

Digital humanities in Spain: Historical perspective and current scenario

Maurizio Toscano; Aroa Rabadán; Salvador Ros; Elena González-Blanco

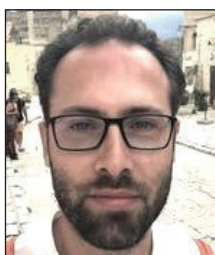
How to cite this article:

Toscano, Maurizio; Rabadán, Aroa; Ros, Salvador; González-Blanco, Elena (2020). "Digital humanities in Spain: Historical perspective and current scenario". *Profesional de la información*, v. 29, n. 6, e290601.

<https://doi.org/10.3145/epi.2020.nov.01>

Manuscript received on 8th May 2020

Accepted on 13th July 2020



Maurizio Toscano ✉

<https://orcid.org/0000-0001-5418-3859>

Universidad de Granada
Escuela Internacional de Posgrado
Avenida de Madrid, 13
18071 Granada, Spain
maurizio.toscano@gmail.com



Aroa Rabadán

<https://orcid.org/0000-0001-6830-8827>

Universidad Complutense de Madrid
Facultad de Filología
Plaza Menéndez Pelayo, Edificio A
28040 Madrid, Spain
aroaraba@ucm.es



Salvador Ros

<https://orcid.org/0000-0001-6330-4958>

Universidad Nacional de Educación a Distancia (UNED). ETSI de Informática
Juan del Rosal, 16
28040 Madrid, Spain
sros@scc.uned.es



Elena González-Blanco

<https://orcid.org/0000-0002-0448-1812>

IE Business School
Human Sciences and Technology
María de Molina, 4
28006 Madrid, Spain
egonzalezblanco@faculty.ie.edu

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.

Keywords

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

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 with 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 *Clarín*. 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 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). Applying this approach to digital humanities in Spain is 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.

“ 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 ”

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.

“ 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 ”

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 intention 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 publicly funded research projects. Firstly, bibliography was extracted by looking for the keyword “humanidades digitales” [digital humanities] both in *Dialnet* and in *ÍNDICES-CSIC*⁴. 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, TEI, 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.

“ 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* ”

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.

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

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

As can be observed, both researchers and projects have a key data source, respectively bibliography and *Agencia Estatal de Investigación (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, as well as the procedures used to analyse it. <https://github.com/linhd-postdata/desir>

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 DOI and keep track of future versions.

<https://doi.org/10.5281/zenodo.3893546>

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.

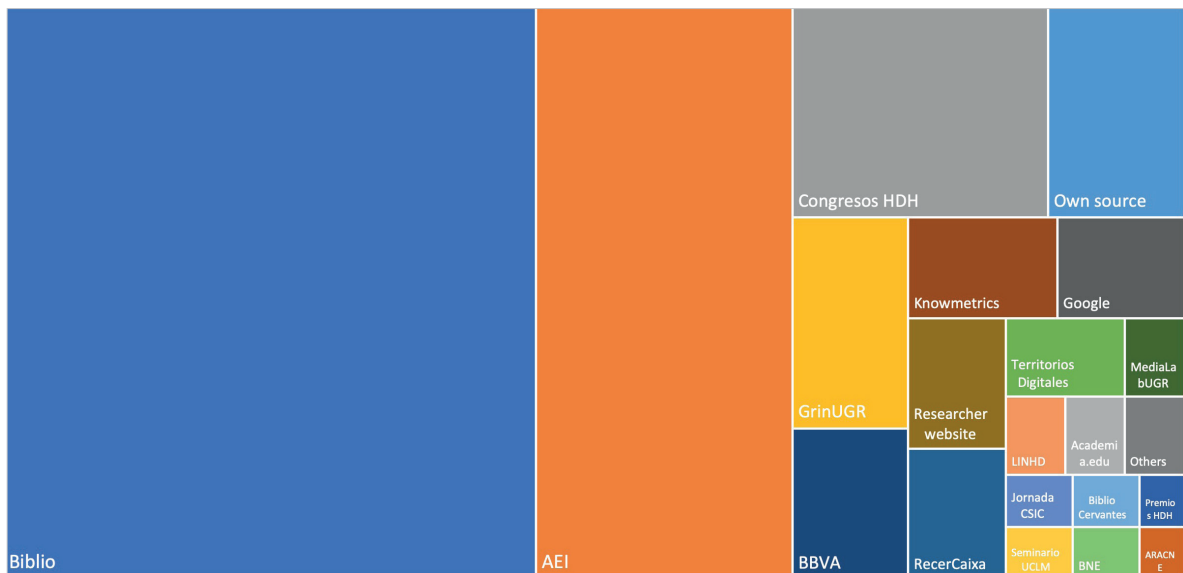


Figure 1. Data sources presented as a treemap chart, weighted according to their contribution to the researcher table.

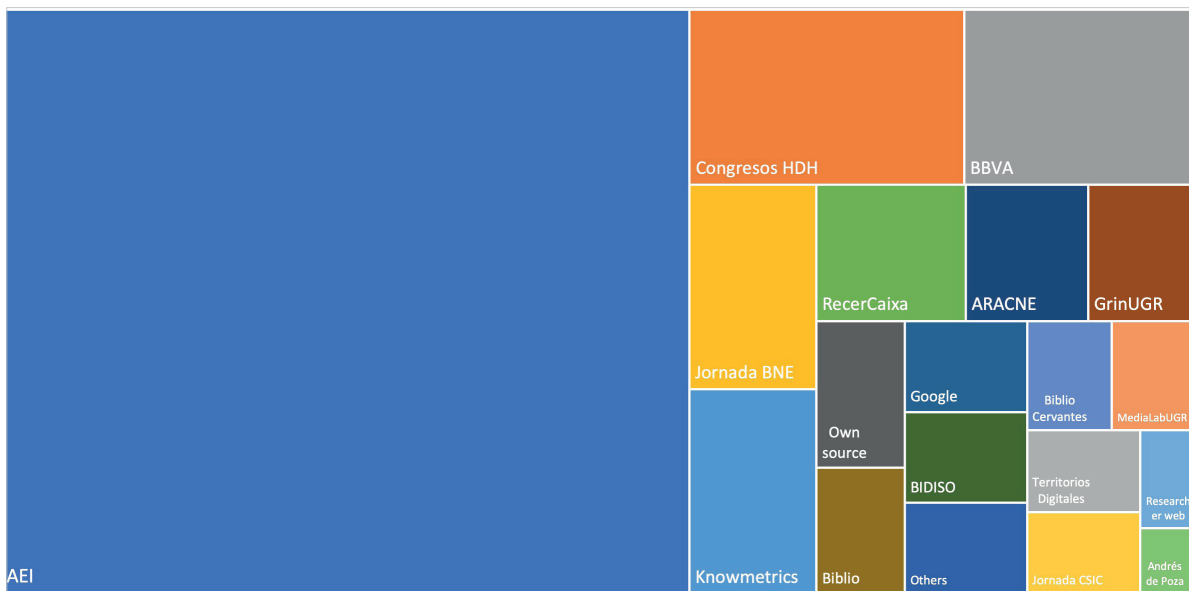


Figure 2. Data sources presented as a treemap chart, weighted according to their contribution to the project table.

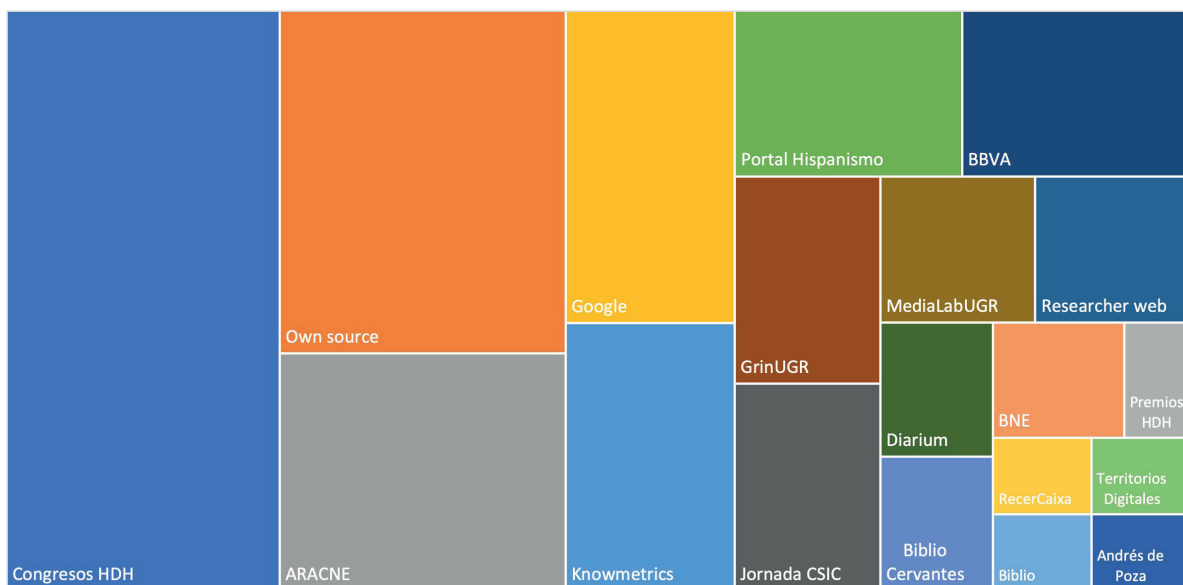


Figure 3. Data sources presented as a treemap chart, weighted according to their contribution to the resource table.

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; 10 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 expected⁶ (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 2001.

<https://www.ine.es/jaxi/Datos.htm?path=/t14/p057/64-01/10/&file=d1004.px>

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 2019⁷.

Table 2. Total researchers and gender percentage in Spain for all disciplines, humanities and digital humanities.

| Researchers in Spain | Males: units (ratio) | Females: units (ratio) | Totals |
|----------------------------------|----------------------|------------------------|---------|
| All disciplines ⁸ | 85,762 (61.2%) | 54,358 (38.8%) | 140,120 |
| Humanities ⁹ | 11,149 (59.8%) | 7,501 (40.2%) | 18,650 |
| Digital humanities ¹⁰ | 305 (52.9%) | 272 (47.1%) | 577 |

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%).

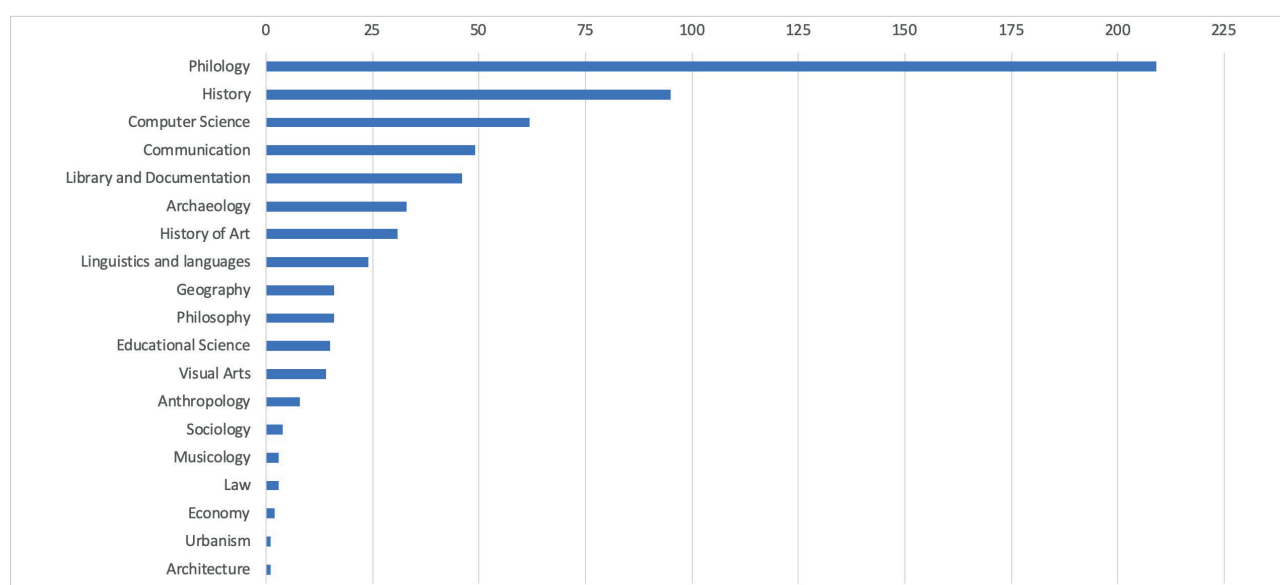


Figure 4. Disciplinary backgrounds of researchers working in digital humanities in Spain.

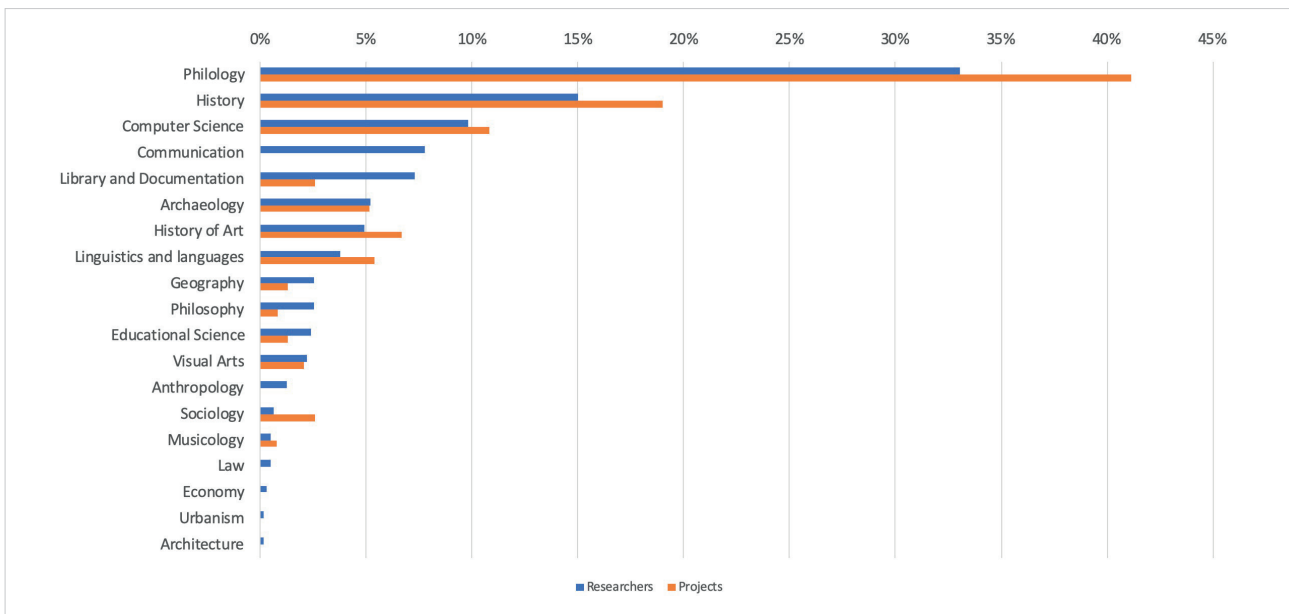


Figure 5. Comparison between researchers' and projects' disciplinary backgrounds.

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.

If this disciplinary classification is used to compare researchers with projects¹¹, 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).

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. Howe-

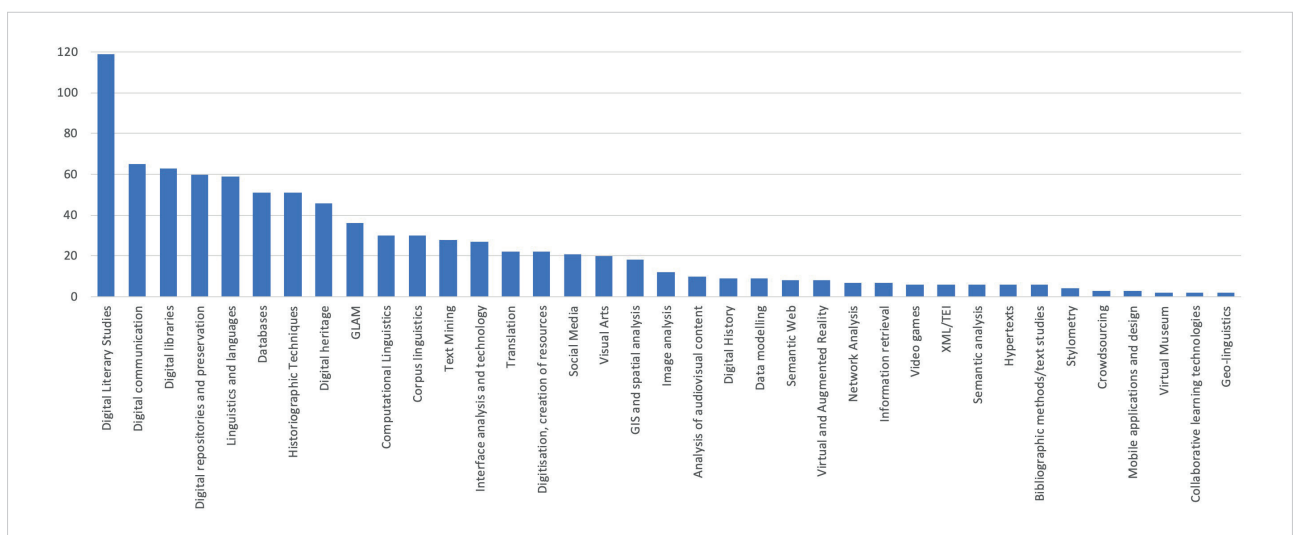


Figure 6. Researchers classified according to their preferred research topics, with up to five each.

ver, 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).

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 majors 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 ego-centric networks (Golbeck, 2013) for (A) Text mining; (B) XML/TEI; (C) Video games; and (D) GIS and spatial analysis (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 are 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.

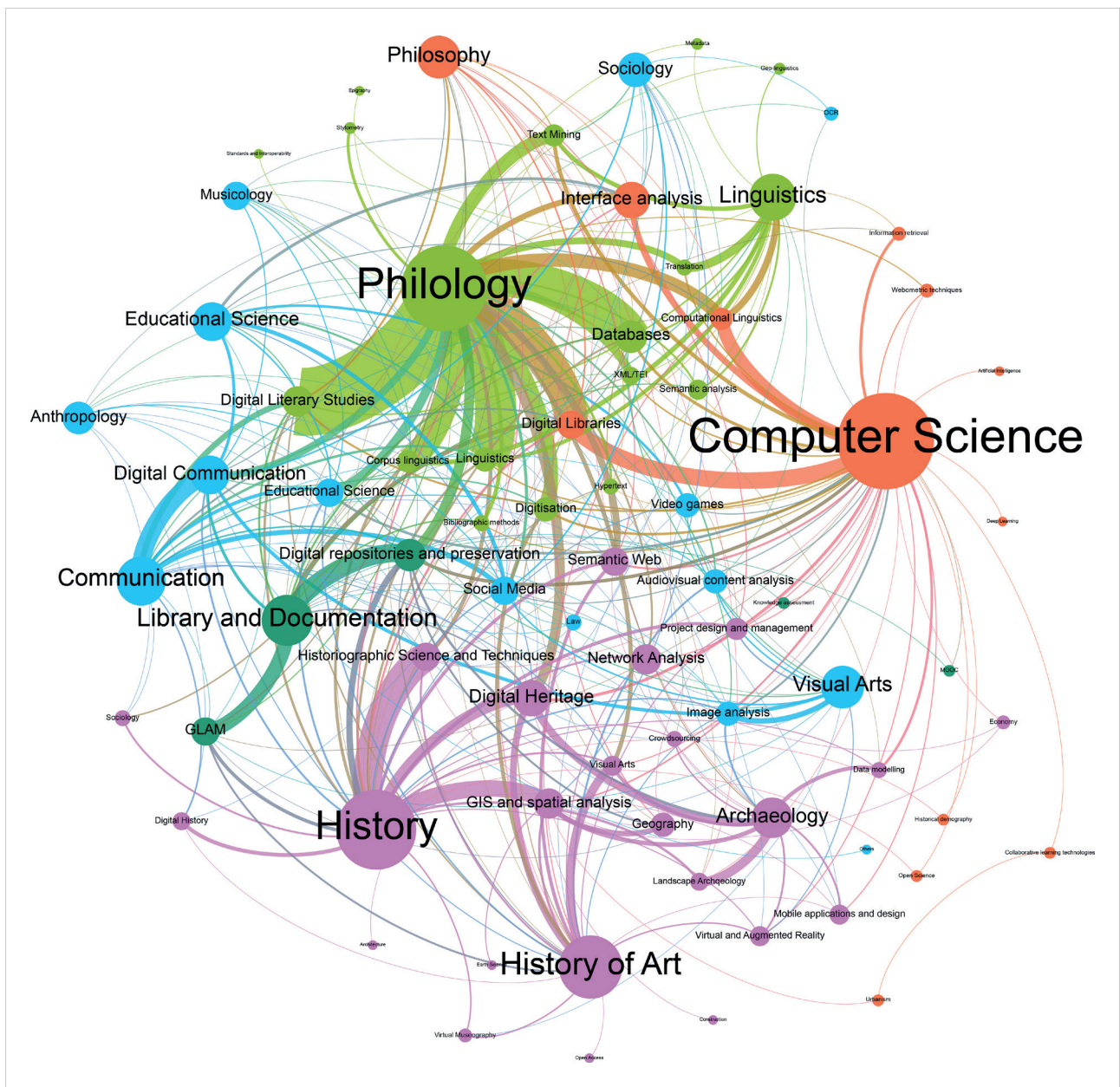


Figure 7. Network analysis of disciplines and research topics. Node size corresponds to the total number of connections, colour to the community, and edge thickness to the degree of connectivity between the two nodes.

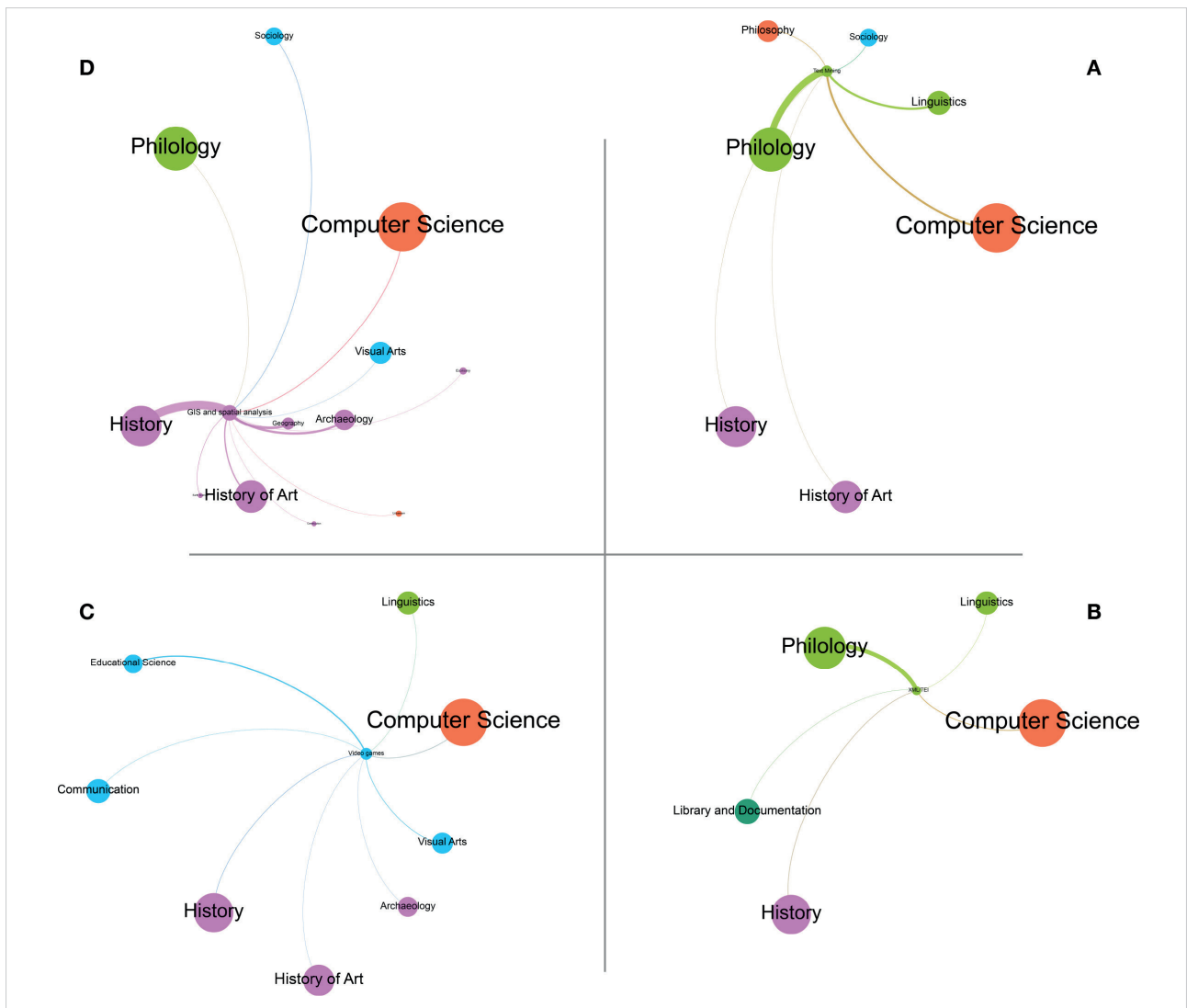


Figure 8. Ego networks derived from Figure 7, for four research topics: (A) Text mining; (B) XML/TEI; (C) Video Games; (D) GIS and spatial analysis.

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 apparently 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% represents 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.

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

| Research and higher education centre | # of researchers | Percentage |
|--|------------------|------------|
| <i>Universidad Complutense de Madrid</i> | 62 | 10.7 |
| <i>Universitat de Barcelona</i> | 41 | 7.1 |
| <i>Universidad de Granada</i> | 40 | 6.9 |
| <i>Universidad de Santiago de Compostela</i> | 28 | 4.8 |
| <i>Universidad Nacional de Educación a Distancia (UNED)</i> | 28 | 4.8 |
| <i>Consejo Superior de Investigaciones Científicas (CSIC)¹²</i> | 26 | 4.5 |
| <i>Universitat Autònoma de Barcelona</i> | 22 | 3.8 |
| <i>Universidad de Salamanca</i> | 19 | 3.3 |
| <i>Universidad de Sevilla</i> | 19 | 3.3 |
| Total | 285 | 49.2 |

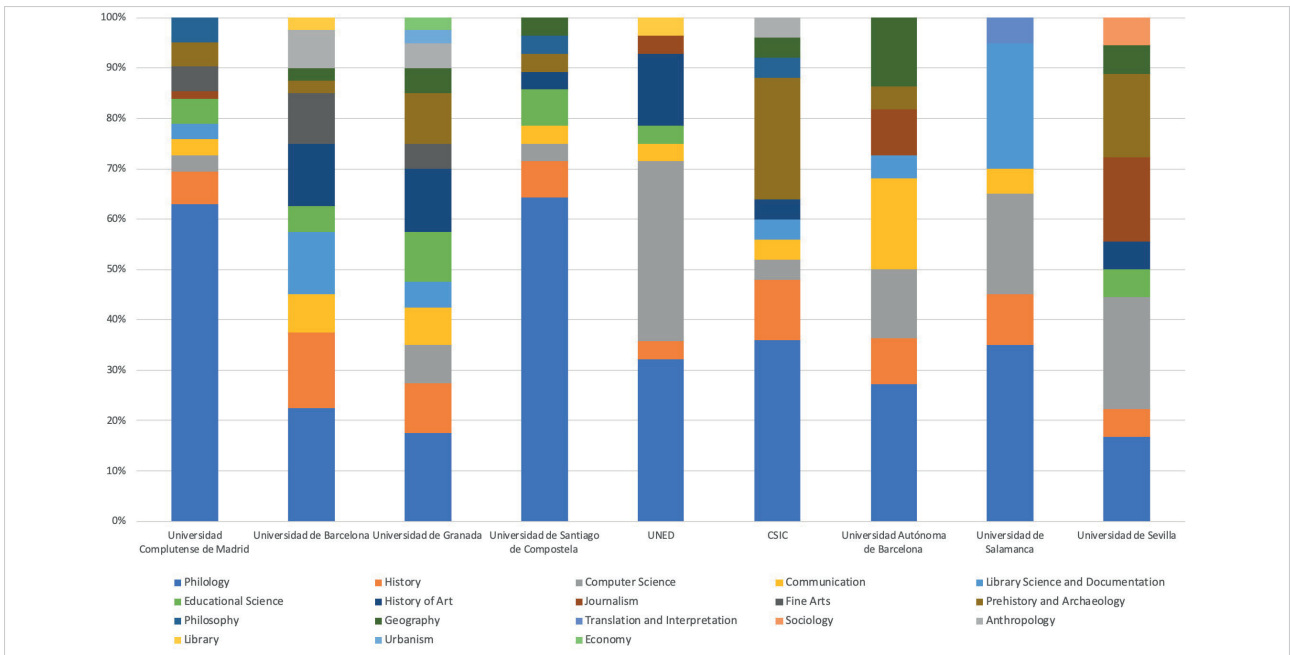


Figure 9. Relative weight of each department in the nine main research centres. To favour comparison, each group of values is stacked at 100%.

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)¹³. 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 departments.

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

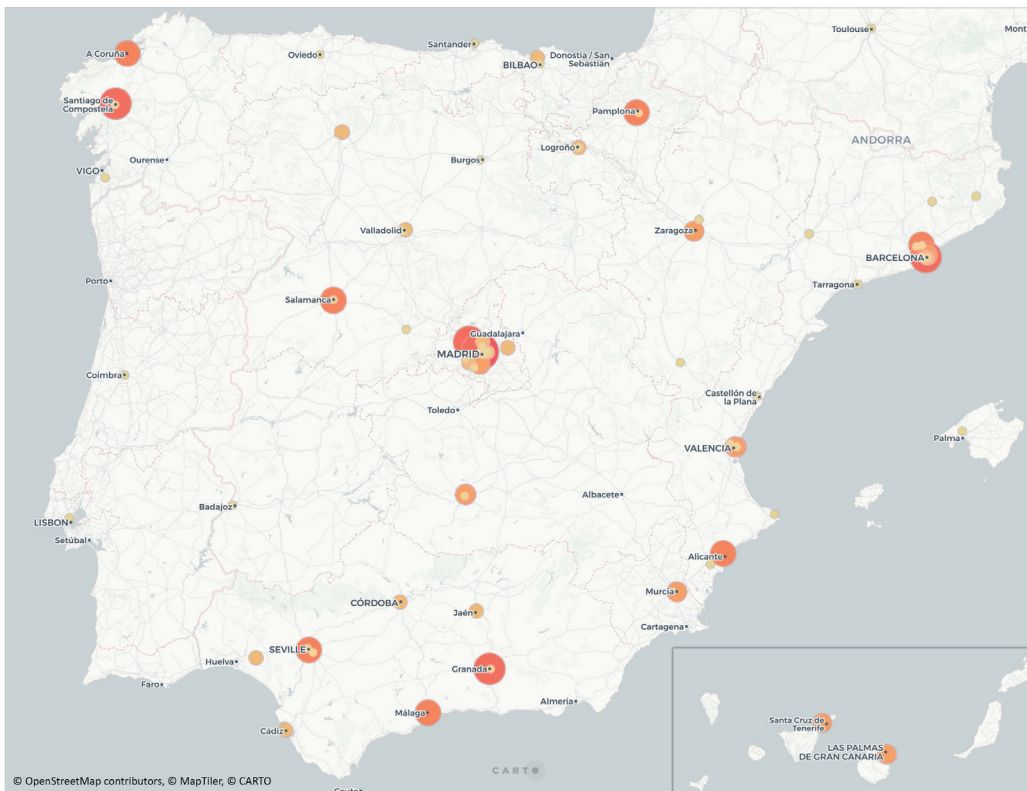


Figure 10. Overview of the distribution of researchers in Spain, produced by geocoding the affiliated institutions address and then creating centroids of points sharing the same location. Online map available at: <http://sl.ugr.es/DHmap>

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).

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 centres¹⁴. 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.

Table 4. Masters focused on digital humanities currently available in Spain.

| Master title | Organising institution | Since |
|---|---|-------|
| <i>Humanitats i patrimoni digitals</i> ¹⁵ | <i>Universitat Autònoma de Barcelona / Universitat Pompeu Fabra</i> | 2018 |
| <i>Máster en historia y humanidades digitales</i> | <i>Universidad Pablo de Olavide</i> | 2019 |
| <i>Tecnologías de la información para la sociedad digital: humanidades y ciencias</i> | <i>Universidad Nacional de Educación a Distancia (UNED)</i> | 2019 |
| <i>Language analysis and processing</i> | <i>Universidad del País Vasco</i> | 2019 |
| <i>Máster de gestión de contenidos digitales</i> | <i>Universitat de Barcelona</i> | 2019 |
| <i>Màster d'humanitats digitals</i> | <i>Universitat de Barcelona</i> | 2019 |
| <i>Máster en letras digitales: estudios avanzados en textualidades electrónicas</i> | <i>Universidad Complutense de Madrid</i> | 2019 |
| <i>Máster universitario en humanidades digitales</i> | <i>Universidad Internacional de La Rioja (UNIR)</i> | 2020 |
| <i>Humanidades digitales para un mundo sostenible</i> | <i>Universidad Autónoma de Madrid</i> | 2020 |
| <i>Máster en humanidades digitales</i> | <i>Universidad Nacional de Educación a Distancia (UNED)</i> | 2021 |

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 funding¹⁶. 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.

The same data classified by the starting date of each project helps to understand the evolution of funding and its relationship with 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

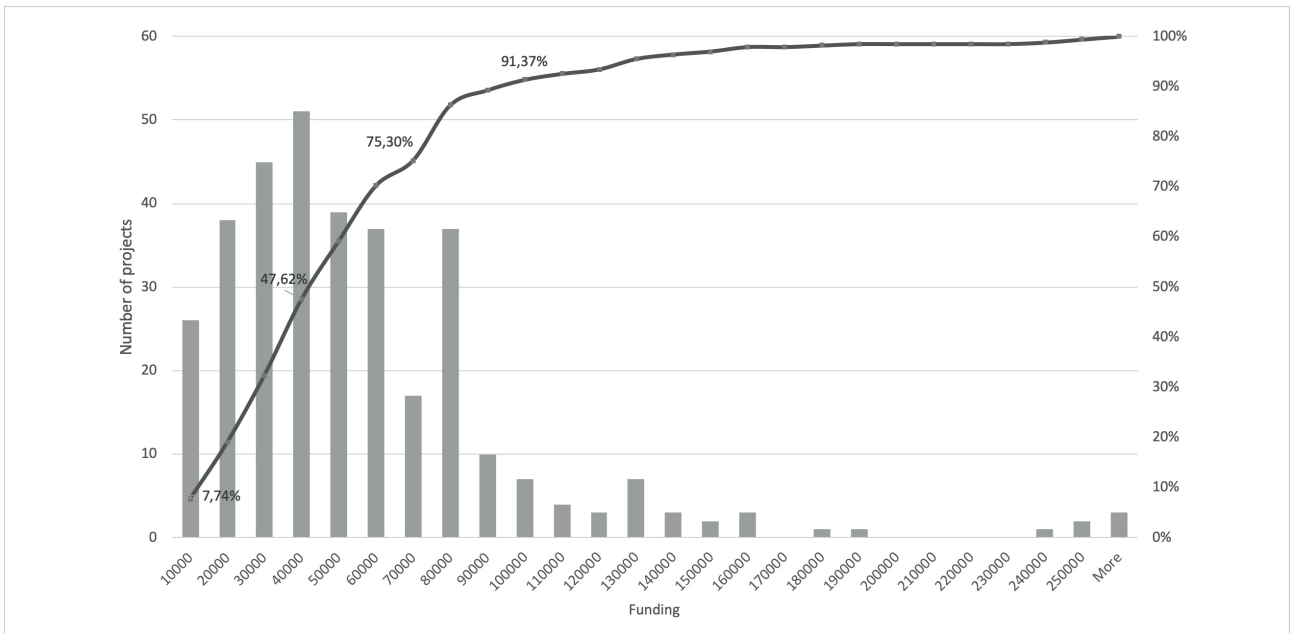


Figure 11. Distribution of projects according to funding received, with cumulative curve.

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 value¹⁷, 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 the 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 Postdata*, coordinated by the *UNED*:

<https://cordis.europa.eu/project/id/679528>

and by the *ERC Starting grant MapModern*, coordinated by the *UOC*:

<https://cordis.europa.eu/project/id/803860>

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

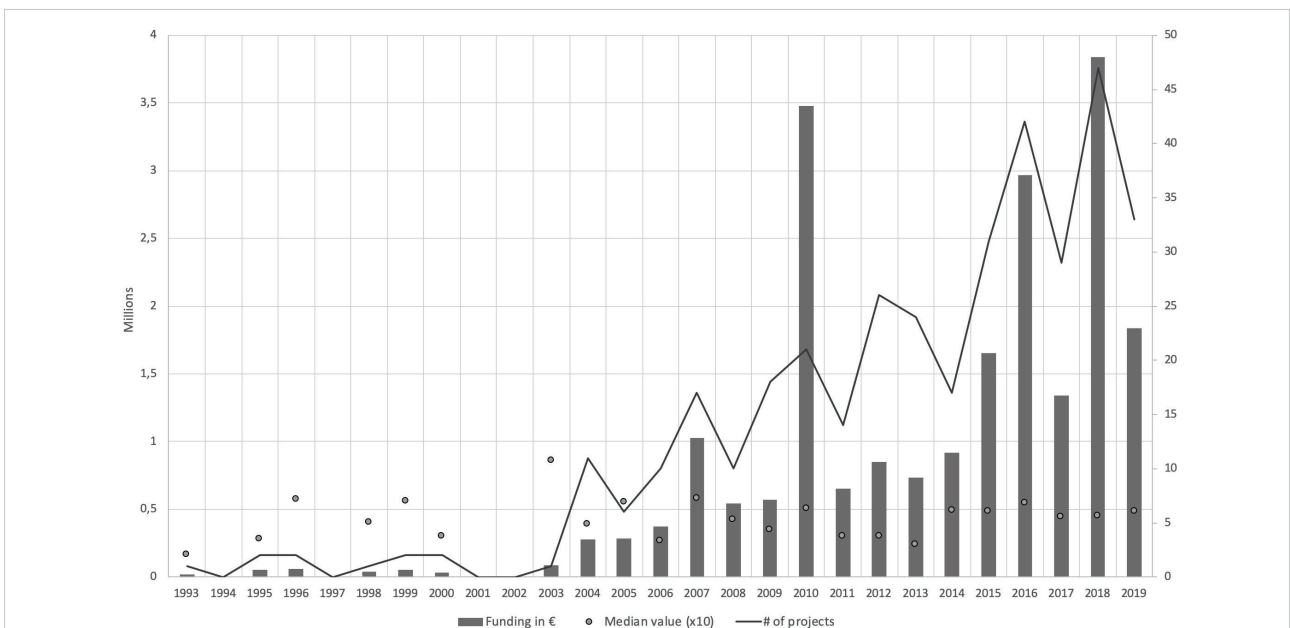


Figure 12. Number of research projects and total funding received during the period (1993-2019).

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 Marín, 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).

“ In the last five years, there seems to have been a supply-driven market that stimulated the emergence of research proposals in digital humanities, partly to take advantage of the funding opportunities ”

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* (López-Poza, 2014) was created, whose first international congress was held in 2013: <https://humanidadesdigitaleshispanicas.es>

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 Centre*, etc.): <https://www.clarin.eu/news/congratulations-our-first-clarin-knowledge-centre>

and with the publication of various studies (Rojas-Castro, 2013; González-Blanco, 2013; Ortega; Eunice-Gutiérrez, 2014; Baraibar-Echeverría, 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, Cataluña), 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 principle 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 contri-

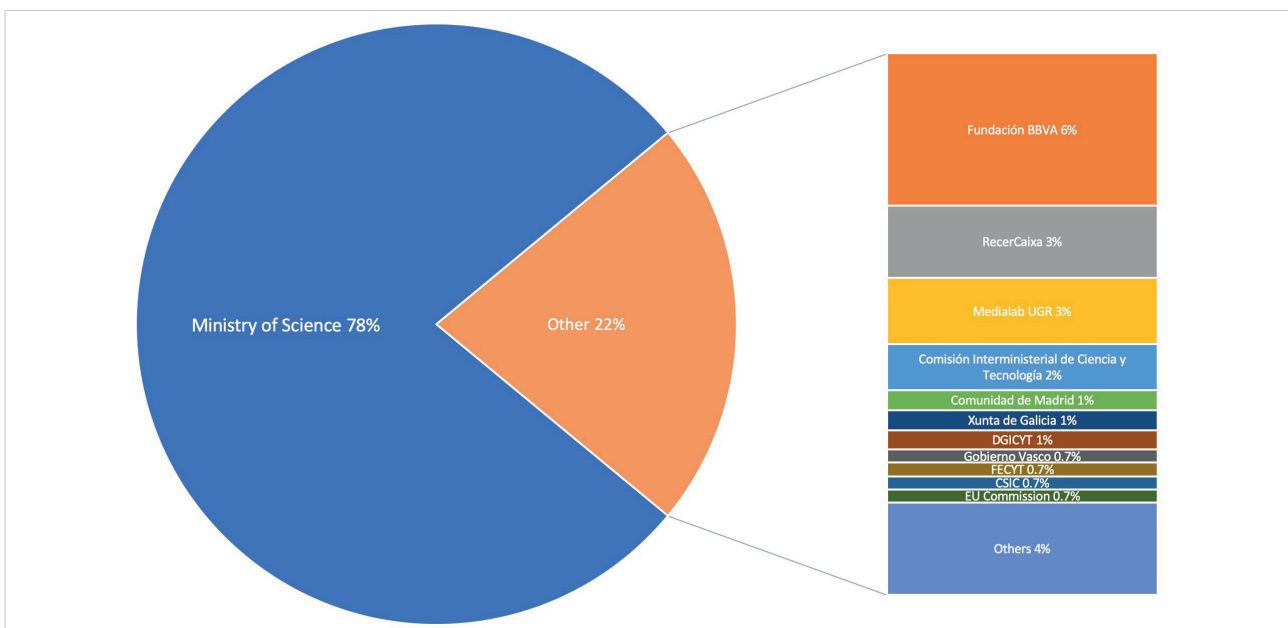


Figure 13. Funding agencies according to number of projects.

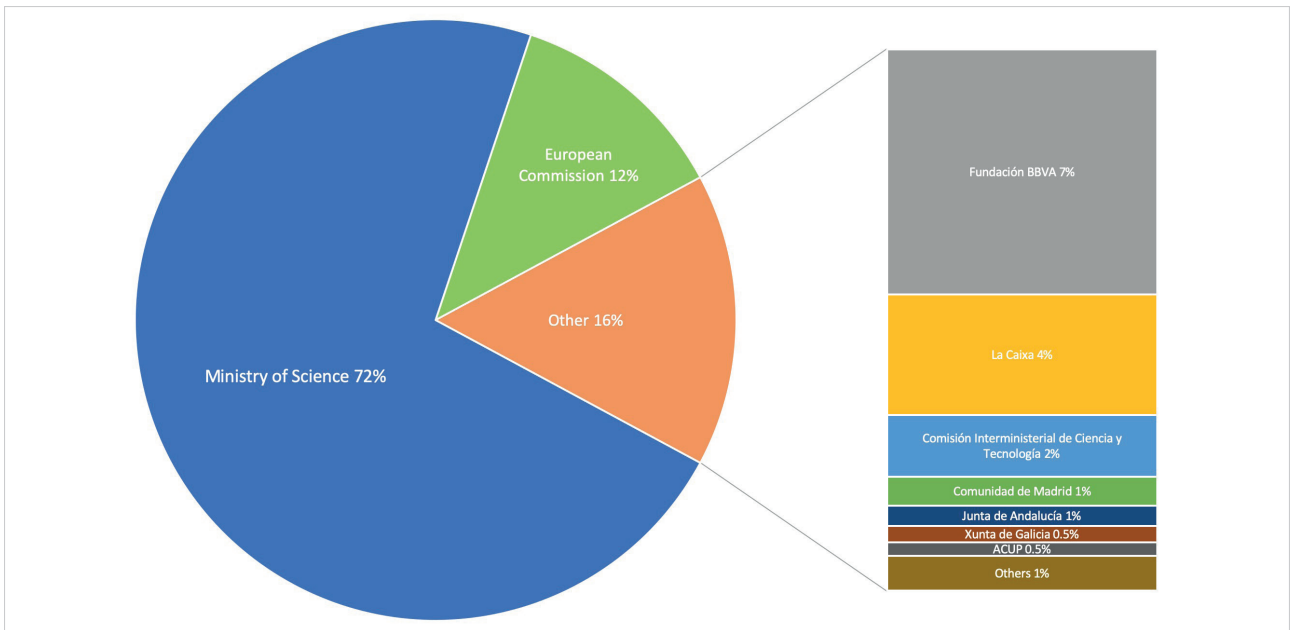


Figure 14. Funding agencies according to total funding.

butions 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 Science*¹⁸) 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.

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.

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 Casti-

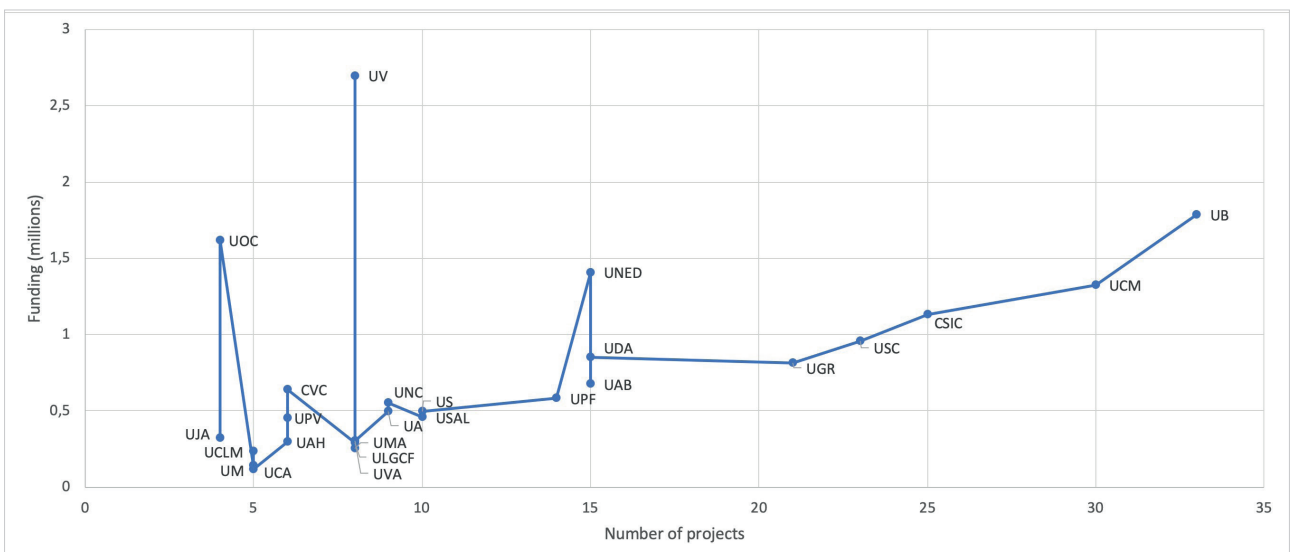


Figure 15. Scatterplot of first twenty-five research centres by number of projects and total funding obtained during the whole analysed period (1993-2019).

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

| 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 Cientificas | 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 Educacion 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 |

Ila 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.

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 (Ram-say, 2012; Enders, 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 that have been analysed against scholarly primitives for a poster presented at *Dariah Annual Event 2020* (Toscano et al., 2020). 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 *Clarín*, the two main European infrastructures dedicated to digital humanities.

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*

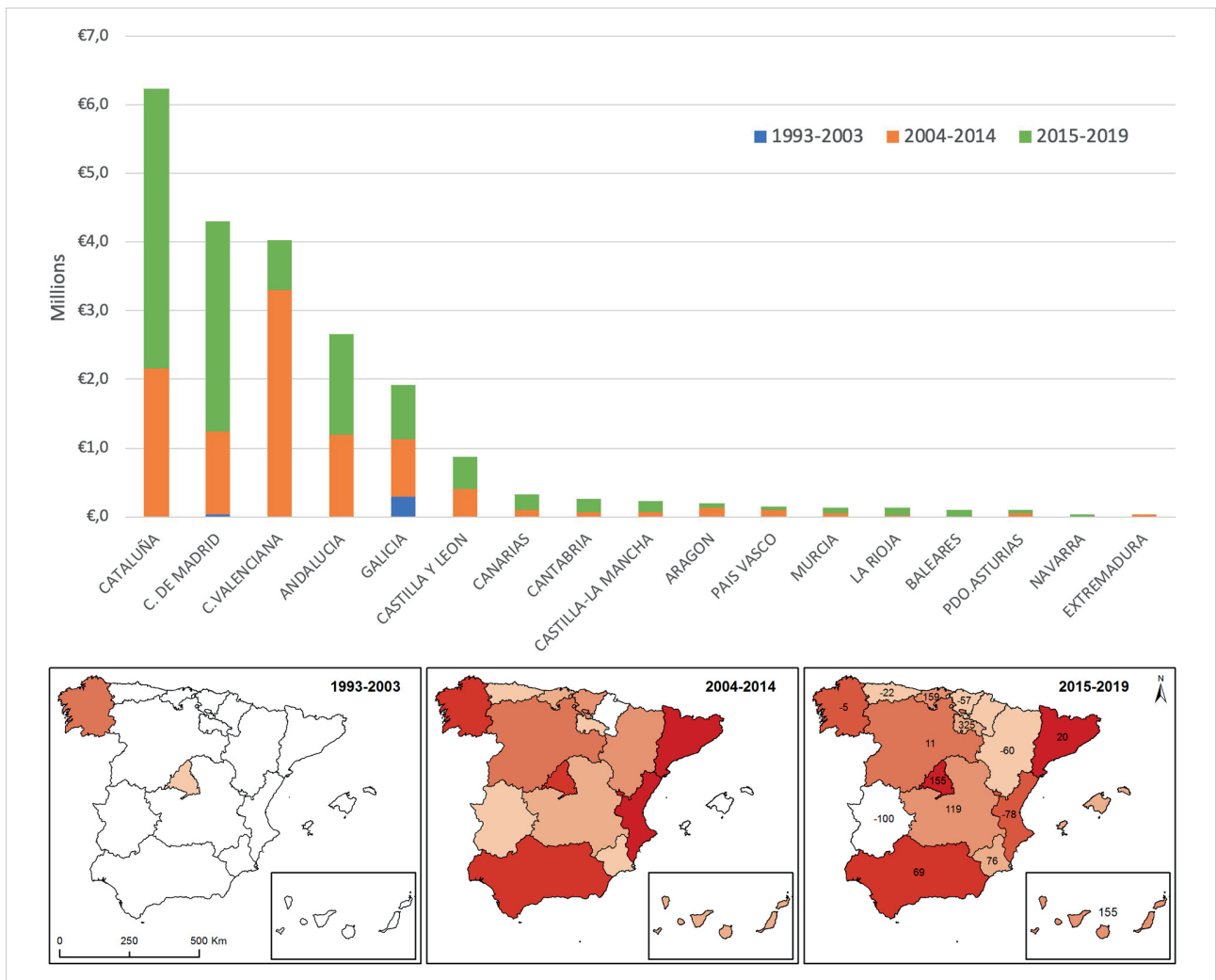


Figure 16. Geographical distribution of funding across Spanish regions during the previously identified three phases.

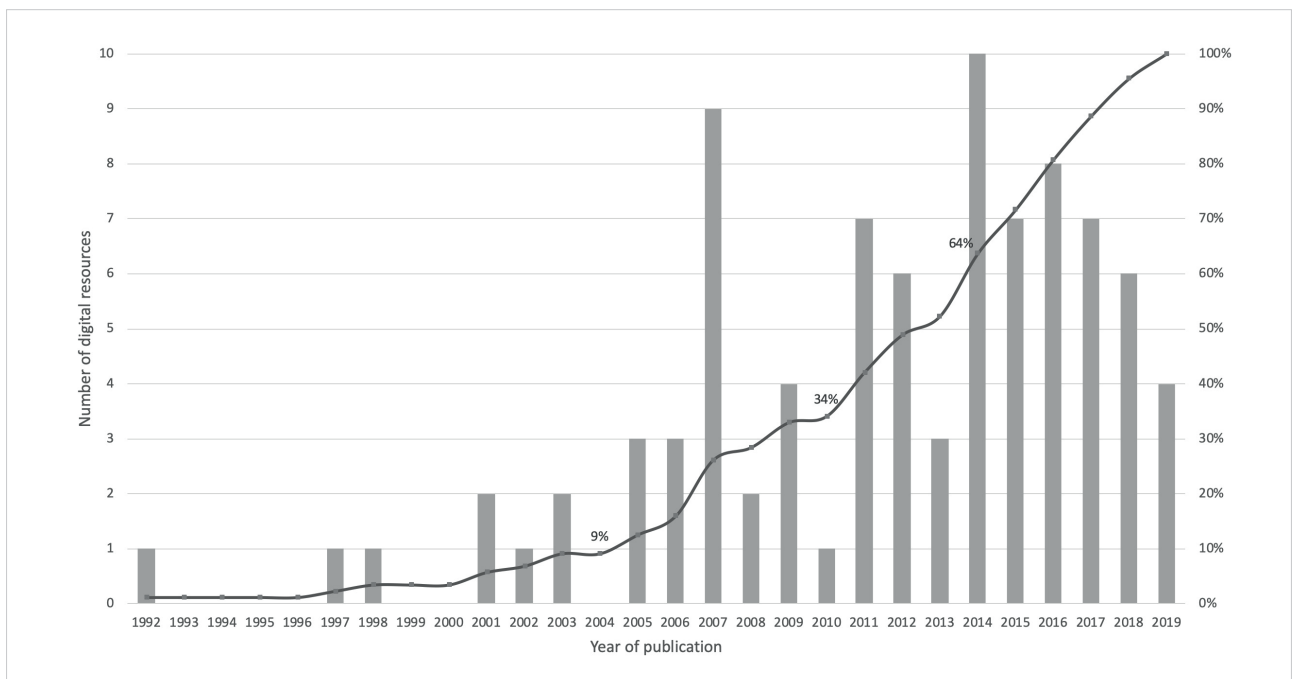


Figure 17. Histogram showing digital resources sorted by publishing date, with cumulative curve.

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.

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

| | 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 |

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 resources, 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 laboratories¹⁹ 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 or no 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 the highest groups in the 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 indexes²⁰ 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.

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

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

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-García; 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 *Clarín*. 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 *Clarín* 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/centers>

MapaHD (data have been merged into *GrinUGR*),
GrinUGR, <https://grinugr.org/open-data>

and, more recently, *Red Knowmetrics*
<https://red.knowmetrics.org/mapa>

among others.

2. In this regard, examples include the online resource *What are DH*:
<http://whatisdigitalhumanities.com>

and the *Twitter* bot @DHDefined, which randomly display 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. 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* (**Dallas et al., 2017**).

7. *Científicas en cifras 2017* (Puy-Rodríguez et al., 2019), Charts 1.11 and 1.12, data from 2016.
8. Total researchers in 2018:
<https://www.ine.es/consul/serie.do?s=186-24&c=2&nult=50>
- Female researchers in 2018:
<https://www.ine.es/consul/serie.do?s=186-26>
9. 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).
10. Data coming from our research (referring to 2019).
11. Most projects came already tagged according to *AEI* classification, although some matching has been necessary to consolidate a common vocabulary.
12. 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.
13. For the *CSIC*, for consistency with universities departments, we used researchers' disciplinary backgrounds instead of research centres.
14. An overview of the masters currently available in Spain is provided by the *Humanidades Digitales Hispánicas* on its website:
<https://humanidadesdigitaleshispanicas.es/formacion-hd>
15. The *Universitat Autònoma de Barcelona* launched in the autumn of 2013 a *Máster en Humanidades Digitales* that is not available anymore.
16. As said earlier, in section 2.3, the vast majority of these data come from the *AEI*, the Spanish State Research Agency.
17. 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.
18. For consistency, we used the label "Ministry of Science" to group the various denomination given across the analysed period to the ministry in charge of the national research funding programme.
19. 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).
20. *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". *Artefactos*, v. 7, n. 1, pp. 13-28.
<https://doi.org/10.14201/art2018711328>
- Baraibar-Echeverría, Álvaro** (2014). *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". *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". *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; Sánchez-González, María (eds.). *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.
<https://doi.org/10.5749/minnesota/9780816677948.003.0010>
- Robinson-García, Nicolás; Amat, Carlos 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*, v. 41, n. 2, a201.
<https://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". *Caracteres*, 2, pp. 10-52.
- Romero-Frías, Esteban; Del-Barrio-García, Salvador** (2014). "Una visión de las humanidades digitales a través de sus centros". *Profesional de la información*, v. 23, n. 5, pp. 485-492.
<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-Contreras, 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.
<http://www.hsozkult.de/debate/id/diskussionen-2449>
- Svensson, Patrik** (2010). "The landscape of digital humanities". *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". *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>