teresa y Calvo Tello, José (2017). *Biblioteca Electrónica Textual del Teatro en Español (BETTE)*, La Rioja: UNIR https://bit.ly/2SGtwhB

Martínez Carro, Elena y Santa María, Teresa (2019). Biblioteca electrónica textual del teatro español (1868-1936) e investigación con grafos. *Revista de Humanidades Digitales* 3, 23-45. http://revistas.uned.es/index.php/RHD/article/view/23144/

*Mnemosine*. http://repositorios.fdi.ucm.es/mnemosine/queesmnemosine.php

Pedraza Jiménez, Felipe Blas y Rodríguez Cáceres, Milagros (1983). *Manual de literatura española. 7: Época del realismo.* Berriozar: Cénlit.

Rodríguez-Ortega, Nuria (2018). Cinco ejes para pensar las humanidades digitales como proyecto
de un nuevo humanismo digital. *Artnodes* 22: 1-6. http://dx.doi.org/10.7238/a.v0i22.3263

Romero López, Dolores (2014). Hacia la *Smartlibrary*: *Mnemosine*, una biblioteca digital de textos literarios raros y olvidados de la Edad de Plata (1868-1936). Fase I. En Sagrario López Poza y Nieves Pena Sueiro (eds.), *Humanidades Digitales: desafíos, logros y perspectivas de futuro* (pp. 411-422). Universidade da Coruña: SIELAE.

Romero López, Dolores (2015). Mujeres traductoras en la Edad de Plata (1868-1939): Identidad moderna y *affidamento. Hermeneus: Revista de la Facultad de Traducción e Interpretación de Soria,* 17, 179-207.

Ruiz Ramón, Francisco (1986). *Historia del teatro español. Siglo XX.* Madrid: Cátedra.

Ruiz Ramón, Francisco (1996). *Historia del teatro español (desde sus orígenes hasta 1900)*. Madrid: Cátedra.

Santa María, María Teresa (2021). Los grafos y la edición digital: Propuesta para enseñar Literatura en Secundaria y Bachillerato. *Revista Internacional de Pedagogía e Innovación Educativa,* 2(1), 139–162. https://doi.org/10.51660/ripie.v2i1.81

Thion Soriano-Mollá, Dolores (2021). Emilia Pardo Bazán, una intelectual moderna, también de la Edad de Plata. *Feminismo/s,* 37, 53-80. https://doi.org/10.14198/fem.2021.37.03

# III

# LAS VÍAS DEL MODELADO, DE LOS DATOS Y DE LA INFORMACIÓN

# IDENTIFICANDO EL SUJETO DEL PAISAJE:
## MODELADO E INTEGRACIÓN DE LA DIMENSIÓN SOCIAL PARA UNA (RE)LECTURA DE LA GEOGRAFÍA ALTOMEDIEVAL

Sonia Medina Gordo, Esther Travé Allepuz

**Resumen**

El estudio del poblamiento y estructuración del espacio durante la alta Edad Media presenta una variedad de dinámicas que condicionan enormemente la forma en que concebimos unos paisajes en pleno proceso de transformación. En este sentido, la valoración de diferentes registros informativos ha supuesto un avance significativo para la superación de viejos modelos interpretativos del periodo, pero también ha puesto de relieve la difícil relación existente entre las disciplinas que recuperan este pasado. El problema no es puramente terminológico, sino que surge de la manera en que comprendemos y modelamos los datos y la información que manejamos. Nuestra labor como investigadores debería poder capacitar la identificación e interrelación de estos conocimientos desde entornos de trabajo propicios a la colaboración, promoviendo una verdadera comunicación entre disciplinas. El propósito de este trabajo es procurar una equiparación epistemológica, particularizando en la dimensión social

Sonia Medina Gordo, Universitat de Barcelona
Grup de Recerca d'Arqueologia Medieval i Postmedieval (GRAMP.UB)
ORCID: https://orcid.org/0000-0002-3921-3025
sonia.medina@ub.edu

Esther Travé Allepuz, Universitat de Barcelona
Grup de Recerca d'Arqueologia Medieval i Postmedieval (GRAMP.UB)
ORCID: https://orcid.org/0000-0002-6769-4487
esther.trave@ub.edu

que participa de este paisaje altomedieval a través de su modelado en diferentes niveles para poder llevar a cabo interpretaciones transversales e interoperables de esta etapa.

**Palabras clave:** Modelado de datos, Actor, rol, semántica, interoperabilidad

**Abstract**
The study of peopling and the organisation of space during Early Middle Ages presents a variety of dynamics that greatly determine the way we conceive landscapes in the midst of a process of transformation. The evaluation of different sources of information has been a significant step forward in overcoming old interpretative models for the period, yet it has also highlighted the difficult relationship between the disciplines that recover this past. The problem is not purely terminological but arises from the way we understand and model data and the information we handle. Our target as researchers should be to enable the identification and interrelation of this knowledge from collaborative perspectives, thus promoting real communication between disciplines. This paper aims at seeking an epistemological equalisation, focusing on the social dimension involved in this early medieval historical landscape through its modelling at different levels in order to carry out transversal and interoperable interpretations of the period.

**Key words:** Data modelling, Actor, Role, Semantics, Interoperability

## 1. INTRODUCCIÓN

El periodo comprendido entre los siglos V-XI se presenta como una etapa de transición larga y compleja, con unos modelos de ocupación y jerarquización del espacio norte peninsular desiguales en función de las diferentes tradiciones romanas, suevas, visigodas, musulmanas, asturleonesas o cispirenaicas. Estas nuevas realidades de poblamiento y organización social del espacio no solamente parecen manifestar procesos de identidad local en la transformación interna de las mismas sociedades que los habitan, sino que se están produciendo además sobre escenarios históricamente marginales, como lo fueron las zonas de montaña o de valle. De esta manera, uno de los principales problemas que ofrece esta geografía altomedieval es una falta de interpretación homogénea, fruto en ocasiones del contexto, las observaciones y los registros disponibles para cada zona. Esto sugiere que todavía necesitamos una sistematización y formalización de los distintos marcos interpretativos que se forman a partir de los diferentes registros explotados.

Dentro de esta variedad, la documentación escrita ha supuesto durante muchos años la base de información privilegiada sobre la que se ha cimentado la narrativa propia de estos siglos. Hablamos de un corpus relativamente poco abundante, pero marcadamente accesible ya que cuenta con una larga tradición de historiadores que han editado y estudiado casi en su totalidad el conjunto de textos del que disponemos. Esta trayectoria es la que, de hecho, se contrapone al registro arqueológico, que es cuantitativamente más abundante, pero de difícil manejo para una buena parte de medievalistas dada su complejidad, escasa publicación y su alta fragmentación territorial. A pesar de esta situación, el uso de las fuentes materiales es lo que no en pocas ocasiones nos ha permitido cuestionar o matizar ciertas lecturas a veces profundamente arraigadas en la tradición académica (Martínez Sopena, 2004) y, al mismo tiempo, en convivencia con los elementos geográficos y topográficos de cada territorio,

nos está ayudando a ampliar y consolidar los conocimientos que incorporamos para este pasado.

Dentro de este panorama, en las últimas décadas se ha puesto de relieve una doble tendencia que, hasta cierto punto, está vertebrando los planteamientos metodológicos y de gestión de la información de muchas investigaciones actuales. Por un lado, la disponibilidad de una amplia multiplicidad de fuentes y registros, junto con los beneficios evidentes que un uso conjunto de los mismos conlleva en la producción de conocimiento de este pasado medieval, ha propiciado que la relación entre las disciplinas implicadas sea cada vez más estrecha, y el uso combinado de los datos que cada una de ellas emplea se haya naturalizado en las tareas de nuestro día a día. No obstante, esta familiarización no está trayendo consigo una dialéctica real y equidistante entre datos, y, en muchas ocasiones, esta comunicación deficiente termina produciendo unas contradicciones y distancias entre discursos difícilmente salvables (Zadora-Rio, 1995; Quirós, 2007).

Por otro lado, y directamente relacionado con los avances digitales y las implicaciones que el *giro computacional* ha tenido dentro de las Humanidades (Berry, 2011), asistimos a un desarrollo prometedor en la investigación altomedieval a través de los medios tecnológicos actuales que, sin embargo, es todavía muy afín a la datificación y cuantificación de las fuentes y su información; unas prácticas que principalmente hacen del contenido de las mismas un conocimiento objetivo sólo en apariencia, a la par que interpretable, analizable, representable y compartible en términos matemáticos (Kitchin, 2014).

A pesar de que ambos procedimientos promueven una colaboración y un mayor acercamiento de nuestras bases de conocimiento a otros investigadores desconocedores quizás de la temática y especialidad que cada uno de nosotros manejamos, no se termina de resolver una reclamada reflexión en torno a la modelación de la información y del conocimiento en estos nuevos contextos

pluridisciplinares. Porque, de hecho, aquí la cuestión no pasa por preguntarnos de qué herramientas digitales nos debemos valer para crear y vincular los datos provenientes de diferentes fuentes, ni tampoco si su cuantificación para crear o ampliar conocimiento sobre el periodo es pertinente y provechoso. Más bien, el desafío reside en *potenciar* nuestra propia capacidad de abstracción y representación de conceptos para hacerlos aprovechables a toda ciencia que estudia el pasado, y realizar esta tarea siendo conscientes de que el dato que trabajamos en Humanidades es también complejo, interpretado, no siempre preciso y, sobre todo, muy dependiente del contexto (Owens, 2011; Schöch, 2013).

Teniendo presente estas cuestiones, creemos que el avance en las investigaciones de esta etapa histórica debería considerar y replantear las antedichas tendencias y reforzar el aprovechamiento de los aspectos cualitativos —y no tanto cuantitativos— de la información que manejamos. Mediante este planteamiento, el contenido de las fuentes debería poder permitirnos un razonamiento de la información en términos no numéricos. Esto significa que los datos dejarían de ser simples unidades limitadas a ser medibles o cuantificables, para pasar a ser conceptos o *entidades* sobre los que poder razonar, realizar interpretaciones y atribuir significados.

Entre los beneficios que ello supone, queremos destacar la posibilidad de proporcionar un mayor grado de consenso y compartición de este conocimiento, en tanto en cuanto una descripción explícita de la información facilita su reutilización e integración dentro de otros modelos y formas de gestión del conocimiento promoviendo así estrategias y pautas de trabajo basadas en el intercambio y la cooperación. Unos criterios para los que, además, todavía es necesario un trabajo de integración de modelos sociales en las interpretaciones históricas, tanto desde la propia identificación del componente humano en los registros (Tuominen, Hyvönen y Leskinen, 2018) como de la definición formal de sus relaciones (Bearman, Moody y Faris, 2002).

En este trabajo queremos plantear este tipo de cuestiones a partir de la experiencia que nos brinda el modelado de datos, y cómo este acercamiento nos está permitiendo ampliar el grado de consenso entre las distintas disciplinas que trabajan el pasado medieval. Para ello, nos centramos en los aspectos sociales que conforman un paisaje, presentando unos resultados preliminares que toman como base un modelo conceptual surgido de la experiencia que nos proporciona la Arqueología en los estudios de paisaje y territorio. Todos los datos y la información generada están siendo gestionados a través de herramientas basadas en tecnologías de información actuales siguiendo una lógica relacional y espacial a fin de integrar y comprender transversalmente las dinámicas y propiedades que caracterizan la geografía de la alta Edad Media.

## 2. OBSERVANDO EL PAISAJE A TRAVÉS DE LA SOCIEDAD ALTOMEDIEVAL

Para entender los cambios producidos en los territorios situados en el tercio norte peninsular, y con ellos, el nuevo modelo de sociedad que caracterizará una buena parte del periodo medieval es necesaria la lectura en paralelo de varios factores generados y desarrollados muy tempranamente, en los albores de esta etapa histórica, y en la que el uso de las fuentes de información disponibles no siempre resulta sencillo. Nuestras interpretaciones deben atender cuestiones como el grado desigual de romanización y cristianización de todo el conjunto peninsular; la relación de las sociedades situadas al este y oeste de un emergente reino Astur, y mediante qué dinámicas se vincularán las mismas a este entramado estatal en formación; y, también, cómo las propias características y transformaciones sociales internas de estos pueblos permiten la superación y desarticulación de unos vínculos de encuadramiento supralocal que terminarán determinándose en la implantación de una nueva estructura económica y social, sustentada sobre

un campesinado articulado en torno a comunidades de aldea (Mínguez, 2004: 171-175; Estepa Díez, 2009).

Precisamente, uno de los problemas más debatidos de estas comunidades aldeanas es la definición de su forma y origen. Para aquellos medievalistas acostumbrados al manejo de la documentación escrita, este tipo de hábitat parece evolucionar en consonancia con formas y elementos propios de un espacio feudal, que limitarían la libertad de los campesinos a los poderes señoriales. La ordenación de la producción y las formas de poblamiento del momento, por tanto, son elementos que responderían a este tipo de modelos surgidos a partir del siglo XI, en el momento de implantación de parroquias, castillos y necrópolis, y la sociedad medieval sería una consecuencia directa de la implantación del feudalismo en estos lugares. Por su parte, aquellos otros medievalistas más familiarizados con la lectura e interpretación de los registros materiales acercan unas lecturas del fenómeno a momentos más tempranos. Para ellos, estos hábitats serían territorialmente más dispersos en origen, pero mantendrían unos esquemas sociales y una mentalidad comunal que les permitirían articular y gestionar perfectamente sus territorios. El proceso descrito por los primeros respondería, desde el punto de vista de los segundos, a unas pautas de reordenamiento y transformación del paisaje, y no su origen (Zadora-Rio, 1995; Quirós, 2007).

Las dinámicas sociales y políticas que van a terminar configurando cada paisaje, por otro lado, no se pueden entender sin la incorporación de factores como los ritmos y elementos de transformación en cada territorio. Por ejemplo, trabajos como los de J. C. Sánchez (2013) nos están permitiendo reseguir mejor la trayectoria histórica en los sectores más occidentales de la Península, cuyos precedentes romanos, suevos y después visigodos jugaron un papel determinante en su incorporación en centros de poder emergentes, como lo es el asturleonés. En cambio, esta incorporación no es tan clara para los territorios orientales, donde el legado romano-visigodo parece que apenas tuvo presencia o

pudo haber quedado prácticamente borrado (Mínguez, 2019: 193-194). Aquí, la integración de estos espacios y las comunidades que los habitan será más tardía, cuando se documente un mayor protagonismo de los reyes leoneses en el reparto de bienes y privilegios en estos espacios. Pero para reseguir estos procesos, como decimos, debemos poder ser capaces de acudir a las fuentes desde perspectivas integradoras y críticas con la información.

## 3. EL USO DE LAS FUENTES: ENTRE SÍNTESIS INTERPRETATIVAS Y COLECCIONES DE DATOS

La Historia y las Ciencias de la Información y de la Documentación identifican y clasifican dos tipos de fuentes: las fuentes primarias y las fuentes secundarias. El uso del término *primarias* como objeto y punto de partida desde donde obtener el conocimiento del pasado refuerza la idea de un tipo de información cuyo conocimiento no ha estado todavía interpretado, ya que su contenido está compuesto de manera casi exclusiva por aquellos elementos que son los que nos han sobrevivido del pasado (Thibodeau, 2019). En cambio, la noción de *secundarias* pretende acentuar el aprovechamiento y reflexión sobre las primeras, y que por su estrecha y directa vinculación nos pueden servir a nosotros como base o refutación para la elaboración de nuestras interpretaciones. Adicionalmente, podríamos considerar un tercer tipo de fuentes, las denominadas fuentes terciarias, que a nivel de obtención y construcción de conocimiento no nos aporta ninguna idea adicional o de refuerzo, pero su empleo en nuestras investigaciones nos ayuda a ubicar y reconocer las dos primeras en tanto que nos las recompila en un único espacio (Fjordback Søndergaard, Andresen y Hjørland, 2003).

Para el periodo que nos concierne, contamos con dos registros sin los que el conocimiento de esta etapa habría resultado una tarea del todo impracticable: el registro escrito y el registro material. En relación con el primero, disponemos de un fondo documental que nos está testimoniando muchas de las actividades

llevadas a cabo por los centros monásticos y —aunque en menor medida— por los sectores laicos de la sociedad. Hablamos de documentos que dan prueba de eventos tan habituales del momento como lo eran las donaciones, las compraventas, la fundación de nuevos lugares o la concesión de exenciones u otros privilegios dirigidos a quienes protagonizaban y se beneficiaban de todas estas mercedes (Escalona, 2017).

A dichas colecciones escritas debemos añadir otro tipo de textos, con un carácter narrativo, que nos permiten ilustrar en paralelo los diferentes ritmos expansivos y de integración territorial de las sociedades del tercio norte peninsular. No obstante, para este caso las crónicas castellanas solamente nos ofrecen el panorama de actuación de los monarcas astur-leoneses, ajustándose a un discurso de legitimación de la figura regia como el personaje principal y heredero de la tradición goda, que es lo que le promueve a avanzar y reincorporar por derecho los espacios ocupados por los musulmanes, y además hacerlo apoyándose en otros individuos distinguidos en los relatos por tener un cierto poder de alcance local (Isla, 2017).

En lo tocante a los datos materiales, su valoración como fuente de información para la caracterización de algunos aspectos relativos a las dinámicas de población o la caracterización de los espacios de hábitat es bastante reciente. En esta tarea ha sido necesaria, por una parte, la consolidación de la Arqueología Medieval como una disciplina capaz de producir un discurso propio y sólido, y, por otra, la incorporación progresiva de datos nuevos, alejados de los tradicionales campos de investigación que durante mucho tiempo se habían centrado en informar sobre las necrópolis, las iglesias, los eremitorios y las cerámicas del periodo (Gutiérrez, 2006). El reto que presenta esta materialidad es que queda plasmada sobre una variedad de formatos escasamente normalizados, y que además abarcan desde la confección de cartas arqueológicas para la localización y conocimiento de los yacimientos de un determinado territorio, hasta la redacción de

una memoria técnica en la que idealmente se recoge por escrito y de manera gráfica (con planos, fotografías o dibujos) toda la información emanada durante las labores campo.

En este panorama, la incorporación de medios tecnológicos para el procesamiento de la información parece estar brindando a los medievalistas la posibilidad de un trabajo con las fuentes diferente y actualizado. A tal respecto, uno de los elementos cada vez más demandados en los proyectos de investigación actuales es el diseño de un plan de gestión de los datos en el que se especifique el ciclo de vida completo de los mismos en relación con el estudio. Generalmente, la documentación y gestión de la información se realiza sobre los vestigios primarios, elaborando una base de datos diseñada para tal fin. Un ejemplo genuinamente ilustrativo de este tipo de actuaciones sobre el legado sociocultural de las sociedades del pasado puede ser el proyecto *Procesos Judiciales en las sociedades medievales del norte peninsular* (PRJ)[1], cuya base de datos tuvo por objetivo proporcionar de manera pública un corpus documental con información (cuantificada) acerca de disputas o procesos judiciales de las sociedades septentrionales de la Península Ibérica durante la Edad Media.

## 4. LA EXPERIENCIA DEL MODELADO, LA GESTIÓN DIGITAL Y LA INTEGRACIÓN DE LOS DATOS

En base al contexto previamente expuesto, nuestro estudio se está vertebrando sobre un plan de gestión que ha priorizado, no tanto la creación de nuevos datos, como el aprovechamiento de los que ya disponemos para (re)interpretar los paisajes altomedievales y las sociedades que los significan. Las fuentes manejadas para tal fin son de naturaleza escrita, narrativa, arqueológica y gráfica, con independencia de su tipología primaria o secundaria. En este sentido, el objetivo que perseguimos se ha centrado en la definición

1. http://prj.csic.es [consulta: 9 de marzo de 2022].

semántica de los datos para la integración de la información emanada de cada una de ellas, procurando acercar un conocimiento holístico del periodo.

### 4. 1. LA SEMÁNTICA DE LO SOCIAL: DEFINICIÓN DEL MODELO CONCEPTUAL

La definición de ontologías o el diseño de modelos conceptuales para modelar el conocimiento no son campos desconocidos en Humanidades. Las investigaciones recientes cuentan con una dilatada experiencia de proyectos y estudios que han incorporado en sus planes de investigación la elaboración de este tipo de representaciones del conocimiento. Uno de los modelos de referencia conceptual mejor conocido en el ámbito del patrimonio cultural y la documentación de museos es el modelo CIDOC, que ha llegado a consolidarse como todo un estándar para la integración y el intercambio de información patrimonial. Esto es debido en parte por la semántica que conforma el modelo, que permite desde un alto nivel de abstracción la posibilidad de extraer información de diferentes fuentes o recursos con independencia del soporte o ciencia que las trabaje. Sin embargo, precisamente porque modela la información desde niveles de abstracción altos han surgido otros modelos de carácter más específico que han hecho extensible su aplicación sobre estudios más delimitados.

En el caso de los modelos que trabajan la dimensión social, éstos se han orientado principalmente a la extracción de información de sujetos o colectivos; por tanto, han buscado conceptualizar a las personas. Un buen ejemplo de ello es el modelo Bio CRM, un primer acercamiento en la conceptualización de personajes para obras de carácter biográfico o prosopográfico (Tuominen Hyvönen y Leskinen. 2018). Lo interesante de este modelo es que descompone al individuo, detallando aspectos como sus atributos, sus relaciones con otras personas, y aquellos eventos en los que su implicación define el rol que mantienen. Algo parecido podemos encontrar

en el modelo propuesto por K. Thibodeau, quien ha diseñado una manera más concreta de identificar el grado de intervención de la agencia humana en los eventos, especialmente cuando ésta es de carácter activo (Thibodeau, 2019: 8-11). Paralelo a ambos, parece estar desarrollándose también otra extensión de CIDOC, en este caso CRMsoc, que pretende documentar el fenómeno social en un sentido mucho más amplio (Alamercery *et al.*, 2019).

Sin embargo, a pesar de todos estos esfuerzos, la experiencia por modelar *lo social* desde los estudios de corte histórico parece estar menos perfeccionada que otras parcelas temáticas, como ahora la espacialidad (Hiebel *et al.*, 2014) o la temporalidad (May, 2020). Además, todavía falta extender y ampliar los conocimientos de campos como el de las redes sociales sobre fenómenos no necesariamente sociales, así como en otras disciplinas (como ya se está haciendo en Arqueología) (vid. Donnellan, 2020; Burghmans y Peeples, 2017). Parte de estas limitaciones son las que hemos procurado matizar a través de la conceptualización de la dimensión social entendida como un ámbito que participa del paisaje.

Hemos optado por trabajar en base a un modelo conceptual cuya aplicación práctica cuenta ya con un cierto recorrido (Mauri, 2006; Travé, del Fresno y Mauri, 2020) y está presente en herramientas de gestión integral del patrimonio o proyectos de investigación como lo son HORAI[2] o GREYWARE (Medina, del Fresno y Travé, 2022; Travé *et al.*, 2021b). Este modelo al que hacemos referencia se vertebra sobre los conceptos de Unidad Topográfica (UT), como la evidencia de una acción o una situación localizable en tiempo y espacio; Unidad Estratigráfica (UE), una evidencia similar para la que es necesario la documentación de su materialidad; y Actor (Ac), el participante individual o corporativo, pasivo o activo, de este tipo de acciones. Un aspecto que debemos destacar de este modelo es la propia capacidad de ampliar sobre

2. https://horai.es/ [Consulta: 4 de marzo de 2022]

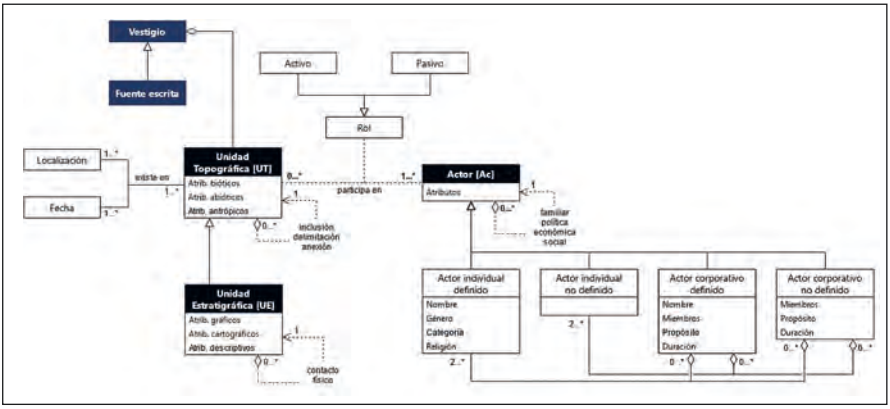



Figura 1. Diagrama UML del modelo conceptual UT/Ac. Se ha refinado el concepto de *Actor* y el de *vestigio* ajustándonos a las necesidades del estudio.

otras fuentes y disciplinas conceptos notablemente asumidos y usados en arqueología por medio del concepto de Unidad Topográfica. Pero también por su manera de incorporar nociones particularmente difíciles en muchos modelos abstractos, como las entidades o los acontecimientos, lo que nos permite refinar los conceptos en función de los objetivos que perseguimos para cada estudio (González-Pérez y Martín-Rodilla, 2015).

### 4. 1. 1. Detallando al individuo y al colectivo

La identificación de un Actor como individual o corporativo depende del número de participantes *reconocidos* en las fuentes. Conceptualizar de esta manera la información nos obliga a tener que detallar su definición ya que, si bien es cierto que esta distinción parece clara al principio, podemos encontrarnos con situaciones en las que esto no parece tan evidente. Para ilustrar estas apreciaciones, facilitamos a continuación un extracto de la Crónica de Alfonso III, en su versión rotense. En este texto se nos narra la victoria de Alfonso II sobre Mugait y su ejército, así como el traslado de la capitalidad del reino a Oviedo en paralelo a la construcción y dotación de numerosas basílicas e iglesias:

> Hunctus est in regno predictus rex magnus Adefonsus XVIII Kalendas Octobris era quo supra. Anno regni eius tertio Arabum exercitus ingressus est Asturias cum duce quendam nomine Mugait. Qui in loco Lutos ab Astores prebenti cum idem duce suo LXX milia sunt interfecti. Iste solium suum Oueto firmauit. Baselicam quoque in honore Domini et Saluatoris nostri Ihesu Xpi cum bis seno numero apostolorum altaris adiungens, sibe eclesiam hob honorem sancte Marie semper uirginis cum singulis hinc atque inde titulis miro opere atque forti instructione fabricauit: etiam aliam ecclesiam beatissimi Tirsi martiris prope domum sancti Saluatoris fundauit, necnon satis procul a palatium edificauit eclesiam in honorem sancti Iuliani et Baselisse cum uinis altaribus magno opere et mirauili conpositione locauit; nam et regia palatia, balnea, promtuaria atque uniuersa stipendia formauit et instruere precepit." (Gil, Moralejo, y Ruiz de la Peña, 1985)

En negrita hemos resaltado algunos de los actores que nos interesan para entender las diferencias entre unidades. Por ejemplo, no nos resulta difícil identificar rápidamente que tanto el rey Alfonso como Mugait son ambos actores individuales dentro de la narración. De hecho, tanto uno como el otro parecen estar actuando a título personal, lo que nos ha llevado a reconocerlos conceptualmente como un *Actor individual definido*. Ahora bien, términos como "ejército de árabes" o "astures", en un principio, dado el número de individuos que se necesitan para la composición de estos grupos, intuitivamente podrían ser entendidos como actores corporativos. Sin embargo, este tipo de actores están actuando como una única entidad; es decir, en ningún momento de la narración podemos saber con exactitud el número concreto de hombres que conformaron el ejército de los árabes que entraron en Asturias con Mugait, ni menos conocer a cada uno de los astures. Dada esta falta de determinación de sus integrantes, esta unidad queda por tanto esquematizada en el modelo como un *Actor individual no definido*. Individual, precisamente, porque a pesar de

que claramente son un colectivo formado por varios individuos, actúan como un ente unitario.

En lo relativo a los actores corporativos, su definición detallada está sujeta a una identificación explícita o implícita en los documentos. Por ejemplo, en el extracto anterior, "los santos Julián y Basilia" son una muestra clara de *Actor corporativo definido*. En este caso, la unidad queda descrita por medio del sustantivo "santos", conociendo además quienes integran de manera individual estos santos; Julián y Basilia. Es por ello por lo que podemos definirlos como tal. En cambio, en el caso del *Actor corporativo no definido* su identificación aquí es algo menos intuitiva. En el texto solamente podemos identificar una unidad de este tipo; los invasores de Asturias, que viene a ser el ejército de los árabes junto con Mugait. Aquí no podemos decir que se trata de un Actor corporativo definido ya que su denominación como "invasores de Asturias" no se recoge en el documento de manera explícita, sino que es una deducción a la que hemos llegado nosotros por el propio contexto, que es el que en última instancia nos permite saber que ambos actores están llevando a cabo un mismo propósito, en un mismo momento, pero sin que ello sea nominalmente identificable. Esto no pasa con los santos, quienes además de ser denominados como tal, comparten una misma acción común —su advocación— en un mismo momento.

**4. 1. 2. La importancia del contexto, las relaciones y las interacciones**

Otro aspecto que destacamos del modelo son las propias relaciones entre las unidades de información, para las que, además, podemos derivar el rol de los Actores. En nuestro modelo, las únicas relaciones que pueden mantener los Actores son, o entre ellos mismos, o bien con las Unidades Topográficas. En cambio, su vinculación con las Unidades Estratigráficas no es posible puesto que este tipo de dato no nos puede informar acerca de los Actores, como ya se ha señalado en otros trabajos (Travé, del Fresno y Mauri, 2020). En lo referente a las interacciones con las

Unidades Topográficas, conviene reflexionar brevemente acerca del concepto de rol. Toda modelización considera por separado la caracterización y uso de dicho elemento. Por un lado, el rol puede entenderse como un atributo producto de una relación binaria entre dos objetos (A, B), y, por tanto, ser una construcción puramente ontológica. Por el contrario, esta noción podría estar incorporando una característica que le es inherente a cada objeto, con independencia de toda relación existente dentro del sistema (Guarino, 1992; Steinmann, 2000).

La noción que nosotros ofrecemos, y que nos servirá de base para la declaración de las asociaciones posibles entre cada uno de los elementos, asume explícita y separadamente el rol como un atributo derivado de una relación. Se diferencia, por tanto, de la forma de representación propia de la unidad. En nuestro modelo, este rasgo queda perfectamente recogido como un atributo más que define al Actor. Partiendo de estos preceptos, en su vinculación con toda Unidad Topográfica, a cada Actor se le puede establecer un papel activo o pasivo sobre la primera. De esta manera, en contextos como la donación de bienes o la atribución de derechos, no solamente podremos documentar y ubicar de una manera más o menos exacta el conjunto de elementos que se enajenan, sino que al mismo tiempo podremos establecer y vincular al (Ac) propietario, al (Ac) donatario y al (Ac) beneficiario sobre una zona determinada. En lo relativo al paisaje y su articulación, trabajar estos aspectos nos ayuda a relacionar quiénes son los individuos que participan en este tipo de dinámicas, resiguiendo en paralelo ciertos aspectos como las relaciones de poder o el aprovechamiento de espacios de uso colectivo.

Por su parte, las relaciones entre Actores también nos permiten establecer el propio rol de la unidad. Por ejemplo, los vínculos familiares pueden determinar la paternidad o maternidad de un Actor con respecto a otro, al mismo tiempo que determina la filiación del segundo con respecto al primero. En este sentido, como podemos ver, no solamente estamos determinando el propio

rol del individuo, sino que además establecemos implícitamente la propia direccionalidad de la relación. De igual manera, en aquellas otras asociaciones de carácter político o social, nos encontramos con otro amplio abanico de posibilidades sistematizadas según el contexto del que son producidas. Así, es más sencillo de registrar el papel de ciertos personajes dentro del entramado que conformará el paisaje altomedieval, como pudiera ser el caso de los *boni homine* con respecto a las comunidades campesinas (Luis Corral y Pérez, 2017), o del monarca en relación con otros personajes notables del periodo, ya sean laicos como eclesiásticos (Carvajal, 2017; Carvajal y Narbarte, 2019; Escalona, 2016).

## 4. 2. ORGANIZANDO LA INFORMACIÓN: EL MODELO LÓGICO RELACIONAL

El esquema organizativo que hemos ajustado para el manejo de la información es el modelo relacional definido por E. F. Codd (1970). En este tipo de diseños, los datos quedan ordenados y se gestionan a través de la definición de tablas o *relaciones*. El precursor de este modelo lógico definía formalmente estas relaciones como un subconjunto del producto cartesiano de *n* conjuntos (Codd, 1970: 379), por lo que las tablas son, en esencia, un conjunto de datos ordenado en columnas y filas (Wickham, 2014). La vinculación entre cada tabla se determina a través de sus columnas o *atributos*; esto es, los atributos que identifican unívocamente a cada uno de los registros de una tabla (clave principal) son usados en otra para determinar este nexo (clave foránea). Por tanto, es necesario asegurarse de que cada una de nuestras tablas vaya a tener un identificador de valor único que defina de manera clara e inequívoca cada fila para poder hacer efectiva su relación en otra tabla, en la que el valor relacionado seguramente se corresponderá con varias filas o *registros*.

Partiendo de las clases que hemos definido en la Figura 1, y centrándonos en el caso de las crónicas, nuestro modelo organizativo de la información valora un total de 10 tablas, y que

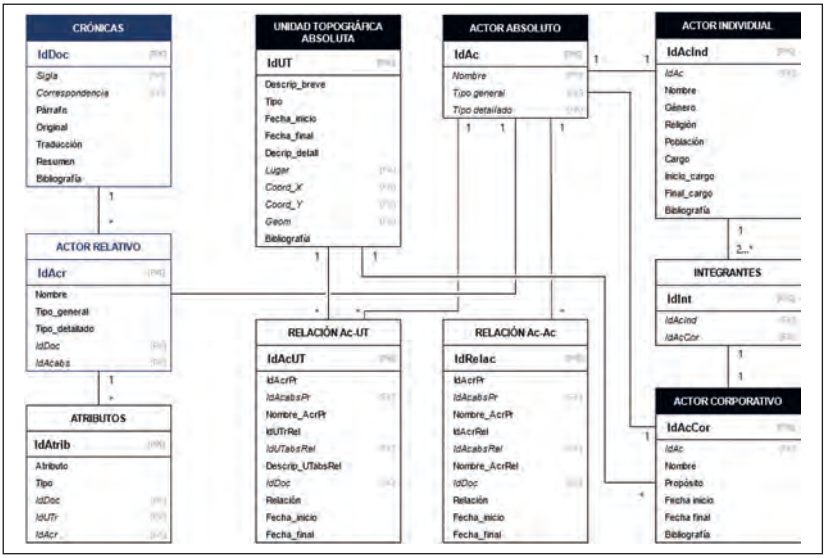



Figura 2. Diagrama UML del modelo lógico según Codd (1970) ajustado a nuestro sistema de gestión de la información.

son las que se recogen en la Figura 2. Estas tablas definen, por un lado, un tipo concreto de fuente escrita (Tabla *Crónica*), y por el otro las propias unidades mínimas de información de nuestro sistema. En este paquete de unidades, conviene diferenciar entre la información específica de sus relaciones (Tablas *Relación Ac-UT* y *Relación Ac-Ac*), las unidades (Tablas *Unidad Topográfica Absoluta*, *Actor Absoluto*, *Actor Individual*, *Actor Corporativo*), la propia etiquetación de la información sobre la fuente (Tabla *Actor Relativo*) y los atributos de los actores identificados en las fuentes (Tablas *Atributos* e *Integrantes*). En conjunto, definimos un grupo de tablas que fundamentalmente nos estructuran e interrelacionan la información correspondiente con datos sobre los Actores. Conviene señalar, además, que esta organización es independiente de cualquier herramienta o medio digital concreto, ya que simplemente estamos ajustando los conceptos del modelo anterior a una lógica relacional tal y como la define Codd.

## 4. 3. DEPENDENCIAS INFORMÁTICAS A TRAVÉS DEL MODELO FÍSICO

En el marco específico de este trabajo, estamos haciendo uso de dos Sistemas Gestores de Bases de Datos (SGBD): Filemaker Pro 17 y PostgreSQL 14. En el primer caso, esta herramienta no solamente nos permite gestionar la información según el modelo relacional de Codd, sino que además integra un Sistema Gestor de Contenido (SGC) para la elaboración de interfaces de usuario. Esta prestación nos permite mayor usabilidad en el momento de manipular la base de datos, sin necesidad de realizar cambios de actualización o consulta de la información mediante sentencias, ni tampoco su visualización en formato tabular. En el caso de PostgreSQL, su integración en el propio sistema nos permite suplir una carencia clara de FileMaker, y es la posibilidad de manipular datos espaciales; particularmente los valores geométricos. Si bien la mayoría de las veces estamos tratando con unidades que pueden ser identificadas como un punto o que carecen de coordenadas exactas, también estamos lidiando con fuentes en las que la geometría y ubicación son determinantes a la hora de extraer información y relacionarla con otros datos.

Respecto a la usabilidad a la que hemos hecho referencia, no solamente responde a la comodidad de consulta de los datos, sino que también proyectamos el repositorio como un entorno amigable, intuitivo y adaptado a los privilegios y roles de los diferentes tipos de usuarios que pudieran manipular la información. Este diseño se sustenta sobre los cuatro elementos básicos que conforman FileMaker, que son el archivo, la tabla, la presentación y el objeto. En la Figura 3a podemos apreciar que la tabla *Crónicas* cuenta con dos presentaciones o *interfaces*: una para la caracterización de las fuentes, y otra para el etiquetado de unidades. En esta segunda interfaz convergen en un mismo espacio diferentes objetos para la introducción o visualización de la información de otras tablas. Desde aquí, solamente podemos movernos entre los registros de la tabla, pero con la ayuda de ciertos objetos (los portales y los campos

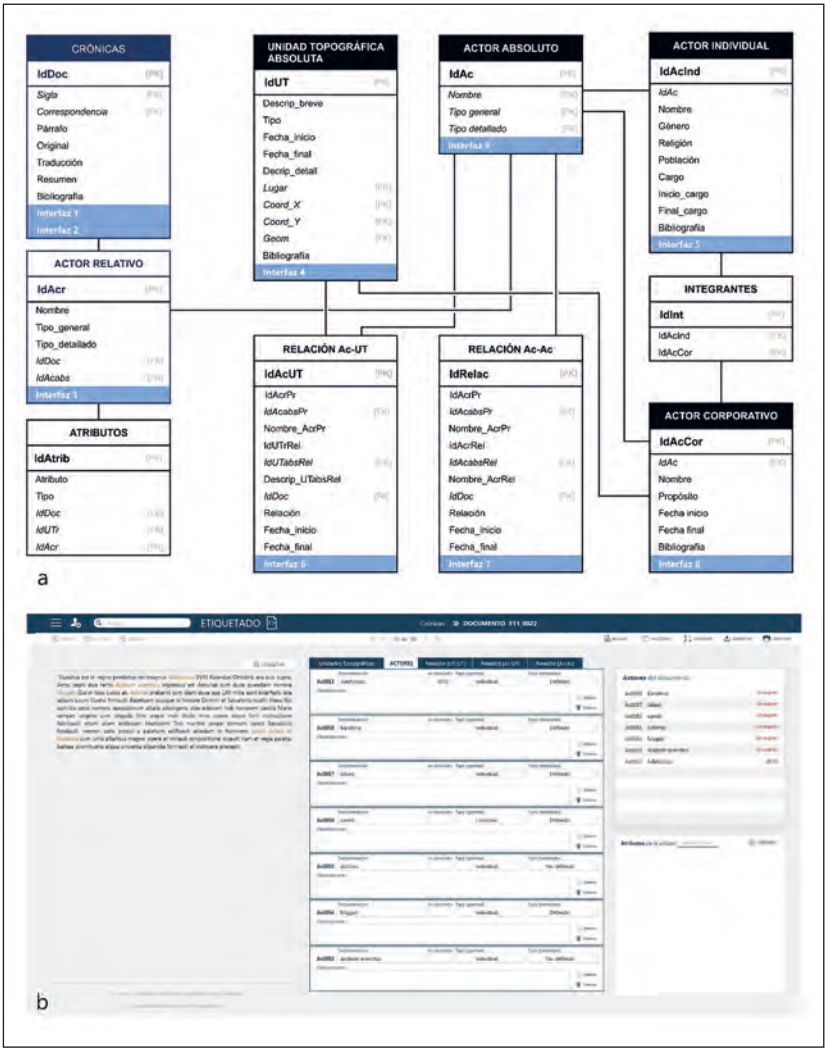



Figura 3. Diagrama de la estructura física de la base de datos (a) e *Interfaz* 2 (b) de la tabla *Crónicas.*

relacionados) es posible administrar la información de otras tablas en una misma vista y trabajar más rápidamente con los datos.

Por su parte, el funcionamiento de PostgreSQL toma como base dos ideas básicas. Por una parte, que el carácter holístico de un paisaje comprende diferentes atributos que, interrelacionados, nos permiten su caracterización y la recuperación de los diferentes conocimientos que se tienen sobre el mismo (Escalona, 2019). Por otra parte, que la forma en que percibimos y procesamos la información de un espacio determina genuinamente la forma en que razonamos sobre el mismo (Egenhofer y Mark, 1995), siendo así indispensables aspectos como la escala de análisis, el nivel de detalle y precisión de las entidades definidas, así como el tipo de información que queremos almacenar. En la actualidad contamos con propuestas elaboradas sobre una teoría consolidada (Turner, 2006; Dabaut y Carrer, 2020) enfocadas precisamente a percibir y analizar todos los elementos del paisaje en su conjunto. Mediante este SGBD, estamos integrando el potencial que ofrece el uso de los valores espaciales (como los que consignamos a las Unidades Topográficas) y los valores no espaciales (como los que pueden definirse en los Actores) para dar forma a los territorios a través de sus huellas, sean o no tangibles sobre el terreno.

Por último, para que ambos sistemas puedan relacionarse entre sí de manera directa hemos hecho uso de controladores ODBC. Este canal de conexión, para el caso de las bases de datos que usan el mismo sistema gestor, no es necesario ya que el programa no necesita adaptar ningún puente entre repositorios. En cambio, este requisito sí ha sido de obligado cumplimiento en nuestro estudio porque trabajamos con programas diferentes y queríamos manejar los datos sin necesidad de migrarlos a través de archivos tipo CSV. Esta decisión se ha tomado en base a la preservación de la calidad de los datos, especialmente en términos de consistencia. De hecho, la habilitación de estos conectores también nos está sirviendo para conectar ambos SGBD a otros programas orientados a otros fines, como los Sistemas de Información Geográfica u otros entornos dedicados a la estadística, como es el caso de RStudio.

## 5. MUCHO MÁS QUE LISTAS, NÚMEROS Y ESTADÍSTICAS: INVESTIGACIÓN CUALITATIVA DE LOS DATOS

Nuestro estudio busca un conocimiento sobre el pasado medieval compartido e integrado, y en particular, de la sociedad que habita y caracteriza estos espacios inmersos en diferentes dinámicas. Para ello, hemos tratado de reformular la significación de los resultados surgidos de la acumulación y el análisis de la información, desarrollando este tipo de inferencias en los propios datos y reflexionando desde su propia naturaleza. Específicamente, esta forma de reconocer, interpretar y recuperar la información de las fuentes ha procurado una integración de las mismas desde diferentes niveles, manteniendo en todo momento la posibilidad de interoperar con terceras personas en cada estadio y modelo; el conceptual, el lógico y el físico.

La integración de los datos y de la información resulta un impedimento puramente técnico, mientras que la integración de conocimientos requiere de unos procedimientos y metodologías que se fundamenten en la propia semántica. En la actualidad, existen propuestas prometedoras desde las Humanidades Digitales de este tipo de puentes de interconexión entre fuentes de conocimiento mediante las prácticas de refinado gradual de modelos (González-Pérez y Martín-Rodilla, 2015). Ello nos está permitiendo reconocer y desarrollar mejores estrategias de diálogo entre los datos, superando de manera efectiva cualquier tipo de silo informativo como de estancamiento o caducidad de la información una vez finalizado el estudio o el proyecto del que haya surgido nuestro modelo. El modelo que presentamos, que toma como base los conceptos de Unidad Topográfica, Unidad Estratigráfica y Actor, propicia precisamente este tipo de trabajo con la información porque posibilita un acercamiento multinivel, ampliando, detallando o proponiendo equivalencias en los registros que construyen el pasado.

Otra vía genuinamente novedosa es la que aprovecha estas definiciones de corte semántico para el desarrollo de propuestas

cercanas a campos como los de la Inteligencia Artificial o la lingüística computacional. En este trabajo hemos ofrecido una aproximación y uso de nuestro modelo de datos a partir de las fuentes escritas, más proclives a informarnos de este tipo de actividades sociales. No obstante, los Actores podemos documentarlos también en otro tipo de registros informativos, como puedan ser las fotografías o los vídeos. Además, el volumen de información generada en un proyecto puede limitar nuestra manera de situarnos con respecto a la misma; desde lecturas cercanas a aproximaciones más panorámicas. En este sentido, la incorporación de ciertas técnicas basadas en sistemas de procesamiento natural del lenguaje, como el *geoparsing* (Gregory *et al.*, 2015), a los estudios de corte histórico y territorial, o de detección de objetos (Lang y Ommer, 2021), particularmente potenciados desde disciplinas como la Historia del Arte, están ofreciendo resultados prometedores, especialmente en materia de adquisición de conocimiento cuantitativo.

Todas estas propuestas nos sugieren una última reflexión en torno a los datos y su manejo a través de los medios digitales actuales. La generación de conocimiento presupone la extracción de información nueva, ya sea porque no se conocía con anterioridad (conocimiento cuantitativo) o porque ha habido una mejor comprensión de esta misma información (conocimiento cualitativo) (Verschoof-van der Vaart y Lambers, 2021: 23-26). Esta creación de conocimiento a partir del desconocimiento (Huggett, 2020), en el primero de los casos, se centra en los datos incompletos (factores epistemológicos), mientras que el segundo profundiza en aspectos como la vaguedad o complejidad de los mismos (limitaciones ontológicas) (Wylie, 2008: 184-185). Tanto uno como el otro acercamiento parecen estar reformulando la manera en que diseñamos los marcos de nuestras investigaciones, acotando la percepción del dato como aquello del que derivamos información en función del valor otorgado a la cantidad (Small/Big Data) o al tiempo que dedicamos en registrarlo y darle sentido (Slow/Fast Data) (Huggett, 2022). Este estudio contribuye también en el reconocimiento de estos acercamientos a las fuentes, y

ha querido resaltar las ventajas que ofrece una mayor introspección sobre las mismas.

## 6. CONCLUSIONES Y VÍAS FUTURAS

El estudio de la formación de nuevas realidades poblacionales y de estructuración social del espacio durante la alta Edad Media plantea todavía hoy una serie de problemas interpretativos de calado. Por una parte, estos acercamientos se realizan a partir de un uso combinado de diferentes fuentes de información disponibles para analizar el fenómeno, lo que está poniendo sobre la mesa la difícil relación que existe entre las diferentes disciplinas que permiten la recuperación de estos vestigios. Por otra parte, la introducción de las nuevas tecnologías en el avance de conocimiento de esta etapa medieval se ha centrado particularmente en la adquisición de la información en forma de datos cuantificables, sin ofrecer una reflexión acerca de los mismos. Todo ello nos está ofreciendo un panorama en el que hay una falta de equiparación de conocimientos evidente, y en el que las investigaciones parecen estar reformulando sin un consenso claro la forma en la que nos relacionamos con la información.

La dimensión social en los paisajes altomedievales es un marco temático en el que este tipo de problemas y situaciones no solamente se están dando, sino en el que además falta por realizar un trabajo considerable a nivel teórico, interpretativo y conceptual. Las investigaciones actuales han priorizado el sujeto como figura única de análisis, pero dejan incógnitas con respecto a las relaciones de éste con el entorno. En este sentido, creemos que una vía imprescindible en la que se debería profundizar es la definición formal y sistemática de las relaciones sociales, partiendo tanto de modelos abstractos como específicos. Mediante esta valoración, no solamente se ampliaría el conocimiento formal de este ámbito, sino que conseguiríamos dilucidar con mayor claridad la implicación de este tipo de informaciones con nociones algo más complejas, como lo son las entidades o los acontecimientos.

**Agradecimientos**
La investigación desarrollada en este trabajo es posible gracias a la financiación recibida de la Generalitat de Catalunya a través de su Secretaria d'Universitats i Recerca, del Fondo Social Europeo, mediante una beca predoctoral (2022 FI_B1 00021). Nuestro trabajo se incluye en la línea de investigación de Arqueología del Paisaje del Grupo de Investigación en Arqueología Medieval y Postmedieval (2021-SGR-00236). El desarrollo metodológico se ha beneficiado también de los trabajos desarrollados en el proyecto GREYWARE (PID2019-103896RJ-I00) financiado por el Ministerio de Ciencia e Innovación y la Agencia estatal de investigación (MCIN/AEI/10.13039/501100011033).

## REFERENCIAS BIBLIOGRÁFICAS

Alamercery, Vincent; Beretta, Francesco; Bruseker, George; Doerr, Martin; Sanderson, Robert; Velios, Athanasios *et al.* (2019). *Definition of the CRMsoc. An Extension of CIDOC CRM to support social documentation.* [Borrador]. Disponible en: https://cidoc-crm.org/crmsoc/fm_releases.

Bearman, Peter; Moody, James y Faris, Robert (2002). Networks and History. *Complexity*, 8(1): 61-71.

Berry, David M. (2011). The *Computational Turn*: Thinking about the Digital Humanities. *Machine Culture*, 12. Disponible en: https://culturemachine.net/the-digital-humanities-beyond-computing/.

Brughmans, Tom y Peeples, Matt (2017). Trends in Archaeological Network Research: A Bibliometric Analysis. *Journal of Historical Network Research*, 1(1), 1-24. Doi: 10.25517/jhnr.v1i1.10.

Carvajal, Álvaro (2017). *Bajo la máscara del "regnum". La monarquía asturleonesa en León (854-1037)*. Madrid: Consejo Superior de Investigaciones Científicas.

Carvajal, Álvaro y Narbarte, Josu (2019). Royal power and proprietary churches in the eleventh-century Kingdom of Pamplona. *Journal of Medieval Iberian Studies*, 11(2): 115-134. Doi: 10.1080/17546559.2019.1566760.

Codd, Edgar F. (1970). A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, 13(6): 377-387. Doi: 10.1145/362384.362685.

Dabaut, Niels y Carrer, Francesco (2020). Historic Landscape Characterisation: Technical Approaches Beyond Theory. *Landscapes*, 21(2): 152-167. Doi: 10.1080/14662035.2020.1993562.

Donnellan, Lieve (2020). *Archaeological Networks and Social Interaction*. Routledge.

Egenhofer, Max J.; Mark, David M. (1995). Naive Geography. En: A. U. Frank y W. Kuhn (eds). *Spatial Information Theory: A Theoretical Basis for GIS.* Springer, 1-15. Doi:10.1007/3-540-60392-1_1.

Escalona, Julio (2016). In the name of a distant king: representing royal authority in the county of Castile, c. 900–1038. *Early Medieval Europe*, 24(1): 74-102.

Escalona, Julio (2017). Cartularios, memoria y discurso en la Castilla medieval. En: E. López Ojeda (coord.). *La memoria del poder, el poder de la memoria. XXVII Semana de Estudios Medievales. Nájera 25 a 29 de julio de 2016*. Logroño: Instituto de Estudios Riojanos, 163-203.

Escalona, Julio (2019). Dense Local Knowledge: Grounding Local to Supralocal Relationships in Tenth Century Castile. J. Escalona, O. Vésteinsson y S. Brookes (eds.). *Polity and Neighbourhood in Early Medieval Europe*. Turnhout: Brepols, 351-379.

Estepa Díez, Carlos (2009). La Castilla primitiva (750-931): Condes, territorios y villas. *Territorio, Sociedad y Poder*, 2: 261-278.

Fjordback Søndergaard, Trine; Andersen, Jack y Hjørland, Birger (2003). Documents and the Communication of Scientific and Scholarly Information. Revising and Updating the UNISIST Model. *Journal of Documentation*, 59(3): 278-320.

Gil, Juan; Moralejo, José L. y Ruiz de la Peña, Juan I. (1985). *Crónicas asturianas.* Gijón: Servicio de Publicaciones de la Universidad de Oviedo.

González-Pérez, César y Martín-Rodilla, Patricia (2015). Integration of Archaeological Datasets through the Gradual Refinement of Models. En: F. Giligny, F. Djindjian, L. Costa, P. Moscati y S. Robert (eds.). *21st Century Archaeology: Concepts, Methods and Tools - Proceedings of the 42nd Annual Conference on Computer Applications and Quantitative Methods in Archaeology*. Archaeopress, 193-204.

Gregory, Ian; Donaldson, Christofer; Murrieta-Flores, Patricia y Rayson, Paul (2015). Geoparsing, GIS, and textual analysis: current developments in spatial humanities research. *International Journal of humanities and Arts Computing*, 9(1): 1-14. Doi: 10.3366/ijhac.2015.0135.

Guarino, Nicola (1992). Concepts, Attributes, and Arbitrary Relations Some Linguistic and Ontological Criteria for Structuring Knowledge Bases. *Data & Knowledge Engineering*, 8: 249-261.

Gutiérrez, José Avelino (2006). Sobre la transición del sistema antiguo al feudal: una revisión arqueológica del altomedievo hispano. *Territorio, sociedad y poder*, 1: 53-78.

Hiebel, Gerald; Doerr, Martin; Hanke, Klaus y Masur, Anja (2014). How to put archaeological geometric data into context? Representing mining history research with CIDOC CRM and extensions. *International Journal of Heritage in the Digital Era*, 3(3): 557-577.

Huggett, Jeremey (2020). Capturing the Silences in Digital Archaeological Knowledge. *Information*, 11(278). Doi: 10.3390/info11050278.

Huggett, Jeremy (2022). Is less more? Slow Data and Datafication in Archaeology. En: K. Garstki (ed.). *Critical Archaeology in the Digital Age*. Cotsen Institute of Archaeology Press, 97-110.

Isla, Amancio (2017). Poder regio y memoria escrita: las crónicas regias altomedievales. En: E. López Ojeda (coord.). *La memoria del poder, el poder de la memoria. XXVII Semana de Estudios Medievales. Nájera 25 a 29 de julio de 2016*. Logroño: Instituto de Estudios Riojanos, 54-67.

Kitchin, Rob (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, April-June. Doi: 10.1177/2053951714528481.

Long, Sabine y Ommer, Björn (2021). Transforming Information Into Knowledge: How Computational Methods Reshape Art History. *Digital Humanities Quaterly*, 15(3). Disponible en: http://www.digitalhumanities.org/dhq/vol/15/3/000560/000560.html

Luis Corral, Fernando y Pérez, María (2017). Negotiating Fines in the Early Middle Ages: Local Communities, Mediators and the Instrumentalization of Justice in the Kingdom of León. *Al-Masāq*, 29(2): 172-185. Doi:10.1080/09503110.2017.1349980.

Martínez Sopena, Pascual (2004). Tradiciones y tendencias en el Medievalismo español. *BUCEMA*, 8. Doi:10.4000/cem.931.

Mauri, Alfred (2006). *La configuració del paisatge medieval: el comtat de Barcelona fins al segle XI* [Tesis doctoral]. Barcelona: Universitat de Barcelona. Disponible en: https://www.tdx.cat/handle/10803/2065#page=1.

May, Keith (2020). The Matrix: Connecting Time and Space in archaeological stratigraphic records and archives. *Internet Archaeology*, 55. Doi:10.11141/ia.55.8.

Medina, Sonia; del Fresno, Pablo y Travé, Esther (2022). Bridging the gap: A semantic-based architecture to bring together the archaeological record and written evidence [Poster]. *28th EAA Annual Meeting (Budapest). Digital Humanities and Medieval Archaeology: Re-Integrating Isolated, Fragmented and Overlooked Archaeological Evidence*. Doi: 10.13140/RG.2.2.17798.45122.

Mínguez, José María (2004). En torno a la génesis de las sociedades peninsulares altomedievales. Reflexiones y nuevas propuestas. *Studia Historica. Historia Medieval*, 22: 169-188.

Mínguez, José María (2019). Los aspectos diferenciales de la expansión astur-leonesa en los territorios orientales. En: F. J. Fernández Conde, J. M. Mínguez y E. Portela (eds.) *El reino de Hispania (siglos VIII-XII). Teoría y práctica del poder*. Madrid: Akal, 143-267.

Owens, Trevor (2011). Defining Data for Humanists: Text, Artifact, Information or Evidence?. *Journal of Digital Humanities*, 1(1). Disponible en: http://journalofdigitalhumanities.org/1-1/defining-data-for-humanists-by-trevor-owens/.

Quirós, Juan Antonio (2007). Las aldeas de los historiadores y de los arqueólogos en la Alta Edad Media del norte peninsular. *Territorio, Sociedad y Poder*, 2: 63-86.

Sánchez Pardo, José Carlos (2013). Power and rural landscapes in early medieval Galicia (400–900 AD): towards a re-incorporation of the archaeology into the historical narrative. *Early Medieval Europe*, 21(2): 140-168. Doi: 10.1111/emed.12013.

Schöch, Christof (2013). Big? Smart? Clean? Messy? Data in the Humanities. *Journal of Digital Humanities*, 2(3). Disponible en: http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/.

Steinman, Friedrich (2000). On the representation of roles in object-oriented and conceptual modelling. *Data & Knowledge Engineering*, 35: 83-106.

Thibodeau, Kenneth (2019). The Construction of the Past: Towards a Theory of Knowing the Past. *Information*, 10(332). Doi: 10.3390/info10110332.

Travé, Esther; del Fresno, Pablo y Mauri, Alfred (2020). Ontology-Mediated Historical Data Modeling: Theoretical and Practical Tools for an Integrated Construction of the Past. *Information*, 11(182). Doi: 10.3390/info11040182.

Travé, Esther; Medina, Sonia; del Fresno, Pablo; Vicens, Joan y Mauri, Alfred (2021): "Towards an Ontology-Driven Information System for Archaeological Pottery Studies: The Greyware Experience". *Applied Sciences*, 11(17):7989. Doi:10.3390/app11177989.

Tuominen, Jouni; Hyvönen, Eero y Leskinen, Petri (2018). Bio CRM: A Data Model for Representing Biographical Data for Prosopographical Research. En: A. Fokkens, S. ter Braake, R. Sluijter, P. Arthur y E. Wandl-Vogt (eds.). *Proceedings of the Second Conference on Biographical Data in a Digital World 2017. CEUR Workshop Proceedings; Vol. 2119.* RWTH Aachen, 59-66. Disponible en: http://ceur-ws.org/Vol-2119/paper10.pdf.

Turner, Sam (2006). Historic Landscape Characterisation: A landscape archaeology for research, management and planning. *Landscape Research*, 31:4: 385-398. Doi: 10.1080/01426390601004376.

Verschoof-van der Vaart, Wouter B.; Lambers, Karsten (2021). Applying automated object detection in archaeological practice: A case study from the southern Netherlands. *Archaeological Prospection*, 29(15): 15-31. Doi: 10.1002/arp.1833.

Wickham, Hadley (2014). Tidy Data. *Journal of Statistical Software*, 59(10): 1-23.

Wylie, Alison (2008). Mapping ignorance in archaeology. The Advances of Historical Hindsight. En: R. N. Proctor y L. Schiebinger (eds.). *Agnotology. The Making and Unmaking of Ignorance*. Stanford: Stanford University Press, 183-198.

Zadora-Rio, Elisabeht (1995). Le village des historiens e le village des archéologues. En: Elisabeth Mornet (dir.). *Campagnes médiévales. L'homme et son espace. Études offertes à Robert Fossier*. Paris: Publications de la Sorbonne,145-153.

# HILAME.
## DE LA PERSONA AL DATO, DEL DATO A LA PERSONA

Agurtzane Paz Moro, Teresa Jular Pérez-Alfaro

**Resumen**

En este trabajo se presenta la evolución de «HILAME (Hidalgos, Labradoras, Mercaderes)» desde su creación en 2016 hasta el momento actual, partiendo de su proceso creativo y sus presupuestos metodológicos y finalizando con los objetivos a futuro del proyecto. Además, se enfoca desde dos disciplinas interrelacionadas, dos puntos de vista complementarios, el de la Historia y el del Diseño.

La principal aportación es mostrar el estudio de caso desarrollado en 2021 con la participación de cinco alumnos de la Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU) quienes se han implicado activamente en el proyecto trabajando con una metodología, alternativa a la original y bautizada como «orfebrería de datos». Una experiencia de transferencia y aprendizajes, con aportaciones, tanto para la investigación histórica como para la mejora de la plataforma que la recoge.

**Palabras clave:** Historia, Edad Media, Prosopografía, Humanidades Digitales, Diseño.

Agurtzane Paz Moro, Universidad Pública de Navarra (UPNA)
0000-0002-6891-906X
agurtzane.paz@unavarra.es
Teresa Jular Pérez-Alfaro, XLI design+thinking
hola@teresajular.com

**Abstract**
This paper presents the evolution of «HILAME (Hidalgos, Labradoras, Mercaderes)» from its creation in 2016 to the present, starting from its creative process and its methodological assumptions and ending with the future objectives of the project. In addition, it focuses from two interrelated disciplines, two complementary points of view, that of History and that of Design.

The main contribution is to show the case study developed in 2021 with the participation of five students from the Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU) who have been actively involved in the project, working with a methodology alternative to the traditional one that we have called "data goldsmith". An experience of transfer and learning, with contributions for historical research and for the improvement of the platform that collects it.

**Keywords:** History, Middle Ages, Prosopography, Digital Humanities, Design.

## 1. PRESENTACIÓN DEL PROYECTO

HILAME, acrónimo de Hidalgos, Labradoras, Mercaderes, es un proyecto de Humanidades Digitales de carácter prosopográfico que busca recopilar los datos biográficos de todas aquellas personas que aparecen reflejadas en las fuentes documentales de época medieval e inicios de la moderna con el objeto de recomponer, en la medida de lo posible, sus transcursos vitales. Así, nuestro objetivo es reconstruir las redes sociales configuradas en el periodo y espacio escogido, inicialmente los territorios de la cornisa cantábrica, al final de la Edad Media.

Los tres términos recogidos en su acrónimo no han sido escogidos al azar, sino que presentan la dicotomía —campo, ciudad— de la sociedad medieval, añadiéndole además una perspectiva de género, en consonancia con la relevancia que paulatinamente está adquiriendo la Historia de las Mujeres entre la historiografía actual. A grandes rasgos, la palabra Hidalgos representa el ámbito rural, estructurado en torno a los señores que lo gobernaban. De hecho, el primer aspecto que abordó HILAME fue el conflicto banderizo que enfrentó a los linajes vizcaínos bajomedievales (Guerra, 1930). Posteriormente, la información sobre estos y otros linajes se ha ido ampliando con la inclusión de fuentes que completan otros aspectos de su *modus vivendi*: espacios de poder, fuentes de renta, patrimonio, entre otros (Goicolea *et al.*, 2005; Dacosta *et al.*, 2010; Munita *et al.*, 2014).

Asimismo, la referencia a las Labradoras encarna la perspectiva de género. La elección del término tiene su origen en el análisis que José Ramón Díaz de Durana (2004: 249-254) realizó sobre las labradoras alavesas, en el que dedica varios epígrafes a exponer los debates generados en torno a la condición social y las obligaciones fiscales de las labradoras que contraían nupcias con hidalgos en la Llanada alavesa, una práctica relativamente frecuente. No obstante, la mayor aportación sobre las mujeres que integran la base de datos procede de dos fuentes monásticas femeninas alavesas: los monasterios de Santa María de Barría y San Juan de Quejana (Paz, 2013 y 2017).




Figura 1. La plataforma HILAME en Internet. Fuente: www.hilame.info

Por último, los Mercaderes simbolizan el mundo urbano, el más infrarrepresentado hasta la fecha en nuestra plataforma web, circunstancia que, como se verá, estamos solventando con los resultados derivados del proyecto que aquí presentamos.

También es relevante la delimitación espacio cronológica del proyecto, establecida en relación a los objetivos de investigación de los grupos en los que el mismo se inserta[1]. HILAME está codirigido por José Ramón Díaz de Durana y Arsenio Dacosta, e integrado por un grupo que viene trabajando en los últimos años en el análisis de las transformaciones económicas, sociales, políticas, institucionales, culturales e ideológicas que acontecieron en los territorios del actual País Vasco y otros limítrofes entre

1. Proyecto de investigación «Violencia y transformaciones sociales en el nordeste de la Corona de Castilla (1200-1525)», financiado por el Ministerio de Ciencia e Innovación (ref. PID2021-124356NB-I00), y grupo de investigación «Sociedades, procesos, culturas (siglos VIII a XVIII)» financiado por el Gobierno Vasco (ref. IT1465-22).

los siglos XIV y XVIII. En este contexto, uno de sus principales ámbitos de estudio es la conflictividad banderiza característica de los siglos bajomedievales. Por eso, el estudio prosopográfico de sus protagonistas y de sus redes sociales resulta un elemento de análisis histórico determinante para entender y profundizar en el conocimiento del proceso.

El trabajo prosopográfico parte de la combinación de dos conceptos, el de «contexto» y el de «factoide». El contexto es el fragmento textual tomado de una fuente escrita, en este caso del periodo bajomedieval, que contiene una determinada información. Es un concepto fundamental para HILAME por dos motivos: por ofrecer al investigador la información susceptible de ser procesada y por garantizar que el usuario pueda verificar la información que está consultando en su contexto original. Mientras que el factoide (Tinti, 2003), tal como lo introduce Dion Smythe en la *Prosopography of Byzantine Empire* (PBE) (Smythe, 2007), es una forma de categorizar las informaciones que aparecen en los testimonios medievales referidas a un individuo, normalizándolas para su tratamiento masivo y para la correlación de coincidencias. Estos factoides están divididos en diversas categorías que nos permiten documentar desde cargos y oficios individuales o colectivos hasta eventos vitales, políticos o sociales, sin obviar los vínculos jurídicos y familiares de cada uno de los individuos documentados en las fuentes primarias.

El procesamiento informático para el análisis heurístico de la documentación, la comunicación pública de resultados en tiempo real, el debate crítico sobre las Humanidades Digitales y la búsqueda de una difícil pero deseable transdisciplinariedad, entendida como la mayor interacción posible entre las disciplinas en juego y estadio siguiente a la interdisciplinariedad establecida de partida en HILAME, también forman parte de los objetivos de este proyecto de investigación.

El proyecto reúne a medievalistas y modernistas expertos en distintos campos y a historiadores e historiadoras de la lengua

y la literatura familiarizados con el análisis de la documentación, las crónicas y la producción literaria de estas épocas. Entre las personas del equipo de Humanidades, la investigadora y docente Agurtzane Paz, experta en el tratamiento prosopográfico de las fuentes medievales trabajadas. A ellos se sumó un pequeño grupo de diseñadores e informáticos externos, incorporados desde el comienzo para, todos juntos, cocrear una cultura de proyecto innovadora, acorde a los objetivos y principios del mismo. Entre estos profesionales, Teresa Jular, responsable de la coordinación de las tareas técnicas y la comunicación digital.

## 2. CRONOLOGÍA DEL PROYECTO

HILAME nació en 2016 como proyecto orientado a la Ciencia Abierta, con unos principios de transparencia y conocimiento compartidos comunicados a través de múltiples canales, publicaciones y eventos. Desde sus inicios, el proyecto ha aspirado a la convergencia e interoperabilidad con otros (Charle, 1985; Millet, 1987; Boute, 2002), logro conseguido en 2020 con *Scripta manent*[2]. Esto ha permitido incorporar nuevas fuentes documentales y varios cientos de nuevos individuos a la base de datos principal, además de sumar una nueva base de datos secundaria a las existentes: «la quinta bola» (*véase Figura 2*).

El proceso de ideación y desarrollo del proyecto HILAME incluyó muy pronto asuntos relacionados con el diseño y la comunicación, abordados colaborativamente: la definición del nombre, la identidad gráfica, la arquitectura de la información para las interfaces web y algunas primeras aplicaciones para la difusión pública del mismo. Se creó la sede en Internet, www.hilame.info,

2. Proyecto de investigación coordinado por Cristina Jular Pérez-Alfaro (IH-CCHS, CSIC) «Scripta manent III: de registros privados a textos públicos. Un archivo medieval en la Red» financiado por el Ministerio de Ciencia e Innovación (ref. PID2020-116104RB-I00)




Figura 2. Las bases de datos principal y secundarias de HILAME
Fuente: www.hilame.info

con licencia *creative commons* by-nc-sa, que detalla el proyecto, presenta al equipo investigador y muestra el trabajo en proceso en varias bases de datos complementarias, cinco hasta el momento.

Paralelamente, se desarrolló, en PHP, una base de datos prosopográfica a medida para la alimentación de las bases de datos públicas por parte de documentalistas entrenados en la misma y, más recientemente, un importador de datos tabulados en Excel y desarrollado también en PHP, como método alternativo para una carga masiva de datos de documentalistas de otros proyectos.

Esto proporciona un acceso inmediato a los resultados a medida que se desarrolla la investigación, en tiempo real, algo poco frecuente en el mundo académico, al menos en aquel momento, hace más de un lustro.

Asimismo, se han concebido tutoriales para facilitar su uso a distintos perfiles de interesados, se incluyen referencias webgráficas y bibliográficas, con algunos libros completos accesibles para su descarga, así como audiovisuales de divulgación científica para mejor entendimiento del contexto histórico en

el que se desenvolvían las personas identificadas a partir de las fuentes históricas tratadas.

Una infografía actualizada con frecuencia, «HILAME en cifras», nos permite resumir el proyecto en una sola hoja (*véase Figura 3*).

Se han creado perfiles en diferentes canales para compartir y dar visibilidad pública a HILAME: un perfil en Academia.edu, que recoge artículos y eventos científicos relacionados con el proyecto; otro en Twitter, desde el que se comunican avances del proyecto y de los miembros del equipo en tiempo real, se dialoga sobre #HumanidadesDigitales en general y sobre #Historia en la #EdadMedia de manera específica; un canal en YouTube que, creado en 2016 con varias entrevistas sobre el concepto de hidalguía y las guerras banderizas, se ha ido enriqueciendo con ponencias y talleres disponibles para personas interesadas en la evolución del proyecto. El más reciente es un canal de podcast «HILAME habla» en ivoox, estrenado en septiembre de 2021 para audioentrevistas y contenidos en clave periodística. Para el seguimiento de resultados, nos apoyamos en las Analytics de Google y de Twitter, y en otras herramientas abiertas, como los analizadores Tweet Topic Explorer y Orbit.

El equipo HILAME ha participado en diversos eventos y congresos que han dado como resultado publicaciones sobre la evolución del proyecto en sus distintas vertientes (Dacosta y Díaz de Durana, 2017; Dacosta y Jular, 2018; Jular y Dacosta, 2018; Dacosta y Jular, 2020). Como aportación genuina, cabría destacar que el valor del Diseño ha entrado a formar parte de los debates académicos especializados en Historia y la Historia ha sido protagonista en actividades y publicaciones propias del ecosistema del Diseño.

En este sentido, sin ánimo de ser exhaustivas, creemos relevante reseñar algunos desafíos que hemos abordado en los últimos años, por ser una apuesta a futuro y por reflejar el carácter innovador del proyecto. Por un lado, el desafío de la

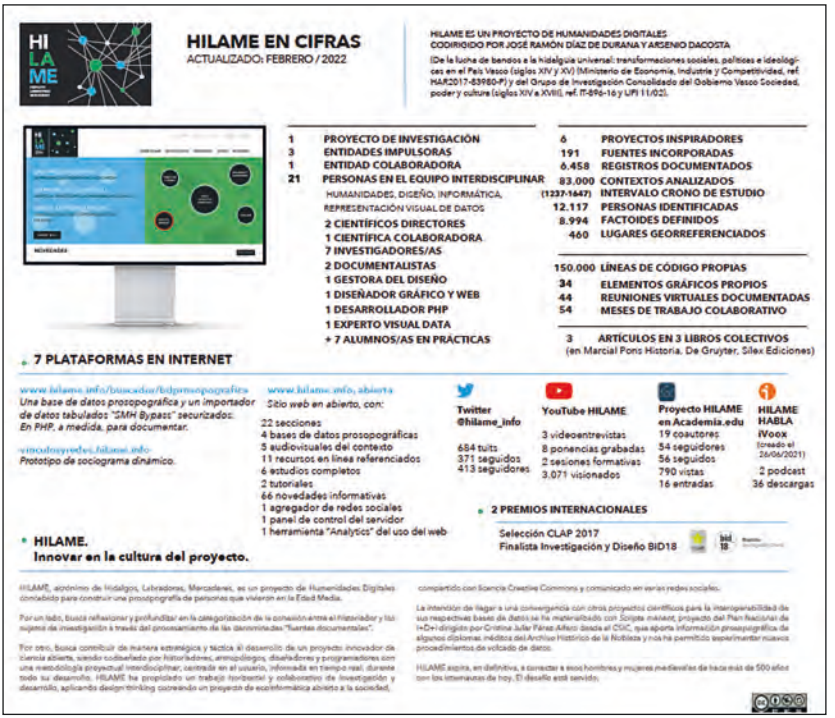



Figura 3. Infografía «HILAME EN CIFRAS. febrero 2022»
Fuente: xli design+thinking para HILAME

*n-dimensionalidad*, nuevos modos de interrogar a nuestras bases de datos, innovar con ayuda del pensamiento visual y la representación interactiva de la información extraída de las fuentes documentales. Por otro, la interoperabilidad con el proyecto *Scripta manent*, que ha permitido el enriquecimiento mutuo de ambas plataformas. Finalmente, la introducción de mejoras para la usabilidad y la proyección pública de HILAME (Baraibar, 2014), aunque sin olvidar nuestro principal propósito: facilitar la producción científica. Esta última iniciativa es el punto de partida del estudio de caso que aquí presentamos.

## 3. TESTEO, VALIDACIÓN Y MEJORAS EN EL SITIO WEB DE HILAME

Durante el confinamiento y meses posteriores, abordamos el desarrollo de mejoras en la usabilidad de HILAME, tarea que vino precedida de una llamada al testeo y recepción de validaciones y sugerencias. Fueron invitadas una decena de personas de variados perfiles, concretamente colegas no implicados en el proyecto, estudiantes y algunos miembros del equipo, con los que dialogamos sobre distintos aspectos del sitio web público: clima visual, organización de la información, interfaces, bases de datos, referencias y otros contenidos didácticos... Aprovechamos el momento que vivíamos, sin actividad pública, para hacer un análisis no normativo pero sí transparente, para contar con aportaciones de audiencias complementarias antes de abordar los cambios.

Además de ampliar las Novedades, que ahora recogen no solo las actividades del proyecto en sí sino también las actividades de los distintos componentes del equipo, se hicieron modificaciones destinadas a facilitar la consulta de los contenidos

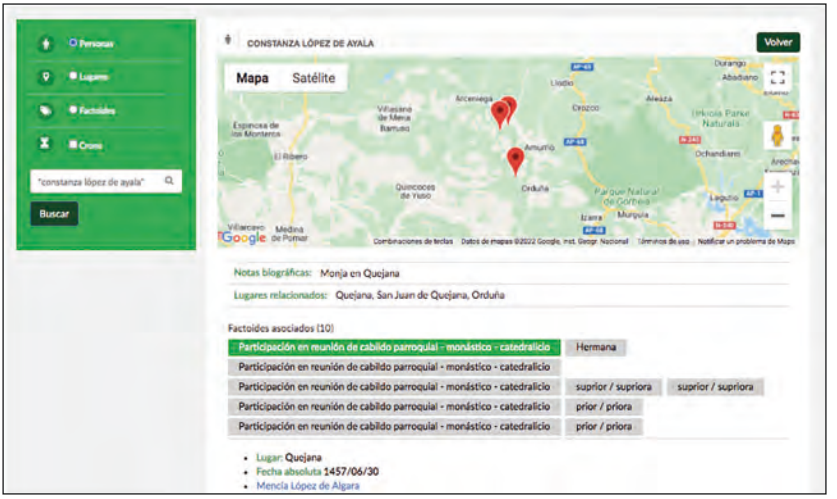



Figura 4. Resultado de una consulta con texto exacto en base datos principal HILAME. Fuente: https://www.hilame.info/buscador/

en las bases de datos principal y secundarias, donde mejoramos sensiblemente la precisión de las consultas y la organización de los resultados (*véase Figura 4*).

Y es precisamente en este proceso de mejorar HILAME —como paso previo y paralelo al mismo— donde se incorporó el nuevo proyecto de transferencia que presentamos en forma de comunicación en el congreso de HDH de 2021, con el mismo título que este artículo: «HILAME. De la persona al dato, del dato a la persona».

En este contexto, no se trataba de una transferencia de conocimientos en sentido estricto, sino un ejercicio de transparencia y de comunicación de nuestro trabajo, que fue transmitido a un grupo de cinco estudiantes que cursaban el segundo ciclo del grado de Historia en la Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU).

## 4. LA TRANSFERENCIA: ESTUDIO DE CASO

Este grupo de estudiantes se integró en el proyecto a través de un programa de prácticas voluntarias que la UPV/EHU ofrece a los estudiantes de segundo ciclo de su grado de Historia, particularmente a aquellos interesados en la Historia Medieval. El programa se desarrolló entre noviembre de 2020 y abril de 2021 (se repitió entre las mismas fechas del curso siguiente, el 2021-2022, y ya hay tres aspirantes para continuar el trabajo en el curso actual, 2022-2023), con una asignación mínima de 150 horas (aunque ampliable en otras 150 más), que permite la convalidación de una asignatura. Inicialmente, el programa se desarrolló en tres fases. En primer lugar, una participación activa en la evaluación de la plataforma a través de un cuestionario preparado para la ocasión. El cuestionario constaba de treinta preguntas, que podríamos clasificar en tres categorías.

Por un lado, las dirigidas a favorecer el conocimiento de la plataforma y detectar su amigabilidad para un usuario ajeno al proyecto y a los recursos y canales ofrecidos en Internet. Este conjunto de preguntas nos permitió alcanzar dos objetivos: que los

estudiantes se familiarizaran con el funcionamiento del sitio web y que nos proporcionaran información sobre su experiencia personal respecto al acceso a los contenidos, la claridad de los mismos y la idoneidad de la oferta web de HILAME, entre otras cuestiones. Además de las preguntas para identificar su comportamiento como usuarios, se incluyeron otras en las que se buscó su implicación en el proyecto, permitiéndoles aportar una perspectiva de mejora.

Asimismo, les planteamos preguntas destinadas al testeo de los buscadores. Para ello, proyectamos la realización de consultas concretas de linajes o individuos ya documentados en la plataforma. La resolución de cada pregunta requería indagar en sus diversas bases de datos, la principal y las otras cuatro que le acompañan: Oñacinos y gamboínos, *Scripta manent*, Linajes y Vínculos y redes. Y también en las opciones de búsqueda (general o por linajes, fuentes o contextos) de la propia base de datos principal.

En última instancia, se incorporaron preguntas destinadas al apartado donde se encuentran los recursos audiovisuales:




Figura 5. Primer nivel de la organización de factoides. Fuente: Proyecto HILAME

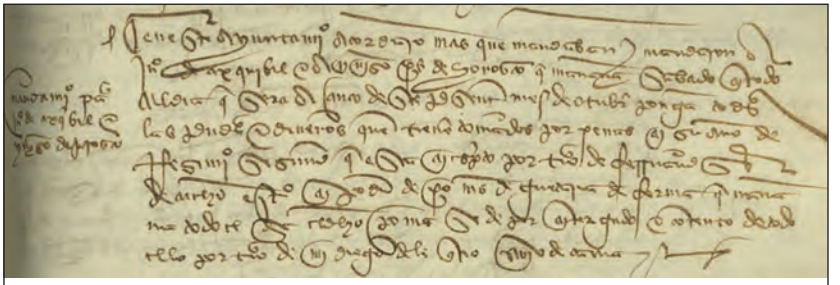
vídeos varios, publicaciones relacionadas con el contenido de la prosopografía y también con la metodología de trabajo en HILAME.

Esta iniciativa, de gran valor por aportar miradas nuevas, nos proporcionó una visión más imparcial de su funcionalidad y una perspectiva a nivel de usuario de las necesidades a cubrir. Por tanto, consideramos que ha sido una buena actividad para implementar las mejoras en HILAME y avanzar en su dimensión pública.

En una segunda fase, buscamos dotar a los estudiantes de los instrumentos necesarios para el desarrollo de su trabajo, pues todos ellos carecían de formación previa en HILAME en todos sus aspectos (proyecto, procesamiento de fuentes documentales o metodología de trabajo) y en la prosopografía como método de análisis histórico. Para ello, elaboramos un ciclo formativo, titulado *Jornadas formativas sobre Humanidades Digitales: el proyecto HILAME*, online debido a las circunstancias y dividido en dos sesiones[3].

La primera sesión se destinó a presentar el proyecto, introduciendo algunos de sus signos identitarios. Primeramente, el «diseño ecoinformático» (Dacosta y Jular, 2018), que implica la integración de profesionales de diferentes especialidades como historiadores, antropólogos, diseñadores, técnicos informáticos, etc., en la génesis, el desarrollo del proyecto y el seguimiento documentado de los avances del mismo. En segundo lugar, el fundamento teórico de la plataforma y su conceptualización desde el punto de vista histórico: la prosopografía y el uso de factoides para la descripción de individuos. Y, finalmente, la metodología a emplear para el volcado de datos, la que hemos denominado «orfebrería de datos», fundamentada en la selección de un contexto que nos permite situar en tiempo y lugar a cada uno de los sujetos históricos citados en la documentación a manejar, así como establecer sus vínculos sociales, completando su ficha de forma manual y personalizada (*véase Figura 5*).

3. Accesible en https://www.youtube.com/channel/UChQ390pDZ56d9LloSBKya_A [Enlace visitado el 19 de septiembre de 2022]

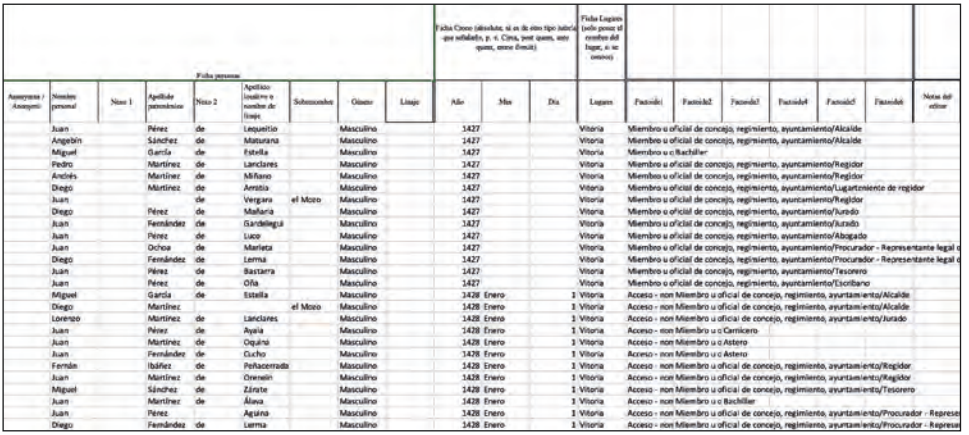


<Mandamiento para Johan de Axquibil e Ynnigo de Horozco> (*al margen, resumen del contenido*). En este ayuntamiento, acordaron más que mandaban e mandaron a Johan de Axquibil e a Ynnigo Pérez de Horozco que mannana sábado, en todo el día, que será a çinco de este presente mes de otubre, pongan todas las prendas e dineros que tienen tomados por penas en su anno de regimiento, segúnd que está escrito por testimonio de Ferrnand Sáez de Cucho, escrivano, en poder de Pero Martínez de Guetaria, de forma que mannana, todo el día, el dicho Pero Martínez se dé por entregado e contento de todo ello. Por testimonio de mí, Diego de Lequetio, escrivano de cámara.

Figura 6. Ejemplo de las Actas Municipales de Vitoria. Estructura, grafía y transcripción
Fuente: Archivo Municipal de Vitoria (A.M.V.), Sección 12, Legajo 20, fol. 94r.º. Documento digitalizado en https://www.vitoria-gasteiz.org/wb021/was/contenidoAction.do?idioma=es&uid=u721bcfc_1357a52d91b___7fef [Enlace visitado el 19 de septiembre de 2022]

La segunda sesión se centró en la labor que iban a tener que desarrollar, profundizando en la metodología de trabajo específica, las fuentes y sus características, así como los problemas y dificultades más habituales que podrían encontrar.

En este contexto, optamos por procesar un tipo documental concreto: las Actas Municipales de Vitoria-Gasteiz (A.M.V.), que están digitalizadas y en acceso abierto a través de la web pública del Ayuntamiento de Vitoria. Es esta una documentación de gran valor para el estudio de la sociedad y del funcionamiento urbano, cuyo primer manuscrito se retrotrae a 1428 y que se conserva ininterrumpidamente desde septiembre de 1479 hasta la actualidad, si bien es cierto que la digitalización solo se extiende hasta 1775, incorporando también los años finales de la guerra de la Independencia (1812-1814), coincidentes con la Batalla de Vitoria (21 de junio de 1813).



| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | Ficha Cnoo (absoluta, si es de otro tipo indicar que señalen, p. e. Cinu, yest quem, anty quem, cnoo fenit) | | | Ficha Lugares (solo poner el nombre del lugar, si se conoce) | | | | | | | | | |
| | | | | Ficha personas | | | | | | | | | | | | | | | | | | | |
| Asentamiento / Asentamiento | Nombre personal | Nexo 1 | Apellido patronímico | Nexo 2 | Apellido toponímico o nombre de linaje | Sobrenombre | Género | Linaje | | Año | Mes | Día | Lugares | Factoide1 | Factoide2 | Factoide3 | Factoide4 | Factoide5 | Factoide6 | | Notas del editor |
| | Juan | | Pérez | de | Lequeitio | | Masculino | | | 1427 | | | Vitoria | Miembro u oficial de concejo, regimiento, ayuntamiento/Alcalde | | | | | | | |
| | Angebin | | Sánchez | de | Maturana | | Masculino | | | 1427 | | | Vitoria | Miembro u oficial de concejo, regimiento, ayuntamiento/Alcalde | | | | | | | |
| | Miguel | | García | de | Estella | | Masculino | | | 1427 | | | Vitoria | Miembro u o Bachiller | | | | | | | |
| | Pedro | | Martínez | de | Lanclares | | Masculino | | | 1427 | | | Vitoria | Miembro u oficial de concejo, regimiento, ayuntamiento/Regidor | | | | | | | |
| | Andrés | | Martínez | de | Miñano | | Masculino | | | 1427 | | | Vitoria | Miembro u oficial de concejo, regimiento, ayuntamiento/Regidor | | | | | | | |
| | Diego | | Martínez | de | Armiño | | Masculino | | | 1427 | | | Vitoria | Miembro u oficial de concejo, regimiento, ayuntamiento/Lugartiniente de regidor | | | | | | | |
| | Juan | | | de | Vergara | el Mozo | Masculino | | | 1427 | | | Vitoria | Miembro u oficial de concejo, regimiento, ayuntamiento/Regidor | | | | | | | |
| | Diego | | Pérez | de | Mañaria | | Masculino | | | 1427 | | | Vitoria | Miembro u oficial de concejo, regimiento, ayuntamiento/Jurado | | | | | | | |
| | Juan | | Fernández | de | Gardelegui | | Masculino | | | 1427 | | | Vitoria | Miembro u oficial de concejo, regimiento, ayuntamiento/Jurado | | | | | | | |
| | Juan | | Pérez | de | Luco | | Masculino | | | 1427 | | | Vitoria | Miembro u oficial de concejo, regimiento, ayuntamiento/Abogado | | | | | | | |
| | Juan | | Ochoa | de | Marieta | | Masculino | | | 1427 | | | Vitoria | Miembro u oficial de concejo, regimiento, ayuntamiento/Procurador - Representante legal | | | | | | | |
| | Diego | | Fernández | de | Lerma | | Masculino | | | 1427 | | | Vitoria | Miembro u oficial de concejo, regimiento, ayuntamiento/Procurador - Representante legal | | | | | | | |
| | Juan | | Pérez | de | Bastarra | | Masculino | | | 1427 | | | Vitoria | Miembro u oficial de concejo, regimiento, ayuntamiento/Tesorero | | | | | | | |
| | Juan | | Pérez | de | Oña | | Masculino | | | 1427 | | | Vitoria | Miembro u oficial de concejo, regimiento, ayuntamiento/Escribano | | | | | | | |
| | Miguel | | García | de | Estella | | Masculino | | | 1428 | Enero | 1 | Vitoria | Acceso - non Miembro u oficial de concejo, regimiento, ayuntamiento/Alcalde | | | | | | | |
| | Diego | | Martínez | | | el Mozo | Masculino | | | 1428 | Enero | 1 | Vitoria | Acceso - non Miembro u oficial de concejo, regimiento, ayuntamiento/Alcalde | | | | | | | |
| | Lorenzo | | Martínez | de | Lanclares | | Masculino | | | 1428 | Enero | 1 | Vitoria | Acceso - non Miembro u oficial de concejo, regimiento, ayuntamiento/Jurado | | | | | | | |
| | Juan | | Pérez | de | Ayala | | Masculino | | | 1428 | Enero | 1 | Vitoria | Acceso - non Miembro u o Carnicero | | | | | | | |
| | Juan | | Martínez | de | Oquina | | Masculino | | | 1428 | Enero | 1 | Vitoria | Acceso - non Miembro u o Astero | | | | | | | |
| | Juan | | Fernández | de | Cucho | | Masculino | | | 1428 | Enero | 1 | Vitoria | Acceso - non Miembro u o Astero | | | | | | | |
| | Fernán | | Ibáñez | de | Peñacerrada | | Masculino | | | 1428 | Enero | 1 | Vitoria | Acceso - non Miembro u oficial de concejo, regimiento, ayuntamiento/Regidor | | | | | | | |
| | Juan | | Martínez | de | Orneain | | Masculino | | | 1428 | Enero | 1 | Vitoria | Acceso - non Miembro u oficial de concejo, regimiento, ayuntamiento/Regidor | | | | | | | |
| | Miguel | | Sánchez | de | Zárate | | Masculino | | | 1428 | Enero | 1 | Vitoria | Acceso - non Miembro u oficial de concejo, regimiento, ayuntamiento/Tesorero | | | | | | | |
| | Juan | | Martínez | de | Álava | | Masculino | | | 1428 | Enero | 1 | Vitoria | Acceso - non Miembro u o Bachiller | | | | | | | |
| | Juan | | Pérez | de | Aguina | | Masculino | | | 1428 | Enero | 1 | Vitoria | Acceso - non Miembro u oficial de concejo, regimiento, ayuntamiento/Procurador - Represe | | | | | | | |
| | Diego | | Fernández | de | Lerma | | Masculino | | | 1428 | Enero | 1 | Vitoria | Acceso - non Miembro u oficial de concejo, regimiento, ayuntamiento/Procurador - Represe | | | | | | | |

Figura 7. Detalle de plantilla Excel para el procesamiento de datos
Fuente: Proyecto HILAME

Asimismo, presenta unas características formales bien definidas para la fecha: la periodicidad anual del curso político, con el nombramiento de oficiales el 29 de septiembre —día de San Miguel—; la estructura de los acuerdos tomados en cada reunión concejil, que aportan la información necesaria y están precedidos de un resumen de su contenido; y el tipo gráfico, cortesana que está evolucionando hacia procesal, poco cuidado y repleto de abreviaturas (*véase Figura 6*).

A pesar de ser una fuente inédita, a excepción de un documento previo, el que recoge los acuerdos municipales del año 1428 y parte de los del 29 (Díaz de Durana, 1984), su edición documental es un proyecto en ciernes de nuestro grupo de investigación. Por tanto, hemos podido facilitar al alumnado una transcripción provisional, aunque pulida, de los manuscritos a procesar. La metodología escogida para el procesamiento documental ha seguido un sistema indirecto, consistente en el volcado de datos en una plantilla de Excel que nos permite describir un gran volumen de individuos de forma seriada con el empleo de pocos factoides (*véase Figura 7*).

Posteriormente, los datos ya importados son revisados por los documentalistas del equipo. Esta es una labor imprescindible

para poder solucionar los problemas asociados a la metodología de trabajo, que pone a la persona en primer lugar y relega los factoides. Así, resulta esencial establecer los vínculos jurídicos y lazos familiares entre individuos (Dacosta y Jular, 2020) y unificar bajo un denominador común los registros correspondientes a un mismo individuo, recurriendo para ello a una función de nuestra herramienta que hemos denominado «el fusionador de personas».

Esta metodología de procesamiento documental ya había sido previamente empleada en la importación de datos de las bases secundarias «Oñacinos y gamboínos» y «*Scripta manent*», lo que corrobora su eficacia y pertinencia para ser aplicada también en este caso.

Finalmente, acompañados por la tutora en consultas y revisiones periódicas del trabajo, los estudiantes han llevado a cabo el procesamiento documental completo, que ha sido particularmente enriquecedor desde dos perspectivas.

Por un lado, desde la perspectiva de los alumnos, convirtiéndose en una herramienta idónea para profundizar en su formación medievalista, especialmente entre aquellos interesados en la historia bajomedieval, pues favorecemos un primer contacto estrecho con las fuentes primarias y su tratamiento. De este modo, gracias a la transcripción que reciben y al acceso abierto a los documentos digitalizados, se familiarizan con la metodología de edición de fuentes, su estructura, el léxico propio del periodo, los tipos escriturarios de finales de la Edad Media, los usos y costumbres urbanos, el funcionamiento de una organización concejil de envergadura, las competencias que asumían sus oficiales, etc. En definitiva, una inmersión en la metodología de trabajo para el análisis histórico en el periodo que transcurre entre la Baja Edad Media y la Modernidad.

Por otro lado, desde el punto de vista del grupo de investigación y del proyecto, ha sido una gran oportunidad para ampliar los datos disponibles en la plataforma con registros de las Actas de Vitoria. Tal y como hemos señalado en la presentación

del proyecto, este es un objetivo esencial por cuanto el mundo urbano es el más infrarrepresentado hasta la fecha. Así, las Actas Municipales de Vitoria contribuyen a completar y diversificar aún más el contenido de HILAME, ilustrando las peculiaridades de la ciudad a través de los oficios concejiles y las familias que los ostentaban, sus profesiones y actividades económicas principales, los conflictos judiciales en los que se veían inmersos, la participación de las minorías étnicas —especialmente los judíos— en diversas actividades socioeconómicas, entre otras cuestiones.

## 5. VALORACIÓN Y RESULTADOS

Al finalizar esta experiencia piloto (julio 2021), los alumnos han tenido ocasión de valorar la iniciativa a través de un breve cuestionario predeterminado por el propio programa de prácticas. Sus respuestas nos permiten afirmar que la experiencia ha sido positiva y ha proporcionado a nuestros estudiantes una herramienta de inmersión en la historia medieval. De hecho, entre sus comentarios destaca su utilidad para facilitar el acercamiento a unas fuentes primarias que entienden como instrumento para conocer la sociedad medieval en su conjunto y, particularmente el mundo urbano, las características socioeconómicas de su población y el funcionamiento del concejo.

Junto a esta valoración positiva, también han señalado algunos aspectos susceptibles de ser mejorados y todos coinciden en la necesidad de recibir una formación más específica que ejemplifique con más detalle la metodología de trabajo. De hecho, las mayores dificultades que han tenido que afrontar los estudiantes se han manifestado esencialmente en el aspecto metodológico, siendo la selección de contextos y la asignación de factoides los mayores escollos. En este sentido, consideramos que estos son problemas inherentes a la metodología de trabajo, al desconocimiento de la plataforma y a su inexperiencia. Para resolverlos y superar las dificultades iniciales, es para lo que se ofreció la orientación personalizada antes mencionada.

Además, hemos concluido el volcado de todos los individuos contenidos en las Actas Municipales de la villa/ciudad de Vitoria desde 1479 a 1484, incluyendo los del manuscrito de los años 1428 y 1429, así como el conjunto de individuos correspondientes a los años 1479-80 y 1480-81. El resultado ha sido la introducción de 586 nuevas personas en HILAME, lo que ha implicado incorporar una relación completa de los oficiales concejiles de ese periodo y de otros vecinos pertenecientes a gremios o minorías étnicas. Esta circunstancia aporta una mirada de conjunto a esos datos, facilitando el análisis de la organización socioeconómica de la Vitoria de finales del siglo XV, sus usos y costumbres, las familias preeminentes, la conflictividad política y social, entre otras muchas cuestiones. Este es un resultado notable: sin la participación de estos cinco alumnos de la UPV/EHU no habríamos alcanzado un grado de procesamiento documental tan detallado.

## 6. COLOFÓN

Ciertamente, la experiencia anterior nos ha ayudado a mejorar la usabilidad de nuestra web, ha enriquecido nuestras bases de datos con nuevos registros y nos ha mostrado un camino para que nuestra investigación y nuestros planteamientos metodológicos puedan aportar más. Por tanto, aprovechamos la ocasión para agradecer su trabajo a Eva Valdivia, Julen Plágaro, Mikel Bilbao, Óscar González y Pablo Ortega, precursores en esta iniciativa.

HILAME es un proyecto ambicioso en lo tocante a sus objetivos, una búsqueda constante de innovación científica para un mayor retorno social. Por eso, nuestros próximos movimientos podrían resumirse así:

Por un lado, vamos a ampliar el *corpus* de individuos contenidos en la documentación publicada, incluyendo en la medida de lo posible a las mujeres de nuestras fuentes. En este sentido, cabe destacar que la documentación editada para los territorios del norte peninsular en la baja Edad Media y particularmente para

el País Vasco es profusa. Además de las colecciones documentales editadas por el grupo de investigación de la Universidad del País Vasco dirigido por José Ramón Díaz de Durana, son imprescindibles los libros publicados por la Sociedad de Estudios Vascos-Eusko Ikaskuntza, impulsora y protagonista de esta práctica con su colección Fuentes Documentales del País Vasco, que se compone (hasta la fecha) de 156 volúmenes[4].

Por otro lado, buscamos progresar en la usabilidad y transferencia de nuestra plataforma web promoviendo la convergencia con otros proyectos afines y la interoperabilidad entre las bases de datos respectivas, propiciando un enriquecimiento mutuo.

Además, tenemos gran interés en avanzar en la representación visual e interactiva del resultado de nuestras investigaciones. Seguir la línea de investigación que iniciamos con el prototipo Gephi «Vínculos y redes», donde elegimos el género como una primera dimensión a tratar, para llegar a una n-dimensionalidad que nos deje más satisfechas en nuestra intención de continuar aportando al desarrollo de las Humanidades Digitales de manera abierta, con una herramienta que, además de responder a nuestras consultas, nos proponga nuevos interrogantes sobre el pasado.

Y por último, pero no menos importante, queremos seguir aportando a los desafíos que nos plantean desde, entre otros aspectos, la interdisciplinariedad y la transdisciplinariedad, las innovaciones tecnológicas y sociales, la ciencia abierta, la sostenibilidad, el retorno a la sociedad y el trabajo colectivo en red, en ámbitos nacional e internacional, tal y como hemos pretendido hacer con esta contribución.

4. Accesibles en https://www.eusko-ikaskuntza.eus/es/publicaciones/fuentes-documentales-medievales-del-pais-vasco/bi-127/ [Enlace visitado el 19 de septiembre de 2022]

## REFERENCIAS BIBLIOGRÁFICAS


Baraibar, Álvaro (2014). *Visibilidad y divulgación de la investigación desde las Humanidades digitales: experiencias y proyectos*. Pamplona: Universidad de Navarra.

Boute, Bruno (2002). Towards more uniform database structures for prosopographical research: work in progress in University history – the example of the Lovanienses Database. En: K.S.B. Keats-Rohan (ed.). *Resourcing Sources. The Use of Computers in Developing Prosopographical Methodology*. Oxford: University of Oxford, pp. 35-48.

Charle, Christophe (1985). Problèmes de traitement informatique d'une enquête sur trois élites en 1901. En: Hélène Millet (ed.). *Informatique et prosopographie*. París: CNRS, pp. 233-246.

Dacosta, Arsenio y Díaz de Durana, José Ramón (2017). Prosopografía y bases de datos. Desafíos teóricos y metodológicos para el estudio de la Edad Media. En: Ana Isabel Carrasco Manchado (dir.). *El historiador frente a las palabras. Lenguaje, poder y política en la sociedad medieval: nuevas herramientas y propuestas*. Lugo: Axac, pp. 191-217.

Dacosta, Arsenio y Jular, Teresa (2018). Hidalgos en la escalera del diseño. En: Arsenio Dacosta, Cristina Jular y José Ramón Díaz de Durana (eds.). *Hidalgos e hidalguía en la Península Ibérica (siglos XII-XV)*. Madrid: Marcial Pons, pp. 429-446.

Dacosta, Arsenio y Jular, Teresa (2020). HILAME ante los desafíos de las Humanidades Digitales: convergencia e interoperabilidad. En: Sandra de la Torre, Ekaitz Etxeberria y José Ramón Díaz de Durana (coords.). *Valer más en la tierra, Poder, violencia y linaje en el País Vasco bajomedieval*. Madrid: Sílex, pp. 87-100.

Dacosta, Arsenio; Lema, José Ángel; Munita, José Antonio y Díaz de Durana, José Ramón (2010). *Poder y privilegio. Nuevos textos para el estudio de la nobleza vizcaína al final de la Edad Media (1416-1527)*. Bilbao: EHU Press.

Díaz de Durana, José Ramón (1984). *Vitoria a finales de la Edad Media (1428-1476)*. Vitoria: Diputación Foral de Álava.

Díaz de Durana, José Ramón (2004). *La otra nobleza, escuderos e hidalgos sin nombre y sin historia. Hidalgos e hidalguía universal en el País Vasco al final de la Edad Media (1250-1525)*. Bilbao: EHU Press.

Goicolea, Francisco Javier; Villanueva, Eider; Lema José Ángel; Fernández de Larrea, Jon Andoni; Munita, José Antonio y Díaz de Durana, José Ramón (2005). *Honra de hidalgos, yugo de labradores: nuevos textos para el estudio de la sociedad rural alavesa (1332-1521)*. Bilbao: EHU Press.

Guerra, Juan Carlos (1930). *Oñacinos y gamboinos. Rol de banderizos vascos, con la mención de las familias pobladoras de Bilbao en los siglos XIV y XV*. San Sebastián: Joaquín Muñoz Baroja.

Julár, Teresa y Dacosta, Arsenio (2018). HILAME (Hidalgos, Labradoras, Mercaderes): procesamiento y visualización de datos prosopográficos. En: *Humanidades Digitales. Miradas hacia la Edad Media*. Berlín: De Gruyter, pp. 79-89.

Millet, Hélène (1987). From sources to data. The construction of a prosopographical data-bank. En: Peter Denley y Deian Hopkin (eds.). *History and computing*. Manchester: Manchester University Press, pp. 63-67.

Munita, José Antonio; Dacosta, Arsenio; Lema José Ángel; Paz, Agurtzane y Díaz de Durana, José Ramón (2014). *«En tiempo de ruidos e bandos». Nuevos textos para el estudio de los linajes vizcaínos: los Barroeta de la merindad de Marquina (1355-1547)*. Bilbao: EHU Press.

Paz Moro, Agurtzane (2013). *El monasterio de Barría. Historia y documentos (1232-1524)*. Bilbao: EHU Press.

Paz Moro, Agurtzane (2017). *Colección documental del monasterio de San Juan de Quejana (Álava, 1332-1525)*. Bilbao: EHU Press.

Smythe, Dion C. (2007). A whiter shade of Pale: issues and possibilities in prosopography. En: K. S. B. Keats-Rohan (ed.). *Prosopography approaches and applications: a handbook*. Oxford: University of Oxford, pp. 127-137.

Tinti, Francesca (2003). Facing factoids. En: *King's College London, Second PASE Colloquium, 26 April 2003.*

# ESTRATEGIAS DE RECUPERACIÓN DE INFORMACIÓN MEDIANTE *POOLING* PARA LA CONSTRUCCIÓN DE COLECCIONES DE REFERENCIA DESDE REDES SOCIALES

## CASO DE ESTUDIO DURANTE LAS TENSIONES RACIALES DE 2020

Patricia Martín-Rodilla, David Otero

**Resumen:**
Las redes sociales son una valiosa fuente de información para documentar los procesos de surgimiento o evolución en patrimonio cultural, también para obtener una instantánea en tiempo real de un tema de investigación dentro del mismo. Muchos investigadores en este campo emplean las redes sociales para estudiar estos procesos, creando, para ello, colecciones que constituyen archivos digitales, potencialmente reutilizables y de interés para otros investigadores. Sin embargo, las técnicas de recuperación y de archivo empleadas actualmente con este objetivo requieren mucho tiempo y esfuerzo, dificultando la reproducibilidad, la evaluación y la liberación de las colecciones. Combinando ideas de Recuperación de Información con técnicas emergentes utilizadas en el campo del archivo, podemos mejorar este proceso y reducir el esfuerzo requerido para construir colecciones nuevas. Concretamente, el pooling es una técnica ampliamente empleada en el campo de la Recuperación de Información para extraer una muestra

Patricia Martín-Rodilla, Information Retrieval Lab, Centro de Investigación CITIC, Universidade da Coruña
0000-0002-1540-883X
patricia.martin.rodilla@udc.es

David Otero, Information Retrieval Lab, Centro de Investigación CITIC, Universidade da Coruña
0000-0003-1139-0449
david.otero.freijeiro@udc.es

de documentos (publicaciones en el caso de las redes sociales) de un conjunto más grande, de tal manera que sea una muestra completa e imparcial de documentos relevantes sobre un tema determinado. Esto es especialmente útil en el caso de las redes sociales, debido a la enorme cantidad de publicaciones que se crean en estas cada día. Con este método podemos crear una colección de referencia evitando anotar todo el corpus de documentos. En este artículo presentamos la aplicación de esta metodología para la creación de colecciones de referencia a partir del contenido de una red social. Mostramos como, con el método presentado, podemos crear nuevas colecciones reduciendo drásticamente el número de documentos anotados con respecto al tamaño del corpus entero y, por tanto, reducir el esfuerzo y el tiempo requeridos. Finalmente, validamos la metodología presentada mediante la creación de una colección de referencia sobre los ataques sobre el patrimonio que ocurrieron durante 2020 motivados por protestas antirracistas.

**Palabras clave:** redes sociales; racismo; colección; pooling; recuperación de información.

**Abstract:**
Social networks are a valuable source of information for documenting the processes of emergence or evolution in cultural heritage and obtaining a real-time snapshot of a research topic within it. Many researchers in this field use social networks to study these processes, creating collections that constitute digital archives, potentially reusable and of interest to other researchers. However, the retrieval and archiving techniques currently employed for this purpose are time and effort-consuming, making reproducibility, evaluation and release of the collections difficult. Combining ideas from Information Retrieval with emerging techniques used in the archival field can improve this process and reduce the effort required to build new collections. Specifically, pooling is a technique widely used in the Information Retrieval field to extract a sample of documents (publications in the case of social networks) from a larger set, such that it is a complete and unbiased sample of relevant documents on a given topic. This is especially useful in the case of social networks due to the enormous number of posts created on social networks every day. With this method, we can create a reference collection avoiding annotating the entire corpus of documents. In this article, apply this methodology for the creation of

reference collections from the content of a social network. We show how, with the presented method, we can create new collections by drastically reducing the number of annotated documents with respect to the size of the entire corpus and, therefore, reducing the effort and time required. Finally, we validate the presented methodology by creating a reference collection on heritage attacks that occurred in 2020 motivated by anti-racist protests.

**Keywords:** social media; racism; collection; pooling; information retrieval

## 1. INTRODUCCIÓN: LAS REDES SOCIALES COMO FUENTE DE COLECCIONES DE REFERENCIA Y ARCHIVOS SOCIAES

Determinados procesos sociales de especial relevancia en humanidades, como los procesos de memoria e identidad, lingüísticos, colaborativos o patrimoniales, tienen su reflejo de diferentes maneras en nuestra sociedad, como, por ejemplo, en las redes sociales. En estas, los usuarios expresan sus opiniones, debaten, comparten contenido e incluso planifican acciones, como manifestaciones o procesos participativos.

El estudio de estos procesos sociales, en este caso, relacionados con el patrimonio cultural, se ha centrado en el análisis del contenido generado en las redes sociales. En estas, los usuarios crean contenido en diversos formatos (publicaciones en formato textual, pero también todo tipo de contenido multimedia como fotos, vídeos u otro tipo de contenido). En estos contextos, numerosos investigadores estudian este contenido generado en redes sociales desde una perspectiva archivística (con sus implicaciones metodológicas, conceptuales y disciplinares), siendo posible encontrar numerosos ejemplos de archivos nativos digitales desde redes sociales (Pybus, 2013; Pybus, 2015; Acker & Brubaker, 2014) que buscan recuperar, preservar y poner a disposición de la comunidad partes de esa interacción en redes sociales, lo cual es interesante para otros investigadores para su posterior estudio.

Este nuevo enfoque, denominado en ocasiones "archivo social" (Ruiz Gómez & Vallès, 2020), permite "archivar las redes sociales" (Vallès A. M., 2020), creando colecciones de referencia para el estudio de diversas temáticas y su reflejo e impacto en redes sociales (Gehl, 2009; Perpinyà-Morera & Cid-Leal, 2020; Zubia Muxika, 2018). Un ejemplo reciente de esto son los archivos sociales desde Twitter creados para estudiar aspectos de violencia de género (Vallès & Soler Jiménez, 2016; Ruiz Gómez & Vallès, 2020) o en contextos de enfermedad mental (Parapar, Martin-Rodilla, Losada, & Crestani, 2021).

Para esta tarea, existen protocolos que permiten la búsqueda, recuperación y archivo de páginas webs (Schefbeck, Spiliotopoulos, & Risse, 2012) o correspondencia electrónica desde un punto de vista archivístico. Sin embargo, en el caso de las redes sociales, los procesos de recuperación, búsqueda e indexación de información en dominios humanísticos siguen realizándose en su mayor parte de forma manual o semi-manual (Schefbeck, Spiliotopoulos, & Risse, 2012; Mordell, 2019), o bien usando técnicas computacionales no adaptadas a la naturaleza de la información que recuperan ni a su dominio (Winters & Prescott, 2019). Esto convierte a las tareas de recuperación en tareas muy costosas en tiempo y recursos. Además, la mayor parte de los protocolos de recuperación aplicados para crear colecciones desde redes sociales en el ámbito humanístico no abordan desde su creación aspectos de reproductibilidad (impidiendo replicar el proceso a otros investigadores) ni de privacidad y seguridad de la información, como los procesos de anonimización (McNealy, 2011; Obodoruku, 2016).

La Recuperación de Información (RI), como disciplina computacional, ha abordado la creación de colecciones de referencia desde sus inicios, desarrollando numerosas técnicas que permiten realizar estas tareas de forma más eficiente, y también elaborando metodologías para evaluar la calidad de las colecciones resultantes. Combinando técnicas de RI con los protocolos emergentes de archivos sociales, algunos de los problemas identificados

anteriormente pueden mejorarse sustantivamente. En concreto, para evaluar sistemas de recuperación y sus colecciones resultantes, así como para evitar que un asesor humano tenga que revisar todos los documentos de la colección creada (Losada, Parapar, & Barreiro, 2017) surge la técnica denominada *pooling,* como método para extraer una muestra de un conjunto de documentos (*posts* o publicaciones en el caso de redes sociales), obteniendo el conjunto más completo e imparcial de documentos relevantes sobre un tema determinado. Esto es especialmente útil en las redes sociales, debido a la gran cantidad de *posts* existentes y temas tratados por el mismo usuario o en el mismo hilo o publicación.

Este artículo ilustra la aplicación de *pooling*, una metodología ampliamente usada en el campo de la Recuperación de Información, para la construcción de nuevas colecciones de referencia desde redes sociales. Con esta metodología logramos construir nuevas colecciones mientras se reduce el esfuerzo necesario en anotar los documentos de esta. Como caso real para una investigación en curso, se creó una colección de referencia desde la red social Reddit para el estudio del impacto en la comunidad hispanohablante en redes sociales del movimiento *Black Lives Matter* y las revueltas a raíz del asesinato de *George Floyd* en el año 2020.

## 2. ANTECEDENTES

### 2.1. *COMPUTATIONAL ARCHIVAL SCIENCE*, ARCHIVOS DIGITALES Y REDES SOCIALES

Podemos identificar una conexión, como áreas de investigación, entre la Recuperación de Información, disciplina que trata con la representación, almacenamiento, organización y acceso a la información desde una perspectiva computacional (Baeza-Yates & Berthier, 1999), y las prácticas archivísticas, entendidas estas como un área de conocimiento relacionada con el procesamiento, acceso, análisis y almacenamiento de registros archivísticos (Marciano, y otros, 2018).

De la identificación y desarrollo de este tipo de conexión surge el concepto de *Computational Archival Science*, un campo multidisciplinar que integra métodos computacionales con prácticas archivísticas para la creación, procesamiento, preservación, análisis, acceso a y almacenamiento de grandes registros archivísticos (Marciano, y otros, 2018).

Esta conexión interdisciplinar es crucial a la hora de manejar archivos digitales, donde cada registro presenta una naturaleza digital intrínseca (por ejemplo, una publicación en una red social). En este contexto, podemos encontrar archivos nacidos completamente desde una perspectiva digital, entendiendo la web entera como una fuente para estos archivos (Abdallah, y otros, 2017), o también usando las redes sociales como fuente (Acker & Brubaker, 2014).

Además, este tipo de trabajos se relacionan también de alguna manera con la teoría del modelo continuo de registros (Upward, 2005). Esta teoría establece que cualquier registro debe ser gestionado sin fases o protocolos estrictos desde su creación. La manera de ver esta gestión de registros como un flujo continuo es beneficiosa cuando trabajamos con información extraída de las redes sociales. Por tanto, recuperar datos de las redes sociales y crear colecciones de referencia sobre estas nos permite trabajar con esta teoría

Sin embargo, todos estos trabajos presentan algunos problemas en cuanto a las técnicas de recuperación que se emplean a la hora de construirlos. Entre ellos, hemos identificado los siguientes: (i) algunos archivos son creados utilizando técnicas manuales o semi-manuales (Schefbeck, Spiliotopoulos, & Risse, 2012), incluyendo una recuperación automática de los documentos, pero con una segunda fase manual en la que es necesario que los expertos anoten la colección entera, complicando la reproducibilidad, la evaluación y la liberación de los resultados; (ii) algunos se recuperan utilizando tecnologías que no están adaptadas al dominio en cuestión (Mordell, 2019; Winters & Prescott, 2019); (iii) la mayoría de ellos no tienen en cuenta cuestiones de privacidad y de seguridad (McNealy, 2011; Obodoruku, 2016).

En este trabajo identificamos una serie de problemas en trabajos en el campo de la archivística a la hora de crear nuevas colecciones de referencia y, más concretamente, cuando estas son creadas a partir de contenido recuperado en redes sociales. Por otro lado, la disciplina de RI nos ofrece algunas técnicas especialmente diseñadas para crear colecciones con documentos anotados por asesores expertos. En este trabajo proponemos la aplicación de alguna de estas técnicas para solucionar los problemas expuestos anteriormente.

En la siguiente sección presentamos una breve introducción a algunas técnicas empleadas en el campo de la RI y nuestra propuesta para construir nuevas colecciones de referencia, solucionando algunos de los problemas identificados.

## 2.2. CREACIÓN DE NUEVAS COLECCIONES CON *POOLING*

La Recuperación de Información es una disciplina tradicionalmente ligada a una estricta evaluación en el desarrollo de nuevos modelos que permitan recuperar información de diversas fuentes y formatos con criterios concretos y técnicas semiautomáticos o automáticas. La evaluación de cualquier modelo creado dentro de ella requiere el uso de lo que se conoce como colecciones de prueba. Estas están formadas, entre otras cosas, por un corpus de documentos y unos juicios de relevancia (anotaciones hechas por expertos humanos) que establecen qué documentos son relevantes con respecto a una consulta dada. Estos últimos son muy costosos de obtener, ya que requiere que los documentos sean revisados por los anotadores. Debido a la importancia del rol que juegan estas colecciones en la evaluación y la necesidad de elaborar cada vez colecciones con un corpus de documentos más grandes (del orden de miles de millones de documentos), muchos investigadores han trabajado en elaborar estrategias para reducir el esfuerzo y los recursos que requiere construir colecciones nuevas. Concretamente, en una metodología que se conoce con el nombre de *pooling*.

El *pooling* es una metodología madura en el campo de la RI que se emplea para seleccionar un subconjunto del corpus original de la colección (en nuestro caso de *posts* desde redes sociales) (Sanderson, 2010). Una vez seleccionado este subconjunto (el *pool)*, se presentan estos documentos a asesores humanos que los evalúan como relevante o no para la consulta enviada. Si el *pooling* está bien diseñado y aplicado, se puede asumir que los documentos seleccionados son una representación imparcial de los documentos del resto del corpus. Así, se pueden anotar documentos de una colección de entre un corpus más grande a la vez que se mantienen los recursos y el tiempo necesarios en unos rangos abordables, ya que no es necesario revisar y anotar todos los documentos de la colección.

El desarrollo de estrategias para seleccionar qué elementos de un conjunto más grande deben ser revisados y anotados por evaluadores expertos es un tema muy importante más allá del campo de la Recuperación de Información. Por ejemplo, esto es un aspecto crucial a la hora de anotar publicaciones en archivos elaborados desde redes sociales, debido a la gran cantidad de publicaciones que se generan en ellas cada día. Por este motivo, creemos que puede ser útil aplicar estrategias del campo de la RI en otros campos donde existe una problemática similar.

En particular, en este artículo proponemos la aplicación de *pooling* para construir nuevas colecciones de referenciade las redes sociales. Mediante el uso de esta técnica, podemos reducir significativamente el número de publicaciones anotadas por los expertos con respecto al tamaño total del corpus. Además, el uso de esta estrategia permite que, una vez los documentos están anotados por los asesores expertos, estos se puedan emplear para recuperar documentos similares a los juzgados como relevantes, lo que asegura la calidad de la colección en cuanto a su adecuación a la temática pertinente.

En la siguiente sección describimos la metodología propuesta, y todos los pasos necesarios para aplicarla para la construcción de una nueva colección de referencia.

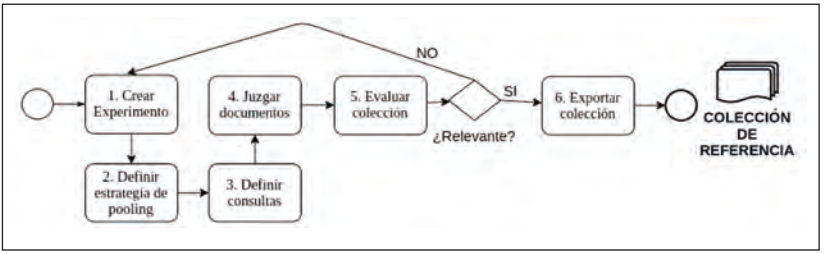



Figura 1: Flujo de trabajo para construir una nueva colección de referencia.

## 3. PROPUESTA Y METODOLOGÍA

En esta sección mostramos la aplicación de la metodología explicada anteriormente para construir colecciones de referencia desde redes sociales. Con esta metodología podemos resolver algunos de los problemas identificados en la Sección 2, reduciendo así mismo los recursos necesarios para anotar las publicaciones recuperadas desde redes sociales.

En este caso hemos diseñado una adaptación de la metodología que se emplea tradicionalmente en el campo de RI para construir nuevas colecciones. Una explicación detallada de esta se puede encontrar aquí (Sanderson, 2010). Esta incluye, normalmente, la organización de una tarea compartida *(shared task)* en la que participan equipos e investigadores de todo el mundo. En este caso, proponemos una adaptación de esta metodología que automatiza algunas de las fases de la original.

En concreto, la metodología que proponemos en este trabajo se detalla en la Figura 1.

Como se puede observar en esta figura, el resultado final es una colección de referencia que incluye las publicaciones seleccionadas como relevantes, así como el resto de las publicaciones que se obtuvieron.

En este contexto, un experimento es cada una de las instancias en las que desarrollamos el flujo de trabajo de la Figura 1, y abarca la selección de la estrategia de *pooling* (fase 2), la definición de las necesidades de información y consultas (fase 3),

el juicio de los documentos (fase 4) y, finalmente, la creación de la colección final (fases 5 y 6).

Dentro de cada experimento, el primer paso, la fase 2, consiste en la selección de la estrategia de *pooling*. Ésta dicta el orden en el que los documentos serán presentados a los asesores expertos según diferentes criterios. La selección de esta estrategia es un paso esencial en la construcción de la colección final. En este caso hemos empleado una estrategia basada en Bandidos Bayesianos, MaxMean. Esta se ha demostrado como una estrategia muy eficaz para construir nuevas colecciones en el campo de la Recuperación de Información (Losada, Parapar, & Barreiro, 2017).

A continuación, en la fase 3, el siguiente paso es definir las consultas, que se emplearán para la búsqueda de los documentos (*posts* en redes sociales). Estas consultas están normalmente formadas por varios términos clave que representan las necesidades de información de la colección, como se verá más adelante en el caso de estudio. Una vez seleccionada la estrategia de *pooling* y definidas las consultas que se utilizarán para recuperar documentos, el siguiente paso es establecer los juicios de los documentos seleccionados (fase 4). Estos juicios pueden ser emitidos por una persona, o un grupo de personas (expertos en el dominio y temática en el que se encuadra la colección que se construirá finalmente). Los juicios consisten en que cada asesor experto determina si un documento es relevante o no para la necesidad de información que se definió anteriormente. Una vez los juicios se han completado, el último paso consiste en revisar si el contenido de la colección se adecúa a las necesidades de información que se establecieron al principio (fase 5). Finalmente, los documentos que forman la colección en conjunto con los juicios establecidos por los expertos, la colección ya puede ser liberada y exportada para su uso en el futuro (fase 6).

## 4. CASO DE ESTUDIO Y RESULTADOS

Durante el último año, nuestro grupo de investigación ha aplicado la metodología propuesta (ver Figura 1) para la construcción de colecciones de referencia en diversos ámbitos y proyectos de investigación en humanidades digitales. Entre ellos, una línea de interés era el seguimiento del impacto en redes sociales de las tensiones y protestas provocadas por la muerte de *George Floyd* el 25 de mayo de 2020.

Durante el año 2020 tuvieron lugar protestas alrededor del mundo, protestas que comenzaron como respuesta a la muerte de *George Floyd* en los Estados Unidos. Este evento dio lugar a numerosas protestas y altercados, en los que se atacaron y vandalizaron diversos elementos patrimoniales conmemorativos, como placas y estatuas, con la motivación de que estos representaban figuras e ideas racistas y genocidas.

Esta conexión entre el patrimonio y el racismo ya ha sido ampliamente estudiada en el pasado, pudiendo encontrar trabajos de múltiples disciplinas, como la filosofía (Arday, 2020; Davidson-Harden, 2020; Samayeen, AdrianWong, & McCarthy, 2020) o la antropología y la historia (Gibson & Reich, 2017; Ince, Rojas, & Davis, 2017).

Por esto, estos ataques ocurridos durante el año 2020 reflejan procesos sociales y patrimoniales de una naturaleza diversa que, en nuestra opinión, requieren de la existencia de colecciones de referencia que puedan emplearse posteriormente para un estudio más profundo desde diferentes perspectivas, campos de estudio y metodologías.

En concreto, los acontecimientos ligados a la muerte de *George Floyd* en el año 2020 provocaron la aparición de movimientos paralelos en redes sociales en los que los usuarios comentaban y discutían sus opiniones sobre estas problemáticas. La relevancia de la captura en tiempo y forma de todas estas opiniones y movimientos organizados desde redes sociales sobre la temática nos llevó a aplicar la metodología presentada arriba,

y así crear una colección de referencia que permita estudiar estos procesos patrimoniales y culturales.

Por otro lado, este tema es probablemente uno de los temas relacionados con entidades patrimoniales que más atención recibe de la opinión pública, los medios de comunicación e los investigadores de diferentes disciplinas (Arday, 2020; Davidson-Harden, 2020; Samayeen, AdrianWong, & McCarthy, 2020). Se ha demostrado también el importante papel que juegan las redes sociales a la hora de influenciar la opinión pública (los procesos electorales en Estados Unidos, por ejemplo) como un medio del llamamiento a la acción y la organización de plataformas y colectivos, todos ellos relacionados con los ataques a entidades patrimoniales. Por tanto, en nuestro grupo de investigación creemos que la preservación de toda esta información sobre estos procesos patrimoniales es una perfecta motivación para construir una nueva colección de referencia y así validar la metodología propuesta.

Para ello, y usando la red social *Reddit* (Reddit, 2020) como fuente, se aplicó en el pasado la metodología descrita para esta misma temática, con la principal diferencia de que los textos estaban escritos en inglés, creando así una colección que está actualmente disponible (Otero, Martín-Rodilla, & Parapar, 2020; Otero, Martín-Rodilla, & Parapar, 2021). La elección de *Reddit* viene dada por su alto número de usuarios en Estados Unidos (foco de las revueltas), su especialización por *subreddits* temáticos (haciendo más efectivo el *pooling*) y por su organización tipo blog, que permite a sus usuarios expresar opiniones largas sin limitación de caracteres. Es importante mencionar que la metodología no depende de la red social que se utilice como fuente de datos, puede adaptarse fácilmente para ser aplicada a otras fuentes, como por ejemplo otras redes sociales como Twitter. Además, también podría adaptarse para recuperar textos de varias fuentes a la vez, aunando publicaciones de diferente naturaleza (en Twitter, por ejemplo, existe una limitación que no permite escribir publicaciones largas). En este caso, para la colección en

español, hemos empleado *Reddit* también como fuente, por los mismos motivos expuestos anteriormente.

En esta sección presentamos el caso de estudio y la colección resultante para español, explicando paso a paso la metodología aplicada para su creación.

### 4.1. DEFINICIÓN DE TÉRMINOS DE BÚSQUEDA

El proceso de diseño de consultas y selección terminológica constituye un proceso clave en la calidad posterior de la colección, y requiere de la colaboración interdisciplinar para su éxito. Algunas de sus implicaciones a nivel multilingüe y de calidad de la colección han sido discutidas previamente (Otero, Parapar, & Barreiro, 2019).

Para encontrar publicaciones relevantes con respecto a nuestra necesidad de información, queríamos encontrar publicaciones que hiciesen referencia a la muerte de *George Floyd* o, por otro lado, a una entidad patrimonial y a la opinión de los usuarios de la red social sobre la misma. Por estos motivos, elaboramos una lista de términos que incluyese palabras relacionadas con entidades patrimoniales y palabras relacionadas con la opinión de las personas en cuanto a las protestas antirracistas y los conflictos que se derivaron de ellas, así como términos que se relacionasen con la muerte de *George Floyd*.

La Tabla 1 muestra los términos empleados para la creación de consultas que después se emplearon para obtener las publicaciones de *Reddit*. La selección de términos se realizó como un proceso en paralelo a la colección en inglés (Otero, Martín-Rodilla, & Parapar, 2020; Otero, Martín-Rodilla, & Parapar, 2021), así como guiado por diversas pruebas previas de cobertura (seleccionando, ante varias alternativas y sinónimos, el término que arrojaba mayores resultados en la red social). La facilidad para automatizar la posterior recuperación de información a partir de las consultas diseñadas permite el rediseño continuo de las consultas a modo de ensayo-error, así como la ampliación de esta lista de términos y la replicación de este proceso para ajustar cada colección a las necesidades de investigación, idiomáticas y regionales en cada caso.

Cabe destacar también en este punto la posibilidad de expandir en sucesivas consultas los términos asociados a esta lista mostrada en la Tabla 1, para así obtener consultas con una mayor amplitud léxica, una idea que planeamos explotar en el futuro.

| | |
|---|---|
| racismo | vandalizado |
| protesta | *vandalismo* |
| racista | *esclavo* |
| George Floyd | *esclavitud* |
| blm | *policía* |
| Black lives matter | *violencia* |
| estatua | *monumento* |
| busto | *revuelta* |
| placa | |

**Tabla 1: Términos en español seleccionados en la fase de definición de consultas.**

## 4.2. JUICIOS DE RELEVANCIA SOBRE LAS PUBLICACIONES

Las consultas definidas dieron como resultado 219 (el tamaño del *pool*) publicaciones diferentes para juzgar. Una vez recuperadas estas publicaciones, el siguiente paso es realizar las anotaciones de relevancia o no relevancia con respecto a nuestra necesidad de información inicial, en este caso, publicaciones relacionadas con las revueltas ocasionadas por la muerte de *George Floyd*.

Los juicios fueron hechos por dos evaluadores a partes iguales, a los que se les presentó una publicación a la vez. Estos evaluadores, aunque de perfiles disciplinares diferentes, ambos están relacionados con el tratamiento de información sobre patrimonio cultural. Esto se hizo así para garantizar el equilibrio a la hora de emitir los juicios. Por otro lado, y para garantizar aún más estandarización en los juicios, a los evaluadores se les facilitó, previamente al comienzo del proceso, un documento en el que se explicitaban los criterios para establecer la relevancia de las publicaciones. Esto se hizo para evitar introducir sesgos

personales a la hora de emitir los juicios y para darles a los asesores una guía simple para anotar las publicaciones. Concretamente, solo se anotaron como publicaciones relevantes los *posts* en español que contenían referencias a la muerte de *George Floyd* y sus consecuencias (revueltas, manifestaciones, detenciones y ataques a entidades patrimoniales, etc.). Este proceso duró en torno a una hora por evaluador. Una vez emitidos los juicios, siguiendo la metodología, recuperamos todos los hilos a los que pertenecían los posts que se anotaron como relevantes.

El hecho de que la metodología permita reproducir el experto y construir la misma colección permite que en el futuro se puedan incorporar más asesores expertos al proceso y así poder realizar un estudio sobre el nivel de acuerdo en los juicios emitidos. Creemos que este tipo de estudio en relación con los procesos patrimoniales que se relacionan con el contenido obtenido de redes sociales puede ser interesante desde un punto de vista de investigación.

### 4.3. COLECCIÓN RESULTANTE

Después del proceso de definición de consultas y juicio experto, a partir de los *posts* juzgados como positivos se recuperaron un total de 77 hilos diferentes, con un total de **19 303** *posts* en español en la colección de referencia. Esto nos da una media de 250 *posts* incorporados a la colección por hilo, recuperando información de *subreddits* de toda Latinoamérica, España y comunidades hispanas en otros países, con al menos 15 países diferentes identificados.

Esta gran cantidad de hilos diferentes asegura que la metodología permite recuperar e incorporar a la colección publicaciones no solo de hilos explícitos sobre el movimiento *Black Lives Matter*, sino también de muchos hilos con diferentes temas centrales que hagan referencia a estos eventos en algún momento. Esto es importante para asegurar la cobertura del tema dentro de la red social, no solo recuperando aquellos hilos de conversación explícitos (que pueden reflejar puntos de vista que están excesivamente polarizados o dirigidos por asociaciones o

personas directamente implicadas en los conflictos) sino también llegando a hilos con temáticas diferentes al tema principal, tratando las protestas en algún momento. Estas conversaciones pueden ser menos polarizadas e incluyen ciudadanos con perfiles y responsabilidades más diversos.

Así mismo, los *posts* integrantes fueron escritos entre el 25 de mayo de 2020 (día del fallecimiento de *George Floyd*) y el 10 de abril del 2021. Nótese que solo fue necesario juzgar los 219 *posts* del *pool* inicial durante la fase de juicio (evitando la anotación completa de la colección) para recuperar casi 1 año de actividad en Reddit de la comunidad hispanohablante acerca de las protestas antirracismo. La colección está disponible para su descarga en este vínculo[1].

## 5. DISCUSIÓN Y CONCLUSIONES

En este trabajo hemos propuesto una metodología que permite la construcción de colecciones de referencia desde redes sociales como archivos sociales relevantes en humanidades. Con esta hemos logrado dejar atrás algunos de los problemas más comunes a la hora de construir colecciones de referencia. Además, hemos validado nuestra metodología mediante la construcción de una nueva colección sobre un tema estrechamente relacionado con el patrimonio cultural que ha atraído mucha atención desde la sociedad en los últimos meses. Presentamos aquí una colección en español desde Reddit que recupera, monitoriza, documenta, conserva y permite el estudio evolutivo en esta red social del fenómeno de protesta antirracismo desencadenado por la muerte de *George Floyd* en mayo de 2020.

Como análisis crítico, cabe discutir varios aspectos de la propuesta (y de la propia colección). En primer lugar, es necesario un estudio en profundidad del nivel de acuerdo entre anotadores

1. https://www.dc.fi.udc.es/~david/hdh2021

en la fase de *pooling* para garantizar cierta ausencia de sesgos en el diseño de consultas y en la fase de juicio posterior. Este aspecto está ampliamente desarrollado en la literatura (Cohen, 1960; Fleiss, Levin, & Paik, 1981) y también desde el punto de vista de RI como disciplina (Hripcsak & Rothschild, 2005; Artstein R., 2017; Artstein & Poesio, 2008), por lo que planeamos ampliar el número de anotadores e incorporarlo en el futuro.

En segundo lugar, somos conscientes que el estudio de un fenómeno tan complejo como el racismo y/o el movimiento *Black Lives Matter* van mucho más allá de esta colección que presentamos. También de la necesidad de ampliar a más redes sociales el estudio, y de discutir la representatividad social en las mismas. Sin embargo, creemos que, debido a la generalización de las redes sociales como una fuente adicional en dichos estudios, esta colección (y la metodología que subyace), permite complementar estudios del fenómeno desde diferentes perspectivas (sociológicas, antropológicos, históricas, patrimoniales, de medios de comunicación etc.) con otras fuentes y/o técnicas que los investigadores usen, así como posibilita el uso de técnicas computacionales sobre ella (procesado de lenguaje natural, *text mining* o *Social Network* (Kelly & Kittle Autry, 2011), entre otros). Un estudio comparativo tanto a nivel metodológico como entre las colecciones resultantes en inglés (Otero, Martín-Rodilla, & Parapar, 2020) y en español (el que aquí se presenta) arrojará en el futuro datos sobre las implicaciones de la lengua empleada para la creación de estas colecciones desde redes sociales.

Destacamos así mismo que la metodología presentada permite replicar el proceso de creación de la colección con las mismas consultas y fuentes en *Reddit*, pudiendo, por ejemplo, ampliar la colección en rangos de fechas para estudiar el fenómeno de manera diacrónica. En el caso específico de *Black Lives Matter*, ciertos eventos como la celebración del juicio por la muerte de *George Floyd* en abril de 2021 pueden hacer resurgir el movimiento y la actividad en redes sociales, por lo que su ampliación puede ser interesante desde el punto de vista investigador.

Es importante también reseñar que, aunque la técnica de *pooling* y la metodología empleada resuelve algunos de los problemas detectados en cuanto a la aplicación de técnicas de recuperación de información en este tipo de trabajos, no se pretende con ello automatizar la tarea de crear colecciones de referencia (como archivos nativos digitales) a partir de las redes sociales. Somos conscientes de que las mejoras en el enfoque tecnológico que ofrece deben ser aplicadas desde una perspectiva transdisciplinar, donde sea la colaboración de expertos en el dominio de las colecciones, en archivística y documentación y en recuperación de información la que permita la mejora de este tipo de colecciones.

## 6. TRABAJO FUTURO

Planeamos extender el trabajo presentado en este artículo a través de varias líneas de investigación. La primera idea que exploraremos será ampliar la colección presentada con la información de perfil de los usuarios cuyos posts forman parte de la colección, lo que permite una investigación más profunda acerca de perfiles demográficos y/o psico-sociales de usuarios dentro de la colección, pero teniendo en cuenta desde el diseño inicial de la colección protocolos específicos de anonimización y seguridad. Otra línea que planeamos estudiar es realizar una comparación entre la colección en inglés y la colección en español, presentada en este artículo. Sobre estas dos colecciones podemos estudiar sus similitudes y diferencias en cuanto a los temas que se tratan en las publicaciones, la edad, el género y las intenciones a la hora de publicar textos a favor o en contra de las protestas antirracistas y los ataques a entidades patrimoniales, entre otros muchos aspectos que se pueden comparar. Otro aspecto interesante es investigar el nivel de acuerdo o desacuerdo en las anotaciones de los expertos. En este caso, podrían expandirse las anotaciones ya existentes con anotaciones de otros expertos, y analizar el nivel de

acuerdo de estas en función del contenido de las publicaciones. Por otro lado, podríamos explorar formas de mejorar la metodología empleada para construir la colección. Una manera de mejorarla podría ser explorar el uso de técnicas más avanzadas de RI para obtener mejores publicaciones de las redes sociales. Otra línea que planeamos investigar es el uso de bases de datos léxicas, tesauros u otros recursos lingüísticos que permitan expandir y mejorar las consultas empleadas para recuperar las publicaciones. Por ejemplo, se podrían expandir en sucesivas consultas los términos asociados a esta lista mostrada en la Tabla 1, incluyendo, para cada término, aquellos con los que mantiene relaciones de hiperonimia, hiponimia o meronimia, entre otras. Este enfoque permitirá consultas con mayor amplitud léxica, usando para ello bases de datos léxicas en el idioma deseado (en este caso para español) a modo de estructuras synsets, como en el caso de Wordnet[2]. Otro enfoque para la posible expansión de las consultas en la creación y uso de tesauros especializados en la temática deseada, aunando estudios referenciados anteriormente sobre la relación entre patrimonio y racismo y su léxico especifico. Sin embargo, el uso de estos métodos automáticos o semiautomáticos para la expansión de las consultas no ha demostrado para este caso la recuperación de publicaciones en redes sociales para la colección más relevantes sobre la temática tratada, motivo por el cual dejamos esta aproximación como aspecto a trabajar en el futuro.

2. George A. Miller (1995). WordNet: A Lexical Database for English. Communications of the ACM Vol. 38, No. 11: 39-41 and Christiane Fellbaum (1998, ed.) *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press. WordNet: https://wordnet.princeton.edu/