

This is a pre-print version of the following research article:

Alvite-Díez, M.L., & Barrionuevo, L. (2021). Confluence between library and information science and digital humanities in Spain. Methodologies, standards and collections. *Journal of Documentation*, 77(1), 41-68. <https://doi.org/10.1108/JD-02-2020-0030>

## Confluence between *Library and Information Science* and *Digital Humanities* in Spain. Methodologies, standards, and collections

**María-Luisa Alvite-Díez**

University of León, Spain

[luisa.alvite@unileon.es](mailto:luisa.alvite@unileon.es)

**Leticia Barrionuevo**

University of León, Spain

[buffl@unileon.es](mailto:buffl@unileon.es)

### Abstract

**Purpose** – The purpose of this paper is to study the relevance of heritage collections and the convergence of methodologies and standards traditionally linked to Library and Information Science (LIS) in the development of digital humanities (DH) research in Spain.

**Design/methodology/approach** – This paper is based on a systematic review of scientific publications that are representative of DH in Spain and were published between 2013 and 2018. The analysis considered doctoral theses, journal articles and conference papers.

**Findings** – The results highlight the synergies between documentary heritage, Library and Information Science and digital humanities. However, it appears that there is a scarcity of scientific literature to support the confluence of LIS and DH and a limited formal connection between heritage institutions and the areas of academia that reuse and enrich these source collections.

**Research limitations/implications** – The review of representative scholarly DH publications was mainly based on the metadata that describe the content of articles, thesis, and conference papers. This work relies on the thematic indexing (descriptors and keywords) of the analysed documents but their level of quality and consistency is very diverse.

**Originality/value** – The topic of the study has not been explored before and this work could contribute to the international debate on the interrelation and complementarity between Library and Information Science and digital humanities. In addition, this paper shows the contribution that standards and documentary methodologies make to projects in which technology is applied to humanities disciplines. We propose that there is an urgent need to strengthen the “scientific relationships” between heritage institutions, as well as enhancing links between the academic field of DH and LIS in order to improve teaching and research strategies in conjunction.

**Keywords** Digital collections, Digital Humanities, Documentary methodologies, Documentary standards, Library and Information Science

**Paper type** Research paper

## Introduction

The development of Digital Humanities (DH) has been linked, from its origins, with the concepts of collaboration and interdisciplinarity. In addition, the scientific literature has shown, amongst the interrelated set of disciplines associated with DH initiatives are related, a further close relationship with heritage collections found in libraries, archives and museum, and in parallel, with Library and Information Science (LIS) and the standards and methodologies traditionally linked to this discipline.

Combining subject knowledge and computer science, Digital Humanities are inherently interdisciplinary, data-driven, generally collaborative and heavily indebted to libraries and LIS (Poole, 2017). While international studies have been completed to analyse shared areas of interest and mutual impact of DH and LIS (Balck, 2016; Burghardt *et al.*, 2015; Frank, 2016; Gladney, 2012; Hobohm, 2016; Robinson *et al.*, 2015; Sula, 2013; White and Gilbert, 2016), this epistemological and procedural permeability has not been studied in Spain, a country which has seen many DH projects taking off in recent years.

This work aims to study the relevance of Spanish heritage collections and the confluence of methodologies and standards traditionally linked to LIS in the development of research in Digital Humanities in Spain. To this end, a bibliographic review of scientific publications on DH in the period from 2013 to 2018 is carried out. Based on the review, conclusions and suggestions for improvement are presented.

The subject of this work is a theme that has not been researched previously and we feel that it is of clear interest to both disciplines, contributing to the international debate concerning the interrelation and complementarity of Library and Information Science and Digital Humanities.

The article is arranged into five sections. Following from this Introduction justifying the interest and scope of the work, we present the Background to our work, including the contextual and epistemological framework of the study. The Objectives and Methodology section outlines the aims of the investigation, as well as the methodology employed and its various phases. The next section is dedicated to the analysis of results, organised into three sets according to the documentary typology object of study. Finally, we present our conclusions and reflect on the findings of our study.

## Background

This section presents a review of the academic literature, discussing the relevance of digital heritage collections and the concomitance of LIS in the development of DH. Likewise, it contextualizes the current state of digital collections and DH in Spain.

### ***Digital Collections***

The heritage collections of libraries, archives and museums are a priority type of infrastructure when it comes to the provision of content that can be used and deployed in digital humanities teaching and research. Efforts at mass digitization of collections in recent decades and the ease with which these collections can be accessed have contributed to an acknowledgement of the role played by these institutions and should result in the use of library and archival content in DH initiatives.

Digital humanities projects are usually the outcome of collaborative work and generally have a multidisciplinary and multilingual relevance. They seek to create rigorous resources that contribute to scientific development in diverse disciplines: language studies, literature, philosophy, history, art, bibliography, and so on. In many of these projects, original sources have been analysed; texts and other forms of cultural expression have been digitized; links to

the source collections have been established; and agreements have been entered into with libraries and archives that keep the originals for the purpose of heritage dissemination. As Steiner *et al.*, (2014), point out, the huge value of these cultural heritage treasures, available online thanks to large scale digitisation, is not fully appreciated. Digital collections often lack the qualities to enable quantitative and qualitative analysis in great depth. Digital humanities initiatives have a primary role in improving the quality of access to collections of cultural heritage and in providing new virtual environments for their investigation which support the conceptualisation, visualisation, and analysis of information, and, indeed, collaborative projects. It is unsurprising that Spence (2014) calls our attention to the need for research on the scientific relationship between digital libraries and digital humanities.

Terras, a leading researcher in this area, points out that digital humanities' rise as a discipline is inevitable and would have happened whatever name it was given. To refer to her own work, however, she uses the term Heritage Science and notes that DH will have to work with the Heritage Science academic community in the future (Terras, 2016). For Svensson (2014) there is an unbreakable link between what is termed "digital cultural heritage" which is concerned with digital management and access to cultural heritage in a generic sense. Robinson, Priego, and Bawden (2015) show how a good part of DH research is carried out in connection or collaboration with institutions that host museum or document collections. Dalbello (2011), in her overview of the formative stages of DH, highlights the value of the reflections of these pioneers centering on meaning and reading in digital libraries in order to discursively construct the DH as whole.

Likewise, the origin of DH must to a large extent be placed within a set of practices and a community of professionals that in turn emerged in libraries and archives. In their study, Clement *et al.*, (2013) underline the deepening trust in building user-builder communities of archivists, librarians and researchers in the humanities who work with technology. The recommendations concerning the use of XML format to ensure the reusability of information and, specifically, the TEI (Text Encoding Initiative) initiative, which is vital for the humanities, emerged from the work of digital academic-publishing programmes, many of which were based in libraries. The same is true of several pioneering DH projects, including important work done on digital archives, interface design, and text analysis (Hockey, 2004, 2006; Posner, 2013).

According to Zeng (2019), when heritage materials are worked with in DH research, the most challenging stage is essentially the first one, when decisions are made about what data can be collected and how. The items that are the object of interest may be documents, artefacts and other types of information-bearing objects. They might or might not be digitized; they may or may not be text based; and they could exist in various formats and media. For researchers who need heritage data that are difficult to obtain by trawling the Web, the data and information resources provided by archives, libraries and museums are of extraordinary value.

Nonetheless, although the report *Europe's Digital Humanities Landscape* (Wilms, Derven, O'Dwyer, Lingstadt, and Verbeke, 2019, p. 12), which gathers data from fifty-four institutions in twenty European countries, supports the unquestionable reliability and role of physical collections in research, it draws attention to the insufficiency of digitized objects and materials as well as to the licenses that govern possible use of them. For her part, Terras (2015) points out that, despite the enormous levels of investment in the digitization of cultural and heritage materials, it is very difficult to access statistics and reports on the ways in which digitized collections have been used. She notes that although the environment in which digitized content is created and provided has changed radically in the last twenty years, allowing the sharing and reuse of digital data and encouraging new advances in research, we are still at an initial stage, one in which problems with granting of licences and open access to cultural and heritage content persist.

## ***Digital Humanities and Library and Information Science***

Beyond the obvious importance of heritage collections in the development of DH, it would be worthwhile to investigate the permeability of DH projects when it comes to the methodologies, technologies, and standards inherent to LIS. The epistemological reach of LIS is prevalent within the context of DH and is not always entirely transparent (Svensson, 2014). In their definition of DH Bradley *et al.*, (2018) emphasise the idea of community, specifically, “scholarly learning community with a convergence on shared methods and values”, similarly the idea of a “Community of Practice” is central for Poole and Garwood (2018), that is to say, DH can be conceived as a field in which the potential for collaboration is great. For Rodríguez-Ortega, (2018), the practice of Digital Humanities has amplified the research sphere of the humanities scholar, who, alongside what they may have traditionally considered to be their field of activity, now engage in the production of new artefacts in order to take forward their research work and undertake cultural analyses (digital resources, databases, digital libraries, analytical systems, corpus linguistics, user interaction platforms, etc.).

Sula (2013) demonstrates the growing interest in DH within LIS publications, as well as a major overlap of interests, competences and institutional structures between the two disciplines. Wang (2018), meanwhile, describes digital humanities as an interdisciplinary movement and the result of a dynamic dialogue between humanistic exploration and digital means. Authors such as Dalbello (2011), Koltay (2016), Poole (2017) o Robinson *et al.*, (2015) reinforce this idea of common interests, suggesting that DH and LIS complement one another to a very great degree since both disciplines focus on documents and recorded information. Going further, Poole and Garwood (2018b) refer to archivists and librarians as natural allies in DH and their work gives support to the premise that interdisciplinary collaboration online nurtures connections between researchers in LIS and DH to create sustainable knowledge bases which are also more innovative than those which either discipline could produce in isolation.

Two important bibliometric studies provide essential information concerning the state of research in DH and verify some of the characteristics with which this field has defined itself. Tang *et al.*, (2017) examine the level of diversity in topic and intellectual cohesion in the literature published between 1989 and 2014. The results show that publications in the field of DH continue to grow in number, its diversity and coherence, two distinguishing characteristics of an interdisciplinary approach, continue to be reinforced. The network of co-authors observed, however, remained fairly fragmentary, with collaborations limited on the basis of language and geography, and with few international collaborations. The work of Wang (2018), for its part, analysed 803 entries on the database Web of Science, (WoS) and reaffirms the interdisciplinary nature of this material showing how Digital humanities are a research field closely related to history, cultural and literary heritage and Library and Information Science. The work also points to the progressive growth in scientific productivity, and its significant acceleration in the period 2006-2016. The most productive institutions were found to be Anglo-American and the dominant language was English (88.67%).

With respect to the techniques and skills associated with DH, Poole, (2017) addresses complementary working methods that encompass digital humanities, among which he indicates: digitization, crowdsourcing, archives and databases, digital curation, texts, editing, visualization, geospatiality, gaming and code. Robinson *et al.*, (2015), meanwhile, point out the following topics of interest shared between DH and LIS: search and retrieval; digital archives and libraries; description of metadata and resources; ontologies, classifications and taxonomies; publication and dissemination; open access; linked data; management and conservation of collections; portals and repositories; bibliography; digitization; preservation; interaction and user experience; interfaces and navigation; cultural heritage; information visualization; big data and data mining; and bibliometrics. Among the skills found in DH

research teams, Garwood and Poole (2019) recently listed ones related to: network analysis, visualization, text coding and data and metadata management.

Besides, the current state of technology means that we face a new challenge requiring heritage institutions to broaden the value of these digital collections so that they can go from being resource repositories to being data collections that can be shared, linked to, enriched and used better. Without doubt, and despite the long road ahead, considerable advances have been made on the part of heritage institutions in publishing data, metadata and vocabularies using the “Linked Open Data” model (Smith-Yoshimura, 2018), which is based on a firm commitment to the free reuse of said data. However, as Borgman (2009) comments, the humanities community, “should continue to clarify their choices of data and data sources”. Zeng (2019) describes this challenge as a passage from digitizing to “datafying”, which means converting heritage materials not only into machine-readable resources, but also into machine-processable ones. In the case of digital humanities, this approach emphasizes processes for transforming unstructured data into structured and semi-structured data. Terras *et al.*, (2018) refer to large-scale data exploitation, and since many institutions provide collections as data, this is a new phase in providing libraries’ content to more effectively support humanities specialists in developing methods and approaches that take advantage of these new research opportunities.

These technological developments are related to the results of Given and Willson’s work (2018) in which they observe two new research practices among academics in the humanities disciplines: the design of digital tools that are tailored to their specific needs and the preparation of data in formats that are storable, accessible, analysable and usable in the writing process. Similarly, the literature draws attention to humanities research’s successful embrace of the concept of “smart data”, which refers to transforming and analysing data to discover its hidden potential (Schöch, 2013; Zeng, 2017, 2019). Semantic Web technologies (RDF, Linked Open Data, OWL) allow data to be shared and reused from different domains and communities and provide the architecture for working with large amounts of data. This provides a new dimension to analytical data modeling activities (like semantic reasoning), which can be applied across heterogeneous datasets. The Semantic Web challenges for humanities scholars in promoting a more meaning-orientated approach to data representation (Oldman *et al.*, 2015; Thorsen and Pattuelli, 2016; Zeng, 2017).

### ***Heritage collections and DH in Spain***

The volume of bibliographic and documentary heritage in Spain is enormous and digitisation has increased the dissemination and visibility of a significant part. Of particular importance, we should mention The Spanish National Library, the digital collections coordinated by the Ministry of Culture and the exceptional resources of the State Archives, fundamental in preserving the memories of Europe, America, and the Pacific, up until the 18<sup>th</sup> century. At the same time, over the last few years, huge efforts have been made to constantly update technologies and standards, in order to achieve alignment with the semantic web, through the publication of both RDF controlled vocabularies and bibliographic collections using the format of Linked Open Data (Candela *et al.*, 2019; Carrato Mena, 2008; Lafuente Urién, 2016; Sánchez Mairena, 2017; Vila-Suero *et al.*, 2013).

In Spain, the beginning of the use of computer technologies within the humanities disciplines can be traced to the 1970s and their deployment in language-studies-related projects such as Philobiblon or Admyte. In the nineties, projects grew in number and extended to other disciplines, though initiatives were disparate and disconnected. In the twenty-first century, the first digital humanities programmes and studies on digital literature began; the bibliography indicates 2011 as the turning point when the growth of group events and meetings fostered the creation of the *HDH: Humanidades Digitales Hispánicas. Sociedad Internacional* (Hispanic

Digital Humanities: International Society), whose existence became official in 2012 (Iribarren, 2017; López Poza, 2015; Rodríguez Yunta, 2014; Rojas Castro, 2013; Spence and González-Blanco García, 2014).

Humanities research projects that have led to the construction of specialized digital archives, libraries and databases began from the mid-1990s, with libraries', archives', and museums' traditional or digital resources used as their starting point. In the field of language studies in particular, these digital collections have been giving shape to personalized platforms integrating digital heritage objects that have been described, analysed, contextualized and enriched for academic purposes. However, until now, there has been no attempt to make a systematic study of the reuse of digitized heritage collections nor the possible relationships between researchers in the humanities and the LIS community. Hispanists like Romero López (2018), explicitly recognise the role of digital memory in the digital humanities revolution: one memory that leads to more and better knowledge, but also to the reuse and sharing of knowledge on a scale unimaginable until now promoting the development of what this researcher terms "smart culture".

Focusing on the English-speaking environment, Poole (2017) underlines the relevant and growing impact of DH in the LIS community and vice versa and calls for rigorous study of the relationship between LIS/libraries and digital humanities. DH are a "making" culture and libraries can encourage the creation of creative digital products and projects. Digital humanities allow libraries to be better at aligning support services and infrastructure with humanities specialists' new research and communication behaviours; they can serve as an agent of change within the library (Rockenbach, 2013) Zhang, Liu, and Mathews (2015) conclude that DH communities and digital libraries share common challenges and tasks. Along these lines, Koltay (2016) talks about epistemologies and interests that are shared between DH and LIS, and he refers to the two disciplines' functions associated with the academic use of written/recorded information.

Within this contextual framework, this study aims to shed light on the impact of digital collections in DH research and the methodological permeability of LIS and DH through the work of academics in this discipline, which in Spain, is still in its inception.

## **Objectives and methodology**

The general aim of this study is to ascertain the role of the reuse of archival and library collections in the undertaking of digital humanities research in Spain and identify the transfer of documentary techniques in its studies and projects. Based on our examination, we present our conclusions and make suggestions for improvement.

The specific objectives are as follows:

- Investigate the actual importance of Spanish heritage collections in the undertaking of Hispanic digital humanities projects.
- Understand the confluence and interrelation of digital repositories in DH developments.
- Observe the direct participation of libraries and archives in humanities projects.
- Assess the interconnection of documentary standards, technologies, and methodologies in DH work.

The period that we study corresponds to the six-year period 2013-2018. This time period has been considered in response to the vital role in the rise of DH in Spain that the literature (López Poza, 2015; Rojas Castro, 2013; Spence and González-Blanco García, 2014) attributes to

the founding of the HDH<sup>1</sup>. HDH was founded in 2012 and its work bore fruit in the form of biennial international congresses, the first of which was held in 2013.

This study's methodology is based on a literature review of representative scholarly DH publications carried out in Spain between 2013 and 2018. The types of work included were papers presented at the HDH congresses, doctoral theses, and journal articles. Our work took place over four phases:

Phase 1: Collection of the scientific literature.

We included:

- a) Papers presented at the HDH congresses of 2013, 2015 and 2017. We took the view that these events are Spain's most representative meeting point when it comes to finding out about DH projects, developments, and advances, meaning that analysing the works presented at them would be essential to this study's aims.
- b) Doctoral theses from between the 2012/2013 and 2017/2018 academic years located in the Spanish Ministry of Education's Teseo database<sup>2</sup>. This is currently Spain's official, exhaustive database in which doctoral theses are collected. However, Teseo's search engine is not as powerful as one would hope for. Moreover, the database's thesaurus is based on the UNESCO Classification, and therefore its use is not appropriate for our research topic. Consequently, we set up a battery of searches with the aim of retrieving all theses possibly on digital humanities.

The search strategies entailed combining terms within the database's only two fields that are searchable: Title and Abstract. We set the time period covered by each search to that mentioned above. We performed the following searches:

- humanidades AND digitales  
(digital AND humanities)
- colecciones AND digitales  
(digital AND collections)
- humanidades AND tecnologías  
(humanities AND technologies)
- documentación AND digital  
(digital AND documentation)
- estándares AND documentación  
(documentation AND standards)
- humanidades AND archivos  
(humanities AND archives)
- "bibliotecas digitales" (PHRASE)  
(“digital libraries”)
- "archivos digitales" (PHRASE)  
(“digital archives”)

- c) Articles from academic journals. We included in the six-year research-funding period used as the time frame the journals that the scientific literature has identified as publications belonging to the Digital Humanities discipline (Iribarren, 2017; Rodríguez Yunta, 2014; Wang, 2018). The titles included in Paul Spence's<sup>3</sup> blog on Digital Humanities were also taken into consideration. Eleven titles were ultimately included.

In this phase limitations were identified in connection with the keywords used in the subject descriptions for the documents under analysis which were, in many cases too general. Wherever possible, we have gained access to the full text version of documents in order to determine the nature of the subject matter with more precision. In the analysis

section, this process is outlined and the lack of consistency in the use of indexing terms is evaluated in the results of this investigation.

Phase 2: Application of selection criteria to shape the collection to be studied.

The following guidelines were established:

- Contributions that specify the use of library and archival collections.
- Works that refer to the methodologies traditionally ascribed to LIS.
- Studies in which documentary standards are used.
- Papers from documentary institutions.

Phase 3: Systematic review of the selected literature.

We focused on the following elements:

- The researchers' university affiliations.
- Documentary institutions involved in DH research.
- Heritage collections used in studies / projects.
- Themes dealt with.
- Documentary methodologies and standards used.

Phase 4: Analysis and discussion

In this phase we will refer to the results of the bibliometric studies of Tang *et al.*, (2017) and Wang (2018) in order to assess possible correlations in terms of themes and standards seen in our study with respect to the analysis of international productivity within DH.

## Results and discussion

In this section, we present an analysis of our study's results. The section is organized into three sets that are based on the types of documents reviewed. We discuss the results in final part of the section.

### *Conference papers*

As we mentioned in the Methodology section, we selected papers presented at the three editions of the HDH congresses for which there existed full proceedings or summaries of them at the time of writing this article—that is, those from 2013, 2015 and 2017.

#### *Volume of papers*

The first HDH congress<sup>4</sup> was organized by the University of La Coruña and had the title: *Humanidades digitales: desafíos, logros y perspectivas de futuro (Digital Humanities: Challenges, achievements and perspectives on the future)*. The congress took place between 9 and 12 July 2013. The programme includes sixty-nine papers, though the total number of papers published in the journal *Janus* (Appendix 1; 2014)<sup>5</sup> after peer review of them was thirty-eight.

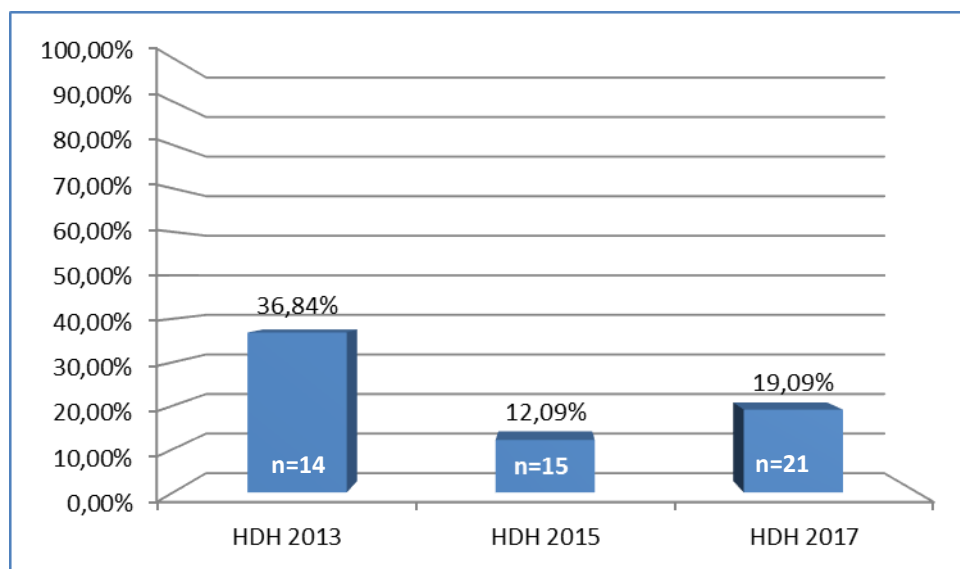
In 2015, the HDH Congress was overseen by Spain's National University of Distance Education (Universidad Nacional de Educación a Distancia; UNED). Held between 5 and 7 October 2015, its title was: *Humanidades digitales: Innovación, globalización e impacto (Digital Humanities: Innovation, globalization, and impact)*<sup>6</sup>. The programme lists four plenary lectures, forty-nine long papers, thirty-two short ones, seven panels and thirteen posters. An ebook of abstracts containing 124 works was published<sup>7</sup>.



Finally, the third HDH congress, *Sociedades, políticas, saberes (Societies, Policies, Knowledge)*<sup>8</sup> was organized by the University of Malaga and held between 18 and 20 October 2017. The programme lists three plenary lectures, forty-eight long papers, thirty short ones, six panels, nineteen posters and two art projects. A book of abstracts covering 110 works was published<sup>9</sup>.

In selecting papers, we opted for a restrictive criterion: in unclear cases, we preferred to exclude the paper. In the case of the 2013 congress, we analysed the full texts of the works, ensuring the selection was rigorous. However, in the cases of the second and third congresses, it was only possible to analyse the abstracts, which varied greatly in terms of length and quality. Moreover, the assigned keywords were in many cases of scarce significance. Specifically, those for the 2015 HDH congress were limited by the event organizers and were extremely generic. We counted panels as a single paper.

In total, fifty papers from the three congresses covered by our analysis made it through our selection filter: fourteen from HDH 2013, fifteen from HDH 2015 and twenty-one from HDH 2017. However, as Figure 1 shows, the thematic representativeness of the investigated content was greatest in the case of the 2013 congress.



**Figure 1. HDH congress papers analysed and their representativeness**

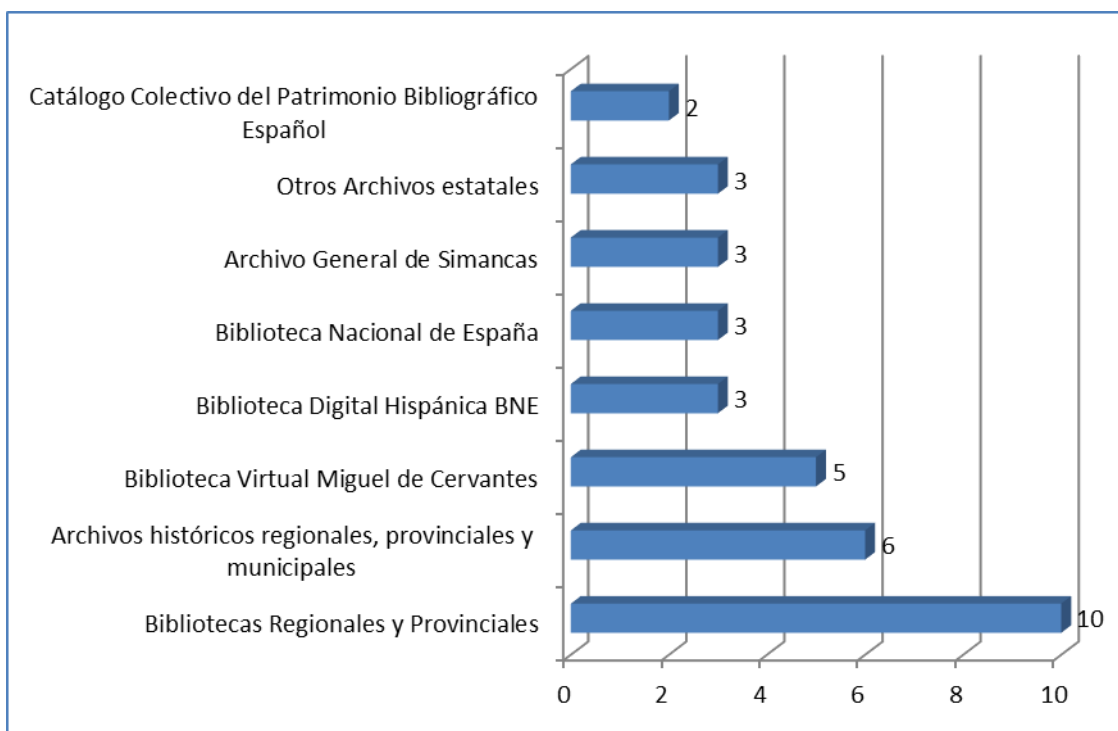
### *Universities*

In terms of the researchers' affiliations, thirty-two institutions were identified. The institution at the top of the list of relevant papers in our study is the University of La Coruña with eight papers that cover, among other things, projects such as Inventories and Libraries of the Spanish Golden Age (*Inventarios y bibliotecas del siglo de oro*; IBSO), Spanish Golden Age Digital Library (*Biblioteca Digital Siglo de Oro*; BIDISO), *Relaciones de Sucesos* Digital Library and Catalogue (*Catálogo y Biblioteca Digital de Relaciones de Sucesos*; CBDRS) and *Symbola*. In next place comes the Complutense University of Madrid with six papers that present projects such as *Mnemosine* (Digital Library of the Other Spanish Silver Age) and *Diachronic and Interactive Linguistic Atlas of the Community of Madrid* (*Atlas lingüístico diacrónico e interactivo de la Comunidad de Madrid*; ALDICAM-CM). UNED appears in four interuniversity papers, and the universities of Alicante and Santiago de Compostela each figure in three papers from the congresses analysed. In addition, we recorded contributions from institutions outside of Spain such as the universities of Bologna, Würzburg and Poitiers. These were generally authored by Hispanic language, literature or culture researchers.

*Source collections –collections used*

Among the documentary collections used in the papers on digital humanities presented at the three congresses, as can be seen in Figure 2, the Miguel de Cervantes Virtual Library is the individual collection that stands out<sup>10</sup>. It is worth highlighting the express reference made to an agreement; a paper on the Valle-Inclán Digital Archive states: “Created pursuant to a collaboration agreement between the University of Santiago de Compostela and the Miguel de Cervantes Virtual Library”. In the case of the Corpus of Castilian Poetry, a University of Alicante project, the paper in question states: “The texts have been taken from the Sonnet Library of the Miguel de Cervantes Virtual Library”.

Another noteworthy detail is that the collections of the National Library of Spain<sup>11</sup> and its Hispanic Digital Library<sup>12</sup> are each mentioned in three papers.



**Figure 2. Documentary source collections at the HDH congresses**

*Documentary institutions*

Table 1 shows the Spanish institutions that made contributions at the HDH congresses. These are four papers that present technological developments intended to disseminate and preserve documentary heritage of various kinds.

**Table 1. Papers presented by documentary institutions at the HDH congresses**

Congress	Institution	Title
HDH 2015	University of Barcelona. CRAI	From Digitization to Preservation: An Integrated Approach to the Dissemination of the Universitat de Barcelona’s Documentary Heritage (De la digitalización a la preservación: un tratamiento integrado para la difusión del patrimonio documental de la Universitat de Barcelona.).
HDH 2015	Institut Cartogràfic i Geològic of Catalonia	Georeferencing of Historical Cartography with the Community’s Help: The Experience of the

		Catalonia Map Collection (ICGC) (Georeferenciación de cartografía antigua con la ayuda de la comunidad: la experiencia de la Cartoteca de Cataluña (ICGC)).
HDH 2015	Archival Information and Documentation Centre (State Archives Branch of the Spanish Ministry of Education, Culture and Sport)	The Role of Archives in the Digital Humanities: The Current Picture (El papel de los archivos en las Humanidades Digitales. Estado de la cuestión).
HDH 2017	National Library of Spain	Use of the Music Collection -the National Library of Spain's Sheet Music and Audio Recordings- (La reutilización de la colección de Música -partituras y registros sonoros de la Biblioteca Nacional de España-).

What might be highlighted here is the fact that featuring among the institutions that participated in the 2017 congress are the National Library of Colombia, which presented the paper “Humanizing the Digital Humanities: The Challenge facing the National Library of Colombia” (“Humanizar las Humanidades Digitales: el reto de la Biblioteca Nacional de Colombia”); the National Library of Uruguay, which presented “Omeka and the Digital Publication of Delmira Agustini’s notebooks” (“Omeka y la edición digital de los cuadernos de Delmira Agustini”); and, finally, Stanford University Libraries (in collaboration with Spanish researchers), which presented “Leveraging the International Interoperability Framework and Mirador,” a paper that described an interoperability framework for managing cultural heritage images.

*Themes*

In terms of the themes addressed in the papers from the HDH congresses, the previously described disparities in the level of consistency with which key words were assigned should be taken into account. These disparities had a particular impact on the indexing terms used for the 2015 congress.

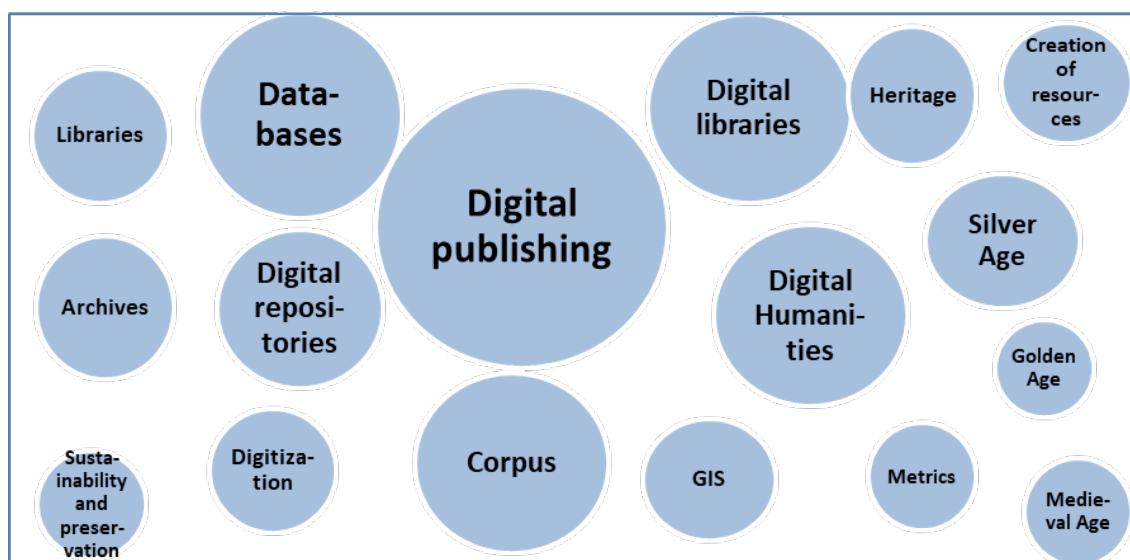


Figure 3. Themes at the HDH congresses

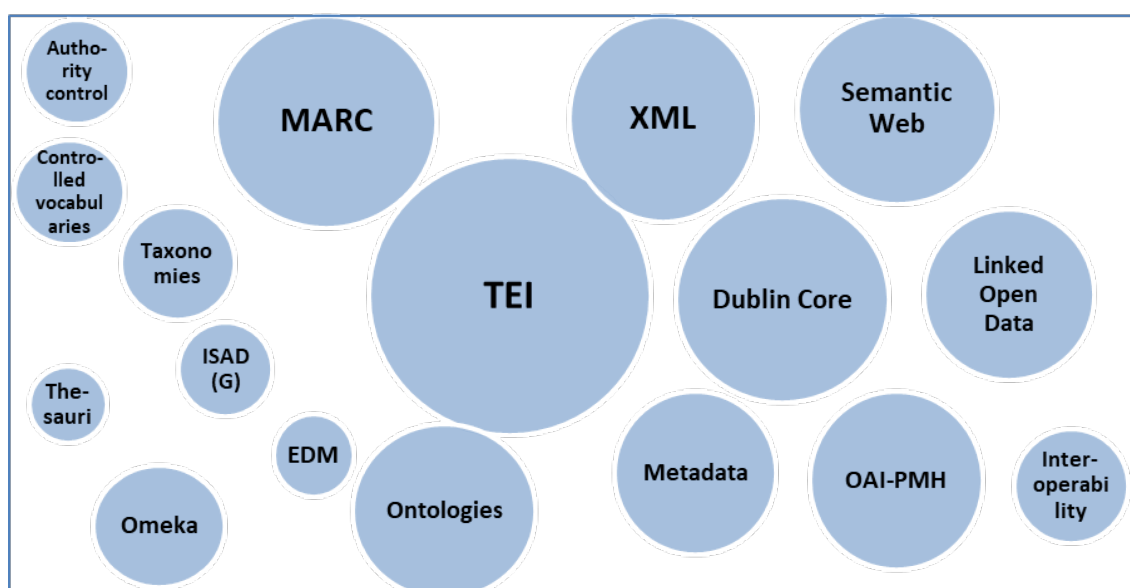
As Figure 3 shows, it is clear that the most frequently used keyword is “digital publishing,” which appears in ten of the papers, and is a fundamental concept to DH referring to the use of digital media in the cultural sphere in a way that is both cross and inter-disciplinary. This word is followed by “databases”, which appeared in seven, and “corpus”, “digital libraries”, and

“digital humanities” which each appeared in six papers. The predominant historical periods are the Spanish Golden Age (sixteenth century) and the Silver Age (1900-1939).

We can state that the main subjects collected reflect, to a large extent, themes of interest common to LIS and DH. It is perhaps worth talking about an umbrella category ‘digital cultural heritage’, which would cover the majority of the collected research topics: “libraries”, “archives”, “heritage”, “digital publishing”, “digitization”, “digital libraries”, “digital repositories”, and which underpins the intersection of study aims in LIS and DH. It seems striking that in papers submitted to the congresses, “digital humanities” is used as an indexing term, perhaps because the author / authors were reinforcing their contribution’s unequivocal place in the discipline.

#### *Documentary methodologies and standards*

Figure 4 shows a clear correlation with the identified thematic terms.



**Figure 4. Standards and methodologies at the HDH Congresses**

For the most part, what we find are language-studies projects in which digital publishing is addressed. As a result, the TEI (Text Encoding Initiative) is mentioned in nine of the analysed papers, while references to the Semantic Web and MARC format were found in seven contributions. Six texts refer to XML and Dublin Core and five mention OAI-PMH. Regarding the Semantic Web, it is noteworthy that four of the papers discuss the development of ontologies and of linked open data.

References to software mostly concern relational databases, though we did identify projects that are based on software that is widely implemented in the LIS environment. For instance, Omeka is referred to in three of the texts, and we also found mentions of the use of ICA-ATOM, KOHA, and Content DM.

## Doctoral theses

### Volume of defended theses

Table 2 presents the results of the different search strategies applied to the TESEO database that we described in the Methodology section. It also shows the number of theses that, following a reviewing of the abstract and the keywords included in the corresponding descriptive matter, were consistent with the established selection criteria.

**Table 2. Theses retrieved and theses' objects of study**

Search strategy	Number of theses found	Number of theses selected
humanidades AND digitales (digital AND humanities)	33	3
coleccionen AND digitales (digital AND collections)	23	2
humanidades AND tecnologías (humanities AND technologies)	52	0
documentación AND digital (digital AND documentation)	245	7
estándares AND documentación (documentation AND standards)	58	1
humanidades AND archivos (humanities AND archives)	35	0
"bibliotecas digitales" ("digital libraries")	23	4
"archivos digitales" ("digital archives")	7	0

Seventeen theses on the topic of focus were selected. Table 3 sets out their titles in order of the year of thesis defence. The results seem to reveal a stable trend from 2013 in terms of the number of theses in which the confluence of digital humanities and library and information science is observed. The most prolific years were 2016 and 2017, in each of which four theses were defended.

**Table 3. Theses' object of study**

Title	University	Department	Year
References to the Digestive System in Medical Texts Printed in Spain during the Renaissance and Retrieved from Digital Libraries (Referencias al aparato digestivo en los textos de materia médica impresos en España durante el Renacimiento y recuperados en Bibliotecas digitales)	University of Murcia	Internal Medicine	2013
Spatial Heritage Documentation and Preservation of Information: Needs, Possibilities, Strategies and Standards (Documentación espacial del patrimonio: preservación de la información. Necesidades, posibilidades, estrategias y estándares)	University of the Basque Country / Euskal Herriko Unibertsitatea	Geography, Prehistory and Archaeology	2014
Graphic Documentation Concerning and Applied to the Management, Conservation and Dissemination of Architectural Heritage: A Case Study of Villa de Ágreda (Soria) (Documentación gráfica del patrimonio arquitectónico aplicado a su gestión, conservación y difusión. El caso de estudio de la Villa de Ágreda (Soria))	Polytechnic University of Catalonia	Architectural Graphic Expression I	2014
Diffuse Linguistic Recommendation Systems for the Dissemination of Information in Digital Libraries	University of Granada	Artificial Intelligence and	2014

(Sistemas de recomendaciones lingüísticos difusos para la difusión de información en Bibliotecas digitales)		Computer Science	
Evaluating Digital Press Archives: The Case of the Main Ibero-American Newspapers (Evaluación de las hemerotecas de la prensa digital. El caso de los principales Diarios iberoamericanos)	Complutense University of Madrid	Library and Information Science	2015
Integration of Semantic Content into a Web Portal of Valencian Scientists and Humanists: Vestigium (Integración de contenidos semánticos en un portal web de científicos y humanistas valencianos: Vestigium)	Saint Vincent Martyr Catholic University of Valencia	Library and Information Science	2015
New Strategies for Conserving Art Collections with Technological Elements: Methodological Proposals from the Digital Humanities (Nuevas estrategias para la conservación de colecciones de arte con elementos tecnológicos: propuestas metodológicas de humanidades digitales)	Complutense University of Madrid	Painting and Conservation – Restoration	2015
The Study, Cataloguing and Digitization of the Artistic Bindings of the San Fernando Royal Academy of Fine Arts (Estudio, catalogación y digitalización de las encuadernaciones artísticas de la Real Academia de Bellas Artes de San Fernando)	Complutense University of Madrid	Library and Information Science	2016
The Evaluation of Cuban Digital Arts and Humanities Libraries: The case of the University of the Arts Library (La evaluación de las bibliotecas digitales de artes y humanidades cubanas: el caso de la Biblioteca de la Universidad de las Artes)	University of Granada	Library and Information Science	2016
Digital Humanities: University Repositories and Their Analysis through Naval History (Humanidades digitales: los repositorios universitarios y su análisis a través de la historia naval)	University of Murcia	Modern, Contemporary and American History	2016
IT Solutions and Languages for a Digital Repertoire of Spanish Medieval Metrics (Lenguajes y soluciones informáticas para un repertorio digital de la métrica medieval española)	National University of Distance Education	Spanish Literature and Literary Theory	2016
Geometric Documentation and the Digital Dissemination of Architectural Heritage Elements: The Capitals of the Royal Alcázar of Seville (Documentación geométrica y difusión digital de elementos arquitectónicos patrimoniales: capiteles del Real Alcázar de Sevilla)	University of Seville	Graphic Engineering	2017
Graphic Documentation and Revitalization of Historical Heritage through New Technologies (Documentación gráfica y dinamización del patrimonio histórico mediante nuevas tecnologías)	University of Seville	Graphic Engineering	2017
New Technologies Applied to Heritage Documentation: Guadix Alcazaba - Píñar Castle (Nuevas tecnologías aplicadas a la Documentación del Patrimonio. La Alcazaba de Guadix- El Castillo de Píñar)	University of Granada	Medieval History and Historiographic Technical Sciences	2017
Recognition and Classification of Entities Named in Medieval (Fourteenth Century) Documents: The <i>Becerrero de las Behetrías</i> (Reconocimiento y clasificación de entidades nombradas en documentos medievales (s. XIV): Libro Becerrero de las Behetrías)	University Carlos III of Madrid	Library and Information Science	2017
Application of 3D Technology to Techniques for the Documentation, Conservation and Restoration of Cultural Assets (Aplicación de la tecnología 3D a las técnicas de documentación, conservación y	Complutense University of Madrid	Painting and Conservation – Restoration	2018

restauración de bienes culturales)			
The Application of GIS Tools to Improve Retrieval of Granada's Historical Graphic Documentation: Georeferencing and Geolocation (Documentación histórica gráfica granadina. Aplicación de herramientas SIG para mejorar su recuperación: georreferenciación y geolocalización)	University of Granada	Library and Information Science	2018

### *Universities and Departments*

As can be seen in Table 3, the Complutense University of Madrid and the University of Granada are the institutions with the greatest output, with four theses each, followed by the University of Seville with two doctoral theses. As for the departments that supervised the theses, six are LIS department and six are humanities departments (Geography, Prehistory and Archaeology; Painting and Conservation - Restoration; Modern, Contemporary and American History; Spanish Literature and Literary Theory; Medieval History and Historiographic Technical Sciences). We can state that there is a balanced representation of the two sets.

It is worth mentioning that four of the theses originate from Engineering departments: Architectural Graphic Expression (Polytechnic University of Catalonia), Artificial Intelligence and Computer Science (University of Granada) and Graphic Engineering in the case of the two theses completed at the University of Seville.

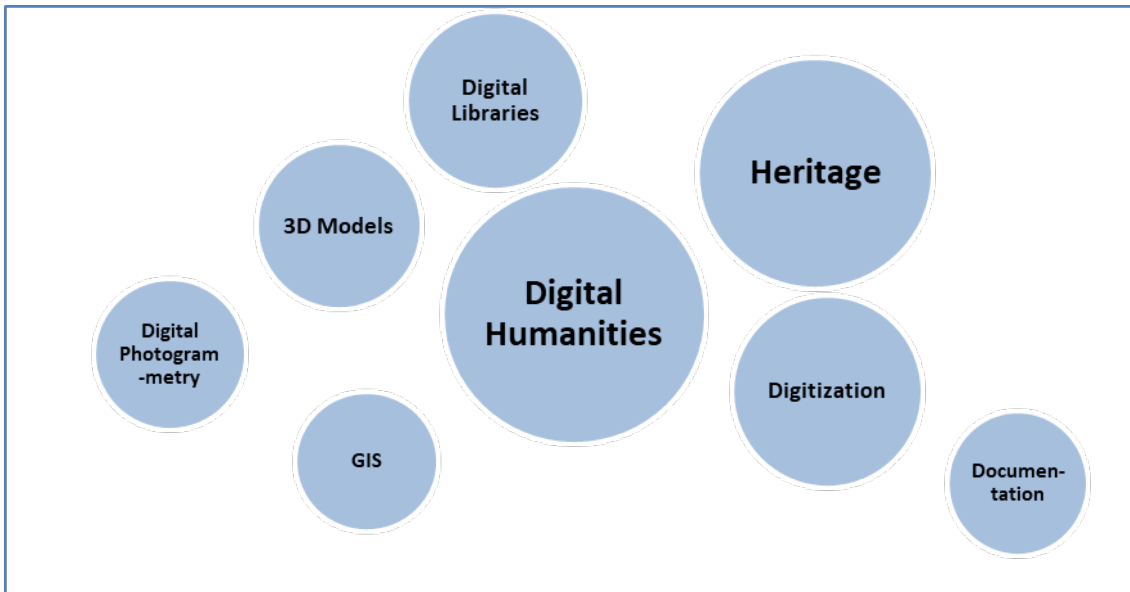
### *Documentary source collections*

The following documentary source collections are mentioned particularly prominently in the theses:

- Engraving Archive of the San Fernando Royal Academy of Fine Arts.
- Municipal Historical Archive of Granada.
- Library/Archive of the San Fernando Royal Academy of Fine Arts.
- Digital Library of Valencian Scientists and Scholars.
- Digital Library of the University of the Arts of Cuba.

### *Themes*

To compile themes and standards, we took into consideration the keywords that appear in the descriptive matter for each thesis in the TESEO database, as well as the relevant terms observed upon reading the corresponding abstract. As we noted in the Methodology section, TESEO's indexing is too generic, and its descriptors are not sufficiently representative of the theses' content (see Figure 5).



**Figure 5. Themes in the doctoral theses**

In the case of the doctoral theses, there are two themes that clearly stand out: “digital humanities” and “heritage.” These were found in seven of the theses analysed, meaning that 41% of the doctoral theses analysed correspond to these subjects. The keywords “digitization” and “digital libraries” each appear in four of the works (23.5%), with the theme of “3D models” appearing with the same frequency. Furthermore, the keywords “geographic information systems,” “digital photogrammetry” and “documentation” appear in three of the theses.

It is worth pointing out that in the semantic field of digital collections, in addition to “digital libraries,” we found: “digital repositories,” “digital press” and “virtual newspaper library.” Also striking is the use of disciplines/subdisciplines as descriptive keywords for the thesis subject. In addition to “documentation,” which appeared four times, we observed: “machine translation,” “computational linguistics,” “heraldry,” “codicology” and “literature.”

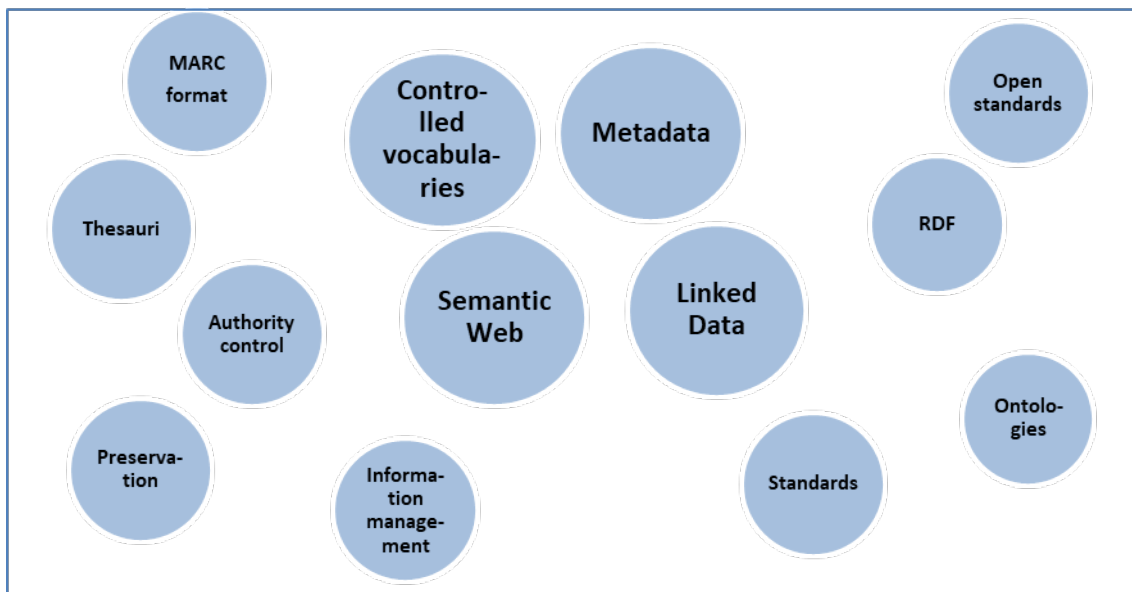
Regarding work methodologies generally associated with DH, in addition to the aforementioned term “geographic information systems” in three of the studies, we also observed “geolocation,” “georeferencing,” “3D models,” “augmented reality” and “virtual reality.” Likewise, the following descriptors more typical of language studies were identified: “data mining,” “corpus” and “linguistic modelling.”

#### *Standards and Methodologies*

“Semantic Web” and “linked data” featured in two of the theses that we analysed. Within this same field, the keywords “RDF” and “ontologies” were also identified.

And the terms “metadata” and “authority control” were included in two of the theses. The keywords “thesauri,” “controlled vocabularies” and “MARC format” also appear.





**Figure 6. Standards and methodologies in doctoral theses**

As Figure 6 shows, we also recorded more generic terms: “preservation,” “information management,” “standards” and “open standards.”

The thesis *Documentación espacial del patrimonio* [...], defended in 2014, describes the implementation of different metadata schemes to develop a document preservation and management system aimed at protecting them and ensuring their accessibility over time, guaranteeing their authenticity, and supporting access management, privacy and intellectual property rights. It is also worth mentioning the practical application that the author assigns to the MARC format as a standard used to encode fairly frequently occurring documents that, due to the needs of the management system, are transferred from the organization where they were generated to a library.

In the thesis *Integración de contenidos semánticos en un portal web de científicos y humanistas valencianos* (2015), which focuses on the development of the website of the digital library proposed in the research, the Drupal Content Management System is used. This open-source software, which conforms to Semantic Web guidelines and allows the publication of content with linked data technologies and RDF tagging functions, also incorporates terminology-control tools such as thesauri and knowledge classification systems in the form of taxonomies and ontologies. Likewise, the doctoral thesis *Nuevas estrategias para la conservación de colecciones de arte con elementos tecnológicos* [...], defended in 2015, advocates the use of controlled vocabularies and instruments for representing content. In this case, thesauri are used, on the one hand, to describe different types of relationships such as the place of birth of an artist, the cities where he or she worked, his or her contacts with other artists and the dates when those relationships came about. On the other hand, the author proposes the creation of an ontology that allows the different entities that make up contemporary artworks with a great variety of technological elements to be defined and understood.

Semantic Web technologies are also applied to linguistics and, more specifically, to the field of digital catalogues for metrics. This is what is highlighted in the thesis entitled *Lenguajes y soluciones informáticas para un repertorio digital de la métrica medieval española*, which proposes the use of linked open data and linked open vocabularies as a strategy for standardization and homogenization.

## Journals

### Volume

Within the literature indicated in the Methodology section, fourteen journal titles were considered. However, since analysis of the articles required us to understand them, three titles were excluded owing to the language in which they are written: *Bibliothek Forschung und Praxis*, whose articles are in German, and *Historical Social Research Historische Sozialforschung*, which is mainly in German. And *Computers and the Humanities (C&H)*, one of the first publications that appeared in the field of digital humanities and whose title changed in 2005, was not analysed, given that its articles are no longer accessible online.

In the end, the content from the 2013-2018 period of eleven journal titles that are representative of the digital humanities discipline was examined. Table 4 details the volumes/issue numbers, the articles surveyed and the articles in which the use of heritage documentary collections or documentary standards / methodologies plays an important role.

**Table 4. Titles of journals analysed**

Journals	Vols./issues. reviewed	Articles reviewed	Articles selected
Artnodes	10	132	0
Characters: Cultural and Critical Studies of the Digital Sphere (Caracteres: estudios culturales y críticos de la esfera digital)	11	123	3
DIGITUM: A relational perspective on culture and society	8	63	0
Digital Humanities Quarterly	23	243	0
Digital Scholarship in the Humanities (whose title used to be Literary and Linguistic Computing)	27	372	0
Frontiers in Digital Humanities	Not organized according to volumes	66	0
International Journal of Humanities and Arts Computing	10	106	0
Janus	6	31	0
Journal of Digital Humanities	4	47	0
Digital Humanities Journal (Revista de Humanidades Digitales)	2	30	3
Scriptum Digital	6	45	2

Given that Spain was set as the territorial scope for our object of study, we excluded articles found in the *International Journal of Humanities and Arts Computing* and the *Revista de Humanidades Digitales* whose themes related to other geographical locations. However, we decided to include among the selected articles “Knowledge Relations in the Digital Library of New Spanish Thought” (“Las relaciones de conocimiento en la biblioteca digital del pensamiento novohispano”), given that it relates to seventeenth-century texts from New Spain.

**Table 5. Articles analysed**

Title	University	Department	Year	Journal
<b>Knowledge Relations in the Digital Library of New Spanish Thought (Las relaciones de conocimiento en la biblioteca digital del pensamiento novohispano)</b>	National Autonomous University of Mexico	Philosophy and Humanities	2014	<i>Caracteres: estudios culturales y críticos de la esfera digital</i>

<b>The Don Quixote Image Bank (1605-1915): A Digital Tool for Understanding Quixotic Iconography (El banco de imágenes del Quijote: 1605-1915: una herramienta digital para conocer la iconografía quijotesca)</b>	Complutense University of Madrid	Department of Romance, French, Italian and Translation Studies	2015	<i>Caracteres: estudios culturales y críticos de la esfera digital</i>
<b>The Digital Library of Old Spanish Texts (BIDTEA) (La Biblioteca Digital de Textos del Español Antiguo (BIDTEA))</b>	College of the Holy Cross	Spanish Department	2015	<i>Scriptum Digital</i>
<b>Seven Digital Publishing Challenges for Documentary Sources (Siete retos en edición digital para las fuentes documentales)</b>	King's College London	Department of Digital Humanities	2015	<i>Scriptum Digital</i>
<b>Approaching the Digital Artistic Catalogue in Spain: Concepts, Context and Preliminary Analysis (Aproximación al catálogo artístico digital en España: conceptos, contexto y análisis preliminar)</b>	University of Granada	History of Art	2016	<i>Caracteres: estudios culturales y críticos de la esfera digital</i>
<b>The Roman Epigraphy of Augusta Emerita beyond the Museum: Digitization, 3D Modelling and Dissemination via Mobile Devices (La epigrafía romana de Augusta Emerita más allá del Museo: digitalización, modelización 3D y difusión a través de dispositivos móviles)</b>	University of Las Palmas de Gran Canaria	University Institute of Textual Analysis and Applications (IATEXT) and University Institute of Applied Microelectronics	2017	<i>Revista de Humanidades Digitales</i>
<b>The Digitization of the Silver Age: A Quantitative Analysis of Its State (Estado de la digitalización de la Edad de Plata: un análisis cuantitativo)</b>	University of Würzburg (Germany)	Digital Humanities and Spanish Philology	2017	<i>Revista de Humanidades Digitales</i>
<b>Representation and Organization of Digital Documents: Details and Practice regarding DIMH Ontology (Representación y organización de documentos digitales: detalles y práctica sobre la ontología DIMH)</b>	National University of Distance Education	Department of Computer Languages and Systems	2017	<i>Revista de Humanidades Digitales</i>

Only eight journal articles made it past the established selection filter. The number is so small that it does not allow us to assess the potential evolution over time of the themes studied. As Table 5 shows, the articles were found in three journals: *Caracteres: estudios culturales y críticos de la esfera digital*<sup>13</sup> (ISSN 2254-4496), a multilingual academic publication founded in 2011 that pays special attention to literary, linguistic, educational, philosophical and hypermedia studies. *Scriptum Digital*<sup>14</sup> (ISSN 2014-640X), founded in 2012 and published by the Autonomous University of Barcelona's Department of Spanish Language Studies; this is a research-oriented publication on theoretical and methodological aspects related to digital publishing and computerized diachronic corpora of the Iberian Romance languages. Finally, *Revista de Humanidades Digitales*<sup>15</sup> (ISSN 2531-1786), published by Spain's National University of Distance Education, was founded in 2017 on the back of collaboration between Digital Humanities researchers from UNED (Spain), the Argentine National Council of Scientific and Technical Research (Consejo Nacional de Investigaciones Científicas y Técnicas; CONICET) and

the National Autonomous University of Mexico (Universidad Nacional Autónoma de México; UNAM). It publishes academic research from the Digital Humanities.

### *Universities*

The researchers come from four Spanish universities (Complutense University of Madrid, University of Granada, University of Las Palmas de Gran Canaria and the National University of Distance Education), two European ones (King's College London and University of Würzburg) and two North American ones (National Autonomous University of Mexico and College of the Holy Cross).

Interdisciplinary collaboration particularly stands out in the case of the article “La epigrafía romana de Augusta Emerita más allá del Museo: digitalización, modelización 3D y difusión a través de dispositivos móviles.” Its authors are researchers from the University Institutes of Textual Analysis and Applications (IATEX) and Applied Microelectronics, both of which are located at the University of Las Palmas de Gran Canaria. The research contained in the article “Representación y organización de documentos digitales: detalles y práctica sobre la ontología DIMH” was carried out at the Department of Computer Languages and Systems. The remaining investigations come from humanities departments (in the case of King's College London and the University of Würzburg), and specifically from digital humanities departments.

### *Source collections*

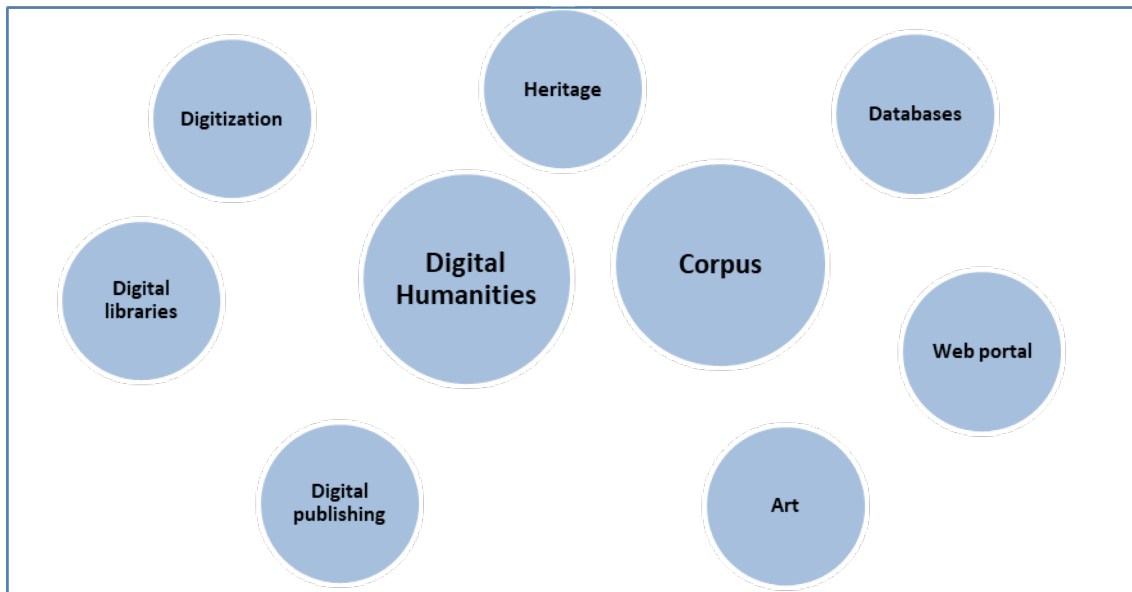
There are numerous references to the collections used in the studies presented. However, no formal agreements or links with documentary institutions are mentioned. Two collections stand out above the rest: the Miguel de Cervantes Virtual Library (mentioned in two of the articles) and the National Library of Spain's Hispanic Digital Library.

In addition to generic mentions, individualized references to the following collections appear: The Cervantes Library of the Miguel de Cervantes University Research Institute (Centre for Cervantes Studies), Alcalá de Henares; the Spanish Bibliographic Heritage Catalogue, National Chalcography (Madrid); Municipal Historical Library (Madrid); Bicentennial Digital Library; Virtual Library of Mexico and Mexican Digital Library.

Although they are not part of our object of this study, the reference made to an important set of museum collections—National Museum of Roman Art (Mérida), San Fernando Royal Academy of Fine Arts (Madrid), Museum of Fine Arts of Valencia, La Castellana Open-Air Sculpture Museum (Madrid) and the Patio Herreriano Museum of Contemporary Spanish Art (Valladolid)—is relevant.

### *Themes*

The most frequent descriptors are “digital humanities” and “corpus,” which were found in three of the articles. The terms “digitization,” “web portal,” “digital libraries,” “digital publishing” and “databases” are each used for two articles. We also find the themes “art” and “heritage” in two of the studies.



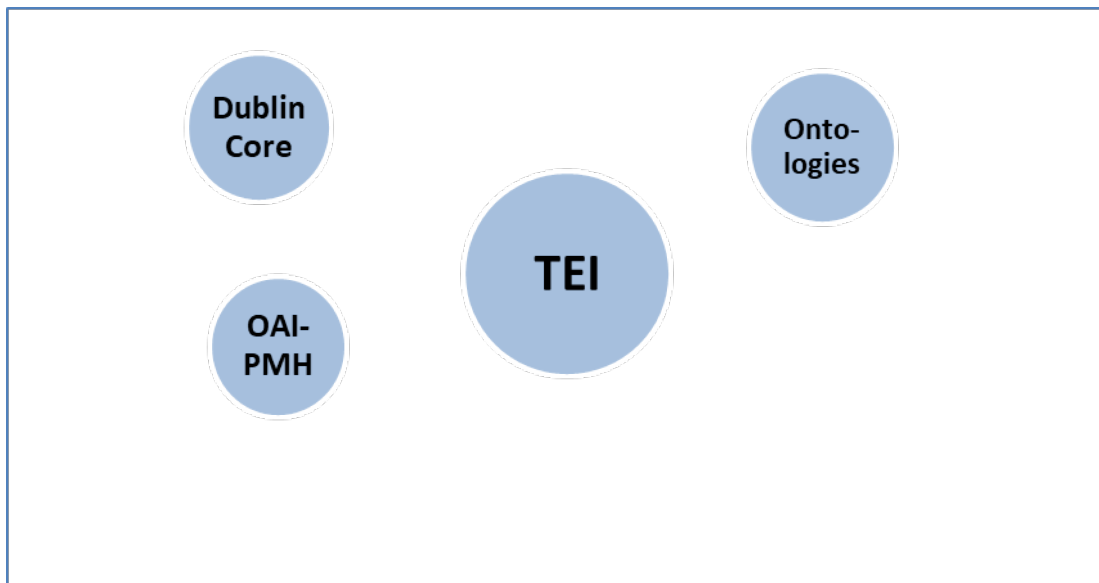
**Figure 7. Themes in the journal articles**

In all cases, what stands out is the idea that the type of study undertaken in the selected articles could only be carried out because of the different digital tools used. The use of data mining to obtain the direct information present in the texts is considered a relevant topic in some of the articles that we analysed. The historical periods that appear in the works examined range from the Roman era, in relation to which the digitization and 3D modelling of inscriptions along with dissemination of the results via mobile devices are addressed, to the medieval period and then on to the second half of the seventeenth century, in the form of the study of historical texts and semipalaeographic transcriptions of them. Within the selected articles, we also observed a combination of themes from different humanities disciplines, a clear example being the article entitled “El banco de imágenes del Quijote: 1605-1915: una herramienta digital para la conocer la iconografía quijotesca,” in which art and literature converge.

#### *Standards and methodologies*

We found the Text Encoding Initiative (TEI) standard in two of the studies analysed. The article “Las relaciones de conocimiento en la Biblioteca Digital del Pensamiento Novohispano” proposes TEI / XML markup for a series of documents on the nature of comets in the second half of the seventeenth century for subsequent automatic indexing of terms, references, names, and dates. In a similar vein, “Siete retos en edición digital para las fuentes documentales,” which discusses the CHARTA network, an initiative geared towards online publication of a corpus of historical document and texts originating from Spain and Latin America between the twelfth to the nineteenth centuries, puts forward the CHARTA-TEI model, whose objective is to publish documents from the network based on the TEI recommendations with a view to facilitating the preparation of a digital edition as well as the preservation of texts and the possibility of sharing them.

We also observed references to ontologies as formal and digital content-representation systems that facilitate access to and searches, visualization and automatic organization of that content, and we additionally found references to the Dublin Core metadata standard, which is useful for describing and identifying digital documents, and to the OAI-PMH protocol, which was drawn up to facilitate interoperability standards and the dissemination of content online.



**Figure 8. Standards and methodologies in the journal articles**

### *Discussion*

In total, we found seventy-five academic works that conformed to the selection criteria established to understand the interaction between Digital Humanities and Library and Information Science. The largest category in this collection under study is conference papers (50), followed by doctoral theses (17) and, finally, journal articles (8).

Although the results of the analysis responded to the points raised in the objectives, certain limitations must be considered regarding the configuration of the collection under study. These relate to the debatable consistency and depth of the thematic indexing that accompanies the documents. Prior to the selection of the documents, the literature review was based to a large degree on the descriptors and keywords that describe the content of articles, thesis and conference papers. In the case of doctoral theses, the TESEO database's low power search engine probably led to a certain margin of error in the information retrieval, and this has to be taken into account.

Although we surveyed the content of eleven journals that are representative of DH, we found only a very limited number of articles that met the selection criteria established. We believe that this limitation confirms, first of all, the continuation of the tradition of Spain's humanities researchers leaning more towards publishing their research works in books rather than in scientific journals, and, second, the dearth of Spanish research published on this subject in international journals. It also seems that journals dedicated to DH do not feature among the LIS researchers' preferred titles. Undoubtedly, our results can be explained by the short period in which DH research has been undertaken in Spain and, consequently, by the reputation and brief existence of some of this discipline's journals.

These caveats mentioned, we would venture to point out that the works presented at the three editions of the HDH congresses provided us with the most extensive knowledge of the concomitance between DH and LIS in Spain. In particular, it has been possible to more closely assess the use of heritage collections and the transfer of methodologies and documentary standards. Although they were scarce in number, we did find specific mentions of formal collaboration between documentary institutions and DH research projects. The analysed collection of conference papers mostly belong to the field of language studies and pursue research lines intrinsically linked to textual studies. Via the application of technologies, these

research lines have led to the creation of specialized digital repositories and corpora (Alvite-Díez and Pena Sueiro, 2020). These considerations are related to Robinson *et al.*'s (2015) observation in their study that most DH works related to LIS are linked to the creation and use of information and documents, the latter word being understood here in the broad sense of containers of recorded knowledge.

Three heritage collections clearly appear in both the HDH congress papers and in the journal articles: the National Library of Spain (or its subcollections, and especially the Hispanic Digital Library), the Miguel de Cervantes Virtual Library and the Spanish Bibliographic Heritage Collective Catalogue.

In the case of doctoral theses, investigations that focus on specific individualized collections are in all cases defended works from LIS departments, and they all address the application of technologies in document management and the dissemination of documentary heritage.

Regarding the origin of the researchers, the Complutense University of Madrid and UNED make appearances in the conference papers, theses, and journal articles. The set of theses also highlights the University of Granada's productivity. The significant number of studies—both conference papers and journal articles—that are recorded as falling within this subject area and originate from universities located outside of Spain is remarkable. Two of the doctoral theses' content focuses on Latin America.

We did not observe many collaborations between humanities and LIS researchers at the level of authorship. Worth noting, however, is collaboration with Computer Engineering departments, which can clearly be seen in the journal article "La epigrafía romana de Augusta Emérita [...]," a collaboration between a humanities specialist and an engineer. Assessment of the conference papers is less reliable here, since the authors' full affiliation does not always appear. That said, at least nine of the proposals include, in addition to academics from humanities disciplines, researchers from computer engineering or from university computing services. Moreover, the quantity of theses from engineering departments (23.6%) is remarkable.

Two significant points about the journal articles are worth highlighting: on the one hand, the absence of LIS studies, and, on the other hand, the notable presence of researchers not affiliated with Spanish institutions, who authored 50% of the works analysed.

Within the focal years of our study, we found few works from documentary institutions, but it should be stressed that the importance of DH for heritage institutions is confirmed by the participation of the National Library of Spain and the State Archives Branch of the Spanish Ministry of Education, Culture and Sport, both of which are leading institutions within Spain's library and archives system, at the HDH congresses.

There is a multidisciplinary interest in the subject that concerns us, one that is clearly apparent in the case of doctoral theses, where different humanities disciplines, various forms of engineering and, with a consistent presence, LIS are represented.

In the literature analysed, we observed the expected correlation between the thematic terms and the methodologies and standards identified. The most repeated and common subjects in all the research studied here are: "digital humanities," "digital libraries," "digitization" and "heritage."

The results of this study show significant correlation with the author keywords most frequently found in the study by Tang *et al.*, (2017) in the time period 2005-2014. The 38 keywords identified in that bibliometric study are those used with a frequency greater than 10 and of these, 16 (42.10%) appear in our study. In addition, we see a strong correlation with the map of keywords co-occurrence network of Wang (2018). Of the forty-one top descriptors that, in his bibliometric analysis, represent the digital humanities studies that appear in the WoS

database and were published between 1968 and 2017, twenty appeared in our work. The most common keyword is “digital humanities” in both studies cited above, and in the present study focussing on Spain. Amongst all the results of these studies, we find that, overwhelmingly, the most frequent keywords are “digital libraries”, “archives” and “digitization”. In our opinion, this reflects and reinforces the cohesion of the knowledge base in which DH and LIS converge.

Although we located more indexing terms because the greatest number of studies corresponds to papers from the HDH congresses section, we believe that there is a great thematic correspondence across the conference papers and journal articles; in the research presented in the two categories, “digital publishing,” “corpora” and “databases” must be added to the common subjects already mentioned.

The relevance and applications of “Geographic Information Systems,” meanwhile, are raised in three doctoral theses and three conference papers.

Regarding standards and methodologies, we found a great similarity between the conference papers and the doctoral theses, except over the use of “TEI,” which does not appear in the latter but was nevertheless very relevant at the congresses, as well as in two of the journal articles. Looking at all of the scientific output that we assessed, we can observe three large sets in which the interrelation of DH and LIS clearly manifests itself: library and archives standards and techniques, metadata standards and the Semantic Web.

Within the area of libraries and archives standards and techniques we observed the following descriptors: “MARC format,” “authority control,” “controlled vocabularies,” “thesauri,” “ISAD (G),” “information management,” “preservation” and “taxonomies.” Appearing in the set of metadata standards are: “Dublin Core,” “OAI-PMH” and “EDM.” Finally, we might mention a special interest in technologies associated with the Semantic Web, regarding which we found: “Semantic Web,” “XML,” “RDF,” ontologies” and “linked open data.” The presence of keywords linked to semantic technologies is also found in the bibliometric studies we have cited here (Tang *et al.*, 2017; Wang, 2018), without doubt, indicating the transcendence of these technologies in DH research and in the visibility and reuse of cultural heritage (Thorsen and Pattuelli, 2016).

## Conclusions

Through this study, we identified research emanating from Spain and published between 2013 and 2018 in which research objects and methodologies connected to LIS and Digital Humanities converge.

Without neglecting the limitations detailed in the Discussion part of the previous section, we seem to have observed a scarce volume of scientific literature that supports the confluence of LIS and DH. It should be taken into account that in the study by Wang (2018) the category of Information Science and Library Science is indicated within the generic domain of DH in WoS in second position, after Computer Science. We focused on literature potentially belonging to that first thematic category, but the connection does not seem to be as broad as one would expect. We believe that the much more recent emergence of DH research in Spain can perhaps partly explain this situation.

The results highlight the synergies between documentary heritage, Library and Information Science and digital humanities. We believe that these synergies corroborate the idea of a Heritage Science community expressed by Melissa Terras, (2016). However, at the moment, our study shows greater collaboration between humanities researchers and areas of computer engineering in comparison to the expected collaboration between humanities specialists and LIS experts.



With respect to the observable documentary techniques and standards in the collection under analysis, we must undoubtedly highlight the major presence of Semantic Web technologies (interoperability, metadata, ontologies and so on). To a lesser but certainly relevant extent, we detected the application of controlled vocabularies, thesauri, authority control, use of records in MARC format, digitization techniques and so on, which clearly supports concomitance with LIS techniques. Regarding this point about the application of software, methods, documentary standards and so forth, an important field of work can be observed, one that has so far been very much neglected owing to the limited collaboration detected between DH researchers and their LIS counterparts.

The three most frequently used heritage collections, namely the National Library of Spain (or its subcollections, and especially the Hispanic Digital Library), the Miguel de Cervantes Virtual Library and the Spanish Bibliographic Heritage Collective Catalogue, are among the most significant sets of heritage items. However, despite the digitization policies coordinated by the Spanish Ministry of Culture for decades (Carrato Mena, 2008) and even more recent and innovative programmes such as the 2015-2020 Strategic Plan (Biblioteca Nacional de España, 2015), we observed few formal links with documentary institutions in the study. This leads us to consider the existence of a limited “scientific relationship” between heritage institutions and the academic sector that uses and enriches these source collections.

There is an urgent need to strengthen synergies and “scientific relations” between heritage institutions and DH research groups that could converge in the development of projects and applications that strengthen the value of digital collections in the transformation of humanities research and that also result in greater visibility for documentary heritage and the institutions responsible for preservation of it. Along these lines, digitization policies are the result of agreements reached between heritage institutions and humanities research groups and aimed at selecting the documents to be digitized. It is crucial to strengthen the relationships with heritage institutions as well as enhancing links between the academic sphere of DH and LIS in order to improve teaching and research in conjunction. A genuinely robust form of collaboration is needed to obtain the maximum potential from collections and strengthen research in both disciplines.

Likewise, we agree with Koya and Chowdhury (2019) when they emphasize that the field of cultural heritage informatics is evolving a great deal and requires a balanced set of competencies from various disciplines at a theoretical and applied level. This may require a rethink of some of current study curricula as well as promotion of the formation of new interdisciplinary research groups that serve the shared interests of DH and LIS. We are in agreement with Poole and Garwood (2018b) who claim that LIS educators align their curricula regarding current and future research needs and employment opportunities.

Finally, we believe that this work could be supplemented through analysis of Spanish research projects within the digital humanities to ascertain possible lines of joint work, real interdisciplinarity within teams and the possible collaboration of heritage institutions—as well as of staff from academic libraries—within them.

## Notes

[1] HDH: *Humanidades Digitales Hispánicas. Sociedad Internacional* (Hispanic Digital Humanities: International Society) <https://humanidadesdigitaleshispanicas.es/>

[2] <https://www.educacion.gob.es/teseo/>

[3] <http://hd.paulspence.org/recursos/hh-dd-es/>

[4] <https://web.archive.org/web/20131031113017/http://hdh2013.humanidadesdigitales.org/>

[5] <https://www.janusdigital.es/anexo/descargar.htm?id=5>

[6] <https://humanidadesdigitaleshispanicas.es/contribuciones/congreso-hdh-2015/>

[7] <https://humanidadesdigitaleshispanicas.es/contribuciones/congreso-hdh-2015/>

[8] <https://humanidadesdigitaleshispanicas.es/contribuciones/congreso-hdh-2017/>

[9] <https://humanidadesdigitaleshispanicas.es/contribuciones/congreso-hdh-2017/>

- [10] <http://www.cervantesvirtual.com/>  
 [11] <http://www.bne.es/>  
 [12] <http://www.bne.es/es/Catalogos/BibliotecaDigitalHispanica/Inicio/>  
 [13] <http://revistacaracteres.net/>  
 [14] <http://scriptumdigital.org/index.php>  
 [15] <http://revistas.uned.es/index.php/RDH/index>

## References

- Alvite-Díez, M.-L. and Pena Sueiro, N. (2020), “Colecciones digitales patrimoniales especializadas. Estudio de la Red ARACNE”, in Tramullas, J., Garrido-Picazo, P. and Marco-Cuenca, G. (Eds.), *Actas Del IV Congreso ISKO España-Portugal 2019 (XIV Congreso ISKO España)*, Sociedad Internacional para la Organización del Conocimiento (ISKO) – Capítulo Ibérico., Zaragoza, pp. 185–195.
- Balck, S. (2016), “(X)Disziplinarität der Informationswissenschaft”, *LIBREAS. Library Ideas*, Humboldt-Universität zu Berlin, Vol. 30, available at:<https://doi.org/http://dx.doi.org/10.18452/9105> (accessed 25 June 2020).
- Biblioteca Nacional de España. (2015), *Plan Estratégico 2015-2020*, Madrid, available at: [http://www.bne.es/export/sites/BNWEB1/webdocs/LaBNE/PlanEstrategico/Plan\\_estrategico\\_2015-2020.pdf](http://www.bne.es/export/sites/BNWEB1/webdocs/LaBNE/PlanEstrategico/Plan_estrategico_2015-2020.pdf) (accessed 25 June 2020).
- Borgman, C.L. (2009), “The Digital Future is Now: A Call to Action for the Humanities”, *Digital Humanities Quarterly*, Vol. 3 No. 4, available at: <http://digitalhumanities.org//dhq/vol/3/4/000077/000077.html> (accessed 25 June 2020).
- Bradley, A.J., El-Assady, M., Coles, K., Alexander, E., Chen, M., Collins, C., Janicke, S. and Wrisley, D. J. (2018), “Visualization and the Digital Humanities”: IEEE Computer Graphics and Applications, Vol. 38 No. 6, pp. 26–38.
- Burghardt, M., Wolff, C. and Womser-Hacker, C. (2015), “Informationswissenschaft und Digital Humanities”, *Information - Wissenschaft & Praxis*, Vol. 66 No. 5–6, available at:<https://doi.org/10.1515/iwp-2015-0053>.
- Candela, G., Escobar, P., Carrasco, R.C. and Marco-Such, M. (2019), “A linked open data framework to enhance the discoverability and impact of culture heritage”, *Journal of Information Science*, Vol. 45 No. 6, pp. 756–766.
- Carrato Mena, M.A. (2008), “Las colecciones digitales en España: situación actual y perspectivas de futuro”, *El Profesional de La Información*, Vol. 17 No. 4, pp. 418–421.
- Clement, T., Hagenmaier, W. and Levine Knies, J. (2013), “Toward a Notion of the Archive of the Future: Impressions of Practice by Librarians, Archivists, and Digital Humanities Scholars”, *The Library Quarterly*, Vol. 83 No. 2, pp. 112–130.
- Dalbello, M. (2011), “A genealogy of digital humanities”, *Journal of Documentation*, Vol. 67 No. 3, pp. 480–506.
- Frank, I. (2016), “Fortschritt durch Rückschritt”, *LIBREAS. Library Ideas*, Humboldt-Universität zu Berlin, Vol. 30, available at:<https://doi.org/http://dx.doi.org/10.18452/9108> (accessed 25 June 2020).
- Garwood, D.A. and Poole, A.H. (2019), “Pedagogy and public-funded research: an exploratory study of skills in digital humanities projects”, *Journal of Documentation*, Vol. 75 No. 3, pp. 550–576.

- Given, L.M. and Willson, R. (2018), "Information technology and the humanities scholar: Documenting digital research practices", *Journal of the Association for Information Science and Technology*, Vol. 69 No. 6, pp. 807–819.
- Gladney, H.M. (2012), "Long-Term Digital Preservation: A Digital Humanities Topic?", *Historical Social Research / Historische Sozialforschung*, GESIS - Leibniz-Institute for the Social Sciences, Center for Historical Social Research, Vol. 37 No. 3, pp. 201–217.
- Hobohm, H.-C. (2016), "Transdisziplinäre Aspekte der Informationswissenschaften als Kernaufgaben der Informationsberufe", *Informationswissenschaft: Theorie, Methode Und Praxis*, Vol. 4 No. 1, available at: <https://doi.org/10.18755/iw.2016.4> (accessed 25 June 2020).
- Hockey, S. (2004), "The History of Humanities Computing", in Schreibman, S., Siemens, R. and Unsworth, J. (Eds.), *A Companion to Digital Humanities*, Blackwell, Oxford, available at: <http://www.digitalhumanities.org/companion/>.
- Hockey, S. (2006), "The rendering of humanities information in a digital context: Current trends and future developments", *Aslib Proceedings: New Information Perspectives*, Vol. 58 No. 1–2, pp. 89–101.
- Iribarren, T. (2017), "Las Humanidades Digitales", in Gómez Díaz, R., García Rodríguez, A. and Cordón García, J.A. (Eds.), *Fuentes Especializadas en Ciencias Sociales y Humanidades*, Pirámide, Madrid, pp. 440–473.
- Koltay, T. (2016), "Library and information science and the digital humanities: Perceived and real strengths and weaknesses", *Journal of Documentation*, Vol. 72 No. 4, pp. 781–792.
- Koya, K. and Chowdhury, G. (2020), "Cultural Heritage Information Practices and iSchools Education for Achieving Sustainable Development", *Journal of the Association for Information Science and Technology*, Vol. 71 No. 6, pp. 696–710.
- Lafuente Urién, A. (2016), "PARES hacia la web semántica: autoridades y tesauros en los archivos estatales", in *El Lenguaje sobre el Patrimonio. Estándares Documentales para La Descripción y Gestión de Colecciones*, Ministerio de Educación, Cultura y Deporte, Madrid, pp. 181–194.
- López Poza, S. (2015), "Humanidades digitales y literaturas hispánicas: presente y futuro", *Insula: Revista de Letras y Ciencias Humanas*, No. 822, pp. 3–5.
- Oldman, D., Doerr, M. and Gradmann, S. (2015), "Zen and the Art of Linked Data", *A New Companion to Digital Humanities*, John Wiley & Sons, Chichester, UK, pp. 251–273.
- Poole, A.H. (2017), "The conceptual ecology of digital humanities", *Journal of Documentation*, Vol. 73 No. 1, pp. 91–122.
- Poole, A.H. and Garwood, D.A. (2018a), "Interdisciplinary scholarly collaboration in data-intensive, public-funded, international digital humanities project work", *Library and Information Science Research*, Vol. 40 No. 3–4, pp. 184–193.
- Poole, A.H. and Garwood, D.A. (2018b), "'Natural allies': Librarians, archivists, and big data in international digital humanities project work", *Journal of Documentation*, Vol. 74 No. 4, pp. 804–826.
- Posner, M. (2013), "No Half Measures: Overcoming Common Challenges to Doing Digital Humanities in the Library", *Journal of Library Administration*, Vol. 53 No. 1, pp. 43–52.
- Robinson, L., Priego, E. and Bawden, D. (2015), "Library and Information Science and Digital Humanities: Two Disciplines, Joint Future?", in Pehar, F., Schlögl, C. and Wolff, C. (Eds.), *Re:Inventing Information Science in the Networked Society. Proceedings of the 14th*

- International Symposium on Information Science (ISI 2015), Zadar, Croatia, 19th–21st May 2015*, Verlag Werner Hülsbusch, Glückstadt, pp. 44–54.
- Rockenbach, B. (2013), “Introduction. Digital Humanities in Libraries: New Models for Scholarly Engagement”, *Journal of Library Administration*, Vol. 53 No. 1, pp. 1–9.
- Rodríguez-Ortega, N. (2018), “Five central concepts to think of Digital Humanities as a new digital humanism project”, *Artnodes*, No. 22, pp. 1–6.
- Rodríguez Yunta, L. (2014), “Ciberinfraestructura para las humanidades digitales: una oportunidad de desarrollo tecnológico para la biblioteca académica”, *El Profesional de La Información*, Vol. 23 No. 5, pp. 453–462.
- Rojas Castro, A. (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*, Vol. 2 No. 2, pp. 10–53.
- Romero López, D. (2018), “La Edad de Plata en la tarima digital. Hacia la transdisciplinariedad y la cultura smart”, *Artnodes*, No. 22, pp. 112–119.
- Sánchez Mairena, A. (2017), “Pares 2.0”, *Archivamos: Boletín ACAL*, No. 103, pp. 34–35.
- Schöch, C. (2013), “Big? Smart? Clean? Messy? Data in the Humanities”, *Journal of Digital Humanities*, Vol. 2 No. 3, available at: <http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/> (accessed 25 June 2020).
- Smith-Yoshimura, K. (2018), “Analysis of 2018 International Linked Data Survey for Implementers”, *Code4lib Journal*, No. 42, available at: <https://journal.code4lib.org/articles/13867> (accessed 25 June 2020).
- Spence, P. (2014), “Centros y fronteras: el panorama internacional”, in López Poza, S. and Pena Sueiro, N. (Eds.), *Humanidades Digitales: Desafíos, Logros y Perspectivas de Futuro*, A Coruña: Universidade da Coruña, SIELAE, pp. 37–61.
- Spence, P. and González-Blanco García, E. (2014), “A historical perspective on the digital humanities in Spain”, *H-Soz-Kult*, pp. 1–19, available at: <https://www.hsozkult.de/debate/id/diskussionen-2449?language=en> (accessed 25 June 2020).
- Steiner, C.M., Agosti, M., Sweetnam, M.S., Hillemann, E.-C., Orio, N., Ponchia, C., Hampson, C., Munnely, G., Nussbaumer, A., Albert, D. and Conlan, O. (2014), “Evaluating a digital humanities research environment: the CULTURA approach”, *International Journal on Digital Libraries*, Vol. 15 No. 1, pp. 53–70.
- Sula, C.A. (2013), “Digital Humanities and Libraries: A Conceptual Model”, *Journal of Library Administration*, Vol. 53 No. 1, pp. 10–26.
- Svensson, P. (2014), “The Landscape of Digital Humanities”, *Digital Humanities Quarterly*, Vol. 4 No. 1, available at: <http://digitalhumanities.org/dhq/vol/4/1/000080/000080.html> (accessed 25 June 2020).
- Tang, M.-C., Cheng, Y.J. and Chen, K.H. (2017), “A longitudinal study of intellectual cohesion in digital humanities using bibliometric analyses”, *Scientometrics*, Vol. 113 No. 2, pp. 985–1008.
- Terras, M. (2015), “Opening Access to collections: The making and using of open digitised cultural content”, *Online Information Review*, Vol. 39 No. 5, pp. 733–752.
- Terras, M. (2016), “A Decade in Digital Humanities”, *Journal of Siberian Federal University. Humanities & Social Sciences*, Vol. 9 No. 7, pp. 1637–1650.

- Terras, M., Baker, J., Hetherington, J., Beavan, D., Austwick, M.Z., Welsh, A., O'Neill, H., Finley, W., Duke-Williams, O. and Farquhar, A. (2018), "Enabling complex analysis of largescale digital collections: Humanities research, high-performance computing, and transforming access to British Library digital collections", *Digital Scholarship in the Humanities*, Vol. 33 No. 2, pp. 456–466.
- Thorsen, H.K. and Pattuelli, M.C. (2016), "Linked Open Data and the Cultural Heritage Landscape", in Jones, E. and Seikel, M. (Eds.), *Linked Data for Cultural Heritage*, Facet, London, pp. 1–22.
- Vila-Suero, D., Villazón-Terrazas, B. and Gómez-Pérez, A. (2013), "datos.bne.es: A library linked dataset", *Semantic Web*, Vol. 4 No. 3, pp. 307–313.
- Wang, Q. (2018), "Distribution features and intellectual structures of digital humanities: A bibliometric analysis", *Journal of Documentation*, Vol. 74 No. 1, pp. 223–246.
- White, J.W. and Gilbert, H. (Eds.). (2016), *Laying the Foundation: Digital Humanities in Academic Libraries*, Purdue University Press, West Lafayette, Indiana.
- Wilms, L., Derven, C., O'Dwyer, L., Lingstadt, K. and Verbeke, D. (2019), *Europe's Digital Humanities Landscape: A Study From LIBER's Digital Humanities & Digital Cultural Heritage Working Group*, available at:<https://doi.org/10.5281/ZENODO.3247286> (accessed 25 June 2020).
- Zeng, M.L. (2017), "Smart Data for Digital Humanities", *Journal of Data and Information Science*, Vol. 2 No. 1, pp. 1–12.
- Zeng, M.L. (2019), "Semantic enrichment for enhancing LAM data and supporting digital humanities . Review article", *El Profesional de La Información*, Vol. 28 No. 1, pp. 1–35.
- Zhang, Y., Liu, S. and Mathews, E. (2015), "Convergence of digital humanities and digital libraries", *Library Management*, Vol. 36 No. 4–5, pp. 362–377.