



**UNIVERSIDAD
DE GRANADA**

PROGRAMA DE DOCTORADO EN HISTORIA Y ARTES

**Humanidades Digitales en Internet.
Diseño e implementación de
Sistemas de Información Web
para la investigación en
Historia, Arte y Arqueología**

Maurizio Toscano

DIRECTORES

**Enrique Herrera Viedma
Lidia Bocanegra Barbecho**

TESIS DOCTORAL
Granada - 2021

Editor: Universidad de Granada. Tesis Doctorales

Autor: Maurizio Toscano

ISBN: 978-84-1306-980-7

URI: <http://hdl.handle.net/10481/70129>

We Shape Our Tools, and Thereafter Our Tools Shape Us.
John M. Culkin (1967)

ÍNDICE

| | |
|--|------------|
| AGRADECIMIENTOS..... | V |
| RESUMEN..... | VII |
| ABSTRACT | X |
| NORMATIVA..... | XII |
| INDICIOS DE CALIDAD DE LAS PUBLICACIONES | XVI |
| I. INTRODUCCION | 1 |
| <i>I.1. OBJETO DE ESTUDIO.....</i> | <i>3</i> |
| <i>I.2. JUSTIFICACIÓN DE LA INVESTIGACIÓN</i> | <i>6</i> |
| II. OBJETIVOS..... | 8 |
| III. METODOLOGIA | 11 |
| <i>III.1. SOLUCIONES SOFTWARE PARA LAS HUMANIDADES DIGITALES Y EL PATRIMONIO CULTURAL: ESTADO DE LA CUESTIÓN</i> | <i>11</i> |
| <i>III.2. LAS FASES DE TRABAJO.....</i> | <i>30</i> |
| IV. RESULTADOS | 44 |
| <i>IV.1. El CONTEXTO DE LAS HUMANIDADES DIGITALES EN ESPAÑA (1993-2019)</i> | <i>44</i> |
| <i>IV.2. DISEÑO E IMPLEMENTACIÓN DE SISTEMAS DE INFORMACIÓN WEB PERSONALIZADOS</i> | <i>81</i> |
| IV.2.1. Un sistema de información web para la optimización del flujo de trabajo | 82 |
| IV.2.2. Un sistema de información web como atlas temático digital | 94 |
| <i>IV.3. LA WEB COMO PLATAFORMA DE PRESENTACIÓN E INTERCAMBIO DE FUENTES PRIMARIAS.....</i> | <i>109</i> |
| <i>IV.4. LOS SISTEMAS DE INFORMACIÓN WEB PARA EL CROWDSOURCING Y LA CIENCIA CIUDADANA.....</i> | <i>144</i> |
| IV.4.1. Desde datos locales a plataformas en línea | 146 |

| | |
|---|-----|
| IV.4.2. Crowdsourcing en humanidades..... | 164 |
| IV.4.3. Análisis de estrategias de involucración pública..... | 192 |
| V. CONCLUSIONES | 198 |
| VI. BIBLIOGRAFIA | 206 |

AGRADECIMIENTOS

Desde el inicio de esta investigación, he tenido la suerte de compartir momentos de trabajo con muchas personas, que han colaborado, de forma deliberada o involuntaria, a la configuración actual de este texto. Por este motivo me encuentro en deber de expresarles mi gratitud a través de estas pocas líneas.

El primer y más importante agradecimiento va a Lidia, mi mujer, a quien he podido, además, tener como codirectora, por su insustituible e incondicional apoyo. Fue ella que me instó a realizar esta tesis y que siguió de cerca todo su desarrollo; nunca parca de consejos o avara de tiempo. También le debo un *grazie* a Clio, y no solo por estar, que ya sería motivo suficiente, si no por la paciencia demostrada a diario en esperar a que acabara con este, y con muchos otros trabajos, para poderle dedicar tiempo de juego.

Un agradecimiento especial va a Enrique y a Antonio, respectivamente mi director y mi actual tutor. El primero, por acoger desde el principio con interés y simpatía este proyecto, sin dudar nunca del buen éxito del mismo, y por su disponibilidad en asesorarme y animarme frente a los titubeos. El segundo, por su cercanía y disponibilidad en resolver mis dudas y atender mis peticiones, así como por empujarme a realizar una actividad formativa sobre una de las temáticas de la tesis, frente a mi natural recelo hacia la enseñanza.

A los coautores de algunas de las publicaciones que forman parte del capítulo de resultados, quienes se han visto embarcados en aparecer en estas páginas. A Lidia, Lara, Mario, Paul, Salvador, Elena, Aroa e Giuseppe va mi agradecimiento por las fructíferas y enriquecedoras discusiones que han rodeado esos artículos.

A los promotores de los sistemas de información desarrollados en esos años, ya sea los que han acabado como casos de estudio en la tesis, que los muchos que no han encontrado cabida, pero que han representado igualmente una experiencia estimulante y han marcado, cada uno de ellos, un nuevo reto en esta profesión. A John, José María, Margarita, Rafael, Yolanda, M^a Elena, Mario, Paul, Alberto, Aurora, Andrea, Fernando y Mar, va mi agradecimiento por haber depositado su confianza, y sus valiosos datos de investigación, en mi persona y en mis competencias, así como por los largos e intensos diálogos que han acompañado el diseño de esas plataformas.

A los numerosos usuarios que, al utilizar en esos años alguno de los sistemas que he tenido el deber y el placer de acometer, han enviado sus opiniones, ofrecido sus consejos o expresado sus exigencias, colaborado así a mejorarlos y hacerlos más útiles y eficaces.

A las extraordinarias comunidades de desarrolladores y usuarios, en buena medida anónimos, del software utilizado a lo largo de este trabajo, a las cuales he acudido en las fases de implementación técnica de las plataformas para resolver dudas, solucionar

errores y encontrar pistas para implementar nuevas funcionalidades; su ayuda ha sido, en muchos casos, insustituible.

RESUMEN

La presente tesis doctoral, de corte eminentemente metodológico, es fruto de un trabajo de investigación que se genera a partir de una serie de sistemas de información web diseñados y desarrollados por el candidato, en colaboración con investigadores especialistas en su correspondiente dominio disciplinar (Historia, Historia del Arte y Arqueología). Estas plataformas se han utilizado como casos de estudio para valorar el impacto generado, en particular a nivel metodológico y de procedimientos, sobre toda una serie de procesos de investigación, tales como: la obtención y verificación de fuentes y datos; el almacenamiento, la gestión y la recuperación de la información; el flujo de trabajo; la puesta en común y la co-creación de contenidos; la ilustración de resultados.

En el capítulo de introducción se abarca la definición del objeto de estudio: los sistemas de información web, desde la perspectiva de su funcionamiento y aplicación, y a través de las implicaciones metodológicas implícitas en su proceso de planteamiento, modelado, implementación y uso. Asimismo, se definen los límites del contexto de análisis, que son los proyectos y líneas de investigación en Historia, Historia del Arte y Arqueología; se justifica el interés del estudio realizado, en función del rol central que dichos sistemas ocupan en el proceso de digitalización de la investigación, así como por el interés en los enfoques participativos que esos sistemas habilitan, y por la alta demanda, en la última década, por este tipo de artefactos. Finalmente, se describe el encuadre disciplinar de referencia, que se enmarca en los principios de las Humanidades Digitales y de la Historia Pública Digital.

El capítulo dedicado a la metodología está dividido en dos partes. En la primera presentamos una amplia revisión del estado actual en cuanto a soluciones informáticas para la creación de sistemas de información web en humanidades y patrimonio cultural, analizando comparativamente software perteneciente a las siguientes categorías: sistemas de gestión del contenido; sistemas de gestión de activos digitales; entornos virtuales de investigación; plataformas para el *crowdsourcing* y la Ciencia ciudadana. En el segundo apartado, describimos en detalle las diferentes fases de la metodología de trabajo empleada en la tesis: perfilada a partir del ciclo clásico de vida, uso y mantenimiento de un sistema de información, pero ampliada en función de los objetivos específicos planteados.

Los resultados conseguidos se han plasmado en seis publicaciones científicas, cuatro artículos de revistas y dos capítulos de libros, que hemos organizado temáticamente en cuatro bloques.

El análisis histórico del proceso de consolidación de las Humanidades Digitales en España desde los principios de los años noventa hasta la actualidad nos ha permitido: 1) avanzar una primera periodización en tres fases en la evolución de la disciplina; 2) ofrecer una imagen actualizada del panorama actual en términos de investigadores, instituciones implicadas y organismos de financiación; 3) agrupar una amplia

variedad de temas de investigación entorno a cinco grandes grupos disciplinares (Historia, Filología, Comunicación, Informática y Documentación); y 4) confirmar, métricamente, la mayoría de las evidencias detectadas en estudios precedentes, así como establecer nuevas métricas y valores que constituyen una base de referencia para monitorear la evolución de la disciplina en el futuro y facilitar comparaciones con contextos internacionales.

Otro bloque se refiere al análisis del proceso de diseño, definición de funcionalidades y uso de sistemas de información web, personalizados a medida sobre las problemáticas específicas de una amplia variedad de proyectos de investigación. Este estudio ha evidenciado los beneficios en términos de cantidad y variedad de fuentes y datos interrelacionables, sobre la capacidad inferencial y el potencial de recuperación de la información, la ampliación de los sujetos implicados en el proceso de investigación, así como los cambios significativos producidos en el flujo de trabajo en términos de recogida, verificación y publicación en línea de la información.

En el tercer bloque exploramos el potencial de la Web ligado a la presentación dinámica y al intercambio de fuentes primarias de investigación, de cara a la ilustración de artículos científicos. Los resultados son básicamente tres, todos ellos metodológicos y desarrollados a partir de una problemática procedente de la Geoarqueología: 1) se ha ampliado el espectro de las técnicas consolidadas para la digitalización de láminas delgadas, con la introducción de la *manual Whole Slide Imaging* (mWSI), procedente desde la Patología humana; 2) se demuestra el potencial informativo de esas digitalizaciones como fuentes primarias en abierto, que habilitan nuevos procesos como la contextualización de las microfotografías, la verificación en remoto de la lectura estratigráfica y las descripciones colaborativas; 3) se define y propone un protocolo concreto para la ilustración dinámica de resultados de investigación, basados en imágenes de alta resolución.

En el cuarto y último bloque de resultados ponemos en relieve el rol de los sistemas de información web como herramientas habilitadoras de nuevos procesos de Ciencia ciudadana. En primer lugar, se analizan los beneficios en la creación de plataformas web para el enriquecimiento de conjuntos de datos existentes, describiéndose los aspectos a tener en cuenta para optimizar el proceso. Sucesivamente, a través de ejemplos de investigación en Historia Pública, se analiza el impacto de tres innovadoras metodologías de trabajo: el *crowdsourcing*, la co-creación y el trabajo en red. Este último bloque incluye, además de dos publicaciones, la definición de un modelo lógico para la clasificación y el análisis de estrategias de involucración pública, orientadas a fomentar la participación en proyectos colaborativos, cuya publicación se encuentra actualmente en curso.

En las conclusiones, volvemos a reflexionar acerca de cómo los sistemas de información web influyen sobre dos de las principales tendencias reconocibles en el campo disciplinar de las Humanidades Digitales: el enfoque instrumental y el metodológico. De cómo esos sistemas capacitan toda una serie de procesos que afectan

significativamente la generación de nuevo conocimiento. De cómo se está consolidando, como alternativa a los propios sistemas de información web personalizados, el uso de los entornos virtuales de investigación, caracterizados por un elevado grado de customización en el diseño del modelo de datos y en la definición de los vocabularios/taxonomías, junto a una escasa, o nula, posibilidad de adaptación de las funcionalidades. Finalmente, se proporcionan, a partir de una serie de carencias detectadas, unas propuestas concretas: 1) para la mejora de la formación interdisciplinar en Humanidades Digitales; 2) para la creación de unidades técnicas de asesoramiento y apoyo a la normalización del uso de recursos y herramientas digitales en todas las ramas de las humanidades; 3) para el reconocimiento académico y curricular de la actividad ligada al desarrollo de artefactos tecnológicos para la investigación.

ABSTRACT

This doctoral thesis, which is essentially methodological in nature, is the result of a research project originated from a series of web-based information systems designed and developed by the candidate, in collaboration with researchers specialised in their respective disciplines (History, History of Art and Archaeology). These platforms have been used as case studies to assess the methodological and procedural impact on a range of research processes, such as: gathering and verification of sources and data; storage, management and retrieval of information; workflow; sharing and co-creation of content; and illustration of results.

The introduction covers the definition of the object of study: web information systems, from the perspective of their operation and application, and through the methodological consequences implied in the process of their planning, modelling, implementation and use. It also defines the limits of the context of analysis, which are projects and research lines in History, Art History and Archaeology. The interest of this study is justified by the central role that these systems play in the process of digitisation of research, by the interest in participatory approaches that these systems enable, and by the high demand, in the last decade, for this type of artifacts. Finally, it describes the disciplinary framework of reference, which is based on the principles of Digital Humanities and Digital Public History.

The methodology chapter is divided into two parts. In the first, we present a broad review of the current state of the art in terms of software solutions for the creation of web-based information systems in the Humanities and cultural heritage, comparatively analysing software belonging to the following categories: content management systems; digital asset management systems; virtual research environments; platforms for crowdsourcing and Citizen science. In the second section, we describe in detail the different phases of our working methodology, outlined on the basis of the classic life, use and maintenance cycle of an information system, but extended according to the specific objectives of the thesis.

The results achieved have translated into six scientific publications, four journal articles and two book chapters, which can be arranged thematically in four blocks.

The historical analysis of the consolidation process of the Digital Humanities in Spain from the early 1990s to the present day has allowed: 1) to advance a first periodisation in three phases in the evolution of the discipline; 2) to offer an updated picture of the current panorama in terms of researchers, institutions involved and funding bodies; 3) to group a wide variety of research topics around five major disciplinary groups (History, Philology, Communication, Computer Science and Documentation); and 4) to confirm, metrically, most of the evidence detected in previous studies, as well as to establish new metrics and values that constitute a reference base to monitor the evolution of the discipline in the future and facilitate comparisons with international contexts.

Another block refers to the analysis of the process of design, definition of functionalities and use of web information systems, customised on the specific problems of a wide variety of research projects. This study has shown the benefits in terms of the number and variety of interrelated sources and data, on the inferential capacity and information retrieval potential, the broadening of the subjects involved in the research process, as well as the significant changes in the workflow in terms of information collection, verification and online publication.

In the third block, we explore the potential of the Web for the dynamic presentation and exchange of primary research sources for the illustration of scientific articles. The results are basically threefold, fundamentally methodological and developed on the basis of a problem arising from Geoarchaeology: 1) we have broadened the spectrum of established thin-sections digitisation techniques, with the introduction of manual Whole Slide Imaging (mWSI), coming from Human Pathology; 2) we demonstrate the informative potential of these scans as Open Primary Sources, which enable new processes such as microphotographs' contextualisation, remote verification of stratigraphic units or collaborative descriptions; 3) we define and propose a specific protocol for the dynamic illustration of research results, based on high-resolution images.

In the fourth and final block of results, we highlight the role of web information systems as enabling tools for new Citizen Science processes. First of all, the benefits of creating web platforms for the enrichment of existing datasets are analysed, describing the aspects to be considered in order to optimise the process. Subsequently, through three examples of research in Public History, the implications of three new working methodologies are discussed: crowdsourcing, co-creation and networking. This last section includes, in addition to two publications, the definition of a logic model for the classification and analysis of public engagement strategies aimed at encouraging participation in collaborative projects, the publication of which is currently in progress.

In the conclusions, we return again on how web information systems influence two of the main trends recognisable in Digital Humanities: the instrumental and the methodological approach. On how these systems enable a whole series of processes that significantly affect the generation of new knowledge. On how the use of virtual research environments, characterised by a high degree of customisation in the design of the data model and in the definition of vocabularies/taxonomies, combined with little or no possibilities of tailoring the functionalities, is becoming a consolidated alternative to custom web information systems. Finally, based on a series of detected weaknesses, specific proposals are made: 1) for the improvement of interdisciplinary training in Digital Humanities; 2) for the creation of technical units to advise and support in the use of digital resources and tools in all branches of the Humanities; 3) for the academic and curricular recognition of the activity linked to the development of technological tools for research.

NORMATIVA

La presente tesis doctoral ha sido elaborada de acuerdo con la modalidad "agrupación de publicaciones". Se han incorporado, como resultados de la misma, los trabajos previamente publicados en revistas científicas internacionales indexadas en la base de datos del *SCImago Journal & Country Rank (SJR)* de Scopus o en libros de editoriales indexadas en la base de datos del *Scholarly Publishers Indicators in Humanities and Social Science (SPI)*.

La memoria ha sido redactada de acuerdo con lo establecido en el "TEXTO REFUNDIDO DE LAS NORMAS REGULADORAS DE LAS ENSEÑANZAS OFICIALES DE DOCTORADO Y DEL TÍTULO DE DOCTOR POR LA UNIVERSIDAD DE GRANADA" (aprobadas en Consejo de Gobierno de 2 de mayo de 2012 y modificadas en Consejo de Gobierno de 30 de octubre de 2013 y de 25 de febrero de 2020).

Estas normas establecen, en el Capítulo I: La Tesis Doctoral del Título Segundo: La Tesis Doctoral, lo siguiente:

Artículo 18º. La tesis doctoral

1. La tesis doctoral consistirá en un trabajo original de investigación elaborado por el candidato en cualquier campo del conocimiento que se enmarcará en alguna de las líneas de investigación del Programa de Doctorado en el que está matriculado.
2. La tesis debe capacitar al doctorando para el trabajo autónomo en el ámbito de la I+D+i.
3. La tesis debe contar, al menos, con los siguientes contenidos: título, resumen, introducción, objetivos, metodología, resultados, conclusiones y bibliografía.
4. Una tesis doctoral puede también consistir en el reagrupamiento en una memoria de trabajos de investigación publicados por el doctorando en medios científicos relevantes en su ámbito de conocimiento.

Los artículos que configuren la tesis doctoral deberán estar publicados o aceptados con fecha posterior a la obtención del título de grado y del master, no podrán haber sido utilizados en ninguna tesis anterior.

Si la publicación ha sido realizada por varios autores, además del doctorando, se debe adjuntar la declaración de los restantes autores de no haber presentado dicha publicación en otra tesis doctoral o la renuncia a hacerlo.

En esta modalidad, la tesis debe tener además de los apartados mencionados en el punto anterior, los artículos que la componen, bien integrados como capítulos de la tesis o bien como un Anexo.

El Comité de Dirección de la Escuela de Doctorado correspondiente podrá establecer el número mínimo de artículos necesarios para presentar una tesis en esta modalidad y las condiciones adicionales sobre la calidad de los trabajos (nueva redacción aprobada en Consejo de Gobierno de 30 de octubre de 2013).

5. La tesis podrá ser escrita y, en su caso, defendida, en los idiomas habituales para la comunicación científica en su campo de conocimiento. Si la redacción de la tesis se realiza en otro idioma, deberá incluir un amplio resumen en español.

Por otra parte, la Escuela Internacional de Posgrado de la Universidad de Granada ha dictado lo siguiente.

En la cubierta y en la portada de la memoria figurará:

- Universidad de Granada, junto con el escudo institucional.
- Programa de Doctorado.
- Título de la Tesis.
- Autor/a.

En la tesis doctoral deberá figurar en las páginas iniciales una hoja donde tanto la/el doctoranda/o como sus directores/as garanticen que se han respetados los derechos de otros/otras autores/as a ser citados/as, cuando se han utilizado sus resultados o publicaciones.

Fuente: <http://escuelaposgrado.ugr.es/doctorado/estudiantes/tesis> [consultado: 01 de diciembre de 2020].

En cuanto a la modalidad "**agrupación de publicaciones**" la Escuela Internacional de Posgrado de la Universidad de Granada, ha dictado lo siguiente:

La recomendación del Consejo Asesor de Doctorado es que para utilizar este tipo de formato se utilice un mínimo de tres artículos y que se incluya un informe con el factor de impacto de las publicaciones presentadas. En aquellas áreas en las que no sea aplicable este criterio se sustituirá por las bases relacionadas por la Comisión Nacional Evaluadora de la Actividad Investigadora (CNEAI) para estos campos científicos.

Las Comisiones Académicas de los Programas de Doctorado podrían requerir condiciones adicionales a las reflejadas en esta página web. Debería consultar la web del programa de doctorado y de la Escuela de Doctorado correspondiente para comprobar si tiene publicadas condiciones adicionales.

El/la doctorando/a debe respetar los derechos de propiedad intelectual relativos a la difusión de los artículos utilizados en la tesis doctoral. Los artículos deberán figurar en la versión que permita o autorice la revista.

Fuente: <http://escuelaposgrado.ugr.es/doctorado/estudiantes/tesis> [consultado: 01 de diciembre de 2020].

Respecto a lo mencionado anteriormente, la Comisión Académica del **Programa de Doctorado en Historia y Artes** publica en su página web un apartado específico sobre TESIS POR ARTÍCULOS que recoge lo siguiente.

Según el artículo 18º sobre la tesis doctoral de la normativa de las Normas reguladoras de las enseñanzas oficiales de doctorado y del título de doctor por la Universidad de Granada:

1. La tesis doctoral consistirá en un trabajo original de investigación elaborado por el candidato en cualquier campo del conocimiento que se enmarcará en alguna de las líneas de investigación del Programa de Doctorado en el que está matriculado.
2. La tesis debe capacitar al doctorando para el trabajo autónomo en el ámbito de la I+D+i.
3. La tesis debe contar, al menos, con los siguientes contenidos: título, resumen, introducción, objetivos, metodología, resultados, conclusiones y bibliografía.
4. Una tesis doctoral puede también consistir en el reagrupamiento en una memoria de trabajos de investigación publicados por el doctorando en medios científicos relevantes en su ámbito de conocimiento.

Los artículos que configuren la tesis doctoral deberán estar publicados o aceptados con fecha posterior a la obtención del título de grado y del máster universitario, no podrán haber sido utilizados en ninguna tesis anterior y se deberá hacer mención a la Universidad de Granada a través de la afiliación del doctorando.

El Programa de Doctorado en Historia y Artes autorizará el depósito de la tesis en la modalidad de reagrupamiento de artículos cuando se presenten un mínimo de 3 artículos publicados o aceptados para su publicación, en un medio incluido en los índices de impacto admitidos por el programa de doctorado correspondiente y previa autorización de la Comisión Académica.

Si la publicación ha sido realizada por varios autores, además del doctorando, se debe adjuntar la declaración de los restantes autores de no haber presentado dicha publicación en otra tesis doctoral o la renuncia a hacerlo. Los coautores señalarán el trabajo del doctorando en los mencionados trabajos.

En esta modalidad, la tesis debe tener además de los apartados mencionados en el punto anterior, los artículos que la componen, bien integrados como capítulos de la tesis o bien como un Anexo.

5. La tesis podrá ser escrita y, en su caso, defendida, en los idiomas habituales para la comunicación científica en su campo de conocimiento. Si la redacción de la tesis se realiza en otro idioma, deberá incluir un amplio resumen en español.

Fuente: https://doctorados.ugr.es/historia/pages/organizacion/doctorandos#_doku_tesis_por_articulos

[consultado: 01 de diciembre de 2020]

Respecto a los **índices de impacto admitidos**, se han utilizado como referencia las bases relacionadas por la Comisión Nacional Evaluadora de la Actividad Investigadora en la resolución de 23 de noviembre de 2017, que por el campo de Historia, Geografía y Artes establecen lo siguiente.

Campo 10. Historia, Geografía y Artes

3. En la valoración de los artículos se atenderá al medio de difusión empleado, aceptándose como indicio de calidad la publicación en revistas de reconocida valía. La inclusión de las revistas en bases de datos internacionales se considerará como una referencia de calidad; específicamente las indexadas en las bases de datos de la *Web of Science (Arts and Humanities Citation Index, Social Science Citation Index, Science Citation Index, Journal Citation Reports, y Emerging Sources Citation Index)* o las que ocupen posiciones relevantes en *Scimago Journal Rank*, También se tendrán en cuenta como referencia de calidad las revistas incluidas en la *International Medieval Bibliography, RILMS Abstracts of Music Literature*, etc. Igualmente, se podrán considerar otras revistas nacionales e internacionales que pudiendo estar incluidas en diferentes bases de datos presenten a juicio del comité asesor indicios de calidad equiparables y satisfagan los criterios especificados en el apéndice de esta Resolución. Las revistas electrónicas estarán sujetas a los mismos criterios que las demás. La mera indización o indexación de una publicación en un repertorio o una base de datos sin que incluya índices de gradación no es en sí misma un indicio de calidad.

4. En la valoración de libros y capítulos de libro se considerarán de manera especial el prestigio de la editorial y la existencia de un procedimiento selectivo para la aceptación de originales según sistemas recogidos en el *Scholarly Publishers Indicators*, y atendiendo particularmente a la posición que ocupe en este ranking. Se tendrán en cuenta también el número y carácter de las citas recibidas, las reseñas y críticas en revistas especializadas, la colección, la traducción a otras lenguas, etc. Con estos indicios de calidad serán evaluados los catálogos monográficos de artista o catálogo razonado que deberán reflejar claramente que son fruto de la investigación o de la reflexión documentada sobre estética o teoría del arte. Asimismo, se tendrán en cuenta estos criterios para la evaluación de las actas de congresos. No serán consideradas como nuevas aportaciones las traducciones de la propia obra a otras lenguas.

INDICIOS DE CALIDAD DE LAS PUBLICACIONES

A la vista de la normativa citada anteriormente, dentro del capítulo de resultados se incluyen los textos íntegros de cuatro artículos publicados en revistas de impacto y dos capítulos de libros de editoriales de prestigio, que abordan, desde distintas perspectivas, los objetivos de esta investigación.

En las siguientes tablas se detallan los indicios de calidad de dichas publicaciones, según la información disponible más actualizada o correspondiente al año de publicación, cuando fuera disponible.

PUBLICACIÓN nº 1

| | |
|--------------------|---|
| Autor/es | Toscano Maurizio, Rabadán Aroa, Ros Salvador, González-Blanco Elena |
| Año de publicación | 2020 |
| Título | Digital Humanities in Spain: historical perspective and current scenario |
| Revista | El Profesional de la Información |
| ISSN | 1699-2407 |
| Volumen | v. 29, n. 6 |
| Página desde | 1 |
| Página hasta | 22 |
| DOI-URL | https://doi.org/10.3145/epi.2020.nov.01 |
| Título en español | Las Humanidades Digitales en España: perspectiva histórica y escenario actual |

Indicios de calidad



FACTOR DE IMPACTO Y POSICIÓN EN SU CATEGORÍA DEL MEDIO DE DIFUSIÓN

Datos del **Scimago** Journal & Country Rank (2019):

- Impact Factor: 0,48
- Categorías y cuartiles: Cultural Studies (Q1); Communication (Q2); Information Systems (Q2); Library and Information Sciences (Q2)

Datos de la Clasificación Integrada de Revistas Científicas (**CIRC**): Clasificación A

Sello de calidad de las revistas científicas españolas **FECYT** 2020

MIAR ICDS: 10,9

Revista indexada en:

- Academic search complete (Ebsco)
- Academic search premier (Ebsco)
- CIRC
- Communication Source (Ebsco)
- Informe Académico (Cengage)
- Dialnet (Universidad de La Rioja)
- ERIH Plus
- FRANCIS
- ÍndICES CSIC
- Google Scholar
- **Web of Science** - Journal Citation Reports (JCR 1,58)
- Inspec, Information services in physics, electronics and computing
- ISOC, Índice español de ciencias sociales y humanidades (Iedcyt)
- ISTA, Information science and technology abstracts (Ebsco)
- Latindex (Catálogo)
- Lisa, Library and information science abstracts (CSA)
- Lista, Library, information science & technology abstracts (Ebsco)
- **Scopus** (Elsevier)
- Recyt, Repositorio español de ciencia y tecnología (FECYT)

PUBLICACIÓN nº 2

| | |
|---------------------|---|
| Autor/es | Toscano Maurizio |
| Año de publicación | 2018 |
| Título | Andalucía y América. Plataforma colaborativa sobre patrimonio artístico iberoamericano |
| Revista | Quiroga. Revista de Patrimonio Iberoamericano |
| ISSN | 2254-7037 |
| Volumen | 14 |
| Página desde | 36 |
| Página hasta | 43 |
| DOI-URL | https://doi.org/10.5281/zenodo.3403745 |
| Indicios de calidad | <p>FACTOR DE IMPACTO Y POSICIÓN EN SU CATEGORÍA DEL MEDIO DE DIFUSIÓN</p>  <p>Datos del Scimago Journal & Country Rank (2018):</p> <ul style="list-style-type: none"> - Impact Factor: 0,111 - Categorías y cuartiles: Visual Arts and Performing Arts (Q2); Conservation (Q2); Museology (Q2) <p>Datos de la Clasificación Integrada de Revistas Científicas (CIRC): Clasificación B</p> <p>MIAR ICDS: 9,4</p> <p>Revista indexada en:</p> <ul style="list-style-type: none"> - Academic search premier (Ebsco) - Clasificación Integrada de Revistas Científicas – CIRC - Communication Source (Ebsco) - Informe Académico (Cengage) - Dialnet (Universidad de La Rioja) - Fuente Academica Plus (Ebsco) - Directory of Open Access Journals (DOAJ) - ERIH Plus - Resh CSIC - ÍnDICES CSIC - Google Scholar - Web of Science - Emerging Sources Citation Index (Clarivate) - ISOC, Índice español de ciencias sociales y humanidades (Iedcyt) - Latindex (Catálogo) - REDIB - Scopus (Elsevier) - Recyt, Repositorio español de ciencia y tecnología (FECYT) |

PUBLICACIÓN nº 3

| | |
|---------------------|---|
| Autor/es | Toscano Maurizio, Romagnoli Giuseppe |
| Año de publicación | 2020 |
| Título | Atlante dei siti fortificati della provincia di Viterbo, Italia (X-XV secolo). Fonti e metodi per la ricostruzione della rete insediativa bassomedievale |
| Libro | Navarro Palazón, J. y García Pulido, L. (eds), <i>Defensive architecture of the Mediterranean</i> , vol. X |
| ISBN | 978-84-338-6645-5 |
| Editoriales | Universidad de Granada y Universidad Politécnica de Valencia |
| Página desde | 471 |
| Página hasta | 478 |
| DOI-URL | https://dx.doi.org/10.4995/FORTMED2020.2020.11545 |
| Título en español | Atlas de yacimientos fortificados en la provincia de Viterbo, Italia (siglo X-XV). Fuentes y métodos para la reconstrucción de la red de asentamientos medievales |
| Indicios de calidad | <p>PRESTIGIO EDITORIAL Y PROCEDIMIENTO DE SELECCIÓN DE ORIGINALES</p> <p>Scholarly Publisher Indicators (SPI, 2018):</p> <ul style="list-style-type: none"> - Ranking general: 33 de 272 (Q1) – ICEE 7.739 - Especialización temática Historia: 18 de 122 (Q1) – ICEE 1.296 - Especialización temática Arqueología y Prehistoria: 12 de 26 (Q2) – ICEE 0.161 <div style="background-color: black; color: white; padding: 5px; text-align: center;"> <p>SPI Scholarly Publishers Indicators In Humanities and Social Sciences</p> </div> <p>Volumen indexado en:</p> <ul style="list-style-type: none"> - Dialnet https://dialnet.unirioja.es/congreso/5393 - Conference Proceedings Citation Index (Web of Science Core Collection) <p>Procedimientos de selección para la aceptación de originales:</p> <ul style="list-style-type: none"> - informe de especialistas externos a la editorial |

PUBLICACIÓN nº 4

| | |
|---------------------|---|
| Autor/es | Gutiérrez-Rodríguez Mario, Toscano Maurizio, Goldberg Paul |
| Año de publicación | 2018 |
| Título | High-resolution dynamic illustrations in soil micromorphology: a proposal for presenting and sharing primary research data in publication |
| Revista | Journal of Archaeological Science: Reports |
| ISSN | 2352-409X |
| Volumen | 20 |
| Página desde | 565 |
| Página hasta | 575 |
| DOI-URL | https://doi.org/10.1016/j.jasrep.2018.05.025 |
| Indicios de calidad | <p>FACTOR DE IMPACTO Y POSICIÓN EN SU CATEGORÍA DEL MEDIO DE DIFUSIÓN</p>  <p>Datos del Scimago Journal & Country Rank (2018)</p> <ul style="list-style-type: none"> - Impact Factor: 0,79 - Categorías y cuartiles: Archeology (Q1); Archeology (arts and humanities) (Q1); History (Q1) <p>Datos de la Clasificación Integrada de Revistas Científicas (CIRC): Clasificación A</p> <p>MIAR ICDS: 8,2</p> <p>Revista indexada en:</p> <ul style="list-style-type: none"> - Academic search premier (Ebsco) - Clasificación Integrada de Revistas Científicas – CIRC - Communication Source (Ebsco) - ERIH Plus - Google Scholar - Index Islamicus (Brill) - Arts and Humanities Citation Index (Clarivate) - Web of Science - Emerging Sources Citation Index (Clarivate) - Scopus (Elsevier) |

PUBLICACIÓN nº 5

| | |
|---------------------|--|
| Autor/es | Toscano Maurizio |
| Año de publicación | 2018 |
| Título | Where the researcher cannot get. Open platforms to collaborate with citizens on cultural heritage research data |
| Libro | Romero Frías, E. y Bocanegra Barbecho, L. (eds), Ciencias Sociales y Humanidades Digitales aplicadas: casos de estudio y perspectivas críticas |
| ISBN | 978-84-338-6318-8 |
| Editoriales | Universidad de Granada |
| Página desde | 538 |
| Página hasta | 560 |
| DOI-URL | http://doi.org/10.5281/zenodo.3758663 |
| Título en español | Donde el investigador no puede llegar. Plataformas web de colaboración con la sociedad sobre datos de investigación en abierto |
| Indicios de calidad | <p>PRESTIGIO EDITORIAL Y PROCEDIMIENTO DE SELECCIÓN DE ORIGINALES</p> <p>Scholarly Publisher Indicators (SPI, 2018):</p> <ul style="list-style-type: none"> - Ranking general: 33 de 272 (Q1) – ICEE 7.739 - Especialización temática Historia: 18 de 122 (Q1) – ICEE 1.296 - Especialización temática Arqueología y Prehistoria: 12 de 26 (Q2) – ICEE 0.161 <div style="text-align: center; background-color: black; color: white; padding: 5px; margin: 10px 0;">  </div> <p>Volumen indexado en:</p> <ul style="list-style-type: none"> - Dialnet https://dialnet.unirioja.es/servlet/libro?codigo=741086 <p>Reseñas:</p> <ul style="list-style-type: none"> - Ignacio Jesús Serrano Contreras, en <i>Quiroga. Revista de Patrimonio Iberoamericano</i>, Num. 14, Universidad de Granada, 2018, ISSN: 2254-7037, pp. 164-165. https://dialnet.unirioja.es/servlet/articulo?codigo=6824898 <p>Procedimientos de selección para la aceptación de originales:</p> <ul style="list-style-type: none"> - informe de un comité de lectura de la editorial - informe de especialistas externos a la editorial |

| | |
|---------------------|---|
| Autor/es | Bocanegra Barbecho Lidia, Toscano Maurizio, Delgado Anés Lara |
| Año de publicación | 2017 |
| Título | Co-creación, participación y redes sociales para hacer historia. Ciencia con y para la sociedad |
| Revista | Historia y Comunicación Social |
| ISSN | 1137-0734 |
| Volumen | 22-2 |
| Página desde | 325 |
| Página hasta | 346 |
| DOI-URL | https://doi.org/10.5209/HICS.57847 |
| Indicios de calidad | <p>FACTOR DE IMPACTO Y POSICIÓN EN SU CATEGORÍA DEL MEDIO DE DIFUSIÓN</p>  <p>Datos del Scimago Journal & Country Rank (2017):</p> <ul style="list-style-type: none"> - Impact Factor: 0,218 - Categorías y cuartiles: History (Q1); Communication (Q3); Sociology and Political Science (Q3) <p>Datos de la Clasificación Integrada de Revistas Científicas (CIRC): clasificación A</p> <p>MIAR ICDS: 10,9</p> <p>Revista indexada en:</p> <ul style="list-style-type: none"> - Clasificación Integrada de Revistas Científicas – CIRC - Dialnet (Universidad de La Rioja) - Fuente Academica Plus (Ebsco) - American History and Life (Ebsco) - Historical Abstracts (Ebsco) - Political Science Complete (Ebsco) - ABI/INFORM (ProQuest) - Sociological abstracts (ProQuest) - Worldwide Political Science Abstracts (ProQuest) - MLA - Modern Language Association Database - Directory of Open Access Journals (DOAJ) - ERIH Plus - Resh CSIC - ÍNDICES CSIC - Google Scholar - Web of Science - Arts and Humanities Citation Index (Clarivate) - Journal Citation Reports, JCR (Clarivate Analytics) - ISOC, Índice español de ciencias sociales y humanidades (Iedcyt) - Latindex (Catálogo) - Scopus (Elsevier) |

I. INTRODUCCIÓN

La tesis doctoral que aquí se presenta – en formato de compendio de publicaciones – es fruto de un trabajo de investigación que empieza a gestarse con anterioridad a la matriculación del candidato en el Programa de Doctorado de la Universidad de Granada, y toma forma a partir de la actividad profesional del doctorando, arqueólogo con un master en la gestión informática de los datos de investigación en humanidades. Actualmente se identifica este perfil híbrido como Humanista Digital, por la definición que en la última década adquieren aquellos investigadores que incorporan herramientas informáticas en las metodologías propias de las disciplinas humanísticas tradicionales.

En concreto, la tesis estudia la manera en que los sistemas de información web¹ influyen, o tienen el potencial de influir, en cómo los investigadores en Historia, Arte o Arqueología organizan, recogen y verifican fuentes y datos, comparten y difunden su trabajo, así como las implicaciones asociadas. Las diferentes etapas implicadas en la creación y uso de un sistema de información web – modelación conceptual, implementación informática, flujo de datos, dinámica de trabajo, publicación, co-creación entre investigadores y colaboración en línea con la sociedad – han sido analizadas por el candidato a través de una serie de publicaciones desarrolladas a lo largo del periodo 2017-2020. Algunas de ellas, en concreto cuatro sobre seis, han sido llevadas a cabo en coautoría con otros investigadores: ocho en total entre arqueólogos, historiadores, filólogos e informáticos; lo que habla de cierta necesidad y capacidad de establecer colaboraciones a nivel multidisciplinar e interdisciplinar. Todas ellas se han revelado importantes para atajar temas de investigación fronterizos, que se colocan en un espacio del conocimiento compartido entre disciplinas humanísticas, tecnologías de la información y de la comunicación y sociedad (figura 1).

¹ A lo largo de este trabajo, para referirnos a los sistemas de información web, utilizaremos indistintamente la forma extensa o el acrónimo SIW.

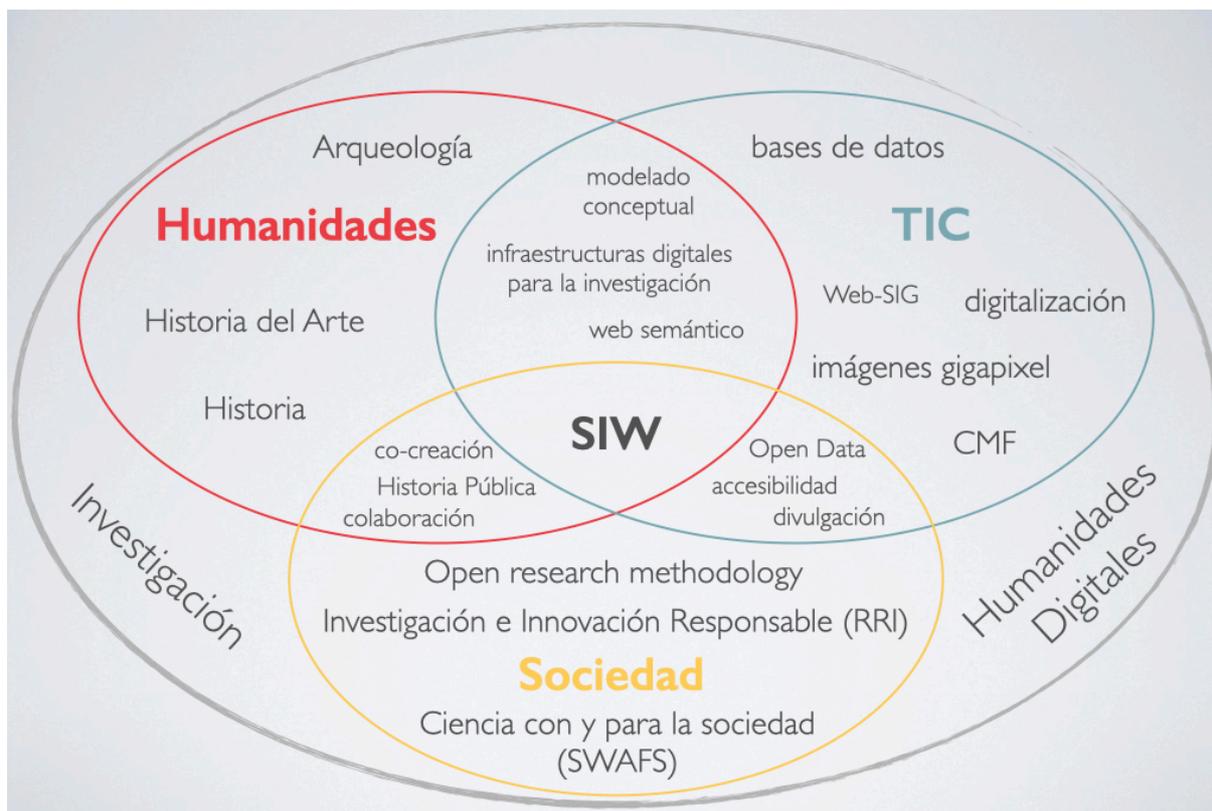


Figura 1. Esquema conceptual de la relación entre los diferentes componentes involucrados en la presente tesis doctoral. Fuente: elaboración propia.

En su conjunto, el presente trabajo de investigación doctoral se ha planteado con una perspectiva más horizontal que vertical, con el fin de que primara la amplitud de temas frente a la exhaustividad de análisis de los diferentes argumentos tratados. Ha sido concebido de esta manera para ofrecer una visión unitaria y de conjunto sobre varios aspectos implicados en el uso de los sistemas de información personalizados y para adaptarse a un corte fundamentalmente metodológico, más que histórico o temático.

Nuestra ambición con este trabajo no es aportar grandes novedades relativas al diseño, construcción y publicación en línea de sistemas de información, sino más bien analizar las consecuencias, a nivel metodológico, de estos procesos sobre las tendencias historiográficas tradicionales; así como las respuestas que pueden propiciar a partir de necesidades concebidas y generadas desde proyectos de investigación en humanidades.

I.1. OBJETO DE ESTUDIO

El objeto de estudio de este trabajo de investigación son los sistemas de información web, no tanto a nivel teórico o de desarrollo informático, sino a través de las implicaciones metodológicas implícitas en el proceso de diseño, desarrollo y uso, en el marco de proyectos de investigación en Historia, Historia del Arte y Arqueología. El contexto de análisis son los proyectos focalizados en la investigación en ciencias humanas y patrimonio cultural, así como en la divulgación y transferencia de los resultados, normalmente promovidos por grupos universitarios. Dentro de esta categoría, y en función de los usuarios involucrados, se han distinguido al menos tres tipologías: proyectos individuales, proyectos de colaboración entre investigadores o grupos y proyectos de colaboración entre investigadores y sociedad.

Han quedado por tanto excluidas de este análisis las iniciativas en las que el desarrollo de un sistema de información esté focalizado en la catalogación en sí de objetos o fuentes documentales, ajena a un determinado objetivo de investigación. Estos proyectos son normalmente promovidos por archivos, museos o bibliotecas², que por sus propios fines institucionales carecen de una determinada perspectiva de estudio y entienden el proceso de digitalización y gestión informatizada de las fuentes como una labor de servicio, cuyo objetivo principal es satisfacer las necesidades de un público culturalmente interesado más que proporcionar recursos para una investigación determinada.

En esta amplia distinción se esconden, por supuesto, muchos puntos en común entre estos dos enfoques. Ambos hacen uso de tecnologías similares para construir y gestionar colecciones digitales (véase apartado III.1), y la principal comunidad de usuarios de archivos y bibliotecas siguen siendo los investigadores en ciencias humanas. No obstante, las diferencias entre estas dos tipologías de iniciativas son lo suficientemente definidas como para que nos centramos exclusivamente en las infraestructuras informáticas desarrolladas en el marco de proyectos de investigación, como una actividad distinta en el ámbito de las humanidades.

Existe claramente una tensión, en la arquitectura de la información, entre estandarización, compatibilidad e interoperabilidad por un lado y personalización, atomización y segmentación de los datos por el otro, con argumentos válidos a favor de ambas tendencias (González-Pérez 2018). En este trabajo, nos hemos concentrados sobre la segunda tendencia, ya que permite concretar, particularizar y personalizar mucho más el modelo de datos y las funcionalidades del sistema, para conseguir finalmente más potencia inferencial y más capacidad investigadora, con el fin que la herramienta informática desarrollada ayude concretamente en el planteamiento, desarrollo y aproximación progresiva desde las preguntas hacia los resultados de investigación. Todos los sistemas de información web que se describen y analizan en este trabajo no están planteados para un enfoque global, sino que miran a responder a las necesidades concretas de proyectos de investigación específicos y sobre ellas se conforman.

² Conocidos con el acrónimo inglés G.L.A.M.: *galleries, libraries, archives* y *museums*.

Un sistema de información es un conjunto de personas, datos, procesos y tecnología que interactúan para recoger, procesar, almacenar y proveer la información necesaria para el correcto funcionamiento de una organización (Whitten *at al.* 2004). Nacen en ámbito empresarial para gestionar, almacenar, procesar y recuperar los datos que lo componen, con el fin de producir información. Se trata de entornos colaborativos por definición, que abogan hacia la puesta en común y el trabajo en grupo, cada vez más perentorio en humanidades. No nacen con la Web, pero encuentran hoy en día en Internet su hábitat ideal, gracias a un entorno de navegación amigable, que emplea herramientas con las que los usuarios están ampliamente familiarizados, como los navegadores y los métodos de búsqueda, y es capaz de operar de forma prácticamente independiente, ya sea a nivel hardware que software, del sistema operativo lado servidor y cliente. En el contexto de la investigación científica, el uso de la Web resulta finalmente un componente sustancial, que habilita y favorece la abertura del sistema al exterior, hacia otros grupos de investigación o con la sociedad.

En buena medida, el trabajo de investigación se ha generado a partir de una serie de casos de estudio desarrollados por el autor de esta tesis en colaboración con investigadores de diferentes disciplinas, expertos en su correspondiente dominio disciplinar. Incluso si el análisis se basa sobre casos concretos, los contextos temáticos y cronológicos son muy variados y los resultados y conclusiones se pueden aplicar a otros ámbitos disciplinares y a diferentes temas de estudio.

El encuadre disciplinar desde que abordamos el análisis de los sistemas de información web es aquel de las Humanidades Digitales. Se trata de una tendencia historiográfica relativamente reciente que, en este contexto, resulta apropiado definir a través de las palabras de Stephen Ramsay que, en una provocadora charla en 2011 frente a la Asociación de Lenguas Modernas de América (MLA) titulada “Who’s in and Who’s Out”, otorga especial importancia al proceso de creación y producción de nuevos contenidos y artefactos, afirmando: “Personally, I think Digital Humanities is about building things. [. . .] If you are not making anything, you are not — in my less-than-three-minute opinion — a digital humanist”³ (Ramsay 2013a). Básicamente, según Ramsay, los humanistas digitales son en primer lugar constructores, hacedores: no sólo “investigan” cosas; “hacen” cosas. Y añade, en su otra intervención “On building” (Ramsay 2013b), “there’s always been a profound—and profoundly exciting and enabling—commonality to everyone who finds their way to dh [digital humanities]. And that commonality, I think, involves moving from reading and critiquing to building and making”⁴. Nosotros compartimos esta visión de los humanistas digitales como hacedores, y defendemos que es desde la creación de cosas concretas, y a través de la

³ Publicado originalmente el 8 de enero de 2011 en <http://stephenramsay.us/text/2011/01/08/whos-in-and-whos-out.html> y actualmente visible a través de la *Wayback Machine* <http://web.archive.org/web/20121015012254/http://stephenramsay.us/text/2011/01/08/whos-in-and-whos-out.html>.

⁴ Publicado originalmente el 11 de enero de 2011 en <http://stephenramsay.us/text/2011/01/11/on-building.html> y actualmente visible a través de la *Wayback Machine* <http://web.archive.org/web/20121015012052/http://stephenramsay.us/text/2011/01/11/on-building.html>.

resolución de problemas y necesidades específicas de investigación, que la reflexión metodológica y epistemológica se genera y se desarrolla.

Resumiendo, se han elegido los sistemas de información web como objeto de estudio por sus características multifacéticas. Constituyen una pieza central en el proceso de digitalización de la investigación que, a través de las diferentes fases de planteamiento, modelado, implementación, recogida-almacenamiento-interrelación de datos, análisis y publicación, propicia y en cierto modo obliga el planteamiento de nuevas preguntas y la reformulación de aquellas ya planteadas, lo que en definitiva constituye la cuestión nuclear y el rol de las Humanidades Digitales en el mundo actual (Rodríguez-Ortega 2018).

I.2. JUSTIFICACIÓN DE LA INVESTIGACIÓN

Las Humanidades Digitales (Schreibman *et al.* 2004; Burdick *et al.*, 2012; Schreibman *et al.* 2016) constituyen una tendencia destacada en la investigación a nivel nacional e internacional, ya sea como objeto de estudio que como temática preferente de financiación. Dentro de este marco disciplinar y metodológico, se ha planteado el presente proyecto de investigación, tomando los sistemas de información web como el eje vertebrador para analizar el proceso de informatización de la investigación en humanidades. Estos sistemas constituyen en muchos casos el punto de partida de ese recorrido y la primera ocasión en la que el investigador que utiliza métodos tradicionales se enfrenta con una serie de problemáticas y oportunidades planteadas por el cambio digital. A la vez, las bases de datos que sustentan esos sistemas de información ocupan un lugar preferente porque guardan lo más preciado, es decir los datos, que una vez almacenados de manera estructurada pueden llegar a explotarse y reelaborarse a lo largo de décadas.

Los SIW, al contrario de los datos, que en principio son para siempre, son hijos de sus tiempos y de las tecnologías disponibles y empleadas en un determinado momento, que son contingentes y sufren una rápida evolución, también en función de los dispositivos que se utilizan para interactuar con ellos. Cualquier análisis que se haga al respecto es una instantánea del estado actual de la evolución de la tecnología en ese determinado sector. Por este motivo, ha sido necesario realizar, en primer lugar, una revisión de las soluciones software actualmente disponibles en el mercado y más frecuentemente utilizadas en humanidades o para la gestión de recursos materiales e inmateriales del patrimonio cultural, que sirva también como estado actual de la cuestión (véase apartado III.1).

La presente tesis doctoral tiene un corte prevalentemente metodológico y enfoca los SIW desde la perspectiva del funcionamiento y de la aplicación, más que desde el desarrollo del software en sí. La elección del objeto de estudio se justifica en buena medida por la trascendencia que el planteamiento de un sistema de información adquiere sobre el proyecto que lo genera, así como por su potencial proyección fuera del ámbito académico, con el fin de responder a necesidades de co-creación y participación ciudadana. Este análisis está, asimismo, justificado por la alta demanda que existe por este tipo de artefactos en proyectos de investigación en humanidades (véase apartado IV.1. - 3.4 *Digital resources for research and dissemination*). Se observa un creciente interés en la comunidad investigadora en tomar ventaja de los recursos proporcionados por la aplicación de las TIC en Humanidades y, en particular, en el uso de recursos web para el rescate, la organización, la búsqueda y el intercambio de los datos de investigación. Este interés se ve reflejado en la progresiva consolidación de las Humanidades Digitales (véase apartado VI.1).

Por un lado, las humanidades necesitan soluciones para ayudar a gestionar las crecientes cantidades de datos e información que se están generando, y respaldar las prácticas asociadas a la generación de conocimiento a través de enfoques cuantitativos, que

han crecido progresivamente en popularidad. Cada vez más, la investigación en humanidades depende de cómo se configura, se estructura y se gestiona la información y de cómo, a partir de ella, se genera el conocimiento mediante procesos interpretativos.

Por otro lado, los enfoques participativos han hecho mella en las tendencias actuales de la historiografía, en particular en el marco metodológico de la Historia Pública⁵ y de la Historia Pública Digital (Gallini y Noiret 2011; Noiret 2015a; 2015b; 2018; Pons 2018; Mira y Zurita 2020). A la vez, se ha detectado un creciente interés de la sociedad misma por participar en proyectos colaborativos en humanidades y patrimonio cultural (Bocanegra Barbecho 2020), a través de las redes sociales y en el marco de la iniciativa SwafS del programa H2020⁶, lo que justifica proporcionar un marco metodológico que ayude en el planteamiento de los mismos.

En definitiva, se trata de un trabajo que mira en poner a disposición del investigador en humanidades las herramientas necesarias no tanto para diseñar o desarrollar en propio los artefactos informáticos que pueda necesitar, si no más bien para comprender y sopesar su utilidad y así evaluar cuándo y cómo sea necesario plantearse el desarrollo de un sistema personalizado para la gestión informatizada de los datos, las implicaciones en el desarrollo del mismo, así como sus potencialidades, limitaciones y consecuencias.

⁵ El *National Council on Public History* (NCPH) en EEUU, país de origen y consolidación de esa disciplina, la define como un movimiento, una metodología y un enfoque que promueve el estudio y la práctica en colaboración de la historia. Esta colaboración es cada vez más habilitada y mediada a través de infraestructuras digitales, y de ahí el campo más específico de Historia Pública Digital. National Council on Public History, 2020. *How do we define Public History?* <https://ncph.org/what-is-public-history/about-the-field/>.

⁶ <https://ec.europa.eu/programmes/horizon2020/en/h2020-section/science-and-society>.

II. OBJETIVOS

El objetivo general de este trabajo de investigación era analizar el impacto que el uso de los SIW ha tenido y tiene sobre el proceso de investigación en humanidades, en relación con la estructuración y el flujo de gestión de los datos, su almacenamiento, su recuperación y su presentación; así como sobre la dinámica de trabajo entre investigadores y de colaboración con la sociedad en su conjunto.

Para alcanzar este objetivo general, se han planteado una serie de objetivos específicos, que se han ido desarrollando de manera individual a través de las publicaciones que constituyen el núcleo de los resultados de la presente tesis doctoral. Estos objetivos específicos han sido los siguientes:

1. Estudiar la evolución y el escenario actual de las Humanidades Digitales en España, para identificar y describir el contexto de investigadores, instituciones y proyectos a través de los cuales se han desarrollado, y se mantienen, un número relevante de recursos digitales para la investigación.
2. Analizar el proceso de diseño e implementación de un sistema de información web personalizado a medida, que se adapte a las características de las fuentes históricas objeto de estudio, al proceso de recogida y verificación de los datos y a los objetivos propios de la investigación que lo genera.
3. Evaluar el potencial de la Web como plataforma de presentación e intercambio de fuentes primarias de investigación y para la ilustración dinámica de publicaciones científicas.
4. Analizar, en diferentes ámbitos disciplinares, las implicaciones y consecuencias que conlleva el almacenamiento local y restringido de los datos de investigación, frente al uso de plataformas abiertas y en línea, desarrolladas para la colaboración a distintos niveles entre investigadores, grupos de investigación y la sociedad en general.
5. Examinar las oportunidades que el uso de los sistemas de información web, y en general las tecnologías de la información y de la comunicación (TIC), ofrecen a proyectos de *crowdsourcing*⁷ y Ciencia ciudadana.

⁷ Término polisémico, derivado del inglés *crowd* (multitud) y *outsourcing* (recursos externos), que se podría traducir en español como “colaboración abierta y distribuida” o “externalización abierta de tareas” [Fuente: Wikipedia]. En su estudio, Estelles-Arolas y Gonzalez-Ladron-de-Guevara relevan 40 definiciones diferentes desde 32 artículos distintos publicados entre 2006 y 2011 (2006, 2; 2008, 7; 2009, 4; 2010, 10; 2011, 9), algunos de ellos de los mismos autores (Estelles-Arolas y Gonzalez-Ladron-de-Guevara 2012).

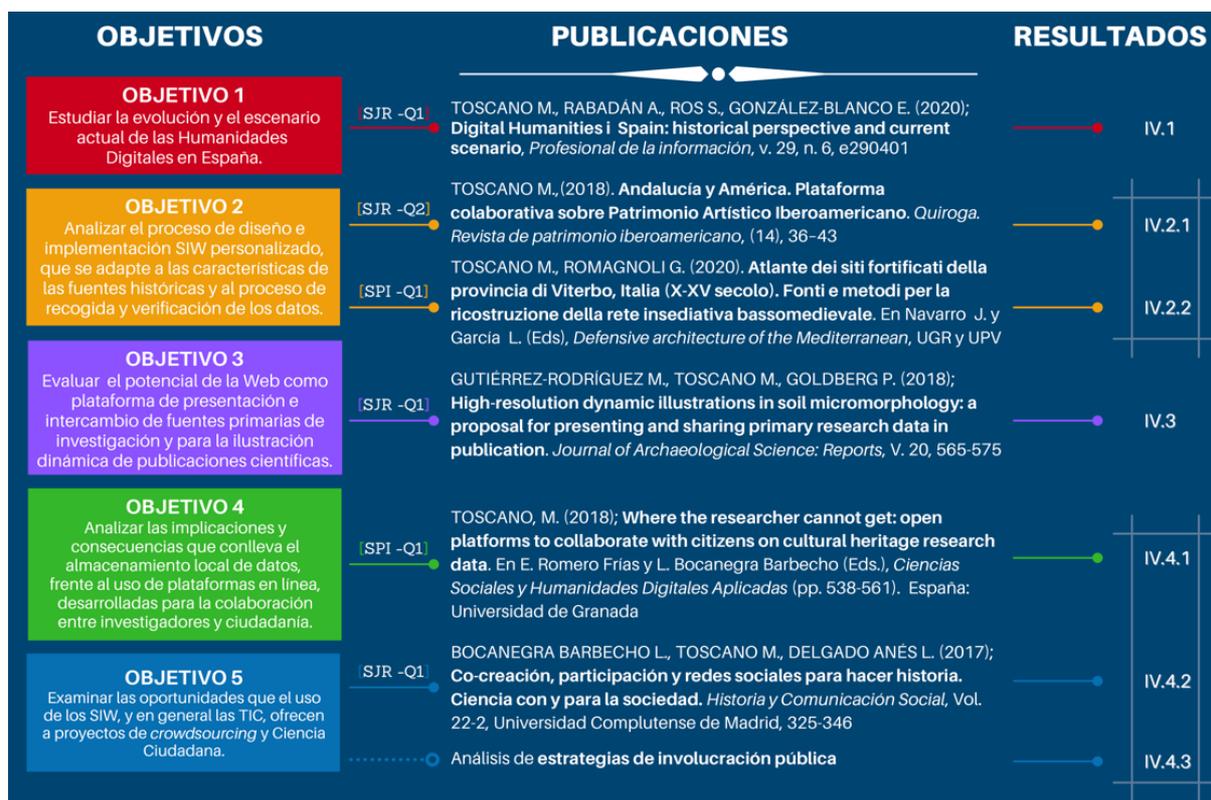


Figura 2. Relación entre objetivos, publicaciones y resultados. Fuente: elaboración propia.

A cada objetivo corresponden uno o dos resultados de investigación, plasmados a través de publicaciones científicas, publicadas o en curso de elaboración (figura 2). Todas ellas analizan y describen un aspecto de la investigación a partir de sistemas de información creados por el autor de este trabajo. Otras experiencias de creación de SIW, llevadas a cabo en paralelo con la elaboración de la tesis, no se han incorporado en publicaciones científicas y por este motivo no se incluyen en este trabajo, pero han generado reflexiones relevantes que han permitido mejorar y completar los resultados que aquí se presentan.

Como se ha podido apreciar a través del esquema conceptual de la figura 1, los tres ámbitos interconectados que estructuran esta tesis son las humanidades, las tecnologías de la información y la comunicación y la sociedad. En humanidades, nos hemos centrado sobre todo en problemáticas de investigación procedentes de la Historia, de la Historia del Arte y de la Arqueología. Desde las TIC hemos tomado herramientas y métodos para la organización, manipulación y presentación de los datos. Finalmente, hemos abordado el tema de la involucración de la sociedad en el proceso de investigación a través de los conceptos transversales de Investigación e Innovación Responsable (RRI por su sigla en inglés, *Responsible Research and Innovation*) y Ciencia con y para la sociedad (SWAFS por su sigla en inglés, *Science with and for society*), términos consensuados y utilizados por la Comisión Europea para los proyectos marcos de financiación de la investigación. Todas estas componentes, y otras más que se quedan al margen de esta investigación, como el tema de la propiedad intelectual y de la protección de los datos, se integran a través de los sistemas de información web, que en este

sentido juegan un rol central en el proceso de consolidación y maduración de las Humanidades Digitales.

En el capítulo de resultados (véase capítulo IV) volveremos reiterativamente sobre el esquema conceptual de figura 1, cada vez que introduciremos, a partir de las diferentes publicaciones, las componentes principales, así como los diferentes temas específicos de investigación involucrados.

III. METODOLOGÍA

A nivel metodológico, el presente trabajo de investigación se enmarca dentro del ámbito disciplinar de las Humanidades Digitales. Partiendo desde problemáticas propias de disciplinas como la Historia, la Arqueología y la Historia del Arte, y a través del conocimiento de herramientas y métodos propios de las tecnologías de la información y la comunicación, se proporcionan soluciones digitales a problemáticas de gestión, co-creación, consulta y divulgación de conjuntos de datos de investigación. En el primer apartado de este capítulo presentamos el estado de la cuestión en cuanto a soluciones informáticas para la creación de sistemas de información web en humanidades y patrimonio cultural. En el segundo apartado, describimos en detalle la metodología empleada en el presente trabajo para alcanzar los objetivos planteados (véase capítulo II).

III.1. SOLUCIONES SOFTWARE PARA LAS HUMANIDADES DIGITALES Y EL PATRIMONIO CULTURAL: ESTADO DE LA CUESTIÓN

En el proceso de desarrollo de un sistema de información web, una vez realizado el diseño del modelo de datos y evaluadas las funcionalidades necesarias para cubrir las exigencias específicas (véase apartado III.2), es necesario elegir el entorno de trabajo más apropiado para la sucesiva implementación a nivel informático. En este apartado nos centraremos concretamente sobre este aspecto, a través de una revisión crítica de las soluciones software existentes para la implementación de sistemas de información web en humanidades y gestión del patrimonio cultural. Lo haremos a través de una comparación de distintos paquetes que suelen utilizarse en esos ámbitos, sean ellos genéricos o desarrollados para un dominio específico, y teniendo en cuenta sus principales funcionalidades, ventajas e inconvenientes.

Con el fin de ofrecer una panorámica suficientemente amplia y coherente de las herramientas actualmente disponibles en el mercado, hemos establecido unos **criterios básicos de selección**:

- software gratuito y de código abierto (FOSS⁸), que permita modificaciones, integraciones y personalizaciones a medida;
- en desarrollo y mantenimiento activo;
- con evidencia de casos de uso y aplicaciones demostradas en el ámbito de las humanidades o del patrimonio cultural.

Existe una cantidad relevante de soluciones tecnológicas desarrolladas y mantenidas para dar respuesta a la necesidad de gestionar información (y documentación) estructurada en la Web, y mientras que algunas son claramente alternativas entre sí, otras cubren diferentes aspectos, con funcionalidades que se solapan sólo en parte, y son

⁸ Free and Open Source Software: https://es.wikipedia.org/wiki/Software_libre_y_de_c%C3%B3digo_abierto.

más difícilmente comparables. Un listado representativo, aunque incompleto y temporal, de aplicaciones FOSS utilizadas en el campo de las Humanidades Digitales comprende: MediaWiki, Drupal, BackdropCMS, WissKI, Mukurtu, Omeka Classic, Omeka S, Fedora Commons, Islandora, AtoM, Arches, CollectiveAccess, ResearchSpace, nodedgoat, AlchemyCMS, Zooniverse, Pybossa, EnrichEuropeana y CrowdHeritage.

Podemos agrupar esos softwares en cuatro diferentes **categorías** (tabla 1): los sistemas de gestión del contenido (conocidos como CMS, según el acrónimo inglés de *Content Management Systems*⁹); los sistemas de gestión de activos digitales (conocidos como DAM, según el acrónimo inglés de *Digital Assets Management*¹⁰); los entornos virtuales de investigación (conocidos como VRE, según el acrónimo inglés de *Virtual Research Environment*¹¹); las plataformas para el *crowdsourcing* y la colaboración científica con la sociedad.

Como veremos a continuación, todas las herramientas seleccionadas proporcionan al desarrollador un entorno de trabajo preparado para la configuración, con un mayor o menor margen de customización. Se han excluido de esta revisión, a pesar de ser soluciones muy populares, paquetes software que ofrecen por defecto una muy limitada personalización y donde la gestión del contenido desestructurado prima sobre la gestión y la manipulación de datos y metadatos, como es el caso de Joomla¹² o WordPress¹³. Al otro margen del espectro de posibilidades para el desarrollo de un sistema de información web se encuentra la opción de la programación desde cero de la plataforma, combinando un Sistema de Gestión de Bases de Datos¹⁴ (conocidos como DBMS, según el acrónimo inglés de *Database Management Systems*) con un *back end* y un *front end*¹⁵ desarrollados a medida. En particular, resulta frecuente la combinación denominada LAMP, que es la sigla de Linux (sistema operativo), Apache (servidor web), MySQL (base de datos relacional) y PHP (lenguaje de programación). En principio, este enfoque permite el grado más alto de personalización y proporciona un mayor nivel de seguridad frente a ataques externos, ya que estas soluciones están menos sujetas, aunque no exentas, a la revelación pública de vulnerabilidades¹⁶ y a la

⁹ https://en.wikipedia.org/wiki/Content_management_system.

¹⁰ https://en.wikipedia.org/wiki/Digital_asset_management.

¹¹ https://en.wikipedia.org/wiki/Virtual_research_environment.

¹² <https://www.joomla.org/>.

¹³ <https://wordpress.com/>.

¹⁴ Los SGBD, o DBMS, con arquitectura relacional más populares en el mercado son: MySQL, MariaDB, SQLite, PostgreSQL, Microsoft SQL Server y Oracle. También existe la posibilidad de utilizar bases de datos orientadas a objetos.

¹⁵ Según el contexto, para referirse a *front end* se usan otros términos como frontal o interfaz de usuario, mientras que a *back end* se le llama servidor, motor o modo administrador. En diseño de software, el *front end* es la parte del software que interactúa con los usuarios y el *back end* es la parte que procesa la entrada desde el *front end* [Fuente: Wikipedia].

¹⁶ La revelación de agujeros de seguridad o vulnerabilidades representa el principal riesgo para la seguridad de un sistema de información basado en un CMS o DAM, ya que expone a ataques automatizados en larga escala, que pueden comprometer incluso sistemas que no son objetivos directos en la ejecución de código malicioso. Famosos en este sentido son los casos denominados Drupalgeddon 1 y 2, ocurridos respectivamente en 2014 (<https://www.drupal.org/forum/newsletters/security-advisories->

necesidad de instalar actualizaciones de seguridad (véase apartado III.2). Por otro lado, requiere de competencias más especializadas para la implementación del sistema, lo que se traduce en un coste más alto, en particular para el desarrollo inicial, pero también para el simple mantenimiento adaptativo a medio-largo plazo. Expone, además, a un mayor riesgo frente al denominado *bus factor*¹⁷, ya que la completa personalización de un sistema conlleva un grado más alto de dependencia con respecto al desarrollador inicial y reduce la posibilidad de transferencia del mantenimiento/ampliación del mismo a otra persona o grupo. Otro coste añadido resulta de la necesidad de producir y mantener documentación sobre el código generado y de hacerse cargo, en propio, de asegurar la compatibilidad del sistema con librerías externas y tecnologías lado servidor, que inevitablemente se irían actualizando en el tiempo.

En definitiva, hemos elegido, en el contexto de esta tesis, y por consecuencia en esta revisión, limitarnos al uso de soluciones software existentes, pero altamente adaptables. El **principio** que nos ha guiado en esta elección ha sido privilegiar, por las razones expresadas anteriormente, el uso de código común para proyectos personalizados en lugar de código personalizado para proyectos que en muchos casos comparten, de base, características y funcionalidades comunes.

[for-drupal-core/2014-10-15/sa-core-2014-005-drupal-core-sql](https://www.drupal.org/sa-core-2014-005-drupal-core-sql)) y en 2018 (<https://www.drupal.org/sa-core-2018-002>), en los que se descubrieron vulnerabilidades que posibilitaban el hacheo automático de plataformas no actualizadas basadas en Drupal y la creación en pocas horas de puertas traseras que no venían eliminadas por la posterior actualización del software (<https://www.drupal.org/PSA-2014-003>).

¹⁷ https://en.wikipedia.org/wiki/Bus_factor.

| Categoría | Software | Versión analizada | Actualizada el | Lanzamiento | Lenguajes de programación | Desarrollador | Licencia |
|---------------------------|-------------------|-------------------|----------------|-------------|------------------------------|-------------------------------|------------------|
| Content Management System | Media Wiki | 1.35 | 25/9/20 | 2003 | PHP, JavaScript | Fundación Wikimedia | GPLv2+ |
| | Drupal | 7.77 | 03/12/20 | 2011* | PHP, JavaScript | Asociación Drupal | GPLv2/v3 |
| | Drupal | 8.9.11 | 03/12/20 | 2015* | PHP, JavaScript | Asociación Drupal | GPLv2/v3 |
| | Backdrop | 1.17.4 | 26/11/20 | 2015 | PHP, JavaScript | Backdrop PMC | GPLv2+ |
| | Mukurtu | 2.1.6 | 23/11/20 | 2012 | PHP, JavaScript | CDSC - WSU | GPLv3 |
| | WissKI | 8.x-2.3 | 06/11/20 | 2012 | PHP, JavaScript | FAU-GNM-ZFMK | GPLv2+ |
| | Omeka | 2.6.1 | 02/05/18 | 2008 | PHP | CHNM - GMU | GPL |
| | Omeka S | 3.0.1 | 20/10/20 | 2019 | PHP, JavaScript | CHNM - GMU | GPL |
| | Alchemy | 5.0 | 17/07/20 | 2010 | Ruby | C. Fregin, T. von Deyen | BDS |
| Digital Assets Management | Fedora | 5.1 | 26/07/19 | 2003 | Java | DuraSpace | Apache 2.0 |
| | Islandora | 7.x-1.13 | 03/09/19 | 2006 | PHP, JavaScript, Java | Islandora Foundation | GPL |
| | AtoM | 1.3.1** | 27/05/13 | 2008 | PHP, XSLT, Javascript | Artefactual Systems | GPLv3 |
| | Arches | 5.1 | 27/10/20 | 2013 | Python, Javascript | Getty CI - World Monum.s Fund | GPLv3 |
| | Collective Access | 1.7.9 | 01/01/21 | 2009 | PHP, JavaScript | Whirl-i-Gig | GPLv3 |
| VRE*** | Research Space | 3.4.0 | 20/02/20 | 2018 | JavaScript, TypeScript, Java | The British Museum | GPLv3+ |
| | Nodegoat | 7.3 | 17/09/19 | 2013 | PHP, JavaScript, C++ | lab1100 | GPLv3 |
| Crowdsourcing | Zoomuni-verse | en línea | 20/12/20 | 2009 | JavaScript, Ruby | Fingerprint Digital Media | Apache 2.0 |
| | Pybossa | 3.3.0 | 20/12/20 | 2013 | Python | Scifabric | GPLv3+ |
| | Enrich Europeana | 1.0 | 01/12/20 | 2019 | PHP, JavaScript, Java | EnrichEuropean project | GPLv2/ EUPLV.1.2 |
| | Crowd Heritage | 1.0 | 05/12/20 | 2019 | JavaScript | AILS-Lab | EUPL V.1.2 |

* Para Drupal indicamos la fecha de publicación de cada versión en vez que la fecha inicial de lanzamiento del programa, que fue 2001.

** La versión 2.0 se encuentra actualmente en desarrollo, sin una fecha de publicación estimada.

*** Virtual Research Environment.

Tabla 1. Características de las plataformas revisadas. Fuente: elaboración propia.

MediaWiki¹⁸ es el software de base de Wikipedia y de otros proyectos de la Fundación Wikimedia (Wikcionario, Wikilibros, etc). Se lanzó en 2003 y desde entonces ha tenido una gran expansión y se ha utilizado en un gran número de wikis, con contenidos muy diferentes. Su uso en proyectos de Humanidades Digitales es bastante extendido¹⁹, por ejemplo para actividades de transcripción colaborativa (Causer y Wallace 2012)²⁰, catálogos de bienes patrimoniales (Chevalier *et al.* 2012)²¹, corpus de textos (Rutherford *et al.* 2018) y todo tipo de colecciones web, en particular a través de su extensión Semantic MediaWiki²², que añade funcionalidades para el manejo de datos estructurados, el soporte para SQL y SPARQL y el control granular de permisos.

Drupal²³ es una plataforma software nacida en 2001: modular, multipropósito y altamente configurable. Es el tercer CMS más popular²⁴, con una cuota de mercado del 4,7%, después de WordPress, que ocupa casi el 60%, y de Joomla (6,7%). Con respecto a sus competidores, de los que se diferencia por una más pronunciada curva de aprendizaje y una mayor flexibilidad, se considera más bien un CMF (*Content Management Framework*), es decir un entorno para el desarrollo de aplicaciones web personalizadas. En las últimas versiones²⁵ integra por defecto funcionalidades para el manejo y la consulta de datos estructurados. Sus puntos de fuerza son varios: el alto número de extensiones disponibles (módulos²⁶), que en algunos casos cubren necesidades muy relevantes en el ámbito de las Humanidades Digitales²⁷; la posibilidad de combinar e integrar módulos para la creación de funcionalidades adicionales; sus sólidas capacidades de mapeo y geocodificación²⁸; una amplísima comunidad de desarrolladores y usuarios. Drupal es actualmente utilizado por varias importantes universidades, grandes ONG's, sitios web gubernamentales y de alto valor informativo²⁹. Además, su uso ha sido muy extendido en proyectos de Humanidades Digitales (Dombrowski 2016a;

¹⁸ <https://www.mediawiki.org/wiki/MediaWiki>.

¹⁹ Para una panorámica de proyectos en Humanidades Digitales basados en MediaWiki véase https://www.semantic-mediawiki.org/wiki/Projects_in_eHumanities_running_Semantic_MediaWiki.

²⁰ El *Transcription Desk* del *Bentham Project*:

http://transcribe-bentham.ucl.ac.uk/td/Transcribe_Bentham.

²¹ *Corpus architecturae religiosae europeae (IV-X saec.) [CARE]*: <https://care.huma-num.fr/care/index.php?title=Accueil>.

²² https://www.semantic-mediawiki.org/wiki/Semantic_MediaWiki.

²³ <https://www.drupal.org/>.

²⁴ Open-Source CMS market share: <https://www.opensourcecms.com/cms-market-share/>.

²⁵ A partir de la 7 incorpora en el *core* el módulo *Content Creation Kit (CCK)* para la creación de modelos de datos personalizados; a partir de las 8 incorpora el módulo *Views*, para la configuración asistida de vistas dinámicas en SQL.

²⁶ Actualmente existen 46.483 módulos disponibles de los cuales cerca de la mitad se encuentra en mantenimiento activo y es compatible con versiones superiores a la 7 (https://www.drupal.org/project/project_module).

²⁷ Por ejemplos, módulos cómo: *Features*, *Feeds*, *TEICHI*, *Partial Date*, *Autocomplete Deluxe*, *Views data export*, etc.

²⁸ <https://groups.drupal.org/mapping>.

²⁹ Entre muchas otras: Harvard, Stanford, Oxford, UNAM, UGR; GreenPeace, la Cruz Roja, Oxfam, Amnesty; gobiernos de EEUU, Argentina, Bélgica, Francia, Nueva Zelanda, España, Reino Unido; Fox News, The Weather Channel, NBC, France24.

2016b), sobre todo en su versión 7, lanzada en 2011 y cuyo soporte oficial ha sido recientemente extendido hasta finales de 2022. El lanzamiento de su versión 8, en 2015, ha introducido unos cambios relevantes en la arquitectura de base³⁰, orientados a cubrir sobre todo las necesidades de grandes empresas, lo que ha incrementado el coste de desarrollo y los requisitos de alojamiento web, en particular con respecto al uso de memoria volátil lado servidor. En general, esos cambios han frenado el grado de adopción de Drupal 8³¹ (figura 3) y, en especial, su uso en proyectos de Humanidades Digitales³², y han supuesto el nacimiento, en enero de 2015, de un *fork*³³ derivado: BackdropCMS. La versión 9, que mantiene una marcada continuidad con la versión 8, ha sido lanzada el pasado 3 de junio de 2020, pero su uso es aún marginal.

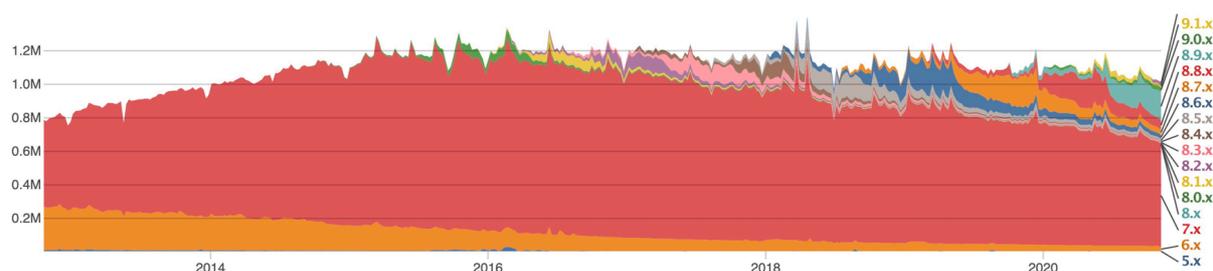


Figura 3. Grado de adopción y uso de las diferentes versiones de Drupal, desde la 5 hasta la 9.1. Se puede apreciar como el uso de la versión 7, con el 66% del total, sigue siendo prevalente. Fuente: <https://www.drupal.org/project/usage/drupal>.

Como hemos señalado, **BackdropCMS** nace como un *fork* de Drupal 7, del cual ha mejorado muchas de las deficiencias bajo los principios de personalización, asequibilidad, rapidez, retro-compatibilidad y extensibilidad³⁴. BackdropCMS, en la hoja de ruta de su desarrollo, pone mucho énfasis en la experiencia a nivel editorial (flujo de trabajo) y en la flexibilidad de la arquitectura de la información, lo que, unido con la herencia de Drupal 7 en cuanto a modelación de datos y creación de vistas dinámicas, lo hace particularmente atractivo en el contexto de las Humanidades Digitales y para el desarrollo de SIW personalizados³⁵. Actualmente, su mayor desventaja es el nivel

³⁰ En particular, la adopción del framework *Symfony*, del sistema de gestión de paquetes *PHP Composer* y del lenguaje de programación orientado a objetos, en lugar de la programación por procedimientos.

³¹ La versión 7 es utilizada por el 66% de todos los sitios web desarrollados en Drupal, mientras que la versión 8, después de seis años, ha alcanzado sólo el 28% del total (<https://w3techs.com/technologies/details/cm-drupal>).

³² <https://digitalhumanities.stanford.edu/enterprise-tools-and-dh>.

³³ En ingeniería de software se considera una bifurcación (en inglés *fork*) al desarrollo de un proyecto informático tomando como base un código fuente que ya existe o a la ramificación de un proyecto madre en varios proyectos que son independientes entre sí y que cuentan con objetivos o desarrolladores diferentes [Fuente: Wikipedia].

³⁴ <https://backdropcms.org/philosophy>.

³⁵ <https://forhumanists.tamu.edu/backdrop>.

de adopción³⁶, aún escaso si comparado con Drupal, y la baja disponibilidad de módulos³⁷.

Un fenómeno interesante en el contexto de Drupal son las distribuciones³⁸: versiones pre-configuradas que unen un conjunto de módulos para satisfacer las necesidades de un ámbito de aplicación específico, algunos de los cuales muy relevantes para nuestro contexto. Se trata, en cierto modo, de un nuevo software que utiliza Drupal como componente de base. Dos de las distribuciones más interesantes para proyectos en patrimonio cultural y Humanidades Digitales son Mukurtu³⁹ y WissKI⁴⁰.

Mukurtu es un sistema de gestión de contenidos y una herramienta para proporcionar acceso digital al patrimonio cultural, co-diseñada para satisfacer las particulares necesidades de las comunidades indígenas para gestionar, dar a conocer y compartir su patrimonio a nivel digital, de manera culturalmente relevante y ética. Ha sido desarrollada y es mantenida por el *Center for Digital Scholarship and Curation* de la *Washington State University* como una plataforma FOSS. Mukurtu está orientada en particular hacia la curación y el uso de metadatos para describir y etiquetar correctamente el patrimonio indígena de una comunidad, proporcionando un acceso diferencial a la información para los miembros de la comunidad y para el público en general, a través de los que denomina “protocolos culturales”⁴¹ (Wilberg 2014). Su plan de desarrollo futuro, que es público y ha sido recientemente actualizado⁴², conlleva el pasaje a Drupal 8 con la versión 4 de Mukurtu, prevista para finales de 2021, y un periodo de migración de un año para que los usuarios puedan migrar a esa nueva versión antes que finalice el soporte oficial a Drupal 7, a finales de 2022.

WissKI es definido por sus creadores como un entorno virtual de investigación basado en la web (VRE) y un sistema de gestión de contenidos (CMS) que, a pesar de ser multiusuario, mantiene un enfoque explícito en el ámbito de las Humanidades Digitales. Ha sido desarrollado por tres instituciones científicas, concretamente el Grupo de Investigación de Humanidades Digitales del Departamento de Ciencias Informáticas de la Universidad Friedrich Alexander de Erlangen-Nuremberg (FAU), el Departamento de Informática de Museos del Germanisches Nationalmuseum (GNM), y el Grupo de Informática de la Biodiversidad del Zoologisches Forschungsmuseum Alexander Koenig (ZFMK). Implementa conceptos de web semántica y se orienta claramente hacia

³⁶ BackdropCMS cuenta actualmente con una estadística de uso de 1.464 páginas web (<https://backdropcms.org/project/usage/backdrop>).

³⁷ Actualmente 533 (<https://backdropcms.org/modules>), aunque más de 70 módulos muy populares han sido incorporado en el *core* (<https://backdropcms.org/upgrade-from-drupal/features-added-core>) y casi la totalidad de los 100 módulos más utilizados en Drupal ya han sido migrados (<https://backdropcms.org/upgrade-from-drupal/top-100-drupal-7-modules>).

³⁸ Existen en la actualidad 1380 distribuciones oficiales https://www.drupal.org/project/project_distribution.

³⁹ <https://mukurtu.org/>.

⁴⁰ <http://wiss-ki.eu/>.

⁴¹ <https://mukurtu.org/support/what-are-communities-cultural-protocols-and-categories/>.

⁴² <https://us12.campaign-archive.com/?u=0b0b343dc81346ec97cfc51d9&id=acbdafa77c>.

proyectos científicos que recogen, almacenan, gestionan y comunican conocimiento relacionado con resultados de investigación, con la mirada puesta en la interoperabilidad a largo plazo (Scholz y Goerz 2012). Incorpora la ontología Erlangen CRM⁴³, que es una implementación OWL-DL 1.0 (Web Ontology Language) de CIDOC-CRM⁴⁴, pero soporta la importación de cualquier otra ontología. Como almacén de RDF⁴⁵ utiliza el *triplestore* ARC14, con un punto SPARQL para la consulta de datos.

Desde la perspectiva de las Humanidades Digitales, **Omeka** se considera una de las principales alternativas a Drupal y sus distribuciones, pero como veremos a continuación su estructura está diseñada sobre todo para gestionar colecciones digitales de objetos, enriquecidos con metadatos. Ha sido desarrollada a partir de 2007 por el *Roy Rosenzweig Center for History and New Media* de la Universidad George Mason⁴⁶, como una solución software para la catalogación y presentación de objetos culturales procedentes de bibliotecas, museos, archivos, colecciones y exposiciones virtuales (Alcaraz Martínez 2012). Se encuentra actualmente disponible en dos diferentes versiones: Omeka Classic⁴⁷ y Omeka S⁴⁸. La primera, orientada hacia proyectos individuales y el sector educativo, está basada en el estándar de metadatos Dublin Core y sus funcionalidades pueden ampliarse a través de *plugins*. Las posibilidades de extensión son mucho más limitadas que en Drupal, pero existen módulos muy relevantes para nuestro contexto de interés, tales como *Neatline*, para incorporar mapas y líneas temporales, y *Scripto*, una herramienta para las transcripciones colaborativas. Omeka S, cuya versión 1.0 fue publicada a finales de 2017, está orientada hacia un ámbito más institucional, como un recurso centralizado desde que desarrollar múltiples plataformas. Además, no está limitada a Dublin Core, sino que puede implementar múltiples vocabularios/ontologías y facilita la vinculación de datos abiertos enlazados. En ambas versiones, la arquitectura del modelo de datos basada en estándares de metadatos constituye la principal diferencia con Drupal, donde el modelo de datos es completamente personalizable. Además de almacenar los ítems directamente en la base de datos de Omeka, existe la posibilidad de importar objetos digitales (archivos XML, como TEI o EAD, imágenes o videos) desde un repositorio Fedora Commons⁴⁹ como datos enlazados, a través del módulo FedoraConnector⁵⁰ (Ethan Gruber 2020).

En esta panorámica, merece una mención el software **AlchemyCMS**⁵¹. Se trata del sistema de gestión de contenidos utilizado por *Europeana Virtual Exhibitions*⁵². A nivel informático, es una plataforma *headless*, es decir que proporciona solo un *back-end*

⁴³ <http://erlangen-crm.org/>.

⁴⁴ Es la norma internacional para el intercambio controlado de información sobre el patrimonio cultural (<http://www.cidoc-crm.org/>).

⁴⁵ RDF es un modelo estándar para el intercambio de datos en la Web.

⁴⁶ <https://rrchnm.org/>.

⁴⁷ <https://omeka.org/classic/>.

⁴⁸ <https://omeka.org/s/>.

⁴⁹ Acrónimo de *Flexible Extensible Digital Object Repository Architecture*.

⁵⁰ <https://omeka.org/s/modules/FedoraConnector/>.

⁵¹ <https://alchemy-cms.com/>.

⁵² <https://github.com/europeana/europeana-virtual-exhibitions>.

construido como un repositorio de contenido estructurado, que suministra a través de una RESTful API. Por su propia arquitectura, se centra en almacenar contenido puro, sin lenguaje de marcado, estilo o formato. No proporciona una plataforma lista para usarse, pero pone a disposición los componentes necesarios para desarrollar un CMS a medida. Integra la tecnología ImageMagick⁵³ y por eso es particularmente apta para la creación de sistemas basados en imágenes. El proyecto empezó como un aplicativo propietario en 2007, bajo el nombre washAPP, y se convirtió en una iniciativa de código abierto en 2010⁵⁴. Ofrece una documentación muy amplia y estructurada y cuenta con una sólida y activa base de desarrolladores⁵⁵.

Pasando a la categoría de sistemas de gestión de activos digitales, **Fedora Commons**⁵⁶ representa una solución software muy consolidada, orientada a la creación de colecciones digitales y a la preservación de recursos a largo plazo. Su desarrollo se inicia a principios de la década del 2000 por las universidades de Cornell y Virginia y, como otros sistemas analizados aquí, es extensible y modular. Además de almacenar todo tipo de contenido y metadatos, también tiene la capacidad de acceder a los datos a través de RESTful APIs y proporcionar búsqueda nativa RDF a través de SPARQL. Su público objetivo es más bien institucional, en particular bibliotecas y archivos que quieran preservar y poner a disposición, internamente y externamente, colecciones grandes y complejas de materiales históricos y culturales. También se utiliza para ofrecer acceso especializado a datos científicos proporcionados por parte de organizaciones académicas y de patrimonio cultural, universidades, instituciones de investigación y organismos gubernamentales (Lagoze *et al.* 2006).

Desde la combinación de Fedora y Drupal nace **Islandora**⁵⁷, un *framework* FOSS diseñado para ayudar instituciones y organizaciones a gestionar y localizar colecciones de objetos digitales de manera colaborativa con su público. Originalmente desarrollado por la *University of Prince Edward Island's Robertson Library*, es ahora mantenido a través de una fundación internacional⁵⁸. Utiliza Fedora como repositorio para el almacenamiento de los objetos digitales, el control de acceso y el *versioning*; Drupal para la creación de la interfaz web, la interacción con los usuarios, la creación de los metadatos, las taxonomías y la expansibilidad a través de módulos⁵⁹. Islandora añade a la combinación entre Fedora y Drupal toda una serie de módulos específicos para la configuración de la plataforma y la interoperabilidad con servicios externos (microservicios y servicios de almacenamiento). Puede almacenar y visualizar una muy amplia variedad de objetos: imágenes, audios, videos, PDFs, colecciones, contenido paginado (libros,

⁵³ <https://imagemagick.org/index.php>.

⁵⁴ <https://dev.to/alchemycms/the-past-present-and-future-of-alchemy-cms-1f6f>.

⁵⁵ https://github.com/AlchemyCMS/alchemy_cms.

⁵⁶ <https://duraspace.org/fedora/>.

⁵⁷ <https://islandora.ca/>.

⁵⁸ <https://islandora.ca/if>.

⁵⁹ El motor de indexación y búsqueda Solr (<https://lucene.apache.org/solr/>), a pesar de ser presentado como una componente de base en la misma página del proyecto, sigue siendo un modulo opcional.

periódicos, seriales) y archivos binarios (Ruest y Stapelfeldt 2014; Stapelfeldt y Moses 2013).

AtoM es la abreviación de *Access to Memory*. Se trata de una aplicación de código abierto creada para la descripción y catalogación de fondos de archivos, basada en normas estándares. Se construyó originalmente con el apoyo del Consejo Internacional de Archivos⁶⁰, para fomentar la adopción de estándares internacionales. Como otras aplicaciones en esta reseña, está pensada para un público institucional y soporta instalaciones multirepositorios, donde pueden contribuir varios organismos. Pone mucho énfasis en las herramientas de importación y exportación de datos e incorpora varios estándares de intercambio de metadatos (EAD, EAC-CPF, CSV y SKOS). Proporciona un buen soporte multilingüaje, incluyendo también idiomas en alfabetos no latinos. Está focalizada en la descripción de objetos digitales según plantillas estándares y se puede integrar, para el almacenamiento, con otras plataformas como DSpace, ArchivesSpace, Fedora, Islandora o contentDM.

Volviendo al campo más específico de la gestión del patrimonio cultural, una solución muy completa es **Arches**⁶¹, desarrollada conjuntamente por el *Getty Conservation Institute* (GCI) y el *World Monuments Fund* (WMF), con una amplia participación de profesionales del patrimonio a nivel internacional. Está pensada para un uso prevalentemente institucional, con el fin de crear inventarios digitales que describan tipo, ubicación, extensión, períodos culturales, materiales y condiciones de los recursos patrimoniales y permitan establecer relaciones entre ellos. La plataforma Arches incluye: un sistema para gestionar, definir y estructurar los datos; varias herramientas de consulta y visualización para buscar, señalar y representar elementos (también a nivel espacial, temporal y *fuzzy*); una herramienta de gestión de proyectos y tareas, para organizar la edición de datos, estableciendo flujos de trabajo personalizados. Soporta el uso de múltiples ontologías, a partir de CIDOC-CRM, y la definición de vocabularios controlados (Carlisle *et al.* 2014). Además, sigue los estándares del Open Geospatial Consortium (OGC)⁶², lo que significa que la plataforma es compatible con las aplicaciones GIS de escritorio. Desde la versión 4, ofrece la aplicación móvil “Arches Collector”⁶³, destinada a proyectos de recopilación y edición de datos en campo. Recientemente, las capacidades básicas y las tecnologías subyacentes de Arches se han ampliado para abarcar una más amplia gama de casos de uso⁶⁴, por ejemplo, a través de la iniciativa “Arches for Science”, cuya versión 1.0 está prevista para el verano de 2021, con el propósito de soportar la comunidad de investigadores en el campo de la ciencia y conservación del patrimonio (*Heritage Science*).

⁶⁰ <https://www.ica.org/es>.

⁶¹ <https://www.archesproject.org/>.

⁶² <https://www.ogc.org/standards>.

⁶³ <http://archesproject.org/collector>.

⁶⁴ Otros casos de uso documentados son la cartografía del patrimonio en riesgo, la vigilancia de los efectos del aumento del nivel del mar en los recursos costeros, la investigación arqueológica, la gestión y publicación de información sobre la procedencia de los objetos de arte, etc.

Otras soluciones web de código abierto utilizadas en el campo de las Humanidades Digitales y del patrimonio cultural, lo suficientemente flexibles como para implementar esquemas de metadatos relacionales y específicos para cada proyecto, incluyendo vocabularios normalizados, son Collective Access⁶⁵, ResearchSpace⁶⁶ y nodegoat⁶⁷.

El software **Collective Access** empezó a gestarse en 2003 bajo el nombre OpenCollection, desarrollado por la compañía Whirl-i-Gig⁶⁸, que pero lanzó su primera versión pública sólo en 2007 (Alcaraz Martínez 2014). Está pensado para administrar y publicar en la Web grandes y heterogéneas colecciones de bibliotecas, museos y archivos, que requieren soporte para diferentes estándares de metadatos (Dublin Core, PBCore y VRA Core vienen integrados), tesauros de referencia (*Library of Congress subject headings*, Getty, GeoNames), formatos de medios (imágenes, video, audio, documentos multi-páginas, escaneos 3D) y una variedad de opciones de visualización (zoom tesselado para imágenes en alta resolución, anotaciones, líneas dinámicas del tiempo, mapas, etc.). A nivel de infraestructura informática, Collective Access se basa en dos componentes principales: Providence⁶⁹, cómo interfaz de catalogación *back end*, y Pawtucket⁷⁰, cómo plataforma de consulta y publicación *front end*.

ResearchSpace es un entorno virtual de investigación (VRE) desarrollado por el British Museum en colaboración con la empresa alemana Metaphacts⁷¹ y se encuentra actualmente en fase de *pre-release*⁷². Está focalizado en modelos de datos basado en grafos y emplea tecnologías de la Web semántica para apoyar la investigación en patrimonio cultural. Pone el énfasis en la representación gráfica del conocimiento, más que en la catalogación, y promueve la investigación colaborativa. En palabras de sus promotores: "The project's long-term goal is to build a community of researchers that open and share their data, knowledge, research practices, and arguments with each other" (Oldman y Tanase 2018). Proporciona un entorno integrado de herramientas diseñadas para reflejar los métodos de investigación, incluyendo búsqueda semántica basada en CIDOC-CRM, portapapeles y anotaciones semánticas, herramientas de manipulación y anotación de imágenes utilizando IIIF⁷³ y mapas del conocimiento. A pesar de encontrarse aún en un estado temprano de desarrollo, se ha documentado su uso en varias disciplinas⁷⁴, por ejemplo: en Arqueología, el proyecto Gravitate⁷⁵, focalizado

⁶⁵ <https://collectiveaccess.org/>.

⁶⁶ <https://www.researchspace.org/>.

⁶⁷ <https://nodegoat.net/>.

⁶⁸ <http://www.whirl-i-gig.com/>.

⁶⁹ <https://github.com/collectiveaccess/providence/releases>.

⁷⁰ <https://github.com/collectiveaccess/pawtucket2/releases>.

⁷¹ <https://metaphacts.com/>.

⁷² <https://github.com/researchspace/researchspace/releases>.

⁷³ El *International Image Interoperability Framework* (IIIF) define varias interfaces de programación de aplicaciones que proporcionan un método normalizado de descripción y suministro de imágenes a través de la Web, así como "metadatos basados en la presentación", sobre secuencias estructuradas de imágenes [Fuente: Wikipedia].

⁷⁴ <https://www.researchspace.org/projects.html>.

⁷⁵ <http://gravitate-project.eu/>.

en ofrecer una solución informática para establecer relaciones, culturales o físicas, entre artefactos patrimoniales; en Historia del Arte, el proyecto Late Hokusai⁷⁶; en Historia de la Ciencia, el proyecto CorpusTracer⁷⁷ (Kräutli y Valleriani 2018).

Nodegoat comparte con ResearchSpace el enfoque orientado a la visualización dinámica de la información y a la generación de conocimiento. Se trata de un entorno virtual de investigación desarrollado desde 2011 por LAB1100⁷⁸, una spin-off de la Universidad de Ámsterdam. Incluye herramientas de gestión de datos, de análisis de redes y de visualización gráfica. Nodegoat está directamente pensado para la investigación, más que para la clasificación y preservación de contenido. Permite a los usuarios co-crear conjuntos de datos basados en modelos diseñados a medida. En el mismo entorno de gestión, ofrece módulos para analizar y visualizar esos conjuntos de datos de forma relacional, diacrónica y espacial (Bree y Kessels 2013). Ha sido utilizado en varias disciplinas humanísticas para una amplia pluralidad de casos de uso⁷⁹, en particular en proyectos donde las preguntas de investigación estaban muy ligadas a la combinación de las variables de espacio y tiempo. Una versión gratuita, pero limitada, de nodegoat está disponible en modalidad Software como un Servicio (ScuS o SaaS por el acrónimo inglés de *Software as a Service*)⁸⁰, mientras que las otras opciones de uso prevén la suscripción de un contrato de servicio anual con LAB1100. Existe la posibilidad de descarga libre de todo el código fuente⁸¹, pero más bien con el fin de garantizar la transparencia del proyecto que para un uso independiente.

Entre las disciplinas humanísticas objeto de la presente tesis doctoral, sin duda la Arqueología es la que, por la mole de datos producidos por la propia actividad arqueológica de campo, sea esa de excavación o de prospección, posee la más larga trayectoria en el uso de sistemas informáticos para la gestión de la información (Martlew 1984; Macchi Jánica 1999; Di Giacomo y Scardozzi 2014; Bertoldi *et al.* 2015). Merece la pena señalar al menos dos soluciones informáticas desarrolladas para la gestión y presentación de datos en esa disciplina: **Archaeological Recording Kit (ARK)**⁸² y **Open Context**⁸³. ARK es un paquete software de código abierto que incluye herramientas de edición, creación, visualización e intercambio de datos. Su relevancia viene principalmente por ser la tecnología utilizada en la plataforma Fasti Online⁸⁴, una base de datos en línea de excavaciones arqueológicas, prospecciones y proyectos de conservación arqueológica, muy relevante en particular para el contexto italiano. A pesar de esto, hemos decidido no incluirla en esta reseña porque la última versión estable (v1.1⁸⁵) ha

⁷⁶ <https://www.latehokusai.org/>.

⁷⁷ <https://sphaera.mpiwg-berlin.mpg.de/>.

⁷⁸ <https://lab1100.com/>.

⁷⁹ <https://nodegoat.net/usecases>.

⁸⁰ <https://nodegoat.net/products>.

⁸¹ <https://github.com/nodegoat/nodegoat>.

⁸² <https://ark.lparchaeology.com/>.

⁸³ <https://opencontext.org/>.

⁸⁴ <http://www.fastionline.org/>. Fasti Online es un proyecto de la *International Association of Classical Archaeology* (AIAC) y del *Center for the Study of Ancient Italy of the University of Texas at Austin* (CSAI).

⁸⁵ <https://sourceforge.net/projects/arkdb/files/>.

sido publicada en 2014 y no es compatible con las actuales versiones de PHP. La versión 2.0 se encuentra en desarrollo desde hace tiempo, pero el repositorio no presenta cambios en los últimos tres años⁸⁶. Open Context es una plataforma en línea de gestión de datos de investigación y servicios de publicación. Existe únicamente como “Software como un Servicio”, con un modelo de negocio basado en la publicación y almacenamiento de datos⁸⁷. Al no ser un software FOSS, queda también fuera de los criterios de selección de la presente reseña.

El último grupo de software que pasamos a analizar reúne soluciones informáticas orientadas hacia la creación de plataformas para el *crowdsourcing* y la Ciencia ciudadana. En particular, hemos seleccionado cuatro: Zooniverse⁸⁸, Pybossa⁸⁹, EnrichEuropeana⁹⁰ y CrowdHeritage⁹¹.

Zooniverse es un portal nacido en 2009 desde el proyecto colaborativo Galaxy Zoo⁹². No se limita al ámbito de las humanidades, sino que abraza proyectos desde cualquier ámbito de la investigación científica. Está basado en la colaboración abierta entre pares y ofrece la posibilidad de crear proyectos para la recolección y manipulación de datos involucrando la ciudadanía. Cuenta con una muy amplia comunidad de usuarios, que en 2019 llegaba a 1,6 millones de voluntarios⁹³. A pesar de publicar buena parte de su código en abierto⁹⁴, está pensado para usarse como un servicio a través de su plataforma web. Ofrece herramientas *ad hoc* para asignar tareas a los participantes, así como la posibilidad de cargar conjuntos de datos (normalmente imágenes) sobre los que se pide a los voluntarios que realicen operaciones; la estructura de su modelo de datos no se puede adaptar a casos específicos (Simpson *et al.* 2014).

Pybossa⁹⁵ es un *framework* de código abierto para el desarrollo de plataformas de recogida, análisis y enriquecimiento de datos en entornos colaborativos, lanzado en 2013 por Scifabric⁹⁶, una compañía con sede en España que sigue ocupándose de su desarrollo. Ofrece funcionalidades de transcripción de documentos manuscritos; transcripción de vídeo y audio; análisis de objetos multimedia; geolocalización de archivos; identificación y etiquetado de objetos, caras, sonidos, etc., en imágenes, audio y vídeo; identificación de secciones de obras digitalizadas, tales como animales, personas, análisis de sentimientos, etc.; enriquecimiento de registros y archivos con metadatos. Su ámbito de aplicación no se limita a proyectos de Ciencia ciudadana, pero muchos de sus casos de uso proceden de las humanidades. En particular, hay que señalar que la

⁸⁶ <https://github.com/lparchaeology/ark2>.

⁸⁷ <https://opencontext.org/about/estimate>.

⁸⁸ <http://www.zooniverse.org/>.

⁸⁹ <https://pybossa.com/>.

⁹⁰ <https://pro.europeana.eu/project/enrich-europeana>.

⁹¹ <https://pro.europeana.eu/project/crowd-heritage> y <https://crowdheritage.eu/en>.

⁹² https://es.wikipedia.org/wiki/Galaxy_Zoo.

⁹³ <https://es.wikipedia.org/wiki/Zooniverse>.

⁹⁴ <https://github.com/zooniverse>.

⁹⁵ <http://pybossa.com/>.

⁹⁶ <https://scifabric.com>.

tecnología ofrecida por Pybossa constituye la infraestructura base de la iniciativa Comunidad BNE⁹⁷, promovida en 2019 por la Biblioteca Nacional de España en colaboración con Red.es⁹⁸ (Sánchez Nogales 2019). Comunidad BNE es una plataforma de trabajo colaborativo, abierto y distribuido, sobre las colecciones digitales y datos de la Biblioteca Nacional de España, similar a Zooniverse, pero con dos diferencias importantes: los investigadores no pueden aportar sus propios datos; se pueden sugerir nuevos proyectos, pero no crearlos directamente. A los usuarios se le pueden asignar tareas de identificación de objetos, transcripción de textos y audios, georreferenciación, marcado y etiquetado, enriquecimiento de metadatos y corrección OCR. BNElab⁹⁹, que está a cargo de su desarrollo, ha añadido nuevas funcionalidades a la tecnología de base que, a su vez, ha puesto a disposición para su libre reutilización¹⁰⁰.

EnrichEuropeana es el nombre de un paquete software desarrollado en el marco del homónimo proyecto, financiado por Europeana y la Comisión Europea entre 2018 y 2019 con el fin de renovar la infraestructura tecnológica de la plataforma Transcribathon¹⁰¹ (Sciotti 2019). Como otras plataformas para el *crowdsourcing*, ofrece funcionalidades para transcribir, anotar y georreferenciar documentos históricos, pero tiene un ámbito de aplicación más restringido que las otras descritas anteriormente. A nivel tecnológico, está compuesta por tres módulos¹⁰²: una aplicación *front end* compuesta por una versión personalizada de WordPress y un visor IIF; una API para el enriquecimiento semántico; una infraestructura de intercambio de datos para la transferencia de información entre Europeana y la plataforma Transcribathon.

La plataforma **CrowdHeritage** es la última iniciativa en orden de tiempo promovida por Europeana en este campo. Fue creada con el fin de utilizar el *crowdsourcing* para mejorar la calidad de los metadatos del patrimonio cultural digital indexado en Europeana, así como en otras bases de datos de instituciones dedicadas al patrimonio que cuenten con APIs públicas. Ha sido desarrollada por la Universidad Nacional Técnica de Atenas, en colaboración con la Asociación Europea *Fashion Heritage*, la Asociación *MICHAEL Culture*, el ministerio francés de la cultura y la Fundación Europeana. A diferencia de otras soluciones informáticas en este ámbito, el tipo de herramientas que pone a disposición es algo más restringido, focalizándose en funcionalidades de anotación del contenido, marcado semántico a través de tesauros y vocabularios, análisis de colores y geo-tagging colaborativo. La participación de los usuarios se promueve a través de la gamificación, utilizando una tabla de clasificación y asignando insignias. En principio, todo el código está disponible en abierto¹⁰³ y puede ser

⁹⁷ <https://comunidad.bne.es/>.

⁹⁸ <https://red.es/redes/>.

⁹⁹ <https://bnelab.bne.es/>.

¹⁰⁰ <https://github.com/BNELab/comunidadbne>.

¹⁰¹ <https://europeana.transcribathon.eu/>; la nueva plataforma ha sido inaugurada el 24 de septiembre 2019.

¹⁰² <https://github.com/EnrichEuropeana>.

¹⁰³ <https://github.com/ails-lab/crowdheritage>.

utilizado para implementar plataformas similares, aunque no sea ese el objetivo de la iniciativa.

Una vez realizada la descripción general de las diferentes aplicaciones seleccionadas, en la tabla 2 mostramos una comparativa de sus principales características. Hemos elegido una serie de funcionalidades, entre más básicas y más avanzadas, que suelen ser necesarias para la creación de sistemas de información web y que los creadores/usuarios de estos sistemas suelen requerir.

| Software | Formatos básicos* | Formatos avanzados** | Mapas | Multilinguaje | Modelo de datos personalizable | Estándares para metadatos | Datos enlazados | Búsqueda básica | Búsqueda avanzada | Importación / exportación | Extensible |
|-------------------|-------------------|----------------------|-------|---------------|--------------------------------|---------------------------|-----------------|-----------------|-------------------|---------------------------|------------|
| MediaWiki | X | E | E | X | E | E | E | X | Elasticsearch | X | X |
| Drupal 7 | E | E | E | X | X | L | | X | Solr | E | X |
| Drupal 8/9 | X | E | E | X | X | L | | X | Solr | E | X |
| BackdropCMS | X | | L | X | X | | | X | | X | X |
| WissKI | X | E | E | X*** | X | X | X | X | Solr | X | X |
| Mukurtu | X | E | X | X*** | X | X | X | X | Solr | X | X |
| Omeka | X | E | L | X | L | X | | X | | X | X |
| Omeka S | X | E | L | X | L | X | X | X | Solr | X | X |
| AlchemyCMS | X | | | L | L | | | X | Ferret | | L |
| Fedora | X | X | | | X | X | X | Solr | Solr | X | X |
| Islandora | X | X | X | X*** | X | X | X | Solr | Solr | X | X |
| Arches | X | X | X | X | X | X | X | X | Elasticsearch | X | X |
| AtoM | X | | | X | | X | | X | X | X | X |
| Collective Access | X | X | X | X | X | X | | X | X | X | |
| ResearchSpace | X | | X | | X | X | X | X | X | X | |
| Nodegoat | X | | X | | X | X | X | X | X | X | |
| Zoomuniverse | X | | X | L | | | | X | | X | |
| Pybossa | X | X | X | L | L | | | X | | X | X |
| Enrich Europeana | X | | | | | X | L | X | | | |
| Crowd Heritage | X | | | X | | X | X | X | | | |

Legenda: E = extensión; L = limitado.

* PDFs, imágenes, audios y videos. ** Objetos 3D, colecciones, contenido paginado, binarios, KML, etc.

*** Soporte multilinguaje basado en Drupal, pero no disponible para la interfaz de la propia distribución.

Tabla 2. Comparativa de funcionalidades de las plataformas analizadas. Fuente: elaboración propia.

El riesgo intrínseco en la elaboración de tablas comparativas es exceder en la simplificación a beneficios de la visión de conjunto. En muchos casos, la calidad y la exhaustividad en la implementación de una cierta componente son muy distintas en un software respecto a otro y señalar la simple presencia/ausencia puede resultar reductivo. Para compensar, en parte, esta limitación hemos introducido dos variantes: “L” por limitado, para indicar los casos de una implementación parcial o experimental; “E” por extensión, si la funcionalidad depende de un paquete externo (módulo o *plugin*), lo que de por sí no habla de su calidad, pero sí de su nivel de integración y perspectiva de mantenimiento. A pesar de esos matices, para completar la reseña es necesario añadir una serie de valoraciones más cualitativas.

Una primera distinción necesaria hay que realizarla entre las plataformas agnósticas y aquellas desarrolladas para un ámbito de aplicación específico. Pertenecen claramente al primer grupo: MediaWiki, Drupal, Fedora, Islandora, Collective Access, Alchemy, Zooniverse y Pybossa. Están pensadas, por otro lado, para un ámbito de aplicación más concreto: BackdropCMS (pequeñas empresas y ONG); WissKI, Nodegoat y Research Space (investigación en Humanidades Digitales); Mukurtu (patrimonio cultural de comunidades indígenas); Omeka (colecciones digitales); Arches (inventarios patrimoniales); AtoM (fondos de archivos digitales), EnrichEuropeana y CrowdHeritage (transcripción y etiquetado de colecciones indexadas en Europeana).

Otra variable que afecta significativamente las características de algunas plataformas frente a otras es el público objetivo. Podemos identificar dos tendencias principales: plataformas desarrolladas para grandes empresas y clientes institucionales, como los GLAM, por un lado; soluciones web para proyectos individuales, por otro. El primer grupo (ej. Fedora, Islandora, Collective Access, OmekaS, Arches, AtoM) está orientado hacia la creación de repositorios institucionales y aboga por una clara preferencia hacia modelos de datos basado en estándares para metadatos, la interoperabilidad (Datos enlazados y APIs) y la incorporación de protocolos del sector de referencia (ej. AtoM), frente a la personalización de la estructura y de las funcionalidades. En principio, soluciones con estas características quedan al margen del objeto de estudio de la presente tesis doctoral (véase apartado I.1), pero han sido incorporadas en este análisis comparativo porque en los últimos años han ido incorporando más opciones de configuración a medida, marcando un interesante camino hacia soluciones híbridas. Algunas de ellas en particular, como WissKI e Islandora, son muy prometedoras: permiten mantener un cierto nivel de compatibilidad e interoperabilidad con los estándares más consolidados, incorporando por ejemplo Dublin Core y CIDOC-CRM, y a la vez personalizar el modelo de datos y desarrollar ontologías propias encima de los estándares comunes. Como consecuencia de esta ampliación de funcionalidades, se ha incrementado también el peso y la complejidad del código fuente, así como el número de dependencias externas, y el panorama ofrece ya pocas opciones de fácil instalación y con unos requisitos básicos en cuanto a alojamiento web. Entre estas últimas podemos nombrar: MediaWiki, BackdropCMS, Omeka Classic y Mukurtu.

Con respecto a la trayectoria de desarrollo, solo dos, ResearchSpace y CrowdHeritage, son aplicativos nuevos aparecidos en los últimos cinco años, ya que OmekaS, BackdropCMS y EnrichEuropeana son evoluciones de plataformas preexistentes. En general, son todas soluciones muy maduras, que en los años han ido ampliando el abanico de funcionalidades y aumentando el grado de solape entre ellas. El nivel de adopción es muy variado y difícilmente comparable, sobre todo entre aplicaciones agnósticas y especializadas. La competencia entre ellas por cuotas de mercado ha sido hasta ahora beneficiosa, pero es previsible que con el tiempo alguna acabe afirmándose como solución privilegiada y aparezcan otras nuevas.

Finalmente, es necesario realizar una última distinción entre los proyectos que publican su código fuente en abierto para favorecer su distribución, descarga y uso y aquellos que utilizan el *open source* principalmente para asegurar la transparencia de la iniciativa y la estabilidad del código a largo-plazo. En el primer caso, el proceso de instalación suele ser relativamente simple o muy bien documentado y, si la aplicación necesita de dependencias software externas, suelen ser relativamente comunes y ampliamente utilizadas en entornos de desarrollo web. Entran en este grupo la mayoría de las soluciones analizadas en esta reseña, como: MediaWiki, Drupal, Backdrop, WissKI, Mukurtu, Omeka, Alchemy, Fedora, Islandora, Arches, AtoM y CollectiveAccess. En cambio, al segundo grupo pertenecen aplicaciones que suelen ofrecerse en la modalidad Scus (Software como un Servicio) y que proporcionan apoyo activo a su comunidad de usuarios (ej. Zoomuniverse, EnrichEuropeana y CrowdHeritage) o al desarrollo de los proyectos que albergan (ej. Nodegoat, ResearchSpace y Pybossa).

Después de valorar las características de las soluciones informáticas seleccionadas, podemos concluir que todas las plataformas analizadas tienen sus ventajas y sus desventajas y es necesario adquirir un conocimiento actualizado del panorama disponible en cada momento con el fin de elegir la opción más adecuada para cubrir las necesidades identificadas en cada caso. Al estado actual, Drupal, de por sí o a través de un paquete software que lo incorpora (ej. WissKI, Mukurtu o Islandora), ocupa un nicho muy específico y sigue representando una de las mejores opciones disponibles para nuestro ámbito de aplicación. Su nivel de flexibilidad en cuanto a modelo de datos e incorporación de funcionalidades avanzadas a través de módulos lo hace comparable con soluciones completamente personalizadas, pero a la vez proporciona un *front end* amigable y cuenta con una muy amplia comunidad de soporte. Se adapta ya sea a la gestión de conjuntos de datos estructurados de investigación, que a la creación de repositorios de objetos digitales y a la gestión de contenido, sobre todo en ambientes colaborativos. WissKI en particular, gracias a su *back end* semántico y su larga trayectoria de desarrollo (2012-2021), es una opción muy válida para proyectos de Humanidades Digitales de cierta envergadura. Por ejemplo, en el contexto de la presente tesis, WissKI ha sido la opción elegida para el atlas descrito en el apartado IV.2.2.

Para proyectos más pequeños y con un presupuesto reducido o ausente de cara al mantenimiento, la alternativa de BackdropCMS es muy adecuada y solo queda la duda

sobre su sostenibilidad a largo plazo, considerando el relativamente bajo nivel de adopción hasta la fecha.

Fuera de la galaxia Drupal, los productos más interesantes en el contexto de las Humanidades Digitales son tres. Omeka, que pero mantiene en ambas versiones un enfoque prácticamente exclusivo sobre colecciones digitales con metadatos estándares. Nodogot, por el amplio espectro de funcionalidades pre-configuradas que cubren desde la modelación de datos a la publicación en línea, pasando por la co-creación, la visualización espacio-temporal y el análisis de redes. Pybossa, para proyectos de Ciencia ciudadana que prevean la involucración de la sociedad a través de la asignación de tareas pre-configuradas.

Para un público más institucional, no podemos evitar de poner en evidencia la plataforma Arches, que con las últimas versiones ha ampliado su ámbito de aplicación más allá de la gestión del patrimonio cultural. Presenta una aplicación móvil dedicada para el trabajo de campo, unas funcionalidades muy sólidas en cuanto a georreferenciación y mapeo, el control granular de los permisos de acceso, la gestión integral y flexible del flujo de trabajo y mantiene publica una hoja de ruta bien definida en cuanto a desarrollo para los próximos años¹⁰⁴.

¹⁰⁴ <https://www.archesproject.org/roadmap/>.

III.2. LAS FASES DE TRABAJO

Como hemos comentado anteriormente, el presente trabajo de investigación doctoral se define a partir de un enfoque esencialmente metodológico, en la medida en que el objeto de estudio en sí es el proceso de diseño e implementación de sistemas de información web para el almacenaje, la gestión y la recuperación de datos estructurados, así como las consecuencias que el uso de esos sistemas producen en el proceso mismo de investigación y en función de su apertura al exterior, ya sea hacia otros grupos de investigación que hacia la sociedad. Este enfoque se completa con una perspectiva predominantemente empírica y práctica, que produce sus resultados a partir de diferentes casos de estudio, desarrollados en varios ámbitos disciplinares (Historia, Arqueología e Historia del Arte), y cuyas características e implicaciones se describen y se analizan a través de una serie de publicaciones científicas incorporadas en el capítulo de resultados.

Para alcanzar los objetivos descritos anteriormente (véase capítulo II), se ha perfilado una metodología de trabajo en diferentes fases que, a partir del ciclo clásico de vida, uso y mantenimiento de un sistema de información (Senn 1992; Schach 1996), lo amplía en función de los objetivos planteados. En concreto, hemos identificado e implementado las siguientes etapas:

1. planificación;
2. conocimiento del dominio;
3. identificación de conjuntos de datos preexistentes;
4. diseño del modelo de datos;
5. creación de las taxonomías;
6. definición de las funcionalidades y del flujo de datos;
7. implementación informática;
8. pruebas;
9. despliegue;
10. entrada, enriquecimiento y verificación de datos;
11. mantenimiento.

Este conjunto de fases ha dado lugar a un protocolo de actuación que ha funcionado como hoja de ruta en el diseño e implementación de varios SIW. Mientras que las etapas centrales son consustanciales al proceso, otras se han llevado a cabo sólo en casos concretos. Además, en la práctica, algunas de ellas no están perfectamente diferenciadas entre sí y suelen producirse en paralelo. A continuación, nos centraremos en detallar las diferentes etapas y su peso relativo a nivel metodológico, con el objetivo de comunicar la complejidad del proceso en su totalidad.

III.2.1. Planificación

La primera fase, de **planificación** (1), corresponde a la investigación preliminar sobre la viabilidad del proyecto de desarrollo del SIW. Nos hemos focalizado sobre la factibilidad técnica, en términos de recursos humanos involucrados en el proceso, factibilidad económica, en términos de costes de desarrollo y mantenimiento, y factibilidad

temporal, para valorar los plazos necesarios desde la planificación al despliegue de la plataforma.

III.2.2. Conocimiento del dominio

La etapa de **conocimiento del dominio** (2), o ámbito del proyecto, se ha dedicado a examinar el campo disciplinar y el objeto de estudio que el SIW se proponía de organizar. En el marco de referencia de las disciplinas humanísticas involucradas en la presente tesis, para cada caso de estudio, esta fase ha sido necesaria para acotar el tema concreto de investigación, así como el enfoque teórico-metodológico que subyace la investigación, las posibles implicaciones multidisciplinares e interdisciplinares, así como las principales fuentes de datos con las que se iba a trabajar.

III.2.3. Identificación de conjuntos de datos preexistentes

Con respecto a las fuentes, sean ellas artefactos históricos de cualquier índole (ej. documentos escritos, obras de artes, objetos de la vida doméstica, etc.), muestras (ej. láminas delgadas) o simplemente datos (ej. unidades estratigráficas), ha sido necesario en cada caso evaluar la existencia de **conjuntos de datos preexistentes** (3) al planteamiento del sistema de información. Como veremos en detalle en el apartado IV.4.1, en muchos proyectos la realización de un SIW nace en respuesta a las dificultades que una gestión desestructurada de los datos pone de manifiesto, en particular en fase de procesamiento y de análisis. Cuando efectivamente existían datos preexistentes, ha sido imprescindible valorar a fondo la calidad de los mismos, cómo estaba organizada la información, los formatos de almacenamiento, la existencia y la completitud de los metadatos. En función del volumen de datos, se ha optado por una importación manual o por la creación de *scripts* de importación semi-automática, ya que muy difícilmente el modelo de datos era directamente compatible.

III.2.4. Diseño del modelo de datos

La fase de arquitectura de la información y **elaboración del modelo de datos** (4) representa indudablemente una de las etapas con mayor implicaciones metodológicas y epistemológicas del proceso (Fernández 2006). En ella, hemos tenido que combinar los conocimientos relativos al ámbito de aplicación y métodos de investigación, por un lado, con los conocimientos sobre los criterios de modelación de la información y los principios de estructuración de bases de datos por el otro. Defendemos que un buen modelo de datos es el fruto de un diálogo, una interlocución, entre el investigador, o grupo de investigación, que lo necesita y el experto en modelación de datos y tecnologías de la información que lo desarrolla, en un proceso interactivo e iterativo. No se ha tratado, en ningún caso, de un simple encargo técnico externalizado. Además, en nuestra experiencia, ese diálogo multidisciplinar resulta mucho más fructífero y efectivo si al menos una de las partes controla ambos campos, el humanístico y el informático, es decir si es una persona interdisciplinar. Un diálogo necesita, por definición, de un lenguaje común que de por sí no existe entre profesionales de las TIC y humanistas, ya que trabajan en ámbitos disciplinares que tradicionalmente no comparten conceptos ni métodos. Ese lenguaje hay que construirlo y la presencia en el grupo de trabajo de

perfiles híbridos, especialistas en Humanidades Digitales, resulta ser clave para cubrir ese *gap* de comunicación y competencias recíprocas. La experiencia de este trabajo de investigación lo confirma ampliamente. Así como nos ha confirmado que, si los investigadores en humanidades no colaboran a diseñar y modelar los entornos en los que trabajarán, simplemente no podrán utilizarlos.

El objetivo de esta fase ha sido, por cada caso, la creación de un modelo que permitiera a un investigador dotado del conocimiento sobre el ámbito de investigación en objeto de utilizar el sistema de información. Cada modelo ha tenido que adaptarse y representar la realidad objeto de estudio. Se ha convertido, en este sentido, en una herramienta de investigación del dominio de aplicación para determinar, por ejemplo, la unidad mínima de información que fuese necesario distinguir, almacenar y procesar. A más granularidad de información equivale más capacidad de análisis, más posibilidades de explotación, de generar filtros y de cuantificar; pero implica una entrada de datos más laboriosa y posibles asimetrías en el nivel mínimo de detalle alcanzable en diferentes entidades.

A la vez, cada modelo ha sido una representación simplificada y parcial de la realidad. Como en un dibujo, o en un mapa, no se puede representar el objeto de estudio tal y como es, ni es necesario. La creación de un modelo mira en hacer más entendible la realidad y eso pasa por determinar qué detalles de las entidades a representar se conservan y cuáles se descartan. También una foto es una simplificación de la realidad, pero su objetivo no es simplificar, sino más bien el contrario: representar cuantos más detalles posibles, y por esto no se puede considerar un modelo.

Implícito en esta fase de modelación es el riesgo de inversión de modelado (González-Pérez 1999), que se determina cuando se diseña un modelo que modifica la realidad investigada, en vez que representarla, por motivos ajenos a la propia metodología de investigación. Las principales causas pueden ser la voluntad de trasladar al modelo limitaciones que proceden de cómo los ordenadores almacenan la información o, peor aún, condicionantes propios de la solución software elegida para la implementación del sistema de información (véase apartado III.1).

La elaboración de cada modelo de datos se ha apoyado, en buena medida, sobre la propia metodología de investigación y ha obligado a revisarla e incluso a ponerla en discusión, explicitando los procesos de captación, entrada y consulta de datos, producción de la información y generación del conocimiento. En este sentido, en muchos casos, se ha demostrado ser la fase con mayores consecuencias para el proyecto de investigación en sí, que se ha beneficiado de las reflexiones generadas durante la elaboración del modelo conceptual.

En definitiva, para la elaboración de las diferentes estructuras, hemos tenido que identificar: 1) qué tipo de entidades era necesario representar (las tablas de la base de datos); 2) qué atributos, o características, describían esas entidades (las columnas de cada tabla); 3) las propiedades y dominios de esos atributos, es decir la tipología de valores,

los rangos y las restricciones de integridad (los condicionantes de los valores de cada celda); 4) las relaciones, o interacciones, existentes entre las entidades, así como su cardinalidad.

III.2.5. Creación de las taxonomías

Una parte integrante del proceso de modelado ha sido determinar qué atributos necesitaban de la **creación de taxonomías** personalizadas (5) (Currás 2005; García-Marco 2007). En el contexto de este trabajo, hemos elegido el uso del término taxonomías como el más apropiado para describir los listados jerárquicos de términos utilizados para clasificar la información en unidades conceptuales, y reducir la heterogeneidad semántica, frente a alternativas como "vocabularios controlados", "tesauros" y "ontologías"¹⁰⁵. El objetivo de esta fase ha sido la creación de listados de categorías que describieran atributos relevantes para la investigación. El principio que ha regido este proceso ha sido que las definiciones fueran tan específicas como fuera posible, para así elevar el grado de atomización de la información y poder responder a preguntas más concretas, y tan generales como fuera necesario, para permitir asignar valores también en los casos de información incompleta o fragmentaria, algo frecuente en las disciplinas que se ocupan del estudio del pasado. Como veremos más adelante, a nivel informático, el uso y la definición de las taxonomías juega además un rol clave en la indexación y recuperación de la información, en particular para los análisis cuantitativos. En la práctica, las fases relativas al modelado conceptual (4) y a la elaboración de taxonomías (5) se han llevado a cabo en paralelo y, según la complejidad de la temática, han necesitado entre dos y seis reuniones de alrededor de dos horas cada una. El resultado final del proceso, en cada caso, ha sido la elaboración de un diagrama entidad-relación (Chen 1976), necesario para poder sucesivamente diseñar el modelo físico de la base de datos relacional¹⁰⁶. Ese esquema visualiza las entidades a representar, cómo se relacionan entre sí y con qué atributos y valores se describen.

Como ejemplo visual de un **esquema entidad-relación**, en la figura 4 presentamos el caso de un proyecto sobre ajueres domésticos y enseres de casa del reino de Granada en el siglo XVIII, a partir principalmente de la documentación notarial, dotes e

¹⁰⁵ Los vocabularios se diferencian por carecer de estructura y de jerarquía y por limitarse al simple control del lenguaje. Los tesauros añaden a las taxonomías los términos relacionados, no preferentes y preferentes, además de relaciones asociativas y de sinonimia, y son apropiados para contextos en los que los usuarios del sistema no controlan el ámbito de aplicación. Por otro lado, en Ciencia de la información, una ontología es un modelo de datos semánticos que se utiliza para describir conceptos en un determinado dominio del conocimiento, sus propiedades y las relaciones entre esos conceptos. Según el W3C (<https://www.w3.org/standards/semanticweb/ontology>), "There is no clear division between what is referred to as 'vocabularies' and 'ontologies'. The trend is to use the word 'ontology' for more complex, and possibly quite formal collection of terms, whereas 'vocabulary' is used when such strict formalism is not necessarily used or only in a very loose sense".

¹⁰⁶ La tipología relacional es la tipología de base de datos utilizada en este trabajo. Existen otras tipologías, como las bases de datos orientadas a objetos, las jerárquicas, las orientadas a grafos, etc. (Silberschatz *et al.* 2019).

inventarios, conservada en el Archivo Histórico de Protocolos de Granada¹⁰⁷. El objetivo de la investigación era conocer cómo se vestía la casa, es decir, que consideraban sus habitantes como necesario y adecuado, y cómo ello contribuía a la configuración de dicho espacio en procesos culturales de adaptación-resistencia, frente a los condicionantes del contexto cronológico de referencia¹⁰⁸.

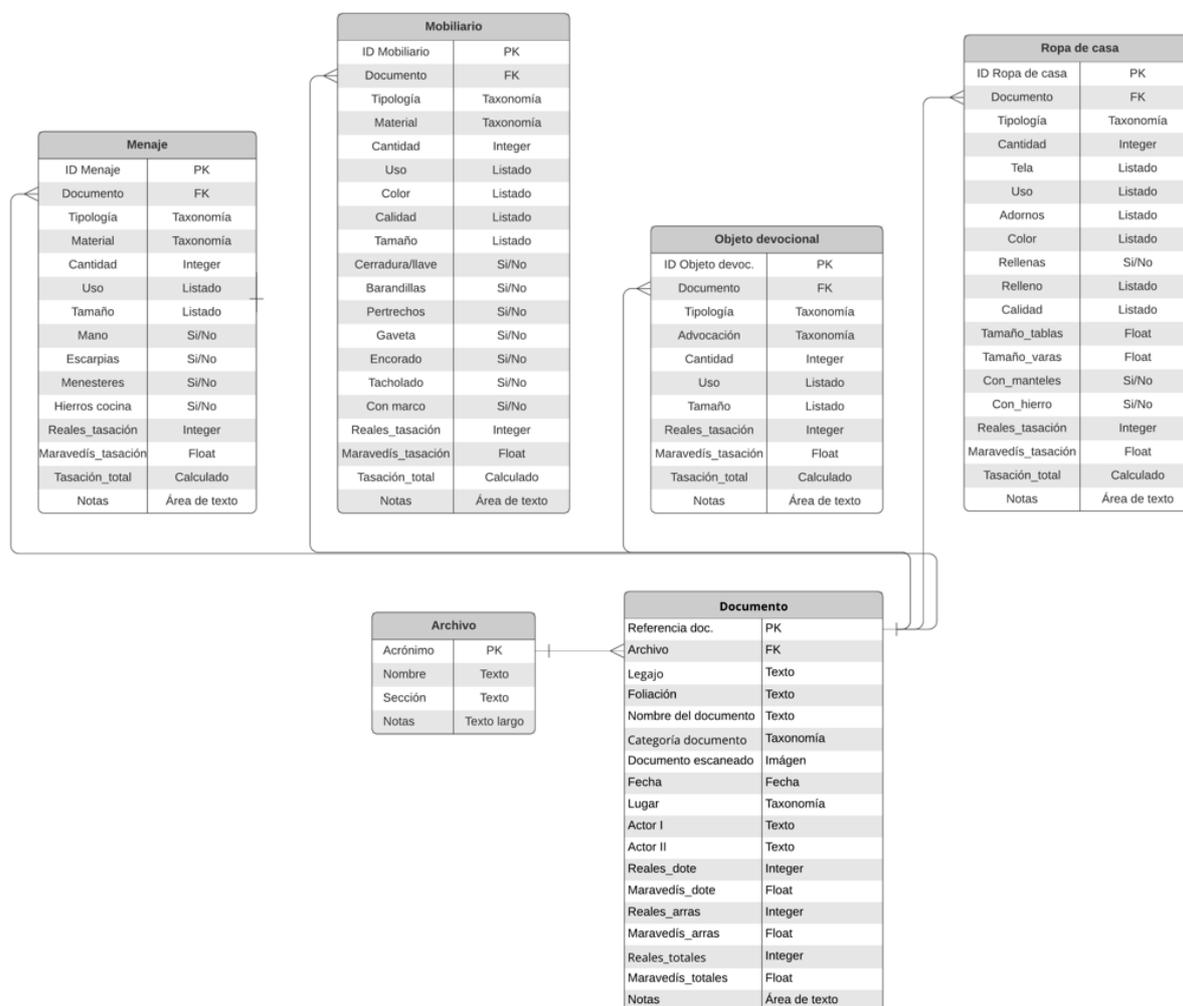


Figura 4. Esquema entidad-relación del modelo de datos sobre ajueres históricos, con atributos. Fuente: proyecto “Ajueres domésticos”.

En este caso concreto, se han identificado seis diferentes entidades: archivo, documento, menaje, mobiliario, objeto devocional y ropa de casa. Las primeras dos, archivo y documento, son propedéuticas al registro de las entidades objeto de investigación (menaje, mobiliario, objeto devocional y ropa de casa) y sirven para agruparlas según su fuente de información (documento) y la colocación de la misma (archivo). Todas las relaciones son de cardinalidad 1:N, es decir que un registro de la tabla primaria se

¹⁰⁷ El modelo conceptual para este SIW se ha generado en colaboración con la investigadora principal del proyecto, Margarita M. Birriel Salcedo, profesora titular del Departamento de Historia Moderna y de América de la Universidad de Granada, experta en Historia Moderna y estudios de las mujeres.

¹⁰⁸ Actualmente, la base de datos cuenta con 1140 fichas registradas, correspondientes respectivamente a: 2 archivos, 63 documentos, 344 objetos de menaje, 385 muebles, 48 objetos devocionales y 298 fichas correspondientes a ropa de casa.

puede asociar a uno o varios registros de la tabla secundaria. Por ejemplo: cada archivo conserva muchos documentos, pero cada documento está presente en un solo archivo; cada objeto se describe en un solo documento, pero cada documento menciona varios objetos. La elaboración conjunta de taxonomías jerárquicas de términos controlados (figura 5) para clasificar las diferentes tipologías de objetos, materiales, colores, etc. ha llevado a un intenso proceso de revisión bibliográfica y de las fuentes, que ha sido muy beneficioso para el propio proyecto de investigación.

| Categoría documento | Tipología menaje | Tipología mobiliario | | Material | | Tipología Ropa de casa | Tipología objeto devocional |
|-----------------------|------------------|----------------------|---------------------|----------|----------|------------------------|-----------------------------|
| Carta de dote y arras | Alcuza | Arca | | tela | | Almohada | Crucifijo |
| Carta de dote | Almirez | Artesa | | | lienzo | Cabecera | Grabado |
| Inventario | Asador | Artesón | | | red | Cernadero | Lienzo |
| Testamento | Bañil | Bañil | | | indiana | Cobertor | Retablo |
| Tasación | Cacico | Bufete | | | estambre | Colcha | |
| Carta de capital | Caldera | Cama | | | tafetán | Colchón | pintura basta |
| | Caldero | | | | tiradizo | Cortina | Escultura |
| | Campanilla | | alfarjía | madera | | Jergón | Lámina |
| | Cazuela | | encordelada | | nogal | Mantel | |
| | Cedazo | Cama de colgar | | | palo | Paño de manos | |
| | Chocolatera | Cantarera | | | moral | Sábanas | |
| | Cuchara | Cielo de cama | | | junco | Servilletas | |
| | Espetera | Cofre | | | castaño | Tendido | |
| | Espumadera | Cuadro | | | pino | Toalla | |
| | Jicara | | Grabado | metal | | | |
| | Llares | | Lamina | | hojalata | | |
| | Olla | | Lienzo | | plata | | |
| | Orza | | Pintura | | pelitre | | |
| | Parillas | Delanteras de cama | | | hierro | | |
| | Perol | Escabel | | | bronce | | |
| | Plato | Escaño | | | cobre | | |
| | Rallo | Escritorio | | | azófar | | |
| | Rasera | Espejo | | | acero | | |
| | Sartén | Esteras | | | | | |
| | Tarro | Estrado | | | | | |
| | Tenazas | Iluminación | | | | | |
| | Trébedes | | Candelabro | | | | |
| | Vidriado | | Candelerero | | | | |
| | | | Candil | | | | |
| | | | Velador | | | | |
| | | | Velón | | | | |
| | | Mesa | | | | | |
| | | Rodapiés | | | | | |
| | | Silla | | | | | |
| | | | baqueta de moscovia | | | | |
| | | | de anea | | | | |
| | | Tarima | | | | | |

Figura 5. Taxonomías del modelo de datos sobre ajueres históricos. Fuente: proyecto “Ajueres domésticos”.

III.2.6. Definición de las funcionalidades y del flujo de datos

Volviendo al protocolo general de trabajo, una vez concretada la estructura del sistema en cuanto a tablas, campos, relaciones y taxonomías, ha sido necesario pasar a la **definición de las funcionalidades** (6), que en última instancia son lo que hacen del sistema una herramienta útil. Existe, de hecho, una estrecha relación entre estructura y función en la medida en que “una estructura permite o impide ciertas funciones; una función propone o incluso necesita de una estructura” (figura 6) (González-Pérez 1999).

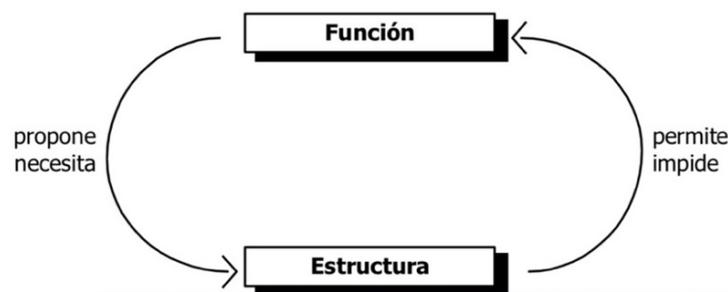


Figura 6. Relaciones entre estructura y función. Fuente: González-Pérez 1999.

Una importante ventaja de un SIW personalizado, frente a soluciones preconfeccionadas, es la posibilidad de incorporar solo las funcionalidades necesarias al proyecto, simplificando la interfaz usuario y optimizando el flujo de datos. Una de las funcionalidades indefectibles es, sin duda, el **sistema de búsqueda**, o de recuperación de la información, que en buena medida se apoya en la propia clasificación de conceptos definida a través de las taxonomías (5). En los casos de estudio analizados a continuación (véase capítulo IV) se ha optado, en general, por dos diferentes estrategias: un sistema de búsqueda por filtros, es decir por valores preestablecidos para determinar una combinación de criterios que los resultados tienen que cumplir; o un sistema por facetas, que son categorías interdependientes, y organizadas jerárquicamente, que permiten delimitar subconjuntos de resultados. Para un sistema de información web basado en datos de investigación estructurados, consideramos ambas opciones muy efectivas y preferibles con respecto a los sistemas de búsqueda que analizan las consultas del usuario para determinar los resultados en función de la relevancia y del contexto, como los que encontramos en los motores de búsqueda web. Por ejemplo, los actuales algoritmos de búsqueda de Google, RankBrain y BERT, están optimizados para analizar el lenguaje natural, en función de la semántica de la consulta del usuario, mientras que las dos opciones descritas anteriormente (filtros y facetas) permiten determinar criterios reproducibles basados en las propiedades o atributos de los datos, con el fin de aislar resultados que comparten determinadas características comunes. Además, la creación de un sistema de búsqueda estructurado obliga, en la etapa de determinación de las funcionalidades (6), a explicitar las preguntas con las que se quieren interrogar los datos. Se convierte, en definitiva, en una ulterior herramienta que favorece la reflexión, con el fin de determinar subconjuntos coherentes de datos a través de un proceso de aproximaciones progresivas.

Otra importante ventaja en el uso de una base de datos para almacenar la información es la posibilidad de generar **vistas dinámicas**¹⁰⁹, ya sea para visualizar que para cuantificar datos. Se trata de consultas preconfiguradas que se presentan como una tabla virtual, una vez seleccionados, entre otros, el origen, los campos a visualizar, los criterios de filtrado y de ordenación. Para volver al ejemplo de los ajueres domésticos, una de las funcionalidades identificadas, y resuelta a través de varias vistas preconfiguradas, ha sido la necesidad de calcular automáticamente el valor pecuniario de las diferentes clases de objetos descritos en cada documento de dotes y arras. A continuación,

¹⁰⁹ [https://es.wikipedia.org/wiki/Vista_\(base_de_datos\)](https://es.wikipedia.org/wiki/Vista_(base_de_datos)).

en la figura 7, se puede apreciar la ficha de un documento, con los metadatos a la izquierda y las páginas digitalizadas del mismo a la derecha, con la funcionalidad lupa para permitir la lectura directa. Abajo, los enlaces para crear objetos relacionados y, a continuación, las mencionadas tablas para la cuantificación automática del valor económico de cada agrupación.

AHPGR-DO Restábal. José Sáenz Diente. 1752-1753, f. 51r-f. 52v

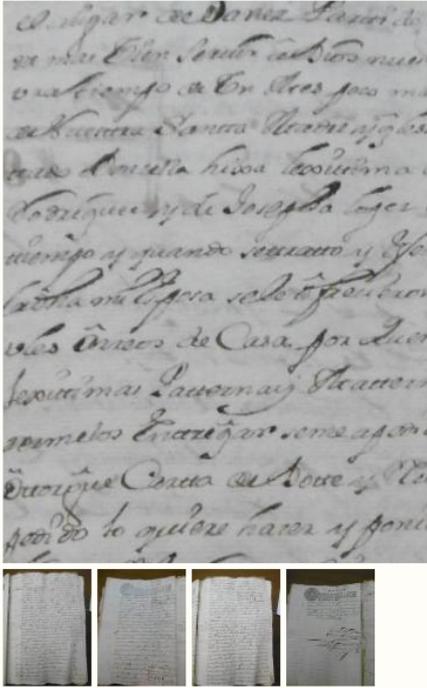
Ver Editar Gestionar presentación Devel

Post date: Martes, Enero 15, 2019 - 12:33
Archivo y sección: AHPGR-DO
Nombre del documento: Teodora Rodríguez López, su dote y arras. Contra, Juan López, su marido
Fecha: Sábado, Marzo 18, 1752
Categoría documento: Carta de dote y arras
Lugar: Izbor
Actor I: Teodora Rodríguez López
Actor II: Juan López

▼ Valor dote
Reales: 170

▼ Valor arras
Reales: 1 100

▼ Valor total
Reales: 1 270



Añadir Menaje Añadir Mobiliario Añadir Objeto devocional Añadir Ropa de casa

Objetos en el documento

Ropa de casa: 3 objetos en total ⚙️

| Tipología | Cantidad | Reales | Maravedís |
|------------------|----------|------------|-------------|
| Almohada | 2 | 10 | |
| Sábanas | 2 | 70 | |
| Colchón | 1 | 57 | |
| Total SUM | 5 | 137 | 0.00 |

Mobiliario: 4 objetos en total

| Tipología | Material | Cantidad | Reales | Maravedís |
|---------------------|--------------|----------|-----------|-------------|
| Silla, de anea | | 2 | 4 | |
| Iluminación, Candel | | 1 | | |
| Rodapiés | | 1 | 10 | |
| Cama, alfarjía | madera, pino | 1 | 11 | |
| Total SUM | | 5 | 25 | 0.00 |

Menaje: 1 objetos en total

| Tipología | Material | Cantidad | Reales | Maravedís |
|------------------|----------|----------|----------|-------------|
| Trébedes | | 1 | 8 | |
| Total SUM | | 1 | 8 | 0.00 |

Figura 7. Ficha documento, en la que aparecen las tablas resumen de los objetos catalogados, con las cuantificaciones automáticas. Fuente: proyecto “Ajuares domésticos”; <https://proyectoajuaresdomesticos.es/>.

El **sistema de control de versiones** es otra funcionalidad que se ha implementado en todos los sistemas de información con autoría y edición concurrente, en particular aquellos abiertos a la colaboración con usuarios externos no especializados. De hecho, de cara a las aportaciones externas, contar con un sistema de control de versiones permite la verificación a posteriori de dichas aportaciones, favoreciendo estrategias de involucración de la ciudadanía no mediadas (véase apartados IV.4.2 y IV.4.3). En cambio, para las aportaciones internas al grupo de investigación, esta funcionalidad habilita la moderación de los datos en diferentes estados, desde la versión borrador a la versión pública verificada (véase apartado IV.2.1).

En esta fase de obtención de requisitos es cuando se ha tenido además que valorar si el sistema necesitase de estar abierto a la colaboración externa, en cuanto a aportaciones de nuevos datos o modificación/enriquecimiento de datos existentes. En esos casos, ha sido necesario incorporar funcionalidades adicionales como, por ejemplo, la definición de roles con permisos restringidos, un sistema de notificación para nuevas aportaciones, filtros *anti-spam*, tablas de clasificación de colaboradores con insignias, tableros de anuncios, etc.

III.2.7. Implementación informática

Para pasar a la fase de **implementación informática** (7) ha sido necesario valorar, en función del modelo de datos diseñado (4) y de las funcionalidades previamente identificadas (6), la solución software más adecuada o que necesitara del nivel mínimo de personalización para poder responder a las necesidades del proyecto (véase apartado III.1). En el contexto de esta investigación, el entorno de trabajo utilizado con más frecuencia ha sido Drupal, en particular en su versión 7, con el que se ha podido dar respuesta a una amplia variedad de casos de usos (figura 8). Como hemos visto anteriormente, se trata de un software de código abierto, que cuenta con una comunidad de desarrollo muy amplia y activa, y con una consolidada trayectoria de uso en humanidades¹¹⁰. Se ha elegido por su flexibilidad a la hora de describir y visualizar entidades con una estructura compleja, su capacidad de gestionar taxonomías, así como por la amplitud de módulos disponibles y la facilidad para desarrollar nuevas extensiones necesarias para cubrir necesidades especiales.

¹¹⁰ Para una descripción de Drupal y referencias al amplio uso en contextos relacionados con las Humanidades Digitales, véase el apartado III.1.



Figura 8. Infografía de las fases nucleares de elaboración de un sistema de información web personalizado. Fuente: elaboración propia.

III.2.8-9. Pruebas y despliegue

La fase de **pruebas** (8) se ha llevado a cabo en paralelo con la implementación informática (7), ya que se ha sometido a prueba cada sistema a medida que se iban incorporando nuevas funcionalidades, para verificar que no crearan conflictos con las existentes (pruebas de regresión). Una vez implementada una primera versión completa del sistema, denominada versión *alpha*, las pruebas se han orientado a simular el comportamiento de los usuarios y el flujo de datos previsto, con el objetivo de detectar y posiblemente corregir los errores introducidos en la fase de desarrollo y testear la usabilidad general del sistema. Esta fase, normalmente, no ha implicado los usuarios finales y se ha realizado en autonomía.

Para los SIW privados, es decir aquellos pensado exclusivamente para un uso interno al equipo de investigación, una vez concluidas las pruebas de la versión *alpha* se ha pasado directamente al **despliegue** (9), con la abertura del sistema a los usuarios del mismo. En los otros casos, cuando estaba prevista la involucración de usuarios externos, en particular el público en general (véase apartado IV.4.2), ya sea para la simple consulta o para la aportación y manipulación de datos, se ha realizado una ulterior fase de pruebas *beta*, con una controlada implicación de colaboradores externos, para limitar al máximo las posibles disfuncionalidades. En estos casos, se ha podido comprobar que el comportamiento de los usuarios no especializados es en cierta medida impredecible y obliga a desarrollar sistemas mucho más cerrados, en los que es aconsejable reducir el flujo de datos a un número limitado y lineal de pasajes.

III.2.10. Entrada, enriquecimiento y verificación de datos

La **entrada de datos** y el **enriquecimiento** (10) de los mismos con metadatos son etapas de la fase de uso del sistema y su ejecución corresponde a los usuarios finales del mismo. En el ciclo de vida de un sistema de información es indudablemente la fase que más se extiende en el tiempo. Sobre todo al principio, resulta fundamental para testar la bondad del modelo de datos y la efectividad de las funcionalidades implementadas, y poder llevar a cabo el necesario mantenimiento correctivo (11).

La **verificación de datos**, lejos de ser un momento puntual del proceso, se encuadra en el flujo de trabajo como un proceso rutinario que acompaña el uso de la plataforma. La configuración de distintos roles, con permisos de lectura y edición diferenciados, ha permitido, en los casos donde fuera necesario, implementar sistemas de revisiones basados en diferentes estados que cada ficha puede asumir en el proceso (ej. borrador, lista para revisar, corregida, publicada) y en una serie de transacciones entre ellos (tabla 3).

| Nombre de la transición | desde | hasta |
|-------------------------|--------------------|--------------------|
| Enviar a revisión | Borrador | Lista para revisar |
| Rechazada | Lista para revisar | Borrador |
| Comentarios resueltos | Lista para revisar | Corregida |
| Publicada | Lista para revisar | Publicada |
| Publicada | Corregida | Publicada |

Tabla 3. Ejemplos de transiciones entre estados de verificación de una ficha de registro.

Esta funcionalidad se ha revelado particularmente útil para los SIW utilizados por grupos de investigación relativamente grandes y jerarquizados, en los que los integrantes poseen responsabilidades y competencias diferenciadas. Por otro lado, en los sistemas más horizontales, tipo wiki, la verificación de los datos es normalmente un proceso que se realiza a posteriori, por lotes o a través de la propia comunidad pública de usuarios (véase apartado IV.4.2), sobre datos ya publicados, y se apoya en el mencionado sistema de control de versiones (véase apartado III.2.6).

III.2.11. Mantenimiento

Otro proceso rutinario, pero de naturaleza discontinua, que se ha tenido en cuenta desde la fase de planificación (1), es el de **mantenimiento** (11) (Sommerville 2005). Para la sostenibilidad de un sistema de información que reside en la Web, se trata de un aspecto central y muchas veces infravalorado, en particular en el contexto de proyectos de menor envergadura.

En primer lugar, ha sido necesario llevar a cabo el **mantenimiento correctivo**, que se dedica a corregir los errores que los usuarios puedan detectar durante el uso del sistema o que puedan identificarse a raíz de auditorías de seguridad. Por definición, todo contenido publicado en la Web es accesible desde cualquier nodo conectado a la red, y por eso queda expuesto a la identificación de posibles fallos de seguridad. En sistemas creados a partir de una distribución software concreta (véase apartado III.1), la mejor protección contra la revelación pública, y la subsecuente posible explotación malévol, de fallos de seguridad es la instalación repentina de los parches de seguridad publicados, en particular aquellos relativos a la componente *core*. En este sentido, Drupal, como solución software utilizada para la implementación de buena parte de los sistemas analizados en este trabajo, no proporciona, al estado actual, un sistema de actualización automática de la componente *core*, aunque sea una funcionalidad estratégica en desarrollo¹¹¹. BackdropCMS, que como hemos visto anteriormente es un *fork* de Drupal orientado hacia pequeñas empresas, organizaciones sin ánimo de lucro y proyectos con pocos recursos (véase apartado III.1), incorpora una versión manual de esta funcionalidad desde la versión 1.12¹¹². En definitiva, el objetivo de conseguir una cierta autogestión en el mantenimiento de sistemas de información web, en particular aquellos desarrollados para pequeños proyectos, está aún por conseguir, pero hay iniciativas que trabajan en esa dirección.

Además de ser un entorno accesible, la Web es también un espacio tecnológico en continua transformación y actualización, ya sea lado servidor que lado cliente. Esa continua evolución provoca, cada cierto tiempo, incompatibilidades de los sistemas desarrollados con las tecnologías web que sustentan su alojamiento en un servidor. Un ejemplo representativo en este sentido es la evolución del lenguaje PHP, en absoluto

¹¹¹ <https://www.drupal.org/about/strategic-initiatives/automatic-updates>.

¹¹² <https://backdropcms.org/news/backdrop-1-12-released>.

el más popular en la Web¹¹³. La transición desde la versión 5, soportada hasta 2018, a la versión 7, publicada en 2015, supuso relevantes mejoras, que pero causaron incompatibilidades con las aplicaciones existentes que obligan a una actualización manual del código fuente. Aún hoy en día, alrededor del 40% de todos los sitios web en PHP utilizan la versión 5.6 o anterior, lo que nos indica la relevancia que puede llevar la falta de **mantenimiento adaptativo**, que es como se define el conjunto de tareas dedicadas a asegurar esa compatibilidad. En la elección de la solución software más adecuada para desarrollar un SIW, es fundamental tener en cuenta la presencia de una amplia y activa comunidad de usuarios y desarrolladores, ya que la existencia de la primera crea las condiciones necesarias para que los segundos actúen en la resolución de las incompatibilidades sobrevenidas. Las tareas de mantenimiento adaptativo pueden reducirse eligiendo servicios de alojamiento web que pongan cierto énfasis sobre la retro-compatibilidad, tales como Reclaim Hosting¹¹⁴ o Webfaction¹¹⁵.

Otra faceta de la misma problemática es asegurar la compatibilidad en el tiempo con servicios externos, incorporados en el sistema, pero ofertados por terceras partes (mapas, identificación a través de redes sociales, indexación, etc.). Es emblemático, en este sentido, el caso de Google Maps¹¹⁶. Su API se publicó a mitad de 2005 y se convirtió en pocos años en la solución de facto para el mapeo en la Web, integrada en la gran mayoría de plataformas. En mayo de 2018¹¹⁷, Google anunció un cambio significativo en el servicio, que incluía desde entonces la necesidad de disponer, para poder seguir usando su API, de una cuenta de facturación en la plataforma Google Cloud, además de una reducción significativa de la cuota de uso gratuito. Ese cambio supuso la automática ruptura de todos los mapas dinámicos presentes en páginas web que no contaban con un servicio de mantenimiento adaptativo.

El tercer aspecto que se ha tenido que tener en cuenta en la parte final del ciclo de vida de cada sistema de información desarrollado es el **mantenimiento perfectivo**, que se ejecuta una vez transcurrido un tiempo desde el despliegue (9) de la primera versión de la plataforma. Se trata de la fase en la que se plantean y se incorporan nuevas funcionalidades o modificaciones en la estructura, cuya necesidad se ha identificado con el uso o para responder a una ampliación del ámbito de estudio. En esta fase, que puede también ser continua e iterativa en los métodos de desarrollo ágiles¹¹⁸, juegan un rol importante las lecciones aprendidas con el uso del sistema y la recogida de opiniones por parte de los diferentes grupos de usuarios involucrados.

¹¹³ Según los datos de W3Techs, PHP es usado por el 79.1% de todos los sitios web con un lenguaje de programación lado servidor conocido. Fuente: <https://w3techs.com/technologies/details/pl-php>.

¹¹⁴ <https://reclaimhosting.com/>.

¹¹⁵ <https://www.webfaction.com/>. Desde principios de 2021, los servicios de Webfaction están siendo migrados al proveedor de alojamiento tsoHost.

¹¹⁶ <https://developers.google.com/maps/documentation?hl=es>.

¹¹⁷ <https://cloud.google.com/blog/products/maps-platform/introducing-google-maps-platform>.

¹¹⁸ https://en.wikipedia.org/wiki/Agile_software_development.

En definitiva, el mantenimiento repite "en miniatura" el ciclo completo de vida de un sistema de información, y de ahí que el mismo se pueda definir un ciclo y no simplemente una progresión de etapas.

En paralelo con la elaboración de la presente tesis doctoral se han diseñado otros modelos de datos muy distintos y con diferentes grados de complejidad, que sucesivamente se han implementado en sistemas de información web actualmente en uso en proyectos de investigación muy dispares, cuyo análisis y descripción aquí no ha encontrado cabida. Sin embargo, en este capítulo, así como en las conclusiones del presente trabajo, se recogen una serie de reflexiones generadas también a partir del desarrollo de esos sistemas. Consideraciones que nos ha permitido, por un lado, afinar la metodología de trabajo y, por el otro, verificar la aplicabilidad de los principios y de las prácticas que aquí se exponen a otras disciplinas de estudio, desde la Antropología Cultural y el estudio de género, a la micromorfología arqueológica y el estudio del biodeterioro del patrimonio cultural (Bolívar-Galiano *et al.* 2020).

En conclusión de este apartado, consideramos que la cantidad y calidad de procesos necesarios para el diseño y la implementación de un sistema de información web permiten considerarlo como una de las principales herramientas a disposición del humanista digital para implementar procesos de digitalización de la investigación. Constituyen además una oportunidad para poner a prueba y verificar las preguntas de investigación, así como para plantear nuevas hipótesis y generar nuevas modalidades de trabajo colaborativo. Un buen sistema de información web representa finalmente el mejor presupuesto hacia la producción de buenos conjuntos de datos.

Más allá del protocolo de trabajo descrito, algunas de las publicaciones que presentamos a continuación han requerido de metodologías propias de recogida y procesamiento de datos. Para esos casos, véase las secciones metodológicas de las correspondientes publicaciones.

IV. RESULTADOS

Tal y como se ha descrito en el capítulo de introducción, los resultados de investigación de esta tesis doctoral han sido desarrollados y se presentan a través de un total de seis publicaciones. En concreto: cuatro artículos en revistas científicas de impacto y dos capítulos de libros en editoriales de prestigio. El respectivo idioma de cada publicación ha sido elegido según la audiencia preferente y en función del tema objeto de estudio: tres trabajos se han redactados en inglés, dos en español y uno en italiano.

En cada apartado de este capítulo de resultados se presentan los trabajos correspondientes, con el mismo texto, imágenes, gráficos, notas y anexos de la publicación original¹¹⁹, pero siguiendo el formato de maquetación de esta tesis, según la normativa de la Escuela Internacional de Posgrado de la Universidad de Granada. Cada publicación está precedida por un texto explicativo del rol que la misma ocupa en el desarrollo global del trabajo, así como de los principales resultados obtenidos en función de los objetivos de investigación planteados (véase capítulo II).

IV.1. EL CONTEXTO DE LAS HUMANIDADES DIGITALES EN ESPAÑA (1993-2019)

El primer artículo que presentamos se ha planteado para adquirir una visión de conjunto sobre el panorama de las Humanidades Digitales en España, en cuanto a investigadores y centros de investigación, disciplinas implicadas y temas de investigación de interés, principales organismos de financiación y evolución de su inversión, desde los principios de los años noventa hasta la actualidad. En definitiva, el objetivo ha sido mapear los promotores y hacedores de recursos digitales. Entre estos artefactos se encuentran numerosos sistemas de información web y, en general, infraestructuras digitales para la investigación, entendidas éstas últimas como la combinación e integración entre los recursos de información digital, las herramientas analíticas y de visualización y la comunidad activa de investigadores que colaboran a través de proyectos de investigación.

En el contexto de la presente tesis doctoral, el rol de esta publicación es proporcionar la trama de fondo sobre la que se mueven los investigadores en Humanidades Digitales en España; así como facilitar al lector un marco general y una serie de puntos de referencia para la lectura y el encuadramiento de los demás resultados de investigación. Introducimos así, con este trabajo, el análisis del primer componente del esquema conceptual representado en la figura 1: las disciplinas humanísticas afectadas por el giro digital.

El mapeo se ha llevado a cabo mediante la combinación de dos principales tipologías de fuentes de datos: 1) información disponible en línea sobre congresos, seminarios,

¹¹⁹ Siguiendo este criterio, la numeración de los sub-apartados, así como de las figuras y de las notas, es interna a cada publicación y no sigue la numeración general de la tesis.

convocatorias temáticas, mapas participativos, etc.; 2) conjuntos masivos de datos sobre publicaciones científicas (procedentes de las bases de datos bibliográficas Dialnet e ÍNDICES-CSIC) y proyectos de investigación financiados (procedentes de la Agencia Estatal de Investigación). Podemos identificar una serie de resultados a destacar: 1) se ha propuesto una primera periodización en tres fases en la evolución y consolidación de la disciplina; 2) se han identificado cinco grandes grupos disciplinarios que unen un amplio conjunto de temas transversales de investigación; 3) se han detectado tendencias y patrones en el desarrollo de recursos digitales para la investigación; 4) se han establecido nuevas métricas y valores que constituyen una base de referencia para monitorear la evolución de las Humanidades Digitales en España y favorecer futuras comparaciones, tanto a lo largo del tiempo como con otros contextos a nivel europeo e internacional.

Finalmente, en las conclusiones del artículo, se proporcionan una serie de recomendaciones que avanza propuestas concretas sobre un conjunto de problemáticas identificadas. Esas van desde el reconocimiento disciplinar, las nuevas competencias y los perfiles híbridos, la colaboración y la coautoría, hasta la cuestión abierta de la participación de España en las grandes infraestructuras europeas dedicadas a la investigación en humanidades (DARIAH y CLARIN); así como la falta de reconocimiento, por parte de las agencias de evaluación de la actividad investigadora, por actividades como el diseño, la implementación y la codificación de bases de datos, el desarrollo de infraestructuras de datos espaciales o de herramientas de análisis de datos, temas centrales en el presente estudio, y que nosotros defendemos como parte integrante de los nuevos procesos de investigación en humanidades y por eso merecedoras de reconocimiento científico y curricular.

La realización de este artículo se ha llevado a cabo en el marco de una colaboración con el proyecto H2020 DESIR (*DARIAH ERIC Sustainability Refined*¹²⁰), liderado por los investigadores Elena González-Blanco¹²¹ y Salvador Ros, y con la contribución de la investigadora pre-doctoral Aroa Rabadán, que ha participado en la fase de recogida de datos. En este contexto, el candidato ha colaborado con los demás autores en las fases de conceptualización del trabajo, búsqueda bibliográfica, recogida y curación de datos. Además, ha llevado a cabo en autonomía las siguientes tareas: creación del sistema de información para la recogida de datos, definición de la metodología, análisis cualitativo y cuantitativo de los datos, elaboración de gráficos, redacción del borrador inicial, revisión formal, correspondencia con el comité editorial de la revista e implementación de las sugerencias de los revisores.

¹²⁰ Uno de los objetivos del proyecto era propiciar la adhesión de España a DARIAH. El proceso de adhesión se encuentra actualmente en curso en el marco del proyecto INTELE: INfraestructura de TECnologías del Lenguaje <http://ixa2.si.ehu.eus/intele/>.

¹²¹ Destacar el rol de Elena González-Blanco en el ámbito de las Humanidades Digitales en España y a nivel internacional, siendo la primera investigadora española en obtener una ERC- *Starting Grant* (2016 - Proyecto POSTDATA) y una ERC - *Proof of Concept* (2020 - Proyecto LyrAics) sobre esta disciplina.

El conjunto de datos recogidos en esta investigación¹²² (Toscano y Díaz 2020) ha sido sucesivamente utilizado para realizar un análisis sobre las *scholarly primitives*, es decir de las “basic functions common to scholarly activity across disciplines, over time, and independent of theoretical orientation” (Unsworth 2000), desde la perspectiva de la Humanidades Digitales. Los resultados de este estudio, liderado por el mismo autor de la presente tesis doctoral, han sido presentados en la conferencia internacional de la Infraestructura Europea para la Investigación Digital en Arte y Humanidades DARI AH2020 y han recibido el premio al mejor póster¹²³ (Toscano *at al.* 2020).

¹²² Disponibles en GitHub <https://github.com/linhd-postdata/desir/tree/v1.0> y en Zenodo <https://doi.org/10.5281/zenodo.3893546>, como Jupyter Notebooks.

¹²³ <https://www.virtualdariah2020.dariah.eu/>.

Referencia bibliográfica

Toscano, M., Rabadán, A., Ros, S., & González-Blanco, E. (2020). Digital humanities in Spain: Historical perspective and current scenario. *El profesional de la información*, 29(6), e290601. <https://doi.org/10.3145/epi.2020.nov.01>

Título original

Digital Humanities in Spain: historical perspective and current scenario

Título en español

Humanidades digitales en España: perspectiva histórica y escenario actual

Abstract

The objective of this study was to provide the global community of interested scholars with an updated understanding of digital humanities in Spain, in terms of researchers and research centres, disciplines involved and research topics of interest, trends in digital resources development, main funding bodies and the evolution of their investment since the early nineties. One of the characteristics that differentiates this study from previous approaches is the information used to carry out the research. It combines large datasets of publicly available data from trusted sources with a handpicked selection of records grouping information scattered over the Web. Most of the evidence detected by other studies has been numerically confirmed. At the same time, the new metrics and values established constitute a reference base for monitoring the future evolution of the discipline, and thus favour comparisons. Half of the researchers were found to be affiliated to only nine institutions, whereas the other half of them were scattered across 84 locations. Department affiliation showed a varied pattern of the different degrees of specialization in each institution. Although the major historic role played by Philology was confirmed, the rising interest of other areas of the Humanities and Social Science produces a wider picture, which helped to identify five large clusters of research topics, centred on major disciplines. The quantitative analysis of funding, a dimension almost unexplored in the Humanities, proved to be a valuable way to assess the discipline and its historical evolution. In fact, it revealed interesting trends that led to our proposal of a three-phase periodization in the consolidation of digital humanities in Spain. The paper concludes with a set of recommendations regarding how to deal with issues that could harm the future development of this research area and the role that Spanish researchers can play in the international context.

Resumen

El objetivo de este estudio era proporcionar a la comunidad internacional de estudiosos interesados un conocimiento actualizado de las Humanidades Digitales en España, en cuanto a investigadores y centros de investigación, las disciplinas implicadas y los temas de investigación de interés, las tendencias en el desarrollo de los recursos digitales, los principales organismos de financiación y la evolución de su inversión desde principios de los años noventa hasta la actualidad. Una de las características que diferencian este estudio de los anteriores es la información utilizada para llevar a cabo la investigación. Se han combinado grandes conjuntos de datos de acceso público

procedentes de fuentes de confianza con una selección minuciosa de registros que agrupan información dispersa en la Web. La mayoría de las evidencias detectadas por otros estudios ha sido confirmada numéricamente. Al mismo tiempo, las nuevas métricas y valores establecidos constituyen una base de referencia para el seguimiento de la evolución futura de la disciplina y, por lo tanto, favorecen las comparaciones. Se ha comprobado que la mitad de los investigadores está afiliada a sólo nueve instituciones, mientras que la otra mitad está dispersa en 84 centro de investigación. La afiliación a los departamentos ha revelado un patrón variado de los diferentes grados de especialización de cada institución. Aunque se confirmó el importante papel histórico desempeñado por la Filología, el creciente interés de otras esferas del conocimiento en Humanidades y Ciencias Sociales produce un panorama más amplio, que ayudó a identificar cinco grandes grupos de temas de investigación, centrados en las principales disciplinas. El análisis cuantitativo de la financiación, una dimensión casi inexplorada en las humanidades, demostró ser una forma valiosa de evaluar la disciplina y su evolución histórica. De hecho, ha revelado interesantes tendencias que nos han llevado a realizar nuestra propuesta de periodización en tres fases en la consolidación de las Humanidades Digitales en España. El trabajo concluye con un conjunto de recomendaciones sobre cómo abordar las cuestiones que pueden perjudicar el desarrollo futuro de esta área de investigación y el papel que pueden desempeñar los investigadores españoles en el contexto internacional.

Keywords

Digital humanities; Humanities; Digital scholarship; Digital tools for research; Open data; Investment in research; Funding; Historical analysis; Metrics; Spain.

Palabras claves

Humanidades digitales; Humanidades; Investigación digital; Herramientas digitales para la investigación; Datos en abierto; Inversión en investigación; Financiación; Análisis histórico; Métricas; España.

Funding

This article is a result of the project DESIR (Dariah ERIC Sustainability Refined), who received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 731081, and the project POSTDATA, who received funding from the European Research Council (ERC) in the European Union's Horizon 2020 research and innovation programme under grant agreement No. 679528.

Acknowledgements

We thank *Agencia Estatal de Investigación* for providing data extracted from the public database of call resolutions; Aitor Díaz Medina for the assistance in the creation of the Jupyter Notebooks; Wenceslao Arroyo Machado for suggesting improvements in the network analysis graph; Dr. Pamela Faber for proofreading the text; Dr. Enrique Herrera-Viedma and Dr. María Dolores Romero López for reading an earlier draft of this manuscript and provide meaningful counsels, and the anonymous reviewers for their valuable feedback and suggestions.

1. Introduction

In the last ten years, Digital Humanities (DH) has become a field of interest in Spain even though, in comparison to other European countries, this interest was somewhat later in arriving. In fact, it has become a leading research trend, either as a field of study or as a preferential funding topic. At the same time, because of its novelty, it is now under scrutiny by the research community and government institutions because the investment return is still unclear. Neither is the role that Spanish researchers can play within European research infrastructures, such as DARIAH and CLARIN. In order to provide the global community of scholars working in this field with a greater understanding of the current Spanish scenario, LINHD (*Laboratorio de Innovación en Humanidades Digitales - UNED*) promoted a research study on the evolution of Digital Humanities in Spain over the last 25 years approximately, using a quantitative approach. The objective of the study was to identify researchers in the field of Digital Humanities and to explore their funding, institutional affiliations, research topics, and digital resources developed.

In the past, there have been similar initiatives, which have resulted in internationally or regionally focused maps¹, some of which are still available online. These maps vary in regard to geographical coverage and the type of data mapped (**Ortega; Eunice-Gutiérrez**, 2014; **Romero-Frías; Del-Barrio-García**, 2014). Others overviews on the subject have chosen a different approach, using bibliography or other sources to identify the most relevant steps in the evolution and consolidation of this research topic in the Spain (**Rojas-Castro**, 2013; **González-Blanco**, 2013; **Spence; González-Blanco**, 2014). Rather than a comprehensive historical review, this paper offers a complementary view, based on an analysis of the most recent data and previously untapped sources of information.

This research was very data-oriented since it primarily assessed the quantitative dimension of Digital Humanities in Spain. This six-month study was conducted in three phases. During Phase I (August-September 2019), the mapping methodology was developed. A shared Web information system was put in place and the data model was defined. In Phase II (October-December 2019) the core data were collected and verified. During phase III (January-February 2019), the results were analysed, and this paper written.

In order to obtain an overview of a research field in a certain context, a common practice is to use scientific outputs (i.e. research papers, books, and reports). This approach to Digital Humanities in Spain was problematic for three reasons. Firstly, the main databases for consulting the scientific production of a discipline (e.g. Web of Science and Scopus) are not representative of Digital Humanities because most reference journals are not indexed there or have been incorporated very recently. The second issue relates to the Digital Humanities label, whose use is not widespread across the whole spectrum of Humanities. This is particularly the case in Spain, where it is mainly linked to literary studies (Philology and Linguistics). It is even less frequently used in

other fields with a long tradition of computer applications and methods, such as Archaeology and Cultural Heritage. Without a doubt, the community of scholars actually working in the discipline is significantly larger than the community of scholars that identify themselves with the Digital Humanities label. Thirdly, a somewhat less mentioned but also relevant issue is the varying usage of the term in fields such as Medicine, Psychology, or Economics, which further complicates bibliometric analysis. Finally, problems with defining Digital Humanities² multiplies the difficulties in assessing whether a project or a researcher should be included in the final dataset.

2. Materials and methods

2.1 OBJECT OF THE RESEARCH

This paper describes the results pertaining to the current state of digital research infrastructures in the field of Arts and Humanities in Spain. Digital research infrastructures combine and integrate digital information resources, analytical and visualization tools, and an active community of researchers, collaborating within or across privately or publicly funded research projects. For this reason, we subdivided the object of the research into five main entities: researchers, projects, resources, bibliography, and postgraduate courses.

The researcher record represents the main entity that ties everything together, as Principal Investigator or member of research projects, promoter of digital resources, and author of research papers. Through the researchers' affiliation, we were also able to identify laboratories and research groups working in the field of Digital Humanities, as well as the institutions that they belong to. Furthermore, the background of researchers and their main research topics were also recorded. Project records were key elements in mapping two essential data: (1) evolution of research topics over time; (2) funding. Digital resources and bibliography were considered to be conceptually similar entities, and were regarded as research project outputs, whereas education was included to highlight a recent trend that promoted postgraduate courses, mainly master's courses at different universities.

2.2 EXTENT OF THE MAPPING

Several factors influenced and helped to define the research topics and questions. Key considerations were DESIR (DARIAH-ERIC Sustainability Refined) project goals (i.e. preparing the groundwork for establishing DARIAH membership in new countries and measuring the acceptance of DARIAH in new communities), and the types of information that might be useful to support a Spanish candidacy in the DARIAH European Infrastructure (i.e. number of potential users, research groups, and institutions).

A topic of some debate was which criteria should be used to include a researcher in the dataset. Finally, inclusion was based on researcher affiliation with a Spanish research centre (private or public) or participation in a research project funded by a

Spanish institution. Following these criteria, we also included 14 researchers, working in foreign universities (i.e. Paris, Lisbon, Miami, etc.), whereas we excluded researchers working abroad in foreign-funded projects, even if their research object was within our scope (i.e. digital humanists working on Spanish literature from the United Kingdom). Regarding the timespan considered, records cover the period from 1992³ to 2019. Our plan is to keep the dataset open for the inclusion of data, which will be periodically analysed.

2.3 DATA SOURCES

In the context of the H2020-DESIR project, an online survey on the topic of this paper was first conducted in 2018. However, participation was scattered, and results were unsatisfactory. This time we chose a different approach in order to ensure data consistency and efficient results in a shorter timeframe. The vast majority of research projects and resources in Digital Humanities are directly available or at least presented on the Web, because of their digital nature. In addition, most researchers have a more or less complete online profile. Bibliography is now easily available online, thanks to a multitude of databases (i.e. Dialnet, ÍnDICES-CSIC, Humanities Full Text - Wilson, MLA International Bibliography, etc.) and academic social profiles on platforms such as Google Scholar, Academia, or ResearchGate. For these reasons, our methodology involved two phases: (1) manually gathering information online without the direct involvement of the researchers; (2) identifying and importing information in bulk from publicly available datasets. This approach also facilitated ongoing quality and criteria control, which would have been a very time-consuming task if carried out afterwards on researcher-generated data.

The two publicly available datasets used were bibliography and public funded research projects. Firstly, bibliography was extracted by looking for the keyword “humanidades digitales” [digital humanities] both in Dialnet and in ÍnDICES-CSIC⁴. The searches in these two databases returned 409 and 352 records, respectively. Once the duplicates and off-topic records were cleaned, the final dataset consisted of 309 publications from 427 authors. In this context, bibliography was only used as a source of information, to identify authors and projects in the field, and was not analysed per se. Such a study would have been difficult because Digital Humanities literature is scattered over a wide variety of sources, is often not indexed, and is not well represented in databases, such as Web of Science and Scopus.

Secondly, to access the results of publicly funded research projects, the *Agencia Estatal de Investigación* – AEI – [Spanish Government Research Agency] was contacted. The AEI, which is in charge of coordinating research project evaluations in Spain, provided details of 993 projects from its own database, corresponding to the following keywords (in Spanish): Digital Humanities, Digital Libraries, Digital History, Computational Linguistics, Ontologies, Digital Heritage, GIS, SPARQL, TEL, ICT, Semantic Web, XML. Since some of these keywords were quite generic and not strictly within our scope, it was necessary to reduce the dataset to 246 records before its importation.

Despite the central role of these two datasets, other information sources were also consulted, particularly those related to Digital Humanities, which provided data from otherwise hidden entities. Data were manually gathered from Digital Humanities events, such as seminars and conferences; user-generated online collections, such as GrinUGR and RedKnowmetrics; and specific grants, like those offered each year by the Fundación BBVA. In addition, since this information is interlinked, each dataset was also used as a source of additional data. For example, project records helped to create many researcher records and so on⁵. The information sources were recorded for each entity type. The full spectrum of the data sources used is listed in Table 1 as well as in Figures 1-3.

| Source | Acronym | # of records |
|--|----------------------|--------------|
| Agencia Estatal de Investigación | AEI | 344 |
| Bibliografía | Biblio | 271 |
| Congresos Asociación Humanidades Digitales Hispánicas | Congresos HDH | 95 |
| Own source | Own source | 45 |
| Fundación BBVA | BBVA | 42 |
| GrinUGR | GrinUGR | 33 |
| Red Knowmetrics | Knowmetrics | 33 |
| Google search | Google | 24 |
| RecerCaixa | RecerCaixa | 23 |
| ARACNE | ARACNE | 20 |
| Proyectos de Humanidades Digitales en la BNE | BNE | 19 |
| Researcher personal website | Researcher website | 17 |
| Congreso Territorio Digitales | Territorio Digitales | 14 |
| I Jornada científico-técnica en Humanidades Digitales en el CSIC | Jornada CSIC | 12 |
| Medialab UGR projects | Medialab UGR | 12 |
| Laboratorio de Innovación de Humanidades Digitales | LINHD | 11 |
| Biblioteca Virtual Miguel de Cervantes | Biblio Cervantes | 10 |
| Portal del Hispanismo | Portal Hispanismo | 7 |
| BIDISO research group | BIDISO | 6 |
| Academia.edu | Academia.edu | 4 |
| Asociación Humanidades Digitales Hispánicas awards | Premios HDH | 4 |
| Diarium (USAL) | Diarium | 4 |
| Seminario Historia y Humanidades Digitales UCLM 2017 | Seminario UCLM | 4 |
| Portal Andrés de Poza | Andrés de Poza | 4 |
| Proyecto ReMetCa | ReMetCa | 1 |
| Consorcio Madroño | Madroño | 1 |
| CORDIS | CORDIS | 1 |

Table 1. Full range of data source used in the research.

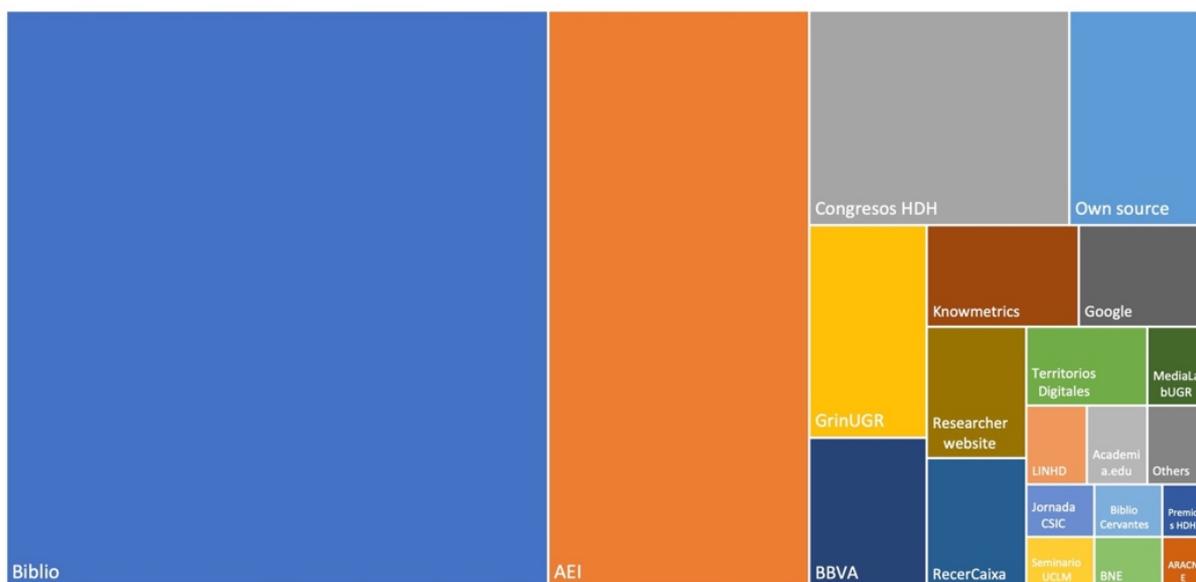


Figure 1. Data sources presented as a treemap chart, weighted according to their contribution to the researcher table.

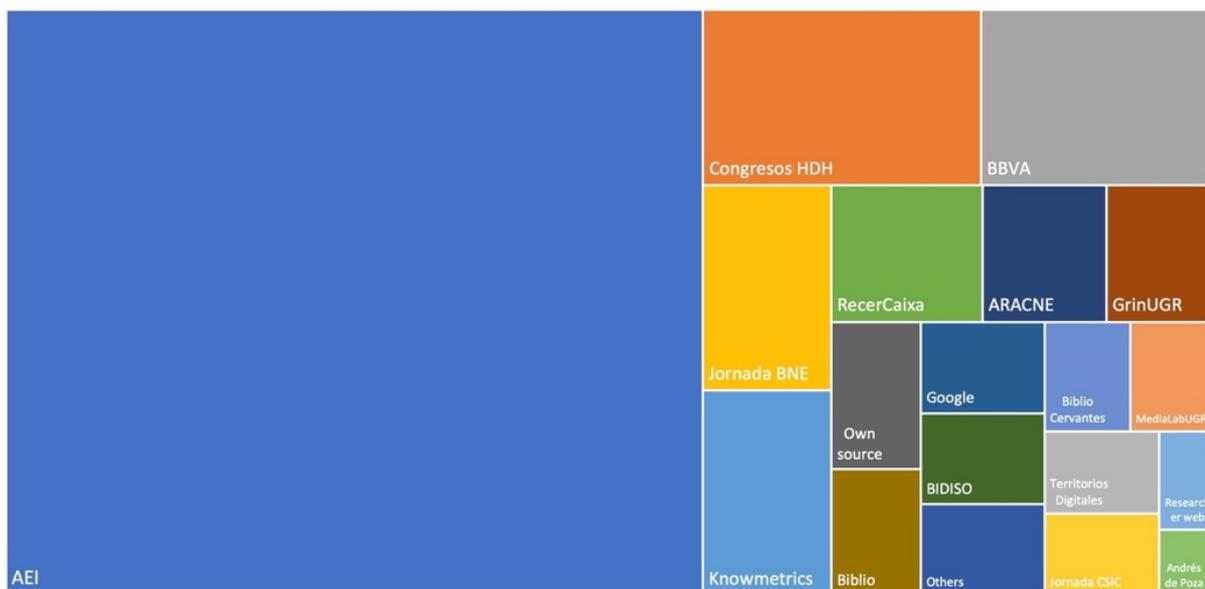


Figure 2. Data sources presented as a treemap chart, weighted according to their contribution to the project table.

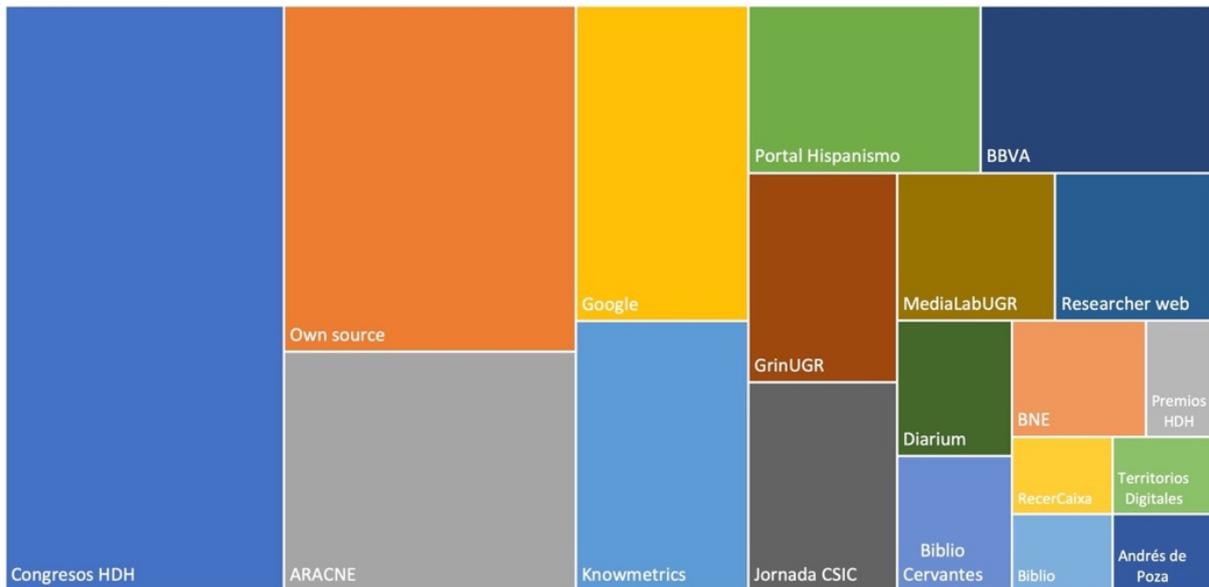


Figure 3. Data sources presented as a treemap chart, weighted according to their contribution to the resource table.

As can be observed, both researchers and projects have a key data source, respectively bibliography and AEI data, whereas digital resource records were filled with a more balanced range of sources.

2.4 DATA ANALYSIS AND VISUALIZATION

From the very beginning of this research, our plan was to share Open Data and to document the analytical workflow used so that it could be verified and reproduced. For this purpose, we used an Open Science approach. Accordingly, we published the final dataset in CSV format on a GitHub repository (<https://github.com/linhd-post-data/desir>), as well as the procedures used to analyse it. These included Jupiter Notebooks combining Markdown descriptions and Python code. Once the final version of the data and code for this paper were obtained, we linked the GitHub repository to Zenodo to obtain a DOI⁶ and keep track of future versions. Although this approach is gaining traction, thanks to the focus set by the European Commission on Open Data and Open Science (Walsh, 2016), it is still not very popular in the Humanities. In addition to making the results accessible and reproducible by separating the data from the code used, it will make it much easier to renew all charts with an updated dataset in the future. It will also allow other researchers to replicate them and even manipulate the process, via a standard browser and a virtual environment, such as Binder.

To analyse the relationship between disciplines and research topics, network analysis was used. The Leiden algorithm (Traag *et al.*, 2019) was applied, using modularity method and resolution of 1, to calculate communities. The average degree distribution, as well as the betweenness and closeness centrality were also calculated.

3. Discussion of results

The results obtained are presented and discussed in the following sections. They are based on each mapped entity: (1) researchers (gender ratio, disciplinary background, research topics); (2) research institutions and higher education centres (institutional and departmental affiliation, geographical distribution, postgraduate courses); (3) research projects (investment historical evolution, funders, institutional distribution of funding); (4) digital resources (trends, types, sustainability). A final section of the discussion focuses on a series of issues detected from the analysis of collected data.

3.1 RESEARCHERS

The total volume of records is 1,359, distributed as follows: 577 researchers; 368 projects; 309 bibliographical records; 88 resources; 9 post-graduate courses; and 8 journals. As for the researchers, 305 were males (52.9%) and 272 females (47.1%). This ratio, as compared to the general gender balance among researchers in Spain across all disciplines (61.2% of males and 38.8% of females) or limited to Humanities (59.8% of males and 40.2% of females), reflects a significantly higher female presence than expected⁷ (Table 2). Unfortunately, updated figures regarding researchers in Spain, by discipline, are not easy to find since the most recent data provided by the National Institute of Statistics (INE) date back to 2001⁸. More recent figures from the INE only apply to all researchers or are categorized by gender. We thus combined them with the percentage of researchers by discipline taken from the report *Científicas en cifras 2017*, presented by the FECYT (*Fundación Española para la Ciencia y la Tecnología*) in 2019⁹.

| Researchers in Spain | Males: units (ratio) | Females: units (ratio) | Totals |
|----------------------------------|----------------------|------------------------|---------|
| All disciplines ¹⁰ | 85,762 (61.2%) | 54,358 (38.8%) | 140,120 |
| Humanities ¹¹ | 11,149 (59.8%) | 7,501 (40.2%) | 18,650 |
| Digital Humanities ¹² | 305 (52.9%) | 272 (47.1%) | 577 |

Table 2. Total researchers and gender percentage in Spain for all disciplines, Humanities and Digital Humanities.

The total number of 577 researchers working in the field of Digital Humanities in Spain may seem high for a relatively new discipline. However, if we restrict our study to researchers in disciplines belonging to Humanities, and thus exclude Computer Science, Library Science, and Communication, the number decreases to 421 (only 2.25% of the total number of researchers in Humanities). At the same time, this number encompasses both those researchers with clear and consolidated research interests in Digital Humanities as well as those with incidental contributions, which are limited to a paper or collaboration in a research project, or those only marginally involved in the discipline, mainly because of its multidisciplinary nature. Others, for example, the youngest researchers or independent researchers, unaffiliated with a research institution, whose work is less visible, may not be included although they are actively working in this field.

We classified researchers by discipline (Figure 4) and discovered a clear prevalence of scholars from Philology, followed (as expected) by History and Computer Science. These three disciplines represent a clear majority of researchers (57.7%). Regarding the other disciplines, there are two groups: (1) 21.5% related to other disciplines in Humanities (i.e. Archaeology, History of Arts, Linguistics, Geography, Philosophy and Visual Arts); (2) 17.5% pertaining to Communication, Library Science, Documentation and Educational Science. Despite the evidently skewed distribution, it is highly relevant that there are almost twenty different disciplines involved. This is additional proof of the interdisciplinary scope of Digital Humanities.

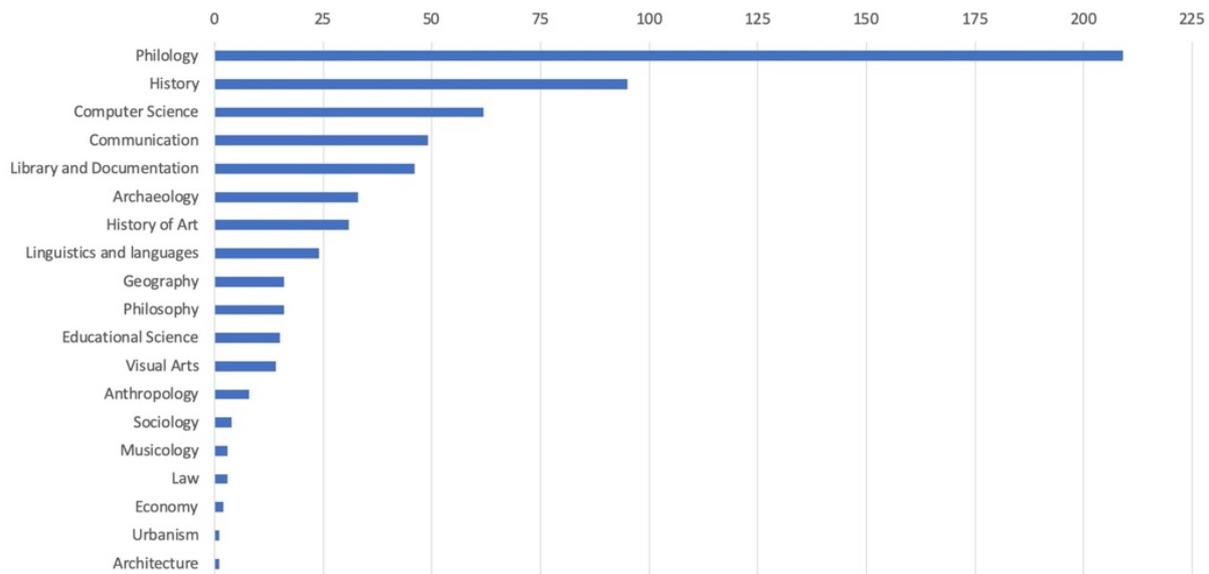


Figure 4. Disciplinary backgrounds of researchers working in Digital Humanities in Spain.

If this disciplinary classification is used to compare researchers with projects¹³, the information in Figure 5 reflects a very strong and predictable correlation between the percentage of records in each field (Pearson index 0.96). However, certain differences can be observed.

For example, there is a higher percentage of projects compared to researchers in Philology, History, Computer Science, Linguistics and History of Arts, and a lower percentage of projects in Communication, Library Science, Geography or Philosophy. These differences are understandable since researchers from core disciplines tend to lead projects in Digital Humanities. In some cases, they have directed several projects, whereas others from ancillary/peripheral fields, who have joined this study area more recently, have contributed other ways (e.g. co-authoring publications rather than leading research projects).

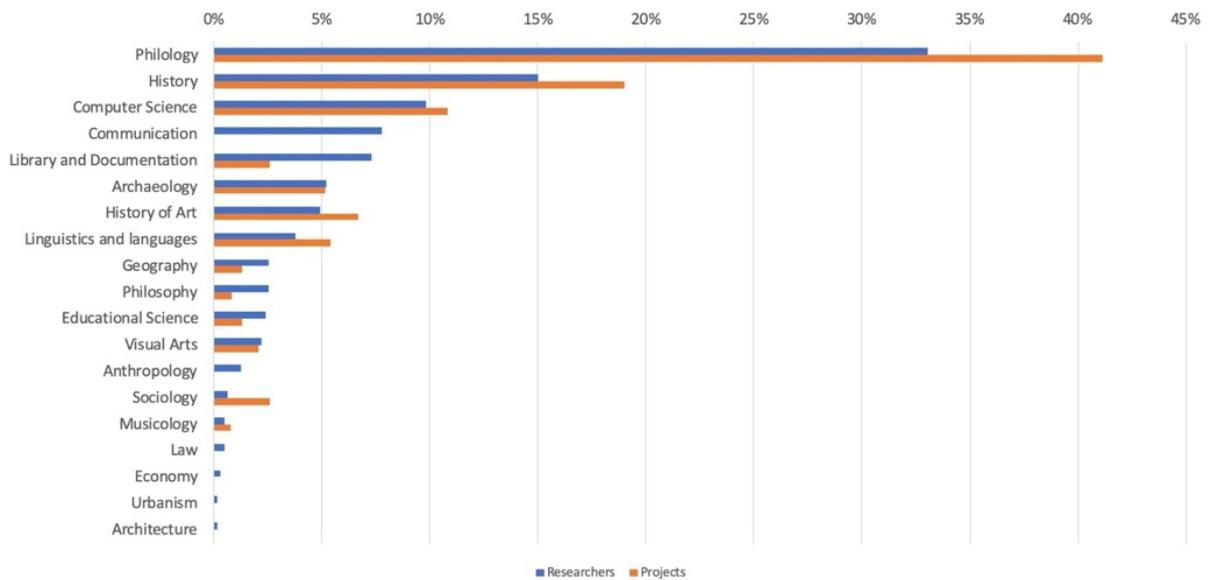


Figure 5. Comparison between researchers and projects disciplinary backgrounds.

Researchers have also been classified according to up to five research topics (Figure 6). This classification proved to be risky and with a relatively high degree of subjectivity, mixing topics that are more or less relevant to the discipline. However, this provided an overview of the topics that attracted a higher or lower number of researchers. A striking number of researchers were found to be working in the field of Digital Literary Studies, something previously observed in Spain (Fiormonte, 2014; Aibar-Puentes, 2018).

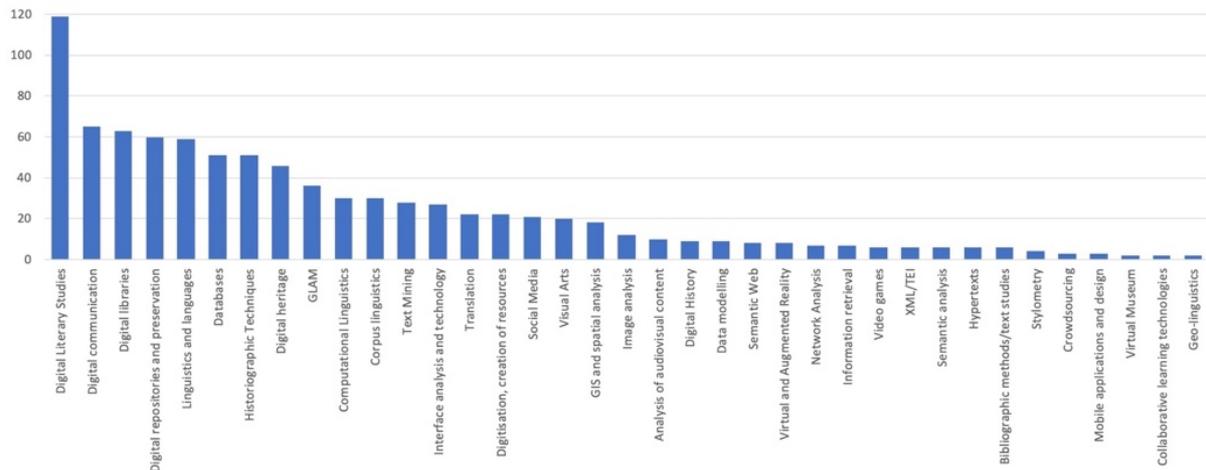


Figure 6. Researchers classified according to their preferred research topics, with up to five each.

Connecting disciplines with research topics, both for researchers and projects, provided interesting results (Figure 7). The resulting graph, composed of 75 nodes and 324 edges, presents five communities that were identified by using Leiden algorithm (Traag *et al.*, 2019). They are grouped in five major disciplinary groups: History (32% of nodes), Philology (24%), Communication (20%), Computer Science (17.3%) and Documentation (6.7%). Computer Science, despite being the fourth community in terms of nodes, is the discipline connected to the highest number of research topics

(13.3%). This was hardly surprising, considering its transversal role in the framework of Digital Humanities. Very few topics are linked to only one cluster, whereas the vast majority are connected to two or more disciplinary areas. To visualize this high degree of interdisciplinarity, we isolated egocentric networks (Golbeck, 2013) for (A) Text mining; (B) XML/TEI; (C) Video games; and GIS (D) (Figure 8). Both A and B are consolidated topics in Digital Humanities that, despite being closely related to Philology, are connected to four different clusters. GIS and spatial analysis, a topic that has become very popular in Archaeology since the late 1970s, is now gaining traction in other fields such as History of Arts and Philology, where it has been respectively used, for example, to locate the distribution of artwork or to trace exiled writers. Video games is also a very recent trend that has gained attention as a research topic in Education, Communication, and Humanities because of its potential in terms of three-dimensional digital representation and user engagement.

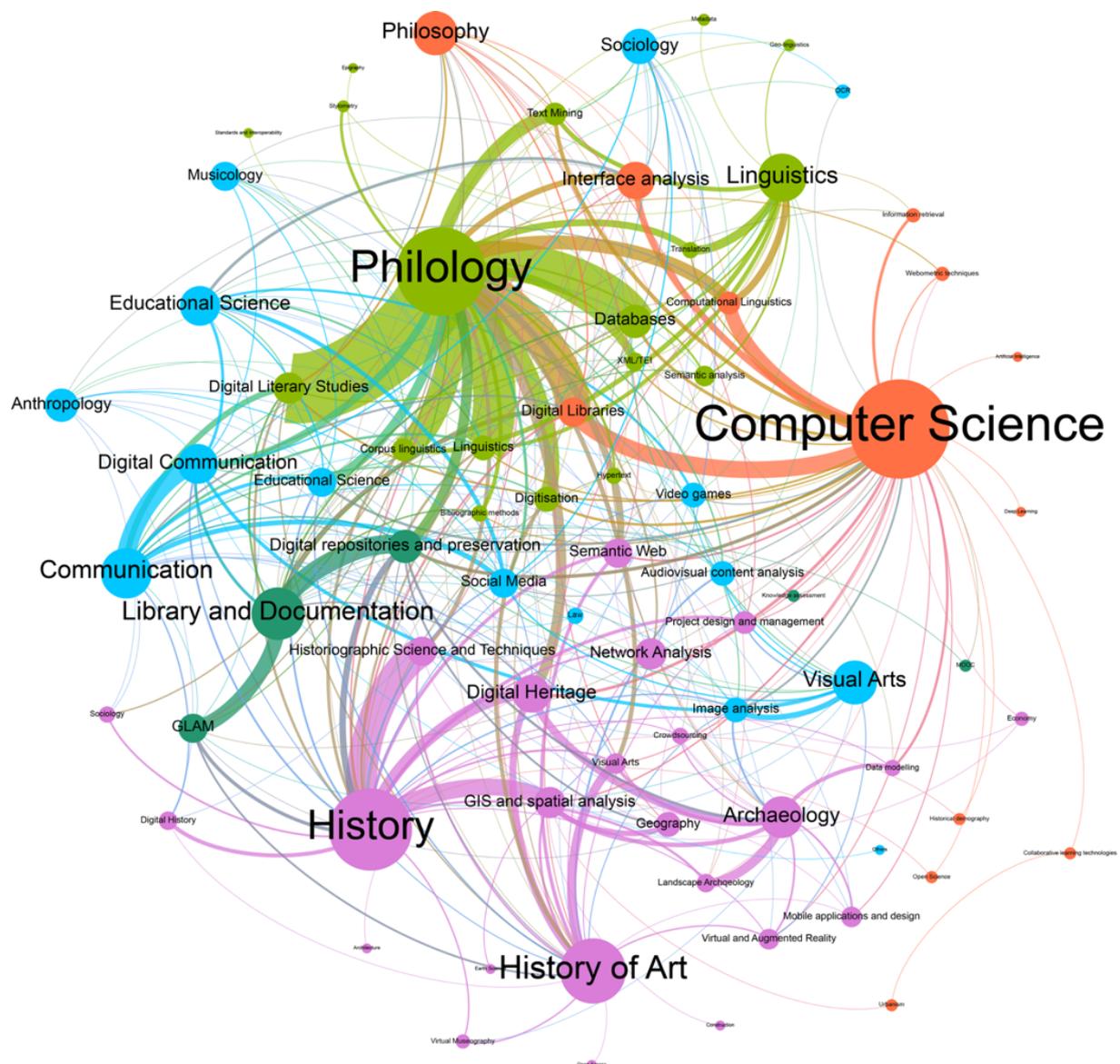


Figure 7. Network analysis of disciplines and research topics. Node size corresponds to the total number of connections, colour to the community, and edge thickness to the degree of connectivity between the two nodes.

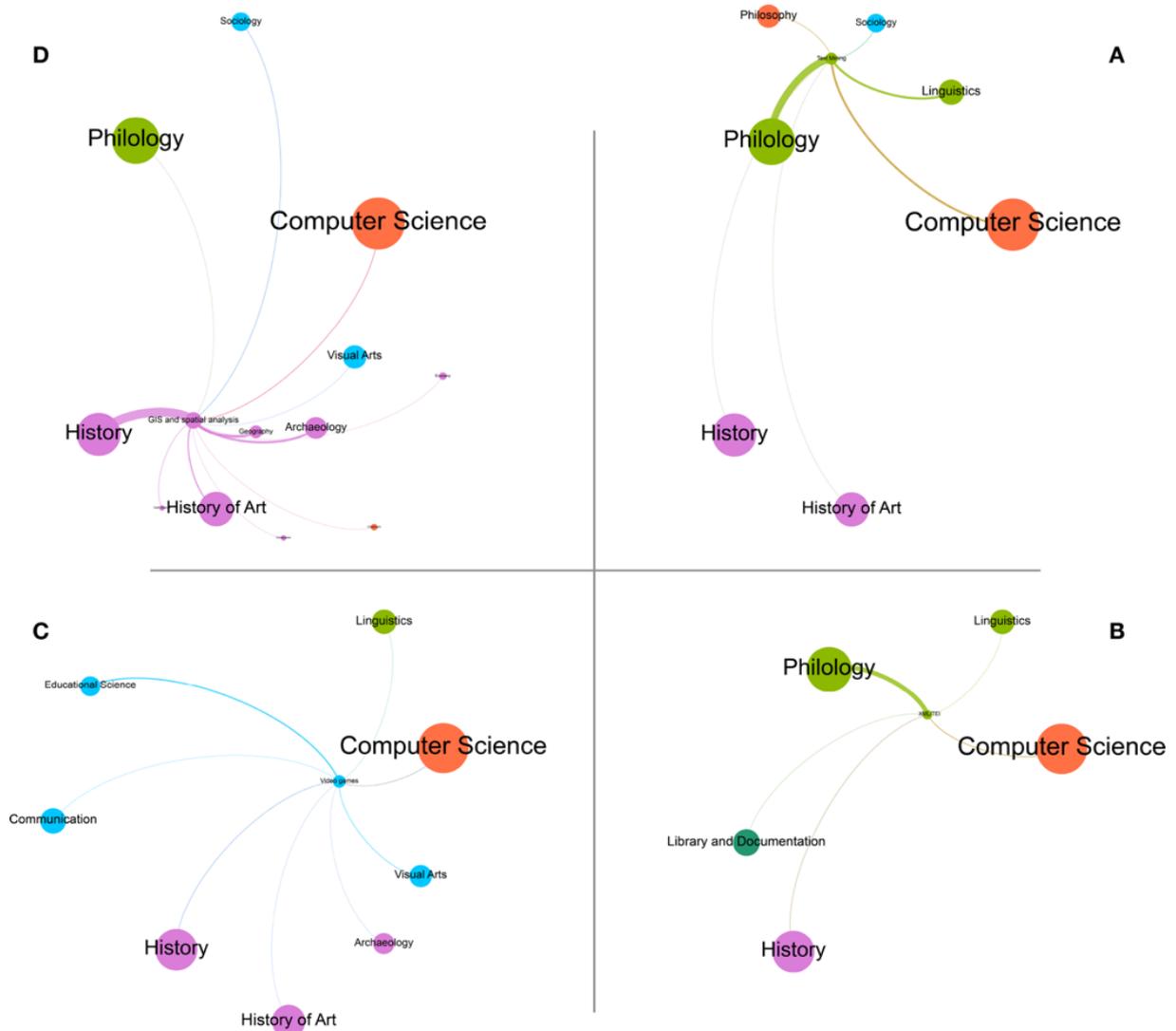


Figure 8. Ego networks derived from Figure 7, for four research topics: (A) Text mining; (B) XML/TEI; (C) Video Games; (D) GIS and spatial analysis.

3.2 RESEARCH AND HIGHER EDUCATION CENTRES

Another type of important information extracted from researcher records is their institutional affiliation (Table 3). This allowed us to draw a map of the distribution of research and higher education centres in Spain. Remarkably, half of the researchers (49.2%) belong to only nine institutions, whereas the other half is scattered across 84 institutions, with 41 who are the only digital humanities researcher in their centre. In total, we recorded 93 institutions, with a mean of 6.2 researchers each, median of 2 and a standard deviation of 10. The relatively high standard deviation can be explained by the fact that the top 3%, represent 24.7% of all researchers in the dataset. The distribution shows a marked asymmetry because 44% of the institutions only have one researcher associated and 63% of the total between one and three.

| Research and Higher education centre | # of researchers | Percentage |
|--|------------------|------------|
| Universidad Complutense de Madrid | 62 | 10.7 |
| Universidad de Barcelona | 41 | 7.1 |
| Universidad de Granada | 40 | 6.9 |
| Universidad de Santiago de Compostela | 28 | 4.8 |
| Universidad Nacional de Educación a Distancia (UNED) | 28 | 4.8 |
| Consejo Superior de Investigaciones Científicas (CSIC) ¹⁴ | 26 | 4.5 |
| Universidad Autónoma de Barcelona | 22 | 3.8 |
| Universidad de Salamanca | 19 | 3.3 |
| Universidad de Sevilla | 19 | 3.3 |
| TOTAL | 285 | 49.2 |

Table 3. The nine major centres working on Digital Humanities in Spain, according to the number of affiliated researchers.

Considering only this subset of institutions representing half of all researchers in Digital Humanities in Spain, we classified the relative weight of each department (Figure 9)¹⁵. According to the chart, even though Philology is the department with the most researchers in almost every case, Philology is more predominant at the Complutense of Madrid or Santiago de Compostela than, for example, Seville or Granada. The University of Granada has the most variation with researchers from 13 different departments.

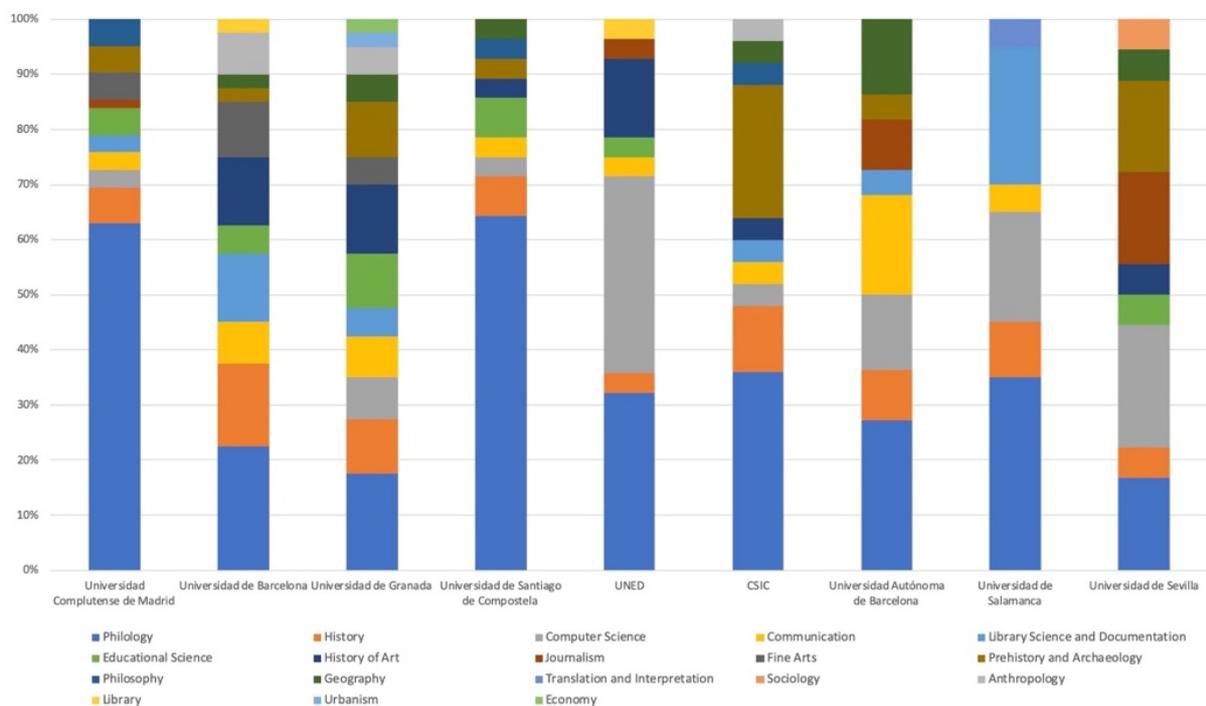


Figure 9. Relative weight of each department in the nine main research centres. To favour comparison, each group of values is stacked at 100%.

Other striking aspects include the weight of Computer Science at the UNED, Archaeology at the CSIC and Library Science at Salamanca. Philology and History are the only two departments with researchers in all nine institutions. It should be highlighted that the CSIC is an anomaly. Although it was regarded as a single institution, its

researchers actually belong to eight different centres: 14 to the *Centro de Ciencias Humanas y Sociales* (CCHS) in Madrid, whereas the rest work in research institutes in Santiago de Compostela, Barcelona, and Granada.

The following map presents an overview of the distribution of researchers in Spain (Figure 10). As mentioned in section 2.2, the map also includes 14 researchers working in Digital Humanities at foreign institutions, but linked to a Spanish funded research project. The map was produced by first geocoding to the coordinates of the address of the institutions. Centroids were then created of all points sharing the same location. The resulting layer was styled, both for point size and point colour, using the Jenks natural breaks classification method, in order to reduce the variance within classes and maximize the variance between classes (Jenks, 1967).

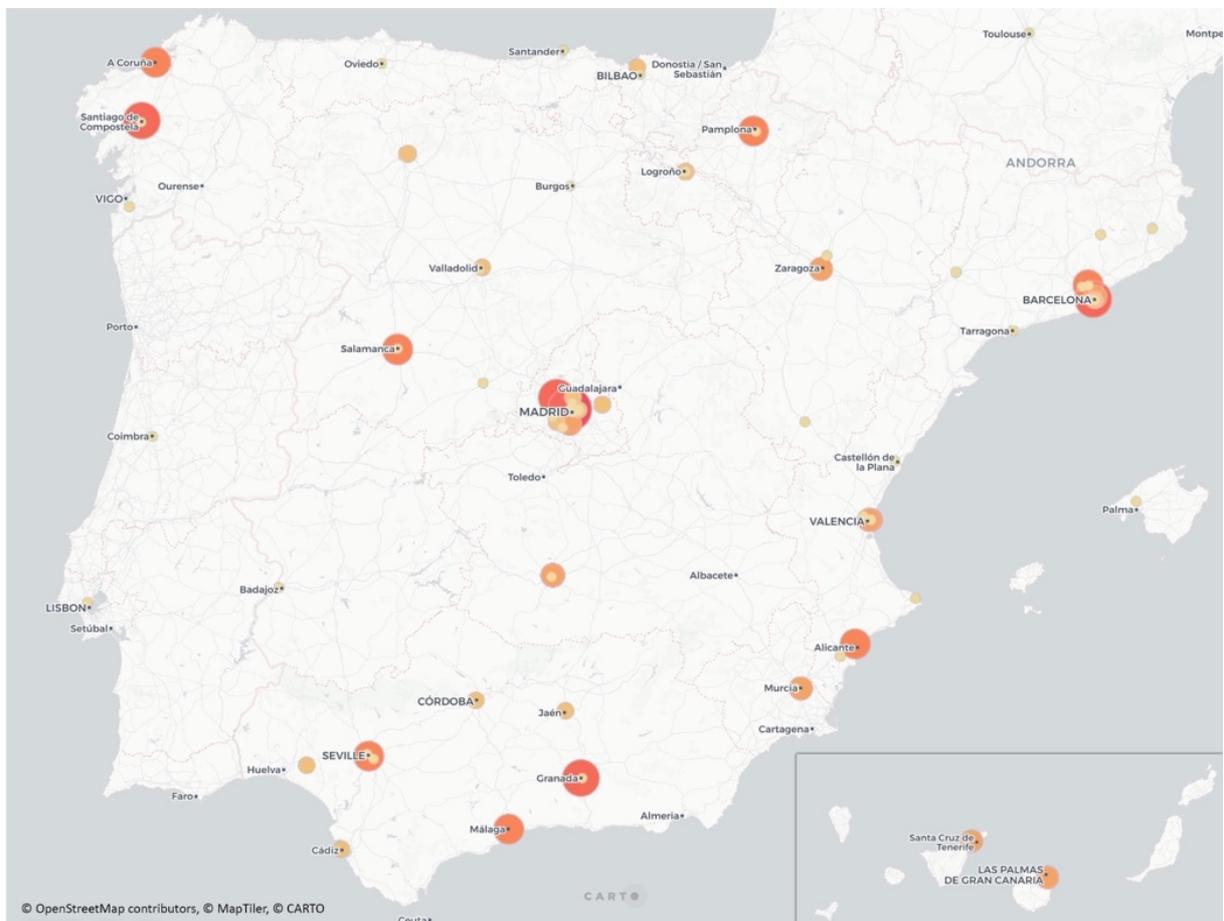


Figure 10. Overview of the distribution of researchers in Spain, produced by geocoding the affiliated institutions address and then creating centroids of points sharing the same location. Online map available at <http://sl.ugr.es/DHmap>.

Considering this scenario, it comes as no surprise that more than half of the masters focused on Digital Humanities currently available in Spain (Table 4) are offered by this top group of research centres¹⁶. When they do not have a wide scope, they reflect the relative weight of their departments (Figure 8). In fact, five embrace Digital Humanities from a broad spectrum, while the rest focuses on more specific topics, such as the following: digital content management, digital heritage, natural processing language, digital literature, cybercrime, communication, and digital education. The first master

of this kind in Spain was launched, probably ahead of its time, by the *Universidad de Castilla la Mancha* in 2006 and lasted until 2011.

| Master title | Organising institution | Since |
|--|---|-------|
| Humanitats i Patrimoni Digitals ¹⁷ | Universidad Autónoma de Barcelona / Universidad Pompeu Fabra | 2018 |
| Máster en Historia y Humanidades Digitales | Universidad Pablo de Olavide | 2019 |
| Tecnologías de la información para la Sociedad digital: Humanidades y Ciencias | Universidad Nacional de Educación a Distancia UNED | 2019 |
| Language Analysis and Processing | Universidad del País Vasco | 2019 |
| Máster de Gestión de Contenidos Digitales | Universidad de Barcelona | 2019 |
| Màster d'Humanitats Digitals | Universidad de Barcelona | 2019 |
| Máster en Letras Digitales: estudios avanzados en textualidades electrónicas | Universidad Complutense de Madrid | 2019 |
| Máster Universitario en Humanidades Digitales | Universidad Internacional de La Rioja UNIR | 2020 |
| Humanidades Digitales para un mundo sostenible | Universidad Autónoma de Madrid | 2020 |
| Máster en Humanidades Digitales | Universidad Nacional de Educación a Distancia UNED | 2021 |

Table 4. Masters focused on Digital Humanities currently available in Spain.

This wide selection of new courses is rapidly changing the educational offer in Digital Humanities in Spain. However, it is still limited to postgraduate training, and thus does not introduce the discipline to the new generations of undergraduate students pursuing degrees in Humanities. The impact that these courses will have on opening new lines of research still remains to be seen. Other important issues are how they will evolve and adapt to the evolution of the discipline, not to mention, the kind of digital humanities profile envisaged. For example, the question is whether this profile will be more focused on proficiency in the use of digital tools, on critical DH, or on engineering and development, or on a combination of some or all of these. Even though it is still too early to judge its degree of employability as well as its level of knowledge transfer to the private sector and civil society, in our opinion, these will be the distinguishing factors that will make a difference in their competitiveness and eventual survival.

3.3 RESEARCH PROJECTS

Another way to analyse the evolution of Digital Humanities research in Spain is to focus on research project funding¹⁸. Out of 368 project records, 337 have data about funding, for a total of 21,673,676 €, with a maximum and minimum amount, respectively, of 2,300,000 € and 1,200 €, a mean of 64,313 €, and a median of 42,350 €. Once ranked, the distribution shows a clear majority of small to medium-size projects, with very few exceptions (Figure 11). Half of the projects received less than 42,350 € in funding, with over 90% receiving less than 100,000 €. Considering that projects last an average of 2.8 years, the resources available per year are quite limited. Around this core group of small to medium-size projects, there are also a very few large projects and a

relatively high number of micro-projects, 18 in total, with funding of up to 5,000 €. These micro-projects were mostly focused on the development of a minor digital resource or the digitalization of a small collection, rather than on pure research.

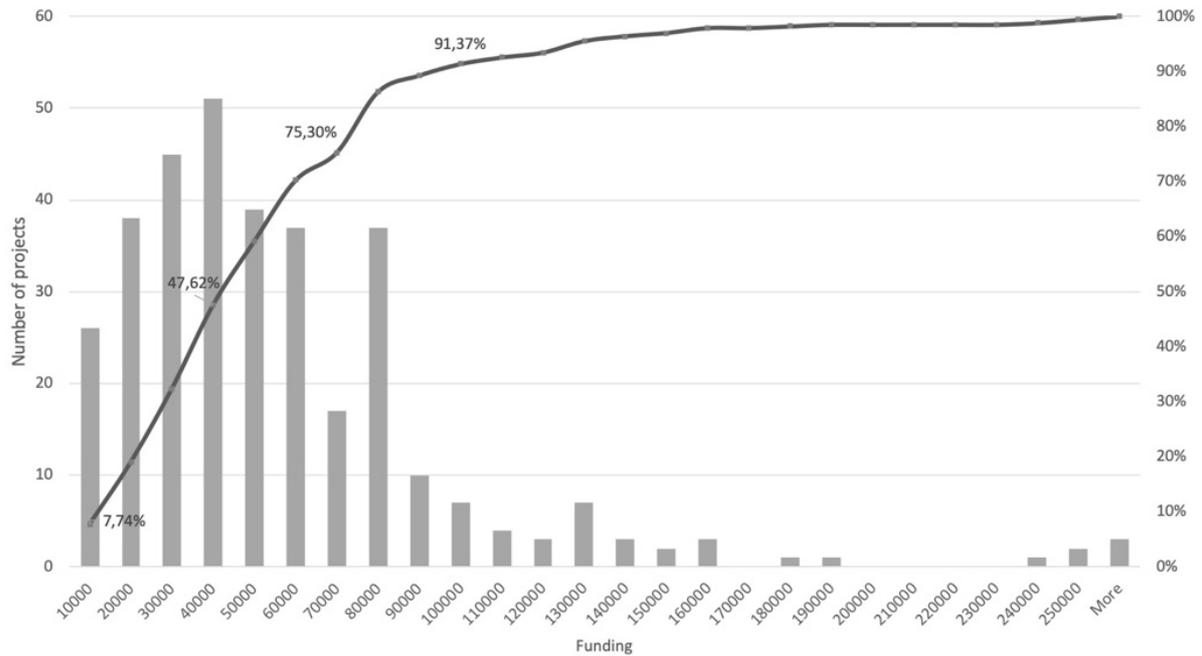


Figure 11. Distribution of projects according to funding received, with cumulative curve.

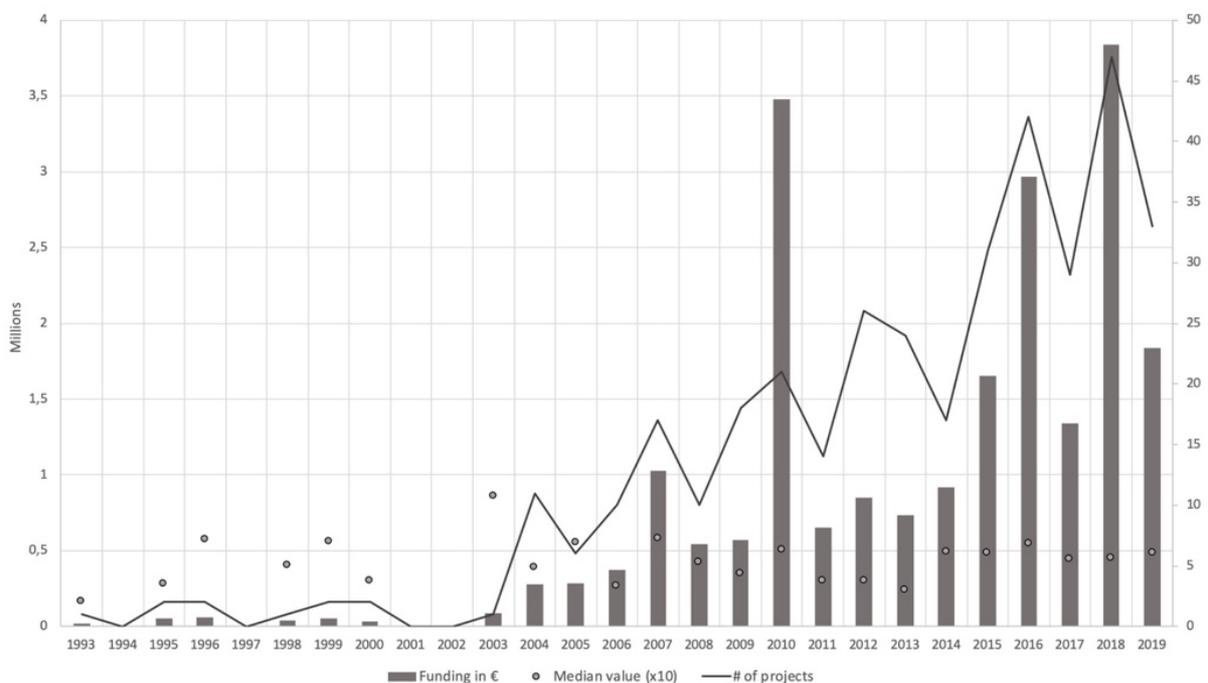


Figure 12. Number of research projects and total funding received during the period (1993-2019).

The same data classified by the starting date of each project helps to understand the evolution of funding and its effects on the discipline (Figure 12). Based on the plotting data of 337 projects over the past 26 years, the following three stages can be observed: (1) [1993-2003], the initial phase; (2) [2004-2014], a period characterised by a rapid

growth in the number of projects, followed by a steady increase; (3) [2015, 2019] a five-year period during which there was a substantial increase in funding, closely related to the increase in the number of funded projects. This is reflected in the fact that there is only a slight fluctuation in the median value¹⁹, which ranges from 57,3970 € to 24,000 €, with an average value of 43,334 €.

The trend of the broken line, with peaks and troughs occurring every three years, corresponds approximately to the average life cycle of the research projects. This is confirmed by the fact that many of the proposals were generated by only a few groups to support their research lines. Looking at the principal investigators of these projects, various examples confirm this pattern. In general, there is a very strong correlation between number of projects and total funding (Pearson index of 0.82), with only two exceptions. For 2010, the peak is caused by a very large network project on the heritage of classical Spanish theatre, coordinated by the University of Valencia. In 2016 and 2018, the anomalies are caused by the ERC Starting Grant POSTDATA²⁰, coordinated by the UNED, and by the ERC Starting Grant MapModern²¹, coordinated by the UOC.

In our data, the initial pioneering phase is characterized by discontinuous funding, with projects mainly funded by local institutions and led basically by two researchers: Juan Manuel Vicent García, from the History Institute of the CSIC, and Sagrario López Poza, principal investigator of the BIDISO research group at the *Universidade da Coruña*. The latter, in particular, led six of the eleven projects during the pre-2004 phase. The role that other researchers, such as Francisco Marcos Marin, Antonio Rodríguez de las Heras, Concha Sanz Miguel or José Manuel Lucía Megías, played in the introduction of Digital Humanities in Spain has been analysed in previous work to which we refer for a complementary view (Rojas-Castro, 2013; González-Blanco, 2013; Spence; González-Blanco, 2014).

During the second phase, when the discipline expanded nationally, there was an increase in the number of institutions (from two to 10 and then to over 20). The Ministry of Science became the main funding agency since it allocated resources for 89% of the projects in this period. This is a 27% increase from the previous period. There were no more gaps in the sequence (i.e. years with 0 projects financed). During this phase, in 2012, the *Asociación de Humanidades Digitales Hispánicas*²² (López-Poza, 2014) was created, whose first international congress was held in 2013.

The transition between the second and third phases (2013-2015) in Spain was a significant step of nationwide awareness and interest in the role of Digital Humanities. It was hardly an accident that it coincided with the organization of a series of initiatives (i.e. DíaHD (Robles-Gómez *et al.*, 2016), GRISO congress in A Coruña, THATCamp in Granada, Digital Humanities conference in Salamanca, the creation of the Spanish CLARIN K Centre²³, etc.) and with the publication of various studies (Rojas-Castro, 2013; González-Blanco, 2013; Ortega; Eunice-Gutiérrez, 2014; Baraibar-Echeverria, 2014; Romero-Frías; Del-Barrio-García, 2014; Spence; González-Blanco, 2015), which

reflected on and further consolidated this new discipline. These important initiatives laid the groundwork for the third phase.

This third period (2015-2019) is characterized by a veritable explosion in the number of funding agencies, which rose from two or three per year to nine. In economic terms, there seems to be a supply-driven market that stimulates the emergence of research proposals in Digital Humanities, partly to take advantage of the funding opportunities. In this period, thematic calls emerge at different levels: Fundación BBVA (national level), RecerCaixa (regional level, Catalunya), and Medialab UGR (institutional level, UGR). Interestingly, the BBVA call, initially launched for Humanities in 2014, changed its name in 2015 to Digital Humanities. Another characteristic of this third phase was a more equal gender distribution of resources, compared to the higher prevalence of resources (70% of the total) allocated to male principal investigators in the second phase.

The rapid growth in the use of the “Digital Humanities” label in this period transformed the term from an unknown expression to a buzzword. However, all too often, it was indiscriminately applied to any kind of project that minimally represented content on the Web. This was probably caused by the late and sudden arrival of digital transformation in a context that had not undergone a parallel methodological innovation. This phenomenon led to a certain deterioration in the essence of Digital Humanities as a discipline since it attracted many researchers whose interests and scientific contributions were not related to Humanities, and had no genuine multidisciplinary interest. Nonetheless, this phenomenon, which is also a consequence of the attractiveness of the discipline itself, will eventually lessen in importance.

Regarding the funding agencies during the whole period (1993-2019), the role of public national funding, (especially, the Ministry of Science²⁴) is evidently predominant, despite the variety of funding entities and their increase over the last five years. Both in number of projects (Figure 13) and total funding (Figure 14), the ministry is by large the main funding agency with 77% funded proposals and 72% of the allocated resources.

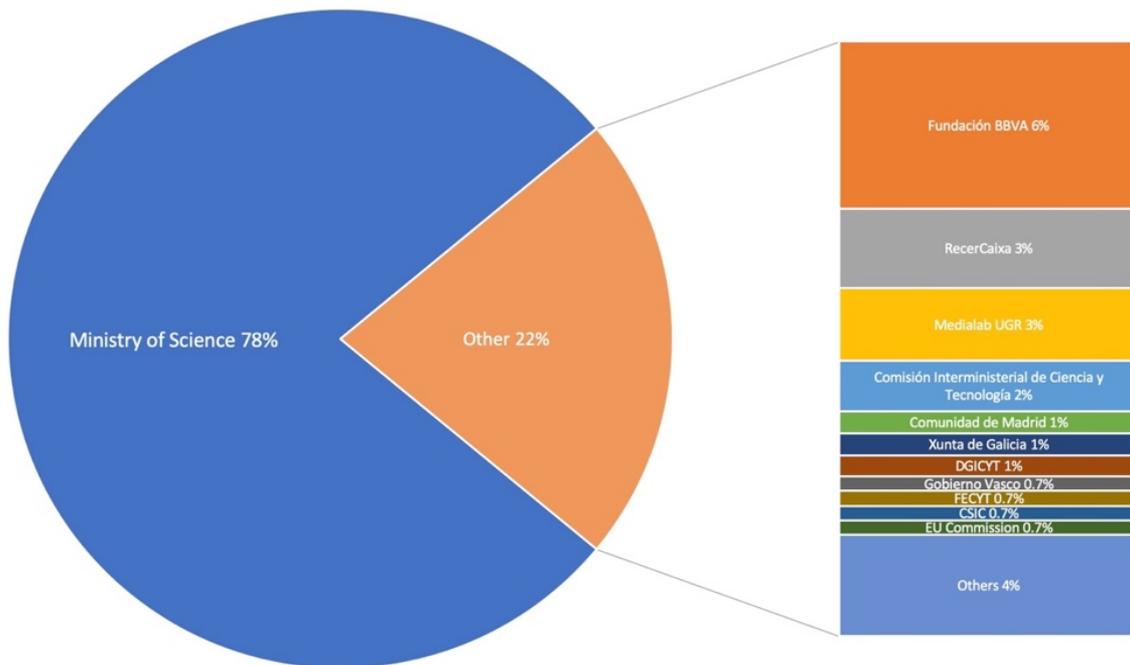


Figure 13. Funding agencies according to number of projects.

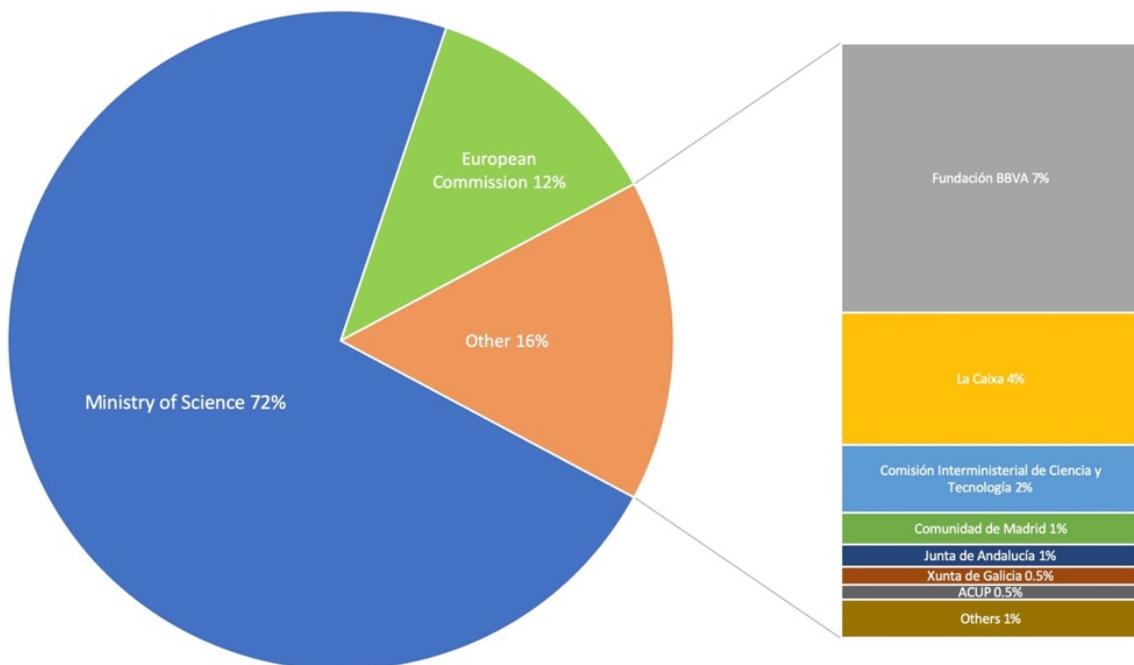


Figure 14. Funding agencies according to total funding.

The next step was to examine the geographical distribution of funding, first by research centre and then by region. When research institutes and higher education centres are sorted by number of projects (Table 5 and Figure 15), this produces a slightly different picture, in comparison to when they were sorted in terms of affiliated researchers. Although there are various similarities, there are also significant differences. The following seven institutions remain at the top of the list though in different positions: the UB, UCM, UGR, UAC, UNED, CSIC, and USC. The *Universidade da Coruña* and *Universitat Pompeu Fabra* move up on the list and catch up to Salamanca and

Seville, who remain immediately behind them. This confirms the accuracy of the first ranking.

| Acronym | Research centre | # of projects | Funding |
|---------|---|---------------|-----------|
| UB | Universitat de Barcelona | 33 | 1,783,391 |
| UCM | Universidad Complutense de Madrid | 30 | 1,321,786 |
| CSIC | Consejo Superior de Investigaciones Científicas | 25 | 1,131,889 |
| USC | Universidade de Santiago de Compostela | 23 | 957,839 |
| UGR | Universidad de Granada | 21 | 812,610 |
| UDA | Universidade da Coruña | 15 | 849,892 |
| UAB | Universitat Autònoma de Barcelona | 15 | 675,461 |
| UNED | Universidad Nacional de Educación a Distancia | 15 | 1,407,619 |
| UPF | Universitat Pompeu Fabra CCT | 14 | 579,920 |
| USAL | Universidad de Salamanca | 10 | 498,276 |
| US | Universidad de Sevilla | 10 | 455,270 |
| UCO | Universidad de Córdoba | 9 | 551,320 |
| UA | Universidad de Alicante | 9 | 496,720 |
| UMA | Universidad de Málaga | 8 | 300,430 |
| UV | Universitat de València | 8 | 2,697,028 |
| UVA | Universidad de Valladolid | 8 | 251,012 |
| ULPGC | Universidad de Las Palmas de Gran Canaria | 8 | 289,950 |
| CVC | Centro de Visión por Computador | 6 | 639,844 |
| UPV | Universitat Politècnica de València | 6 | 453,608 |
| UAH | Universidad de Alcalá | 6 | 293,425 |
| UCA | Universidad de Cádiz | 5 | 118,650 |
| UCLM | Universidad de Castilla-La Mancha | 5 | 231,750 |
| UOC | Universitat Oberta de Catalunya | 4 | 1,616,268 |

Table 5. Research and higher education centres ranked by number of projects and total funding.

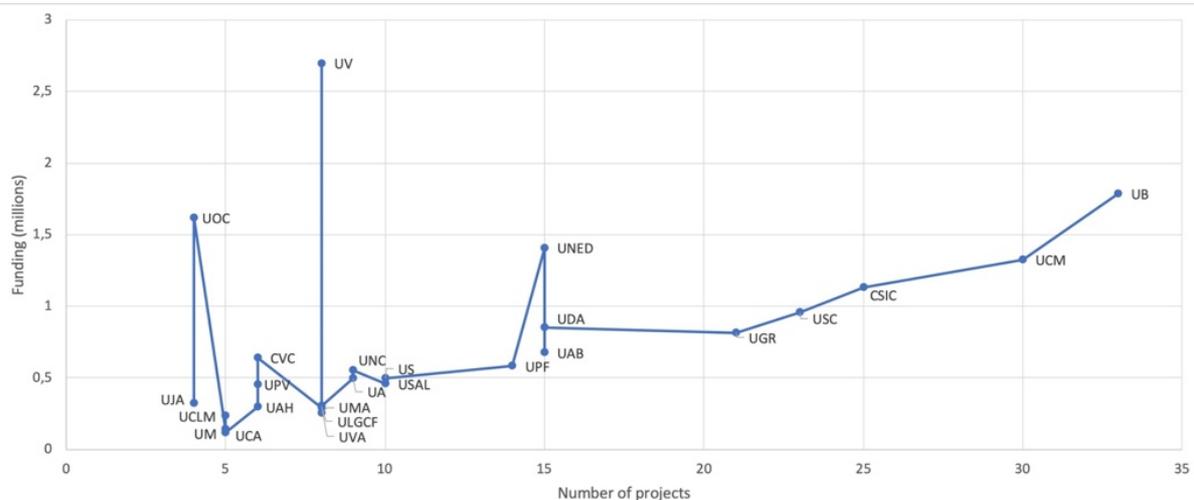


Figure 15. Scatterplot of first twenty-five research centres by number of projects and total funding obtained during the whole analysed period (1993-2019).

In reference to the geographical distribution of funding for the 17 regions in Spain (Figure 16), the three territories that most stand out are Cataluña, Comunidad de Madrid, and Comunidad Valenciana. Cataluña, who was already in the top group in phase two, consolidated its leadership in phase 3 and almost doubled the resources obtained (+89%). Madrid also showed a significant increase (+155%) and almost tripled its funding. Andalucía, Galicia, and to a lesser extent Castilla y León, also received considerable resources, but maintained very similar funding between phase two and three. The other regions lag behind at some distance. With the exception of Valencia, Aragón, País Vasco, Asturias, and Navarre, almost all regions obtained more funding during the third phase, despite being half as long.

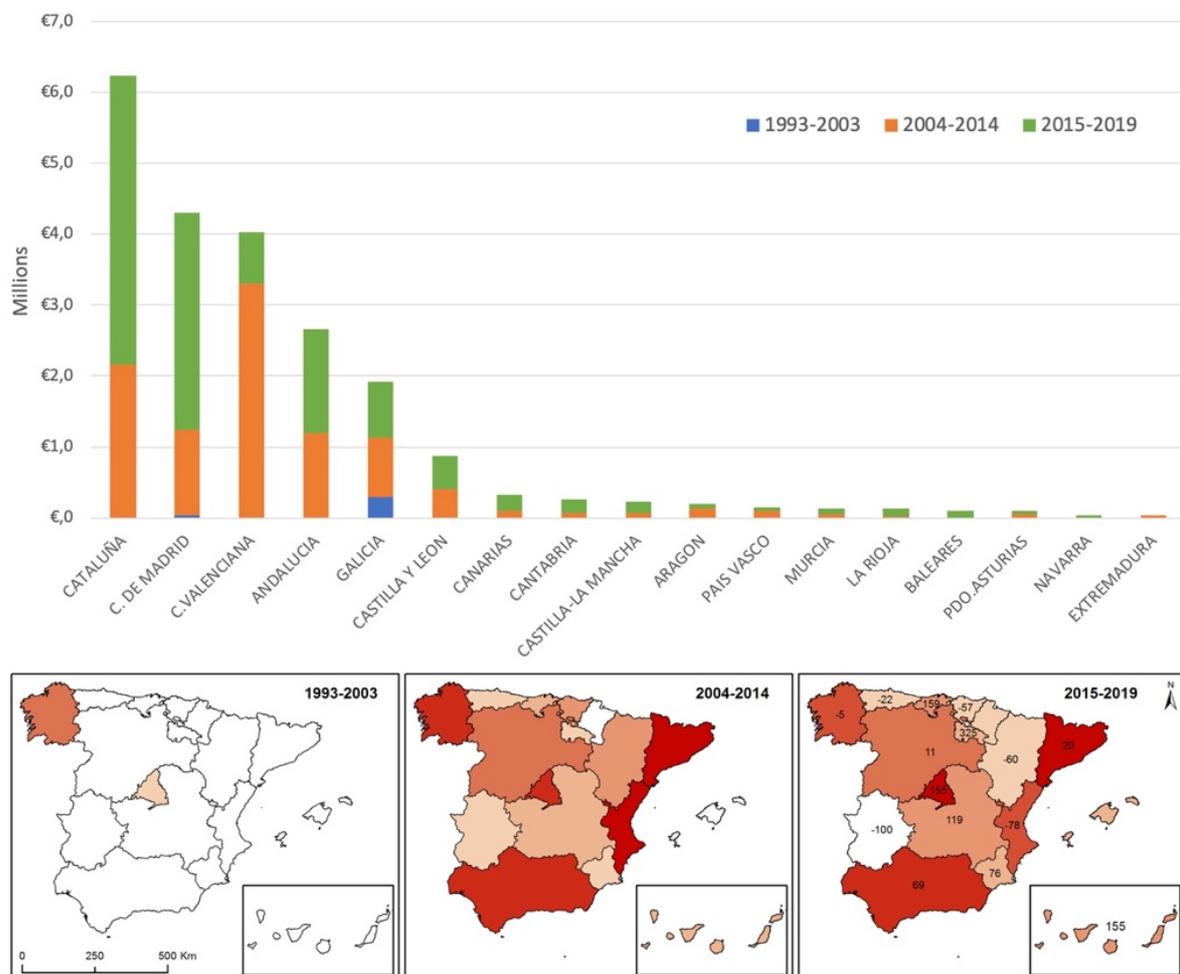


Figure 16. Geographical distribution of funding across Spanish regions during the previously identified three phases.

3.4 DIGITAL RESOURCES FOR RESEARCH AND DISSEMINATION

Digital resources have always been a key element of Digital Humanities, as they are often the final results or the starting point of many research projects. They normally involve generating, collecting or digitizing data that other scholars can reuse for research and training. The value of building and creating digital resources and objects in the Humanities (Ramsay; Rockwell, 2012; Endres, 2017) is still an issue of some debate, but, particularly in Spain, it is a fact that building, hacking, and coding (i.e. a

database, a map, a collection of 3D objects, a text analysis tool) is not considered part of normal research activity and counts for very little or none in the evaluation of academic merits by the ANECA (*Agencia Nacional de Evaluación de la Calidad y Acreditación*).

This research study collected and analysed a total of 88 digital resources developed for research or dissemination. These resources are all web-based, given that desktop solutions or Intranet systems were impossible to identify at this stage. When they are sorted by publishing date (Figure 17), their rate increases significantly over time. This is similar to the trend observed for projects (Figure 12). Only eight resources date back to the pioneering phase (pre-2004), whereas one out of three was published after 2014, though there is a certain decline after the peak in 2014. One possible explanation can be found in a shift in awareness from the creation of customised tools to the willingness to participate in shared international platforms, which favour standards and interoperability. This tendency is clearly harmed by the lack of Spanish participation in DARIAH and CLARIN, the two main European infrastructures dedicated to Digital Humanities.

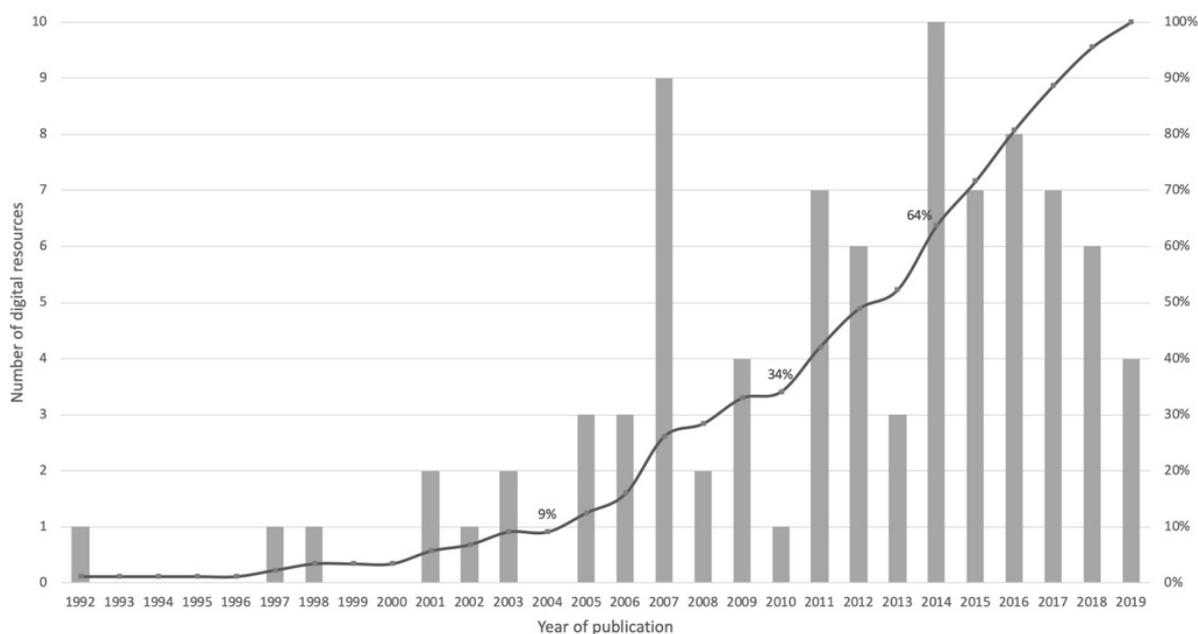


Figure 17. Histogram showing digital resources sorted by publishing date, with cumulative curve.

Digital resources were classified by type, depending on their definition and presentation on the Web, and consistency between resources was also assured (Table 6). Some of these definitions are synonyms or subgroups of each other, though each one has its own peculiarities. This diversity should be maintained because the differences often reflect the disciplinary backgrounds of those promoting the initiative. In such cases where the same resource can be classified in different ways (e.g. as a map, catalogue, and crowdsourcing platform), its classification was based on its main purpose as stated in the project description.

The largest group (72.4%) across all disciplines corresponds to Web information systems. These are platforms that use a Web browser as front-end and a database as back-end to store and retrieve texts, documents or objects (i.e. repositories, corpora, digital libraries, catalogues). This typology is extremely popular and has a long tradition especially in Philology, with examples from the whole chronological spectrum under study. Another popular category is Web portal, a term generally used for websites providing access to different internal or external services, such as the search aggregators ARACNE and LITNET, or for resources oriented toward dissemination and communication rather than research (i.e. Medievalitis).

A distinctive approach is taken by crowdsourcing platforms, developed to foment collaboration with society in collecting research data or improving existing datasets (Toscano, 2018), and Web applications, oriented towards analytical functionalities (e.g. the Pathfinder tool). These categories tend to be more popular in disciplines such as History, Archaeology, and History of Arts, which have seen a significant boost in recent years, with 21 out of 23 artefacts published after 2014.

| | Philology | Linguistics | History | Archaeology | Art History | Heritage | Musicology | Documentation | Communication | Education | TOTALS |
|-----------------|-----------|-------------|----------|-------------|-------------|----------|------------|---------------|---------------|-----------|-----------|
| Digital Library | 12 | | | | | | 1 | 1 | | | 14 |
| Database | 17 | 2 | 3 | 2 | 1 | | 1 | | | | 26 |
| Repository | 3 | | 2 | | 1 | 4 | | 6 | | 1 | 16 |
| Corpus | 1 | | | | | | | | | | 1 |
| Catalogue | 1 | 1 | | 2 | 1 | 1 | | | | | 6 |
| Crowdsourcing | | | 1 | 1 | 1 | | | | | | 3 |
| Portal | 4 | 1 | 2 | 1 | 1 | 1 | | 1 | 1 | | 12 |
| Dictionary | | 1 | | | | | | | | | 1 |
| Map | | | | 1 | | | 1 | | 2 | | 4 |
| Web app | | 1 | | | 1 | | | | | | 2 |
| Mobile app | | | | 2 | | | | | | | 2 |
| TOTALS | 38 | 6 | 8 | 9 | 6 | 6 | 3 | 8 | 3 | 1 | 88 |

Table 6. Digital resources classified according to discipline and typology.

The sustainability and long-term compatibility of all these digital resources is an important issue, both in terms of technological obsolescence and content update. We detected cases where a long series of projects, promoted by the same research group, contributed to the development and maintenance of one or more digital resources (i.e. the BIDISO group promoted the publication of eight digital , thanks to seven research projects, from 1992 to 2019; the *Biblioteca Digital Ovidiana* was supported by five different research projects with the same PI; the *Biblioteca Virtual de la Filología Española* was the result of three different research projects sharing PI and research group). Another

noteworthy example of various projects clearly oriented toward the creation and maintenance of a set of digital catalogues and tools is *TRACE (TRAducción y CEnsura)*, promoted by researchers in Linguistics belonging to four different institutions (the Universidad de Elche, Universidad de León, Universidad del País Vasco, and Universidad de Cantabria).

3.5 DETECTED ISSUES

Despite being a preferred topic of research and funding, Digital Humanities in Spain also suffers from several problems, some of which are reflected in our data: fragmentation, lack of standards, lack of visibility (affecting the discipline worldwide), and lack of appreciation.

Fragmentation

The structure of research groups that predominates in the Humanities is generally small and fragmented. The rigid organization of university departments combined with the variety of disciplines involved (Figure 9) fosters fragmentation. A possible solution would be the creation of cross-cutting and cross-disciplinary laboratories²⁵ and thematic institutes.

Data shows that only a fraction of the mapped projects invested in digital resources and published Open Data, which are among the most effective ways to share results and foster communication and collaboration with other projects using technology for their research.

As seen in Section 3.3, investment is also quite parcelled out and median funding is low, which implies granularity. One way to increase funding, clustering, and international collaboration is by promoting participation in large European consortium projects, in the framework of the coming Horizon Europe Programme.

Lack of standards

One consequence of this lack of shared methodology and tools is the lack of standards. While scientific disciplines, in general, have created and followed standards for research and data structuring, in the Humanities, standards are rarely used to facilitate the sharing of information between different projects, languages or groups. Research results tend to be very local, and interchange standards poorly applied or assumed. The worldwide community in Digital Humanities has developed important standardization initiatives, often in collaboration with libraries. Such initiatives include the TEI (Text Encoding Initiative), the encoding system for humanities texts, and CIDOC-CRM, an ontology for cataloguing works of art and museums. Compared to other European countries, the knowledge and application of these tools are quite limited in Spain. According to our research, only a very limited number of researchers focus on topics related to standards, such as the Semantic Web (7), Data Modelling (9) or TEI (4).

Lack of visibility

The dissemination of results from Spanish researchers in Digital Humanities is still very limited. Their voice is scarcely heard in the international circles that lead the discipline (associations, international congresses, digital humanities infrastructures, etc.). This situation is intensified by the lack of combined strategic support from universities. In terms of publications, the number of specialized journals is low. Those supported by Spanish institutions are mostly new and have low impact factors, which affects the visibility of the papers. In this research, we identified eight journals (Table 7), which explicitly cite Digital Humanities or Humanities in the Digital Age as one of their preferred topics. All but two are less than ten years old, and none belongs to any of the high-impact reference indexes (i.e. FECYT C1, SJR Q1, CIRC A+ or JCR). The shortage of journals with high impact factors is not due to the quality of the papers or to the lack of rigour in the publication process, but rather to the classification of disciplines used by public indexes²⁶ and of the lack of recognition of Digital Humanities as a discipline in itself. For example, a journal on the intersection of Arts and Computer Science will be indexed both under “Arts” and under “Computer Science”, where it will compete against journals with a narrower focus, generally more frequently cited by researchers in each discipline. If there were a category only for Digital Humanities, these journals would have a higher ranking and their papers would have greater visibility. Obviously, this issue is not limited to the context in Spain, but also affects the visibility of Digital Humanities worldwide.

| Title | Publisher | Since | ISSN | Website |
|--|-----------------------------------|-------|-----------|-----------------------|
| Digithum. A relational perspective on culture and society | Universitat Oberta de Catalunya | 1999 | 1575-2275 | digithum.uoc.edu |
| Artnodes: revista de arte, ciencia y tecnología | Universitat Oberta de Catalunya | 2002 | 1695-5951 | artnodes.uoc.edu |
| Caracteres. Estudios culturales y críticos de la esfera digital | Editorial Delirio | 2011 | 2254-4496 | revistacaracteres.net |
| Scriptum Digital. Revista de corpus diacrónicos y edición digital en lenguas iberorrománicas | Universitat Autònoma de Barcelona | 2012 | 2014-640X | scriptumdigital.org |
| JANUS. Estudios sobre el Siglo de Oro | Departamento de Letras UDC | 2012 | 2254-7290 | janusdigital.es |
| ArtyHum Revista Digital de Artes y Humanidades | ArtyHum | 2014 | 2341-4898 | artyhum.com |
| Microtextualidades. Revista Internacional de microrrelato y minificción | Universidad San Pablo CEU | 2017 | 2530-8297 | revistas.uspceu.com |
| Revista de Humanidades Digitales | UNED | 2017 | 2531-1786 | revistas.uned.es |

Table 7. Scientific journals published in Spain, which explicitly state Digital Humanities as one of their reference topics.

Lack of appreciation

Despite these issues, publications are still the major factor used by national evaluation agencies to assess research results. Unfortunately, those projects that invest a lot of resources and time in the creation of digital resources and web projects are

significantly undervalued (as mentioned in Section 3.4). The fact that the creation of such resources is not a quantifiable merit to progress up the academic ladder considerably discourages researchers, especially in their earlier stages, from creating or collaborating in these digital infrastructures, as well as supporting the publication of Open Data, despite the benefits for their research and the whole community.

Though outside the scope of this study, this lack of appreciation is also noticeable in research positions posted by universities and research centres, where disciplinary specialisation is clearly favoured over multidisciplinary knowledge and hybrid profiles. This makes them unattractive to new generations of researchers because they are not very profitable in terms of academic careers.

4. Conclusion

The objective of this research was to describe the current scenario of Digital Humanities in Spain, in terms of researchers and research centres, disciplines involved, research topics of interest, funding agencies, and the evolution of their investment over the last 25 years. Many of the previous considerations regarding the introduction and evolution of Digital Humanities in Spain are still valid today. The study presented in this paper is not an effort to rewrite the same history or make a compendium. Our objective was rather to review and update that evolution by considering other data. This research was carried out from a different perspective and with another methodology of analysis. It has produced complementary results, which provide valuable insights into the evolution of Digital Humanities. In fact, one of the characteristics that differentiates this study from previous approaches is the kind of data used to carry out the research. Our study combined large datasets of publicly available data from trusted sources with a handpicked selection of records in order to cluster information scattered over the Web.

Accordingly, the following three aspects can be highlighted: (1) most of the evidence detected by other studies has been numerically confirmed; (2) quantitative analyses of funding, a dimension almost unexplored in Humanities, proved to be extremely valuable in the assessment of a scientific discipline and its historical evolution; (3) new metrics and values were laid down that constitute a reference base for monitoring the evolution of Digital Humanities in Spain and for favouring comparisons, both over time and with other contexts at both a European and global level.

The following conclusions can be derived from this research:

1. Among researchers, 305 were males (52.9%) and 272 were females (47.1%): a ratio that, compared with the general gender balance among researchers in Spain across all disciplines (61.2% of males and 38.8% of females) or limited to Humanities (59.8% of males and 40.2% of females), reflects a relatively significant higher female presence than expected.

2. The classification of researchers by discipline showed a wide variety of backgrounds (Figure 4), with an evident prevalence of Philologists (36%), followed by Historians (16.5%) and Computer Scientists (10.8%).
3. An analysis of the connections between researcher disciplines and research topics (Figure 7) revealed five large communities (or clusters) centred on the following major disciplines: History (32% of nodes), Philology (24%), Communication (20%), Computer Science (17.3%) and Documentation (6.7%). Computer Science, despite being the fourth community in terms of nodes, is the discipline connected to the highest number of research topics (13.3%), which is to be expected considering its transversal role in the framework of Digital Humanities.
4. According to institutional affiliation (Table 3), half of the researchers (49.2%) belong to just nine institutions, while the other half is scattered over 84 locations, with 41 apparently isolated in their centres. In total, we recorded 93 institutions, with a mean of 6.2 researchers each, a median of 2 and a standard deviation of 10.
5. Department affiliation (Figure 9) shows a much more varied and complex pattern of the relative degree of specialization in each institution: i.e. the relative weight of Philology is higher at the Complutense of Madrid or at Santiago de Compostela than, for example, in Granada, where, conversely, we find researchers from a wide range of departments (13). Other strong associations are Computer Science at the UNED, Archaeology at the CSIC, and Library Science at Salamanca.
6. Our analysis of the investment in 337 research projects over the last 25 years (Figure 12) highlighted interesting trends that reflected a three-phase periodization in the consolidation of Digital Humanities in Spain. Average project funding was found to be 64,313 €, but the median was much lower, at 42,350 €; 5% of projects received up to 5000 € (micro-projects) and over 90% were within the 100,000 € mark (Figure 11).
7. In reference to funding agencies (Figures 12-13), the role of public ministerial investment was evidently predominant, despite a significant variety of funding agencies (26) and 10% of resources from the private sector. The government alone accounted for 77% of the funded proposals and 72% of the allocated resources.
8. Disciplines, such as Philology, Linguistics, and Library Science, evidenced a longer tradition than others, such as History, Archaeology, and History of Arts, in developing digital resources for research. The most frequently developed artefacts (72.4%) were different types of database (i.e. digital libraries, catalogues, repositories, etc.), whereas a shift was observed after 2014 towards more analytical or participative tools. Sustainability of the longest-lived platforms has been mostly guaranteed by a series of nationally funded projects.

In conclusion, despite the issues detected, the Digital Humanities community in Spain is mature, healthy and vibrant, and continues to grow in popularity and in interdisciplinarity. However, in order to maintain its momentum, consolidate its role in national

academia and become a protagonist in the global debate on the current and future role of the Humanities, a number of measures should be taken, at a high institutional level.

Digital Humanities should become a knowledge area that is officially recognised at the ministerial level. This will enhance its value and make it more attractive for the researchers, who wish to pursue it. This would also improve the indexing and ranking of thematic journals, at least at a national level, i.e. including Digital Humanities in the multidisciplinary block of ÍNDICES-CSIC thematic classification, used by FECYT to rank journals and award its prestigious seal of quality (**Sanz-Casado et al.**, 2020).

University students in the Humanities need a new set of technological and computing skills, both to enhance their research and to broaden their job opportunities and employability in the labour market. One way to achieve this is to revise and expand existing degree programs or to promote interdisciplinary double degrees (i.e. combining Humanities with Computer Science, Communication or Documentation). Either of these two options involves the incorporation of new teaching profiles and the adaptation of existing ones. In the short term, this would lead to the creation of new posts for teachers, researchers, technicians and technologists.

The multidisciplinary nature of Digital Humanities research makes increasingly difficult for digital humanists to conduct their research independently. Therefore, the digital turn is affecting how researchers in the Humanities publish their work, both in terms of multi-authored papers and supplementary material (code and open data) that is likely to be produced in collaboration with other peers. The reality is that Spanish evaluation agencies continue to apply an administrative limitation to the number of authors, particularly in the humanities and social sciences, despite recent evidence indicating that the works with the highest number of authors are those with the highest level of internationalization and impact (**Robinson-Garcia; Amat**, 2018).

The government should support Digital Humanities Centres and Labs (physical, virtual, and distributed) with funding, specifically for this purpose. It should also foster centres specializing in one or more areas, to complement each other and foment competitiveness both at the national and international level. This will further strengthen and consolidate the role of these labs as service centres, supporting digital research, training, resource building, and project development within their own institutions.

Finally, in the medium to long term, the government, along with universities and research centres, should arrange to jointly finance and guarantee Spain's participation in European infrastructures dedicated to Digital Humanities, such as DARIAH and CLARIN. These are part of the European Research Infrastructure Consortium (ERIC) as they maintain and implement technology-based infrastructures. DARIAH widely targets resources for research and training in Humanities and Arts while CLARIN has a narrower scope that focuses on Language and Computational linguistics. Both consortiums help researchers to explore and apply ICT-based methods to analyse and interpret their data. All economic efforts in this regard will provide a significant

investment return, which will enhance the visibility of national research at the European level and create opportunities for international collaboration. This, in turn, will increase the sustainability of digital research project results. It will also lead to the acceptance of standards and commonly used interoperable solutions, which will guarantee access to research data, tools, and services through the infrastructure. Spanish scholars and institutions will thus receive more funding within consortium projects at the European level.

Our future research will complement these metrics with a more in-depth analysis of sources, also including aspects that have been barely touched in this paper, like the international dimension of Digital Humanities in Spain. For this reason, this database will be kept active, updated, and publicly available for consultation and further analysis. This will be possible thanks to the following three activities: (1) refinement of the controlled vocabularies used (i.e. research topics and resources classification), enriched with the feedback received; (2) revision of records in collaboration with the Digital Humanities community to fill in the gaps, correct errors, and record additional variables; (3) inclusion of new records, either in collaboration with students and researchers, or by importing successive versions of bulk datasets, from the AEI and the bibliographical databases.

5. Notes

1. CenterNet, <https://dhcenternet.org/centersMapaHD> (data have been merged into GrinUGR), GrinUGR, <https://grinugr.org/open-dataand>, more recently, Red Knowmetrics <https://red.knowmetrics.org/mapaamong> others.
2. In this regard, examples include the online resource What are DH: <http://whatisdigitalhumanities.comand> the Twitter bot @DHDefined, which randomly displays one of the over 800 definitions collected from participants to the Day of DH between 2009 and 2014.
3. 1992 is the date of the oldest digital resource recorded; 1993 the starting year of the oldest project included in the dataset obtained from the AEI; 1994 the publication date of what is considered one of the first books about computing and the humanities in Spain (Marcos-Marín, 1994).
4. We are fully aware that the use of the label “humanidades digitales” is relatively recent, but because this search was focused only on identifying Spanish researchers currently active in this field, and not on analysing the bibliography itself, we decided to perform a limited search.
5. Researchers’ details were not included in the dataset provided by the AEI, because of personal data protection.
6. <https://doi.org/10.5281/zenodo.3893546>.
7. A relatively high presence of female researchers in digital humanities has been observed also in other international surveys, such as the European survey on scholarly practices and digital needs in the arts and humanities (Dallaset *et al.*, 2017).
8. <https://www.ine.es/jaxi/Datos.htm?path=/t14/p057/64-01/10/&file=d1004.px>.
9. Científicas en cifras 2017 (Puy-Rodríguez *et al.*, 2019), Charts 1.11 and 1.12, data from 2016.

10. Total researchers in 2018: <https://www.ine.es/consul/serie.do?s=186-24&c=2&nult=50Female> researchers in 2018: <https://www.ine.es/consul/serie.do?s=186-26>.
11. These data have been calculated based on percentages from charts 1.11 and 1.12 (referring to 2016), according to *Científicas en cifras 2017* (Puy-Rodríguez *et al.*, 2019).
12. Data coming from our research (referring to 2019).
13. Most projects came already tagged according to AEI classification, although some matching has been necessary to consolidate a common vocabulary.
14. A survey on researchers and projects in the field of digital humanities is ongoing (march 2020) within the CSIC and the final results, not available at the time of writing this paper, will most probably increase this figure.
15. For the CSIC, for consistency with universities departments, we used researchers' disciplinary backgrounds instead of research centres.
16. An overview of the masters currently available in Spain is provided by the *Humanidades Digitales Hispánicas* on its website: <https://humanidadesdigitaleshispanicas.es/formacion-hd>.
17. The *Universitat Autònoma de Barcelona* launched in the autumn of 2013 a *Máster en Humanidades Digitales* that is not available anymore.
18. As said earlier, in section 2.3, the vast majority of these data come from the AEI, the Spanish State Research Agency.
19. The median value is a better measure of central tendency in this case, as it reduces the effect of outliers. Values have been multiplied by ten to make them more visible on the primary axis.
20. <https://cordis.europa.eu/project/id/679528>.
21. <https://cordis.europa.eu/project/id/803860>.
22. www.humanidadesdigitales.org.
23. <https://www.clarin.eu/news/congratulations-our-first-clarin-knowledge-centre>.
24. For consistency, we used the label "Ministry of Science" to group the various denominations given across the analysed period to the ministry in charge of the national research funding programme.
25. A recent survey of laboratories set up in the humanities, digital humanities, and media studies within universities across the world, in the period 1983-2018, identified five labs in Spain: LBD, Medialab USAL, iArtHis_Lab, LINHD, and Medialab UGR (Pawlicka-Deger, 2019).
26. Scopus uses ASJC - All Science Journal Classification, while WoS has its own subject categorisation scheme: https://images.webofknowledge.com/images/help/WOS/hp_research_areas_easca.html.

6. References

- Aibar-Puentes, Eduard (2018). "La transformación neoliberal de la ciencia: el caso de las humanidades digitales". In: *Artefactos* 7, n. 1, pp. 13-28.
<http://dx.doi.org/10.14201/art2018711328>
- Baraibar-Echeverria, Álvaro (2014). In: Baraibar-Echeverria, Álvaro (ed.). *Humanidades digitales: Una aproximación transdisciplinar*. A Coruña: Universidade da Coruña, SIELAE.
- Dallas, Costis; Chatzidiako, Nephelie; Benardou, Agiatis; Bender, Michael; Berra, Aurélien; Clivaz, Claire; Cunningham, John; Dabek, Meredith; Garrido, Patricia; González-Blanco, Elena; et al. (2017). *European survey on scholarly practices and digital needs in the arts and humanities*. Zenodo.
<https://doi.org/10.5281/zenodo.260101>
- Endres, Bill (2017). "A Literacy of building: making in the digital humanities". In: Sayers, Jentery (ed.). *Making things and drawing boundaries: experiments in the digital humanities*. Minnesota: University of Minnesota Press, pp. 44-54.
<https://doi.org/10.5749/9781452963778>
- Fiormonte, Domenico (2014). "Digital humanities from a global perspective". In: *Laboratorio dell'ISPF*, XI.
<https://doi.org/10.12862/ispf14L203>
- Golbeck, Jennifer (2013). "Chapter 2. Nodes, edges, and network measures". In: Golbeck, Jennifer (ed.). *Analyzing the social web*. Boston: Morgan Kaufmann, pp. 9-23.
<https://doi.org/10.1016/B978-0-12-405531-5.00002-X>
- González-Blanco, Elena (2013). "Actualidad de las Humanidades Digitales y un ejemplo de ensamblaje poético en la red: ReMetCa". In: *Cuadernos hispanoamericanos*, n. 761, pp. 53-67.
<https://doi.org/10.4000/jtei.1274>
- Jenks, George-Frederick (1967). "The data model concept in statistical mapping". In: *International yearbook of cartography*, n.7, pp.186-190.
- López-Poza, Sagrario (2014). "Humanidades digitales hispánicas." In: Barros-Roel, Rocío (ed.). *Cincuentenario de La Asociación Internacional de Hispanistas*. A Coruña: Universidade da Coruña, Servizo de publicacións, pp. 151-166.
<https://dialnet.unirioja.es/servlet/extart?codigo=4775846>
- Lucía-Megías, José-Manuel (2003). "La Informática Humanística: notas volanderas desde el ámbito hispánico". In: *Incipit*, n. 23, pp. 91-114.
- Lucía-Megías, José-Manuel (2012), *Elogio del texto digital. Claves para interpretar el nuevo paradigma*. Madrid: Fórcola.
- Marcos-Marín, Francisco (1994). *Informática y humanidades*. Madrid: Gredos.
- Ortega, Erika; Eunice-Gutiérrez, Silvia (2014). "MapaHD. Una exploración de las humanidades digitales en español y portugués". In: Romero-Frías, Esteban (ed.); Sánchez-González, María (ed.). *Ciencias sociales y humanidades digitales: técnicas, herramientas y experiencias de e-research e Investigación En Colaboración*. La Laguna, Tenerife: Sociedad Latina de Comunicación Social, pp. 101-128.

- Pawlicka-Deger, Urszula. (2019). *Humanities and media labs around the world*. Zenodo.
<http://doi.org/10.5281/zenodo.2631219>
- Puy-Rodríguez, Ana; Cabello-Valdés, Cecilia; Ruiz-Galán, Olga; Sanmartín-Sola, Joseba; Rodríguez-Méndez, Carlos (2019). *Científicas en cifras 2017*. Fundación Española para la Ciencia y la Tecnología (FECYT): Indicadores I+D+i.
- Ramsay, Stephen; Rockwell, Geoffrey (2012). "Developing Things: Notes toward an Epistemology of Building in the Digital Humanities". In: Gold, Matthew (ed.). *Debates in the digital humanities*. Minnesota: University of Minnesota Press, pp. 75-84.
<http://doi.org/10.5749/minnesota/9780816677948.003.0010>
- Robinson-García, Nicolas; Amat, Carlos B. (2018). "¿Tiene sentido limitar la coautoría científica? No existe inflación de autores en Ciencias Sociales y Educación en España." In: *Revista española de Documentación Científica*, 41(2), a201.
<http://dx.doi.org/10.3989/redc.2018.2.1499>
- Robles-Gómez, Antonio; González-Blanco, Elena; Ros, Salvador; Del-Rio-Riande, Gimena; Hernández, Roberto; Tobarra, Llanos; Caminero, Agustín; Pastor, Rafael (2016). "Researchers' perceptions of DH trends and topics in the English and Spanish-speaking community. DayofDH data as a case study". In: *Digital humanities 2016: Conference abstracts*. Jagiellonian University & Pedagogical University, Kraków, pp. 658-660.
- Rojas-Castro, Antonio (2013). "El mapa y el territorio. Una aproximación histórico-bibliográfica a la emergencia de las humanidades digitales en España". In: *Caracteres*, 2, pp. 10-52.
- Romero-Frías, Esteban; Del-Barrio-García, Salvador (2014). "Una visión de las humanidades digitales a través de sus centros." In: *Profesional de la información*, v. 23, n. 5, pp. 485-92.
<https://doi.org/10.3145/epi.2014.sep.05>
- Sanz-Casado, Elías; Melero, Remedios; Aleixandre-Benavent, Rafael; Codina, Lluís; Coslado-Bernabé, María-de-los-Ángeles; De-Filippo, Daniela; Giménez-Toledo, Elea; Jiménez, Evaristo; Rico-Castro, Pilar (2020). *Metodología de Clasificación de Revistas de Humanidades y Ciencias Sociales con Sello de Calidad*. Fundación Española para la Ciencia y la Tecnología (FECYT): Indicadores I+D+i.
- Spence, Paul; González-Blanco, Elena (2014). "A historical perspective on the digital humanities in Spain". In: *H-Soz-Kult*, November 7th.
www.hsozkult.de/debate/id/diskussionen-2449
- Svensson, Patrik (2010). "The landscape of digital humanities". In: *Digital humanities quarterly*, v. 4, n. 1.
- Toscano, Maurizio. (2018). "Where the researcher cannot get: open platforms to collaborate with citizens on cultural heritage research data". In: Romero-Frías, Esteban (ed.) and Bocanegra Barbecho, Lidia (ed.). *Ciencias sociales y humanidades digitales aplicadas*. Granada: Universidad de Granada, pp. 538-56. New York: Downhill Publishing.
<http://doi.org/10.5281/zenodo.3758663>

- Toscano, Maurizio; Bocanegra Barbecho, Lidia; Ros, Salvador; González-Blanco, Elena (2020). "Insights on scholarly primitives from Digital Humanities research in Spain". Poster presented at the Scholarly primitives - DARIAH annual event 2020, Zagreb, Croatia.
<http://doi.org/10.5281/zenodo.4256689>
- Traag, Vincent; Waltman, Ludo; Van-Eck, Nees-Jan (2019). "From Louvain to Leiden: guaranteeing well-connected communities". In: *Sci Rep*, n. 9, art. 5233.
<https://doi.org/10.1038/s41598-019-41695-z>
- Walsh, Kevin (2016). *Open innovation, open science, open to the world. A vision for Europe*. Luxembourg: Publications Office of the European Union.
<https://doi.org/10.2777/061652>

IV.2. DISEÑO E IMPLEMENTACIÓN DE SISTEMAS DE INFORMACIÓN WEB PERSONALIZADOS

Una vez descrito y analizado el panorama de investigadores y proyectos de investigación en Humanidades Digitales en España, es decir un contexto en el que se plantean y desarrollan herramientas digitales para responder a las necesidades de investigación en humanidades, en este apartado pasamos a analizar ejemplos concretos de diseño e implementación de sistemas de información web personalizados.

El primer caso de estudio se refiere a la creación de un SIW para el manejo y la publicación en línea de fichas de obras de arte de artistas andaluces presentes en el continente americano, y se ha recogido en una publicación como artículo científico. En el mismo, junto con la descripción de la estructura y de las funcionalidades propias del sistema, se prima el análisis de otros dos aspectos centrales en la presente tesis doctoral: por un lado, la creación de plataformas orientadas a la ampliación y mejora de conjuntos de datos existentes (véase apartado IV.4); por el otro, los mecanismos de control y optimización del flujo de trabajo del propio grupo de investigación, en cuanto se refiere a operaciones de recogida, verificación y publicación de los datos.

El segundo caso de estudio habla de un sistema de información geográfica que actúa como elemento aglutinador de una amplia variedad de fuentes multiformato (documentos de archivo, planos cartográficos, aerofotografías, estudios históricos, excavaciones y prospecciones arqueológicas), necesarias todas ellas para la reconstrucción de la malla de asentamientos fortificados de época medieval en una amplia porción de territorio al norte de Roma (Italia).

Este apartado, junto con el siguiente (IV.3), presenta soluciones y propuestas concretas que una metodología digital de investigación puede aportar, a partir del segundo componente del esquema conceptual presentando en la figura 1: las tecnologías de la información y de la comunicación. En esta sección (IV.2), nos centramos en particular en la gestión de la información, mientras que en la próxima (IV.3) analizaremos una solución tecnológica a un problema de comunicación de los resultados de investigación.

IV.2.1. Un sistema de información web para la optimización del flujo de trabajo

En el contexto de la presente tesis doctoral, esta publicación se ha planteado para mostrar los efectos que la integración de un SIW personalizado tiene sobre los procesos de investigación llevados a cabo por parte de un grupo amplio y consolidado, que integra desde catedráticos, hasta investigadores emergentes y en formación. Se analizan, entre otros aspectos, las funcionalidades del SIW “Colabora” en términos de dinámica de trabajo y flujo de datos, ya sea por aportaciones que llegan desde colaboradores externos (especialistas o simples aficionados), que por la incorporación de nuevas fichas de obras de arte por parte de los afiliados al grupo (investigadores, becarios y estudiantes). Las diferentes fases relativas a la entrada y verificación de nuevos datos se han implementado como si se tratara de un proceso editorial (véase apartado III.2.10), a través del cual las fichas del catálogo pasan por varios estadios (borrador, lista para revisar y publicada). En cada fase, los privilegios de acceso y modificación se han establecido en función del rol y de las competencias temáticas que cada investigador cubre en el seno del grupo.

Más allá del flujo de datos, cabe destacar, como resultado de este trabajo, al menos otros dos aspectos que tienen consecuencias metodológicas sobre la ejecución del proceso de investigación: la clasificación del contenido y la recuperación de la información. El primer aspecto, del cual hemos hablado en el capítulo dedicado a la metodología (véase apartado III.2), afecta una parte sustancial de la estructuración conceptual inherente a cualquier sistema de gestión y recuperación de la información, representando una de las cuestiones con mayores consecuencias e implicaciones epistémicas. Su implementación se ha realizado a través de la creación de taxonomías, es decir listados jerárquicos de unidades conceptuales aplicadas a diferentes aspectos: estilo, iconografía, material, técnica, tipología y ubicación. La definición de taxonomías conlleva, además, un beneficio adicional sobre el control del lenguaje, reduciendo de manera considerable la introducción de errores que finalmente afectan la recuperación de la información.

El segundo aspecto está muy relacionado con el primero, en tanto que las taxonomías constituyen también los ejes privilegiados de indexación del contenido, a través de los cuales el usuario puede explorar un catálogo y localizar información. Las funcionalidades de búsqueda avanzada, como las que utilizan combinaciones de filtros u operadores booleanos, son cada vez menos populares debido a la consolidación del modelo de búsqueda a través de un campo único, promovido por los principales motores de búsqueda web. Este hecho no facilita el uso de sistemas de recuperación de la información optimizados para datos estructurados, incluso por parte de profesionales e investigadores. Una posible solución, empleada en este caso de estudio y favorecida por el uso de taxonomías jerárquicas para clasificar variables como el estilo y la localización, es optar por un sistema de búsqueda por facetas (véase apartado III.2).

En el contexto de esta publicación, el candidato ha colaborado estrechamente con varios componentes del equipo de investigación "Andalucía-América: Patrimonio y relaciones artísticas" (Cod.: HUM806) para la elaboración y el diseño del modelo de datos, la definición de las funcionalidades de la plataforma y la elaboración de las taxonomías. Por otro lado, ha llevado a cabo en autonomía las siguientes tareas: implementación informática de la plataforma web, elaboración del núcleo conceptual de la publicación, búsqueda bibliográfica, redacción del borrador inicial del texto, revisión formal, correspondencia con el comité editorial de la revista e implementación de las sugerencias de los revisores.

Referencia bibliográfica

Toscano, M. (2018). Andalucía y América. Plataforma colaborativa sobre patrimonio artístico iberoamericano. *Quiroga: Revista de Patrimonio Iberoamericano*, 14, pp. 36-43. <http://doi.org/10.5281/zenodo.3403745>

Título original

Andalucía y América. Plataforma colaborativa sobre patrimonio artístico iberoamericano

Título en inglés

Andalucía and America. Collaborative platform on Ibero-American artistic heritage

Abstract

The interactive Web platform *Collaborate with Andalucía and America* is a pioneering collaborative initiative, involving researchers and civil society, to extend and improve a catalogue of artworks created by Andalusian artists in Latin America. More recently, it expanded its range of functionalities to become the centralized working environment of the research group promoting the initiative.

Resumen

La plataforma web interactiva *Colabora con Andalucía y América* nace como una iniciativa pionera de colaboración, entre investigadores y con la sociedad, para la ampliación y mejora del catálogo de las obras realizadas por artistas andaluces en América Latina. Recientemente, se amplía para convertirse en la herramienta centralizada de trabajo para el grupo de investigadores promotor de la iniciativa.

Keywords

Web-based Information System, collaborative platform, co-creation, digital humanities, Andalusian artists.

Palabras claves

Sistema de Información Web, plataforma colaborativa, co-creación, humanidades digitales, artistas andaluces.

1. El proyecto

La iniciativa *Colabora con Andalucía y América*¹ (en adelante *Colabora*) se pone en marcha en 2016, en el marco del proyecto de investigación “MUTIS. Patrimonio artístico y relaciones culturales entre Andalucía y América del Sur”², con la voluntad de abrir a la participación ciudadana el amplio catálogo de obras que el grupo de investigación “Andalucía-América: patrimonio cultural y relaciones artísticas”, liderado por el profesor Rafael López Guzmán³, había recopilado a partir de 2008. Se trata de más de 2600 obras de arte realizadas por artistas andaluces en América Latina, registradas en fichas individuales detalladas y que cubren la gran mayoría de los países de América centro-meridional: desde México hasta Chile y Argentina, pasando por Colombia, Cuba,

Perú, Ecuador, Puerto Rico, Bolivia, Guatemala, República Dominicana, Venezuela, Brasil e Uruguay.

El catálogo presenta una recopilación bastante extensa del fenómeno en cuestión, con obras que cubren ampliamente los cinco siglos desde el descubrimiento de América hasta finales del siglo XX. La conciencia de la amplitud de la temática abarcada y del esfuerzo necesario para recopilar las fichas rellenas hasta la fecha, ha constituido el estímulo a poner en marcha esta iniciativa. Con el proyecto web *Colabora* se ha querido crear un instrumento que permita ampliar las posibilidades de acceso a las fuentes y mejorar, a nivel global, el catálogo de obras de arte. Aprovechando de las potencialidades de la Red Internet, hoy es posible recoger aportaciones directamente desde los lugares objetos de estudio o desde otros grupos de investigación que se estén ocupando de temas afines⁴. Las personas que se registran y contribuyen a la iniciativa pueden ser de las más variadas, desde investigadores especialistas en patrimonio artístico americano, apasionados que conozcan alguna de estas obras de primera mano, por cercanía o por algún vínculo específico, hasta simples ciudadanos de a pie que lleguen a conocer la iniciativa y se quieran sumar por afán de co-creación y participación en red⁵.

La plataforma, creada inicialmente para trabajar con y para la sociedad, ha llegado con el tiempo a despertar otro interés por parte del grupo de investigación promotor de la iniciativa, que ha visto en ella la potencialidad de convertirse en su propia herramienta de trabajo. Las funcionalidades necesarias han sido incorporadas muy recientemente, a partir de una revisión del modelo de datos en uso y del flujo de trabajo entre investigadores. Para respetar la sucesión de las fases de desarrollo e implementación de la plataforma, empezaremos describiendo su faceta abierta a la sociedad, para acabar con la ilustración de sus características como espacio privado de trabajo.

2. *Colabora* como plataforma abierta a la sociedad

2.1. EL CATÁLOGO EN ACCESO ABIERTO

Para poder ampliar el catálogo existente, en lugar de distribuir simplemente un formulario por Internet, se ha decidido apostar por una plataforma Web, una herramienta informáticamente más compleja, que brinda varias ventajas en cuanto a posibilidad de ofrecer y recibir información.

En primer lugar, se ha publicado en línea el catálogo completo de obras recopiladas hasta la fecha⁶, con la información estructurada en campos específicos y acompañada por el material ilustrativo. Trasladar el conjunto de fichas y archivos desde la base de datos local, en Access, a la base de datos en línea, en MySQL⁷, ha sido un reto de cierta envergadura, considerada la complejidad de la estructura relacional existente y las características de la herramienta de búsqueda que se quería poner a disposición en la web, de la que hablaremos a continuación.

La decisión de publicar en línea el catálogo en abierto tiene una serie de beneficios⁸: por un lado, facilita la indexación del sitio web por parte de los motores de búsqueda, incrementando su visibilidad; por otro lado, abre la posibilidad de mejorar el catálogo mismo, a través de las contribuciones voluntarias de colaboradores ocasionales, más allá de las posibilidades del grupo de investigación promotor; finalmente, constituye también una manera de ofrecer de forma gratuita una información de calidad a los visitantes de la plataforma, como una contrapartida por su interés en la iniciativa y para fomentar su posible colaboración.

Además, como veremos más adelante, el disponer de una versión en línea de catálogo, ha abierto la posibilidad de convertirlo en un instrumento de trabajo para el propio grupo de investigación, con las ventajas de una moderna herramienta en red: mayor accesibilidad, funcionalidad multiusuario concurrente, control de versiones, compatibilidad con las plataformas móviles, facilidad en compartir la información con otros investigadores, etc.: todas ellas características que constituyen las carencias típicas de una base de datos almacenada localmente.

Para facilitar la consulta del catálogo, se ha implementado una búsqueda por facetas⁹, una potente herramienta que permite a los usuarios explorar las fichas mediante la aplicación combinada de varios filtros: localización, estilo, tipología, autor e iconografía (figura 1). Además, se ha reflejado en los criterios de búsqueda aquella información que estuviese organizada de forma jerárquica, como la localización (País, región, provincia, ciudad), para que resulte más fácil, por ejemplo, encontrar en qué provincias de Bolivia se conservan las 15 representaciones de San Juan de Dios realizada en estilo Barroco (<https://goo.gl/8G942x>) o, al revés, partir desde una determinada ciudad, como por ejemplo Bogotá, y averiguar a qué autores pertenecen las 531 obras catalogadas que ahí se conservan (<https://goo.gl/xrLTTG>).



Buscar obra

Cronología inicial

Escribir solo el año (ej. 1500)

Cronología final

Aplicar **Reiniciar**

Localización

(-) Bolivia
Potosí (7)
Chuquisaca (6)
La Paz (2)

Estilo

(-) Barroco

Tipología

Pinturas de caballete (13)
Esculturas de bulto redondo (1)
Pinturas sobre tabla (1)

Autor

Anónimo (9)
Pérez Holguín, Melchor (2)
Anónimo potosino (1)
Córdoba, Ignacio (1)
Pérez Holguín, Melchor (1)

La búsqueda ha encontrado 15 elementos

San Juan de Dios en contemplación

Autor: Pérez Holguín, Melchor

Cronología: De 1701 hasta 1800

Tipología: Pinturas de caballete

Estilo: Barroco

Localización: Universidad Mayor, Real y Pontificia de San Francisco Xavier Sucre Chuquisaca Bolivia

Elaboración: P07-HUM-03089



San Juan de Dios con el Niño Jesús al hombro

Autor: Anónimo

Cronología: De 1701 hasta 1800

Tipología: Pinturas de caballete

Estilo: Barroco

Localización: Museo Universitario de Charcas Sucre Chuquisaca Bolivia

Elaboración: P07-HUM-03089



San Juan de Dios con el Niño Jesús al hombro

Autor: Anónimo

Cronología: De 1701 hasta 1800

Tipología: Pinturas de caballete

Estilo: Barroco

Localización: Museo Universitario de Charcas Sucre Chuquisaca Bolivia

Elaboración: P07-HUM-03089



San Juan de Dios con el Niño Jesús

Autor: Anónimo

Cronología: De 1667 hasta 1701

Tipología: Pinturas de caballete

Estilo: Barroco

Localización: Museo Convento de la Recoleta Sucre Chuquisaca Bolivia

Elaboración: P07-HUM-03089



Figura 1. Catálogo de obras publicado en línea, con un ejemplo de búsqueda realizada combinando varios filtros.

Finalmente, la plataforma web se ha convertido también en una herramienta de *net-working* del proyecto de investigación, ya que aumenta su visibilidad en red y facilita el acercamiento directo con otros grupos de investigación que quieran involucrarse en la iniciativa. Por este motivo, aún si la información presentada y el público objetivo procede esencialmente de un ámbito hispanohablante, se ha decidido ofrecer la plataforma también en lengua inglesa. Hasta la fecha, la página web ha recibido alrededor de 3.500 visitas y más de una veintena de colaboradores registrados. La analítica señala aún una baja participación desde países de América Latina, con solo Colombia y Argentina entre los diez primeros por número de visitantes, lo que sugiere la importancia de llevar a cabo campañas específicas de difusión de la iniciativa.

2.2. LAS CONTRIBUCIONES EXTERNAS

Siendo el objetivo primario del proyecto la ampliación y mejora del catálogo existente, pasamos ahora a analizar las herramientas desarrolladas para facilitar la participación ciudadana en esta tarea.

Al llegar a la plataforma, el usuario se encuentra con la necesidad de registrarse para poder aportar información, ofreciéndose la posibilidad de identificarse mediante un clic a través de un servicio externo (Facebook o Google) o, más tradicionalmente, a través de la verificación de un correo electrónico. Mantener un acceso identificado conlleva varias ventajas, principalmente ofrecer al colaborador la posibilidad de volver sobre las fichas creadas para mejorarlas y permitir al equipo de gestión de la plataforma de ponerse en contacto con el usuario, además de reconocer el crédito de las aportaciones recibidas, así como limitar el envío de fichas basura (*spam*).

Una vez en el área privada, se invita al usuario a revisar el catálogo en línea, a partir de las obras que se encuentran en su entorno, para colaborar enviando nueva información o mejores fotos de las obras. De hecho, muchas fichas del catálogo no cuentan con ilustraciones o éstas no son de buena calidad, porque no siempre se ha podido acceder directamente a las obras de arte para realizar las fotos; así que se espera que en este aspecto la plataforma *Colabora* pueda aportar nuevas y mejores ilustraciones.

La otra opción de colaboración que se ofrece al usuario es el registro de una nueva obra de arte. El formulario es sencillo y se centra sobre todo en los datos más valiosos para el investigador, es decir una foto de la obra y su localización, aunque deja abierta la posibilidad de aportar más información, como el autor de la obra, la tipología y una pequeña descripción. Con la voluntad de que las fichas se puedan complementar directamente en el lugar de conservación de las mismas (museo, galería, iglesia, colección privada, espacio público, etc.), se ha querido garantizar la plena compatibilidad con las plataformas móviles, así como la posibilidad de subir directamente una foto desde un teléfono inteligente o una tableta (figura 2).

La plataforma notifica cada aportación al equipo de gestión, que revisa la ficha y decide si publicarla en el catálogo en línea, ponerse en contacto con su autor para pedir información adicional o marcarla como no útil, por ser errónea o duplicada. Mientras tanto, las fichas almacenadas se quedan a disposición para que su autor pueda volver sobre ellas y mantener un registro de las aportaciones efectuadas.

The screenshot shows a web form for registering a new artwork. At the top, the URL 'colabora.andaluciayamerica.com' is visible. The form consists of several fields: 'Título de la obra' (Title of the work) with a text input box; 'Autor' (Author) with a text input box; 'Tipología' (Typology) with a dropdown menu currently set to 'Pintura' (Painting); 'Localización *' (Location) with a large text input box; and 'Imagen *' (Image) with a file selection button labeled 'Scogli file ne...to' and a blue button labeled 'Subir al servidor' (Upload to server).

Sube una imagen de la obra o una foto desde tu móvil.
[Más información](#)

Figura 2. Formulario de registro de una nueva obra a través de un aparato móvil.

3. Colabora como herramienta de investigación

3.1. EL MODELO DE DATOS Y LAS FICHAS

El primer paso para convertir la plataforma en una completa herramienta de trabajo para el grupo de investigación fue la revisión del modelo de datos, desarrollado inicialmente para las fichas del catálogo público y las aportaciones externas, el cual tuvo que ampliarse para incorporar una serie de campos útiles para la investigación, pero innecesarios para la consulta pública. En lugar de crear una nueva ficha de trabajo privada, se ha decidido incorporar estos campos a la ficha de catálogo, dejando su visibilidad restringida a los usuarios pertenecientes al grupo de investigación. En definitiva, en la actualidad la base de datos cuenta con tres fichas básicas: una “completa”, que es la que se presenta al investigador; una “reducida”, que coincide con la ficha completa excepto algunos campos de uso privado, y es la que se presenta a través del catálogo público; y una ficha “simplificada”, que hemos descrito anteriormente y se presenta al colaborador externo que quiera señalar la localización de una nueva obra para su incorporación en el catálogo. Igualmente, se ha añadido una “ficha bibliográfica”, relacionada con la ficha completa de obra, para el registro de los volúmenes utilizados como fuente de información.

La revisión de la ficha en uso ha comportado la eliminación de campos donde se guardaba información redundante, de difícil averiguación o simplemente innecesaria para los fines de la investigación, así como la introducción de nuevos, como, por ejemplo, el campo de referencia web como fuente de información.

3.2. EL FLUJO DE TRABAJO

El flujo de trabajo en uso antes de la creación de la plataforma web empezaba con la recopilación de datos que los investigadores llevaban a cabo rellenando una plantilla en un fichero en formato doc. La plantilla presentaba los diferentes campos a través de una serie de etiquetas, organizadas en grupos coherentes (Identificación y localización, Descripción, Documentación, etc.), que básicamente marcaban espacios de texto libre para rellenar. En el mismo fichero se guardaba también la imagen de la obra en cuestión. Los numerosos ficheros Word producidos autónomamente por los investigadores del grupo se centralizaban sucesivamente para el almacenamiento final en una base de datos en Access, desarrollada por el Instituto Andaluz del Patrimonio Histórico para bienes muebles del patrimonio de Andalucía. Los datos eran introducidos manualmente por parte de un único investigador que realizaba al mismo tiempo una tarea de revisión y ajuste del contenido de cada ficha a las características de los campos en la base de datos. Con este flujo de trabajo, frecuente en proyectos de esta índole, los mecanismos de control del lenguaje, fundamentales para las operaciones recuperación y búsqueda de datos, se colocaban en fase de revisión de la ficha y lejos del autor de la misma, dificultando así la posibilidad de resolver dudas utilizando la fuente primaria de la información. Además, la base de datos en Access había sido desarrollada con fines de catalogación y no de investigación, para almacenar ordenadamente cualquier

tipo de bien mueble y no para responder a las exigencias específicas de un proyecto de investigación, que además normalmente evolucionan en el tiempo.

Sin embargo, el nuevo flujo de trabajo ofrece al investigador la posibilidad de rellenar directamente la ficha completa en la base de datos, lo que supone una ventaja notable ya sea en cuanto a la rapidez en la entrada de datos que con respecto a la integridad y consistencia de los mismos, ya que las reglas de control establecidas para cada campo se aplican desde la fase inicial. Cada investigador cuenta con un panel de control personalizado para la visualización de la actividad y el seguimiento de sus fichas, que permanecen en estado “borrador” mientras que se vayan completando. El estado sucesivo de “revisión” indica que la ficha ha sido completada con toda la información disponible y está lista para su incorporación al catálogo, en estado “público”, o para permanecer accesible solo al grupo de investigación, en estado “privado”.

Toda la actividad realizada en la plataforma queda registrada en un historial completo de eventos: acceso de los usuarios, cambios de estado de las fichas y modificaciones realizadas en cada campo, estas últimas, además, reversibles en el tiempo (control de versiones).

3.3. LAS TAXONOMÍAS

Una fase importante en el proceso de revisión de la información existente se ha centrado en las taxonomías, o tesauros, relativas a autores, estilos, técnicas, materiales, tipologías e iconografías. Al limitarse el acceso a la edición de las fichas completas solo a personal investigador, se ha decidido utilizar listas semi-abiertas, que presentan al usuario un listado pre-compilado de elementos seleccionables, así como la posibilidad, si procede, de incorporar nuevas entradas. Tal y como se observa en la figura 3 (parte izquierda), para facilitar el encuentro de elementos existentes, la base ha sido configurada para no tener en cuenta mayúsculas y acentos y buscar correspondencias tanto en el apellido como en el nombre.

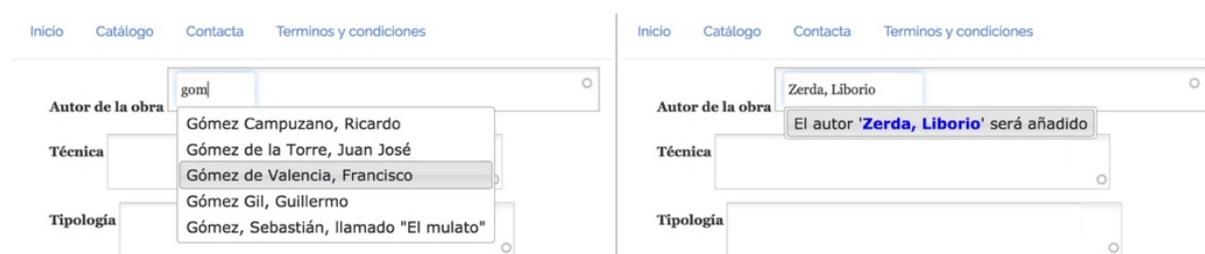


Figura 3. Ejemplo de listado dinámico para el campo Autor de la obra. La base de datos restituye correspondencias seleccionables (izquierda) o facilita el registro de un nuevo autor (derecha).

La revisión ha comportado una reflexión crítica sobre los términos utilizados hasta la fecha, con numerosos cambios y reagrupaciones realizados en función de las necesidades analíticas del proyecto: ulterior ejemplo de cómo la informatización del proceso de investigación y la definición de un modelo de datos personalizado llegan a tener

consecuencias metodológicas y analíticas. En este caso los listados son relativamente simples, pero su rol es fundamental para facilitar la correcta reutilización de la información en diferentes apartados, hacer efectivo el control del lenguaje, destinado a evitar simples errores tipográficos u ocurrencias como el uso de sinónimos para referirse a un mismo concepto, y permitir el análisis cuantitativo de los datos.

3.4. LA RECUPERACIÓN DE LA INFORMACIÓN

Un sistema de gestión de la información no es tal si no cuenta con herramientas que permitan extraer la información almacenada aplicando unos determinados criterios de búsqueda. La plataforma pública dispone de un buscador por facetas que permite combinar varios filtros para encontrar una obra en concreto o para cuantificar los resultados según la tipología, la iconografía, el autor, etc. Acorde a los permisos del usuario que realice la búsqueda, investigador del equipo o simple colaborador, los resultados incluyen solo las fichas públicas o también aquéllas en estado privado o borrador. Otra herramienta de recuperación de la información es la descarga masiva de datos, que se ha incorporado para cuando el investigador necesite procesar los datos a través de programas externos, por ejemplo para realizar tablas y gráficas.

Es necesario hacer hincapié que el disponer de un sistema de información en Internet proporciona una ventaja enorme a la hora de facilitar el acceso rápido y concurrente a los datos, sabiendo que se dispone en cualquier momento de la versión más reciente y completa. Además, mientras se analizan los datos almacenados, se pueden averiguar errores o lagunas y arreglarlos en tiempo real, para que queden inmediatamente disponibles al resto de usuarios. Disponer de copias locales, lo que podría aparecer como una ventaja, daría lugar fácilmente a la duda sobre quién cuenta con la última versión de cada ficha. Con el precedente sistema de almacenamiento en local, las posibilidades de acceso a la información eran muy remotas, llegando a afectar negativamente la calidad del conjunto del *corpus* y el proceso de análisis¹⁰.

4. CONCLUSIONES

En el presente artículo hemos descrito el caso de un sistema de información web para la investigación en historia del arte, creado en principio para extender el alcance geográfico de acceso a las fuentes y ofrecer a la sociedad un amplio catálogo de obras de arte realizadas por artistas andaluces en tierra americana, sucesivamente ampliado para convertirse en la herramienta centralizada de trabajo para el grupo de investigadores promotor de la iniciativa. Más que la descripción detallada de las características propias de la plataforma, el objetivo era ofrecer una panorámica sobre las posibilidades que la adopción de un sistema de información web ofrece a la investigación y algunas claves sobre los aspectos a tener en cuenta en el proceso; evidenciar, en definitiva, cómo el uso y la aplicación de las tecnologías de la información y de la comunicación potencia el desarrollo de las Humanidades Digitales.

A través de la plataforma *Colabora*, hoy en día los componentes del grupo de investigación “Andalucía-América: patrimonio y relaciones artísticas” pueden visualizar su trabajo a la sociedad, compartir datos de forma sencilla, establecer vínculos con otros grupos que trabajan temáticas similares y recibir nuevas fuentes de investigación desde lugares remotos, así como centralizar y optimizar la tarea interna de recopilación de datos, accediendo desde cualquier lugar y a través de cualquier dispositivo, aprovechando de un modelo de datos creados en función de sus necesidades, con un registro automático de la actividad y el control sobre el historial de cambios en cada ficha.

Concluyendo, con este proyecto se ha definido un espacio digital, permanentemente accesible, sobre el patrimonio artístico iberoamericano, con un catálogo de obras en continua ampliación y unas funcionalidades que se irán adaptando a las exigencias de la investigación, que facilita el trabajo en equipo y puede atraer nuevas fuentes de financiación, así como servir de estímulo e inspiración para otros grupos.

NOTAS

¹ <https://colabora.andaluciayamerica.com> [Fecha de acceso: 05/07/2018].

² <http://andaluciayamerica.com> [Fecha de acceso: 05/07/2018].

³ Agradezco cordialmente Rafael López Guzmán y Yolanda Guasch Marí por la disponibilidad y la amabilidad demostrada en las diferentes conversaciones necesarias para la definición del modelo de datos y de las funcionalidades de la plataforma y por el interés en que el presente trabajo se realizara.

⁴ SPIRO, Lisa. “Computing and Communicating Knowledge: Collaborative Approaches to Digital Humanities Projects.”. En: MCGRATH, Laura (Ed.). *Collaborative Approaches to the Digital in English Studies*. Logan: Utah State University Press, 2011, págs. 44-82; BOCANEGRA BARBECHO, Lidia y TOSCANO, Maurizio. “El exilio republicano español: Estudio y recuperación de la memoria a través de la web 2.0. Nuevo enfoque metodológico con el proyecto *e-xiliad@s*”. *Migraciones & Exilios*, 15 (2015), págs. 113-136. [<http://hdl.handle.net/10481/45829>]; BROWN, Susan. “Towards best practices in collaborative online knowledge production”. En: CROMPTON, Constant, LANE, Richard and SIEMENS, Ray (Eds.). *Doing Digital Humanities. Practice, Training, Research*. London: Routledge, 2016, págs. 47-64.

⁵ BOCANEGRA BARBECHO, Lidia; TOSCANO, Maurizio; DELGADO ANÉS, Lara. “Co-creación, participación y redes sociales para hacer historia. Ciencia con y para la sociedad”. *Historia y Comunicación Social* (Universidad Complutense de Madrid), 2 (2017), págs. 325-346. [<http://dx.doi.org/10.5209/HICS.57847>].

⁶ RIDGE, Mia (Ed.) *Crowdsourcing our Cultural Heritage*. Ashgate: Routledge 2014; HEDGES, Mark y Dunn, Stuart. *Academic Crowdsourcing in the Humanities*. Oxford: Chandos Publishing, 2018.

⁷ <https://colabora.andaluciayamerica.com/catalogo> [Fecha de acceso: 05/07/2018].

⁸ Para la creación del sistema de información personalizado se ha elegido el software Open Source Drupal, uno de los sistemas de gestión de contenido de referencia para de

humanidades digitales, ver DOMBROWSKI, Quinn. "Drupal and other content management systems". En: CROMPTON, Constance, LANE, Richard and SIEMENS, Ray (Eds.). *Doing Digital Humanities. Practice, Training, Research*. London: Routledge, 2016, págs. 289-302 y DOMBROWSKI, Quinn. *Drupal for Humanists (Coding for Humanists)*. Texas: A&M University Press, 2016.

⁹TOSCANO, Maurizio, "Where the researcher cannot get. Open Platforms to collaborate with citizens on cultural heritage research data". En: ROMERO FRÍAS, Esteban y BOCANEGRA BARBECHO, Lidia (Coords.). *Ciencias Sociales y Humanidades Digitales aplicadas: casos de estudio y perspectivas críticas*. Granada: Universidad de Granada y Downhill Publishing (NY), 2018, págs. 538-561.

¹⁰https://es.wikipedia.org/wiki/Búsqueda_por_facetas [Fecha de acceso: 05/07/2018].

¹¹TOSCANO, Maurizio. "Desde datos locales a plataformas abiertas: la Web como herramienta de colaboración con la sociedad". En: BOCANEGRA BARBECHO, Lidia y ROMERO FRÍAS, Esteban (Coords.). *Territorios Digitales. Construyendo unas Ciencias Sociales y Humanidades Digitales*. Granada: Universidad de Granada y Downhill Publishing (NY), 2018, págs. 192-193.

IV.2.2. Un sistema de información web como atlas temático digital

La temática histórica objeto del estudio que presentamos en esta sección es la reconstrucción de la red de asentamientos fortificados de época bajomedieval en la provincia de Viterbo (Italia). Con el fin de almacenar la información recolectada y facilitar su consulta y análisis, se ha diseñado y desarrollado un sistema de información geográfica personalizado, que permitiera gestionar la variedad de fuentes involucradas en el proceso.

Más allá del valor de la publicación en términos de reconstrucción de la antigua trama de asentamientos, se trata de una contribución de corte prevalentemente metodológico que se ha planteado en el contexto de esta tesis para ilustrar las diferentes etapas del proceso de investigación vinculado al desarrollo del sistema informático. Se va desde la búsqueda y selección de fuentes históricas, hasta el diseño del modelo de datos (figuras 9 y 10) y la implementación del sistema de información geográfica, cuya primera versión se apoya en una infraestructura informática de escritorio (FileMaker/ArcMap), hasta el planteamiento de las características y necesidades para el traslado del sistema en la Web, sobre una plataforma interoperable basada en datos enlazados¹²⁴. En todo el proceso, el rol del sistema de información personalizado, ya sea durante la fase de recogida de datos, que de cara a su análisis histórico y espacial (Toscano 2008), así como para su posterior publicación, resulta evidente.

En este caso, el formato elegido para la consulta en línea de los datos es el atlas temático, que une el potencial visual y de enfoque conjunto ofrecido por el mapa con las ventajas en términos de consulta y filtrado ofrecidas por el catálogo de fichas individuales. En el protocolo metodológico descrito anteriormente (véase apartado III.2), la presente contribución alcanza hasta el punto 6, es decir antes de la implementación informática en línea, que está previsto realizarse a finales de 2021, en el marco de un proyecto de colaboración con la *Fondazione Carivit Viterbo*.

¹²⁴ La solución tecnológica candidata a cubrir las necesidades individuadas en este proyecto es WissKI (véase apartado III.1).

Progetto di catalogazione dei siti fortificati della Provincia di Viterbo. Scheda sito

| | | |
|--|---|--|
| ID sito <input type="text"/> | Provincia <input type="text"/> | Attestazioni precastrensi <input type="text"/> |
| Toponimo IGM <input type="text"/> | Affidabilità georef. <input type="checkbox"/> | Prima attestazione sito fortificato <input type="text"/> |
| Comune <input type="text"/> | Affidabilità crono. <input type="checkbox"/> | Autore incastellamento <input type="text"/> |
| Località <input type="text"/> | Affidabilità totale <input type="checkbox"/> | Abbandono o decastellamento <input type="text"/> |
| Toponimi storici <input type="text"/> | | Detentore <input type="text"/> |
| Geologia <input type="text"/> | Morfologia <input type="text"/> | cr. iniziale cr. finale Tipologia sito |
| Idrografia principale <input type="text"/> | Idrografia secondaria <input type="text"/> | <input type="text" value="Descrizione"/> |
| Osservazioni <input type="text"/> | | <input type="text"/> |
| | | Definizione condizioni attuali <input type="text"/> |
| | | Descrizione condizioni attuali <input type="text"/> |
| | | Notizie documentarie <input type="text"/> |
| Compilatore <input type="text"/> | | Etichetta <input type="text"/> |

Figura 9. Interfaz usuario de la ficha alfanumérica "asentamiento". Fuente: elaboración propia.

Progetto di catalogazione dei siti fortificati della Provincia di Viterbo. Scheda bibliografia

| | |
|---|---------------------------------------|
| ID_biblio <input type="text"/> <input type="button" value="Genera ID"/> | ID_sito <input type="text"/> |
| Titolo <input type="text"/> | Toponimo IGM <input type="text"/> |
| Autore Cognome <input type="text"/> | Località <input type="text"/> |
| Nome <input type="text"/> | Comune <input type="text"/> |
| Periodico <input type="text"/> | Toponimi storici <input type="text"/> |
| Miscellanea <input type="text"/> | |
| Luogo di edizione <input type="text"/> | |
| Anno <input type="text"/> (es.: 1998 a) | |
| Casa Editrice <input type="text"/> | |
| Pagine <input type="text"/> | |
| Osservazioni <input type="text"/> | |

Figura 10. Interfaz usuario de la ficha alfanumérica "bibliografía". Fuente: elaboración propia.

Para la publicación, el candidato se ha ocupado, en autonomía, de la conceptualización del trabajo, de la redacción del borrador inicial, del apartado gráfico, del desarrollo informático del sistema de información geográfica, de la valoración de las funcionalidades web, de la correspondencia con el comité editorial y de implementar las sugerencias de los revisores. La actualización bibliográfica relativa a las fuentes escritas y a las investigaciones arqueológicas, además de la revisión final del texto, se ha realizado en coautoría.

Referencia bibliográfica

Toscano, M., & Romagnoli, G. (2020). Atlante dei siti fortificati della provincia di Viterbo, Italia (X-XV secolo). Fonti e metodi per la ricostruzione della rete insediativa bassomedievale. X, pp. 471-478, en Navarro Palazón, J. y García Pulido, L. (eds), *Defensive architecture of the Mediterranean*, Universidad de Granada y Universidad Politécnica de Valencia <https://doi.org/10.4995/FORTMED2020.2020.11545>

Título original

Atlante dei siti fortificati della provincia di Viterbo, Italia (X-XV secolo). Fonti e metodi per la ricostruzione della rete insediativa bassomedievale

Título en español

Atlas de los sitios fortificados en la provincia de Viterbo, Italia (siglo X-XV). Fuentes y métodos para la reconstrucción de la red de asentamientos de la Baja Edad Media

Abstract

This study addressed the historical phenomenon known as *incastellamento*, in the area of the current province of Viterbo, from a quantitative and geographical perspective. The time period considered was the tenth-fifteenth century. The paper describes the documentary sources, historical maps, aerial images, past studies and archaeological sources that are available to researchers, and which have been used, in good measure, to reconstruct the fortified settlement network. Moreover, the paper explains the methodologies used to identify, store and geocode the whole dataset, which so far comes to a total of 191 fortified settlements. In conclusion, we discuss the main characteristics of the online atlas, intended as an open and interoperable platform to consult, query and retrieve information from the dataset of late-medieval fortified settlements.

Resumen

En este estudio se ha abordado el fenómeno histórico conocido como *incastellamento*, en la zona de la actual provincia de Viterbo, desde una perspectiva cuantitativa y espacial. Se ha considerado el período de tiempo entre el siglo X y el siglo XV. El artículo describe las fuentes documentales, los mapas históricos, las imágenes aéreas, las fuentes arqueológicas y los estudios previos a disposición de los investigadores y que en buena medida se han utilizado, en este contexto, para reconstruir la red de asentamientos fortificados. Además, se explican las metodologías utilizadas para identificar, almacenar y geolocalizar los lugares objetos de estudio, que hasta ahora asciende a un total de 191 asentamientos fortificados. Para concluir, se examinan las principales características del que será el atlas en línea, concebido como una plataforma abierta e interoperable para consultar, interrogar y recuperar información sobre el conjunto de datos históricos almacenados.

Keywords

Middle Ages, fortified settlements, historical gazetteer, data modelling, Web Information System, Linked Open Data, Digital Humanities, Geographic Information System.

Palabras claves

Edad Media, asentamientos fortificados, nomenclátor histórico, modelización de datos, Sistema de Información Web, Datos Abiertos Enlazados, Humanidades Digitales, Sistema de Información Geográfica.

1. Introduzione

Ambito territoriale oggetto del presente studio è lo spazio racchiuso all'interno dei confini della provincia di Viterbo, quali risultano dalla corrente suddivisione amministrativa, l'ambito cronologico quello compreso tra i secoli X e XV d.C. ed il campo d'indagine quello dell'impianto di abitati fortificati, conosciuti generalmente dalle fonti medievali come *castra* o *castella*. L'area selezionata coincide in buona parte con i limiti del Patrimonio di San Pietro in Tuscia, nella denominazione e delimitazione territoriale assunta nell'XI secolo e conservata fino al XVIII.

Ciò che ha mosso l'indagine non è stato l'interesse per il singolo castello, o per quanto da esso si potesse estrapolare in termini di storia militare, architettonica o urbanistica della provincia, ma per la rete unitaria degli insediamenti nel suo insieme, ricostruita attraverso una serie di fonti e metodologie che descriveremo nei prossimi capitoli. Ciò che muove, infine, la proposta di pubblicazione dell'atlante web dei siti fortificati della provincia di Viterbo è la volontà di rendere pubblici i dati raccolti, espliciti gli strumenti e i metodi utilizzati e riproducibili i risultati delle analisi condotte (Toscano, 2008), secondo i principi della Scienza aperta con Dati Aperti (Machado, 2015) e attraverso la pratica dell'informatica umanistica.

2. Fonti per la ricostruzione della rete

Si passano qui di seguito in rapida rassegna le principali fonti archeologiche, storico-documentarie e cartografiche utili per l'ubicazione dei siti incastellati e, più in generale, per la ricostruzione della rete dei castelli nell'ambito territoriale esaminato.

2.1. LE FONTI DOCUMENTARIE

Tra le fonti di matrice ecclesiastica, una posizione preminente spetta alla documentazione dei grandi enti monastici che, a partire dall'altomedioevo, costituirono i principali proprietari fondiari del territorio a Nord di Roma. Se il ruolo di S. Maria di Farfa e di S. Salvatore al Monte Amiata nel territorio altolaziale si riduce notevolmente nell'età dei castelli fino a divenire insignificante, a partire dal X secolo si accresce il ruolo dei grandi monasteri romani. Tra i cartulari editi di chiese e monasteri romani, si possono menzionare quelli dell'Abbazia di San Paolo fuori le Mura (Trifone, 1908-1909), della Basilica Capitolare di S. Pietro in Vaticano (Schiaparelli, 1901-1902), del Monastero di Sant'Anastasio *ad Aquas Salvias* (Giorgi, 1878), del Monastero di San Silvestro *in Capite* (Federici, 1899-1900), dell'Abbazia di Sant'Alessio all'Aventino (Monaci, 1904-1905), dei Monasteri dei SS. Cosma e Damiano *in Mica Aurea* (Fedele, 1898), di S. Maria in Campo Marzio (Carusi, 1948), di Santa Maria *in Via Lata* (Hartmann,

Merones, 1895-1913) e di Santa Maria Nova (Fedele, 1900-1903). A questi importanti nuclei documentari si possono aggiungere le carte di alcuni cenobi umbri, come l'Abbazia di Sassovivo (Cencetti *et al.*, 1973-1975), e altolaziali.

Rivestono una particolare rilevanza anche i fondi degli enti ecclesiastici viterbesi: le carte del Capitolo della Cattedrale di Viterbo (pubblicate integralmente fino al 1299: Egidi, 1906-1907), delle collegiate di SS. Stefano e Bonifacio di Viterbo (Buzzi, 1988), di S. Angelo in Spatha, S. Sisto e dei Conventi di S. Maria della Verità e della Trinità (Savignoni, 1895).

Per quanto concerne le fonti fiscali e di governo, sono fondamentali per la storia del popolamento il *Liber Censuum* della Chiesa Romana (Fabre, Duchesne, 1905-1952), i censi e le collettorie della Santa Sede (Theiner, 1861-1862), i registri delle decime, che coprono, con varie soluzioni di continuità dovute al carattere episodico di questo tipo di riscossione, il periodo che va dal 1274 alla prima metà del XIV secolo (Battelli, 1946), ed i registri delle tassazioni del sale, eseguite semestralmente dai gabellieri dello Stato della Chiesa (Tomassetti, 1897). Ritrovata dall'archivista De Rossi a Siena, e per questo conosciuta col nome di "Romano-Senese", ma relativa al territorio a nord di Roma, è invece una tassazione del sale e del focatico pubblicata dal Tomassetti alla fine dell'Ottocento e datata da questo studioso alla metà del XIV secolo. Sempre per il Trecento, sono di notevole interesse anche le relazioni inviate dal Rettore del Patrimonio al Pontefice residente in Avignone (Antonelli, 1895) e i registri del Cardinal Albornoz del 1364 (Fabre, 1887).

Passando alle fonti di matrice comunale, il fondo diplomatico del Comune di Viterbo (comprendente la *Margarita Viterbese*) costituisce naturalmente una fonte di primario interesse per la ricostruzione dei rapporti delle istituzioni comunali con le comunità soggette e con le signorie di castello del territorio (Carbonetti Vendittelli 1990), così come l'analoga documentazione prodotta dalla città di Orvieto (Fumi 1884) e da Corneto (Supino 1969). Anche per altri centri del Patrimonio di S. Pietro è stato edito un buon numero di fonti civiche, disponibili in varie raccolte. Tra le fonti normative, spiccano gli statuti comunali medievali di Viterbo degli anni 1237-38, 1251-52 e 1469 (Egidi 1930; Buzzi, 2004) e quelli del castello di Fiorentino, una delle comunità soggette tra XIII e XIV secolo al comune viterbese (Egidi, 1910).

Per la ricostruzione della geografia del potere delle grandi famiglie patrizie, i patrimoni meglio documentati sono quelli dei Colonna, degli Orsini e degli Anguillara (Sora, 1906), restando invece quasi priva di documentazione la dominazione feudale della famiglia Normanni-Alberteschi sulla fascia costiera della Tuscia Romana. Sempre riguardo i grandi proprietari terrieri, risultano di grande interesse gli atti pervenuti relativi al passaggio dei beni territoriali all'Ordine di San Giovanni in Gerusalemme dopo la soppressione dei Templari, conservati anch'essi presso l'Archivio Vaticano (Silvestrelli, 1917).

Si segnalano, infine, le cronache quattrocentesche di Viterbo, compilate, per la parte anteriore al XV secolo, in parte sulla base di fonti più antiche (Egidi, 1901).

2.2. LE FONTI CARTOGRAFICHE E AEROFOTOGRAFICHE

Lo spoglio delle mappe storiche del Lazio (Almagià, 1960; Frutaz, 1972) non è infruttuoso, anche se la cartografia riguardante il territorio regionale inizia a diventare significativa solo a partire dal XVI secolo. La prima carta che riproduce in maniera molto precisa ed estesa buona parte del territorio laziale del Patrimonio di San Pietro in Tuscia è quella redatta dal Bellarmati nel 1536, illustrativa però dei soli centri ancora esistenti all'epoca. Dovendo escludere la splendida carta di Eufrosino della Volpaia del 1547, riferita alla sola Campagna Romana, per ottenere una esauriente descrizione della nostra area dobbiamo attendere gli inizi del XVII secolo, con la redazione dell'atlante del Magini, tra l'altro carente in riferimento alla quantità di insediamenti inseriti. Di maggiore interesse ai nostri fini risultano le carte dell'Oddi, del Mattei e dell'Ameti, con numerosi riferimenti agli insediamenti di età classica e medievale, anche distrutti, ed alla rete viaria dell'epoca di redazione. Sempre riferita al territorio del Patrimonio è la carta redatta nel XVIII secolo dai padri Maire e Boscovich, attualmente conservata presso la Calcografia Nazionale, mentre a piccole porzioni di territorio, ma con una ricca dotazione toponomastica, fanno riferimento quella del Ligustri del 1609, relativa ad una porzione del territorio costiero della Tuscia, e quella dipinta da Pietro da Cortona nella Villa Chigi a Castelfusano.

Il catasto geometrico-particellare dello Stato della Chiesa, prodotto in due serie nei periodi 1815-1835 e 1855-1875, conservate presso gli archivi di Stato di Roma e di Viterbo e note rispettivamente con i nomi di Catasto Pontificio e Catasto Gregoriano, rappresenta la prima rappresentazione completa del territorio metricamente attendibile e conserva il patrimonio toponomastico più dettagliato riferibile alle campagne laziali, oltre ad una preziosa immagine latente dell'uso del suolo, precedente agli sconvolgimenti subiti dal paesaggio a partire dalla prima metà del Novecento.

La cartografia in scala 1:25.000, prodotta dall'Istituto Geografico Militare Italiano a partire dal 1875 circa, riveste un importantissimo ruolo come fonte d'indicazioni toponomastiche ormai scomparse o comunque non riportate nella più recente cartografia regionale in scala 1:10.000, tanto da poter essere considerata ormai alla stregua di una "cartografia storica" (figura 1).

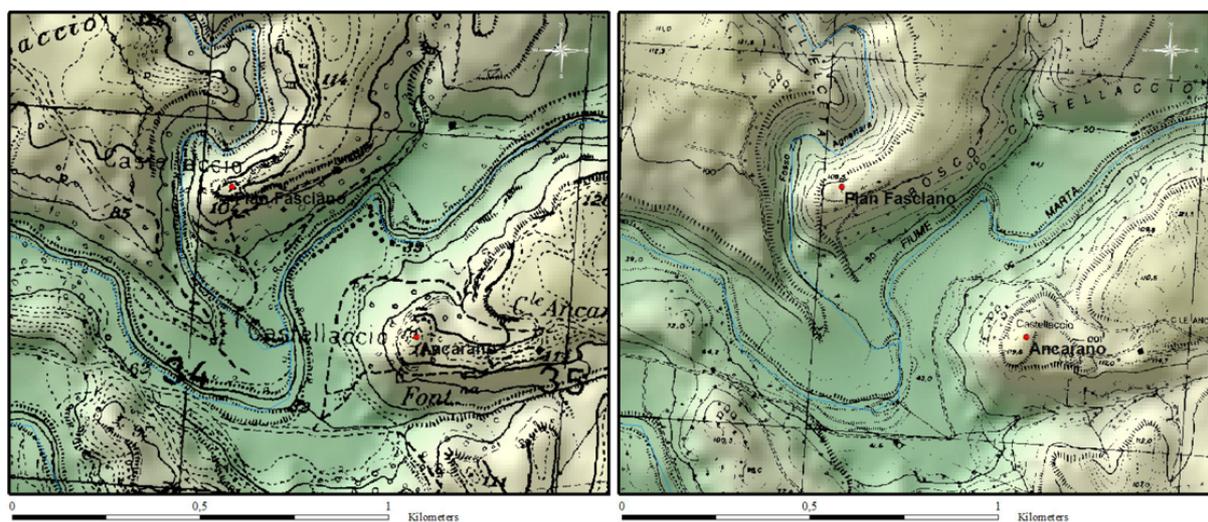


Figura 1. Ricchezza toponimica I.G.M. a confronto con le carte regionali: esempi di Ancarano e Pian Fasciano.

Un ultimo accenno meritano le coperture aerofotografiche del territorio, basilari per la localizzazione di molti degli insediamenti scomparsi del territorio altolaziale. Quella realizzata sull'intera provincia di Viterbo tra il 1938 e il 1940 ai fini della realizzazione del Catasto Italiano, nota come "Ala Littoria" e attualmente conservata presso l'Archivio di Stato di Viterbo, rappresenta da questo punto di vista una fonte di eccezionale interesse, sia per la scala delle riprese, sia per la sua cronologia, riproducendo un'immagine del paesaggio altolaziale antecedente al *boom* edilizio ed alle grandi trasformazioni agrarie del secondo dopoguerra.

2.3. I PRINCIPALI STUDI STORICI

Pur arrivando solo a lambire il territorio altolaziale, il lavoro storico-topografico di Giuseppe Tomassetti sulla Campagna Romana medievale e moderna rappresenta tuttora una importante base documentaria per la ricostruzione degli assetti del territorio a Nord di Roma (Tomassetti, 1910-1956; Carocci, Vendittelli, 2014).

La pubblicazione di P. Toubert sul Lazio medievale (Toubert, 1973), sebbene incentrata sulla Sabina, ha impresso una vera e propria svolta alla ricerca storica sul fenomeno dell'incastellamento nel Lazio, fornendo l'impulso per nuove ricerche di carattere sia storico che archeologico. Nel quarantennio successivo sono mancate sintesi sul territorio altolaziale, sebbene siano apparsi numerosi studi di carattere topografico o storico-documentario, in particolare sull'area circostante il centro urbano di Viterbo tra VIII e XI secolo (Lanconelli, 1992), sul territorio di Sutri (Vendittelli, 2008) e sull'area dei Monti della Tolfa (Vallelonga, 2012).

I repertori dei siti medievali del Lazio curati da G. Silvestrelli (Silvestrelli, 1914), da E. Martinori (Martinori, 1932-1934) e quelli, più recenti, delle sedi umane abbandonate del Patrimonio di San Pietro tra X e XVI secolo (Conti, 1980) e dei siti fortificati del Lazio (Istituto Italiano dei Castelli, 1985), conservano tutt'oggi una certa utilità ai fini di una prima valutazione quantitativa del fenomeno dell'incastellamento, sebbene

siano esclusivamente basati sullo spoglio delle fonti edite. Informazioni utili all'ubicazione dei centri incastellati altolaziali solo reperibili anche in due repertori toponomastici sulla diocesi di Orte e sulla provincia di Viterbo (Del Lungo, 1998; Del Lungo, 1999).

Vanno infine ricordate alcune opere di storiografia municipale, anche datate, quali quelle che hanno avuto ad oggetto le vicende medievali di Viterbo (Pinzi, 1887-1913; Signorelli, 1907-1969) e di Tuscania (Campanari, 1856).

2.4. LE RICERCHE ARCHEOLOGICHE

Evitando di passare in rassegna in modo sistematico le indagini archeologiche (per le quali si rimanda alle sintesi in Molinari, 2010 e Lanconelli, Romagnoli, 2018), si ricorda che le ricognizioni territoriali nell'Etruria Meridionale vantano una lunga tradizione, che risale indietro fino al grande progetto post-unitario della Carta Archeologia d'Italia (Gamurrini *et al.*, 1972), ed hanno contribuito in modo decisivo a scrivere la storia dei paesaggi altolaziali per l'età post-classica. Il progetto del *South Etruria Survey*, condotto dalla *British School at Rome* tra gli anni Cinquanta e Settanta del secolo scorso, ha comportato prospezioni di carattere sistematico in un'ampia regione a Nord di Roma, corrispondente ai territori storici di Veio, Falerii, Sutri e Capena (Duncan, Reynolds 1958; Jones, 1962-1963; Wickham, 1978-1979; Potter, 1985), ed è stato integrato da sondaggi di scavo in alcuni siti campione (Mallet, Whitehouse, 1967; Cameron *et al.*, 1984; Christie, 1991; Potter, King, 1997). I dati relativi al periodo compreso tra la tarda antichità e i secoli centrali del medioevo sono stati sistematicamente riesaminati nell'ambito del *Tiber Valley Project*, coordinato dalla *British School at Rome* a partire dal 1997 (Patterson, 2004; Coarelli, Patterson, 2008). Nel quadro delle ricognizioni territoriali si menzionano anche le recenti indagini condotte nell'ambito del progetto *Carta Archeologica d'Italia*, che mostrano un'apertura – se non uniforme almeno costante – alle emergenze archeologiche di età post-antica (Milioni, 2002; Scardozi, 2004).

Per il suo carattere di completezza, il repertorio degli scavi e delle scoperte archeologiche – prevalentemente di età etrusca e romana – effettuate in Etruria meridionale tra il 1939 e il 1975 nel territorio della Soprintendenza Archeologica dell'Etruria Meridionale può rivelare notizie di interventi su singoli centri, altrimenti difficilmente reperibili (Sommella Mura, 1969; Brunetti Nardi, 1972-1981).

Gli insediamenti rurali medievali dell'entroterra di Civitavecchia e della Teverina viterbese sono stati oggetti di specifiche indagini tematiche condotte nell'ambito dei progetti di scavo delle città medievali di Leopoli-Cencelle presso Civitavecchia (Nardi Combescure, 2002) e di Ferento presso Viterbo (Romagnoli, 2006). Tra gli scavi archeologici in centri incastellati, si possono menzionare almeno quelli condotti a Castel d'Asso presso Viterbo (Güll *et al.*, 2001) e a Pian Fasciano presso Viterbo (Fronti, Romagnoli, 2007).

Anche se non afferente ad un progetto formalizzato, ma comunque raggruppabile in un insieme coerente, un'ampia e interessante casistica, in particolar modo frutto di tesi di laurea e di specializzazione, è offerta dagli studi promossi negli ultimi anni nell'ambito delle cattedre di Archeologia Medievale delle Università di Roma la Sapienza e della Tuscia, che hanno sviluppato filoni di studi sulla facies rupestre dei siti incastellati (De Minicis, 2003-2011) e sull'analisi delle architetture e sulle tecniche costruttive delle strutture fortificate (Andrews, 1978; Chiovelli, 2007).

3. Metodologie di gestione, localizzazione e pubblicazione web degli insediamenti

3.1. L'ARCHITETTURA DEI DATI

Per la raccolta strutturata delle informazioni si è creata, in ambiente GIS, una banca dati relazionale, a partire da un modello dati specifico realizzato per registrare il fenomeno dell'incastellamento, che comprende anche i vincoli d'integrità e i vocabolari necessari alla normalizzazione del linguaggio. In particolare, presentano thesauri i seguenti attributi: comune, tipologia sito, affidabilità (cronologia e georeferenziazione), condizioni attuali e detentore. Completano la scheda sito i campi: toponimi, prima ed ultima menzione, attestazioni pre-castrensi, notizie documentarie e osservazioni. Si sono omessi, in fase di elaborazione del modello dati, gli attributi ottenibili tramite topologia spaziale: coordinate, idrografia, geologia e indice topografico (TPI). Le due tabelle principali sono dedicate l'una ai siti fortificati e l'altra alla relativa bibliografia. Per permettere di gestire l'evoluzione nel tempo della tipologia insediativa e della proprietà, alla tabella sito è stata collegata, attraverso una chiave esterna, la tabella detentori.

3.2. LA MAPPATURA

Il criterio che si è scelto di seguire nell'individuazione degli abitati da schedare è stato quello della presenza in essi, almeno in una determinabile fascia temporale, di strutture fortificate. Si sono esclusi i castelli di fondazione tardoantica che non avessero testimonianze storiche successive al 900 d.C., inserendo invece quelli che restano in vita nel periodo successivo.

Per la geolocalizzazione dei toponimi, la fonte primaria è stata la cartografia IGM, alla quale si è affiancata, per approfondire il dettaglio nel posizionamento dei siti, il modello digitale del terreno e la copertura ortofotografica in bianco e nero e a colori dell'intero territorio provinciale. Per la loro natura di siti fortificati, questi insediamenti erano delimitati da una o più cinte murarie e, in linea di principio, si sarebbero potuti rappresentare mediante l'uso di poligoni. Ciò avrebbe però escluso dalla rappresentazione tutti quei siti dove l'evidenza materiale della fortificazione manca o è frammentaria, e si è quindi prediletto l'uso della geometria puntuale, applicabile a tutti gli insediamenti. Determinata la geometria da utilizzare, è stato necessario definire un metodo per determinare la localizzazione del punto¹: sul cassero/torre, ove visibile nel

caso di insediamenti in vita; sul simbolo :: che nella cartografia IGM indica la presenza di rovine, nei casi di insediamenti abbandonati; sulla quota più alta, nei casi di insediamenti in vita senza evidenti strutture legate all'incastellamento e nei casi di insediamenti abbandonati senza l'indicazione di rovine.

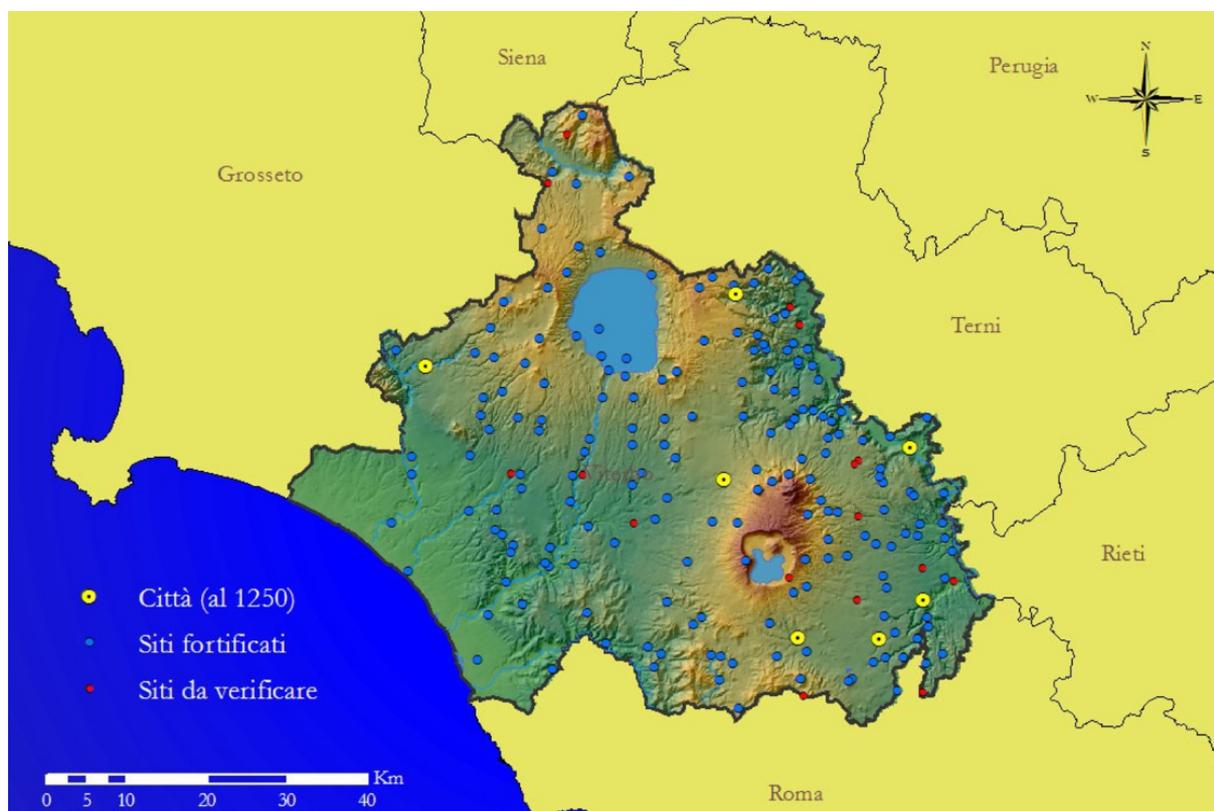


Figura 2. Totale dei siti georeferenziati attualmente presenti nella banca dati.

Il numero totale dei siti fortificati finora censiti e georeferenziati ammonta a 191, comprendendo quelli che hanno assunto o mantenuto configurazione castrale nel periodo storico in esame, ai quali si aggiungono, per il momento, sedici insediamenti da verificare (figura 2).

3.3 IL PORTALE WEB

Per la pubblicazione web della banca dati ottenuta si è optato per la creazione di un portale Web-GIS. Il formato atlante ci permette di combinare i vantaggi della mappa, in termini di visualizzazione, interrogazione e analisi spaziale della rete, con le prerogative di un catalogo, inteso come insieme ordinato di schede individuali e relazionate, con informazioni di dettaglio sui singoli insediamenti. Il visore offrirà l'accesso a tematismi vettoriali di riferimento, livelli di informazione storica e cartografia di base pertinente, integrata attraverso i servizi WMS pubblicati sul Geoportale Nazionale.

La rete ricostruita di insediamenti geolocalizzati costituisce, in definitiva, un indice di nomi geografici (*gazetteer*) e quindi, per favorirne l'accessibilità, si è deciso di utilizzare URIs permanenti per ogni sito, impiegando l'ontologia RDF *Linked Places*², la quale consente di descrivere e illustrare luoghi del passato, così come di temporalizzare

attributi fondamentali, quali: toponimo, geometria, tipologia e relazioni con altri luoghi di rango superiore o inferiore. Questo schema di base, la cui adozione garantisce inoltre la interoperabilità con altri repertori semantici di luoghi geografici, verrà esteso per riflettere le caratteristiche dei dati in esame e i requisiti specifici del progetto. Per la bibliografia, si adotterà invece il formato BIBO (*Bibliographic Ontology*), standard di riferimento del settore.

4. Conclusioni

Il disegno delle caratteristiche e funzionalità del portale web appena descritto costituisce il primo stadio verso la creazione del portale stesso, che occuperà i prossimi mesi di lavoro e per il cui sviluppo restano ancora da definire le soluzioni software da adottare.

La pubblicazione online dei dati, attraverso una piattaforma dinamica, aperta e interoperabile, ci permetterà inoltre di migliorare ed accrescere la banca dati attraverso la collaborazione con la società civile (Toscano 2018), per esempio per includere un apparato fotografico attualmente assente, così come pianificare fasi successive di integrazione e approfondimento attraverso tesi di laurea e di master di studenti dell'Università della Tuscia, ovvero intraprendere iniziative di collaborazione e di studio con altri gruppi di ricerca, per l'ampliamento della rete ai territori limitrofi.

Note

¹ La necessità di criteri di posizionamento del punto sulla carta si amplifica con l'avvento dei sistemi GIS, che separano la cartografia dalla sua scala di realizzazione, creando l'illusione della precisione infinita.

² Il formato Linked Places, v.1.1, sostituisce il Pelagios Gazetteer Interconnection Format (PGIF) come modello dati RDF sia per i progetti Pelagios (Recogito e Peripleo) che per la World-Historical Gazetteer.

Riferimenti bibliografici

- Almagià, R. (1960), *Documenti Cartografici dello Stato Pontificio*, Città del Vaticano.
- Andrews D. (1978), *Medieval masonry in Northern Lazio: its development and uses for dating*. In: Blake H., Potter T.W., Whitehouse D.B. (ed.), *Papers in Italian Archaeology. The Lancaster Seminar*, Oxford, pp. 391-412.
- Antonelli, M. (1895), *Una relazione del Vicario del Patrimonio a Giovanni XXII in Avignone*. *Archivio della R. Società Romana di Storia Patria*, 1895, pp. 447-467.
- Battelli, G., ed. (1946), *Rationes decimarum Italiae nei secoli XIII e XIV: Latium*, Città del Vaticano.
- Brunetti Nardi, G., ed. (1972-1981), *Repertorio degli scavi e delle scoperte archeologiche nell'Etruria Meridionale, II-III (1966-1970 e 1971-1975)*, Roma.
- Buzzi, C., ed. (1988), *Il catasto di S. Stefano di Viterbo*, Roma.
- Buzzi, C., ed. (1998), *Il "Liber quatuor clavium" del Comune di Viterbo*, Roma.
- Buzzi, C., ed. (2004), *Lo Statuto del Comune di Viterbo del 1469*, Roma.
- Cameron, F., Clark, G., Jackson, R.P.J., Johns, C.M., Philpot, S., Potter, T.W., Shepherd, J.D., Stone, M.J., Whitehouse, D.B. (1984), *Il castello di Ponte Nepesino e il confine settentrionale del Ducato di Roma*. *Archeologia Medievale*, XI, pp. 63-147.
- Campanari, S. (1856), *Tuscania e i suoi monumenti. Documenti storici*, Montefiascone.
- Carbonetti Vendittelli, C., ed. (1990), *Liber memorie omnium privilegiorum et instrumentorum et actorum communis Viterbii (1283)*, Roma.
- Carocci, S., Vendittelli, M. (2014), *L'origine della Campagna Romana. Casali, castelli e villaggi nel XII e XIII secolo*, Roma.
- Carusi, E. (1948), *Cartulario di S. Maria in Campo Marzio, 986-1199*, Roma.
- Chiovelli, R. (2007), *Tecniche costruttive medievali. La Tuscia*, Roma.
- Christie N., ed. (1991), *Three South Etrurian Churches: S. Cornelia, S. Rufina and S. Liberato*, London.
- Coarelli F., Patterson H., ed. (2008), *Mercator placidissimus. The Tiber valley in antiquity. New research in the upper and middle river valley (Rome, 27-28 February 2004)*, Roma.
- Conti, S. (1980), *Le sedi umane abbandonate del Patrimonio di S. Pietro in Tuscia*, Firenze.
- De Minicis, E., ed. (2003-2011), *Insedimenti rupestri medievali della Tuscia, I-III*, Roma.
- Del Lungo, S. (1998), *Il territorio dell'antica diocesi di Orte nella toponomastica archeologica, Orte*.
- Del Lungo, S. (1999), *La toponomastica archeologica della Provincia di Viterbo, Tarquinia*.
- Duncan, G., Reynolds, J.M. (1958), *Sutri. Papers of the British School at Rome*, XXVI, pp. 63-134.
- Egidi, P. (1901), *Le Cronache di Viterbo scritte da Francesco D'Andrea*, ASRSP, XXIV, 1901, pp. 197-252, 299-371.

- Egidi, P. (1906-1907), L'archivio della Cattedrale di Viterbo. *Bollettino dell'Istituto storico italiano per il Medio Evo e Archivio Muratoriano*, 27, pp. 7-382; 28, pp. 7-19.
- Egidi, P., ed. (1910), Statuti di Castel Fiorentino. In: F. Tomassetti, V. Federici, P. Egidi (ed.), Statuti della Provincia Romana. Vicovaro, Cave, Roccantica, Ripi, Genazzano, Tivoli, Castel Fiorentino, Roma, pp. 303-359.
- Egidi, P., ed. (1930), Gli Statuti Viterbesi del MCCXXXVII-VIII, MCCLI-II e MCCCLVI, in V. Federici (ed.), Statuti della Provincia Romana. S. Andrea in Selci, Subiaco, Viterbo, Roviano, Anagni, Saccomuro, Aspra Sabina, Roma, pp. 27-282.
- Fabre, P. (1887), Un registre Caméral du Cardinal Albornoz en 1364. Documents pour servir à l'histoire du Patrimonium Beati Petri in Tuscia au quatorzième siècle. *Mélanges d'archéologie et d'histoire publiés par l'Ecole française de Rome et d'Athènes*, VIII, pp. 129-195.
- Fabre, P., Duchesne, L., ed. (1905-1952), *Le Liber Censuum de l'Eglise Romaine*, Paris.
- Fedele, P. (1898-1899), Carte del monastero dei SS. Cosma e Damiano in Mica Aurea, sec. X e XI. *Archivio della R. Società Romana di Storia Patria*, XXI, pp. 459-534; XXII, pp. 25-107.
- Fedele, P. (1900-1903), Tabularium Sanctae Mariae Novae ab anno 982 ad anno 1200, *Archivio della R. Società Romana di Storia Patria*, XXIII, pp. 171-237; XXIV, pp. 159-196, XXV, pp. 169-209; XXVI, pp. 21-141.
- Federici, V. (1899-1900), Regesto del monastero di San Silvestro in Capite. *Archivio della R. Società Romana di Storia Patria*, XXII, pp. 213-300; XXXIII, pp. 67-128.
- Fronti, D., Romagnoli, G. (2007), Indagini archeologiche nel castello di Pian Fasciano (Tuscania). Nota preliminare. In: Corneto medievale: territorio, società, economia e istituzioni religiose. Atti del Convegno di studio (Tarquinia, 24-25 novembre 2007), Tarquinia, pp. 421-454.
- Frutaz, A.P. (1972), *Le carte del Lazio*, Città del Vaticano.
- Fumi, L. (1884), Codice diplomatico della Città d'Orvieto, Firenze.
- Gamurrini, G.F., Cozza, A., Pasqui, A., Mengarelli, R. (1972), Carta archeologica d'Italia (1881-1897). Materiali per l'Etruria e la Sabina, Firenze.
- Giorgi, I. (1878), Il regesto del Monastero di S. Anastasio ad Aquas Salvias. *Archivio della R. Società Romana di Storia Patria*, I, pp. 49-77.
- Güll, P., Fronti, D., Romagnoli, G., Wick, F., 2001, Viterbo. Indagini archeologiche 1997-1998: nuovi dati per la topografia urbana e la cultura materiale. *Archeologia Medievale*, XXVIII, pp. 275-294.
- Hartmann, L.M., Merones, M. ed. (1895-1913), *Ecclesiae S. Mariae in Via Lata Tabularium*, Vienna.
- Istituto Italiano dei Castelli, ed. (1985), *Carta dei luoghi fortificati del Lazio*, Roma.
- Jones, G.D.B. (1962-1963), Capena and the Ager Capenas. *Papers of the British School at Rome*, XXX, pp. 116-207; XXXI, pp. 100-158.
- Lanconelli, A. (1992), Dal "castrum" alla "civitas": il territorio di Viterbo tra VIII e XI secolo. *Società e Storia*, 56, pp. 244-266.

- Lanconelli, A., Romagnoli, G. (2018), *L'incastellamento nel Lazio*. In: A. Augenti e P. Galetti (ed), *Incastellamento. Storia e archeologia. A 40 anni da Les structures di Pierre Toubert*, Spoleto, pp. 379-389.
- Machado, J. (2015), *Open data and open science*. In: Albagli, S. (ed.), *Open Science, Open Questions*, pp. 189-214.
- Mallet, D., Whitehouse, D.B. (1967), Castel Porciano: an abandoned village in the Roman Campagna, *Papers of the British School at Rome*, XXXV, pp. 113-146.
- Milioni, A. (2002), Viterbo. I (IGM F. 136 I SE Capodimonte, 136 II SE La Rocca, 137 III SO La Commenda, 137 III SO Castel d'Asso, 137 II SE Viterbo, 137 IV SO Montefiascone), Viterbo.
- Molinari, A. (2010), Siti rurali e poteri signorili nel Lazio (secc. X-XII). *Archeologia Medievale*, 37, pp. 129-142.
- Monaci, A. (1904-1905), Regesto dell'abbazia di Sant'Alessio all'Aventino. *Archivio della R. Società Romana di Storia Patria*, XXVII, pp. 351-398; XXVIII, pp. 151-200.
- Nardi Combescure, S. (2002), *Paesaggi d'Etruria Meridionale. L'entroterra di Civitavecchia dal II al XV secolo d.C.*, Firenze.
- Patterson, H., ed. (2004), *Bridging the Tiber. Approaches to regional archaeology in the middle Tiber Valley*, London.
- Pinzi, C. (1887-1913), *Storia della città di Viterbo*, Roma.
- Potter, T.W. (1979), *Storia del paesaggio dell'Etruria Meridionale*, Roma 1979.
- Potter, T.W., King, A.C. (1997), *Excavations at the Mola di Monte Gelato. A roman and medieval settlement in South Etruria*, London.
- Romagnoli, G. (2006), *Ferento e la Teverina viterbese. Insediamenti e dinamiche del popolamento tra il X e il XIV secolo*, Viterbo.
- Savignoni, P. (1895), *L'Archivio Storico del Comune di Viterbo*. *Archivio della R. Società Romana di Storia Patria*, XVIII, pp. 5-50.
- Scardozzi, G. (2004), *Ager Ciminus*, Viterbo.
- Schiaparelli, A. (1901-1902), *Le carte antiche dell'Archivio Capitolare di San Pietro in Vaticano*. *Archivio della R. Società Romana di Storia Patria*, XXIV, pp. 393-496, XXV, pp. 273-354.
- Signorelli, G. (1907-1969), *Viterbo nella storia della Chiesa*, Roma.
- Silvestrelli, G. (1914), *Città, castelli e terre della regione romana*, Città di Castello.
- Silvestrelli, G. (1917), *Le chiese e i feudi dell'Ordine dei Templari e dell'Ordine di San Giovanni in Gerusalemme nella Regione Romana*. *Rendiconti della Reale Accademia dei Lincei. Classe di scienze morali, storiche e filologiche*, s. V, XXVI, 5-6, pp. 491-539.
- Sommella Mura, A., ed. (1969), *Repertorio degli scavi e delle scoperte archeologiche nell'Etruria Meridionale, I (1939-1965)*, Roma.
- Sora, V. (1906), *I Conti d'Anguillara dalla loro origine fino al 1465*, *Archivio della R. Società Romana di Storia Patria*, XXIX, pp. 397-442.
- Supino, P. (1969), *La Margherita Cornetana. Regesto dei documenti*, Roma.
- Theiner, A. (1861-1862), *Codex Diplomaticus dominii temporalis S. Sedis*, Paris.

- Tomassetti, G. (1897), Del sale e del focatico del Comune di Roma nel Medioevo, *Archivio della R. Società Romana di Storia Patria*, XX, pp. 313-368.
- Tomassetti, G. (1910-1926), *La Campagna Romana Antica, Medioevale e Moderna*, Roma.
- Toscano, M. (2008), La red de los castillos de la provincia de Viterbo, Italia (siglos X-XV). Primeros resultados de un análisis estadístico-espacial. In: *Arqueología y Territorio Medieval*, Jáen, pp. 73-89.
- Toscano, M. (2018), Where the researcher cannot get. Open Platforms to collaborate with citizens on cultural heritage research data. In: Romero Frías, E., Bocanegra Barbecho, L. (ed), *Ciencias Sociales y Humanidades Digitales aplicadas: casos de estudio y perspectivas críticas*. Granada, pp. 538-561.
- Toubert, P. (1973), *Les structures du Latium médiéval. Le Latium méridional et la Sabine du IXe siècle à la fin du XIIIe siècle*, Roma.
- Trifone, B. (1908-1909), Le carte del Monastero di S. Paolo di Roma dal secolo XI al XV. *Archivio della R. Società Romana di Storia Patria*, XXXI, pp. 267-313; XXXII, pp. 29-106.
- Vallelonga, F. (2012), Insediamenti fortificati di età medievale in un territorio di confine: l'area dei Monti della Tolfa e la valle del Mignone. In: *Il ruolo degli oppida e la difesa del territorio in Etruria: casi di studio e prospettive di ricerca*, Trento, pp. 173-222.
- Vendittelli, M., ed. (2008), *Sutri nel medioevo. Storia, insediamento urbano e territorio*, Roma.
- Wickham, C. (1978-1979), Historical and topographical notes on early medieval South Etruria. *Papers of the British School at Rome*, XLVI, pp. 133-179; XLVII, pp. 66-95.

IV.3. LA WEB COMO PLATAFORMA DE PRESENTACIÓN E INTERCAMBIO DE FUENTES PRIMARIAS

Los beneficios en el uso de la tecnología en las disciplinas humanísticas no se limitan a la gestión de la información, sino que también se aplican, entre otros, al campo de la publicación de los datos y de la comunicación de los resultados de investigación, aportando soluciones innovadoras. En este apartado se ha explorado el potencial de la Web ligado a la presentación dinámica y al intercambio de fuentes primarias de investigación, de cara a la ilustración de artículos científicos. Esta problemática está estrictamente ligada con la más amplia temática objeto de esta tesis. La digitalización de las fuentes primarias, así como su presentación y publicación en línea, constituyen en muchos casos etapas fundamentales en el diseño y en la implementación de sistemas de información web que, una vez desplegados, pueden proporcionar URI¹²⁵ de recursos digitales, útiles para compartir datos, enlazar contenido adicional e ilustrar publicaciones, científicas o divulgativas.

Los resultados de la investigación que presentamos a continuación son principalmente tres, todos fundamentalmente metodológicos y desarrollados a partir de una problemática procedente de la Geoarqueología (Goldberg y Aldeias 2018). Por un lado, en el ámbito de las posibles soluciones para la digitalización de láminas delgadas¹²⁶, se introduce y se aplica la técnica del *Manual Whole Slide Image*¹²⁷ (mWSI), desarrollada en el campo de la Patología humana. Por otro lado, se aporta una reflexión sobre el uso de las imágenes en altísima resolución de láminas delgadas como fuentes primarias que habilitan, en remoto, toda una serie de procesos de investigación hasta el momento inviables porque relegados a la visión directa e individual de la lámina a través de un microscopio físico: verificación en tiempo real de la lectura microestratigráfica, comparación remota de microfacies, contextualización de las microfotografías y descripciones colaborativas. Finalmente, junto con una serie de soluciones comerciales u ofertadas por las propias revistas especializadas, proponemos una opción basada en una combinación de software de código abierto, fácilmente implementable por cualquier investigador.

A pesar de que la propuesta descrita en el artículo esté basada en casos de uso que proceden de la Geoarqueología, la misma resulta fácilmente aplicable a otros campos de las humanidades en los que las fuentes primarias de investigación puedan

¹²⁵ Acrónimo de *Uniform Resource Identifier* (identificador de recursos uniforme): cadena de caracteres que identifica, de forma unívoca, los recursos en una red (Resolución de 19 de febrero de 2013, de la Secretaría de Estado de Administraciones Públicas, por la que se aprueba la Norma Técnica de Interoperabilidad de Reutilización de recursos de la información: [https://www.boe.es/eli/es/res/2013/02/19/\(4\)](https://www.boe.es/eli/es/res/2013/02/19/(4)), Anexo 2).

¹²⁶ En Geoarqueología, una lámina delgada es una preparación que se realiza en el laboratorio sobre una muestra de tierra para ser observada al microscopio petrográfico.

¹²⁷ Imagen de diapositiva completa obtenida manualmente.

digitalizarse en formato de imagen bidimensional (Pons 2011). Por ejemplo, serían candidatos naturales las obras de artes pictóricas¹²⁸, así como las láminas delgadas para el análisis arqueométrico. También los estudios basados en fuentes escritas tomarían ventaja desde la publicación de los documentos objeto de estudio cómo imágenes dinámicas en alta resolución, en calidad de material suplementario o anexo. En algunos de esos casos (ej. lienzos o documentos), las técnicas de digitalización y producción de giga-imágenes variarían considerablemente, en función de las problemáticas conexas a las características propias del soporte material de esas fuentes, pero toda la parte relativa a la propuesta de publicación dinámica de los datos en la Web seguiría siendo válida una vez aplicada en otros campos de estudio.

En el contexto de esta publicación, el candidato ha colaborado con los coautores en la conceptualización del artículo, en la redacción conjunta del borrador inicial así cómo en su revisión y edición y en la implementación de las sugerencias de los revisores. Por otra parte, son de su exclusiva autoría: las soluciones metodológicas y técnicas propuestas; la redacción de las secciones relativas a la digitalización de las láminas y al procesamiento, publicación e intercambio de las imágenes; el contenido de la tabla 1; y la totalidad del Apéndice 1, relativo al material suplementario.

¹²⁸ La empresa Hirox ha recientemente introducido en el mercado el microscopio 3D HRX-01 (<https://hirox-europe.com/products/3d-digital-microscope/hrx-01-new-3d-digital-microscope/>), para la digitalización tridimensional de pinturas, rompiendo con el concepción bidimensionales de las obras pictóricas: <https://web.hirox-europe.com/applications/museum/>.

Referencia bibliográfica

Gutiérrez Rodríguez, M., Toscano, M., & Goldberg, P. (2018). High-resolution dynamic illustrations in soil micromorphology: A proposal for presenting and sharing primary research data in publication. *Journal of Archaeological Science: Reports*, 20, pp. 565-575. <https://doi.org/10.1016/j.jasrep.2018.05.025>

Título original

High-resolution dynamic illustrations in soil micromorphology: a proposal for presenting and sharing primary research data in publication

Título en español

Ilustraciones dinámicas en alta resolución para la micromorfología de suelos: una propuesta para presentar y compartir datos primarios de investigación en publicaciones

Abstract

Geoarchaeology, as a significant discipline within archaeology and the geosciences, operates at many different scales and encompasses several analytical methods and techniques, including what we highlight here: (archaeological) soil micromorphology. Communication practices of micromorphological researchers have not evolved significantly or, at least, at the same rate as archaeology as a whole, which has recently experienced an explosion of technological applications related to documentation. This paper aims to explore and illustrate a new methodological proposal for thin section digitalization and for the presentation of micromorphological data in scientific articles. We started from a premise: micromorphology should be a discipline based on images, since they provide the essential data of micromorphological observation and inference. The proposal combines existing tools and techniques, as well as those from other disciplines, to increase the standard quality of images used in scientific publications and produce dynamic and contextualized illustrations of research results in micromorphology. The final goal of the workflow is to provide keys and tools that allow researchers, as producers of information, to present their data in a more effective, contextualized and informative way as article enrichments or linked illustrations in their papers. To illustrate the procedure, we use two samples from different archaeological contexts: the Palaeolithic site of Roc de Marsal (France) and the Roman city of Baelo Claudia (Spain). The benefits of this proposal are multiple: a) access and objectivity to raw research data, b) improvement in the interaction among micromorphologists, c) spreading of the informative potential of micromorphology to the research community, and, finally, d) making the reader an active participant in the information. Although there are different eligible ways of image scanning and data sharing, the combination of gigapixel thin sections and article enrichments define a step forward in communicating research results and sharing raw data, an opportunity that should be fostered and not be missed.

Resumen

La geoarqueología, como disciplina relevante dentro de la arqueología y las geociencias, opera a muchas escalas diferentes y abarca varios métodos y técnicas analíticas, incluyendo la que destacamos aquí: la micromorfología (arqueológica) de suelos. Las prácticas de comunicación de los investigadores de la micromorfología no han evolucionado de manera significativa o, al menos, al mismo ritmo que la arqueología en su conjunto, que en las últimas décadas ha experimentado una explosión de aplicaciones tecnológicas relacionadas con la documentación. El presente artículo tiene por objeto explorar e ilustrar una nueva propuesta metodológica para la digitalización de secciones delgadas y para la presentación de datos micromorfológicos en artículos científicos. Partimos de una premisa: la micromorfología debe ser una disciplina basada en las imágenes, ya que éstas proporcionan los datos esenciales de observación e inferencia micromorfológica. La propuesta combina herramientas y técnicas existentes, así como las de otras disciplinas, para aumentar la calidad estándar de las imágenes utilizadas en las publicaciones científicas y producir ilustraciones dinámicas y contextualizadas de los resultados de la investigación en micromorfología. El objetivo final del proceso es proporcionar claves e instrumentos que permitan a los investigadores, en su calidad de productores de información, presentar sus datos de manera más eficaz, contextualizada e informativa como enriquecimiento de los artículos (*article enrichment*) o ilustraciones vinculadas en sus trabajos. Para ilustrar el procedimiento, utilizamos dos muestras de diferentes contextos arqueológicos: el yacimiento paleolítico de Roc de Marsal (Francia) y la ciudad romana de Baelo Claudia (España). Los beneficios de esta propuesta son múltiples: a) acceso a los datos brutos de investigación, b) mejora de la interacción entre los micromorfolólogos, c) difusión del potencial informativo de la micromorfología a la comunidad de investigadores y, por último, d) hacer del lector un participante activo de la información. Aunque hay diferentes técnicas para escanear imágenes y compartir datos, la combinación de imágenes gigapíxeles de secciones finas con el enriquecimiento de los artículos define un paso adelante en la comunicación de los resultados de las investigaciones y en el intercambio de datos en bruto: una oportunidad que debe ser fomentada y no debe ser desaprovechada.

Keywords

Geoarchaeology, Archaeological soil micromorphology, Thin section, Article enrichments, Digital humanities, Open data.

Palabras claves

Geoarqueología, Micromorfología arqueológica de suelos, Sección delgada, Enriquecimiento de artículos, Humanidades Digitales, Datos en abiertos.

1. Introduction

Context and contextualized data in time and space are fundamental cornerstones in archaeology, and archaeologists document them in several different ways: photos, profile drawings, field recording sheets, videos, etc. (Green et al., 2014; López et al., 2016; Remondino et al., 2011). In the last decade, there has been an explosion of technological applications affecting different steps of archaeological work along the way, from fieldwork to the lab, and ultimately to publication (De Reu et al., 2014; Prins et al., 2014; Roosevelt et al.). Good examples are the application of drones, photogrammetry and Structure from Motion (SfM) technologies, GIS and the use of collaborative online databases (Albert et al., 2016; Cerrillo-Cuenca, 2017; Cerrillo-Cuenca and Sepúlveda, 2015; Hendrickx et al., 2011; Howey and Brouwer Burg, 2017; Jorayev et al., 2016; McCoy, 2017; Mozas-Calvache et al., 2012).

Archaeological soil micromorphology (hereafter, 'micromorphology' for brevity) employs thin sections made from undisturbed blocks of soil/sediment collected in the field (Courty et al., 1989). The underlying strategy of micromorphology is its use of intact samples of these materials and thus, it conserves the association, internal geometry, and microcontext of all components (minerals, bones, ceramics, etc.) within the sample (Courty et al., 1989; Goldberg 1980; Goldberg and Berna 2010; Goldberg and Macphail 2006; Macphail et al. 1990). Thus, it is possible to distinguish between original depositional aspects of the sediment — grain size/shape/composition, bedding, or coating of grains as in a mudflow — from post-depositional ones associated with diagenesis, as for example, carbonate precipitation/dissolution or phosphatic transformations of the deposits. In addition, features related to large and small-scale human activities (e.g., trampling, sweeping, stabling, agriculture/manuring, and construction) are also discernible in thin sections (Angelucci et al. 2009; Courty et al. 1989; Goldberg and Macphail 2006; Miller et al. 2010; Wattez et al. 1990). The list of applications of this technique in (geo)archaeology, while not endless, has grown dramatically over the last two decades. Micromorphology is progressively used in combination with other *in situ* analytical techniques at the microcontextual level, including Fourier transform infrared spectrometry (FTIR), μ -FTIR, SEM/EDAX, μ -XRF, and magnetic susceptibility (Mentzer and Quade, 2013; Shillito et al., 2009; Weiner et al., 1993; Wouters et al., 2016).

Despite this upsurge in use, communication strategies in micromorphological research have not evolved significantly or, at least, at the same rate as other areas in archaeology. Traditionally, scientific papers on micromorphology include dense descriptions of samples accompanied by tables summarizing and commonly semi-quantifying the contents and features in the thin sections (e.g. composition, texture of vertical distribution, quantification of fabric, and significant pedofeatures). These tables are difficult to understand and digest even for micromorphologists. Goldberg and Aldeias (2016) question whether such tables deliver realistically meaningful, diagnostic, or interpretable information. They are based on relative abundance criteria, however, and

sometimes the absence of elements can be equally diagnostic. On the other hand, it is common to find mosaics of microphotographs. These images, and their legends, generally refer to singular elements of microfacies, which do not help in the understanding of the thin sections as a whole. Finally, and because of the abovementioned causes, there is no or little connection between the micro evidence and the macroscopical archaeological record, from the profile to the whole site.

In this light, we believe that micromorphology needs to update its resources of analysis, media, and the way it communicates its data and interpretations, which, as we show in the following, can be achieved through the application of information technologies, in the spirit of Digital Humanities. This paper aims to explore a new methodological-digital approach for the presentation of data and the use of images in micromorphology. This pragmatic proposal starts from the premise that images and graphic documentation are a very significant and fundamental part of the data. Methods to scan micromorphological thin sections and produce good quality images of entire thin sections were proposed by Arpin et al. (2002) and more recently by Tarquini and Favalli (2010), and Carpentier and Vandermeulen (2016). However, here we attempt to go one step further by examining new ways to obtain very detailed and accurate scans and to present these high-resolution data in scientific publications. We will focus specifically on one fundamental source of data: the thin sections. The proposal combines existing tools and techniques, also coming from other disciplines, to increase the standard quality of images in scientific publications and produce dynamic and contextualized illustrations of research results in micromorphology. The ultimate goal is to provide keys and tools that allow researchers, as producers of information, to present their data in a more effective, contextualized, and informative way, either as article enrichments or as linked illustrations of papers published in peer-reviewed publications.

To illustrate the proposal, we will use two thin sections from two different archaeological contexts. These examples have been chosen to highlight how the use of this procedure can be relatively straightforward to explain and illustrate different types of complex events and site formation processes from diverse chrono-cultural contexts. One thin section comes from Roc de Marsal (France), a Palaeolithic cave containing both geogenic and anthropogenic processes (Aldeias et al., 2012; Goldberg et al., 2012, 2013; Sandgathe et al., 2011). The lower third of the deposits (Layers through 5 – Figure 1B2) is characterized by an abundance of imbricated combustion features mostly composed of charcoal and organic matter, burned bone, and calcareous ashes; some post-depositional alteration in the form of calcite dissolution and phosphatization occurs locally (Figures 3a-c). The second example comes from a classical urban setting, the Roman city of *Baelo Claudia* (Spain), where a suite of human-related micromorphological features is recorded that shows numerous anthropogenic activities and the different usages of a public building through time (Gutiérrez-Rodríguez et al., submitted). The thin section presented here shows a set of microfacies related, firstly, to dumping

of glass production by-products later sealed by an *opus signinum* pavement of a public building, and, finally, dumping and storage of penning residues related to secondary use of this space (Figures 4a-c).

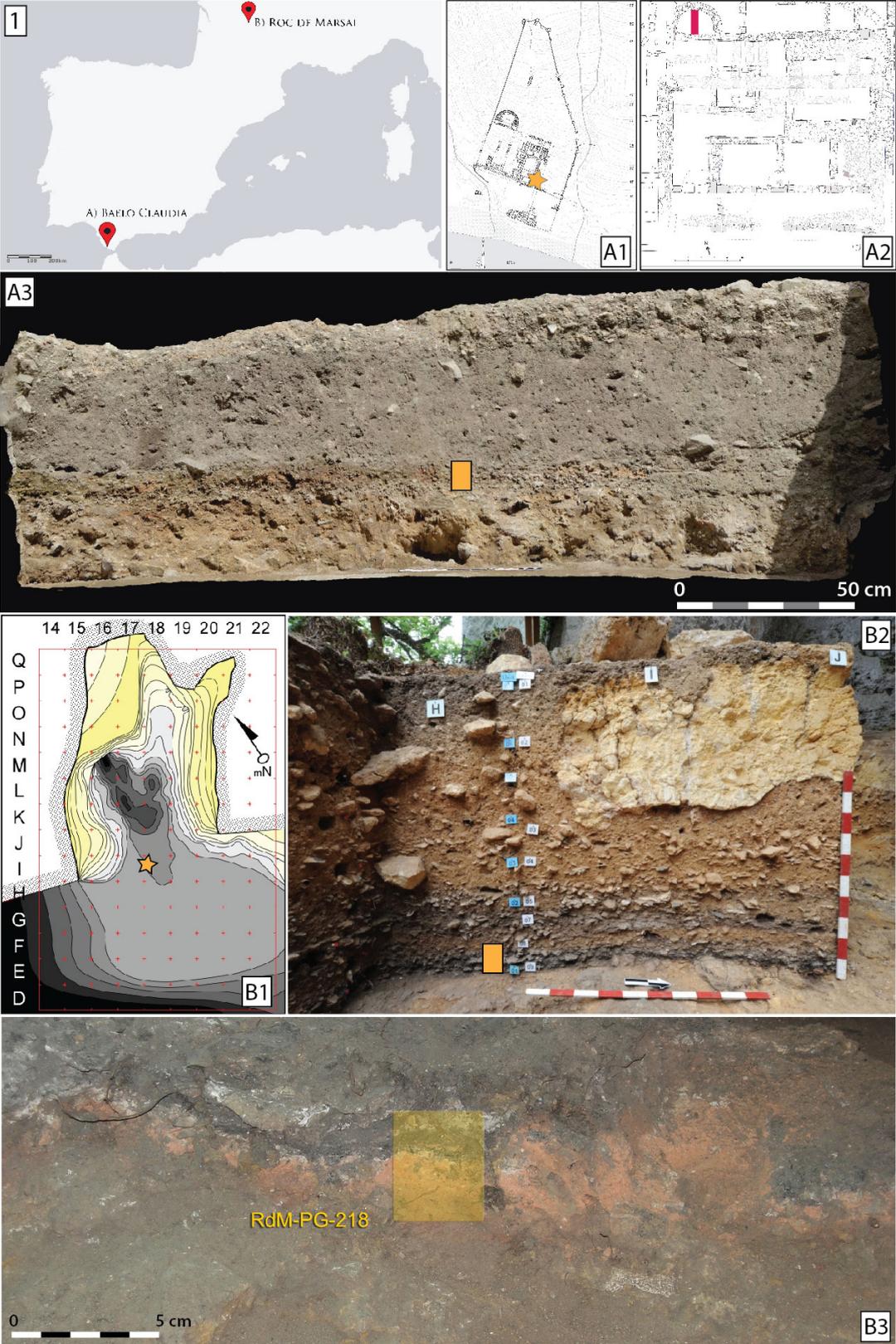


Figure 1: 1. Location of Baelo Claudia and Roc de Marsal; A1. Plan of the Roman city of Baelo Claudia, showing the sampling location (orange star); A2. South-Eastern monument plan

indicating the position of the sampled profile (red line); A3. Photogrammetry of the sampled profile with the indication of the thin section referred in this paper (orange box); B1. Plan of Roc de Marsal Cave; B2. Sampled profile with the indication of the thin section referred in this paper (orange star); B3. Detail of the field context from which the studied thin section comes from (slightly different location from B2 but in the same layer).

2. Materials and methods

2.1 SOIL MICROMORPHOLOGY

Undisturbed and oriented samples were collected from stratigraphic profiles at both sites. Blocks were stabilized by plaster of Paris bandages, and later oven dried for several days at 50°C. Impregnation was carried out under vacuum with polyester resin (Palatal P4-01), styrene monomer, and MEK catalyst. Two thin sections were used for this methodological investigation, and each was examined. They were analyzed under plane-polarized (PPL), cross-polarized (XPL) and oblique incident (OIL) light, as well ultraviolet epi-fluorescence microscopy (UV). Descriptive criteria follow Courty et al. (1989), Stoops (2003), Goldberg and Macphail (2006), and Macphail and Goldberg (2010).

2.2 MICROFACIES ANALYSIS

Archaeological deposits, which contain mixtures of geogenic and anthropogenic inputs, can be quite complex. By their very nature, such deposits include depositional and post-depositional modifications, where lithological changes are common either horizontally or vertically. Thus, it is a methodological advantage to be able to examine them in terms of individually defined lithological packets. Here we use the microfacies concept in the sense of Flügel to refer to the arrangement of sedimentary constituents by distinct and recurrent groups of similar composition and organization within a particular thin section (Flügel 2004). Thus, through microfacies analysis, similar characteristics of lithological composition, geometric association, and post-depositional changes can be grouped, thus allowing us to recognize patterns in different thin sections. This is based on the principle that distinct actions, depositional environments and post-depositional processes produce a particular set of microfacies units that, in turn, can be associated with a specific microfacies type (Courty, 2001; Flügel, 2004; Goldberg et al., 2009; Karkanas & Van de Moortel, 2014; Aldeias & Bicho, 2016).

2.3 THIN SECTION DIGITALIZATION

Thin section digitalization is a complex subject, which is only now being taken into consideration. Haaland et al. (forthcoming) have recently summarized in detail many of the current methods that are used to obtain digital images of thin sections, including their respective pros and cons. They principally focus on techniques of acquiring an image and to a lesser extent the use of such photos. This paper follows along this

trajectory by illustrating how such products can be treated as primary data resources and how they can be shared among the research community. The concept of sharing has two advantages: it provides access to the virtual thin sections both to micromorphologists and archaeologists. For micromorphologists, the ability to examine independently high-resolution, dynamic images of thin sections permits independent verification of interpretations based on these thin sections; at the same time, it expands the pool of accessible basic data for others in interpreting their own slides with comparative data. Moreover, the accessibility to archaeologists of such images fosters communication between disciplines, effectively demystifying the thin section and what it contains to non-practitioners.

There are several available methods for obtaining digital hi resolution images of thin sections (Haaland et al., forthcoming) (Fig. 2). In brief, such techniques include **flatbed scanners** (Arpin et al., 2002), **SLR photography** (Carpentier and Vandermeulen, 2016), and **film scanners** (Haaland et al., forthcoming).

A very recent technological approach, called **MWSI** (Manual Whole Slide Image), acquires the image directly from the microscope, through a camera and desktop software (<http://www.microvisioneer.com/>). It involves using a mechanical stage to systematically move the sample until the entire area of the thin section is covered. The main advantage is that the researcher can acquire the image with the same level of magnification of the microscope, essentially creating a digital copy of the thin section without losing raw data from the original source. The results are ultra-high-resolution scans (above 20,000 dpi), entering the realm of gigapixel macro images: for a standard thin section sample scanning at 50x magnification, the final resolution can reach 2 gigapixels, which means 2,000 megapixels. Images acquired directly from the microscope can be assimilated into raw research data, as they retain the full informative potential of the original sample, one that is not subject to deterioration over time, as happens to thin sections. While this is definitely a positive outcome, it also requires adequate hardware and software for desktop image processing, storage, backups, and Web uploads. Another possible disadvantage to take into consideration is the time spent acquiring the image, as manual scanning is affected by the size of the sample and the level of magnification. In our tests, for example, scanning a thin section of 54 mm x 76 mm at 50x magnification took 9 minutes; the XPL version was produced using the polarizers of the petrographic microscope and the scanning took 15 minutes.

For this paper we present images obtained using three techniques (flatbed scanner, film scanner and MWSI) to illustrate how such digital images can be shared among researchers. For example, being a widely available solution, an Epson Perfection V850 flatbed scanner was used for the Roc de Marsal thin section (see below) (Fig. 3). In a second, we scanned the Baelo Claudia thin section with a film scanner, (Nikon Super CoolScan 9000 ED at the Institute for Archaeological Science, University of Tübingen). (Fig. 4c–d). For both flatbed and film scanners, an XPL version of each thin section can

be obtained by inserting two perpendicularly oriented polarized filters, one on top and the other below the thin section. We followed this procedure to digitise both thin sections from Baelo Claudia and Roc de Marsal.

In order to allow a direct understanding of the potential of this technique in comparison with the other options described above, we decided to rescan the *Baelo Claudia's* sample with MWSI (Figs. 4-7). Thus, in this paper we provide three high-resolution images per each thin section: PPL, XPL and one showing identified microfacies. This allows the reader to have an overall view of the samples and also provides a clearer visualization and understanding of microfacies analysis.

2.4 DIGITAL IMAGE PROCESSING

Processing the images produced in the previous step is crucial both for optimizing their size and for improving their legibility. For best results, we recommend saving the files in TIFF format, using LZW compression. Depending on the methods used to digitise the thin section, this can be done directly or successively using image editing software. The final file size of any of these images will be directly proportional to their resolution: if the image has been acquired with a high-density film scanner it can be in the range of 150-200MB, as opposed to around 1500MB if produced by manual scanning directly from the microscope. Dealing with such large files can be problematic so, if needed, images can be further processed using JPEG compression, quality 95. This will significantly reduce the file size, but can also affect image quality, so you should double check the final results before moving forward. Following this, a standard image processing workflow is applied, cropping the thin section edges and improving the image in terms of brightness, contrast, white balance, and sharpness. Additionally, thin section artefacts (e.g. sample labels) can be removed if they affect the legibility of the geoarchaeological information.

2.5 IMAGE PUBLISHING AND SHARING

The digital copy of the thin section obtained is now a new valuable source of information that, according to its quality and resolution, embraces a large amount of the informative potential of the original thin section, combined with the ease of use and sharing features of a digital file. Anyway, to fully exploit their potential, we need to continue with one final step: publish these images in a Web service.

If the thin sections are used to illustrate an article in a scientific journal or a chapter in a book, the best way is to integrate these images within the article itself, as dynamic illustrations. Some publishers already provide tools that make it easy to include in research papers ultra-high-resolution images, which generally exceed 30,000 by 30,000 pixels.

If this option is not available, or the intention is just in having an online shareable version the thin section, researcher can resort to the use of online commercial platforms, like EasyZoom (<https://www.easyzoom.com>) or Gigamacro (<https://viewer.gigamacro.com/>). They offer the benefits of being easy to use and they generally provide free or freemium and powerful ready-to-use solutions; on the other hand, they bear the disadvantages of being out of the control of both the researcher and the publisher, so their features are not customizable and the consistency of the uploaded resources is not guaranteed in the medium/long term.

Finally, there is also a DIY solution available to the researcher. It involves some further image processing, downloading some code from the Internet, preparing an HTML document, and finally uploading a bunch of files and folders to a Web-server. This solution is a bit more technical, so we have prepared a step-by-step didactic guide, explained with all the necessary source code that can be downloaded as supplementary material, along with a fully working demo package. The procedure can be used to enrich articles published in journals that do not offer their own solution or as an independent possibility for the researcher to create his/her own catalogue of zoomable images, easily shareable, customizable and integratable with other Web instruments. In this sense, this last option offers also some unique advantages. For example, OpenSeadragon, which is the most used Open Source software in this field, is free, backed by a strong and active community of developers, is easily customizable and is extendable. Furthermore, it does not lock your data into a proprietary solution, it supports several input formats, it is light to load, and is mobile compatible. For the specific purpose of displaying geoarchaeological data, its best feature resides in the possibility to add and customize live image filters.

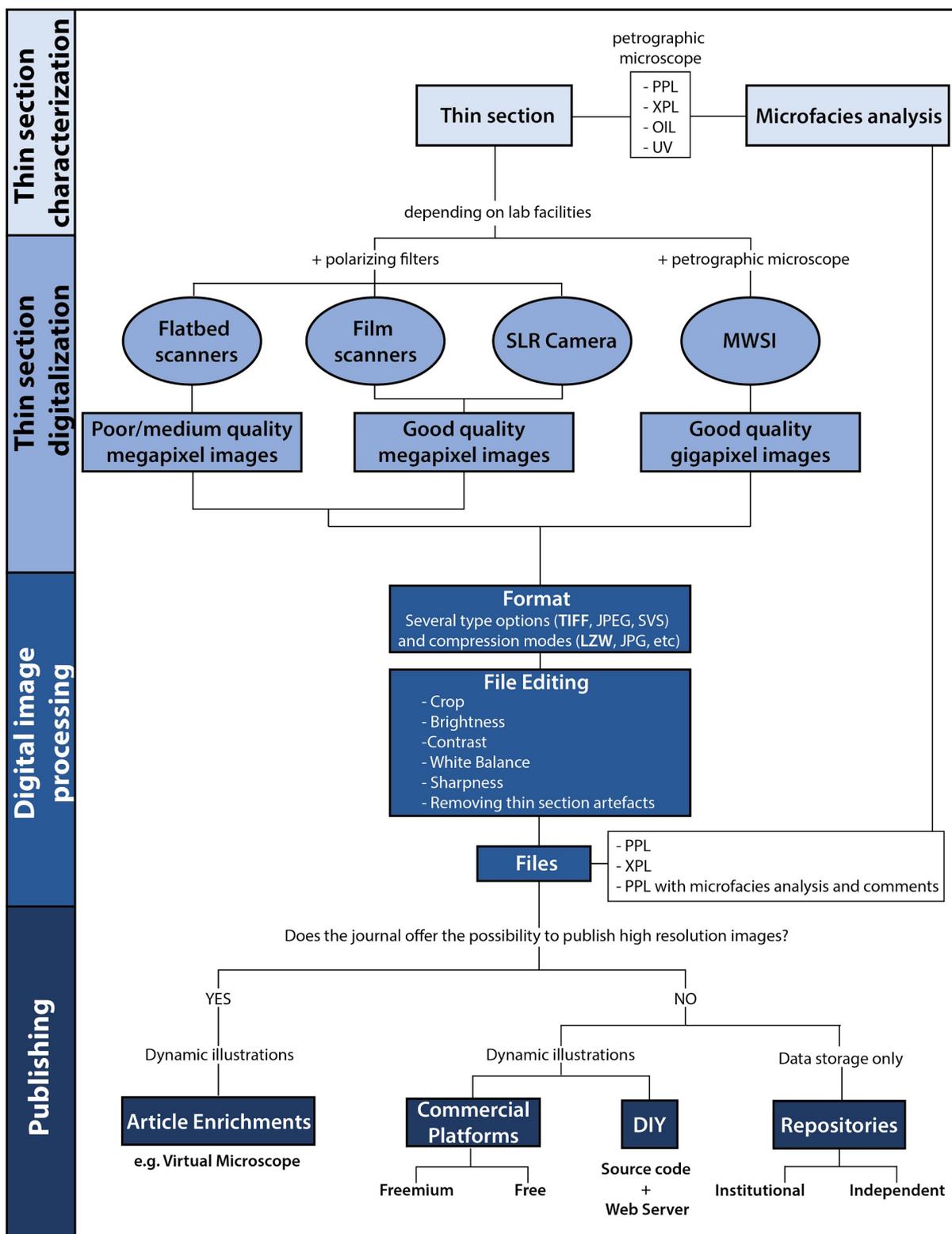


Figure 2. Diagram showing the different possibilities of thin section high-resolution digitalization, editing and publishing

3. Examples of implementation of the procedure

Having reviewed several options of thin section digitalization, editing and publishing, we proceed to analyse two samples, one coming from Roc de Marsal and another from

Baelo Claudia. Textual description and interpretation are combined with dynamic high-resolution illustrations in order to provide a richer and more legible microfacies analysis.

3.1 ROC DE MARSAL (FIGURE 3)



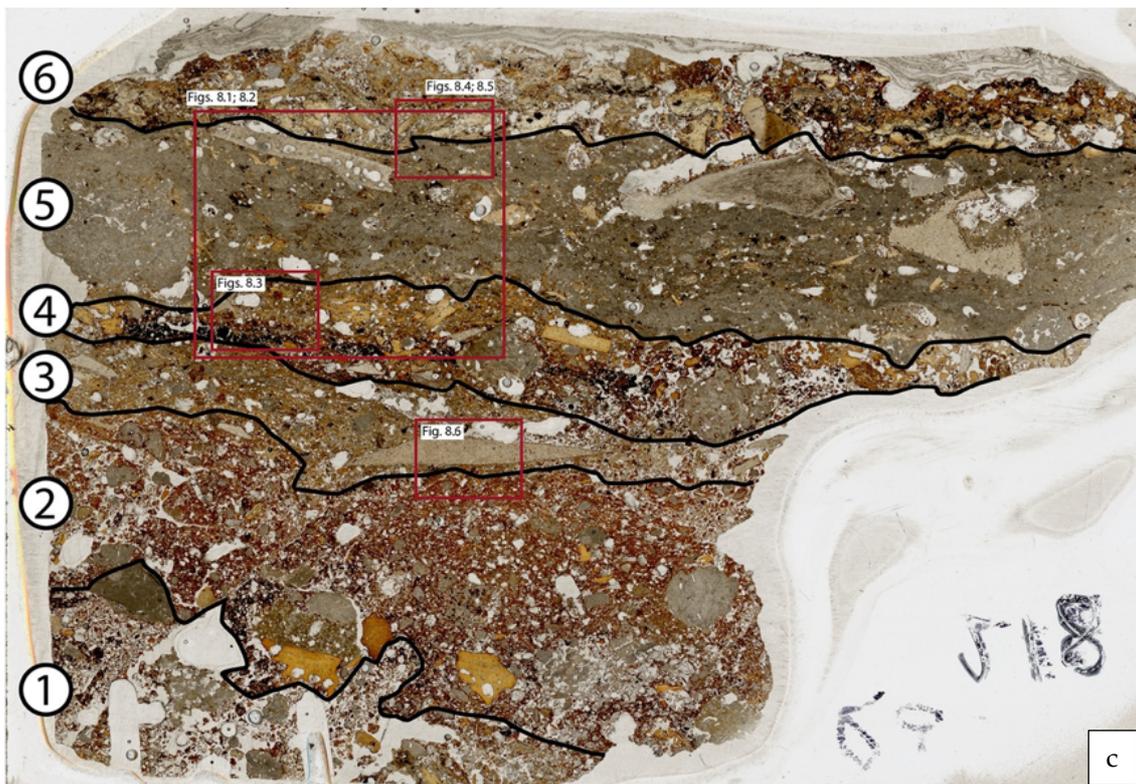


Figure 3. Flatbed scanning of the thin section RdM-PG-318, corresponding with imbricated combustion features mostly composed of charcoal and organic matter, burned bone, and calcareous ashes. Some post-depositional alteration in the form of calcite dissolution and phosphatisation occurs locally. Fig. 3a is a PPL scan of the thin section (<https://goo.gl/Uv38Qv>); fig. 3b is an XPL scan (<https://goo.gl/ZjCqDy>); fig. 3c incorporates microfacies analysis and microphotographs contextualization (<https://goo.gl/hJU6fa>).

Description of the identified microfacies is presented below and illustrated in Figure 3.

mF Unit 1 - Dark reddish brown, with fine fraction composed of fine sand mixed with yellow to reddish orange silty clay (FTIR results show kaolinite smectite/illite and calcite). The coarse fraction is composed of quartz sand, yellow bone fragments (most not calcined), and limestone, some of which are dissolved and phosphatized.

mF Unit 2 - Bright reddish brown, with fine fraction composed of dispersed calcite aggregates (ashes?) impregnated with fine bright reddish brown clay (FTIR shows kaolinite, smectite/illite and calcite). The coarse fraction is composed of quartz sand, bone fragments (FTIR results show that most are not calcined), limestone fragments (some dissolved), and rubefied nodules/ soil aggregates with FTIR showing no kaolinite and some only with smectite/illite.

mF Unit 3 - Calcitic ashes mixed with sand and rubefied soil aggregates. The FTIR analyses show no kaolinite or smectite/illite and point to a temperature at or above

700° C, yellow bone that is not calcined (shown by FTIR), and pale yellow bone that is calcined (FTIR) bone.

mF Unit 4- Calcitic ashes mixed with charred organic matter and yellow bone fragments (FTIR shows no calcined bone).

mF Unit 5- Compact, whitish calcitic ashes mixed with bone fragments. The FTIR results show both calcined and non-calcined bone, but bone is completely calcined at the top; flint, yellow isotropic nodules (with FTIR peak at 1070 cm⁻¹ compatible with amorphous silica), and a few rubefied soil inclusions that do not show FTIR absorption of kaolinite or smectite/illite, suggesting that they were heated to at or above 700 °C (Berna et al., 2007). The ashes exhibit organic and Fe-Mn staining.

mF Unit 6- Brown partially dissolved calcitic ashes with brown Fe/Mn staining, fragments of yellow and charred bones (IR shows some are partially calcined), and elongated domains of isotropic phosphate and dissolved bones.

At a first crude glance, the thin section scan would appear to represent a typical view of a simple *in situ* fireplace: an ashy upper part (unit 5) overlying a blackish band (unit 4) on top of a reddish (rubefied) base (unit 2). However, closer inspection shows that this depiction can be made more nuanced (for details, see Goldberg et al. 2012). Units 5 and 4 indeed represent an intact combustion feature; however, ashy unit 3 intervenes between the charcoal-rich unit 4 and the reddish layer (unit 2; the underlying unit 1 is unaltered cave clay that rests on bedrock). As such, unit 3 appears to represent a previous fire event that produced the underlying rubefication, and not the overlying charcoal-ash couplet of units 4 and 5, respectively. Furthermore, the upper part of unit 5 – which contains burned bone and chert – has been eroded as shown by the sharp contact with the overlying unit 6. In addition, the ashes in the upper part of unit 5 are partially decalcified and replaced with phosphate (dahllite). Both of these features indicate that these ashes were exposed to the surface during a period of abandonment and non-deposition, allowing for mineralogical transformations (dissolution of carbonate, phosphatisation) that were mitigated by bat guano. Thus, there is a temporal gap between the accumulation of units 5 and 6.

3.2 BAELO CLAUDIA (FIGURES 4-7)

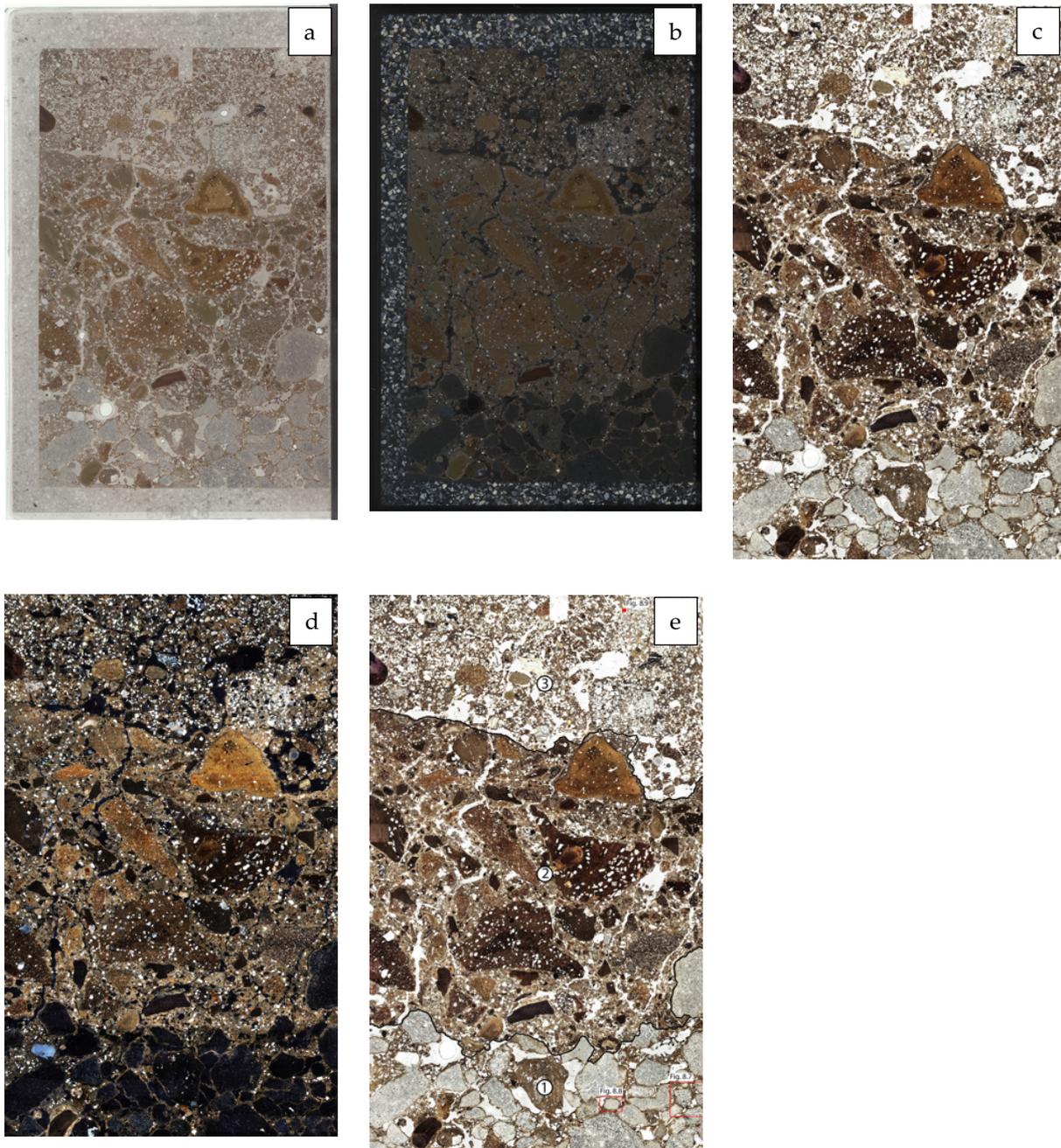


Figure 4. Film and MWSI scanning of the thin section BC-EX-3.2, corresponding to a set of microfacies related, firstly, with dumping of glass production by-products later sealed by an opus signinum pavement of a public building, and, finally, with a secondary use of this space for dumping and storage of penning residues. Fig. 4a is a PPL film scanning of the thin section (<https://goo.gl/fy55Nk>); fig. 4b is a XPL film scanning (<https://goo.gl/kgBvUP>); fig. 4c is a PPL MWSI scanning of the same thin section (<https://goo.gl/emsUW5>); fig 4d is a XPL MWSI scanning (<https://goo.gl/gNbjC2>); fig. 4e is a PPL MWSI scanning that incorporates microfacies analysis and microphotographs contextualization (<https://goo.gl/9W5WmR>).

Description of the identified microfacies is presented below and illustrated in Figures 4-8

mF Unit 1, Well sorted sand to gravel-size glass slags: Coarse material is dominantly composed of sand-size subrounded glass slags with palagonite coatings, and few sand-size rounded mortar and *opus signinum* fragments. They show a single space porphyric c/f related distribution. This microfacies is very porous and mainly composed of compound packing voids. Fresh organic tissues with spherical mesofaunal excrements are localized in voids.

mF Unit 2, Opus signinum: This microfacies is a pavement composed of micritic calcite cement with dominant gravel-size (>2 mm) clasts (temper) that include brick, burned daub and pottery fragments, and sparitic calcarenite with bioclasts in a 60/40 ratio of coarse (gravel and sand) to fine (cement) material.

mF Unit 3, Poorly sorted aeolian sand with penning residues: In this microfacies coarse material is composed of rounded sand-size aeolian quartz grains, foraminifera, shells, beach rock fragments, rounded sand-size quartzite, rounded sand-size microcline and silt-size rounded heavy minerals. The coarse fraction is supported by clays containing silt-size muscovite and silt to sand-size charcoal. Coarse and fine components show a single spaced porphyric c/f related distribution. Groundmass exhibits a crystallitic b-fabric and a faecal spherulite-rich micromass. There are authigenic nodules of phosphate and elongated tubular phytoliths dispersed in the groundmass. Porosity is high, and mainly composed of channels. It shows a channel to spongy microstructure.

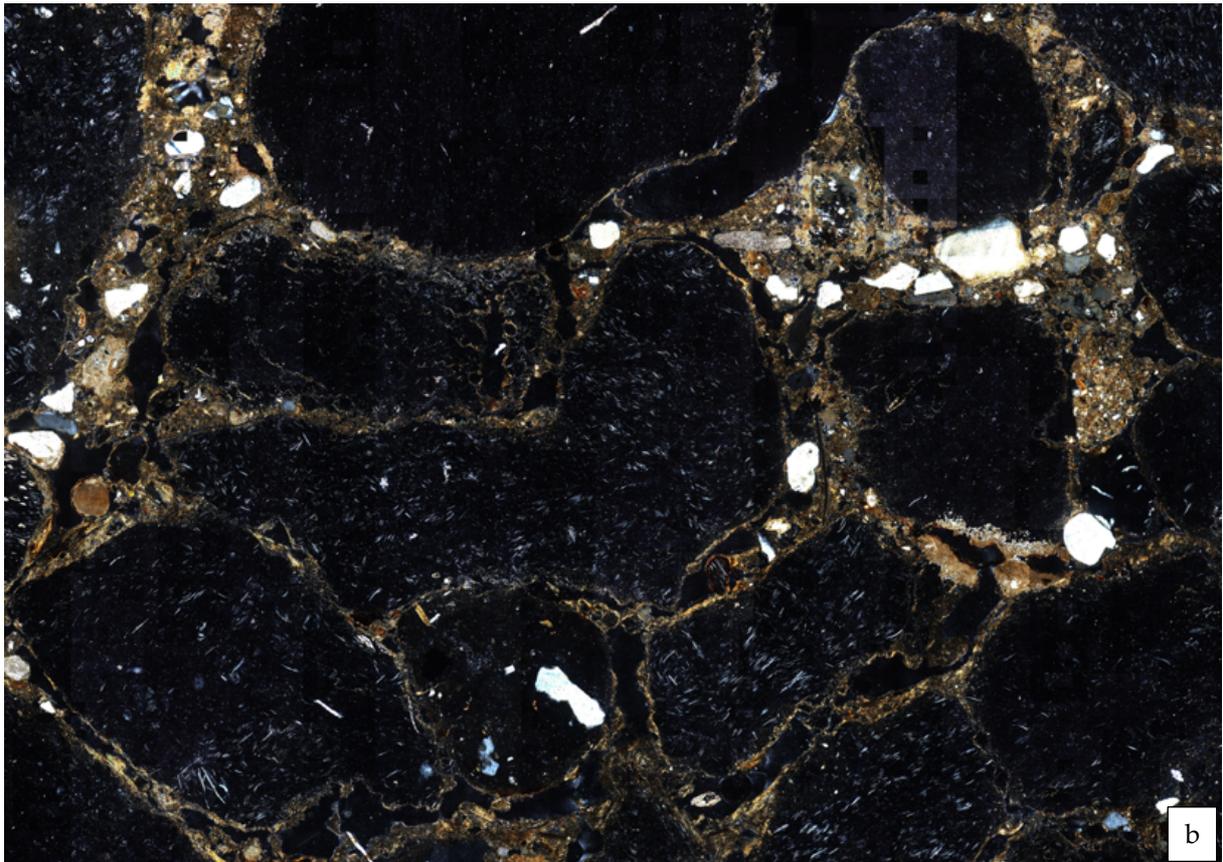
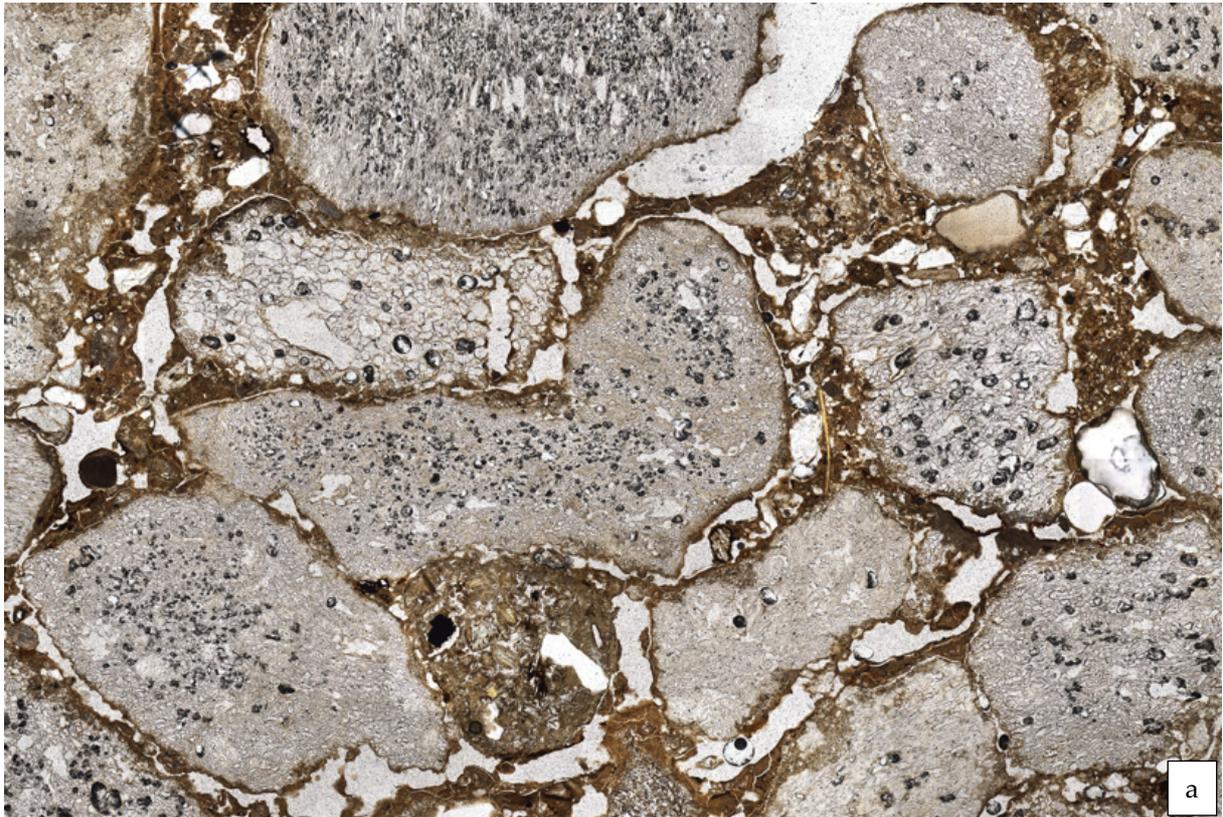


Figure 5. MWSI scanning of microfacies unit 1 (close-up at 100 magnifications). Fig. 5a is a PPL scan (goo.gl/aK9pFC); fig. 5b is an XPL scan (goo.gl/SnYQM1).

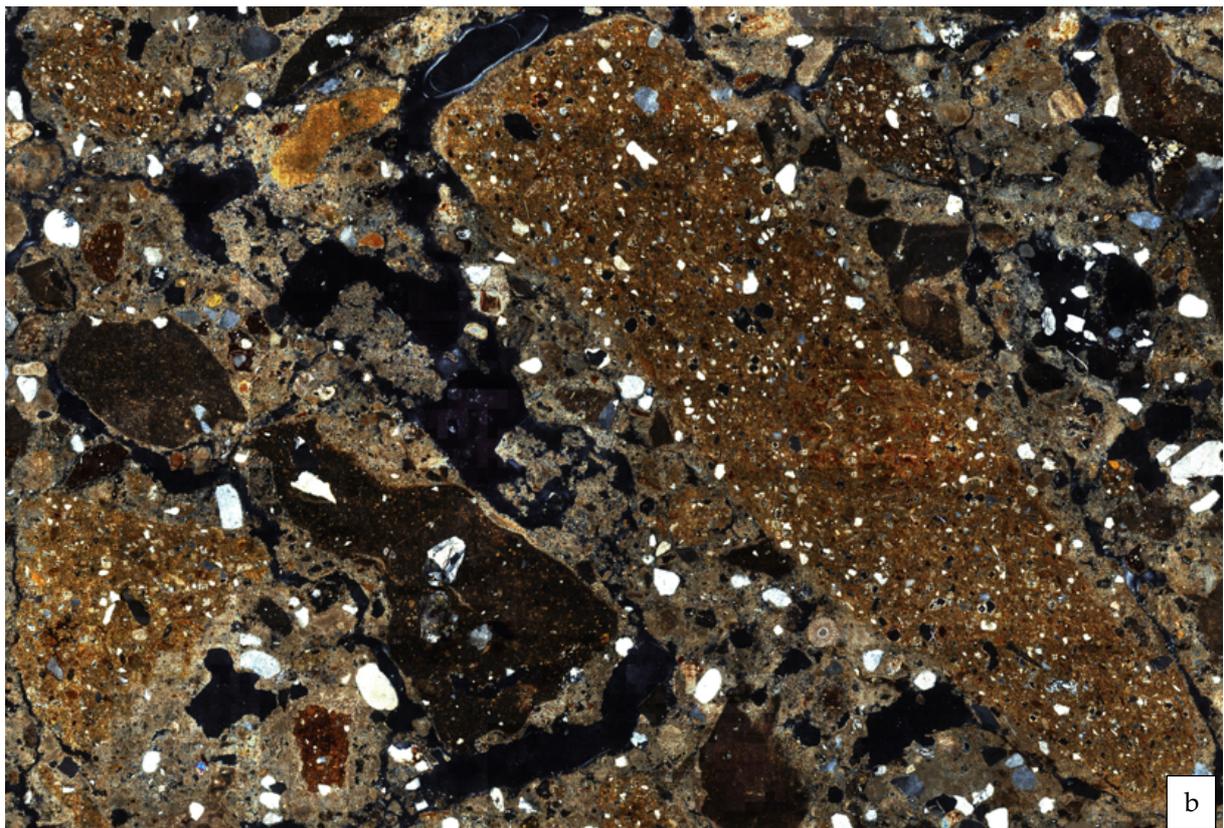
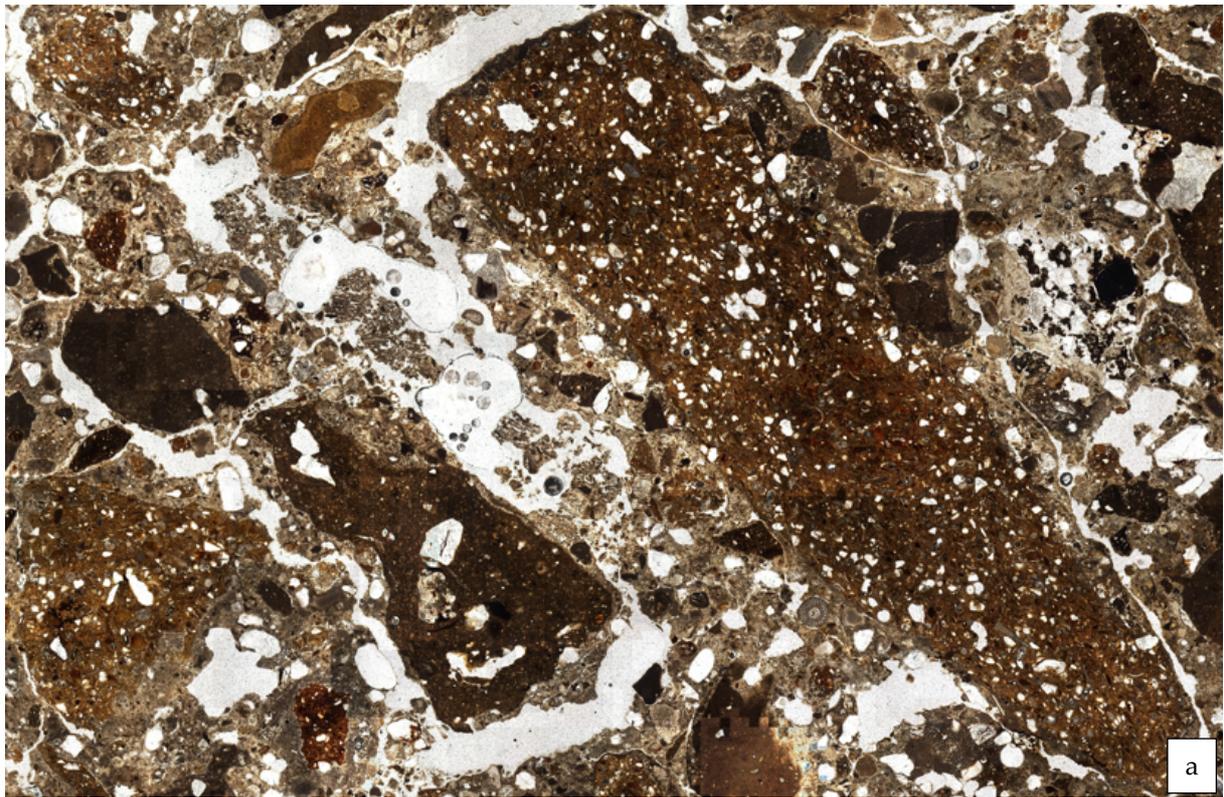


Figure 6. MWSI scanning of microfacies unit 2 (close-up at 100 magnifications). Fig. 6a is a PPL scan; fig. 6b is an XPL scan.

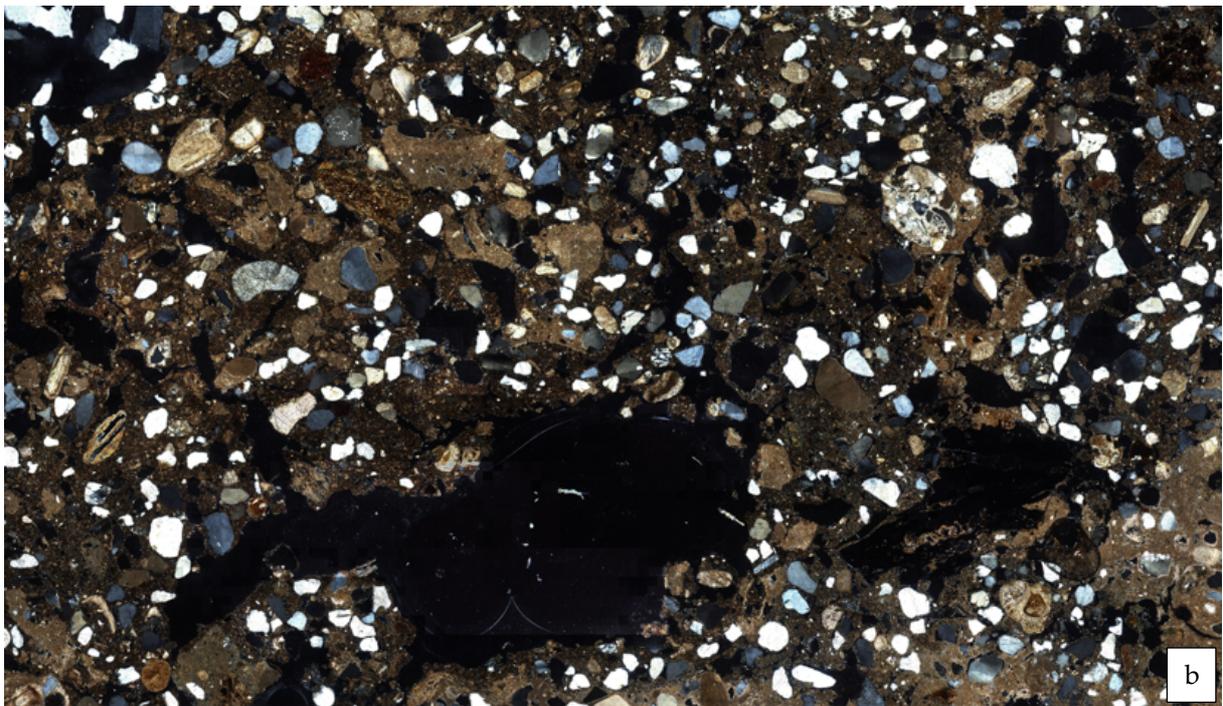


Figure 7. MWSI scanning of microfacies unit 3 (close-up at 100x magnification). Fig. 7a is a PPL scan; fig. 7b is an XPL scan.

The microfacies sequence observed in the Baelo Claudia sample is related to pavement construction in a public building and the amortization of the space for secondary usages once maintenance activities stopped and sedimentation occurred. At the bottom of the sequence, mF Unit 1 is mainly composed by transparent glass slags of fibrous and vesicular pattern (Figs. 4a-e; 5a-b; 8, 7-8). All the fragments show yellow to orange clear, transparent, isotropic, smooth, and concentrically banded coatings of “gel palagonite” (Stroncik & Schmincke, 2002). It is accepted that these features are generated by

an interaction of glassy materials with aqueous solutions at ambient temperature and the influence of microorganisms (Sedov, Stoops, & Shoba, 2010; Thorseth, Furnes, & Tumyr, 1995). Although slags of mF Unit 1 resulted from glass production, their deposition must be understood as a preparation layer of mF Unit 2, the *opus signinum pavement* (Figs. Fig. 4a-e; 6a-b). mF Unit 2 shows a micritic calcite cement and rock fragments and anthropogenic material as binders. Sediment up to this pavement, mf Unit 3, is characterized by a fecal spherulite-rich micromass (Figs. 4a-e; 7a-b; 8, 9). These spherulites, together with authigenic phosphate nodules and phytoliths are some of the major lines of evidence used to identify organic-poor livestock dung. Spherulites, phosphates, and phytoliths are arranged with clay containing silt-size muscovite, aeolian quartz sand and anthropogenic materials. Some of them such as charcoal, bone and fishbone are linked to food processing, consumption, and refuse. Others constitute construction material debris of *opus signinum* and plaster. Root action, evinced by calcitic hypocoatings, and the passage of soil fauna disturbed the integrity of the structure and microstructure of the dung. Synchronic aeolian processes and bioturbation suggest longer exposure, lower frequency of discard and, consequently, long-term formation of a midden deposit related with penning residues management (Matthews et al., 1997; Shillito & Matthews, 2013). Therefore, dung components are dumped and reworked in a midden created inside the former public building after cleaning and maintenance activities had ceased. These stratigraphic changes imply a pronounced reorganization of urban space, introducing rural activities in the monumental area of *Baelo Claudia* during Late Antiquity.

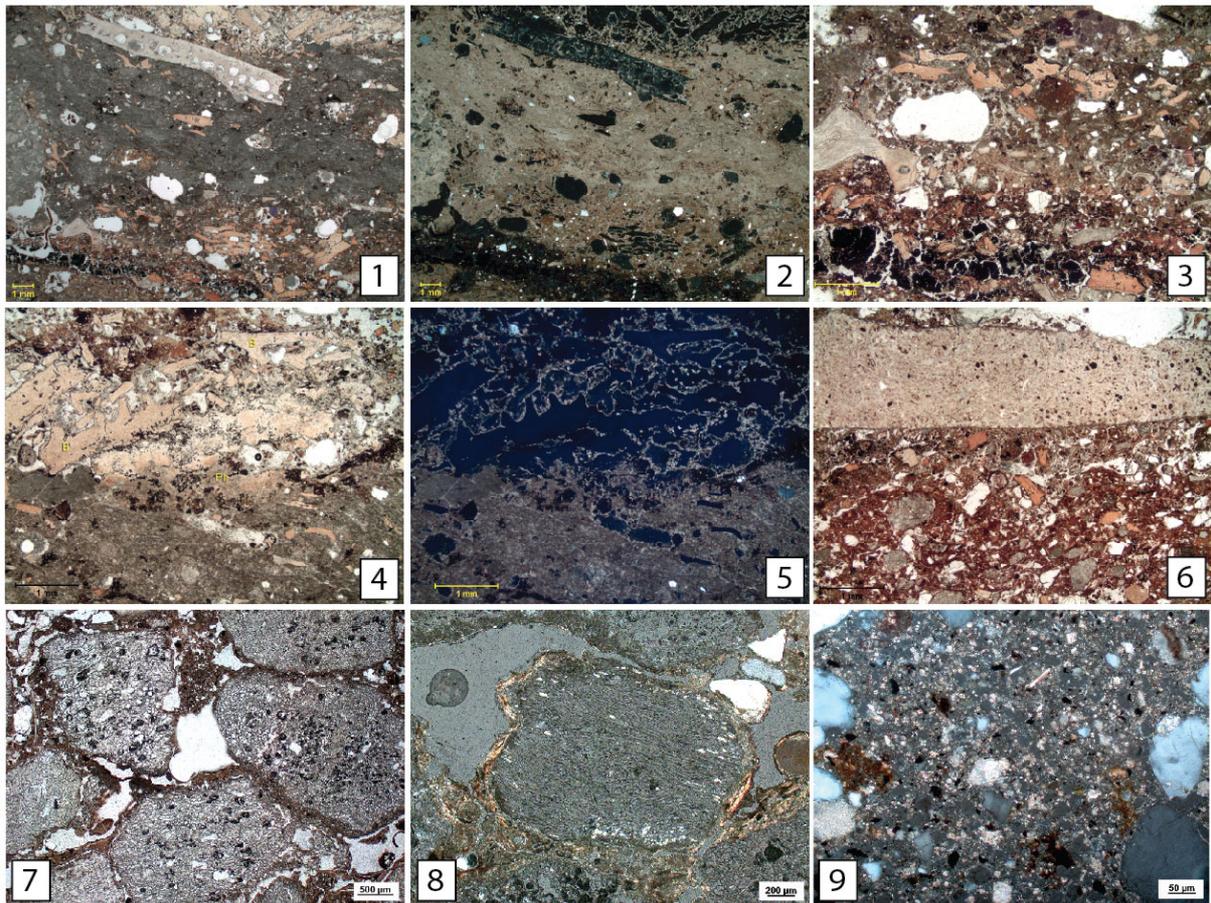


Figure 8: Microphotographs of the studied thin sections, contextualized in figs. 3c and 4e. **1, 2.** PPL view of Units 4, 5, and 6, showing mostly the calcareous ashes of unit 5 with a large burned chert fragment near the top; note the yellowish bone-rich, phosphatized part that are isotropic in the XPL view (**2**); **3.** detailed view of unit 4 and the very base of unit 5 showing charcoal-rich base becoming richer in ash with bone in unit 5, PPL; **4, 5.** Detailed view of the upper part of 1 and 2 with the yellow, phosphatic part of unit 6 in the upper half (PPL), which is isotropic in XPL (**5**); **6.** Chert flake at base of unit 3 resting on rubefied silty clay with sand-size inclusions of bone and limestone grains; **7.** Well sorted sand to gravel-size glass slags; **8.** Same in XPL, showing palagonite coatings; **9.** In this microfacies anthropogenic debris appears in a spherulite-rich micromass suggesting dung reworking and penning residues origin. This small feature is the only one difficult to appreciate in MWSI scans at 100 magnifications (Fig. 7b).

4. Discussion

4.1 ARE WE READY FOR GIGAPIXEL THIN SECTIONS?

The archaeological deposits of Roc de Marsal and *Baelo Claudia* described above were formed through the complex interaction of anthropogenic actions and natural processes. Normally, it is difficult to illustrate with simple thin section microphotographs the complexity of these deposits (associated processes) and their archaeological significance even to the informed (micromorphologists) and the novice (archaeologists and others). Modern digitalization methods, however, have come a long way in bridging this communication gap.

Different digitalization methods (Haaland et al., forthcoming) have their own strengths and weaknesses. All of them are very useful at the so-called mesoscopic level of thin section analysis (Goldberg and Aldeias, 2018), which maintains a visual and analytical continuity from the site as a whole down to the microscale (Courty et al., 1989; Goldberg and Macphail, 2006). All methods allow for microfacies identification, which has been considered as a gap between the macroscopic record and the micro evidence (Goldberg et al., 2009). Likewise, all methods provide files that can be easily shared, uploaded to the Internet, or implemented in publications.

However, these methods show significant differences in terms of image quality and resemblance to the original research data. Whereas flatbed scanners produce medium or poor-quality results, significantly better products can be obtained using high-end film scanners, or the MWSI method; the latter remarkably increases image resolution (DPI and $\mu\text{m}/\text{pixel}$), when scanning at 50 magnifications or at 100 magnifications for close-ups (Table 1). Additionally, when scanning directly from the petrographic microscope, MWSI provides greater chromatic fidelity, especially for the background colour of the glass and transparent minerals in PPL. Thus, it delivers not only better products for microfacies analysis and thin section description, but also for image analysis and GIS implementation.

According to their maximum resolution, film scanner images allow for good identification of gravel to fine sand-size fractions, but they are not suitable for characterizing silt- and clay-size sedimentary components and features. Nevertheless, it is possible to identify features of this size using MWSI due to a significantly higher $\mu\text{m}/\text{pixel}$ cell resolution, especially when scanning at 100 magnifications (Table 1). MWSI provides a very complete image, with different levels of information, but it also exhibits some limits. Scanning occasionally produces artefacts as passage lines, and it slightly modifies the colour of some features, although these can be adjusted with scanning software settings and image post processing.

| | Flatbed scanner | Film scanner | MWSI 50 magnifications | MWSI 100 magnifications |
|--|-----------------|----------------|------------------------|-------------------------|
| Light settings | PPL, XPL, RL | PPL, XPL | PPL, XPL, RL, OIL, UV | PPL, XPL, RL, OIL, UV |
| Resolution (DPI) | 2,400 | 4,000 | ~21,500 | ~43,000 |
| Cell resolution ($\mu\text{m}/\text{pixel}$) | 10.5 | 6.35 | 1.16 | 0.58 |
| Image size (pixels) | 4,700 x 6,600 | 7,600 x 11,900 | 31,600 x 53,000 | 66,500 x 108,500 |
| Megapixels | 31 MP | 90 MP | 1.67 GP | 7.2 GP |
| Approx. compressed TIFF file size | 70 MB | 150 MB | 1.6 GB | 7 GB |

Table 1: Comparison of thin section high-resolution digitalization methods.

What are the ramifications of these methods and images for the discipline in terms of analysis and illustration of the results of our research? A digital, zoomable gigapixel version of the thin section under PPL and XPL, represents a powerful dissemination tool and a new analytical source of information. The zoomable view enhances the possibility of connecting the individual elements within the entire sample over a variety of continuous magnifications. This strategy increases the understanding of how different microfacies relate to each other in the context of the whole thin section. It allows for a continuum of observation from the macroscopic scale up to the microscopic scale (at 40× magnification for samples digitized with a film scanner or a photographic stand, or even 100× for MWSI). Thus, it not only essentially reproduces and enhances observations with the petrographic microscope, but also furnishes additional features such as taking measurements or applying live image filters (e.g., brightness, contrast, colour corrections, etc.). Consequently, a zoomable dynamic view helps in better describing and organizing microfacies or other features. Moreover, these images open up the possibility of carrying out collaborative real-time descriptions that encompass two or more researchers simultaneously observing the thin section; this option is precluded by observing the thin section only through the ocular lens of the microscope in the author's laboratory.

In addition, once published online, with a permanent URL, these images represent another important result, as they offer additional opportunities that micromorphologists, as a community working together, should be able to exploit. They offer the possibility to easily share the resource with other researchers, from large desktop workstations down to portable mobile devices, contrasting PPL and XPL versions, comparing the raw data with the microfacies identification and so on. This strategy raises the limits of the thin section as a physical object, offering something that the traditional series of static thumbnails and weakly contextualized microphotographs has never been able to offer.

Finally, regarding publication, if resolution and image quality is comparable with the view from the microscope, as it is possible using MWSI, those scans will represent a unique combination of research data and article enrichments.

4.2 DIGITAL HUMANITIES FOR GEOARCHAEOLOGY

In recent years, we are witnessing the progressive incorporation of Digital Humanities into the routines of archaeological research (Earley-Spadoni, 2017; Wilson and Edwards, 2016). This philosophy implies a multidisciplinary dimension that entails among others: a) rethinking traditional applications of technological tools; b) new methodological approaches to the study of sediments; c) the complementary use of text and images; d) the application of open formats and open source software, and e) embracing the Open Data initiative in the sharing raw scientific data on the Web (Bocanegra Barbecho, 2016; Burdick et al., 2012; Schreibman et al., 2004; Wouters et al.,

2012). These developments constitute a step forward toward more effective interdisciplinary and trans-institutional collaborations; in some cases, they can produce greater inputs of academic research to society, through co-creation (Bocanegra Barbecho *et al.*, 2017; Toscano, 2018).

We strongly believe that Geoarchaeology can be enriched as a discipline thanks to the implementation of the theoretical and practical fundamentals of Digital Humanities. Sharing of raw data would increase and enrich data interpretation, and would favour the implementation of new common research protocols and general knowledge transfer. In the long term, enhanced geoarchaeology dissemination strategies would stimulate collaboration among researchers, improve communication between micromorphologists and archaeologists, support a better and broader combination of geoarchaeological research in the wider archaeological discourses, and would provide improved resources for teaching and training in the discipline. A greater ease of access to the original raw data through the Internet may foster greater consultation and, therefore, facilitate comparisons and improve the interaction among micromorphologists, offering a better and more reliable understanding of thin section descriptions and interpretations.

We have to remind ourselves that the raw data are the thin sections, which when considered as a whole, foster the contextualization of multiple sets of evidence (microfacies, archaeological materials, pedofeatures, etc.) that can be linked to analytical data sets of other specialists. Thus, providing readily accessible, high-resolution, zoomable scans of these primary sources of information, along with interpretative notes and dynamic tools, converts the reader into an active subject. It delivers greater objectivity to data interpretation and, at the same time, opens up the possibility of re-interpreting and plunging deeper into the sample and the site as a whole.

Recently, one of us in collaboration with Vera Aldeias questioned why does archaeological micromorphology has such relatively little traction in (geo)archaeology (Goldberg & Aldeias, 2016). One of the reasons argued in the paper was that there is not a common language between micromorphology and archaeology (and their practitioners). These products can be easily applied to consultancy, especially for cases in which the specialist in micromorphology needs to create very understandable reports for non-practitioners. We think that this approach can help demystify the discipline. Although the ways to deal with this issue are not simple given the nature and the specificity of micromorphological jargon, providing high-quality interactive images of thin sections is a necessary step for the contextualization of thin sections in archaeological sites. This is the future challenge in communicating micromorphology, and it can be tackled by combining it with detailed planimetry, photogrammetry, GIS techniques and the Web (Goldberg and Berna, 2010; Fisher *et al.*, 2015; Haaland *et al.*, 2017).

In closing, we note that digital technologies advance at an ever-growing rapid pace and so these protocols will and should evolve rapidly to address certain issues.

5. Conclusions

The methodology described in this article is based on a few main overlapping ideas: thin section scanning, scientific paper illustrations and the use of high-resolution images in disseminating primary research data with the research community –both micromorphologists and particularly archaeologists–. The range of possibilities offered from the technologies available nowadays defines the first good practice for our discipline: the acquisition as a standard procedure, of high-definition digital copies of thin sections, with a dpi resolution above 4000 dpi.

The fruits of these high-resolution scans translate not only into having better illustrations for the individual researcher but also the value in their presentation to others. The clear advantages of publishing dynamic high-resolution thin sections are the possibility to read them at different zoom levels, to contextualise sedimentary components, microfacies, descriptions, and microphotographs, and in the end, to pass the bridge from the macroscopic archaeological record to the microarchaeological evidence.

The combination of gigapixel thin section images and article enrichments define a step forward in communicating research results and sharing raw data, an opportunity that should not be missed.

Although this proposal speaks to micromorphologists, the approach presented here demonstrates the potential and applicability to other disciplines that use high definition images of thin sections as a primary source of information, such as mineralogy or petrology. We thus encourage other authors to take advantage of the developments acquiring high-resolution images and sharing their own raw research data.

References

- Albert, R.M., Ruíz, J.A., Sans, A., 2016. PhytCore ODB: a new tool to improve efficiency in the management and exchange of information on phytoliths. *J. Archaeol. Sci.* 68, 98–105. <http://dx.doi.org/10.1016/j.jas.2015.10.014>.
- Aldeias, V., Bicho, N., 2016. Embedded behavior: human activities and the construction of the Mesolithic Shellmound of Cabeço da Amoreira, Muge, Portugal. *Geoarchaeology* 31, 530–549. <http://dx.doi.org/10.1002/gea.21573>.
- Aldeias, V., Goldberg, P., Sandgathe, D., Berna, F., Dibble, H.L., Mcperron, S.P., Turq, A., Rezeq, Z., 2012. Evidence for Neandertal use of fire at Roc de Marsal (France). *J. Archaeol. Sci.* 39, 2414–2423. <http://dx.doi.org/10.1016/j.jas.2012.01.039>.
- Angelucci, D.E., Boschian, G., Fontanals, M., Pedrotti, A., Vergès, J.M., 2009. Shepherds and karst: the use of caves and rock-shelters in the Mediterranean region during the Neolithic. *World Archaeol.* 41, 191–214. <http://dx.doi.org/10.1080/00438240902843659>.
- Arpin, T.L., Mallol, C., Goldberg, P., 2002. Short contribution: a new method of analyzing and documenting micromorphological thin sections using flatbed scanners: applications in geoarchaeological studies. *Geoarchaeology* 17, 305–313. <http://dx.doi.org/10.1002/gea.10014>.
- Benavides López, J.A., Aranda Jiménez, G., Sánchez Romero, M., Alarcón García, E., Fernández Martín, S., Lozano Medina, A., Esquivel Guerrero, J.A., 2016. 3D modelling in archaeology: the application of structure from motion methods to the study of the megalithic necropolis of Panoria (Granada, Spain). *J. Archaeol. Sci. Rep.* 10, 495–506. <http://dx.doi.org/10.1016/j.jasrep.2016.11.022>.
- Bocanegra Barbecho, L., Gea Megías, M., 2016. Capítulo 8. Las Humanidades Digitales y el aprendizaje en acceso abierto: El caso de la comunidad en Historia Digital. In: *Experiencia Mooc: Un Enfoque Hacia El Aprendizaje Digital, La Creación de Contenidos Docentes Y Comunidades Online*. Universidad de Granada, Granada.
- Bocanegra Barbecho, L., Toscano, M., Delgado Anés, L., 2017. Co-creación, participación y redes sociales para hacer historia. *Ciencia con y para la sociedad. Historia y Comunicación Social* 22, 325–346. <http://dx.doi.org/10.5209/HICS.57847>.
- Burdick, A., Drucker, J., Lunenfeld, P., Presner, T., Schnapp, J., 2012. *Digital Humanities*. Massachusetts Institute of Technology Press, Cambridge, Massachusetts. London, England.
- Carpentier, F., Vandermeulen, B., 2016. High-resolution photography for soil micromorphology slide documentation. *Geoarchaeology* 31, 603–607. <https://doi.org/10.1002/gea.21563>.
- Cerrillo-Cuenca, E., 2017. An approach to the automatic surveying of prehistoric barrows through LiDAR. *Quat. Int.* 435, 135–145. <http://dx.doi.org/10.1016/j.quaint.2015.12.099>.

- Cerrillo-Cuenca, E., Sepúlveda, M., 2015. An assessment of methods for the digital enhancement of rock paintings: the rock art from the precordillera of Arica (Chile) as a case study. *J. Archaeol. Sci.* 55, 197–208. <http://dx.doi.org/10.1016/j.jas.2015.01.006>.
- Courty, M.-A., 2001. Microfacies Analysis Assisting Archaeological Stratigraphy. In: Goldberg, P., Holliday, V.T., Ferring, C.R. (Eds.), *Earth Sciences and Archaeology*. Springer US, Boston, MA, pp. 205–239.
- Courty, M.-A., Macphail, R.I., Goldberg, P., 1989. *Soils and Micromorphology in Archaeology*, Cambridge Manuals in Archaeology. Cambridge University Press, Cambridge.
- De Reu, J., De Smedt, P., Herremans, D., Van Meirvenne, M., Laloo, P., De Clercq, W., 2014. On introducing an image-based 3D reconstruction method in archaeological excavation practice. *J. Archaeol. Sci.* 41, 251–262. <http://dx.doi.org/10.1016/j.jas.2013.08.020>.
- Earley-Spadoni, T., 2017. Spatial history, deep mapping and digital storytelling: archaeology's future imagined through an engagement with the digital humanities. *J. Archaeol. Sci.* <http://dx.doi.org/10.1016/j.jas.2017.05.003>. (n.d.).
- Fisher, E.C., Akkaynak, D., Harris, J., Herries, A.I.R., Jacobs, Z., Karkanis, P., Marean, C.W., Mcgrath, J.R., 2015. Technical considerations and methodology for creating high-resolution, color-corrected, and georectified photomosaics of stratigraphic sections at archaeological sites. *J. Archaeol. Sci.* 57, 380–394. <http://dx.doi.org/10.1016/j.jas.2015.02.022>.
- Flügel, E., 2004. *Microfacies of Carbonate Rocks: Analysis, Interpretation and Application*. Springer.
- Goldberg, P., 1980. Micromorphology in archaeology and prehistory. *Paléorient* 6, 159–164.
- Goldberg, P., Aldeias, V., 2018. Why does (archaeological) micromorphology have such little traction in (geo)archaeology? *Archaeol. Anthropol. Sci.* 10 (2), 269–278.
- Goldberg, P., Berna, F., 2010. Micromorphology and context. *Quat. Int.* 214, 56–62. <http://dx.doi.org/10.1016/j.quaint.2009.10.023>.
- Goldberg, P., Macphail, R.I., 2006. *Practical and Theoretical Geoarchaeology*. Blackwell Publishing Ltd.
- Goldberg, P., Miller, C.E., Schiegl, S., Ligouis, B., Berna, F., Conard, N.J., Wadley, L., 2009. Bedding, hearths, and site maintenance in the middle stone age of Sibudu cave, KwaZulu-Natal, South Africa. *Archaeol. Anthropol. Sci.* 1, 95–122. <http://dx.doi.org/10.1007/s12520-009-0008-1>.
- Goldberg, P., Dibble, H., Berna, F., Sandgathe, D., Mcperron, S.J.P., Turq, A., 2012. New evidence on Neandertal use of fire: examples from Roc de Marsal and Pech de l'Azé IV. *Quat. Int.* 247, 325–340. <http://dx.doi.org/10.1016/j.quaint.2010.11.015>.

- Goldberg, P., Aldeias, V., Dibble, H., Mcpherron, S., Sandgathe, D., Turq, A., 2013. Testing the Roc de Marsal Neandertal “burial” with geoarchaeology. *Archaeol. Anthropol. Sci.* 1–11. <http://dx.doi.org/10.1007/s12520-013-0163-2>.
- Green, S., Bevan, A., Shapland, M., 2014. A comparative assessment of structure from motion methods for archaeological research. *J. Archaeol. Sci.* 46, 173–181. <http://dx.doi.org/10.1016/j.jas.2014.02.030>.
- Gutiérrez-Rodríguez, M., Goldberg, P., Brassous, L., Rodríguez-Gutiérrez, O., Martín Peinado, F.J., Orfila, M., 2018. Site formation processes and urban transformations during Late Antiquity from a high-resolution geoarchaeological perspective: Baelo Claudia (Southern Spain). *Geoarchaeology* (submitted).
- Haaland, M.M., Friesem, D.E., Miller, C.E., Henshilwood, C.S., 2017. Heat-induced alteration of glauconitic minerals in the Middle Stone Age levels of Blombos Cave, South Africa: implications for evaluating site structure and burning events. *J. Archaeol. Sci.* 86, 81–100. <http://dx.doi.org/10.1016/j.jas.2017.06.008>.
- Haaland, M.M., Czechowski, M., Carpentier, F., Lejay, M., 2018. Documenting archaeological thin sections in high-resolution: a comparison of methods and discussion of analytical applications. *Geoarchaeology* (forthcoming).
- Hendrickx, M., Gheyle, W., Bonne, J., Bourgeois, J., De Wulf, A., Goossens, R., 2011. The use of stereoscopic images taken from a microdrone for the documentation of heritage – an example from the Tuekta burial mounds in the Russian Altay. *J. Archaeol. Sci.* 38, 2968–2978. <http://dx.doi.org/10.1016/j.jas.2011.06.013>.
- Howey, M.C.L., Brouwer Burg, M., 2017. Assessing the state of archaeological GIS research: unbinding analyses of past landscapes. *J. Archaeol. Sci.* 84, 1–9. <http://dx.doi.org/10.1016/j.jas.2017.05.002>.
- Jorayev, G., Wehr, K., Benito-Calvo, A., Njau, J., de la Torre, I., 2016. Imaging and photogrammetry models of Olduvai Gorge (Tanzania) by unmanned aerial vehicles: a high-resolution digital database for research and conservation of early stone age sites. *J. Archaeol. Sci.* 75, 40–56. <http://dx.doi.org/10.1016/j.jas.2016.08.002>.
- Karkanas, P., Van de Moortel, A., 2014. Micromorphological analysis of sediments at the bronze age site of Mitrou, Central Greece: patterns of floor construction and maintenance. *J. Archaeol. Sci.* 43, 198–213. <http://dx.doi.org/10.1016/j.jas.2014.01.007>.
- Macphail, R.I., Goldberg, P., 2010. 26 - archaeological materials. In: Stoops, G., Marcelino, V., Mees, F. (Eds.), *Interpretation of Micromorphological Features of Soils and Regoliths*. Elsevier, Amsterdam, pp. 589–622.
- Macphail, R.I., Courty, M.-A., Goldberg, P., 1990. Soil micromorphology in archaeology. *Endeavour* 14, 163–171. [http://dx.doi.org/10.1016/0160-9327\(90\)90039-T](http://dx.doi.org/10.1016/0160-9327(90)90039-T).
- Matthews, W., French, C.A.I., Lawrence, T., Cutler, D.F., Jones, M.K., 1997. Microstratigraphic traces of site formation processes and human activities. *World Archaeol.* 29, 281–308. <http://dx.doi.org/10.1080/00438243.1997.9980378>.

- McCoy, M.D., 2017. Geospatial big data and archaeology: prospects and problems too great to ignore. *Journal of archaeological science, archaeological GIS today: persistent challenges, pushing old boundaries, and exploring*. *New Horiz.* 84, 74–94. <http://dx.doi.org/10.1016/j.jas.2017.06.003>.
- Mentzer, S.M., Quade, J., 2013. Compositional and isotopic analytical methods in archaeological micromorphology. *Geoarchaeology* 28, 87–97. <http://dx.doi.org/10.1002/gea.21425>.
- Miller, C.E., Conard, N.J., Goldberg, P., Berna, F., 2010. Dumping, sweeping and trampling: experimental micromorphological analysis of anthro-pogenically modified combustion features. In: Théry-Parisot, I., Chabal, L., Costamagno, S. (Eds.), *The Taphonomy of Burned Organic Residues and Combustion Features in Archaeological Contexts, Proceedings of the Round Table, May 27–29 2008*. *Palethnologie* pp. 25–37.
- Mozas-Calvache, A.T., Pérez-García, J.L., Cardenal-Escarcena, F.J., Mata-Castro, E., Delgado-García, J., 2012. Method for photogrammetric surveying of archaeological sites with light aerial platforms. *J. Archaeol. Sci.* 39, 521–530. <http://dx.doi.org/10.1016/j.jas.2011.10.007>.
- Prins, A.B., Adams, M.J., Homsher, R.S., Ashley, M., 2014. Digital archaeological fieldwork and the Jezreel valley regional project, Israel. *Near East. Archaeol.* 77, 192–197. <http://dx.doi.org/10.5615/neareastarch.77.3.0192>.
- Remondino, F., Barazzetti, L., Nex, F., Scaioni, M., Sarazzi, D., 2011. UAV photogrammetry for mapping and 3D modeling-current status and future perspectives. *Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci.* 38, 25–31.
- Roosevelt, C.H., Cobb, P., Moss, E., Olson, B.R., Unluşoy, S., 2015. Excavation is destruction digitization: advances in archaeological practice. *J. Field Archaeol.* 40, 325–346. <http://dx.doi.org/10.1179/2042458215Y.0000000004>.
- Sandgathe, D.M., Dibble, H.L., Goldberg, P., Mcperron, S., 2011. The Roc de Marsal Neandertal child: a reassessment of its status as a deliberate burial. *J. Hum. Evol.* 61, 243–253.
- Schreibman, S., Siemens, R., Unsworth, J., 2004. (Eds.). *A Companion to Digital Humanities*. Blackwell, Oxford.
- Sedov, S., Stoops, G., Shoba, S., 2010. 13 - regoliths and soils on volcanic ash. In: Stoops, G., Marcelino, V., Mees, F. (Eds.), *Interpretation of Micromorphological Features of Soils and Regoliths*. Elsevier, Amsterdam, pp. 275–303.
- Shillito, L.-M., Matthews, W., 2013. Geoarchaeological investigations of Midden-Formation Processes in the Early to Late Ceramic Neolithic levels at Çatalhöyük, Turkey ca. 8550–8370 cal BP. *Geoarchaeology* 28, 25–49. <http://dx.doi.org/10.1002/gea.21427>.
- Shillito, L.-M., Almond, M.J., Nicholson, J., Pantos, M., Matthews, W., 2009. Rapid characterisation of archaeological midden components using FT-IR spectroscopy, SEM-EDX and micro-XRD. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 73, 133–139. <http://dx.doi.org/10.1016/j.saa.2009.02.004>.

- Stoops, G., 2003. Guidelines for Analysis and Description of Soil and Regolith Thin Sections. Soil Science Society of America Inc., Madison, Wisconsin, USA.
- Stroncik, N.A., Schmincke, H.-U., 2002. Palagonite – a review. *Int. J. Earth Sci.* 91, 680–697. <http://dx.doi.org/10.1007/s00531-001-0238-7>.
- Tarquini, S., Favalli, M., 2010. A microscopic information system (MIS) for petrographic analysis. *Comput. Geosci.* 36, 665–674. <http://dx.doi.org/10.1016/j.cageo.2009.09.017>.
- Thorseth, I.H., Furnes, H., Tumyr, O., 1995. Textural and chemical effects of bacterial activity on basaltic glass: an experimental approach. *Chem. Geol.* 119, 139–160. [http://dx.doi.org/10.1016/0009-2541\(94\)00098-S](http://dx.doi.org/10.1016/0009-2541(94)00098-S).
- Toscano, M., 2018. Where the researcher cannot get. Open platforms to collaborate with citizens on cultural heritage research data. In: *Territorios Digitales. Construyendo unas Ciencias Sociales y Humanidades Digitales*. Universidad de Granada and Downhill Publishing, NY (ISBN 978-84-338-6180-1).
- Wattez, J., Courty, M.A., Macphail, R.I., 1990. Burnt organo-mineral deposits related to animal and human activities in prehistoric caves. In: Douglas, L.A. (Ed.), *Developments in Soil Science, Soil Micro-Morphology: A Basic and Applied Science*. Elsevier, pp. 431–439. [http://dx.doi.org/10.1016/S0166-2481\(08\)70358-1](http://dx.doi.org/10.1016/S0166-2481(08)70358-1).
- Weiner, S., Goldberg, P., Bar-Yosef, O., 1993. Bone preservation in Kebara cave, Israel using on-site Fourier transform infrared spectrometry. *J. Archaeol. Sci.* 20, 613–627. <http://dx.doi.org/10.1006/jasc.1993.1037>.
- Wilson, A.T., Edwards, B., 2016. *Open Source Archaeology, Ethics and Practice*. De Gruyter Open, Berlin. <http://dx.doi.org/10.1515/9783110440171>.
- Wouters, P., Beaulieu, A., Scharnhorst, A., Wyatt, S., 2012. *Virtual Knowledge. Experimenting in the Humanities and the Social Sciences*. Massachusetts Institute of Technology Press, Cambridge, Massachusetts. London, England.
- Wouters, B., Makarona, C., Nys, K., Claeys, P., 2016. Characterization of archaeological metal remains in micromorphological thin sections using μ XRF elemental mapping. *Geoarchaeology* 32, 311–318. <https://doi.org/10.1002/gea.21591>.

SUPPLEMENTARY MATERIAL

1. THE ONLINE PUBLICATION OF HIGH-RESOLUTION IMAGES

As widely explained in the paper, having online a high-resolution digital version of the thin section represents a powerful dissemination and analytical tool. In this supplementary material, we describe a DIY solution, available to the researcher to enrich articles published in journals that do not offer their own solution or as an independent possibility to create his/her own catalogue of zoomable images, easily shareable, customizable and integrable with other Web instruments. This is an efficient solution that involves some further image processing, downloading some code from the Internet, preparing an HTML document, and finally uploading a bunch of files and folders to a web-server. This procedure is a bit more technical, so we have prepared a step-by-step procedure along with a downloadable, fully working demo package.

2. TILED IMAGES PROCESSING

Dealing with the large files produced with high-resolution image scanning may not be an issue for storage, as large hard drives are nowadays very accessible, but it is definitely a problem to serve them on the Web. Fortunately, there is a well-established technique to solve this problem, tiled images, so for our purpose it is just a matter of choosing the appropriate software and the output format. The tiled images technique is widely used on the Web, especially to serve maps and satellite images¹, supporting the display of high-resolution images with a high level of performance. It basically consists in subdividing the original image into smaller squares, tiles, at different levels, each with half the resolution of the previous one, creating a pyramid of images; individual tiles are then rendered and displayed separately, making the process much faster.

There are several software and output formats available that produce similar results, so for simplicity we have chosen to present here just one combination: VIPS and DeepZoom. The DeepZoom format (DZI) is an XML specification developed by Microsoft as an Open Format, so many other applications support it and it is a common standard in this field. It basically consists of a folder containing many folders numbered subsequently, whose final number depends on the original image full resolution. Each internal folder defines a different zoom level and contains from one to thousands of small images, namely the tiles. Microvisioneer supports it out of the box or you can produce Deep Zoom pyramids from any image using the VIPS library².

3. CUSTOM WEB VIEWER

Once you have the folder of tiles and the associated .dzi file, you are ready to display them on the Web using an *ad-hoc* viewer. As often with technologies, there are several

options available, with very different features one to each other. They can be cloud-based, self-hosted, commercial, free or Open Source³. For the purpose of this didactic procedure, we chose to illustrate OpenSeadragon. It is the most used Open Source software in this field and a great choice for several reasons: it is free, it is backed by a strong and active community of developers, it is easily customizable and extendable, it can be self-hosted, so it doesn't lock your data into a proprietary solution, it supports several input formats, it is light and mobile compatible. The source code can be downloaded at <https://openseadragon.github.io/#download> and, as you will see in the following chapter, you don't need to be a developer or an IT specialist in order to use it.

4. STEP-BY-STEP TECHNICAL PROCEDURE

The complete procedure, from the thin section digital image up to the online accessible viewer, involves several steps that we are going to describe here, using as an example a thin section named bc-ex-3-2:

- Reduce the initial file size of each digital version of the thin section (PPL, XPL and annotated), using at least LZW compression for TIFF images, or JPEG format with quality set at least to 90;
- Install the VIPS library, for Windows⁴ there are pre-compiled binaries available;
- Download vips-dev-w64-web-x.y.z.zip from the Download area⁵ and unzip somewhere. Open the command line utility and move to the unzipped folder, inside the *bin* directory;
- At the command-prompt, run
`$ vips dzsaves bc-ex-3-2-ppl.tif bc-ex-3-2-ppl --suffix .jpg[Q=90]`
- This command will run for a while producing the tiles for the image "bc-ex-3-2-ppl.tif" in the folder "bc-ex-3-2-ppl" with JPEG quality set to 90;
- Repeat the same procedure for the other two images, changing the image name and the folder name;
- Group all produced files and folder into a folder named "collection" and move it into a new empty folder named as the thin section;
- Folder structure is essential so, following our example, double check that the DZI files are accessible at: bc-ex-3-2/collection/bc-ex-3-2-ppl.dzi and so on;
- Download the latest OpenSeaDragon version at <https://openseadragon.github.io/> and unzip it inside the "bc-ex-3-2" folder;
- Download the openseadragon-scalebar.js file from <https://github.com/usnistgov/OpenSeadragonScalebar> and place it as well inside the "bc-ex-3-2" folder;
- Create an empty HTML file outside the "bc-ex-3-2-ppl" folder, open it with a text editor and paste this code, replacing text in bold with the name of your thin section:
- Create an empty .html text file just outside that folder and paste in this code, replacing text in **bold**, that refers to sample name:

```

<!DOCTYPE html>
<html>
<head>
  <meta charset='utf-8'>
  <title>bc-ex-3-2 thin section</title>
<script src="bc-ex-3-2/openseadragon.min.js"></script>
<script src="bc-ex-3-2/openseadragon-scalebar.js"></script>
</head>
<body>
<div id="openseadragon1" style="width: 1024px; height: 768px;"></div>
<script type="text/javascript">
var viewer = new OpenSeadragon.Viewer({
id: "openseadragon1",
prefixUrl: "bc-ex-3-2/images/",
sequenceMode: true,
showReferenceStrip: true,
tileSources: ["bc-ex-3-2/collection/bc-ex-3-2-pp1.dzi",
  "bc-ex-3-2/collection/bc-ex-3-2-xpl.dzi",
  "bc-ex-3-2/collection/bc-ex-3-2-annotation.dzi"]
});
viewer.scalebar({
  type: OpenSeadragon.ScalebarType.MICROSCOPY,
  minWidth: "100px",
  pixelsPerMeter: 856580,
  color: "rgb(150, 150, 150)",
  location: OpenSeadragon.ScalebarLocation.TOP_RIGHT
});
</script>
</body>
</html>

```

- Calculate the pixelsPerMeter variable with the following formula: (longest side in pixels/longest side in cm) * 100.

A complete demo package (2.2GB) - images and code - can be downloaded at <http://dx.doi.org/10.17632/2xx36kdvmr.1> (Toscano *et al.* 2021).

The resulting collection of files and folders can then be uploaded to a Web server, linked within the publication or simply shared as any Web resource.

5. NOTES

¹ We all use it frequently with online mapping services, like Google Maps.

² <http://jcupitt.github.io/libvips/API/current/Making-image-pyramids.md.html>.

³ <https://www.easyczoom.com/>, <http://www.zoomify.com/>, <https://viewer.gigamacro.com/>, <https://proscia.com/>, <http://porescale.co.uk/>, https://openseadragon.github.io, <http://www.precipoint.com/microscopy-software/viewpoint/>, <https://qupath.github.io/>, <https://pathcore.com/sedeen/>, etc.

⁴ For unix-like operating systems check installation instructions at <http://jcupitt.github.io/libvips/install.html>. If you have issue using the VIPS library, DZI tiles can be obtained also with the Windows only Photoshop plug-in HD View Utilities, downloadable at <https://www.microsoft.com/en-us/download/details.aspx?id=52341>.

⁵ <https://github.com/jcupitt/libvips/releases>.

IV.4. LOS SISTEMAS DE INFORMACIÓN WEB PARA EL CROWDSOURCING Y LA CIENCIA CIUDADANA

Una vez analizados algunos de los beneficios del uso de las TIC en humanidades, ya sea en relación con la gestión informatizada de datos estructurados que con su comunicación, en este apartado movemos el enfoque hacia el tercer componente del esquema conceptual mostrado en la figura 1, la sociedad, y ampliamos el marco metodológico de referencia hacia la Ciencia ciudadana (Eitzel *et al.* 2017) y la Historia Pública Digital (Noiret 2013; 2015a; 2015b; 2018).

El análisis de las implicaciones y beneficios de la involucración de la ciudadanía, a través de los sistemas de información web, en el proceso de investigación en disciplinas humanísticas, se ha reflejado en dos publicaciones. La primera estudia el tema de la abertura al exterior de conjuntos de datos existentes; no tanto, o no solamente, como datos en abierto (*Open Data*), si no que en relación con la creación de plataformas digitales orientadas a la colaboración, en el marco de una metodología de investigación en abierto (*Open research methodology*). En concreto, se han analizado algunos de los factores que justifican la creación de sistemas de información web y se han descrito los principios a tener en cuenta en el proceso. En la segunda publicación se han abarcado los fenómenos del *crowdsourcing*, de la co-creación con la ciudadanía y del trabajo en red, aplicados a contextos de investigación en Historia Pública.

La posibilidad de contribuir activamente al proceso de investigación acerca la ciencia a la sociedad civil y abre las puertas a nuevos métodos de colaboración y verificación de las fuentes, ayudando a alinear los resultados en I+D+i con las necesidades y expectativas de los ciudadanos, con las exigencias del programa SwafS¹²⁹ de la Comisión Europea y con los principios de la Investigación e Innovación Responsable¹³⁰ (RRI por su acrónimo inglés), que miran a enriquecer el proceso y los resultados de la investigación e innovación con los valores, necesidades y expectativas de la sociedad europea (Bocanegra Barbecho 2020). En concreto, como muestra de este creciente interés por la Ciencia ciudadana, podemos mencionar que desde el inicio del programa H2020 en el año 2014, el programa SwafS ha financiado 205 proyectos¹³¹: un aumento considerable con respecto a la participación en el anterior programa marco FP7, donde se financiaron solo 27 iniciativas de participación pública. En el marco del nuevo programa marco

¹²⁹ *Science with and for Society*, <https://ec.europa.eu/programmes/horizon2020/en/h2020-section/science-and-society>.

¹³⁰ *Responsible Research and Innovation* <https://ec.europa.eu/programmes/horizon2020/en/h2020-section/responsible-research-innovation>, https://en.wikipedia.org/wiki/Responsible_Research_and_Innovation y *Public Engagement and responsible research and innovation* <https://ec.europa.eu/programmes/horizon2020/node/766>.

¹³¹ Información obtenida el 27/12/2020 a través del portal CORDIS, combinando los parámetros de búsqueda 'Projects' y 'Science with and for society': <https://cordis.europa.eu/search?q=contenttype%3D%27project%27%20AND%20programme%2Fcode%3D%27H2020-EU.5.%27&p=1&num=10&srt=Relevance:decreasing>.

*Horizon Europe*¹³² (2021-2027) esos programas, lejos de haber agotado su misión, se mantendrán activos dentro del reto *Widening participation and spreading excellence*.

Este apartado se concluye con la presentación de un modelo teórico para la clasificación y el análisis de estrategias de involucración pública (*participatory engagement strategies*), orientadas a fomentar la participación en proyectos colaborativos, en particular aquellos generados desde el campo de las humanidades y del patrimonio cultural.

¹³² https://ec.europa.eu/info/horizon-europe_en.

IV.4.1. Desde datos locales a plataformas en línea

En esta sección se presentan los resultados de un análisis sobre los beneficios que el uso de sistemas de información web conlleva para el enriquecimiento y completamiento, a través de la colaboración ciudadana o con otros grupos de investigación, de conjuntos de datos ya existentes, precedentemente almacenados localmente. El análisis se realiza a partir de tres casos de estudio procedentes de ámbitos disciplinares y cronológicos muy diferentes: el patrimonio artístico Andaluz en Iberoamérica; los sistemas de riego tradicional de origen medieval en el suroeste de España¹³³; la configuración de las viviendas rurales del siglo XVIII en el Valle de Lecrín (Granada), a través de la documentación de archivo.

El principal resultado de esta publicación son una serie de recomendaciones y buenas prácticas sobre el diseño, la implementación y el mantenimiento de plataformas abiertas a la colaboración externa, así como sobre estrategias para fomentar la participación de la ciudadanía en el proceso de recogida, digitalización y explotación de fuentes históricas. Una ulterior reflexión sobre la temática de la participación pública en proyectos sobre patrimonio cultural, a partir de un contexto mucho más amplio de 110 casos de estudio seleccionados a nivel internacional (Toscano *et al.* 2019), ha sido realizada por el autor de esta tesis a través de un informe científico-técnico (Toscano y Colella 2019), publicado a principio de 2019 en el marco del proyecto internacional REACH-*RE-designing Access to Cultural Heritage for a wider participation in preservation, (re-)use and management of European culture*¹³⁴.

La publicación de este capítulo de libro ha encontrado cabida en el marco de una iniciativa editorial promovida por el MedialabUGR en colaboración con la EditorialUGR, con el objetivo de reflexionar acerca de cómo “las Humanidades y las Ciencias Sociales han empleado los nuevos instrumentos que lo computacional y lo conectado ponen a su servicio para intentar resolver las preguntas de siempre de otra manera y para formular nuevas cuestiones que antes no tenían siquiera sentido”, en palabras de los propios curadores de la monografía (Romero Frías y Bocanegra Barbecho 2018).

En el contexto de esta investigación, el candidato ha colaborado con los investigadores principales de los proyectos utilizados como casos de estudios en la fase de conceptualización del modelo de datos y de definición de las funcionalidades de las plataformas. Por otro lado, ha llevado a cabo en autonomía las siguientes tareas: elaboración del núcleo conceptual de la contribución, búsqueda bibliográfica, elaboración del

¹³³ Este sistema de información web, actualmente accesible en <https://regadiohistorico.es> y desarrollado en primera instancia para el mapeo colaborativo, está siendo sometido, en 2021, a una fase de mantenimiento perfectivo (ver apartado III.2), que prevé la ampliación de sus funcionalidades para habilitar el trabajo telemático en grupo entre estudiantes, en el marco del proyecto de investigación “Aprender a ser científic@s con el regadío histórico” Ref. FCT-19-15217.

¹³⁴ H2020- CULT-COOP-06-2017 - Participatory approaches and social innovation in culture. Grant Agreement No. 769827, https://cordis.europa.eu/project/rcn/212215_es.html.

material gráfico, implementación informática de las plataformas, redacción del borrador inicial del texto, revisión formal, correspondencia con el comité editorial de la monografía e implementación de las sugerencias de los revisores.

Referencia bibliográfica

Toscano, M. (2018). Where the researcher cannot get: open platforms to collaborate with citizens on cultural heritage research data. En E. Romero Frías y L. Bocanegra Barbecho (Eds.), *Ciencias Sociales y Humanidades Digitales Aplicadas* (pp. 538-561). Granada, España: Universidad de Granada [ISBN: 978-84-338-6318-8]. <https://doi.org/10.5281/zenodo.1469337>

Título original

Where the researcher cannot get: open platforms to collaborate with citizens on cultural heritage research data

Título en español

Donde el investigador no puede llegar: plataformas abiertas para colaborar con la sociedad sobre datos de investigación en patrimonio cultural

Abstract

Over the years, many research projects and research groups have accumulated large collections of raw data, nowadays they are normally compiled and stored in local archives, more often on personal computers and sometimes on Intranet servers. Local data represents a missed opportunity for research projects, in particular in the cultural heritage field. The vast majority of these datasets, even those that are already available in a digital format, remain hidden, underutilized and incomplete because they are linked to the effort and limited resources of a single researcher or a specific research group. This paper proposes a shift from Local Data to Open Platforms, from data stored locally and accessible only to a restricted number of researchers to remotely accessible, editable and shareable datasets. Online platforms perfectly fit the needs of a research group that is experiencing the limits of first-hand data collection and verification or that intends to extend the potential area of source collection. To demonstrate the effectiveness of this proposal, three different case studies will be explored. They have all been recently developed as the result of collaborations between the author and university research groups in the humanities. Despite sharing a common background in cultural heritage, they are very different with respect to the object of study. The core of this paper is the discussion, where we analyse the lessons learnt from applying this approach, highlighting its benefits and the intrinsic issues of the process.

Resumen

Muchos proyectos y grupos de investigación han llegado a acumular a lo largo de los años grandes cantidades de datos en bruto, por regla general acumulados y almacenados localmente, a veces en servidores locales o, más a menudo, en ordenadores personales. Los datos locales representan una oportunidad perdida para los proyectos de investigación, en particular en el campo del patrimonio cultural. La gran mayoría de estos conjuntos de datos, incluso aquellos que ya están en formato digital, permanecen ocultos, subutilizados e incompletos porque están vinculados al esfuerzo y los

recursos limitados de un solo investigador o grupo de investigación específico. Esta contribución propone de pasar desde Datos Locales a Plataformas Abiertas, desde datos accesibles solo a un número restringido de investigadores a conjuntos de datos accesibles, editables y compartibles a distancia. Las plataformas en línea se ajustan perfectamente a las necesidades de un grupo de investigación que experimenta los límites de la recopilación y verificación directa de los datos o tiene la intención de ampliar el área potencial de la recopilación de fuentes. Para concretar esta propuesta, se describirán tres diferentes casos de estudio. Todos han sido desarrollados recientemente y son el resultado de la colaboración del autor con diferentes grupos de investigación en humanidades. Incluso si todos están relacionados con el patrimonio cultural, son muy diferentes con respecto al objeto de estudio. El núcleo principal de este documento reside en la discusión, donde se analizan las lecciones aprendidas al aplicar este enfoque, destacando sus beneficios y los problemas intrínsecos en el proceso.

Keywords

Web Information Systems, collaborative platforms, co-creation, digital humanities, Science with and for Society.

Palabras claves

Sistemas de Información Web, plataformas colaborativas, co-creación, humanidades digitales, Ciencia con y para la Sociedad.

1. Background

This paper is the result of the author's professional experience, who over the past few years worked with several researchers and research groups and, as a consequence of these collaborations, built several online collaborative platforms, created to collect unpublished research data or collaborate with other research groups or the civil society over large collections of research data. The author has a background in humanities and cultural heritage, so the case studies described in this paper have been derived from these fields, as well as the proposal and the discussion apply better to projects in human sciences, even if not exclusively. This paper is part of the work that the author is developing within the H2020 European Project REACH - RE-designing Access to Cultural Heritage for a wider participation in preservation, (re-)use and management of European culture¹.

2. Why does local data represent a missed opportunity?

Many research projects and research groups have compiled large collections of raw data over the years, nowadays predominantly in the form of electronic files or databases. Those archives are normally accumulated and stored in local archives, more often on personal computers and sometimes on Intranet servers. This generates several problems, unknown to analogue records, in the form data integrity, backups, and long-term consistency amongst other things. The aim of this paper is not to discuss

these issues, but instead to focus on the missed opportunities that local data represents for research projects, in particular in the field of cultural heritage.

The vast majority of these datasets, even those that are already available in a digital format, remain hidden, underused and incomplete because they are linked to the effort and limited resources of a single researcher or a specific research group.

The publication of raw research data online - Open Data - is becoming increasingly popular², however somewhat slowly within the humanities, and is already a requirement of many European and national calls. This is a very positive initiative, focused on the use, re-use and verification of research data, but it only serves as a partial solution to the local data issue described above, as it does not contribute directly to the improvement, verification and enlargement of those corpora. Data is shared raw, without the necessary tools to allow direct interaction with the general public or other research groups.

3. Why do Open platforms represent a valid instrument of research?

Over the last few decades, Digital Humanities have taken advantage of the widespread increase in the use of the Internet and have developed tools that make it possible to collaborate electronically with citizens (Fresa and Justrell, 2015) and other research groups. In order to deal with the above-mentioned issue, a possible approach is to move from Local Data to Open Platforms, from data stored locally and accessible only to a restricted number of researchers into content available online, where the datasets are accessible, editable and shareable.

Online platforms perfectly fit the needs of a research group that is experiencing the limits of first-hand data collection and verification or that intends to extend the potential area of source collection. They provide several advantages over other instruments of data management. They are collaborative for their own nature, allowing concurrent users working simultaneously, tracking individual activity and allowing version control and restore. When required, they can be managed and customized remotely. They are compatible with widely accepted and well-maintained Web standards, to easily create attractive and easy to use graphic interfaces. In this sense, they belong to an environment, the Web, that the vast majority of potential contributors use on a daily basis for work and during their spare time. The interfaces have a familiar appearance and therefore a complicated manual is not required when interacting with them also to non-specialists. More significantly, they offer built-in instruments to work dynamically on a dataset, either inputting new data, commenting or modifying existing data and uploading any type of archive.

According to the needs of the research project, the platform can be configured to accept real-time changes or through required moderation and approval. We will return to this subject in the discussion, but for now, we can say that allowing real-time changes

produces better engagement since by seeing their input published straightaway, contributors are encouraged to keep contributing; on the other hand, incorporating changes only after administrators' approval improves data quality control.

To effectively involve citizens in scientific research, online live collaboration is a key feature. Theoretically, a standalone PDF form can be sent individually or shared in large social network groups to collect new data, but this practice makes it harder to keep track of new contributions, requires extra management work to consolidate new and existing records and can affect data integrity and version control; additionally, online collaboration offers other benefits, such as creating a sense of community among contributors and foster data sharing. Finally, and most significantly, these collaborative tools can be custom made: not all platforms are made equal, in the same way that not all research projects are made equal. They can be tailored to fit the dataset that is going to be published and to create a separate, simple custom webform for data collection.

When structuring the back-end database, it is crucial to keep the aim of the research project in mind. A clear distinction should be made and different steps should be taken if producing a collection of data is the ultimate goal of the project as opposed to just a necessary step towards analysis and interpretation. In this last case, which is quite common, data collection standards³ can be taken in limited consideration, and the primary focus should be the use of the information within the research context.

4. Case studies

To consolidate this proposal to shift from local data to open platforms and briefly highlight the benefits of this approach and the intrinsic issues of the process, we are going to present three different case studies. They are all recently developed and are the result of the collaboration of the author with different research groups within the humanities. Even if they share a common background in cultural heritage, they are very different with respect to the object of study, one being focused on the shared artistic heritage among Andalucía and Latin America⁴, another on traditional medieval irrigation systems in the provinces of Granada and Almería⁵ and the last on eighteenth century houses in rural villages in the Lecrín Valley (Granada)⁶. For the purpose of this paper, what is important is that they are similar from a technical and methodological point of view, sharing the same issue and choosing a similar path in order to overcome it.

4.1. IBERO-AMERICAN ARTISTIC HERITAGE

For the platform *Colabora con Andalucía y América*⁷ (Toscano, 2018), the local data was represented by a large catalogue of artworks created by Andalusian artists throughout twelve Latin American countries, dating from the Renaissance to the present. Approximately 2,600 records have been collected since 2008, with detailed information about

the artist, the style, the iconography, the chronology and the localization, with the latter organised from country down to the building level. The awareness of the breadth of the subject and of the effort required to compile the files filled to date, has worked as an incentive to launch this initiative. The dataset has been published as individual records, presented in a catalogue view, with an advanced faceted search engine (figure 1). Registered users have two options to collaborate: improving existing records or report a new artwork. Many catalogue records do not have illustrations or they are not high quality, because it has not always been possible to directly access the artworks to take pictures; in those cases, the online platform is an opportunity to get much better illustrations taken directly onsite. On the other hand, the form is very simple and focuses on the most valuable data for the researcher, i.e. a photo of the artwork and its location, even if the possibility of providing more detailed information remains, such as the author, the typology and a short description. In order to facilitate filling in the form directly in front of the artwork, kept in a museum, a gallery, a church, a private collection or in a public space, full compatibility with mobile platforms was ensured, as well as the ability to directly upload a picture from a smartphone or tablet. Finally, to speed up possible contributions, registration has been simplified through the availability of login via social networks (Facebook and Google), basically omitting the user verification step via email. The aim was to create a means to broadening and improving the catalogue of artworks at a global level. Taking advantage of the immense potential of the internet, it is now possible to gather contributions directly from the study areas or from other research groups that are working on related topics.

Catalogue

Search artwork

Initial date

Use just the year (i.e. 1500)

Final date

Location

- (-) Bolivia
- Potosí (19)
- Chuquisaca (10)
- La Paz (8)
- Cochabamba (1)
- Oruro (1)

Style

- (-) Barroco

Typology

- Pinturas de caballete (24)
- Esculturas de bulto redondo (12)
- Pinturas (2)
- Pinturas sobre tabla (1)

Author

- Anónimo (17)
- Cueva, Gaspar de la (8)
- Anónimo potosino (4)
- Pérez Holguín, Melchor (2)
- Anónimo popular (1)
- Anónimo sevillano (1)
- Berrio, Gaspar Miguel de (1)
- Cordoba, Ignacio (1)
- Herrera y Velarde, Francisco de (1)
- Martínez Montañés, Juan (1)
- Pérez Holguín, Melchor (1)
- Zamudio, Bernabé de (1)

Iconography

- San Juan de Dios (15)
- San Francisco Solano (4)
- Divina Pastora (3)
- Jesús atado a la columna (3)
- San Fernando (3)
- Crucifixión de Jesús (2)
- Virgen de la Antigua (2)
- Crucificado (1)
- Dios padre (1)
- Inmaculada Concepción (1)
- San Juan Bautista (1)
- Virgen de la Bella (1)

Search found 39 items

Virgen de la Bella

Author: Anónimo
Date: 1701 to 1800
Typology: Esculturas de bulto redondo
Style: Barroco
Location: Iglesia de San Bartolomé Arani Cochabamba Bolivia
Elaboración: P07-HUM-03089



San Juan de Dios en contemplación

Author: Pérez Holguín, Melchor
Date: 1701 to 1800
Typology: Pinturas de caballete
Style: Barroco
Location: Universidad Mayor, Real y Pontificia de San Francisco Xavier Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089



San Juan de Dios con el Niño Jesús al hombro

Author: Anónimo
Date: 1701 to 1800
Typology: Pinturas de caballete
Style: Barroco
Location: Museo Universitario de Charcas Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089



San Juan de Dios penitente

Author: Anónimo
Date: 1751 to 1800
Typology: Pinturas de caballete
Style: Barroco
Location: Universidad Mayor, Real y Pontificia de San Francisco Xavier Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089



San Juan de Dios con el Niño Jesús al hombro

Author: Anónimo
Date: 1701 to 1800
Typology: Pinturas de caballete
Style: Barroco
Location: Museo Universitario de Charcas Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089



San Juan Bautista

Author: Cueva, Gaspar de la
Date: 1633 to 1638
Typology: Esculturas de bulto redondo
Style: Barroco
Location: Iglesia de San Miguel Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089



San Francisco Solano

Author: Anónimo
Date: 1601 to 1700
Typology: Esculturas de bulto redondo
Style: Barroco
Location: Ermita de San Francisco Solano Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089



San Juan de Dios con el Niño Jesús

Author: Anónimo
Date: 1667 to 1701
Typology: Pinturas de caballete
Style: Barroco
Location: Museo Convento de la Recoleta Sucre Chuquisaca Bolivia
Elaboración: P07-HUM-03089



Figure 1: Colabora con Andalucía y América online platform. Catalogue of 2,600 artworks created by Andalusian artists in Latin America, with faceted search filters in the sidebar. Source: <https://colabora.andaluciayamerica.com/>.

4.2. HISTORICAL IRRIGATION COMMUNITIES IN SOUTHERN-SPAIN

In the project *Regadíos históricos*⁸, core local data has been obtained from a combination of fieldwork surveys and interviews with local irrigation communities. In the field, researchers recorded irrigation channels and cultivated plots with handheld GPS and a portable geodatabase, while also carrying out interviews with local farmers about traditional water management practices. Additionally, to enlarge and enrich the dataset, two other strategies have been carried out: a telephone campaign contacting all irrigation communities in the provinces of Granada and Almeria and a desk-based cartographic survey, based on the *Inventory and Characterization of the Irrigation of Andalusia. Mediterranean and Atlantic Districts* (Inventario, 2008), modified according to well-known historical transformations and more recent modernizations that affected those systems. The resulting dataset that is now available online covers more than 200,000 hectares of irrigated land and counts more than 24,000 km of irrigation ditches, managed by 550 historical irrigation communities, for a total of 1,424 individual records in the database. Considering the cartographic nature of these data, the most appropriate way to present them to the public was using a spatial data infrastructure, more commonly known as a map (figure 2).



Figure 2: Regadíos históricos online platform. The collaborative map showing the example of the irrigation community of Lanteira (Granada), its cultivated area and main irrigation channel (*acequia madre*). Source: <http://regadiohistorico.es>.

Crear Ficha de riego

Nombre de la comunidad de Regantes *

Provincia *

Municipio *

Nombre de la acequia principal

Número de regantes

LOCALIZA EL MUNICIPIO PARA EMPEZAR A DIBUJAR



Utiliza los iconos arriba a la derecha según lo que quieras dibujar: área de riego o acequia. Usa el ratón para desplazarte y la barra de zoom para acercarte.

Hectáreas de riego

Riego eventual

Archivos adjuntos

Añadir archivo nuevo

No file selected.

Los archivos deben ser menores que 25 MB.
Tipos de archivo permitidos: jpg png tif tiff pdf txt doc docx xls xlsx.

Contacto comunidad

Persona de referencia (presidente o secretario)

Teléfono

Información adicional

LEER TÉRMINOS Y CONDICIONES

Figure 3: *Regadíos históricos* online platform. Webform to send new data, with fields to collect structured information, send archives and to draw cultivated plots and irrigation channels on a dynamic map. Source: <http://regadiohistorico.es>.

This product represents an important result by itself, but in order to create a true collaborative map, registered users now have the option to include new records and to comment on existing ones. To facilitate collaboration, a dynamic online form has been setup, divided into two sections (figure 3): a series of text fields to describe the irrigation community, with an additional field to collect files (documents, images,

spreadsheets etc.), and an editable map, where the user is invited to draw on the map the cultivated plots, as polygons, the path of the irrigation channels, as polylines, along with springs and water sources, as points. A short introductory video illustrates the whole process. The concept is similar to many other online collaborative projects that collect user generated geo-data, of which the most widely known is OpenStreetMap⁹, a worldwide collaborative effort to create a free editable map of the world. Registration has been made very easy by combining all the stages into a one step online web form process. The system requests the user email, automatically generates an account, associates the submitted information with it, logs the user in, and sends the username and password via email. In this way, the platform dramatically decreases the initial barrier of user registration and verification, simplifying the whole process of contributing new data.

4.3. EIGHTEENTH CENTURY RURAL HOUSES

In the case of *Las casas del Valle del Lecrín (S. XVIII)*¹⁰, local data derived from a long project carried out independently by the historian Margarita M. Birriel Salcedo, on the sections of the Cadastre of the Ensenada Marquis kept in the historical provincial archive of Granada. The research is focused on rural houses, the information is transcribed directly from the XVIII century manuscripts and registered, until then, using handwritten notes on notebooks organised with colour coded stickers. In this third case study, original data was not only local, but also stored in analogue format, presenting an additional exciting challenge that unfortunately we cannot expand on here, as it is outside the scope of this paper. The database structure has been based on two content types, houses and homeowners, connected by a many-to-many relationship. Recorded data basically concerns the houses, their dependencies, dimensions, locations, neighbouring, owner(s) and census; for the homeowners, their neighbourhood and properties. Primarily, the platform has been designed as an individual research instrument, to improve the recording process and above all the historical analysis, offering search capabilities, structured information and automatic quantifications. In addition, having used an online database from the beginning, instead of a local IT solution, has drastically simplified the process of opening the dataset to the public. Now, anonymous users can browse all published data, which is divided into two main views with several combinable search filters (figure 4), and four galleries, with sketch plans from the cadastre, hypothetical reconstructions of the houses, photos of the villages, houses, annexes and architectural elements. Registration is required to comment on existing records and upload files to the platform: the two means of collaborating available at the moment. Furthermore, to maintain the original objective of the platform - a working tool for the researcher to store and retrieve research data - we have introduced a "Save as draft" option for ongoing records that should not yet be visible to standard users.

Las casas del Valle del Lecrín (S. XVIII)
Catastro del Marqués de la Ensenada

Búsqueda propietarios Búsqueda casas Glosario propietarios Glosario casas Galerias Contacta Acceder

Búsqueda avanzada casas

Tu búsqueda ha encontrado 502 resultados

Tipología 1ª dependencia 2ª dependencia 3ª dependencia 4ª dependencia
- Cualquier - - Cualquier - - Cualquier - - Cualquier - - Cualquier -

Superficie (en metros) Lugar Arrendada Censo Fecha documento
- Cualquier - - Cualquier - - Cualquier - - Cualquier -

Aplica

Reiniciar

Insertar valores min y max

| Casa | Lugar | Barrio | Propietarios | Dependencias | Superficie | Arrendada | Alquiler | Catastro | Total | Censo | Fecha documento |
|--|----------------|----------------|-----------------------------|--------------------------------------|------------|-----------|----------|-------------------|------------------------|-----------------------------|-----------------|
| Casa en Las Albuñuelas Barrio alto | Las Albuñuelas | Barrio alto | 1. PALMA NAVARRO Antonio de | 1. o Planta baja 2. o Planta alta | 44.02 m. | No | | 30 reales y 0 mrs | 150 reales y 33.00 mrs | población, otros | 1752 |
| casa de campo en Las Albuñuelas Pago del Jaral | Las Albuñuelas | Pago del Jaral | 1. JIMÉNEZ Manuel | 1. o Planta baja | 41.92 m. | No | | | 730 reales y 8,00 mrs | habicones, población, otros | 1752 |
| Corral en Las Albuñuelas Barrio alto | Las Albuñuelas | Barrio alto | 1. JIMÉNEZ Manuel | | 44.02 m. | No | | 4 reales y 0 mrs | 730 reales y 8,00 mrs | habicones, población, otros | 1752 |

Figure 4: Las casas del Valle del Lecrín (S. XVIII) online platform. Table of recorded houses, with combinable filters at the top. Source: <http://casaslecrin.es/>.

5. Discussion

Once the issue has been presented, a possible solution proposed and illustrated using several ongoing initiatives, we are now able to discuss some of the lessons learnt from applying this approach.

Considering that all collaboration will be voluntary, it is wise to offer something in exchange. Potential contributors are likely to be people interested in the subject, so the availability of good quality, well-structured open research data will definitely be a valuable asset for the platform visitors and will encourage their engagement. The existing dataset will generate ideas on where and how it can be improved as well as display what kind of new data is useful to search for: users are more likely to contribute if they can preview where their contributions will end up. Additionally, once published, the data will collaborate to increase the visibility of the platform in organic search engine results.

Users behaviour on the platform is unpredictable, so carefully plan the permissions given to registered users, simplify the workflow process to contribute and offer key instructions and help texts, in short chunks written in a font that is easily readable.

When registration is required, as we recommend, sign-up and log-in steps should be as straightforward as possible, to reduce the initial barrier and increase participation¹¹. One possible strategy is the use of social login (i.e. Facebook, Twitter, Google, LinkedIn), basically skipping the user verification step via email. This approach has

several advantages: it reduces spam users, as it requires a valid social media profile, and it can automatically provide additional user information, including a profile photo, place of residence, age, and so on. The downsides are that, since the user verification process is outsourced, additional work may be required to ensure it remains compatible with future changes in the backend API, and if the users later stop using the social network account used to register, they will no longer have access to the website. Additionally, if the project plans to work with schools, directly involving young students in data collection, the use of social login can cause concerns regarding privacy and should therefore be used with caution. Social login has clear value for marketing websites (Goings and Abel, 2013), but for the purposes of an open platform of research data, a better solution comes by integrating the registration process within the form for submitting new data. This way, the web form can be presented directly to the end user, who can begin completing it straightaway, with all the registration and verification steps completed upon submission. In this scenario, in order to keep a trustworthy community of users and receive good quality contributions, a mechanism should be put in place to control and prevent spam, as spamming is inevitable for all public platforms that acquire a certain amount of visibility.

Ownership of contributions and Intellectual Property Rights (IPR) issues should be addressed outlining the terms and conditions of participation in the project. Basic principles are that authorship of individual contributions should be acknowledged and recorded, and the distribution and reuse of the research data collected should always be guaranteed. Some citizen science projects are more likely to provoke IPR and copyright issues, and those should not be underestimated.

As demonstrated for major collaborative websites, including Wikipedia, Yahoo! Groups, YouTube, etc., also online open platforms for research are affected by the 90-9-1 principle, or 1% rule (Nielsen, 2006). It basically suggests that 90% of users are just browsers who never contribute, they just read, 9% of users contribute a little, and 1% of users account for almost all the activity. The ratios might not be accurate, but the general structure still applies and we were able to confirm this behaviour first-hand within the Historic Graves project¹² (Bocanegra Barbecho, Toscano and Delgado Anés, 2017: 329), a collaborative effort to geotag and transcribe all Irish memorials. Keeping this benchmark in mind is useful in order to manage expectations, to estimate the number of passive users needed in order to start receiving a certain amount of contributions, and to implement special features that facilitate multiple contributions by the most engaged users (the top 1%).

Although the platform can function independently, being a readily available resource on the internet, a good practice is to plan and carry out *ad-hoc* campaigns, local events and training days to advertise the initiative and encourage contributions through engagement with the public. Instead of trying to involve individuals on a personal basis, we suggest identifying and engaging with existing communities of citizens or

reference collaborators, linked in some way to the research theme, who can adopt the role of local mediators. The use of online and traditional media to publish press releases about the initiative is also very effective, especially to create peaks of audience, from which long-term users can be acquired. These outreach activities are extremely important, especially at the beginning, as then search engine ranking and word-of-mouth can compensate for the lack of targeted distribution.

In this sense, human sciences research subjects are very well suited to attract the interest of civil society. Local communities sometimes are privileged potential contributors of the initiative, and the research topic often relates to a common cultural heritage, that is sometimes in danger of disappearing. Such is the case, for example, of the project about traditional irrigation channels in Andalucía, where the local communities featured in the dataset are a living research subject, strongly linked with the contemporary process of transformation and modernization in agriculture; or the above-mentioned Irish project where rural communities take care of their own local graveyard, geolocating and transcribing individual headstones (Toscano and Tierney, 2016); or, again, the case of Spanish associations in Argentina, that have had a strong role in the transnational online collaborative project *e-xiliad@s*¹³, focused on retrieving unpublished historical sources about the Spanish republican exile (Bocanegra Barbecho and Toscano, 2015). In cases like these, creating an open collaborative platform is also a way to raise awareness among citizens about the historical and cultural values related to that heritage, and share those values with communities and individuals living further away. On the other hand, establishing collaborations with other research groups can be more challenging, as that is not as common in the humanities as among researchers in natural science; the use of such platforms can stimulate building cooperation and data sharing, encouraging also common research activities.

People who will register and contribute to these initiatives can be greatly varied, from researchers specialised in the object of study, to passionate people related in some way to those places, objects or events, down to ordinary citizens, that come across the initiative and are willing to contribute by a spirit of networking, co-creation and participation. Wherever users come from, the platform should protect data integrity and security, being as open as possible yet as secure as necessary. Anonymous users should be able to browse all content; registered contributors should be allowed to publish content in real-time, comment on existing records or submit new data only through moderation and approval. The latter approach is more common in research, with the intention of keeping a tighter control over data published online, but a warning should be made that it can affect user engagement and can reduce the number of contributions. In every case, it is recommended to implement a versioning system, to keep track of all changes, their authors and timestamps. Being able to restore previous versions of a record and compare the changes between different versions is a much more powerful feature than a simple backup mechanism, which should be in place regardless as a means of protection in order to cover full restores in the event of server outages, data

corruption or security breaches. Finally, an automatic notification system should be put in place as well, to generate communications among the three subjects involved: the user, the platform and the research team. New users, new data, new comments, and new versions of existing records, the availability of security patches: every important event should be notified so that, if needed, appropriate actions can be taken. In our opinion, real-time publishing and editing, combined with a strong versioning mechanism and a solid notification system appears to be the best combination for a collaborative platform, to find a balance between a seamless and engaging experience to access and contribute to the dataset and the control over the platform security and data integrity in the medium-long term.

If the participatory project succeeds, a part-time platform administrator, ideally originating from the research team itself, should also be taken into account, to keep the system running, verify data sent, check for spam and notifications and support users when needed. The commitment of the research group in terms of use and maintenance of the platform in the medium term is important to build trust amongst users, but is also legitimate, at some stage, to consider closing or significantly modifying the platform according to the needs of the research group. In this case, it is highly recommendable to move the whole dataset to an online repository¹⁴, as open data that will support the research findings and can be reused, with appropriate attribution, in other research projects.

6. Conclusion

In conclusion, we can say that online collaborative platforms should be included in the toolkit of today's cultural heritage research, as a valuable option to establish new collaborations, to gather new data and to enrich the quantity and quality of existing datasets (figure 5). There are several keys to success, that we can briefly summarize: develop a user-friendly and feature-rich platform to publish the existing dataset online; choose a modern and clean mobile-friendly layout; implement data integrity mechanisms, automatic notifications and content versioning to keep track of contributions; carry out engaging campaigns, targeted directly at potential contributors, interest groups and local communities; commitment from the research group to disseminate, use and maintain the platform, as well as revise and use, if relevant, the contributions received.

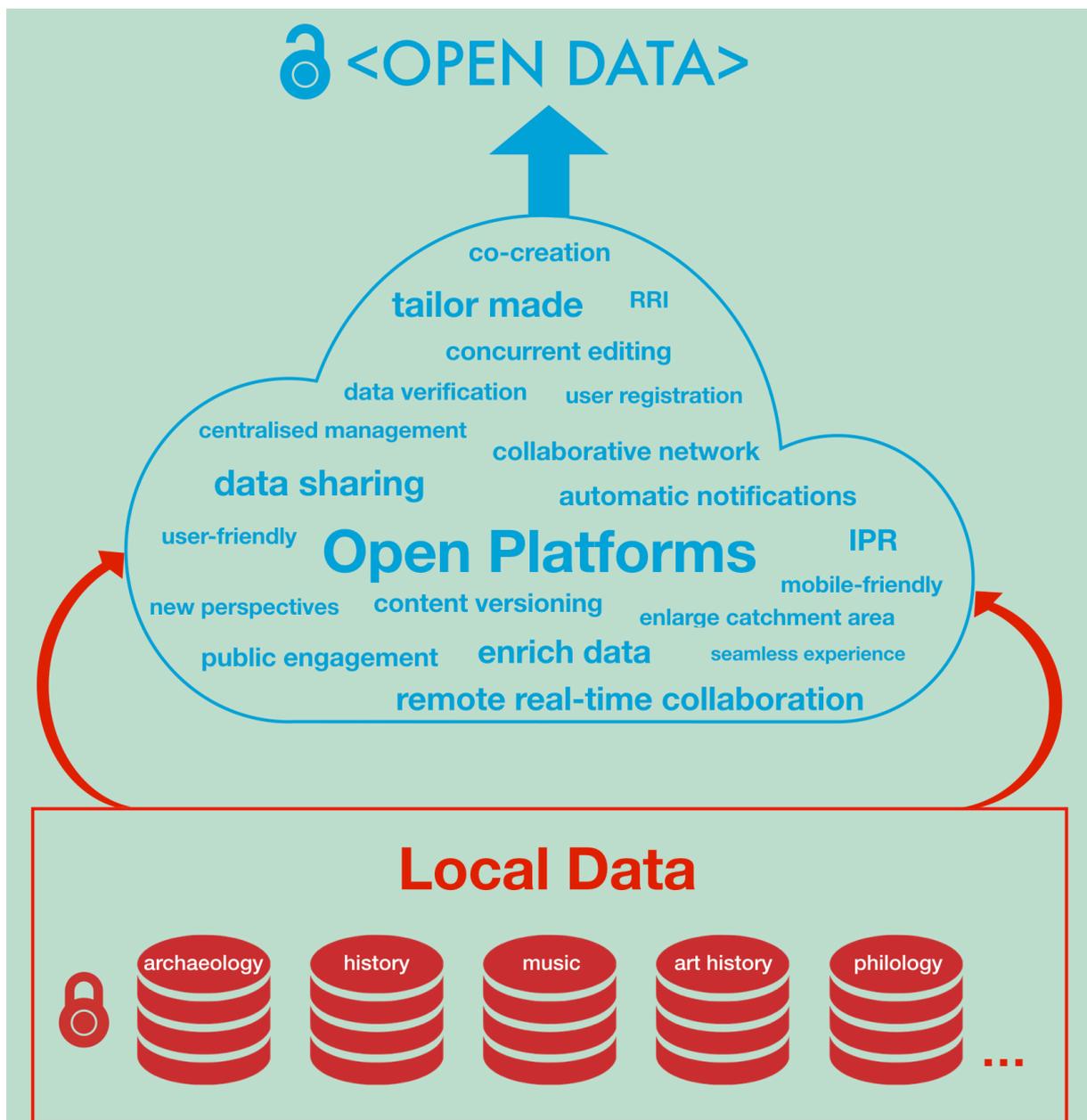


Figure 5: Diagram showing the ideal path from Local Data to Open Data, with the intermediate step of collaborative Open Platforms, and their main benefits. Source: compiled by the autor.

Embracing the Open research methodology has some clear benefits. Moving from local data to open platforms allows researchers to deepen their investigations and expand the collaboration network to other groups, as far as it would have otherwise been impossible, due to logistical, time and resource limits. In addition, this approach enriches the corpus of data with new perspectives and forces the fine-tuning of techniques for verifying information. The ability to actively contribute to the research process brings science closer to civil society and opens new methods of collaboration, helping to align research and development results with the needs and expectations of citizens.

7. Notes

¹ H2020- CULT-COOP-06-2017 - Participatory approaches and social innovation in culture. Grant Agreement No. 769827, https://cordis.europa.eu/project/rcn/212215_es.html.

² The state of Open Research Data in Europe is described by the Open Science Monitor, at <https://ec.europa.eu/research/openscience/index.cfm?pg=home§ion=monitor>.

³ For Humanities and Social Science, main metadata standards have been established by TEI: Text Encoding Initiative (<http://www.tei-c.org/index.xml>), the Visual Resources Association Core (<http://core.vraweb.org>), Dublin Core (<http://dublincore.org/documents/dcmi-terms>) and the Data Documentation Initiative (DDI) (<http://www.ddialliance.org>).

⁴ Research project *MUTIS. Patrimonio artístico y relaciones culturales entre Andalucía y América del Sur*, Ref. HAR2014-57354-P, principal researcher professor Rafael López-Guzmán.

⁵ Research project *Acequias. Proyecto de ciencia participativa basado en el estudio de los valores culturales y ambientales de los sistemas históricos de regadío de Granada y Almería*, Ref. FCT-15-9814, principal researcher professor José María Martín Civantos.

⁶ Research project *Familias, desigualdad social y cambio generacional en la España centro-meridional, 1700-1900*, Ref. HAR2013-48901-C6-6-R, principal researcher professor Francisco García González.

⁷ <http://colabora.andaluciayamerica.com>.

⁸ <http://regadiohistorico.es>.

⁹ <https://www.openstreetmap.org>.

¹⁰ <http://casaslecrin.es>.

¹¹ According to research compiled by WebHostingBuzz in 2013, 86% of users report being bothered by having to create new accounts on websites. Retrieved from <http://www.webhostingbuzz.com/blog/2013/03/21/whos-sharing-what/> [12/02/2018].

¹² <http://historicgraves.com>, developed by John Tierney and the author of this paper at Eachtra Archaeological Projects Ltd.

¹³ <http://exiliadosrepublicanos.info>.

¹⁴ In social science and humanities, some of the most common data repositories are Zenodo (zenodo.org), Harvard Dataverse (dataverse.harvard.edu), Archaeology Data Service (archaeologydataservice.ac.uk), openICPSR (openicpsr.org), figshare (figshare.com) and Mendeley Data (data.mendeley.com). The Registry of Research Data Repositories (re3data.org) is an Open Science tool that offers an overview of existing international repositories for research data.

8. References

- Bocanegra Barbecho, L., Toscano, M., Delgado Anés, L. (2017). *Co-creación, participación y redes sociales para hacer historia. Ciencia con y para la sociedad*, in *Historia y Comunicación Social*, Vol. 22-2, Universidad Complutense de Madrid, 2017, ISSN: 1137-0734 [<http://hdl.handle.net/10481/48545>]
- Bocanegra Barbecho, L., Toscano, M. (2015). El exilio republicano español: Estudio y recuperación de la memoria a través de la web 2.0. Nuevo enfoque metodológico con el proyecto e-xiliad@s, en *Migraciones & Exilios*, 15: 113-136. [<http://hdl.handle.net/10481/45829>]
- Dombrowski, Q. (2016). *Drupal for humanists*. College Station: Texas A&M University Press, ISBN: 9781623494728. <https://lccn.loc.gov/2016009099>
- Fresa, A., Justrel, B. (2015). *Roadmap for Citizen Researchers in the Age of Digital Culture*, CIVIC EPISTEMOLOGIES FP7 project. Retrieved from http://www.civic-epistemologies.eu/wp-content/uploads/2014/08/CE_Roadmap-Handbook.pdf
- Goings, K., Abel, P. (2013). *The Value of Social Login. Solving the Engagement Gap. Insights from Consumer Research*. Retrieved from <http://www1.janrain.com/rs/janrain/images/Industry-Research-Value-of-Social-Login-2013.pdf>
- *Inventario y Caracterización de los Regadíos de Andalucía. Distritos Mediterráneos y Atlánticos*, Junta de Andalucía, Consejería de Agricultura y Pesca. 2008, [Dataset]. Retrieved from <https://goo.gl/7XSzHq>
- Nielsel, J. (2006). *The 90-9-1 Rule for Participation Inequality in Social Media and Online Communities*, Retrieved from <https://www.nngroup.com/articles/participation-inequality> [12/02/2018]
- Toscano, M. (2018). *Andalucía y América. Plataforma colaborativa sobre el patrimonio artístico iberoamericano*, Quiroga. *Revista de Patrimonio Iberoamericano*, Num. 13 (Enero-Junio 2018). Universidad de Granada, ISSN: 2254-7037 – 2018
- Toscano, M., Tierney, J. (2016), *HistoricGraves.com. Explore, Protect, Promote*. Retrieved from <http://historicgraves.com/sites/default/files/presentation/index.html> [12/02/2018]

IV.4.2. Crowdsourcing en humanidades

El artículo que presentamos a continuación analiza el fenómeno del *crowdsourcing* (Carletti *at al.* 2013; Ridge 2013; Ridge 2014; Terras 2016; Hedges y Dunn 2018) y del trabajo en red, empleados por disciplinas humanísticas a través de sistemas de información web personalizados. En investigación científica, el *crowdsourcing*, como fenómeno habilitado a través de la Web, toma forma a partir de 2007 con el proyecto Galaxy Zoo¹³⁵, orientado a la clasificación de imágenes de galaxias desde la *Sloan Digital Sky Survey*. Desde esta iniciativa se genera, en 2009, la que es actualmente la mayor plataforma para el *crowdsourcing* científico multidisciplinar: Zooniverse (véase apartado III.1), portal de Ciencia ciudadana de la *Citizen Science Alliance*¹³⁶. Desde entonces, esa metodología se ha aplicado a una variedad de tareas que podemos enumerar en las siguientes: transcribir o corregir textos manuscritos digitalizados; etiquetar fotos con metadatos para facilitar su recuperación y organización (*social tagging*); realizar entradas de datos estructurados o semi-estructurados; buscar activamente objetos o fuentes adicionales para incorporarlos en colecciones digitales existentes; utilizar curadores no profesionales para montar exposiciones en línea; registrar experiencias personales y recuerdos en forma de historia oral. En general, se pueden distinguir dos tendencias: 1) proyectos que requieren el *crowd* para integrar o enriquecer recursos existentes (ej. metadatos, etiquetados, mejora de OCR, geolocalización de fotos antiguas, etc.); y 2) proyectos que piden al *crowd* que cree o aporte nuevos recursos (ej. microhistoria, recolección de fuentes primarias, etc.).

Como en otros apartados de esta tesis, la reflexión y el análisis de la temática toma impulso desde una serie de casos de estudio que, de manera diferente, han llegado a implicar la sociedad civil en sus actividades. El principal resultado de esta publicación es el análisis conjunto sobre tres nuevas metodologías de trabajo, la co-creación, la participación pública y el uso de las redes sociales, así como las implicaciones y el impacto de su aplicación en el proceso de investigación.

En el marco de este artículo, el autor de la presente tesis doctoral se ha ocupado del diseño y de la implementación de las infraestructuras digitales que subyacen los tres proyectos analizados, en colaboración con los respectivos investigadores para la conceptualización de los modelos y la definición de las funcionalidades de cada plataforma, según los principios metodológicos expresado anteriormente (véase apartado III.2). Además, ha colaborado con los coautores en la conceptualización del artículo, en la redacción conjunta del borrador inicial así como en su revisión y edición y en la implementación de las sugerencias de los revisores. Por otra parte, es de su exclusiva autoría la redacción de la sección relativa al proyecto *Historic Graves*¹³⁷.

¹³⁵ <https://zoo4.galaxyzoo.org>.

¹³⁶ <https://www.citizen-sciencealliance.org/>.

¹³⁷ <https://historicgraves.com>.

Cabe destacar este caso de estudio por su magnitud, su trayectoria y su trascendencia a nivel internacional. Se trata de un sistema de información web desarrollado para el mapeo y la transcripción colaborativa, que cuenta actualmente con un total de más de 100.000 fichas registradas, entre cementerios históricos y tumbas geo-localizadas y transcritas¹³⁸. Recientemente, se ha cruzado el umbral de las 200.000 personas¹³⁹ cuyos datos han sido transcritos, desde los respectivos epitafios, por parte de más de 15.000 usuarios registrados y vienen consultados por una media de alrededor de 20.000 usuarios mensuales. Desde la publicación del artículo a finales de 2017, la iniciativa, que se enmarca en el ámbito disciplinar de la Historia Pública Digital, ha conseguido una serie de hitos que respaldan la bondad del modelo de datos y de la estrategia participativa y co-creativa que la han caracterizado desde su creación y que se analizan a continuación. En 2018, el proyecto fue seleccionado y subvencionado por el *Irish Heritage Council*, en el marco del *European Year of Cultural Heritage* y, gracias a esa financiación, se ha actualizado su infraestructura informática (Toscano 2019b). Asimismo, en 2019, la iniciativa editorial *Visualizing Objects, Places, and Spaces: A Digital Project Handbook*¹⁴⁰ incorporó *Historic Graves* entre una selección de proyectos digitales particularmente significativos por sus fases de implementación y modelo de financiación (Toscano y Tierney 2020). Finalmente, en 2020, la plataforma ha sido seleccionada por el proyecto europeo *Social Sciences and Humanities Open Cloud (SSHOC)*¹⁴¹ para la certificación *CoreTrustSeal*¹⁴², como repositorio de datos en abierto, encontrándose actualmente en fase de evaluación para este prestigioso sello de calidad (véase capítulo V). Muy recientemente, la sostenibilidad del proyecto a medio plazo se ha visto reforzada por su incorporación, como caso de estudio, en la acción de innovación H2020 INCULTUM - *Visiting the Margins. INnovative CULTural ToUrisM in European peripheries*¹⁴³ (2021-2024), focalizada en valorizar el potencial en términos de turismo cultural de zonas marginales y periféricas en Europa.

La trayectoria de este proyecto de co-creación, que el candidato ha podido seguir de cerca de partir de su conceptualización, nos permite también avanzar una serie de lecciones aprendidas sobre la metodología del *crowdsourcing*¹⁴⁴:

1. En primer lugar, podemos confirmar que, aunque con diferentes niveles de implicación, hay un número importante de personas que desean implicarse en la transcripción voluntaria de fuentes históricas (Oomen y Aroyo 2011; Ralli *et al.* 2020).

¹³⁸ En concreto, 105.845 tumbas históricas y 875 cementerios registrados, al 20 de diciembre 2020.

¹³⁹ En concreto, datos personales procedentes de 204.857 personas, a fecha de 20 de diciembre de 2020.

¹⁴⁰ <https://handbook.pubpub.org/>

¹⁴¹ <https://sshopencloud.eu/>.

¹⁴² <https://www.coretrustseal.org/>.

¹⁴³ Convocatoria TRANSFORMATIONS-04-2019-2020 - *Innovative approaches to urban and regional development through cultural tourism*.

¹⁴⁴ Para una panorámica sobre otros proyectos que han utilizado con éxito el *crowdsourcing* en Humanidades, véase Toscano 2019a.

2. El *crowdsourcing* es un trabajo colaborativo entre los voluntarios y los responsables de la iniciativa, quienes tienen que mantener un rol activo ya sea en optimizar el flujo de trabajo, como en favorecer y fomentar la participación de nuevos usuarios y obviamente en resolver incidencias.
3. Aunque el *crowdsourcing* sea una metodología de investigación participativa estrechamente vinculada a la Web, no implica que la interacción con los usuarios ocurra solo en línea: la combinación entre la interacción web y los talleres a nivel local ha contribuido en incrementar y consolidar la participación.
4. La publicación instantánea de las transcripciones ha demostrado ser muy atractiva para los grupos de voluntarios que participan (véase apartado IV.3): ellos ven los resultados de su trabajo de transcripción no mediatos por terceros y pueden compartir de inmediato los registros con familiares y amigos que viven en el extranjero.
5. Las redes sociales pueden jugar un importante rol complementario, en consolidar una comunidad de contribuidores con un fuerte compromiso hacia un objetivo común de creación de nuevas fuentes de información histórica.

Una serie de pruebas de verificación sobre la calidad de las transcripciones, llevadas a cabo en 2017 por el equipo de coordinación del proyecto sobre 280 fichas escogidas aleatoriamente, ha detectado un porcentaje residual de errores (4,6% de fichas con errores de transcripción en fechas; 3,8% con errores ortográficos), muchos de los cuales no relevantes para la localización y lectura de la ficha. Estas cifras han demostrado un alto nivel de calidad en el trabajo realizado por el *crowd*, confirmando los resultados obtenidos por otros estudios (Borromeo *et al.* 2016). Además, en *Historic Graves*, el propio mecanismo de co-creación, donde los usuarios registrados tienen acceso directo a la edición de cada ficha, habilita y favorece la mejora progresiva de las transcripciones. Los usuarios no registrados pueden acudir al formulario de contacto para señalar errores de transcripción (ej. 14 señalizaciones recibidas en 2020). No se han registrado fenómenos de vandalismo digital sobre el contenido de la plataforma.

Referencia bibliográfica

Bocanegra Barbecho, L., Toscano, M., & Delgado Anés, L. (2017). Co-creación, participación y redes sociales para hacer historia. *Ciencia con y para la sociedad. Historia y Comunicación Social*, 22-2, pp. 325-346. <https://doi.org/10.5209/HICS.57847>

Título original

Co-creación, participación y redes sociales para hacer historia. Ciencia con y para la sociedad

Título en inglés

Co-creation, participation and social networks to make history. Science with a for Society

Abstract

The trend towards citizen participation in scientific research is here to stay. Direct involvement of individuals and groups linked with the study theme open up the possibility to access privileged information and new primary sources. It can even modify the objectives and methodology of a project, setting the course in the development of research activities. These factors increase when we talk about humanities and are further enhanced through the use of digital technologies, Web platforms, apps and social networks, that reduce distances and extend the tools available to the researcher.

Resumen

La tendencia hacia la participación ciudadana en los procesos de investigación científica ha llegado para quedarse. La involucración de personas y colectivos vinculados a las temáticas de estudio permite el acceso a información privilegiada a nuevas fuentes primarias; puede llegar incluso a modificar los objetivos y la metodología de un proyecto, marcando el rumbo en el desarrollo de las actividades. Estos factores se incrementan cuando hablamos de proyectos en ciencias humanas y se potencian todavía más a través del uso de las tecnologías digitales, plataformas web, aplicaciones y redes sociales, que recortan las distancias y amplían las herramientas a disposición del investigador.

Keywords

co-creation; participation; social networks; open data; Web platform; open research.

Palabras claves

co-creación; participación; redes sociales; datos abiertos; plataforma web; investigación abierta.

1. Introducción

El presente artículo parte de las experiencias de tres iniciativas que de manera diferente llegaron a implicar la sociedad civil en sus actividades. En algunos casos esta

involucración nació desde la necesidad, en otros fue más planeada desde el principio: en todos ha penetrado hasta el fondo del proyecto, con un efecto directo en la metodología, en los resultados, en la difusión y en el impacto. Los conceptos de co-creación (Waissbluth *et al.*, 2014; Nevens *et al.*, 2013; Jiron, 2015), participación (Arnstein, 1969; Dickson, 2001; Lewenstein, 2003) y redes sociales (Ponce, 2012; Martínez Gutiérrez, 2010) que a través de estos casos de estudio queremos abarcar no son nuevos, pero tampoco están muy bien asentados, especialmente desde el ámbito de las humanidades. Se encuentran en una fase de definición y re-definición conceptual y metodológica en la que queremos entrar aportando al debate reflexiones y técnicas que han evolucionado desde la práctica y las actuaciones en vivo, evitando una reseña bibliográfica que queda fuera del alcance de este artículo. Los casos de estudio, de los cuales hablaremos a continuación, son proyectos que han trabajado con colectivos e individuos muy diferentes, que han aprovechado redes ya existentes y han creado nuevas, donde están teniendo impactos locales y globales muy importantes.

Hablaremos de co-creación, como una metodología de colaboración a diferentes escalas entre vecinos del barrio y desconocidos a miles de kilómetros de distancia sobre una obra colectiva en donde las autorías se difuminan y prima el resultado. Un proceso que se concreta en diferentes técnicas de trabajo en campo y tele-trabajo, en voluntarios y profesionales colaborando de cerca cada uno con sus competencias, con el orgullo de concurrir en crear y compartir un recurso multimedia de historia local y memoria familiar.

Hablaremos de participación ciudadana, más allá de la transferencia y de la difusión de los resultados a la sociedad. Una participación que puede llegar a definir los objetivos mismos de un proyecto de investigación histórica en donde los ciudadanos de a pie son un actor más en el proceso historiográfico y se convierten en agentes para el rescate de fuentes históricas primarias al servicio de la ciencia. La memoria histórica se recompone desde abajo de forma participativa y contributiva.

Y hablaremos de redes sociales, de las muchas que apoyan desde siempre las relaciones interpersonales y que hoy en día lo vuelven a hacer de forma prácticamente ininterrumpida. A partir de las redes locales existentes desde antaño a las cuales los investigadores acuden para conectar con colectivos que poseen información en riesgo de desaparecer, hasta las más recientes redes sociales digitales, horizontales o temáticas, a través de las cuales el investigador llama a la participación y difunde actividades participativas sin ninguna solución de continuidad entre lo digital y lo físico.

En las conclusiones llegaremos a exponer las reflexiones más valiosas, a rescatar los pasajes más significativos que estos proyectos han vivido a lo largo de su trayectoria y finalmente a explicitar nuestra visión sobre cómo hacer trabajar las ciencias humanas con y para la sociedad.

2. Casos de estudio

2.1 HISTORIC GRAVES: UNA PLATAFORMA DE CO-CREACIÓN

Historic Graves nace en Irlanda a finales de 2010 a partir de una idea de John Tierney¹ que ha sido desarrollada a nivel informático por Maurizio Toscano², autor de este artículo, trabajando conjuntamente en la empresa de servicios arqueológicos Eachtra Archaeological Projects Ltd³. Inicialmente la estructura de las fichas se basó en el manual *Guidance for the Care, Conservation and Recording of Historic Graveyards* (O'Brien, 2010) y desde entonces el proyecto ha recogido ideas y sugerencias de varias personas que han colaborado en el mismo. En un principio el objetivo era publicar en línea las fotos de las 363 tumbas del cementerio de Ardmore, en el condado de Waterford, al sur de la República de Irlanda. Rápidamente nos dimos cuenta que la plataforma web que habíamos creado podía alojar, con la misma estructura, cualquier otro cementerio, convirtiéndose en un registro multimedia y estandarizado a nivel nacional. Desde entonces en historicgraves.com se han publicado casi 800 cementerios, más de 80.000 tumbas y se han transcrito los detalles de 159.793 personas⁴.

Originalidad e innovación

La metodología de trabajo surgió desde las competencias arqueológicas comunes a todas las personas que estuvieron involucradas en un principio y se fue refinando en el campo a través de aproximaciones progresivas. En poco tiempo establecimos un protocolo rápido, fácil de aprender y económico que inicia con la numeración creciente, con cinta adhesiva y rotulador, de las lápidas en el campo y sigue con la toma de fotos individuales con cámara GPS o smartphone y la realización de un croquis con la localización recíproca de los elementos principales del cementerio.

En cuanto se llegó a tener un número considerable de fotografías publicadas en línea, se hizo clara la necesidad de digitalizar las inscripciones para introducir un buscador interno y permitir la indexación de las fichas por parte de los motores de búsqueda. Se llegó entonces a calcular el tiempo necesario y nos percatamos que el trabajo de transcripción no era abarcable por el restringido grupo de profesionales que estaba detrás de la iniciativa: se hacía necesario abrir la plataforma a la sociedad civil. Los primeros colectivos que se llegaron a involucrar fueron las comunidades locales, por los estrechos vínculos que muchas de ellas ya tenían con los cementerios de sus alrededores⁵. Son una realidad muy fuerte en el país, sobre todo en la Irlanda rural. De hecho, hasta ahora más de 500 colectivos han colaborado desde abajo con el proyecto, implicándose y dejando su marca bien impresa. En el campo llevan a cabo todo el proceso: desde la toma de las fotos a la lectura de las inscripciones con técnicas no invasivas⁶; a través de la plataforma se encargan de transcribir los textos y completar los detalles de las personas que en ellos aparecen; han llegado incluso a modificar la plataforma web proponiendo cambios en la estructura y en las funcionalidades que han mejorado su

usabilidad. Hasta han marcado la propia cobertura geográfica del proyecto en estos años, impulsando el trabajo en un área más que en otra y difundiendo la iniciativa a través de sus propias redes personales y profesionales. Además, en muchos casos se encargan también de resolver o prevenir los posibles problemas que puedan surgir sobre la privacidad y el uso de estos datos: conocen sus gentes, los vecinos y hasta las sensibilidades individuales de los familiares. Algunas de estas comunidades llevan años colaborando en el proyecto, han empujado la prospección de docenas de cementerios en su área y transcritos miles de inscripciones funerarias; entre ellas nos gustaría citar los *Retired Active Men's Social* en Dublín, Ballyhoura en el condado de Limerick y el colectivo de West Cork (figura 1).



Figura 1. El colectivo de West Cork trabajando en un cementerio rural. Fuente: *Eachtra Archaeological Projects Ltd.*

La sencillez y la economicidad de la metodología descrita anteriormente fue clave para la participación directa de las comunidades locales que no encontraban ningún obstáculo a la hora de sumarse a la iniciativa e involucrarse en el trabajo de campo. Otro aspecto que de manera inesperada fue fundamental para el éxito de Historic Graves fue la rapidez en la publicación del trabajo realizado en los cementerios. En pocos días, o incluso el mismo día, nuestros grupos de voluntarios podían ver en Internet y compartir los resultados de su trabajo, además de seguir completando las fichas desde casa. Una de las claves del éxito fue publicar inmediatamente los datos recopilados en el campo, aún incompletos y fragmentados, para después completar y corregir la información en la plataforma web, a través de un proceso transparente de co-creación y colaboración.

Esta red de comunidades locales constituye la base de nuestra red digital de usuarios que va mucho más allá y se extiende en todos los países que han vivido el fenómeno de la diáspora irlandesa, con alrededor de 9.000 usuarios registrados en la plataforma hasta la fecha. Como en cualquier comunidad digital, el nivel de implicación y de actividad es muy variado y contamos desde colaboradores ocasionales que solamente han transcrito un registro, un 10% del total, hasta el 4% de usuarios más activos que han colaborado en la transcripción de entre 500 y 5.166 fichas individuales, en el caso de nuestro usuario más comprometido (figura 2).

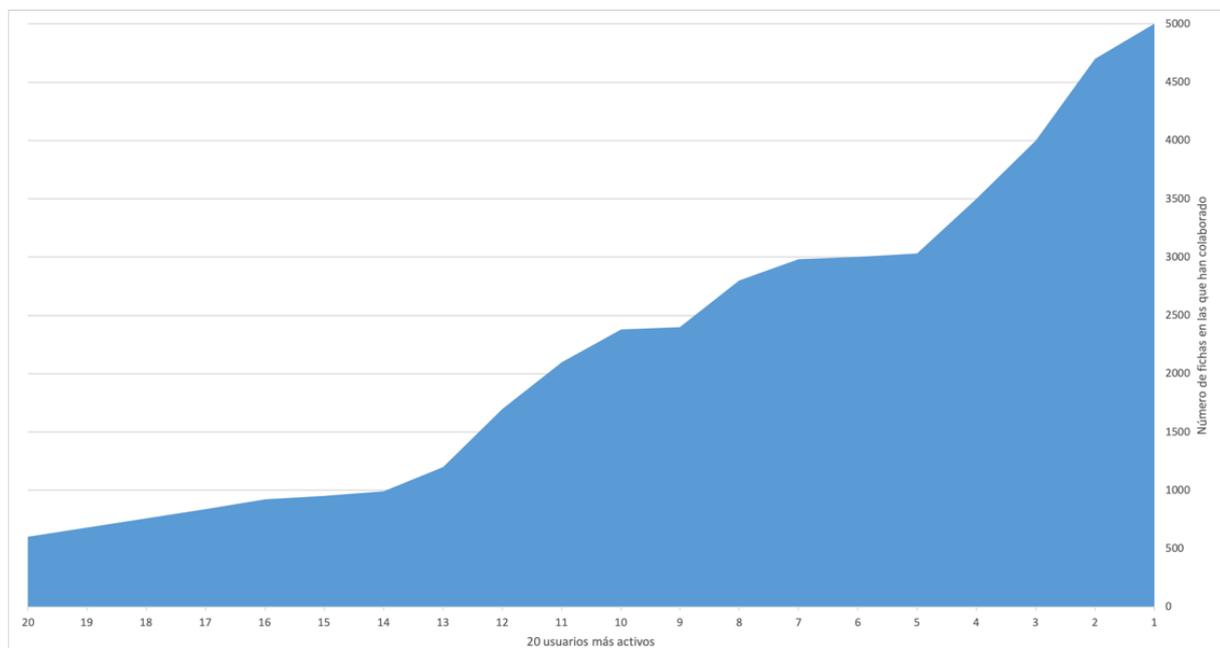


Figura 2. Actividad de los 20 usuarios más activos en la plataforma web. Fuente: elaboración propia.

Cada vez que un usuario modifica algún dato en la ficha de una tumba, el sistema crea una nueva versión y se registra automáticamente la fecha, la hora, el autor y los cambios realizados. Aún si todo queda registrado, la autoría es difusa como en una wiki, el papel de cada colaborador queda difuminado y la colaboración, voluntariamente, no queda evidente al consultar la plataforma. La co-creación es horizontal y se ha producido en todos los niveles de agregación de datos. En el corpus general⁷ han llegado a colaborar activamente hasta la fecha más de 500 usuarios; en los proyectos individuales⁸ de uno o más cementerios, la colaboración en línea normalmente coincide con el grupo de voluntarios de la comunidad local de referencia; en las fichas individuales⁹, en varias ocasiones el trabajo de transcripción realizado inicialmente por un único colaborador es mejorado por otros que detectan errores de lectura o falta de completitud y los arreglan al momento. Como se ha dicho anteriormente, la publicación instantánea de los cambios, sin intermediarios, es muy atractiva para los voluntarios, al poder ver los resultados inmediatos y reales de su colaboración.

En los años hemos implementado y mejorado varios mecanismos de control sobre la calidad y la veracidad de los datos publicados, que actúan a tres escalas: centralizados, automáticos y distribuidos. Los mecanismos centralizados de control se refieren a la creación de las fichas individuales para nuevos cementerios y sus respectivas lápidas, a la asignación de los códigos unívocos y a la subida de las fotografías geo-localizadas que corresponden únicamente al equipo que gestiona la plataforma. Los controles automáticos consisten en la verificación en fase de registro de los usuarios a través de un sistema CAPTCHA¹⁰ y filtros anti-spam para la detección y eliminación de posible publicidad en los campos correspondientes a las transcripciones. Finalmente, los mismos usuarios, registrados o no, suelen detectar autónomamente errores de lectura o de ortografía en las transcripciones y después actúan de dos formas: o bien los corrigen en línea o nos envían un correo al respecto, al cual siempre respondemos invitando a crear una cuenta y aportar los cambios directamente.

La plataforma web historicgraves.com no es una red social en el sentido actual del término, porque en ella no se recogen opiniones, no se publican comentarios y no se comparten hechos personales de los usuarios. En Historic Graves se recopilan datos estructurados y relacionados, que proyectan las comunidades y sus territorios en el pasado, a través de las relaciones familiares, la onomástica y la antroponimia, los estilos del grabado, la evolución del lenguaje y la toponimia y cuanto más queda reflejado en las inscripciones funerarias. Al ser un trabajo co-creativo y comunitario, desde la gestión de la plataforma se reconoce a la misma comunidad la propiedad de los datos, asignándole la prerrogativa de descargar, en formato tabular, el conjunto de datos que han colaborado a recopilar.

Con los años hemos aprendido que los usuarios que a diario entran en la plataforma lo hacen por un propósito común, contribuir en la creación de un *asset* nacional, colaborar en registrar y descifrar las inscripciones que a lo largo de los últimos tres siglos sus antepasados han dejado impresas en la piedra y que son una parte importante de la memoria colectiva de su lugar. Al colaborar en una plataforma en línea, además proporcionan y proyectan esta misma memoria a miles de kilómetros de distancia, donde hoy en día se encuentran a menudo los familiares de estos difuntos (ej. Estados Unidos o Australia).

Con respecto a las muchas otras iniciativas que en el pasado habían abarcado la tarea de documentar y mapear los cementerios históricos en Irlanda, Historic Graves supo tomar ventaja de las nuevas tecnologías al alcance de todos y se diferenció con respecto a las otras plataformas existentes gracias a la participación activa y constante de las comunidades locales, unidas de forma independiente en la co-creación de un recurso colectivo, público y gratuito.

Los medios tradicionales, prensa y telediarios, han hablado en alguna ocasión de la iniciativa a nivel nacional, pero donde la repercusión ha sido más profunda es a nivel local, en las muchas comunidades que en estos años han colaborado a través de la plataforma, con un impacto directo en relación con la memoria y el cuidado de su entorno. En algunos casos este impacto ha evolucionado también en iniciativas paralelas como conferencias temáticas, la edición de libros y jornadas dedicadas al mantenimiento de estos espacios monumentales, en ocasiones completamente abandonados a la incuria, y seguramente en otras actividades desarrolladas de forma independiente de las cuales no tengamos constancia.

El alcance fuera de la República de Irlanda se puede medir a través de dos canales: las estadísticas de visitas a la página y los correos que recibimos a través del formulario de contacto. Desde 2011 la página web ha recibido un total de más de 600.000 visitas, con un *trend* creciente que ha alcanzado la media de 10.000 usuarios mensuales en el último año. Alrededor de la mitad de este tráfico se genera dentro de Irlanda, mientras que el resto se reparte entre los cinco tradicionales destinos de la emigración irlandesa, por orden: Estados Unidos (19%), Gran Bretaña (14%), Australia (6%), Canadá (3%) y Nueva Zelanda (1%). A nivel cualitativo, tenemos que valorar la gran cantidad de correos que recibimos a través de la página web y que en buena medida nos describen historiadores y apasionados que pasan horas leyendo las inscripciones o nos hablan de reencuentros virtuales entre nietos y bisabuelos, estadounidenses o australianos los unos e irlandeses los otros, y de viajes que se planifican para hacer estos reencuentros más reales. Estos mensajes, además de ser un agradecimiento para el esfuerzo colectivo, en buena medida anónimo, que ha convertido Historic Graves en una plataforma de referencia para la búsqueda de inscripciones funerarias en línea, nos hablan del potencial de este recurso para el turismo genealógico. Miles de irlandeses de segunda o tercera generación viven fuera del país, pero mantienen un enlace emocional con la tierra de sus antepasados que Historic Graves puede contribuir a reavivar.

Nos gustaría cerrar la descripción de este caso de estudio con una reflexión sobre el potencial de este valioso corpus de imágenes e inscripciones en campos y disciplinas tan diferentes como la historia local, la iconografía religiosa, las tradiciones fúnebres católica y protestante, la evolución de la escritura y de la ortografía, las diferentes escuelas de canteros y los estilos del grabado hasta la creación de productos independientes como libros y recorridos temáticos. A diferencia de los otros dos casos de estudio que aquí se presentan, Historic Graves no nació como un proyecto de investigación, pero ha creado las bases desde las cuales muchas investigaciones podrían surgir.

2.2 E-XILIAD@S: UN PROYECTO PARTICIPATIVO

En el año 2010 sale a la luz el proyecto digital e-xiliad@s a partir de una idea de Lidia Bocanegra Barbecho¹¹. Se trata de una iniciativa basada en la recogida de fuentes inéditas *on-line*, a nivel internacional, de los exiliados republicanos españoles anónimos¹². Utiliza una plataforma web 2.0 multilingüe (español, inglés y francés) con la siguiente URL: <http://exiliadosrepublicanos.info/>.

Las propias características de movilidad internacional de esta diáspora republicana¹³, junto con la cantidad de gente que se exilió (unos 465.000 en 1939), hace que nos encontremos ante un fenómeno histórico asociado a una gran cantidad de fuentes inéditas esparcidas por casi todo el mundo. Esto explica el porqué de la iniciativa; pero no únicamente, otros factores alimentan también la necesidad de un proyecto de estas características, tales como:

- la dificultad de consultar las fuentes: archivos lejanos, normativas de accesos y costes de reproducción;
- la existencia de material desconocido acerca del exiliado republicano anónimo (nuestro objeto de estudio);
- la todavía escasa documentación en la red internet, debido a la falta de digitalización de muchos archivos públicos y privados (asociaciones, Archivo del Centro Documental de la Memoria Histórica, archivos personales tales como el Fanny Edelman o Archivos institucionales tales como la de los Centros Republicanos en Argentina, por citar algunos ejemplos).

Si bien muchos de esos exiliados retornaron a España entre abril y mayo de 1939, lo cierto es que cerca de unos 200.000 no lo hicieron y entraron, en muchos casos, en una vorágine de búsqueda de países en donde quedarse, sorteando también una II Guerra Mundial. Dicho de otro modo, no siempre se quedaron en el primer país en donde llegaron, para muchos de ellos Francia fue solamente un lugar de paso dirigiéndose posteriormente a México y luego a Argentina, etc. Muchos otros se exiliaron sucesivamente, de manera clandestina, desde España a lo largo del primer franquismo (1939-1959), especialmente a Latinoamérica.

Originalidad e innovación

La característica principal del proyecto e-xiliad@s es su carácter participativo en tanto que se nutre y crece en base de las informaciones y datos que los usuarios ofrecen en línea acerca de exiliados republicanos anónimos. Dentro de la categoría establecida por Bonney en cuanto a la participación pública en la investigación científica, e-xiliad@s se trataría de un proyecto contributivo ya que ha sido diseñado por un científico y para el cual los miembros del público principalmente aportan datos (Bonney, 2009:11). Hay dos tipos de usuarios bien diferenciados en el proyecto:

- usuario activo: quien participa, previo registro en el proyecto mediante acceso identificado, ofreciendo datos inéditos acerca de uno o varios exiliados. Asimismo, este tipo de usuario también aprovecha de la información aportada por otros participantes;
- usuario pasivo: quien consume información del proyecto, pero no ofrece datos.

La propia temática de estudio hace que el público objetivo del proyecto sean familiares o gente que haya tenido un contacto directo con los exiliados republicanos (amistoso o profesional); o bien investigadores o estudiantes que los estén investigando. Desde el inicio, el proyecto tuvo muy claro a quién iba dirigido, desarrollando una metodología y unas estrategias de obtención de datos con el fin de fomentar la participación de ese público objetivo y convertirlo finalmente en usuarios activos de la plataforma.

De esta manera, la principal originalidad e innovación de la iniciativa es la metodología participativa que ha ido desarrollando desde el ámbito de las humanidades digitales y que combina las tecnologías de la información y de la comunicación, la web 2.0 y las redes sociales digitales (figura 3). Esta metodología de recolección de datos desarrollada tiene un conjunto de estrategias bien definidas que trabajan entre sí de una manera interrelacionada. El objetivo final de todas estas estrategias es la de atraer a tantas personas como sea posible a la página web del proyecto, con el fin de convertirlos en usuarios activos de la iniciativa. A continuación, se detallan las estrategias utilizadas:

a) estrategia participativa. El usuario participa en el proceso de la investigación ofreciendo sus recuerdos familiares, de amistades o de estudio, acerca de un exiliado republicano a través de un formulario web creado *ad hoc*. El formulario, vinculado a una base de datos relacional¹⁴, presenta una serie de preguntas, estratégicamente ordenadas para estimular la memoria del usuario, con campos abiertos y listas cerradas para recoger las respuestas (para saber más acerca de cómo está organizada la ficha véase: Bocanegra; Toscano, 2015: 120-125). Caracteriza este formulario interno la facilidad de uso del mismo.

b) Estrategia de visibilidad. Con el fin de obtener el mayor número de datos posible, la visibilidad del proyecto en Internet a través de los principales motores de búsqueda (Google, Bing, Yahoo!) es clave. Por esta razón se implementaron varias técnicas SEO (*Search Engine Optimization*), entre ellas: códigos XHTML y CSS sin errores, metadatos, URLs y enlaces que funcionan correctamente, contenido multilingüe (en español, inglés y francés), la publicación constante de nuevo contenido en la web del proyecto a medida que se reciben nuevos datos y utilización de las redes sociales digitales de la iniciativa (Facebook¹⁵ y Twitter¹⁶) para generar, debatir y difundir este nuevo contenido. Caracteriza estas redes digitales la verticalidad de las mismas, en tanto que se trata de unas redes temáticas, en nuestro caso del exilio republicano español, que utilizan como plataformas digitales unas páginas públicas de Facebook y de Twitter.

c) Estrategia e-commerce. El proyecto tiene un diseño profesional, atractivo y de fácil manejo que incorpora conocimientos de diseño web propios del comercio electrónico: posicionamiento de la información de forma estratégica (imágenes, texto, logo), presentación de los textos en párrafos cortos, utilización de imágenes sugestivas de acuerdo con el argumento. El uso de herramientas en línea de análisis del comportamiento de los usuarios (Google Analytics) permite controlar las rutas seguidas por aquellos que visitan el sitio web, visualizar cuales son las páginas más visitadas y el tiempo de permanencia en cada una de ellas. Usando estos parámetros, entre otros, se puede modificar el contenido y la ubicación del mismo para captar la atención del usuario y favorecer su posible participación.

d) Estrategia de intercambio recíproco. El uso de una estrategia basada en el "yo te doy, tú me das" ayuda a difundir la iniciativa, generando más confianza entre los usuarios activos nuevos y existentes dispuestos a depositar su memoria, o datos del exilio, en el proyecto. ¿Cómo funciona? Al llegar a la plataforma los usuarios ya encuentran disponibles en acceso abierto una serie de contenidos, biografías, testimonios, diarios, referencias biblio/web-gráficas, etc., ofrecidos por otros usuarios o elaborados por los gestores de la plataforma (yo te doy). Este contenido en abierto comunica al potencial colaborador la manera en que sus datos se utilizarían, generando confianza y favoreciendo la creación de uno o más registros en la plataforma (tú me das). Para participar, el usuario anónimo necesita identificarse a través de una cuenta en la que ofrece unos datos de contacto que, a su vez, les abren la puerta a servicios adicionales de la plataforma como un tablón de anuncios para difundir información (por ejemplo, la búsqueda de un exiliado), la posibilidad de conectarse con otros usuarios registrados y hasta el asesoramiento histórico vía correo electrónico, si procede (yo te doy).

Se ha observado, mediante el monitoreo de las nuevas fichas creadas, que la publicación de los datos en acceso abierto, en particular las biografías¹⁷ de los exiliados y las publicaciones científicas en *Open Access*¹⁸, ha coincidido con un aumento en la participación de los usuarios. Así pues, publicar de manera constante nueva información recibida y elaborada no solamente aumenta la visibilidad de la iniciativa proyecto en la red Internet, sino también favorece la participación. Igualmente ayuda a convertir el proyecto en un lugar identitario y de pertenencia del exilio republicano, en donde se recupera la memoria histórica del exiliado anónimo.

e) Estrategia de pertenencia de datos. El usuario activo puede acceder con sus credenciales de acceso, en cualquier momento, a su área personal para añadir nuevos datos o modificar cualquier ficha o archivo previamente enviado. Esta sensación de control sobre los datos otorga cierto convencimiento en el proyecto, ya que el usuario en todo momento es dueño de la información que aporta. El uso y la protección de los datos personales ofrecidos se explica en la sección de *Privacidad*¹⁹ de la iniciativa.

f) Estrategia de compartir igual. Dentro de la filosofía de acceso abierto de la iniciativa, ya sea por los datos recibidos que por la información ofrecida, el proyecto utiliza una

Licencia Creative Commons 3.0: Reconocimiento - NoComercial - CompartirIgual (by-nc-sa). Esta licencia permite el uso no comercial de los datos, permite elaboraciones a partir de los mismos previo reconocimiento de la fuente original e incluye la posibilidad de compartirlos a condición de utilizar una licencia con las mismas condiciones. De esta manera se garantiza la continuidad de utilización de esos datos bajo la misma filosofía. Los derechos de propiedad intelectual, junto con limitación de uso y responsabilidad limitada, se explican en la sección *Términos y Condiciones*²⁰ de la página web del proyecto.

h) Estrategia de ética digital. El proyecto se toma muy en serio la visibilidad profesional en línea del investigador principal, teniendo en cuenta preguntas tales como: ¿qué pueden averiguar sobre él/ella los usuarios en línea? ¿Cómo manejar las solicitudes de amistad de los usuarios en Facebook? ¿A quién seguir en Twitter? El objetivo es establecer límites profesionales entre el investigador y los usuarios participantes, así como controlar, en la medida de lo posible, la información profesional y personal acerca del investigador principal que actualmente se encuentra en Internet. En la misma línea de transparencia y profesionalidad y con el fin de incrementar la confianza en la iniciativa, crucial en un proyecto participativo, los detalles del investigador principal y la información acerca de los colaboradores y de la financiación recibida son claramente visibles en la sección *Créditos*²¹.

Referente a la incógnita de la veracidad de los datos esta es una constante en la propia investigación histórica y no una característica de los proyectos con participación ciudadana. En el caso del proyecto e-xiliad@s, parte de la información que se entrega (ficha cumplimentada) viene auto-verificada por los documentos que la acompañan: visados, pasaportes, certificados de ingresos en campos de concentración, carnets de afiliación, fotografías, etc. Tanto para aquellas fichas con documentación digitalizada asociada, como para aquellas otras que no las tienen, es responsable última de la veracidad de las mismas la persona que las entrega, tal y como se especifica en el apartado de Privacidad y Términos y Condiciones del proyecto. La plataforma e-xiliad@s actúa como un repositorio de datos, como un archivo abierto a aportaciones de información y documentos sin aplicar ex-ante mecanismos de verificación de los mismos. Como con cualquier otra fuente histórica, el control sobre la veracidad de los datos se aplica ex-post, en fase de análisis y elaboración historiográfica.

A título informativo y en la actualidad, la investigadora principal del proyecto ha examinado muchas de esas fichas para analizar temas de movilidad, identidad y cultura publicándose los resultados en revistas científicas y destacándose que nunca se ha encontrado ninguna ficha falsa.



Figura 3. Esquema general del proyecto e-xiliad@s. Fuente: elaboración propia.

Impacto

Hasta ahora hemos hecho referencia a la originalidad del proyecto e-xiliad@s desde un punto de vista de la metodología participativa. Sin embargo, a nivel de beneficio e impacto social el proyecto e-xiliad@s innova de la siguiente manera:

1. recupera la memoria histórica del exiliado republicano anónimo, hasta ahora poco analizada, a diferencia de la trayectoria de los exiliados intelectuales de renombre, de los que se ha podido estudiar bastante por la cantidad de información que de ellos nos ha llegado;
2. ofrece “desde abajo”, desde los propios interesados, de forma pública, gratuita y mediante unos parámetros estandarizados de presentación de la información, un primer repositorio en línea del exiliado republicano;
3. se ha convertido, junto con sus redes sociales digitales, en una zona identitaria y de pertenencia del exilio republicano, un lugar en donde sentirse reflejado y compartir la memoria personal acerca de este momento. Dicho de otro modo, entra a formar parte y ayuda a generar una cultura del exilio republicano en línea (para saber más acerca de esta cultura e identidad en la red internet, véase Bocanegra; Toscano, 2016: 237-252);
4. ayuda a encontrar datos sobre exiliados desaparecidos, ya sea gracias a las fichas publicadas, como a través de la publicación de anuncios y comentarios en el Tablón de anuncios²² y en las redes sociales digitales del proyecto. Por ejemplo, en marzo de 2016, se publicó un anuncio solicitando información acerca del exiliado Ángel Escobio Andraca²³, neurólogo y dirigente comunista original de Santander, quien se exilió en la antigua URSS. Este anuncio recibió varios

comentarios informando acerca de la hija y mujer de este exiliado, así como de una hermana del mismo, Luz Escobio Andraca.

¿Cuales son los resultados cuantitativos del proyecto? Actualmente, la iniciativa cuenta con:

- 469 archivos entre imágenes, documentos oficiales escaneados, artículos de prensa de la época, memorias, poesías y documentos sonoros (entrevistas), en una amplia gama de formatos (figura 4);
- 151 fichas²⁴ cumplimentadas a las que están asociados los archivos anteriormente mencionados;
- 49 anuncios publicados y 30 comentarios en respuesta a las solicitudes realizadas en los mismos;
- 346.988 visitas a la web del proyecto, con 104.373 usuarios únicos y 311.000 páginas visitadas. Las páginas que más tráfico han recibido son, por orden: Historia del Exilio, Bibliografía, Enlaces, Galería de imágenes, El proyecto, Mapa del exilio y Biografías;
- cerca de 1000 seguidores entre Facebook y Twitter, con más de 600 posts y tweets publicados.

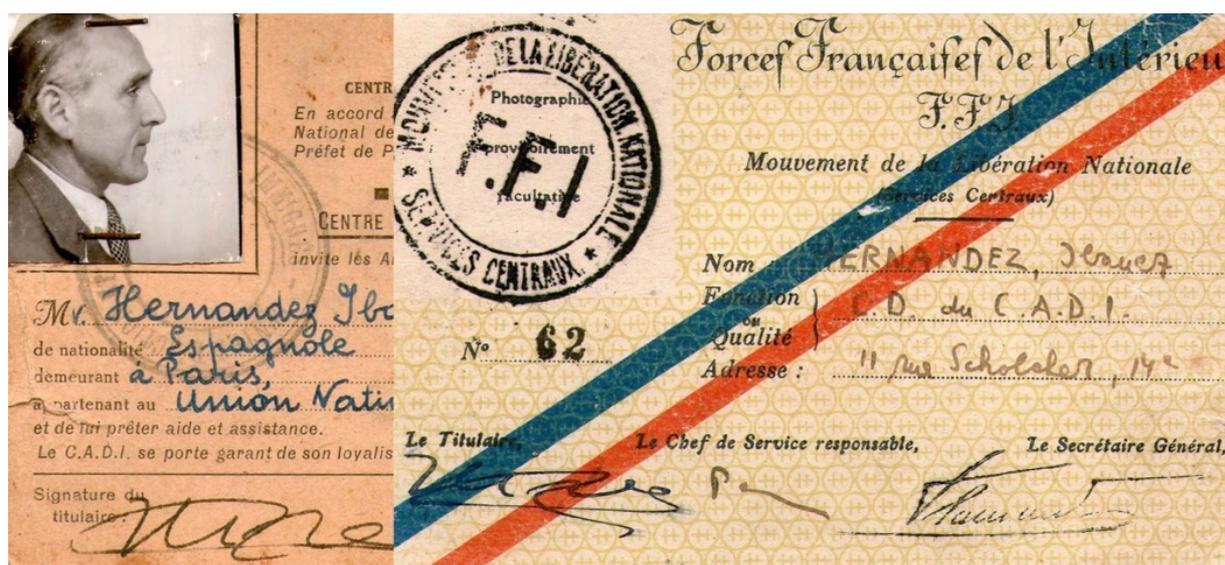


Figura 4. Detalle de los archivos ofrecidos en la ficha del exiliado Julio Hernández Ibañez de Garayo, exiliado primeramente en Francia en 1939 y, posteriormente, a Argentina en 1951 en donde fallecería 28 años más tarde.

2.3 MEMOLA Y EL USO DE LAS REDES SOCIALES

El proyecto MEMOLA²⁵, *MEditerranean MOountainous LAndscapes: an historical approach to cultural heritage based on traditional agrosystems*, estudia desde una perspectiva interdisciplinar los paisajes culturales montañosos de Sierra Nevada (España), Monti di Trapani (Italia), Colli Euganei (Italia) y Vjosa Valley (Albania) desde la tardoantigüedad hasta nuestros días, teniendo como eje central el estudio histórico de dos recursos naturales esenciales: agua y suelo. Cuenta con la participación de diez socios entre universidades, asociaciones y empresas, procedentes de cinco países europeos, bajo la coordinación de la Universidad de Granada. El objetivo principal del proyecto MEMOLA es, a través del estudio del pasado, detectar situaciones problemáticas actuales en las que el conocimiento científico puede aportar respuestas o soluciones para el futuro de las comunidades locales (Martín Civantos, 2015).

Originalidad e innovación

Uno de los pilares centrales del proyecto MEMOLA es hacer hincapié, como investigadores, en el proceso de comunicar, compartir e involucrar a la sociedad, especialmente a la población local. Estos colectivos son memoria viva y poseedores de los conocimientos que han de ser preservados para asegurar el futuro de sus paisajes. Estos grupos sociales, que se implican en la investigación, son también validadores del conocimiento científico. Para llevar a cabo este tipo de estudio ha sido necesario contar con redes sociales off-line y on-line, que faciliten la relación, la colaboración y el intercambio de conocimientos entre los investigadores y el público en general.



Figura 5. Recuperación de la acequia de Barjas (Cañar). Fuente: Lara Delgado Anés, proyecto MEMOLA.

En el 2014 se inició la primera actividad participativa basada en la recuperación y puesta en funcionamiento del sistema de riego de Barjas y del ramal de la Hijuela en Cáñar²⁶, Granada (figura 5). En ella participaron 180 voluntarios y 19 instituciones relacionadas con el patrimonio cultural y natural, entre las que había universidades, centros de investigación, instituciones públicas, centros escolares, asociaciones y colectivos locales.

Para llevar a cabo esta intervención en el territorio, desde el proyecto MEMOLA quedó evidente la necesidad de acudir a las redes sociales históricas ya existentes como son las Comunidades de Regantes locales, para involucrarlas activamente en la iniciativa. Estas comunidades tenían un vínculo emocional con el bien patrimonial, que formaba parte de su memoria y que ha sido posible volver a rescatar y potenciar a través de la restauración de las acequias de regadío.

A través del trabajo colaborativo, entre el proyecto y las comunidades de regantes, se han potenciado las redes existentes y se han creado otras nuevas entre regantes, investigadores y voluntarios. Desde el 2014 al 2016 se han recuperado, en el área de Sierra Nevada, siete acequias en colaboración con siete comunidades de regantes y con la ayuda de 330 voluntarios. Estas intervenciones en el territorio, junto a las comunidades locales, nos han permitido ampliar dichas redes sociales e involucrar nuevos miembros.

A partir de entonces surgen en paralelo otras actividades de difusión, educación, socialización y participación con diferentes colectivos, que abarcan reuniones intergeneracionales, diálogo de saberes, promoción de la organización en conjunto, visitas guiadas, actividades educativas, exposiciones, etc. ampliando así el campo de colaboración e incluyendo otros colectivos. De esta forma, el proyecto MEMOLA promueve la participación social a nivel local, ya sea a través de redes sociales ya en funcionamiento, como utilizando plataformas digitales para crear nuevos círculos sociales que se han ido involucrado en las intervenciones en el territorio, produciéndose así una simbiosis entre analógico y digital.

Hemos creado una red de medios de comunicación, tanto prensa como radio, interesados en las actividades del proyecto. Nuestra *mailing list* cuenta con 153 contactos procedentes de medios locales, provinciales y nacionales. Contamos además con otras listas cuyos perfiles son instituciones, relacionadas con los sistemas históricos y tradicionales de regadío, investigadores, voluntarios, etc.

Otro colectivo al que hemos implicado son los centros escolares de las zonas de estudio, para ello hemos creado una red educativa en cada área (Delgado Anés, 2016). En el caso de Sierra Nevada contamos a día de hoy con trece centros escolares de doce pueblos distintos, de las provincias de Granada y Almería. Esto nos permite estar en contacto directo con los docentes, compartir recursos e informarles sobre las actividades del proyecto.

Las plataformas sociales digitales que vamos a describir a continuación son redes verticales que empleamos como puntos de encuentro de los colectivos involucrados, donde se comparte información relacionada con las actividades y los resultados de los estudios que se realizan en el marco del proyecto. El empleo de un amplio abanico de redes sociales es uno de los factores innovadores del MEMOLA, así como el uso de redes especializadas en una temática muy concreta, como por ejemplo sobre los modelos tridimensionales. Trabajar con este volumen de plataformas digitales es poco habitual en proyectos de investigación y su actualización requiere una cierta inversión de tiempo, pero nos han permitido conectar con diferentes perfiles de la sociedad, ya sea en ámbito local que internacional, proyectando y ampliando el impacto del proyecto.

Facebook

La página Facebook oficial del proyecto MEMOLA²⁷ se abrió en enero del 2014 y a día 10 de enero 2017 cuenta con 1.789 “me gusta” y 1.755 seguidores. Es nuestra principal red social por el público y el volumen de usuarios que nos permite alcanzar. En ella hemos compartido 2.507 fotografías, 65 *slideshows*, 55 videos, 46 eventos y decenas de enlaces.

Nuestros seguidores coinciden en buena medida con las zonas de estudio, un 58% son de España, un 18,5% de Italia y un 8,2% de Albania. La diferencia entre géneros es reducida, con lo que no son datos muy significativos. En cuanto al índice de franja de edad, un 54% de nuestro público son de la generación *Millennials* (de 18 a 34 años), seguido de un 35% de los denominados *Baby busters* (entre 35 y 45 años) y un 10% pertenecientes a la generación de los *Baby boomer* (entre 55 y 77 años).

El público alcanzado en los últimos treinta días de datos disponibles es de 7.659 personas. El mayor número procede de los países que forman parte del proyecto, un 74,5% de España, un 10,7% de Italia y un 3,5% de Reino Unido, con una discrepancia debida al mayor número de actividades realizadas en este periodo en Sierra Nevada. La diferencia de género y las franjas de edades siguen las tendencias globales descritas arriba. Si analizamos los datos desde enero del 2015 hasta diciembre del 2016 vemos que la media de alcance orgánico diario ha sido de 572 personas, con un pico de 9.807 personas y 930 reacciones que coincide con la publicación de las noticias sobre la mezquita de la excavación del Pago del Jarafí el 24 de octubre 2016²⁸.

Respecto a las interacciones los datos geográficos se mantienen mientras se observa una relativa disparidad de género a favor de las mujeres y de los *Baby buster* a nivel generacional, a pesar de no ser nuestra franja principal de seguidores.

Facebook recientemente ha puesto a disposición una herramienta que facilita el vínculo entre diferentes redes sociales. En nuestro caso los usuarios pueden ver los contenidos de nuestras cuentas oficiales de Twitter, YouTube, Google+ e Instagram.

Esto nos permite comunicar el tráfico social y facilitar al usuario acceder por ejemplo a los documentales del canal de YouTube sin tener que salir de la interfaz de Facebook.

Twitter

Desde diciembre de 2013 hasta el 10 de enero de 2017 hemos publicado 1.249 tweets y actualmente contamos con 588 seguidores. Este perfil²⁹ se encuentra vinculado a otras plataformas web del MEMOLA, por lo que automáticamente se publican las entradas que realizamos en inglés en Facebook y los vídeos de nuestro canal de YouTube, lo cual agiliza el trabajo diario de comunicación.

Encontramos diferencias con respecto a Facebook, sobre todo con respecto a la edad de los seguidores que por más de la mitad pertenecen a la generación *Millennial*, un 38% son *Baby buster* y sólo un 2% *Baby boomer*. La procedencia geográfica coincide con los países miembros del proyecto, también si la prevalencia de España se evidencia aún más.

En Twitter hemos comprobado que la audiencia que habitualmente interactúa con nosotros son otros proyectos, investigadores, especialistas y cargos públicos. Esto difiere de Facebook donde el perfil coincide más con una audiencia no especializada, público local y estudiantes universitarios.

Google+

La página de Google+ como red social no tiene mucho uso, aunque es la segunda con más usuarios de Internet ya que se encuentra ligada a productos como Gmail o YouTube. El MEMOLA cuenta con 31 seguidores y el número total de visualizaciones asciende a 97.166. En ella compartimos los enlaces de las publicaciones que realizamos en la web, con objeto de mejorar nuestro posicionamiento en el motor de búsquedas Google, permitiendo así a cualquier usuario encontrar de forma rápida los contenidos del proyecto.

YouTube

El canal YouTube³⁰ se abrió en junio de 2014; en enero de 2017 cuenta con un total de 55 videos, 57 suscriptores y el número de visualizaciones asciende a los 15.245, que corresponden a un total de 36 días y 6 horas de reproducción.

Los videos han recibido un total de 95 “me gustas” y 51 comentarios, han sido compartidos 272 veces y en 107 ocasiones han sido añadidos a listas de reproducción. El público que ha visualizado los documentales presenta una diferencia sustancial entre géneros, con un 69% de hombres y un 31% de mujeres. Los dos grupos principales de nuestros seguidores son, a partes iguales, *Millennials* y *Baby buster*. En cuanto a la procedencia geográfica, nuestros suscriptores y el público alcanzado residen principalmente en España (74%), seguidos de Italia con un porcentaje muy inferior, un 11%; estos datos reflejan en buena medida el idioma propio de los videos ya que 48 de

nuestros documentales están en español y los 7 restantes en italiano, también si procuramos que los más significativos estén subtítulos en inglés.

Los documentales³¹ están siendo una gran herramienta de presentación y resumen de los trabajos que realizamos en las distintas áreas de estudio y tenemos constancia que están siendo utilizados como recurso educativo en institutos y en la universidad. Entre los videos destaca con 3.009 visualizaciones el documental sobre la recuperación de la acequia de Barjas.

Instagram

Es nuestra cuenta más reciente, comenzó en noviembre de 2016. En febrero de 2017 cuenta³² con 121 seguidores, 227 seguidos, 47 posts con 256 *likes*. A través de esta cuenta solemos publicar una fotografía por actividad en inglés, con el propósito de alcanzar también un público fuera de las áreas de estudio. Un dato interesante es que los *tags* que permitieron alcanzar más “me gusta” en nuestros posts fueron *education*, paisaje y *didactic*.

Entre los seguidores un 54% son hombres y un 46% mujeres y se evidencia una marcada prevalencia de *Millennials* con un 75% del total. Los países coinciden con el resto de redes sociales, con grandes diferencias entre ellos, procediendo de España más del 70%.

Sketchfab

Sketchfab es una red social sobre una temática muy concreta: modelos tridimensionales. Comenzamos a usarla en marzo del 2015³³ y hasta febrero de 2017 cuenta con 112 seguidores. Hemos llegado en este tiempo a publicar 12 modelos con un total de 1.800 visitas y 80 *likes*. El perfil del proyecto MEMOLA se encuentra entre las 30 primeras organizaciones relacionadas con el patrimonio.

Esta red social nos permite compartir modelos de las excavaciones arqueológicas. Para visualizar estos 3D el público no necesita tener cuenta, ya que permite visionar los modelos en otras redes sociales como Facebook y Twitter, así como en la página web del proyecto. Aquellos que cuentan con un perfil en Sketchfab podrán darle “me gusta” a los modelos.

Uno de los modelos más populares, con un total de 313 visitas y 13 “me gusta”, es la secuencia estratigráfica del sondeo 50000 de la excavación arqueológica de Lanteria. Sketchfab permite incluir puntos de información dentro del modelo, textos y/o fotografías, lo que enriquece y facilita la difusión de los mismos a todos los públicos. Esta plataforma social ligada a Instagram nos permite crear exposiciones virtuales. En la actualidad contamos con dos realizadas a través de Instamuseum, una sobre fotografías históricas y otra sobre fotografías relacionadas con las recuperaciones de las acequias.

Academia.edu

Academia.edu es una red social académica centrada en intercambiar publicaciones científicas. En abril del 2015 abrimos el perfil³⁴ para compartir artículos científicos relacionados con el proyecto. Actualmente contamos con 51 seguidores, 15 coautores y 614 visualizaciones. El AuthorRank, calculado en función de las recomendaciones recibidas por las publicaciones del proyecto, es de 3,4. Contamos con un total de 17 documentos publicados, de los cuales 12 científicos y 5 educativos.

Desde el 12 de diciembre de 2016 hasta el 10 de febrero de 2017 hemos recibido 130 visitas de 75 usuarios y se han realizado 29 descargas. Estos datos corresponden a 13 países entre los que se encuentran España, Italia, México y Albania. Las búsquedas más utilizadas para localizar el perfil del MEMOLA han sido *archaeology*, *medieval archaeology* y *landscape archaeology*.

Impacto

El trabajo mediante redes sociales analógicas y digitales ha producido un gran impacto a distintas escalas, así como el reconocimiento por parte de otras instituciones. Hemos realizado un estudio a través del Gabinete de Comunicación de la Universidad de Granada y la Oficina para la Gestión de la Comunicación Corporativa, quienes han calculado que la noticia sobre la localización de la mezquita rural en la III Campaña de excavación de Lanteira, acogida por medios provinciales y nacionales, tuvo una audiencia de 5 millones de personas, que correspondería con una inversión potencial en publicidad de casi 50.000 €. Este valor se calculó en función de 24 artículos de prensa y 42 posts en redes sociales, sin embargo el alcance fue aún mayor habiendo registrado 36 artículos y 768 posts. Todas estas cifras corresponden únicamente al hallazgo de la mezquita (figura 6), no a la repercusión total que tuvo en los medios de comunicación la campaña de excavación del 2016.



Figura 6. Uno de los *posts* publicados en redes sociales sobre el hallazgo de la mezquita de Lanteira. Fuente: Twitter.

A nivel de impacto institucional, en el 2015 la recuperación de la acequia de Barjas (Cáñar) fue presentada por la Fundación Cultura de Paz a los premios *Best Practices Award* "Agua para la vida" de la ONU. Nuestra candidatura estaba dentro de la categoría de prácticas de participación pública, educativas, de comunicación y/o sensibilización. Esta misma intervención fue presentada por la asociación cultural Abuxarra a los Premios Hispania Nostra en la categoría Buenas Prácticas en la Conservación del patrimonio cultural y Natural 2015, recibiendo el premio Mención Especial a la intervención en el territorio.

Finalmente, en enero 2016 el Comité Científico Europeo de Humanidades ha seleccionado el MEMOLA como uno de los cinco proyectos más innovadores en investigación en humanidades³⁵, en otras cosas por las fórmulas innovadoras en el desarrollo y organización de los métodos de trabajo, destacando el impacto a nivel de desarrollo local, educación, rutas culturales, documentales y exposiciones.

3. Conclusiones

Muy probablemente el lector de estas páginas esté poco interesado a las temáticas en sí de los cementerios históricos, del exilio republicano español o del estudio de los paisajes históricos de montaña, pero a través de temas tan dispares hemos querido reflejar el impacto y los efectos que ha supuesto abrir la ciencia a un público no especializado.

Las metodologías de participación y co-creación están viviendo un auge y su uso o requerimiento se ha extendido en diferentes esferas, desde el ámbito académico a las convocatorias para fondos europeos o nacionales hasta, incluso, la legislación y las normativas. Sin embargo, su ejecución real ha sido muy limitada, reduciéndose en numerosos casos a una declaración de intenciones o concretándose únicamente en la publicación en acceso abierto de los resultados de la investigación. Involucrar y trabajar desde un plano horizontal con la sociedad requiere tiempo y constancia y los casos que aquí se han presentado cuentan con un desarrollo de entre 4 y 7 años. A pesar de las complicaciones de realizar un trabajo más abierto y participativo, implicar la sociedad ha sentado muy bien y el balance es ampliamente positivo tanto para los investigadores que han participado en estas iniciativas, como para los resultados que de ahí han surgido.

Los tres ejemplos comparten una metodología participativa que se realiza a través del aprovechamiento y la construcción de redes sociales digitales y analógicas que sirven como vehículos a través de los cuales se incorporan al proyecto personas, conocimientos, relaciones y memorias. En Historic Graves las comunidades locales han tomado las riendas de la iniciativa, convirtiendo los gestores de la plataforma en los socios de apoyo para la publicación de nuevos cementerios. En MEMOLA las comunidades de regantes históricas han proporcionado el substrato de conocimientos que ha enriquecido y valorizado las actividades en el campo. Con e-xiliad@s los usuarios han

comprendido el potencial de la participación para dar nombre al exiliado anónimo, recuperando historias de vida de una forma muy social y en horizontal, generando, ellos mismos, la construcción de un espacio digital lleno de contenido en donde sentirse identificados.

En las ciencias humanas, muchas veces los ciudadanos de a pie son los que se encuentran más cerca de las fuentes históricas primarias y llegan incluso a tener un potencial informativo más allá de los datos que puedan ofrecer directamente. Eso los convierte en agentes privilegiados para proyectos que se focalizan en recopilar, organizar, analizar y compartir estas fuentes.

Finalmente queremos proporcionar cuales han sido las claves de éxito de estos proyectos tan sociales y científicos a la vez, para que puedan convertirse en unas herramientas más para los profesionales, los investigadores y las instituciones que quieran impulsar la participación ciudadana en sus actividades. De esta manera, en el proyecto MEMOLA la participación tuvo éxito gracias al vínculo histórico, simbólico y emocional de la comunidad de regantes con el bien patrimonial objeto de estudio y se pudo mantener vivo el interés, a medio largo plazo, a través de la organización en paralelo de otras actividades que promovían el empoderamiento de estos colectivos y la socialización del patrimonio. Ejemplos de estas actividades son las acciones educativas, los diálogos de saberes y las rutas culturales llevadas a cabo, que además repercuten positivamente en el contexto social. En Historic Graves la accesibilidad de la metodología y la inmediatez en la publicación del trabajo realizado, ya sea en el campo que a través de la plataforma, están siendo claramente los motores de la involucración ciudadana. Las comunidades locales no encuentran barreras al acceso, son libres de colaborar sin restricciones prácticamente en todo el contenido publicado y sólo tienen que contar con la intermediación de los gestores de la plataforma para controlar el respeto de la metodología y la calidad de los datos a la hora de publicar nuevas fichas. En el proyecto e-xiliad@s la facilidad de enviar datos por el usuario y la rápida publicación de las fichas es lo que está marcando el ritmo creciente de la participación. Asimismo, un aspecto que ha favorecido la colaboración ha sido el otorgar al usuario la capacidad de decidir si hacer públicos o no los datos que aporta. Lejos de limitar el uso de las fuentes recibidas, este control ha agilizado la participación hasta el punto que casi la totalidad de los usuarios ha optado por la publicación en abierto de las fichas y documentos enviados. Significa esto que la participación tiene también una relación directa con la capacidad de dejar al usuario ser dueño de sus historias.

En definitiva, estos tres proyectos conectan ciencia y sociedad, entendiéndose como tal no solamente a los ciudadanos sino también a colectivos determinados, a instituciones y políticos, comprometiéndolos en una medida u otra en el desarrollo de la investigación y en la realización de actividades. Sí, a esto la Comisión Europea le ha puesto un nombre: *Science with a for Society*.

A nuestro juicio, el camino emprendido hacia modelos de ciencia abierta, participativa e inclusiva marca nuevas posibilidades en la investigación. Llegará a tener consecuencias metodológicas y conceptuales en las disciplinas y en el uso de las fuentes primarias, a revertir o suavizar el concepto de autoría y a valorar los aspectos locales dentro de herramientas de investigación y de difusión siempre más globales.

4. Notas

¹ <https://www.linkedin.com/in/john-tierney-6a609114>.

² <https://www.linkedin.com/in/mauriziotoscano/>.

³ www.eachtra.ie.

⁴ Datos actualizados al 12 de febrero 2017.

⁵ El *Archaeological Survey of Ireland* estima unos 3.433 cementerios en la República de Irlanda (últimos datos disponibles con fecha 25/11/2016 en <https://data.gov.ie/dataset/national-monuments-service-archaeological-survey-of-ireland>).

⁶ El uso de la luz lateral y del papel de calco se han revelado en los años entre las técnicas más exitosas en descifrar las inscripciones antiguas sin afectar el grabado.

⁷ Para una visión global del alcance del proyecto consultar el mapa en <http://historicgraves.com/graveyards-map>.

⁸ Ejemplo de proyecto acerca de diez cementerios en el condado de Roscommon <http://historicgraves.com/project/roscommon-historic-graveyard-project>.

⁹ Una de las muchas fichas fruto de un trabajo colaborativo <http://historicgraves.com/st-filians-roman-catholic/du-sfrc-0042/grave>.

¹⁰ CAPTCHA es la sigla de *Completely Automated Public Turing test to tell Computers and Humans Apart*, es decir un sistema automático para diferenciar entre ordenadores y humanos.

¹¹ Academia.edu: <https://ugr.academia.edu/LidiaBocanegra>; LinkedIn: <https://www.linkedin.com/in/lidiabocanegra>.

¹² El exilio republicano español se produce ya en 1936 a medida que el ejército sublevado, causante de la Guerra Civil española y al mando del General Franco, va ganando terreno. No será hasta enero de 1939, justo antes de la victoria franquista, cuando se produjo el exilio masivo de cerca de medio millón de personas quienes cruzaron la frontera con Francia. Para saber más acerca de este exilio véase: 1939, El éxodo republicano, Proyecto e-xiliad@s. Disponible en: <http://exiliadosrepublicanos.info/es/historia-exilio> (último acceso, 11 de febrero de 2017).

¹³ A partir de 1939, año en que se produjo la diáspora mayoritaria, se reconocen los siguientes principales lugares de destino iniciales: Francia, México, Rusia, Chile, República Dominicana, entre otros.

¹⁴ La base de datos que almacena la información se desarrolló en MySQL mediante el Sistema de Gestión de Contenidos de código abierto Drupal.

¹⁵ <https://www.facebook.com/exiliados.republicanos/>.

¹⁶ @e-xiliadas: <https://twitter.com/exiliadas>.

¹⁷ Las fichas enviadas, que cuenten con el consentimiento previo del usuario, se publican en la sección Biografías del proyecto: <http://www.exiliadosrepublicanos.info/es/biografias>.

¹⁸ Las publicaciones científicas que refieran al proyecto e-xiliad@s, o a análisis científicos realizadas a partir de informaciones específicas soscadas de las fichas, se pueden consultar públicamente a través de la sección Actividades del proyecto, así como en academia.edu: <https://ugr.academia.edu/LidiaBocanegra>.

¹⁹ <http://www.exiliadosrepublicanos.info/es/privacidad>.

²⁰ <http://www.exiliadosrepublicanos.info/es/terminos>.

²¹ <http://www.exiliadosrepublicanos.info/es/creditos>.

²² <http://www.exiliadosrepublicanos.info/es/tablon-anuncios>.

²³ <http://exiliadosrepublicanos.info/es/rabelesius/4494>.

²⁴ Para ver ejemplos de fichas completas véase, por ejemplo, la del exiliado Luis Sáez González: <http://exiliados-republicanos.info/es/content/luis-saez-gonzalez-saturnino-sanchez-gonzalez-francia-10-anos-menos> (último acceso 13/02/2017).

²⁵ La investigación que ha dado lugar a estos resultados ha recibido financiación del Séptimo Programa Marco de la Unión Europea (FP7/2014-2017) en virtud del acuerdo de subvención no613265. Proyecto MEMOLA: “Mediterranean Mountainous Landscapes: an historical approach to cultural landscapes based on traditional agrosystems” www.memolaproject.eu.

²⁶ <http://memolaproject.eu/node/283>.

²⁷ <https://www.facebook.com/MEMOLA.PROJECT/>.

²⁸ goo.gl/D7NacV.

²⁹ <https://twitter.com/MEMOLAProject>.

³⁰ <https://www.youtube.com/user/MEMOLAFP7PROJECT>.

³¹ <http://www.memolaproject.eu/activities/documentaries>.

³² <https://www.instagram.com/memolaproject/>.

³³ <https://sketchfab.com/memolaproject>. El gestor de la cuenta y autor de los modelos tridimensionales es Pablo Romero Pellitero, investigador del proyecto MEMOLA.

³⁴ <https://independent.academia.edu/MEMOLAProject>.

³⁵ goo.gl/m62ILV.

5. Bibliografía

- Arnstein, S. (1969). "A ladder of participation in the USA". En: *Journal of the American Institute of planners*, vol. 35, no 4. p. 216-224.
- Bocanegra Barbecho, L.; Toscano, M. (2015). "El exilio republicano español: estudio y recuperación de la memoria a través de la web 2.0. Nuevo enfoque metodológico con el proyecto e-xiliad@s". En: *Migraciones; Exilios*, no 15, Madrid: AEMIC. p.113-136.
- Bocanegra Barbecho, L.; Toscano, M. (2016). "The Spanish Republican Exile: Identity, Belonging and Memory in the Digital World". En: BOROWIECKI, C.J.; FORBES, N., Fresa, A. (ed.) *Cultural Heritage in a Changing World*, Springer. p. 237-253. http://link.springer.com/chapter/10.1007/978-3-319-29544-2_14/fulltext.html
- Bonney, R. (July 2009). *Public Participation in Scientific Research: Defining the Field and Assessing Its Potential for Informal Science Education*. A CAISE Inquiry Group Report. Washington, D.C.: Center for Advancement of Informal Science Education (CAISE). Disponible en <http://www.birds.cornell.edu/citscitoolkit/publications/CAISE-PPSR-report-2009.pdf> (último acceso el 30 de abril 2017).
- Calvo Hernando, M. (2001). *Divulgación y periodismo científico: entre la claridad y la exactitud*. UNAM.
- Delgado Anés, L. (2016) "Arqueología y Educación: acercando los paisajes culturales a los centros escolares". En: COELHO, I. P.; TORRES, J. B.; GIL, L. S.; RAMOS, T. (coord.). *Entre ciência e cultura: Da interdisciplinaridade à transversalidade da arqueologia. Actas das VIII Jornadas de Jovens em Investigação Arqueológica*. Lisboa: CHAM-FCSH/UNL- UAç y IEM-FCSH/UNL.
- Dickson, David. (2001). "Science the press and the public: from enlightenment to empowerment. 6th International Conference on Public Communication of Science and Technology. European Laboratory for Particle Physics". En: *6th International Conference on Public Communication of Science and Technology*. Ginebra.
- Fernandez Burgueño, P. (2009). *Clasificación de redes sociales*. Blog de Derecho. <http://www.pabloburgueno.com/2009/03/clasificacion-de-redes-sociales/>
- Gutiérrez Martínez, F. (2010). "Las redes sociales: una aproximación conceptual y una reflexión teórica sobre los posibles usos y gratificaciones de este fenómeno digital de masas". En: *TecCom Studies. Estudios de Tecnología y Comunicación*, no 1. Disponible en: <http://www.teccomstudies.com/index.php?journal=teccomstudies&page=article&op=view&path%5B%5D=136&path%5B%5D=117>
- Jirón, P. (2015). "De la participación a la co-creación. Nuevas formas de pensar intervenciones para mejorar el habitar urbano/residencial". En: *Materia Arquitectura*, no 12, Santiago: Universidad San Sebastián. p. 66-75.
- Lewenstein, B. V. (2003). "Models of public communication of science and technology". En: *Public Understanding of Science; Technology*, June. p. 1-11.
- Martín Civantos, J. M^a. (2015). "Landscape archaeology and rural communities: ethnoecology and social involvement". En CHAVARRIA ARNAU, A.;

REYNOLDS, A. (Eds) *Detecting and Understanding Historic Landscapes*. Padova: PCA. p.335-358.

- Nevens, F.; Frantzeskaki, N.; Gorissen, L.; Loorbach, D. (2013). "Urban Transition Labs: co-creating transformative action for sustainable cities". En: *Journal of Cleaner Production*, no 50 (1). p. 111-122
- O'Brien, C. (2010). *Guidance for the Care, Conservation and Recording of Historic Graveyards*. The Heritage Council of Ireland.
- Pasquali, M. (1977). *Comprender la comunicación*. Caracas: Monte Avila.
- Proyecto MEMOLA (2014). "Mediterranean Mountainous Landscapes: an historical approach to cultural heritage based on traditional agrosystems". En: *Agua y Territorio*, nº4, Jaén: Universidad de Jaén. p.146-148.
- Ponce, I. (2012). "Clasificación de redes sociales", En: Monográfico: *Redes Sociales*, Observatorio Tecnológico, Madrid: Ministerio de Educación, Cultura y Deporte. p. 4-10.
- Thomas, S. (2004). "What is Participatory Learning and Action (PLA): An Introduction". *Centre for International Development and Training, London: CIDT*
- Unesco (1999). *Declaración de Santo Domingo. La Ciencia para el siglo XXI: una nueva visión y un marco de acción*. Santo Domingo. Disponible en: <http://www.oei.es/historico/salactsi/santodomingo.htm>
- Weissbluth, M.; Contreras, E.; Galaz, P.; Aguilera, I.; Inostroza, J.; Infante, A.; Gatica, M. (2014). "Co-creación para la innovación: un caso en el sector público chileno". En: *Revista Ingeniería de Sistemas*, no XXVIII. p. 5-26

IV.4.3. Análisis de estrategias de involucración pública

En el contexto del análisis de la participación pública habilitada a través de los sistemas de información web, que ha sido el objeto de estudio de las dos precedentes publicaciones, hemos desarrollado un modelo de análisis de estrategias de involucración pública que se configura como otro resultado del presente trabajo de investigación. Además, como hemos visto anteriormente (véase apartado II.2), en la fase de determinación de los requisitos de un SIW, para la definición de las funcionalidades lado usuarios y la configuración de la interfaz de interacción con el sistema, las estrategias de involucración pública, cuando existen, juegan un papel relevante.

Para que un proyecto participativo sea eficaz es necesario diseñar y llevar a cabo una, o más, estrategias de involucración pública¹⁴⁵ (Rowe y Frewer 2005; Little y Shackel 2014; Bonacini 2018). Es a través de ellas que se favorece y se garantiza la participación del público, componente esencial para la implementación de un enfoque participativo. Las estrategias de involucración pública no siempre están vinculadas a las actividades participativas y pueden aplicarse con diferentes propósitos, pero un proyecto centrado en la colaboración con la sociedad requiere de un plan orientado a cómo atraer la ciudadanía y captar su interés. En este sentido, se puede hacer una clara distinción entre los enfoques participativos que implican la intervención y el empoderamiento de los interesados (como el *crowdsourcing* que hemos analizado en el anterior apartado IV.4.2), y las estrategias puestas en marcha para atraer e involucrar el público.

Aunque lograr la involucración voluntaria de personas ajenas al proyecto, a nivel individual o de comunidades/grupos, es un arte más que una ciencia, el modelo lógico que describimos a continuación permite analizar las principales características de las estrategias de involucración pública. Se trata de un modelo útil ya sea como una primera aproximación hacia una clasificación sistemática, que como herramienta para inspirar y estimular investigadores y profesionales que trabajan en el sector de la Ciencia ciudadana para el diseño de sus propias estrategias.

El modelo se articula en torno a cuatro de los principales **componentes** de un proyecto de investigación participativo: 1) el dominio de investigación; 2) el papel del organizador; 3) el público objetivo y 4) las contribuciones resultantes del proceso de participación. Cada uno de estos componentes se puede clasificar según una correspondiente pareja de **calificadores**: endógeno/exógeno; activo/pasivo; dirigido/no-dirigido; mediado/no-mediado (figura 11). Finalmente, esos calificadores se pueden conectar y producir así una imagen esquemática de cada estrategia, útil para comparar casos y buscar patrones. El modelo puede aplicarse ya sea a actividades existentes, por ejemplo para identificar cadenas comparables, que utilizarse como marco de discusión y planificación para evaluar nuevas estrategias o buscar alternativas.

¹⁴⁵ PES por el acrónimo inglés de *public engagement strategies*.



Figura 11. Esquema de componentes y calificadores del modelo de valoración de estrategias de involucración pública. Fuente: elaboración propia.

Pasamos entonces a analizar en detalle cada elemento. El **primer componente** a evaluar es el dominio, o ámbito de aplicación, del proyecto. Si la estrategia de involucración pública se origina en el contexto temático del propio proyecto, utilizando recursos de su propio objeto de estudio, puede definirse endógena. Por otra parte, una estrategia exógena propicia la participación apoyándose en temáticas externas.

El **segundo componente** describe el papel del organizador en la realización de actividades de fomento de la participación pública: la estrategia es activa si la presencia física y la intervención directa del promotor son necesarias para que la misma se pueda ejecutar. De lo contrario, si la estrategia en sí misma puede funcionar autónomamente una vez establecida, y sólo requiera de ciertos ajustes ocasionales, se define pasiva.

El **tercer elemento** evalúa los destinatarios, las audiencias. Si el público al que se dirige la iniciativa es específico, por ejemplo, se trata de partes interesadas (*stakeholders*) que mantienen vínculos directos con el objeto de estudio abarcado por el proyecto, o si se trata de un público seleccionado en función de la edad, como suele ocurrir en las actividades educativas, la estrategia se define dirigida. En cambio, si la acción está orientada hacia el público en general, y para los organizadores no es relevante quién participará concretamente, entonces se clasifica cómo no dirigida.

El **cuarto y último componente** se centra en el papel que juegan los participantes y, en particular, en el tratamiento de los resultados de su contribución, ya que consideramos esos resultados como parte integrante de la estrategia de involucración del público. Si las contribuciones se utilizan o se incorporan de forma directa, sin pasar por un proceso de filtrado y de revisión, la estrategia se define no mediada. De lo contrario, si las contribuciones se aprovechan sólo después de un proceso de verificación, o la estrategia no prevé que se integren en absoluto al proyecto, el planteamiento se define mediado. Este último componente podría considerarse, a primera vista, como una característica del enfoque participativo más que de la estrategia de involucración pública,

pero la experiencia demuestra que actuar un proceso mediado frente a un proceso no mediado, en el procesamiento de las contribuciones, conlleva consecuencias directas en la participación, por lo que consideramos relevante incorporar ese elemento al modelo.

Ninguno de estos ocho atributos implica un juicio de valor per sé, ya sea positivo o negativo. Como se verá claramente a continuación, según el proyecto en sí, una estrategia pasiva puede ser muy fructífera a largo plazo, un enfoque exógeno puede ser sumamente eficaz y un flujo de trabajo mediado puede ser la única opción disponible.

A título de ejemplo, presentamos aquí brevemente la aplicación de este modelo sobre algunos casos de estudios que ya se han analizado en el contexto de la presente tesis doctoral. Conocer un proyecto en detalle es relevante para poder evaluar las características de la estrategia de involucración ciudadana, ya que, con bastante frecuencia, esas estrategias no se presentan ni se describen como tales y el investigador tiene que seguir un proceso inductivo para determinar las técnicas utilizadas para incrementar el nivel de participación en una actividad.

El proyecto *Historic Graves* (véase apartado IV.4.2) ha aplicado, a lo largo de una década (2010-2020), dos principales estrategias para involucrar las comunidades locales en la actividad de prospección y aumentar la base de usuarios en su plataforma web. El primer enfoque (figura 12) consiste en la participación directa de las comunidades locales (dirigido) durante el trabajo de campo (activo), utilizando el vínculo de conexión familiar con el cementerio local (endógeno). El resultado de la participación se publica directamente en la plataforma en línea (no mediado). Al clicar "Save", las contribuciones de cada usuario en términos de transcripción de imágenes o enriquecimiento de metadatos se incorporan inmediatamente al conjunto de datos y se publican en abierto en la Web. Esta inmediatez de resultados, sin un filtro previo de verificación, ha tenido un fuerte impacto en la participación del público, ya que los colaboradores pueden ver directamente su aportación y compartirla con amigos y familiares, especialmente los que viven en el extranjero. La verificación de la calidad se lleva a cabo a posteriori (véase apartado IV.4.2), y es en buena medida distribuida, es decir a cargo de la comunidad de usuarios que pueden verificar y modificar cada ficha, como en una wiki.

La segunda estrategia (figura 13) utiliza la formación (activo) en tecnologías para la prospección (exógeno) como factor de atracción para captar miembros de la comunidad local que llevan mucho tiempo desempleados (dirigido). En este caso, el resultado de la participación son las fotografías geolocalizadas de cada lápida, que pasan sucesivamente por un proceso de tratamiento y verificación de elementos como la rotación, el nombre del archivo (que genera la clave primaria en el SIW) y la geolocalización, antes de ser subidas en bloque en la plataforma (mediado). Esas verificaciones no se pueden realizar a posteriori ni externalizar.



Figura 12. Estrategia de involucración pública aplicada por el proyecto “Historic Graves”. Fuente: elaboración propia.



Figura 13. Estrategia de involucración pública aplicada por el proyecto “Historic Graves”. Fuente: elaboración propia.

Este segundo caso analizado utiliza un factor de impulso exógeno, pero en general las estrategias endógenas son mucho más comunes y representan la gran mayoría de los 65 casos examinados hasta la fecha, siendo el 81% del total.

Un ejemplo en este sentido es el caso de la plataforma web **Colabora** (véase apartados IV.2.1 y IV.4.1). El enfoque participativo del proyecto (figura 14) prevé, en ese caso, aprovechar de la participación ciudadana para ampliar el catálogo existente sin que los investigadores del grupo tengan que desplazarse hasta el lugar de conservación de las obras de arte, como se venía haciendo antes del desarrollo del SIW. Inicialmente, la idea era publicar en línea un simple formulario de recogida de datos, pero el planteamiento de un sistema de información web abrió la puerta a la publicación en abierto del catálogo completo de fichas recopiladas hasta la fecha¹⁴⁶, que se transformó en la base para diseñar e implementar la estrategia de involucración de la ciudadanía (endógena). En este sentido, el catálogo constituye un recurso en dos direcciones. Por un

¹⁴⁶ En 2017, se trataba de alrededor de 2600 fichas de obras de arte; actualmente (31/12/2021) la base de datos cuenta con 3914 fichas registradas.

lado, una vez indexados sus metadatos por parte de los motores de búsqueda, aumenta la visibilidad SEO¹⁴⁷ de la propia plataforma web, en respuesta a la búsqueda de determinadas obras o artistas. Por otro lado, proporciona a los usuarios que llegan a la página una idea clara del conjunto de datos al cual sus aportaciones serán finalmente incorporadas, y que ellos mismos pueden explotar por sus propios fines de investigación o consulta, propiciando así la participación. Además, en términos de optimización del flujo de trabajo, ofrecer la libre consulta de las fichas colabora en reducir las posibles duplicidades de datos ya existentes. La participación ocurre en línea, de manera autónoma (pasiva), y está abierta a cualquier usuario interesado (no dirigida). Las aportaciones son revisadas por parte del equipo de investigación antes de su incorporación al conjunto público de datos (mediada). Valorando a posteriori el nivel de participación, que ha sido inferior a las expectativas, se ha determinado el elemento no-dirigido como la pieza débil de la estrategia, y se ha planteado así identificar una serie de públicos objetivos a quién dirigir la promoción de la iniciativa, dejando invariables los restantes componentes de la estrategia.



Figura 14. Estrategia de involucración pública aplicada por el proyecto “Colabora con Andalucía y América”. Fuente: elaboración propia.

Comparando estrategias de involucración pública aplicadas a varios proyectos de investigación que utilizan el *crowdsourcing* a través de la Web, el enfoque más común ha resultado ser la combinación endógeno-dirigido-mediado, con ejemplos tanto activos como pasivos, logrando los primeros resultados aparentemente más efectivos. A título de ejemplo, el proyecto **e-xiliad@s** (véase apartado IV.4.2), para recolectar en línea testimonios y fuentes documentales de exiliados republicanos, ha aplicado una estrategia endógena-activa-dirigida-mediada (figura 15), basada en las redes sociales, Facebook y Twitter en particular. Se han creado, en estas redes, perfiles enfocados a la recuperación de la memoria del exilio, se han enlazado los mismos con otras cuentas populares sobre temáticas afines y se ha alimentado y mantenido una comunidad de interesados,

¹⁴⁷ La visibilidad SEO, o visibilidad orgánica, es un término que procede desde el marketing digital y que describe la capacidad de una página web para ser “visible” y alcanzar los primeros puestos en las páginas de resultados de los motores de búsqueda. Fuente: <https://seoalive.com/>.

testigos y familiares a los cuales se ha proporcionado asesoramiento histórico, noticias seleccionadas y recursos temáticos, con el fin de que un cierto número de usuarios se convirtieran en colaboradores activos de la plataforma web del proyecto.



Figura 15. Estrategia de involucración pública aplicada por el proyecto “e-xiliad@s”. Fuente: elaboración propia.

Se ha detectado que varios proyectos de *crowdsourcing*, a la hora de implementar estrategias de involucración pública pasivas, se apoyan en la gamificación. Como casos ejemplares podemos citar las mencionadas iniciativas Comunidad BNE y CrowdHeritage (véase apartado III.1), que utilizan tablas de clasificación e insignias para fomentar la competición entre voluntarios.

Más allá de los ejemplos aquí descritos, el modelo ha sido ya testado sobre un amplio abanico de proyectos participativos (Toscano *et al.* 2019) y, para comprobar su robustez, se está ulteriormente validando en el marco del proyecto de investigación “Co-Historia: análisis de la participación pública en la investigación histórica desde el ámbito de la Ciencia ciudadana”¹⁴⁸. Los resultados completos de este trabajo están pendientes de publicación.

Como hemos visto, en un número significativo de casos, la participación ocurre en línea, o se compagina con algún tipo de actividad a través de la plataforma web, así que definir las estrategias de involucración de la ciudadanía se confirma como uno de los elementos a considerar en el proceso de desarrollo de un sistema de información web, ya sea en las fases de diseño de la estructura y definición de las funcionalidades, que en la fase continua de mantenimiento perfectivo (véase apartado III.2).

¹⁴⁸ Ref. E-HUM-507-UGR18, Programa Operativo FEDER 2018, Universidad de Granada.

V. CONCLUSIONES

Cuando hablamos de humanidades hablamos de un universo muy variado, hecho de disciplinas académicas que trabajan con enfoques, fuentes y tradiciones de estudios muy diferentes; así que para hablar de Humanidades Digitales es necesario definir desde qué **perspectiva** lo hacemos, ya que no hay y no puede haber una visión única¹⁴⁹. En esta tesis doctoral lo hemos hecho, por un lado, desde las problemáticas de investigación propias de la Historia, de la Historia del Arte y de la Arqueología; por el otro, hemos mirado a las tecnologías de la información y de la comunicación desde la perspectiva de los diseñadores y hacedores de herramientas digitales, tomando los sistemas de información web como objeto de la investigación y punto privilegiado de observación de los procesos de digitalización de la actividad investigadora humanística.

En los últimos quince años, la progresiva consolidación, a nivel global, de las Humanidades Digitales ha producido al menos dos tendencias, ligadas entre sí, pero a la vez claramente diferenciadas. Por un lado, la incorporación de herramientas digitales y la aplicación de técnicas cuantitativas en el campo tradicional de la investigación humanística (**enfoque instrumental**); por el otro lado, las Humanidades Digitales como disciplina de estudio, lugar de producción de nuevas preguntas de investigación, generadas a través de la integración de enfoques computacionales y de la hibridación multidisciplinaria, con consecuencias más propiamente epistemológicas (**enfoque metodológico**) (Svensson 2010; Parry 2012).

La **primera tendencia** mira a la normalización en el uso de herramientas, procedimientos y tratamientos de análisis digitales en la investigación humanística. Identifica las Humanidades Digitales como una disciplina de transición, con carácter en cierto modo generacional, destinada a desaparecer en cuanto la alfabetización digital y la adquisición de nuevos hábitos y prácticas se haya extendido a toda la comunidad investigadora. En este sentido, podemos decir que las Humanidades Digitales habrán logrado su objetivo cuando dejarán de existir y la integración de herramientas digitales y métodos computacionales se haya normalizado¹⁵⁰ en la actividad investigadora humanística. Su rol es funcional, pero no por eso menor o efímero. Parafraseando una afirmación de J. M. Lucía Megías en una reciente entrevista (del Río Riande 2018), podemos decir que en un futuro muy cercano las humanidades seguirán siendo tales solo si serán digitales.

En virtud de la **segunda tendencia**, las Humanidades Digitales podrán consolidarse a nivel académico como disciplina autónoma, esencialmente metodológica e

¹⁴⁹ Omitimos voluntariamente en esta sede el debate, ya muy maduro y prácticamente agotado, sobre la definición de las Humanidades Digitales.

¹⁵⁰ Normalización no en el sentido de ajustar a una norma un determinado proceso, sino que dichos procesos se integren en la normalidad de la práctica investigadora; algo que hasta ahora no se ha conseguido con la intensidad necesaria.

interdisciplinar¹⁵¹, aunque su carácter será cambiante en respuesta a las nuevas problemáticas generadas por el avance tecnológico (ej. visión por computador, inteligencia artificial, culturas algorítmicas, etc.). En este sentido, las Humanidades Digitales asumen la tecnología y el pensamiento computacional como el nuevo marco de generación del conocimiento, que obliga a una reflexión crítica del pensamiento humanístico, y viceversa. Juegan, en definitiva, un rol fundamental para la renovación epistemológica de las humanidades (Rodríguez Ortega 2018), y poseen un estatuto disciplinar propio.

Mirando a los **objetivos específicos** planteados al principio de la presente tesis doctoral, y a los resultados conseguidos (véase apartados IV.1-IV.4), podemos decir que los sistemas de información web desempeñan un papel relevante sobre ambas tendencias descritas anteriormente. Como **herramienta informática**, obligan a reconocer los datos como el verdadero material de trabajo de las humanidades, y colaboran a poner de manifiesto el valor de la información como central para el estudio del pasado. Desplazan en cierto modo el foco desde los objetos culturales en sí hacia la información acerca de los mismos (véase apartado III.1); información que los SIW permiten organizar, manipular, co-crear, mapear, visualizar, analizar, compartir y difundir. Su integración en la práctica investigadora se realiza principalmente a través de dos canales: el de los usuarios, que utilizan los SIW como herramienta de consulta y rescate de información; y el de los promotores, que plantean la creación de nuevos sistemas para organizar y explotar los datos de su propia investigación.

Desde la **perspectiva metodológica**, el proceso de formalización y operacionalización impuesto por la naturaleza misma de un sistema informático de gestión de la información, lejos de ser una mera cuestión técnica sin implicaciones intelectuales, se convierte en una oportunidad para realizar una reflexión ulterior sobre el objeto mismo de investigación y sobre las preguntas históricas a las que se quiere responder, poniendo en cuestión la metodología de investigación y el flujo de trabajo; es decir, el proceso mismo de producción del conocimiento. Desde el modelado conceptual hasta el desarrollo de estructuras de metadatos para la descripción codificada de los objetos culturales, pasando por la definición de taxonomías y ontologías y el desarrollo de sistemas de recuperación de la información, muchas de las fases del ciclo de vida de un SIW (véase apartado III.2) son epistémicas y pueden tener consecuencias a nivel metodológico y disciplinar. El uso de los SIW permite, y a la vez obliga, a comprobar el proceso de obtención, verificación y publicación de los datos (véase apartado IV.2.1); amplía la cantidad y calidad de fuentes utilizadas (véase apartado IV.2.2), produciendo así pues resultados más ricos de información; favorece el planteamiento de nuevas preguntas de investigación; ofrece nuevas opciones para difundir datos y resultados de investigación (véase apartado IV.3). En definitiva, como nos advertía Johanna Drucker ya en

¹⁵¹ Compartimos la tesis (Ramírez Sánchez 2016) según la cual el lugar más apropiado para la consolidación de las Humanidades Digitales en los grados de Humanidades sería el campo de las Ciencias y Técnicas Historiográficas.

2009 (Drucker 2009), hemos podido verificar que: “the design of digital tools for scholarship is an intellectual responsibility, not a technical task”.

Además, como hemos conseguido apreciar en los apartados IV.4.1 y IV.4.2, podemos considerar los SIW **plataformas habilitadoras** de nuevos procesos de investigación. Nos referimos en particular a aquellos procedimientos propedéuticos a la incorporación de actores sociales, normalmente ajenos al proceso de rescate de fuentes y enriquecimiento de corpus, o orientados a la colaboración intra- e inter-grupos a nivel global y distribuido, y a la transferencia de los resultados. En este sentido, hemos analizados casos de estudio en los que la presencia de un sistema de información web ha permitido ampliar el proceso de recogida y verificación de fuentes, otros en los que la publicación en línea de los datos existentes ha abierto la puerta a procesos de mapeo o transcripción colaborativa y, sucesivamente, al planteamiento de nuevas funcionalidades didácticas, a partir de proyectos eminentemente de investigación.

Asimismo, a través del análisis del escenario actual de las Humanidades Digitales en España (véase apartado IV.1), hemos podido comprobar que los artefactos digitales en general, y los SIW en particular, han vivido un auge en la última década y siguen ocupando un espacio relevante dentro de la práctica investigadora en humanidades. En la actualidad, los datos analizados nos muestran también cierta tendencia a la consolidación del uso de infraestructuras y repositorios de datos comunes, frente a soluciones personalizadas, en particular en disciplinas con una tradición más consolidada de estándares y protocolos compartidos tales como la Filología, la Lingüística, la Arqueología o el estudio del patrimonio cultural. Esta tendencia, merecedora de ulterior análisis en el futuro, llegará probablemente a cubrir ciertas necesidades, pero difícilmente podrá reemplazar las ventajas ofrecidas por sistemas diseñados acorde a problemáticas de investigación específicas. Otra tendencia hacia soluciones híbridas que hemos podido apreciar es la gradual aparición de los entornos virtuales de investigación (VRE – *Virtual Research Environment*) como soluciones informáticas para la generación, integración, difusión y conservación del conocimiento (véase apartado III.1). Podemos considerar los VRE una nueva tipología de sistemas de información caracterizados por un elevado grado de personalización en el diseño del modelo de datos y en la definición de los vocabularios/taxonomías, junto a una escasa o nula posibilidad de adaptación de las funcionalidades (ej. flujo de datos, opciones de búsquedas, vistas, etc.), que se consideran más generalizadas y aplicables *tout court* a diferentes ámbitos disciplinares y dominios de investigación. Una perspectiva interesante para el sector, en particular para proyectos de mayor envergadura, resultaría del **uso combinado de repositorios y entornos virtuales de investigación**: siendo los primeros el lugar adecuado para el almacenamiento a largo plazo y el intercambio de conjuntos de datos versionados y los segundos el espacio para el procesamiento dinámico y el enriquecimiento con metadatos, en un ambiente eminentemente colaborativo.

Desde otro punto de vista, el desarrollo de este trabajo de investigación nos ha permitido evidenciar una serie de **carencias** que afectan el campo de estudio objeto de esta tesis. En concreto, hemos identificado tres necesidades relevantes: 1) la formación de perfiles interdisciplinares, humanísticos e informáticos; 2) la normalización del uso de recursos y herramientas digitales y computacionales en todas las ramas de las humanidades; 3) el reconocimiento académico y curricular de la actividad ligada al desarrollo de herramientas tecnológicas para la investigación. En relación a estas necesidades detectadas, pasamos ahora a realizar una serie de propuestas y recomendaciones concretas para atajarlas.

Como hemos señalado en el apartado III.2, para el diseño de modelos conceptuales de sistemas de gestión de la información, así como para definir las funcionalidades de cualquier solución informática personalizada a medida sobre un tema de investigación determinado, es recomendable contar con perfiles híbridos, con formación en ambos campos, el humanístico y el informático; es decir **personas interdisciplinares**. Este tipo de formación se podría conseguir revisando y ampliando las titulaciones existentes, promoviendo dobles grados interdisciplinares, que combinen, por ejemplo, Humanidades con Informática, Comunicación o Información y Documentación; o bien fomentado la formación de posgrado a través de másteres complementarios¹⁵². Hemos visto en el apartado IV.1 que la oferta de cursos de posgrado en Humanidades Digitales ha vivido una enorme expansión en España en los dos últimos años. Por otro lado, en la última década, la oferta de doble grados y, en menor medida, de grados conjuntos se ha ido ampliando considerablemente (205 doble grados registrados en España en 2010/2011; 835 en 2018/2019¹⁵³; 1235 en 2020/2021). Según la información disponible en el buscador QEDU del portal del Ministerio de Educación, hay actualmente en España un total de 1235 carreras universitarias dobles¹⁵⁴, de las cuales 156 incluyen disciplinas

¹⁵² En una reciente encuesta sobre necesidades formativas técnicas en Humanidades, promovida por la Asociación de Humanidades Digitales Hispánicas, solo el 10% de los encuestados ha señalado haber adquirido sus competencias en Humanidades Digitales en estudios de grado, mientras que el 73% de los encuestados aluden al autoaprendizaje y el 48% a la formación en línea no reglada. El 54% se declara insatisfecho con su formación técnica. Finalmente, los encuestados señalan tener planes concretos para mejorar su formación en visualización de datos, análisis cuantitativo y cualitativo, programación y desarrollo de bases de datos.

¹⁵³ <https://www.lavanguardia.com/vida/20200809/482479972459/alerta-fiebre-dobles-grads-sustituye-calidad-cantidad.html>.

¹⁵⁴ https://www.educacion.gob.es/notasdecorte/busquedaSimple.action?codTipoEstudio=GRADOD&textTitula=&tipoAcceso=1&nomTipoAcceso=Universidad¬aCorte=&codPresencialidad=&codImpIdiomaExtranjero=&tipoUniv=T&multiselect_codigosUniversidades=&chkEspa%C3%B1a=&multiselect_codigosAut=&codigosProv=00&multiselect_codigosProv=&method%3AAbusquedaSimple=Buscar.

humanísticas¹⁵⁵. De esas, solo una, en la Universidad de Salamanca, ofrece un doble grado en Información y Documentación y en Historia y otra, en la Universidad privada San Pablo-CEU, en Humanidades y Comunicación Digital.

La segunda carencia detectada está ligada a la tendencia instrumental descrita al principio de este capítulo. Las herramientas digitales han llegado a todas las ramas de las humanidades, pero están aún lejos de llegar a todos los investigadores, que en muchos casos viven el universo digital como paralelo al de su propia disciplina, sin ninguna conexión. Por esta razón, mucha de la investigación en humanidades sigue haciéndose al margen de la gestión informatizada de los datos, perjudicando, de esta manera, la generación de nuevo conocimiento. Para normalizar el uso de herramientas digitales y métodos computacionales es necesario contar con **unidades técnicas**, o laboratorios, de Humanidades Digitales - sean ellos físicos, virtuales o mixtos - con una función de servicio en apoyo a la investigación y a la formación técnica y tecnológica. Como hemos visto en el apartado IV.1, en España resulta que únicamente hay cinco laboratorios de esta índole que, además, raramente responden a esa función de apoyo técnico que aquí recomendamos, sino que más bien se focalizan sobre sus propias actividades y líneas de actuación. Su implementación tendría que ser una línea estratégica a nivel institucional, siendo el lugar más coherente para su instalación las bibliotecas de las facultades de Filosofía y Letras o de Humanidades, según el caso. De hecho, las bibliotecas han sido por mucho tiempo los laboratorios de las facultades de Humanidades y tendrían que seguir siéndolo, asumiendo este rol de asesoramiento en la digitalización de la investigación; obviamente, a través de la incorporación de nuevo personal con formación específica. Sus principales funciones de servicio serían: 1) orientación en la elección de soluciones tecnológicas (métodos y técnicas) y en la adopción de buenas prácticas en actividades de investigación y docencia; 2) soporte técnico para la creación de sistemas de información y colecciones digitales, la codificación de metadatos, el diseño de recursos digitales, el análisis cuantitativo y de datos cualitativos, el manejo de datos espaciales, la digitalización de objetos y la creación de prototipos y pruebas de conceptos; 3) desarrollo y mantenimiento de entornos virtuales de investigación para la gestión o el análisis de datos; 4) asesoramiento en la incorporación de enfoques digitales en propuestas y proyectos, para aumentar su competitividad en la atracción de recursos; 5) formación técnica en Humanidades Digitales (en forma de talleres, cursos, prácticas, seminarios, formación de postgrado y del profesorado); 6) asistencia para el almacenamiento y la preservación de datos y proyectos archivados; 7) proporcionar un espacio de experimentación e innovación de nuevos procesos y tendencias. Este modelo no es nuevo; más bien es práctica habitual en los países anglosajones,

¹⁵⁵ https://www.educacion.gob.es/notasdecorte/busquedaSimple.action?codTipoEstudio=GRADODOD&textTitula=&tipoAcceso=1&nomTipoAcceso=Universidad¬aCorte=&codPresencialidad=&codImpIdiomaExtranjero=&codigosAmbitos=219&codigosAmbitos=228&codigosAmbitos=229&codigosAmbitos=209&tipoUniv=T&multiselect_codigosUniversidades=&chkEspana=C&multiselect_codigosAut=&codigosProv=00&multiselect_codigosProv=&method%3AAbusquedaSimple=Buscar.

donde destacan laboratorios de absoluta referencia internacional, junto a una multitud de pequeños centros que trabajan sobre todo a escala local. Entre los ejemplos más paradigmáticos podemos citar el *King's Digital Lab*¹⁵⁶ y el UCLDH¹⁵⁷ en Londres; así como el *Digital Scholarship Lab* de la *Michigan State University*¹⁵⁸, el *Center for Interdisciplinary Digital Research* en Stanford, el Dlab en Berkeley¹⁵⁹, el *Roy Rosenzweig Center for History and New Media*¹⁶⁰ afiliado a la *George Mason University*, el *Maryland Institute for Technology in the Humanities*¹⁶¹, respaldado conjuntamente por la Facultad de Artes y Humanidades y las Bibliotecas de la Universidad de Maryland, o la unidad de *Digital Humanities* de la *Carnegie Mellon University Libraries*¹⁶² en Pittsburg, todos ellos en Estados Unidos¹⁶³.

Finalmente, para concluir, consideramos consecuente remarcar la urgente **necesidad de reconocimiento** de la actividad investigadora basada en procedimientos y métodos digitales en el ámbito de las disciplinas humanísticas (véase apartado IV.1). Nos referimos, en particular y en consonancia con nuestro objeto de estudio, a actividades como: el diseño y la implementación de recursos digitales (sistemas de información, repositorios de datos u objetos culturales, catálogos cartográficos, repertorios iconográficos, etc.); la definición de estándares y protocolos, así como la creación de esquemas de metadatos y taxonomías/ontologías; el desarrollo de aplicaciones orientadas a la sociedad y a la transferencia de conocimiento. La evaluación de esta actividad investigadora es una cuestión problemática, ya que para otorgar un reconocimiento hay, por un lado, que definir un marco de valoración con unos determinados criterios a tener en cuenta y, por el otro, hay que producir documentación que permita valorar dichos criterios. En esa línea se ha expresado ya en 2015 la Asociación de Humanidades Digitales Hispánicas, publicando un documento¹⁶⁴ destinado principalmente a las agencias y a los comités de evaluación de la actividad investigadora y académica, pero desde entonces esta carencia se ha hecho sólo más apremiante.

Para el caso concreto de los sistemas de información web, la referencia obligada son los procedimientos ya existentes para la **certificación de repositorios digitales**. No es este el lugar para una revisión de las normativas de certificación existentes, pero

¹⁵⁶ <https://kdl.kcl.ac.uk/>.

¹⁵⁷ <https://www.ucl.ac.uk/digital-humanities/>.

¹⁵⁸ <https://digitalhumanities.msu.edu/digital-scholarship-lab/>.

¹⁵⁹ <https://dlab.berkeley.edu/consulting>.

¹⁶⁰ <https://rrchnm.org/>.

¹⁶¹ <https://mith.umd.edu/>.

¹⁶² <https://www.library.cmu.edu/research/digital-humanities>.

¹⁶³ Para una panorámica completa, aunque algo desactualizada, del modelo estadounidense, véase *A Survey of Digital Humanities Centers in the United States* (Zorich 2008).

¹⁶⁴ “Documento de recomendaciones para la evaluación y reconocimiento de la investigación llevada a cabo en el ámbito de las Humanidades Digitales”: <https://humanidadesdigitaleshispanicas.es/la-asociacion/documentos/>. Este documento se encuentra actualmente en fase de revisión por parte de una comisión de expertos, en vista de publicar, en 2021, una versión actualizada.

podemos limitarnos a decir que existen al menos tres niveles, con diferentes grados de complejidad (de Lorenzo-Cáceres y Bonal Zazo 2017). El nivel más exigente es ocupado por la norma UNE-ISO 16363:2017¹⁶⁵, que se basa en un proceso muy formal y costoso, que prevé más de un centenar de métricas a cumplir. A un nivel inferior de exigencias se encuentra la normativa *Nestor*¹⁶⁶ (nestor-Seal DIN 31644), que prevé el cumplimiento de treinta y cuatro requisitos y utiliza un proceso de certificación abierto, basado en una autoevaluación que después se somete a un proceso de revisión por pares; se trata de un protocolo que ha tenido un nivel de adopción muy bajo y básicamente limitado al territorio alemán. Basado sobre el mismo procedimiento de evaluación, pero con un conjunto más acotado de dieciséis requisitos¹⁶⁷ (CoreTrustSeal Standards and Certification Board 2019), se encuentra el *CoreTrustSeal*¹⁶⁸, que se está consolidando en los últimos dos años como el primer escalón para la certificación de repositorios digitales. Hay que tener en cuenta que todas estas normativas están pensadas para repositorios institucionales, orientados a la preservación de los datos y no a su manipulación con fines de investigación. A pesar de eso, el caso del *CoreTrustSeal* nos muestra que muchos de los requisitos son asumibles también por sistemas de información web para la investigación, como por ejemplo: los requisitos sobre la definición del tratamiento y enriquecimiento de los datos; la licencia de uso; el control de acceso; la infraestructura organizativa; los procedimientos de almacenamiento y verificación de la calidad de los datos; el flujo de trabajo; los indicadores de identificación unívoca; la infraestructura técnica; las políticas de seguridad. Además, el mismo proceso de autoevaluación es una oportunidad muy valiosa para verificar, a nivel interno, bondades y carencias de un SIW.

En particular para los sistemas abiertos, orientados hacia la sociedad o para la colaboración con otros grupos de investigación, contar con un sello de certificación constituiría un aliciente para que usuarios externos al núcleo promotor depositaran su confianza en aportar nuevos datos, así como en enriquecer y reutilizar los existentes. Además, el *CoreTrustSeal* está pensado para repositorios con un dominio de aplicación específico y una audiencia designada, como suele ser el caso de los sistemas de información personalizados. Sus especificaciones representan ya, al estado actual, una opción válida para la certificación de sistemas de información con vocación colaborativa y, a

¹⁶⁵ Equivalente a la norma ISO 16363:2012, <https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0058850>.

¹⁶⁶ *Network of Expertise in Long-term STORage of Digital Resources* (NESTOR).

¹⁶⁷ Organizados en cuatro categorías: Información de base (R0); Infraestructura organizativa (R1-6); Gestión de objetos digitales (R7-14); Tecnología y seguridad (R15-16).

¹⁶⁸ *CoreTrustSeal* es una organización internacional, de base asociativa, no gubernamental y sin ánimo de lucro, establecida en 2017 con sede en los Países Bajos (<https://www.coretrustseal.org>), bajo el paraguas de la *Research Data Alliance* (<https://www.rd-alliance.org/>).

la vez, constituyen una guía para la definición de nuevos protocolos más concretamente orientados hacia la investigación¹⁶⁹.

Otorgar reconocimiento académico al trabajo relacionado con la generación y publicación de conjuntos de datos de calidad, con sus relativos informes técnicos y código fuente, junto con la valoración de sus métricas de uso, representaría indudablemente un estímulo para la comunidad investigadora y condicionaría favorablemente la actitud hacia el intercambio de información y la ciencia abierta, tan poco consolidada en las disciplinas humanísticas.

¹⁶⁹ La misma *CoreTrustSeal*, en su página web, indica que tiene previsto proporcionar certificación de nivel básico también para herramientas de investigación, como los servicios de datos y el software (<https://www.coretrustseal.org/about/>).

VI. BIBLIOGRAFÍA

- Aibar Puentes, E. (2018). La transformación neoliberal de la ciencia: El caso de las Humanidades Digitales. *ArtefaCToS. Revista de estudios sobre la ciencia y la tecnología*, 7(1), pp. 13-28. <https://doi.org/10.14201/art2018711328>
- Albert, R. M., Ruíz, J. A., & Sans, A. (2016). PhytCore ODB: A new tool to improve efficiency in the management and exchange of information on phytoliths. *Journal of Archaeological Science*, 68, pp. 98-105. <https://doi.org/10.1016/j.jas.2015.10.014>
- Alcaraz Martínez, R. (2012). Omeka: Exposicions virtuals i distribució de col·leccions digitals. *BiD: Textos universitaris de biblioteconomia i documentació*, 28, pp. 1-13.
- Alcaraz Martínez, R. (2014). CollectiveAccess, un sistema de gestión y difusión de colecciones de museos, archivos y bibliotecas. *BiD: Textos Universitaris de Biblioteconomia i Documentació*, 33. <https://doi.org/10.1344/BiD2014.33.23>
- Aldeias, V., & Bicho, N. (2016). Embedded Behavior: Human Activities and the Construction of the Mesolithic Shellmound of Cabeço da Amoreira, Muge, Portugal: Shellmound EMBEDDED BEHAVIOR. *Geoarchaeology*, 31(6), pp. 530-549. <https://doi.org/10.1002/gea.21573>
- Aldeias, V., Goldberg, P., Sandgathe, D., Berna, F., Dibble, H. L., McPherron, S. P., Turq, A., & Rezek, Z. (2012). Evidence for Neandertal use of fire at Roc de Marsal (France). *Journal of Archaeological Science*, 39(7), pp. 2414-2423. <https://doi.org/10.1016/j.jas.2012.01.039>
- Almagià, R. (1960). *Documenti cartografici dello Stato pontificio editi dalla Biblioteca Apostolica Vaticana*. Biblioteca Apostolica Vaticana.
- Andrews, D. (1978). Medieval masonry in Northern Lazio: Its development and uses for dating. En H. Blake, T. W. Potter, & D. B. Whitehouse, *Papers in Italian Archaeology. The Lancaster Seminar* (pp. 391-412). Oxford.
- Angelucci, D. E., Boschian, G., Fontanals, M., Pedrotti, A., & Vergès, J. M. (2009). Shepherds and karst: The use of caves and rock-shelters in the Mediterranean region during the Neolithic. *World Archaeology*, 41(2), pp. 191-214. <https://doi.org/10.1080/00438240902843659>
- Antonelli, M. (1895). Una relazione del Vicario del Patrimonio a Giovanni XXII in Avignone. *Archivio della R. Società Romana di Storia Patria*, XVIII, pp. 447-467.
- Arnstein, S. (1969). A ladder of participation in the USA. *Journal of the American Institute of planners*, 35(4), pp. 216-224.

- Arpin, T. L., Mallol, C., & Goldberg, P. (2002). Short contribution: A new method of analyzing and documenting micromorphological thin sections using flatbed scanners: Applications in geoarchaeological studies. *Geoarchaeology*, 17(3), pp. 305-313. <https://doi.org/10.1002/gea.10014>
- Baraibar Echeverria, Á. B. (2014). *Humanidades digitales: Una aproximación transdisciplinaria*. SIELAE. <https://dialnet.unirioja.es/servlet/libro?codigo=660932>
- Batelli, G. (1946). *Rationes decimrum Italiae nei secoli XIII e XIV: Latium*. Città del Vaticano.
- Benavides López, J. A., Aranda Jiménez, G., Sánchez Romero, M., Alarcón García, E., Fernández Martín, S., Lozano Medina, A., & Esquivel Guerrero, J. A. (2016). 3D modelling in archaeology: The application of Structure from Motion methods to the study of the megalithic necropolis of Panoria (Granada, Spain). *Journal of Archaeological Science: Reports*, 10, pp. 495-506. <https://doi.org/10.1016/j.jasrep.2016.11.022>
- Berna, F., Behar, A., Shahack-Gross, R., Berg, J., Boaretto, E., Gilboa, A., Sharon, I., Shalev, S., Shilstein, S., Yahalom-Mack, N., Zorn, J. R., & Weiner, S. (2007). Sediments exposed to high temperatures: Reconstructing pyrotechnological processes in Late Bronze and Iron Age Strata at Tel Dor (Israel). *Journal of Archaeological Science*, 34(3), pp. 358-373. <https://doi.org/10.1016/j.jas.2006.05.011>
- Bertoldi, S., Fronza, V., & Valenti, M. (2015). Sistemi digitali di documentazione e analisi archeologica. Verso quale direzione? *Archeologia e Calcolatori*, 26, pp. 233-243.
- Bocanegra Barbecho, L. (2020). Ciencia ciudadana y memoria histórica: Nuevas perspectivas historiográficas desde las Humanidades Digitales y la Historia Pública Digital. En J. Caro Saiz, S. Díaz-de la Fuente, V. Ahedo, D. Zurro, M. Madella, J. M. Galán, L. Izquierdo, J. I. Santos, & R. Martínez, *Terra Incognita: Libro blanco sobre transdisciplinariedad y nuevas formas de investigación en el Sistema Español de Ciencia y Tecnología*. PressBooks. <https://doi.org/10.5281/zenodo.4034177>
- Bocanegra Barbecho, L., & Gea Megías, M. (2016). Las Humanidades Digitales y el aprendizaje en acceso abierto: El caso de la comunidad en Historia Digital. En M. Gea Megías, *Experiencias MOOC: un enfoque hacia el aprendizaje digital, la creación de contenidos docentes y comunidades on line*. Universidad de Granada.
- Bocanegra Barbecho, L., & Toscano, M. (2015). El exilio republicano español: Estudio y recuperación de la memoria a través de la web 2.0. Nuevo enfoque metodológico con el proyecto e-xiliad@s. *Migraciones & Exilios*, 15, pp. 113-136.

- Bocanegra Barbecho, L., Toscano, M., & Delgado Anés, L. (2017). Co-creación, participación y redes sociales para hacer historia. Ciencia con y para la sociedad. *Historia y Comunicación Social*, 22-2, pp. 325-346. <https://doi.org/10.5209/HICS.57847>
- Bolívar Galiano, F., Abad Ruiz, C., Sánchez Castillo, P., Toscano, M., & Romero Noguera, J. (2020). Frequent Microalgae in the Fountains of the Alhambra and Generalife: Identification and Creation of a Culture Collection. *Applied Sciences*, 10(18), pp. 6603. <https://doi.org/10.3390/app10186603>
- Bonacini, E. (2018). Heritage Communities, Participation and Co-creation of Cultural Values: The #iziTRAVELSicilia Project. *Museum International*, 70(1-2), pp. 140-153. <https://doi.org/10.1111/muse.12199>
- Bonney, R. (2009). *Public Participation in Scientific Research: Defining the Field and Assessing its Potential for Informal Science Education*. Inquiry Group Report. Washington, D.C.: Center for Advancement of Informal Science Education.
- Borromeo Mae, R. M., Laurent, T., & Toyama, M. (2016). The Influence of Crowd Type and Task Complexity on Crowdsourced Work Quality. *Proceedings of the 20th International Database Engineering & Applications Symposium on - IDEAS '16*, pp. 70-76. <https://doi.org/10.1145/2938503.2938511>
- Bree, P., & Kessels, G. (2013). *Nodegoat: A web-based data management, network analysis & visualisation environment*. <http://nodegoat.net>
- Brown, S. (2016). Towards best practices in collaborative online knowledge production. En C. Crompton, R. Lane, & R. Siemens, *Doing Digital Humanities. Practice, Training, Research*. (pp. 47-64). London: Routledge.
- Brunetti Nardi, G. (1972). *Repertorio degli scavi e delle scoperte archeologiche nell'Etruria Meridionale, II-III (1966-1970 e 1971-1975)*. Roma.
- Burdick, A., Drucker, J., Lunenfeld, P., Presner, T., & Schnapp, J. (2012). *Digital humanities*. Massachusetts Institute of Technology Press, Cambridge, Massachusetts.
- Buzzi, C. (1988). *Il «Catasto» di S. Stefano di Viterbo*. Presso la Società alla Biblioteca Vallicelliana.
- Buzzi, C. (1998). *Il Liber quatuor clavium del Comune di Viterbo*. Istituto storico italiano per il Medio Evo.
- Buzzi, C. (2004). *Lo statuto del comune di Viterbo del 1469*. Istituto Palazzo Borromini.
- Calvo Hernando, M. (2001). *Divulgación y periodismo científico*. UNAM.

- Cameron, F., Clark, G., Jackson, R. P. J., Johns, C. M., Philpot, S., Potter, T. W., Shepherd, J. D., Stone, M. J., & Whitehouse, D. B. (1984). Il castello di Ponte Nepesino e il confine settentrionale del Ducato di Roma. *Archeologia medievale*, 11, pp. 63-147.
- Campanari, S. (1856). *Tuscania e i suoi monumenti: Documenti storici*. Tip. del Seminario.
- Carbonetti Vendittelli, C. (1990). *Liber memorie omnium privilegiorum et instrumentorum et actorum communis Viterbii (1283)*. Roma.
- Carletti, L., Giannachi, G., Price, D., McAuley, D., & Benford, S. (2013). *Digital humanities and crowdsourcing: An exploration*. <https://ore.exeter.ac.uk/repository/handle/10871/17763>
- Carlisle, P. K., Avramides, I., Dalgity, A., & Myers, D. (2014). The Arches Heritage Inventory and Management System: A standards-based approach to the management of cultural heritage information. *CIDOC (International Committee for Documentation of the International Council of Museums)*, 8. <https://www.wmf.org/publication/arches-heritage-inventory-and-management-system-standards-based-approach-management>
- Carocci, S., & Vendittelli, M. (2004). *L'origine della Campagna Romana: Casali, castelli e villaggi nel XII e XIII secolo*. Società Romana di Storia Patria.
- Carpentier, F., & Vandermeulen, B. (2016). High-Resolution Photography for Soil Micromorphology Slide Documentation. *Geoarchaeology*, 31(6), pp. 603-607. <https://doi.org/10.1002/gea.21563>
- Carusi, E. (1948). *Cartulario di S. Maria in Campo Marzio, 986-1199*. Roma.
- Causer, T., & Wallace, V. (2012). Building A Volunteer Community: Results and Findings from Transcribe Bentham. *Digital Humanities Quarterly*, 6(2).
- Cerrillo-Cuenca, E. (2017). An approach to the automatic surveying of prehistoric barrows through LiDAR. *Quaternary International*, 435, pp. 135-145. <https://doi.org/10.1016/j.quaint.2015.12.099>
- Cerrillo-Cuenca, E., & Sepúlveda, M. (2015). An assessment of methods for the digital enhancement of rock paintings: The rock art from the precordillera of Arica (Chile) as a case study. *Journal of Archaeological Science*, 55, pp. 197-208. <https://doi.org/10.1016/j.jas.2015.01.006>
- Chen, P. P-S. (1976). The entity-relationship model: Toward a unified view of data. *ACM Transactions on Database Systems*, 1(1), pp. 9-36. <https://doi.org/10.1145/320434.320440>

- Chevalier, P., Granjon, L., Leclercq, É., Millereux, A., Savonnet, M., & Sapin, C. (2012). Base de Données annotées et Wiki pour la constitution du corpus numérique CARE. *Hortus Artium Medievalium*, 18(1), pp. 27-35. <https://doi.org/10.1484/J.HAM.1.102782>
- Chiovelli, R. (2007). *Tecniche costruttive murarie medievali: La Tuscia*. L'erma di Bretschneider.
- Christie, N. (1991). *Three south etrusian churches: Santa Cornelia, Santa Rufina and San Liberato*. British School at Rome.
- Coarelli, F., & Patterson, H. (2008). *Mercator placidissimus: The Tiber Valley in antiquity: new research in the upper and middle river valley: Rome, 27-28 February 2004*. Quasar.
- Conti, S. (1980). *Le sedi umane abbandonate nel Patrimonio di s. Pietro*. L.S. Olschki.
- CoreTrustSeal Standards and Certification Board. (2019). *CoreTrustSeal Trustworthy Data Repositories Requirements: Extended Guidance 2020–2022*. <https://doi.org/10.5281/ZENODO.3632533>
- Courty, M. A. (2001). Microfacies Analysis Assisting Archaeological Stratigraphy. En P. Goldberg, V. T. Holliday, & C. R. Ferring, *Earth Sciences and Archaeology* (pp. 205-239). Springer US. https://doi.org/10.1007/978-1-4615-1183-0_8
- Currás Puente, E. (2005). *Ontologías, taxonomía y tesauros: Manual de construcción y uso* (3a ed., actualizada y ampliada). Ediciones Trea.
- Dallas, C., Chatzidiakou, N., Benardou, A., Bender, M., Berra, A., Clivaz, C., Cunningham, J., Dabek, M., Garrido, P., Gonzalez-Blanco, E., Hadalin, J., Hughes, L., Immenhauser, B., Joly, A., Kelpšienė, I., Kozak, M., Kuzman, K., Lukin, M., Marinski, I., ... Zebec, T. (2017). *European Survey On Scholarly Practices And Digital Needs In The Arts And Humanities-Highlights Report*. Zenodo. <https://doi.org/10.5281/ZENODO.260101>
- De Minicis, E., ed. (2003-2011). *Insedimenti rupestri medievali della Tuscia, I-III*, Roma.
- De Reu, J., De Smedt, P., Herremans, D., Van Meirvenne, M., Laloo, P., & De Clercq, W. (2014). On introducing an image-based 3D reconstruction method in archaeological excavation practice. *Journal of Archaeological Science*, 41, pp. 251-262. <https://doi.org/10.1016/j.jas.2013.08.020>
- Declaración de Santo Domingo. La Ciencia para el siglo XXI: una nueva visión y un marco de acción*. Santo Domingo. (1999). UNESCO. <https://oei.int/oficinas/secretaria-general>

- Del Lungo, S. (1998). *Il territorio dell'antica diocesi di Orte nella toponomastica archeologica*. Orte.
- Del Lungo, S. (1999). *La toponomastica archeologica della provincia di Viterbo*. Consorzio di Bonifica della Maremma Etrusca.
- Del Rio Riande, G. (2018). En un futuro muy cercano las Humanidades serán Humanidades Digitales o no serán nada. Entrevista a José Manuel Lucía Megías. *Olivar*, 18(28), pp. e041. <https://doi.org/10.24215/18524478e041>
- Delgado Anés, L. (2016). Arqueología y Educación: Acercando los paisajes culturales a los centros escolares. En I. P. Coelho, J. B. Torres, L. S. Gil, & T. Ramos (Eds.), *Actas das VIII Jornadas de Jovens em Investigação Arqueológica*.
- de Lorenzo-Cáceres, M. del P. O., & Bonal Zazo, J. L. (2017). Criterios de certificación y auditoría de repositorios digitales seguros en archivos. En N. Vaquinhas, M. Caixas, & H. Vinagre, *Da produção à preservação informacional: Desafios e oportunidades* (pp. 529-550). Publicações do Cidehus. <https://doi.org/10.4000/books.cidehus.2835>
- Di Giacomo, G., & Scardozzi, G. (2014). GIS cloud per l'archeologia: Strumenti open source per la gestione e condivisione dei dati. *Archeologia e calcolatori*, 25, pp. 93-112.
- Dickson, D. (2001). *Science the press and the public: From enlightenment to empowerment*. 6th International Conference on Public Communication of Science and Technology. European Laboratory for Particle Physics. 6th International Conference on Public Communication of Science and Technology, Ginebra.
- Dombrowski, Q. (2016a). Drupal and other content management systems. En Constance Crompton, R. Lane, & R. Siemens, *Doing Digital Humanities. Practice, Training, Research*. (pp. 289-302). London: Routledge. <https://doi.org/10.4324/9781315707860-30>
- Dombrowski, Q. (2016b). *Drupal for humanists (Coding for Humanists)*. Texas A&M University Press. <http://www.mylibrary.com?id=969517>
- Drucker, J. (2009). Blind Spots. Humanist must plan their digital future. *The Chronicle Review*, 55(30).
- Duncan, G., & Reynolds, J. M. (1958). Sutri. *Papers of the British School at Rome*, 26, pp. 63-134.
- Earley-Spadoni, T. (2017). Spatial History, deep mapping and digital storytelling: Archaeology's future imagined through an engagement with the Digital

Humanities. *Journal of Archaeological Science*, 84, pp. 95-102.
<https://doi.org/10.1016/j.jas.2017.05.003>

Egidi, P. (1901). Le Cronache di Viterbo scritte da Francesco D'Andrea. *ASRSP*, XXIV, pp. 197-251, 299-371.

Egidi, P. (1906). L'archivio della Cattedrale di Viterbo. *Bollettino dell'Istituto storico italiano per il Medio Evo e Archivio Muratoriano*, 27, pp. 7-382.

Egidi, P. (1910). Statuti di Castel Fiorentino. En F. Tomassetti, V. Federici, & P. Egidi, *Statuti della Provincia Romana. Vicovaro, Cave, Roccantica, Ripi, Genazzano, Tivoli, Castel Fioretino*. (pp. 303-359). Roma.

Egidi, P. (1930). Gli Statuti Viterbesi del MCCXXXVII-VIII, MCCLI-II e MCCCLVI. En *Statuti della Provincia Romana. S. Andrea in Selci, Subiaco, Viterbo, Roviano, Anagni, Saccomuro, Aspra Sabina* (pp. 27-282). Roma.

Eitzel, M. V., Cappadonna, J. L., Santos-Lang, C., Duerr, R. E., Virapongse, A., West, S. E., Kyba, C. C. M., Bowser, A., Cooper, C. B., Sforzi, A., Metcalfe, A. N., Harris, E. S., Thiel, M., Haklay, M., Ponciano, L., Roche, J., Ceccaroni, L., Shilling, F. M., Dörler, D., ... Jiang, Q. (2017). Citizen Science Terminology Matters: Exploring Key Terms. *Citizen Science: Theory and Practice*, 2(1), pp. 1-20.
<https://doi.org/10.5334/cstp.96>

Endres, B. (2017). A literacy of building: Making in the digital humanities. En J. Sayers, *Making Things and Drawing Boundaries: Experiments in the Digital Humanities* (Vol. 3, pp. 44-54). University of Minnesota Press.
<https://doi.org/10.5749/9781452963778>

Estellés Arolas, E., & González Ladrón de Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), pp. 189-200.
<https://doi.org/10.1177/0165551512437638>

Fabre, P. (1887). Un registre Caméral du Cardinal Albornoz en 1364. *Mélanges d'archéologie et d'histoire*, 7(1), pp. 129-195. <https://doi.org/10.3406/mefr.1887.6506>

Fabre, P., & Duchesne, L. (1905). *Le Liber censuum de l'église romaine*. A. Fontemoing.

Fedele, P. (1898). Carte del monastero dei SS. Cosma e Damiano in Mica Aurea, sec. X e XI. *Archivio della R. Società Romana di Storia Patria*, XXI, pp. 459-534; XXII, pp. 25-107.

Fedele, P. (1900). Tabulariu Sanctae Mariae Novae ab anno 982 ad anno 1200. *Archivio della R. Società Romana di Storia Patria*, XXIII, pp. 171-237; XXIV, pp. 159-196; XXV, pp. 169-209; XXVI, pp. 21-141.

- Federici, V. (1899). Regesto del monastero di San Silvestro in Capite. *Archivio della R. Società Romana di Storia Patria*, XII, pp. 213-300; XXXIII, pp. 67-128.
- Fernández Alarcón, V. (2006). *Desarrollo de sistemas de información: Una metodología basada en el modelado*. Edicions UPC. <https://upcommons.upc.edu/handle/2099.3/36751>
- Fernández Burgueño, P. (2009). *Clasificación de redes sociales*. <https://www.pablofb.com/2009/03/clasificacion-de-redes-sociales/>
- Fiormonte, D. (2014). Digital Humanities from a global perspective. *Laboratorio dell'ISPF*, XI. <https://doi.org/10.12862/ispf14L203>
- Fisher, E. C., Akkaynak, D., Harris, J., Herries, A. I. R., Jacobs, Z., Karkanias, P., Marean, C. W., & McGrath, J. R. (2015). Technical considerations and methodology for creating high-resolution, color-corrected, and georectified photomosaics of stratigraphic sections at archaeological sites. *Journal of Archaeological Science*, 57, pp. 380-394. <https://doi.org/10.1016/j.jas.2015.02.022>
- Flügel, E. (2004). *Microfacies of Carbonate Rocks: Analysis, Interpretation and Application*. Springer.
- French, C. A. I. (2003). *Geoarchaeology in action: Studies in soil micromorphology and landscape evolution*. Routledge.
- Fresa, A., & Justrel, B. (2015). *Roadmap for Citizen Researchers in the Age of Digital Culture* (pp. 7-60). CIVIC EPISTEMOLOGIES FP7 Project.
- Fronti, D., & Ramagnoli, G. (2007). Indagini archeologiche nel castello di Pian Fasciano (Tuscania). Nota preliminare. *Corneto medievale: territorio, società, economia e istituzioni religiose. Atti del Convegno di studio (Tarquinia, 24-25 novembre 2007)*. Tarquinia, pp. 421-454.
- Frutaz, A. P. (1972). *Le carte del Lazio*. Istituto di Studi Romani.
- Fumi, L. (1884). *Codice diplomatico della città d'Orvieto; documenti e regesti dal secolo XI al XV, e la Carta del popolo: Codice statutario del comune di Orvieto*; Presso G.P. Vieusseux.
- Gallini, S., & Noiret, S. (2011). La historia digital en la era del Web 2.0. Introducción al dossier Historia digital. *Historia crítica*, 43, pp. 16-37.
- Gamurrini, G. F., Cozza, A., Pasqui, A., & Mengarelli, R. (1972). *Carta archeologica d'Italia (1881-1897). Materiali per l'Etruria e la Sabina*. Olschki.

- García Marco, F. J. (2007). Ontologías y organización del conocimiento: Retos y oportunidades para el profesional de la información. *El Profesional de la Información*, 16(6), pp. 541-550. <https://doi.org/10.3145/epi.2007.nov.01>
- Giorgi, I. (1878). Il regesto del Monastero di S. Anastasio ad Aquas Salvias. *Archivio della R. Società Romana di Storia Patria*, I, pp. 49-77.
- Goings, K., & Abel, P. (2013). *The Value of Social Login. Solving the Engagement Gap. Insights from Consumer Research*.
- Golbeck, J. (2013). Nodes, Edges, and Network Measures. En J. Golbeck, *Analyzing the Social Web* (pp. 9-23). Elsevier. <https://doi.org/10.1016/B978-0-12-405531-5.00002-X>
- Goldberg, P. (1980). Micromorphology in Archaeology and Prehistory. *Paléorient*, 6(1), pp. 159-164. <https://doi.org/10.3406/paleo.1980.4265>
- Goldberg, P., & Aldeias, V. (2016). Why does (archaeological) micromorphology have such little traction in (geo)archaeology? *Archaeological and Anthropological Sciences*, 10(2), pp. 269-278. <https://doi.org/10.1007/s12520-016-0353-9>
- Goldberg, P., Aldeias, V., Dibble, H., McPherron, S., Sandgathe, D., & Turq, A. (2013). Testing the Roc de Marsal Neandertal "Burial" with Geoarchaeology. *Archaeological and Anthropological Sciences*, pp. 1005-1015. <https://doi.org/10.1007/s12520-013-0163-2>
- Goldberg, P., & Berna, F. (2010). Micromorphology and context. *Quaternary International*, 214(1-2), 56-62. <https://doi.org/10.1016/j.quaint.2009.10.023>
- Goldberg, P., Dibble, H., Berna, F., Sandgathe, D., McPherron, S. J. P., & Turq, A. (2012). New evidence on Neandertal use of fire: Examples from Roc de Marsal and Pech de l'Azé IV. *Quaternary International*, 247, pp. 325-340. <https://doi.org/10.1016/j.quaint.2010.11.015>
- Goldberg, P., & Macphail, R. I. (2006). *Practical and Theoretical Geoarchaeology*. Blackwell Publishing Ltd. <https://www.ajaonline.org/book-review/550>
- Goldberg, P., Miller, C. E., Schiegl, S., Ligouis, B., Berna, F., Conard, N. J., & Wadley, L. (2009). Bedding, hearths, and site maintenance in the Middle Stone Age of Sibudu Cave, KwaZulu-Natal, South Africa. *Archaeological and Anthropological Sciences*, 1(2), pp. 95-122. <https://doi.org/10.1007/s12520-009-0008-1>
- González Blanco, E. (2013). Actualidad de las humanidades digitales y un ejemplo de ensamblaje poético en la red ReMetCa. *Cuadernos hispanoamericanos*, 761, pp. 53-67. <https://doi.org/10.4000/jtei.1274>

- González Pérez, C. (1999). *Sistemas de Información para Arqueología: Teoría, Metodología y Tecnologías*. Santiago de Compostela.
- González Pérez, C. (2018). *Information Modelling for Archaeology and Anthropology. Software Engineering Principles for Cultural Heritage*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-72652-6>
- Green, S., Bevan, A., & Shapland, M. (2014). A comparative assessment of structure from motion methods for archaeological research. *Journal of Archaeological Science*, 46, pp. 173-181. <https://doi.org/10.1016/j.jas.2014.02.030>
- Gruber, E. (2010). *Building Omeka Exhibits with Fedora Repository Content | Scholars' Lab*. <https://scholarslab.lib.virginia.edu/blog/building-omeka-exhibits-with-fedora-repository-content/>
- Güll, P., Fronti, D., Romagnoli, G., & Wick, F. (2001). Viterbo. Indagini archeologiche 1997-1998: Nuovi dati per la topografia urbana e la cultura materiale. *Archeologia Medievale*, XXVIII, pp. 275-294. <https://doi.org/10.1400/245824>
- Gutiérrez Martínez, F. M. (2010). Las redes sociales: Una aproximación conceptual y una reflexión teórica sobre los posibles usos y gratificaciones de este fenómeno digital de masas. *TecCom Studies: Estudios de Tecnología y Comunicación*, 1, pp. 26-34.
- Gutiérrez Rodríguez, M., Brassous, L., Rodríguez Gutiérrez, O., Martín Peinado, F. J., Orfila, M., & Goldberg, P. (2018). Site formation processes and urban transformations during Late Antiquity from a high-resolution geoarchaeological perspective: Baelo Claudia (Southern Spain). *Geoarchaeology*, 35(2), pp. 258-286. <https://doi.org/10.1002/gea.21769>
- Gutiérrez Rodríguez, M., Toscano, M., & Goldberg, P. (2018). High-resolution dynamic illustrations in soil micromorphology: A proposal for presenting and sharing primary research data in publication. *Journal of Archaeological Science: Reports*, 20, pp. 565-575. <https://doi.org/10.1016/j.jasrep.2018.05.025>
- Haaland, M. M., Czechowski, M., Carpentier, F., Lejay, M., & Vandermeulen, B. (2018). Documenting archaeological thin sections in high-resolution: A comparison of methods and discussion of applications. *Geoarchaeology*, 34(1), pp. 100-114. <https://doi.org/10.1002/gea.21706>
- Haaland, M. M., Friesem, D. E., Miller, C. E., & Henshilwood, C. S. (2017). Heat-induced alteration of glauconitic minerals in the Middle Stone Age levels of Blombos Cave, South Africa: Implications for evaluating site structure and burning

- events. *Journal of Archaeological Science*, 86, pp. 81-100.
<https://doi.org/10.1016/j.jas.2017.06.008>
- Hartmann, L. M., & Merones, M. (1895). *Ecclesiae S. Mariae in Via Lata Tabularium*.
- Hedges, M., & Dunn, S. (2018). *Academic Crowdsourcing in the Humanities*. Chandos Publishing.
- Hendrickx, M., Gheyle, W., Bonne, J., Bourgeois, J., De Wulf, A., & Goossens, R. (2011). The use of stereoscopic images taken from a microdrone for the documentation of heritage – An example from the Tuekta burial mounds in the Russian Altay. *Journal of Archaeological Science*, 38(11), pp. 2968-2978.
<https://doi.org/10.1016/j.jas.2011.06.013>
- Howey, M. C. L., & Brouwer Burg, M. (2017). Assessing the state of archaeological GIS research: Unbinding analyses of past landscapes. *Journal of Archaeological Science*, 84, pp. 1-9. <https://doi.org/10.1016/j.jas.2017.05.002>
- Inventario y Caracterización de los Regadíos de Andalucía. Distritos Mediterráneos y Atlánticos*. (2008). Junta de Andalucía, Consejería de Agricultura y Pesca.
- Istituto Italiano dei Castelli. (1985). *Carta dei luoghi fortificati del Lazio* (1. edizione) [Istituto italiano dei castelli]. Fratelli Palombi editori.
- Jenks, G. F. (1967). The data model concept in statistical mapping. *International Yearbook of Cartography*, 7, pp. 186-190.
- Jirón, P. (2015). De la participación a la co-creación. Nuevas formas de pensar intervenciones para mejorar el habitar urbano/residencial. *Materia Arquitectura*, 12, pp. 66-75.
- Jones, G. D. B. (1962). Capena and the Ager Capenas. *Papers of the British School at Rome*. XXX, pp. 116-207; XXXI, pp. 100-158.
- Jorayev, G., Wehr, K., Benito-Calvo, A., Njau, J., & de la Torre, I. (2016). Imaging and photogrammetry models of Olduvai Gorge (Tanzania) by Unmanned Aerial Vehicles: A high-resolution digital database for research and conservation of Early Stone Age sites. *Journal of Archaeological Science*, 75, pp. 40-56.
<https://doi.org/10.1016/j.jas.2016.08.002>
- Karkanias, P., & Van de Moortel, A. (2014). Micromorphological analysis of sediments at the Bronze Age site of Mitrou, central Greece: Patterns of floor construction and maintenance. *Journal of Archaeological Science*, 43, pp. 198-213.
<https://doi.org/10.1016/j.jas.2014.01.007>

- Kräutli, F., & Valleriani, M. (2018). CorpusTracer: A CIDOC database for tracing knowledge networks. *Digital Scholarship in the Humanities*, 33(2), pp. 336-346. <https://doi.org/10.1093/llc/fqx047>
- Lagoze, C., Payette, S., Shin, E., & Wilper, C. (2006). Fedora: An architecture for complex objects and their relationships. *International Journal on Digital Libraries*, 6(2), pp. 124-138. <https://doi.org/10.1007/s00799-005-0130-3>
- Lanconelli, A, & Romagnoli, G. (2018). L'incastellamento nel Lazio. En A. Augenti & P. Galetti, *Incastellamento. Storia e archeologia. A 40 anni da les structures di Pierre Toubert* (pp. 379-389). Spoleto.
- Lanconelli, Angela. (1992). Dal Castrum alla civitas: Il territorio di Viterbo tra VIII e XI secolo. *Società e storia*, 56, pp. 245-266.
- Lewenstein, B. V. (2003). Models of public communication of science and technology. *Public Understanding of Science; Technology*, pp. 1-11.
- Little, B. J., & Shackel, P. A. (2014). *Archaeology, heritage, and civic engagement: Working toward the public good*. Left Coast Press.
- López Poza, S. (2014). Humanidades digitales hispánicas. *Cincuentenario de la Asociación Internacional de Hispanistas: A Coruña, del 11 al 13 de diciembre de 2012*, pp. 151-166.
- Lucía Megías, J. M. (2003). La «Informática Humanística»: Notas volanderas desde el ámbito hispánico. *Incipit*, 23, pp. 91-114.
- Lucía Megías, J. M. (2012). *Elogio del texto digital: Claves para interpretar el cambio de paradigma*. Fórcola Ediciones.
- Macchi Jánica, G. (1999). *Las bases de datos en la investigación arqueológica* (1a ed). Ediciones Uninorte.
- Machado, J. (2015). Open data and open science. En Sarita Albagli, M. L. Maciel, & A. H. Abdo, *Open science, open issues* (pp. 189-214). Instituto Brasileiro de Informação em Ciência e Tecnologia.
- Macphail, R. I., Courty, M.-A., & Goldberg, P. (1990). Soil micromorphology in Archaeology. *Endeavour*, 14(4), pp. 163-171. [https://doi.org/10.1016/0160-9327\(90\)90039-T](https://doi.org/10.1016/0160-9327(90)90039-T)
- Macphail, R. I., & Goldberg, P. (2010). Archaeological Materials. En G. Stoops, V. Marcelino, & F. Mees, *Interpretation of Micromorphological Features of Soils and Regoliths* (pp. 589-622). Elsevier. <https://doi.org/10.1016/B978-0-444-53156-8.00026-X>

- Mallett, M., & Whitehouse, D. (1967). Castel Porciano: An abandoned Medieval Village of the Roman Campagna. *Papers of the British School at Rome*, 35, pp. 113-146. <https://doi.org/10.1017/S0068246200007613>
- Mallol, C., & Mentzer, S. M. (2017). Contacts under the lens: Perspectives on the role of microstratigraphy in archaeological research. *Archaeological and Anthropological Sciences*, 9(8), pp. 1645-1669. <https://doi.org/10.1007/s12520-015-0288-6>
- Marcos Marín, F. M. (1994). *Informática y humanidades*. Gredos.
- Martín Civantos, J. M. (2015). Landscape archaeology and rural communities: Ethnoecology and social involvement. En A. Chavarria Arnau & A. Reynolds, *Detecting and Understanding Historic Landscapes* (pp. 335-358). PCA.
- Martlew, R. (Ed.). (1984). *Information systems in archaeology*. A. Sutton.
- Matthews, W., French, C. A. I., Lawrence, T., Cutler, D. F., & Jones, M. K. (1997). Microstratigraphic traces of site formation processes and human activities. *World Archaeology*, 29(2), pp. 281-308. <https://doi.org/10.1080/00438243.1997.9980378>
- McCoy, M. D. (2017). Geospatial Big Data and archaeology: Prospects and problems too great to ignore. *Journal of Archaeological Science*, 84, pp. 74-94. <https://doi.org/10.1016/j.jas.2017.06.003>
- Mentzer, S. M., & Quade, J. (2013). Compositional and Isotopic Analytical Methods in Archaeological Micromorphology. *Geoarchaeology*, 28(1), pp. 87-97. <https://doi.org/10.1002/gea.21425>
- Milioni, A. (2002). *Viterbo. I (IGM F. 136 I SE Capodimonte, 136 II SE La Rocca, 137 SO La Commenda, 137 III SO Castel d'Asso, 137 II SE Viterbo, 137 IV SO Montefiascone)*. Viterbo.
- Miller, C. E., Conard, N., Goldberg, P., & Berna, F. (2010). Dumping, sweeping and trampling: Experimental micromorphological analysis of anthropogenically modified combustion features. *Palethnologie*, 2, pp. 25-37.
- Mira Rico, J. A., & Zurita Aldeguer, R. (2020). Historia pública, patrimonio del conflicto y museos. *Her&Mus. Heritage & Museography*, 21, pp. 4-7.
- Molinari, A. (2010). Siti rurali e poteri signorili nel Lazio (secoli X-XIII). *Archeologia medievale*, 37, pp. 129-142.
- Monaci, A. (1904). Regesto dell'abbazia di Sant'Alessio all'Aventino. *Archivio della R. Società Romana di Storia Patria*, XXVII, pp. 351-398; XXVIII, pp. 151-200.

- Mozas Calvache, A. T., Pérez García, J. L., Cardenal Escarcena, F. J., Mata Castro, E., & Delgado García, J. (2012). Method for photogrammetric surveying of archaeological sites with light aerial platforms. *Journal of Archaeological Science*, 39(2), pp. 521-530. <https://doi.org/10.1016/j.jas.2011.10.007>
- Nardi Combescure, S. (2002). *Paesaggi d'Etruria meridionale: L'entroterra di Civitavecchia dal II al XV secolo d.C.* Ed. All'Insegna del Giglio.
- Nevens, F., Frantzeskaki, N., Gorissen, L., & Loorbach, D. (2013). Urban Transition Labs: Co-creating transformative action for sustainable cities. *Journal of Cleaner Production*, 50, pp. 111-122. <https://doi.org/10.1016/j.jclepro.2012.12.001>
- Nielsen, J. (2006). *The 90-9-1 Rule for Participation Inequality in Social Media and Online Communities*. Nielsen Norman Group. <https://www.nngroup.com/articles/participation-inequality/>
- Noiret, S. (2013). Digital History 2.0. En Frederic Clavert & Noiret, Serge, *L'histoire contemporaine à l'ère numérique – Contemporary History in the Digital Age* (pp. 155-190). Peter Lang. <https://orbilu.uni.lu/handle/10993/34981>
- Noiret, S. (2015a). Digital Public History: Bringing the public back in. *Public History Weekly*, 2015(13). <https://doi.org/10.1515/phw-2015-3931>
- Noiret, S. (2015b). Historia digital e Historia Pública. En J. A. Bresciano & T. Gil, *La historiografía ante el giro digital. Reflexiones teóricas y prácticas metodológicas*. (pp. 57-112).
- Noiret, S. (2015c). Historia Pública Digital | Digital Public History. *Liinc em Revista*, 11(1). <https://doi.org/10.18617/liinc.v11i1.797>
- Noiret, S. (2018). Digital Public History. En D. Dean, *A Companion to Public History* (pp. 111-124). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118508930.ch7>
- O'Brien, C. (2010). *Guidance for the Care, Conservation and Recording of Historic Graveyards*. The Heritage Council of Ireland.
- Oldman, D., & Tanase, D. (2018). Reshaping the Knowledge Graph by Connecting Researchers, Data and Practices in ResearchSpace. En D. Vrandečić, K. Bontcheva, M. C. Suárez-Figueroa, V. Presutti, I. Celino, M. Sabou, L.-A. Kaffee, & E. Simperl, *The Semantic Web – ISWC 2018* (Vol. 11137, pp. 325-340). Springer International Publishing. https://doi.org/10.1007/978-3-030-00668-6_20
- Oomen, J., & Aroyo, L. (2011). Crowdsourcing in the cultural heritage domain: Opportunities and challenges. *Proceedings of the 5th International Conference on*

Communities and Technologies - C&T '11, pp. 138-149.
<https://doi.org/10.1145/2103354.2103373>

- Ortega, É., & Eunice Gutiérrez, S. (2014). MapaHD. Una exploración de las Humanidades Digitales en español y portugués. En E Romero Frías & M. Sánchez González, *Ciencias sociales y humanidades digitales: Técnicas, herramientas y experiencias de e-research e investigación en colaboración*. (pp. 101-128). Sociedad Latina de Comunicación social.
- Parry, D. (2012). The Digital Humanities or a Digital Humanism. En M. K. Gold, *Debates in the Digital Humanities* (pp. 429-437). Minneapolis.
- Pasquali, A. (1977). *Comprender la comunicación*. Monte Ávila.
- Patterson, H. (2004). *Bridging the Tiber: Approaches to regional archaeology in the Middle Tiber Valley*. British School at Rome.
- Pawlicka-Deger, U. (2019). *Humanities and Media Labs Around the World*. Zenodo.
<https://doi.org/10.5281/ZENODO.2631219>
- Pinzi, C. (1887). *Storia della città di Viterbo*. Fornin.
- Ponce, I. (2012). *Clasificación de las redes sociales*. MONOGRÁFICO: Redes Sociales - Clasificación de redes sociales | Observatorio Tecnológico. <http://recursostic.educacion.es/observatorio/web/ca/internet/web-20/1043-redes-sociales?start=3>
- Pons, A. (2011). "Guardar como". La historia y las fuentes digitales. *Historia Crítica*, 43, pp. 38-61. <https://doi.org/10.7440/histcrit43.2011.04>
- Pons, A. (2018). El pasado fue analógico, el futuro es digital. Nuevas formas de escritura histórica. *Ayer*, 110, pp. 19-50.
- Potter, Timothy W., King, A. C., Allason-Jones, L., Andrews, R. D., & Warren, K. (1997). *Excavations at the Mola di Monte Gelato: A Roman and Medieval settlement in South Etruria*. British School at Rome.
- Potter, T.W. (1979). *Storia del paesaggio dell'Etruria Meridionale*. British School at Rome.
- Prins, A. B., Adams, M. J., Homsher, R. S., & Ashley, M. (2014). Digital Archaeological Fieldwork and the Jezreel Valley Regional Project, Israel. *Near Eastern Archaeology*, 77(3), pp. 192-197. <https://doi.org/10.5615/neareastarch.77.3.0192>
- Proyecto MEMOLA. (2014). Mediterranean Mountainous Landscape: An historical approach to cultural heritage based on traditional agrosystems. *Agua y Territorio*, 4, pp. 146-148.

- Puy Rodríguez, A., Cabello Valdés, C., Ruiz Galán, O., Sanmartín Sola, J., & Rodríguez Méndez, C. (2019). *Científicas en cifras 2017*. Fundación Española para la Ciencia y la Tecnología (FECYT): indicadores I+D+i.
- Ralli, M., Bekiaris, S., Kaldeli, E., Menis-Mastromichalakis, O., Sofou, N., Tzouvaras, V., & Stamou, G. (2020). CrowdHeritage: Improving the quality of Cultural Heritage through crowdsourcing methods. *2020 15th International Workshop on Semantic and Social Media Adaptation and Personalization (SMA)*, pp. 1-6. <https://doi.org/10.1109/SMAP49528.2020.9248463>
- Ramírez Sánchez, M. R. (2016). Las ciencias y técnicas historiográficas en el contexto de las humanidades digitales: Oportunidades para su desarrollo. En A. Marchant Rivera & L. C. Barco Cebrián, «*Dicebamus hesterna die...*»: *Estudios en homenaje a los profesores Pedro J. Arroyal Espigares y M^a Teresa Martín Palma*, 2016. (pp. 365-392). ENCASA Ediciones y Publicaciones.
- Ramsay, S. (2013a). On Building. En M. Terras, J. Nyhan, & E. Vanhoutte, *Defining Digital Humanities: A Reader* (1.^a ed., pp. 242-245). Routledge. <https://doi.org/10.4324/9781315576251>
- Ramsay, S. (2013b). Who's In and Who's Out. En M. Terras, J. Nyhan, & E. Vanhoutte, *Defining Digital Humanities: A Reader* (1.^a ed., pp. 239-241). Routledge. <https://doi.org/10.4324/9781315576251>
- Ramsay, S., & Rockwell, G. (2012). Developing Things: Notes toward an Epistemology of Building in the Digital Humanities. En M. K. Gold, *Debates in the Digital Humanities* (pp. 75-84). University of Minnesota Press. <https://doi.org/10.5749/minnesota/9780816677948.003.0010>
- Rapp, G. R., & Hill, C. L. (2006). *Geoarchaeology: The earth-science approach to archaeological interpretation*. Yale University Press.
- Remondino, F., Barazzetti, L., Nex, F., Scaioni, M., & Sarazzi, D. (2011). UAV photogrammetry for mapping and 3d modeling—Current status and future perspectives. *Int. Arch. Photogramm. Remote. Sens.*, 38 (1C22), pp. 25-31. <https://doi.org/DOI: 10.5194/isprsarchives-XXXVIII-1-C22-25-2011>
- Renfrew, C. (1976). Introduction. En D. A. Davidson & M. L. Shackley, *Geoarchaeology: Earth science and the past*. Duckworth.
- Ridge, M. (2013). From Tagging to Theorizing: Deepening Engagement with Cultural Heritage through Crowdsourcing. *Curator: The Museum Journal*, 56(4), pp. 435-450. <https://doi.org/10.1111/cura.12046>

- Ridge, M. (2014). *Crowdsourcing our Cultural Heritage*. Ashgate: Routledge.
- Robinson Garcia, N., & Amat, C. B. (2018). ¿Tiene sentido limitar la coautoría científica? No existe inflación de autores en Ciencias Sociales y Educación en España. *Revista española de Documentación Científica*, 41(2), pp. 201. <https://doi.org/10.3989/redc.2018.2.1499>
- Robles Gómez, A., González Blanco, E., Ros, S., del Rio Riande, G., Hernández, R., Tobarra, L., Caminero, A., & Pastor, R. (2016). *Researcher's perceptions of DH trends and topics in the English and Spanish speaking community. DayofDH data as a case study*". pp. 658-660. <https://www.academica.org/gimena.delrio.riande/5>
- Rodríguez Ortega, N. (2018). Cinco ejes para pensar las humanidades digitales como proyecto de un nuevo humanismo digital. *Artnodes*, 22, pp. 1-6. <https://doi.org/10.7238/a.v0i22.3263>
- Rojas Castro, A. R. (2013). El mapa y el territorio: Una aproximación histórico-bibliográfica a la emergencia de las Humanidades Digitales en España. *Caracteres: estudios culturales y críticos de la esfera digital*, 2(2), pp. 10-53.
- Romagnoli, G. (2006). *Ferento e la Teverina viterbese: Insediamenti e dinamiche del popolamento tra il X e il XIV secolo*. Università della Tuscia.
- Romero Frías, E., & Bocanegra Barbecho, L. (2018). *Ciencias sociales y Humanidades Digitales aplicadas. Casos de estudio y perspectivas críticas*. Universidad de Granada y Downhill Publishing. <https://digibug.ugr.es/handle/10481/53480>
- Romero Frías, E., & Del Barrio García, S. (2014). Una visión de las humanidades digitales a través de sus centros. *El Profesional de la Información*, 23(5), pp. 485-492. <https://doi.org/10.3145/epi.2014.sep.05>
- Roosevelt, C. H., Cobb, P., Moss, E., Olson, B. R., & Ünlüsoy, S. (2015). Excavation is Destruction Digitization: Advances in Archaeological Practice. *Journal of Field Archaeology*, 40(3), pp. 325-346. <https://doi.org/10.1179/2042458215Y.0000000004>
- Rowe, G., & Frewer, L. J. (2005). A Typology of Public Engagement Mechanisms. *Science, Technology, & Human Values*, 30(2), pp. 251-290. <https://doi.org/10.1177/0162243904271724>
- Ruest, N., & Stapelfeldt, K. (2014). Introduction to Islandora. *YorkSpace Repository*, 93.
- Rutherford, E., Hemati, W., & Mehler, A. (2018). *Corpus2Wiki: A MediaWiki based Annotation & Visualisation Tool for the Digital Humanities*. pp. 1-8. <https://doi.org/10.18420/INFDH2018-08>

- Sánchez Nogales, E. (2019). ComunidadBNE: El proyecto de enriquecimiento colaborativo de la Biblioteca Nacional de España. *Mi biblioteca: La revista del mundo bibliotecario*, 59, pp. 68-72.
- Sandgathe, D. M., Dibble, H. L., Goldberg, P., & McPherron, S. P. (2011). The Roc de Marsal Neandertal child: A reassessment of its status as a deliberate burial. *Journal of Human Evolution*, 61(3), pp. 243-253. <https://doi.org/10.1016/j.jhevol.2011.04.003>
- Sanz Casado, E., Melero, R., Aleixandre Benavent, R., Codina, L., Coslado Bernabé, M. A., de Filippo, D., Giménez Toledo, E., Jiménez Contreras, E., & Rico Castro, P. (2020). *Metodología de clasificación de revistas de humanidades y ciencias sociales con Sello de Calidad*. Fundación Española para la Ciencia y la Tecnología (FECYT): indicadores I+D+i.
- Savignoni, P. (1895). L'Archivio Storico del Comune di Viterbo: Documenti per la storia di Roma nel medio evo (1169-1495). *Archivio della R. Società Romana di Storia Patria*, XVIII, pp. 5-50.
- Scardozzi, G. (2004). *Ager Ciminius*. Viterbo.
- Schach, S. R. (1996). *Classical and object-oriented software engineering*. Irwin.
- Schiaparelli, L. (1901). Le carte antiche dell'Archivio Capitolare di San Pietro in Vaticano. *Archivio della R. Società Romana di Storia Patria*, XXIV, pp. 273-496; XXV, pp. 273-354.
- Schiffer, M. (1985). Is There a «Pompeii Premise» in Archaeology? *Journal of Anthropological Research*, 41, pp. 18-41. <https://doi.org/10.1086/jar.41.1.3630269>
- Schiffer, M. B. (1987). *Formation Processes of the Archaeological Record by Michael*. University of New Mexico Press.
- Scholz, M., & Görz, G. (2012). WissKI: A Virtual Research Environment for Cultural Heritage. *ECAI*, pp. 1017-1018. <https://doi.org/10.3233/978-1-61499-098-7-1017>
- Schreibman, S., Siemens, R. G., & Unsworth, J. (Eds.). (2016). *A New Companion to Digital Humanities*. Wiley/Blackwell.
- Schreibman, S., Siemens, R., & Unsworth, J. (2004). *A Companion to Digital Humanities (Blackwell Companions to Literature and Culture)*. Blackwell Publishing Professional. <https://doi.org/10.1002/9780470999875>
- Sciotti, E. (2019). Il Transcribathon: Un nuovo approccio alle lettere manoscritte risalenti alla Grande Guerra. *DigItalia*, 2(0), pp. 116-122.

- Sedov, S., Stoops, G., & Shoba, S. (2010). Regoliths and Soils on Volcanic Ash. En G. Stoops, V. Marcelino, & F. Mees, *Interpretation of Micromorphological Features of Soils and Regoliths* (pp. 275-303). Elsevier. <https://doi.org/10.1016/B978-0-444-53156-8.00013-1>
- Senn, J. A. (1992). *Análisis y diseño de sistemas de información*. McGraw-Hill.
- Shillito, L. M., Almond, M. J., Nicholson, J., Pantos, M., & Matthews, W. (2009). Rapid characterisation of archaeological midden components using FT-IR spectroscopy, SEM-EDX and micro-XRD. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 73(1), pp. 133-139. <https://doi.org/10.1016/j.saa.2009.02.004>
- Shillito, L.-M., & Matthews, W. (2013). Geoarchaeological Investigations of Midden-Formation Processes in the Early to Late Ceramic Neolithic Levels at Çatalhöyük, Turkey. *Geoarchaeology*, 28(1), pp. 25-49. <https://doi.org/10.1002/gea.21427>
- Signorelli, G. (1907). *Viterbo nella storia della Chiesa*. Tipografia Cionfi.
- Silberschatz, A., Korth, H. F., & Sudarshan, S. (2019). *Database system concepts*. McGraw-Hill.
- Silvestrelli, G. (1914). *Città, castelli e terre della regione romana*. Città di Castello.
- Silvestrelli, G. (1917). Le chiese e i feudi dell'Ordine dei Templari e dell'Ordine di San Giovanni in Gerusalemme nella Regione Romana. *Rendiconti della Reale Accademia dei Lincei. Classe di scienze morali, storiche e filologiche*, V (XXVI, 5-6), pp. 491-539.
- Simpson, R., Page, K. R., & De Roure, D. (2014). Zooniverse: Observing the world's largest citizen science platform. *Proceedings of the 23rd International Conference on World Wide Web*, pp. 1049-1054. <https://doi.org/10.1145/2567948.2579215>
- Sommella Mura, A. (1969). *Repertorio degli scavi e delle scoperte archeologiche nell'Etruria Meridionale, I (1939-1965)*. Roma.
- Sommerville, I. (2005). *Ingeniería del software*. Pearson Education.
- Sora, V. (1906). I conti d'Anguillara dalla loro origine fino al 1465. *Archivio della R. Società Romana di Storia Patria*, XXIX, pp. 397-442.
- Spence, P., & González Blanco, E. (2014). *A historical perspective on the digital humanities in Spain*. H-Soz-Kult. Kommunikation und Fachinformation für die Geschichtswissenschaften. <http://www.hsozkult.de/debate/id/diskussionen-2449>

- Spiro, L. (2011). Computing and Communicating Knowledge: Collaborative Approaches to Digital Humanities Projects. En L. McGrath, *Collaborative Approaches to the Digital in English Studies*. (pp. 44-82). Logan: Utah State University Press.
- Stapelfeldt, K., & Moses, D. (2013). Islandora and TEI: Current and Emerging Applications/Approaches. *Journal of the Text Encoding Initiative, Issue 5*. <https://doi.org/10.4000/jtei.790>
- Stoops, G., & Vepraskas, M. J. (2003). *Guidelines for analysis and description of soil and regolith thin sections*. Soil Science Society of America.
- Stroncik, N. A., & Schmincke, H.-U. (2002). Palagonite – a review. *International Journal of Earth Sciences*, 91(4), pp. 680-697. <https://doi.org/10.1007/s00531-001-0238-7>
- Supino, P. (1969). *La Margarita Cornetana Regesto dei documenti*. Società romana di storia patria.
- Svensson, P. (2010). The Landscape of Digital Humanities. *Digital Humanities Quarterly*, 004(1). <http://digitalhumanities.org/dhq/vol/4/1/000080/000080.html>
- Tarquini, S., & Favalli, M. (2010). A microscopic information system (MIS) for petrographic analysis. *Computers & Geosciences*, 36(5), pp. 665-674. <https://doi.org/10.1016/j.cageo.2009.09.017>
- Terras, M. M. (2016). Crowdsourcing in the Digital Humanities. En S. Schreibman & R. Siemens, *Companion to Digital Humanities II* (pp. 420-439). Wiley-Blackwell. <http://eu.wiley.com/WileyCDA/WileyTitle/productCd-1118680596.html>
- Theiner, A. (1861). *Codex diplomaticus domini temporalis S. Sedis: 756-1334*. Imprimerie du Vatican.
- Thomas, S. (2004). What is Participatory Learning and Action (PLA): An Introduction. *Centre for International Development and Training, London: CIDT*.
- Thorseth, I. H., Furnes, H., & Tumyr, O. (1995). Textural and chemical effects of bacterial activity on basaltic glass: An experimental approach. *Chemical Geology*, 119(1-4), pp. 139-160. [https://doi.org/10.1016/0009-2541\(94\)00098-S](https://doi.org/10.1016/0009-2541(94)00098-S)
- Tomassetti, G. (1897). Del sale e del focatico del Comune di Roma nel Medioevo. *Archivio della R. Società Romana di Storia Patria*, XX, pp. 313-368.
- Tomassetti, G. (1910). *La Campagna Romana Antica, Medioevale e Moderna*. Loescher.

- Toscano, M. (2008). La red de los castillos de la provincia de Viterbo, Italia (siglos X-XV): *Arqueología y Territorio Medieval*, 15, pp. 73-89. <https://doi.org/10.17561/aytm.v15i0.1495>
- Toscano, M. (2018a). Andalucía y América. Plataforma colaborativa sobre patrimonio artístico iberoamericano. *Quiroga: Revista de Patrimonio Iberoamericano*, 14, pp. 36-43. <http://doi.org/10.5281/zenodo.3403745>
- Toscano, M. (2018b). Desde datos locales a plataformas abiertas: La Web como herramienta de colaboración con la sociedad. En L. Bocanegra Barbecho & E. Romero Frías, *Territorios Digitales. Cosntruyendo unas Ciencias Sociales y Humanidades Digitales*. (pp. 192-193). Universidad de Granada y Downhill Publishing.
- Toscano, M. (2018c). Where the researcher cannot get: Open platforms to collaborate with citizens on cultural heritage research data. En E Romero Frías & L. Bocanegra Barbecho, *Ciencias Sociales y Humanidades Digitales Aplicadas* (pp. 538-561). Granada, España: Universidad de Granada. [ISBN: 978-84-338-6318-8]. <https://doi.org/10.5281/zenodo.1469337>
- Toscano, M. (2019a). *Crowdsourcing en Historia*. La investigación en la era digital. <https://digibug.ugr.es/handle/10481/55656>
- Toscano, M. (2019b). Crowdsourcing Irish History. *3rd National Conference of the Italian Association of Public History – AIPH*, University of Campania Vanvitelli (Santa Maria Capua Vetere, Caserta) 24th-28th of June 2019. <http://hdl.handle.net/10481/56247>
- Toscano, M., & Bocanegra Barbecho, L. (2016). The Spanish Republican Exile: Identity, Belonging and Memory in the Digital World. En K. J. Borowiecki, N. Forbes, & A. Fresa, *Cultural Heritage in a Changing World* (pp. 237-253). Springer International Publishing. https://doi.org/10.1007/978-3-319-29544-2_14
- Toscano, M., & Colella, S. (2019). *Good practices of social participation in cultural heritage* (1.0). Zenodo. <https://doi.org/10.5281/ZENODO.3725500>
- Toscano, M., Gutiérrez-Rodríguez, M., & Goldberg, P. (2021). *Ultra-resolution images - demo package*, Mendeley Data, V1. <http://dx.doi.org/10.17632/2xx36kdvmr.1>
- Toscano, M., Rabadán, A., Ros, S., & González-Blanco, E. (2020). Digital humanities in Spain: Historical perspective and current scenario. *El profesional de la información*, 29(6), e290601. <https://doi.org/10.3145/epi.2020.nov.01>
- Toscano, M., & Romagnoli, G. (2020). Atlante dei siti fortificati della provincia di Viterbo, Italia (X-XV secolo). Fonti e metodi per la ricostruzione della rete

- insediativa bassomedievale. X, pp. 471-478, en Navarro Palazón, J. y García Pulido, L. (eds), *Defensive architecture of the Mediterranean*, Universidad de Granada y Universidad Politécnica de Valencia. <https://doi.org/10.4995/FORTMED2020.2020.11545>
- Toscano, M., & Tierney, J. (2016). *HistoricGraves.com. Explore, Protect, Promote*. HistoricGraves.com. <https://historicgraves.com/sites/default/files/presentation/index.html#/>
- Toscano, M., & Tierney, J. (2020). Historic Graves. *Visualizing Objects, Places, and Spaces: A Digital Project Handbook*. <https://doi.org/10.21428/51bee781.4c0eec46>
- Toscano, M., Bocanegra Barbecho, L., Ros, S., & Gonzalez Blanco, E. (2020). Insights on scholarly primitives from Digital Humanities research in Spain. *Scholarly Primitives-Dariah Annual Event 2020. Zagreb, Croatia*, pp. 538-556. <https://doi.org/10.5281/ZENODO.4256689>
- Toscano, M., Colella, S., Delgado, L., Debródi, B., Krajiček, J., Berlekamp, F., György, E., Fresa, A., Hoffmann, Robert, & Martín Civantos, José María. (2019). *Participatory activities good practices in the field of cultural heritage (REACH project) (1.2.0)*. Zenodo. <https://doi.org/10.5281/ZENODO.3415123>
- Toscano, M., & Díaz, A. (2020). *Mapping digital humanities in Spain - 1993-2019 (v1.0)*. Zenodo. <https://doi.org/10.5281/ZENODO.3893546>
- Toubert, P. (1973). *Les structures du Latium médiéval le Latium méridional et la Sabine du IXe siècle à la fin du XIIe siècle*. Ecole française de Rome.
- Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, 9(1), pp. 5233. <https://doi.org/10.1038/s41598-019-41695-z>
- Trifone, B. (1908). Le carte del Monastero di S. Paolo di roma dal secolo XI al XV. *Archivio della R. Società Romana di Storia Patria*, XXXI, pp. 267-313; XXXII, pp. 29-106.
- Vallelonga, F. (2012). Insediamenti fortificati di età medievale in un territorio di confine: L'area dei Monti della Tolfa e la valle del Mignone. *Aristonothos. Rivista Di Studi Sul Mediterraneo Antico*, 5, pp. 173-221. <https://doi.org/10.6092/2037-4488/2243>
- Vendittelli, M. (2008). *Sutri nel Medioevo. Storia, insediamento urbano e territorio*.

- Waissbluth, M., Contreras, E., Galaz, P., Aguilera, I., Inostroza, J., Infante, A., & Gatica, M. (2014). Co-creación para la innovación: Un caso en el sector público chileno. *Revista de Ingeniería de Sistemas, XXVIII*, pp. 5-26.
- Walsh, K. (2016). *Open innovation, open science, open to the world: A vision for Europe*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2777/061652>
- Wattez, J., Courty, M. A., & Macphail, R. I. (1990). Burnt Organo-Mineral Deposits Related to Animal and Human Activities in Prehistoric Caves. En Douglas L.A, *Developments in Soil Science, Soil Micro-Morphology: A Basic and Applied Science Elsevier* (Vol. 19, pp. 431-439). Elsevier. [https://doi.org/10.1016/S0166-2481\(08\)70358-1](https://doi.org/10.1016/S0166-2481(08)70358-1)
- Weiner, S. (2010). *Microarchaeology: Beyond the Visible Archaeological Record*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511811210>
- Weiner, S., Goldberg, P., & Bar-Yosef, O. (1993). Bone Preservation in Kebara Cave, Israel using On-Site Fourier Transform Infrared Spectrometry. *Journal of Archaeological Science, 20*(6), pp. 613-627. <https://doi.org/10.1006/jasc.1993.1037>
- Whitten, J. L., Bentley, L. D., & Dittman, K. C. (2004). *Systems Analysis and Design Methods*. McGraw-Hill Irwin.
- Wiberg, D. (2014). *MUKURTU: Information retrieval system engineered for indigenous individuals and communities. I*. pp. 1-13.
- Wickham, C. (1978). Historical and topographical notes on early medieval South Etruria. *Papers of the British School at Rome, XLVI*, pp. 133-179; *XLVII*, pp. 66-95.
- Wilson, A. T., & Edwards, B. (2016). *Open Source Archaeology: Ethics and Practice*. De Gruyter Open. <https://doi.org/10.1515/9783110440171>
- Wouters, B., Makarona, C., Nys, K., & Claeys, P. (2016). Characterization of Archaeological Metal Remains in Micromorphological Thin Sections Using μ XRF Elemental Mapping: MXRF ELEMENTAL MAPPING OF ARCHAEOLOGICAL METALS. *Geoarchaeology, 32*(2), pp. 311-318. <https://doi.org/10.1002/gea.21591>
- Wouters, P., Beaulieu, A., Scharnhorst, A., & Wyatt, S. (2012). *Virtual Knowledge: Experimenting in the Humanities and the Social Sciences*. MIT Press.
- Zorich, D. (2008). *A survey of digital humanities centers in the United States*. Council on Library and Information Resources.