Scientometrics (2017) 113:195–217
DOI 10.1007/s11192-017-2486-8

Co-word analysis and thematic landscapes in Spanish information science literature, 1985–2014

Carlos Olmeda-Gómez[1] · Maria-Antonia Ovalle-Perandones[1] · Antonio Perianes-Rodríguez[1]

Received: 20 February 2017 / Published online: 14 August 2017

Abstract This paper discusses the thematic backdrop for Spanish library and information science output. It draws from Web of Science records on papers authored by researchers at Spanish institutions and published under the category 'Information Science & Library Science' between 1985 and 2014. Two analytical techniques were used, one based on co-keyword and the other on document co-citation networks. Burst detection was applied to noun phrases and references of the intellectual base. Co-citation analysis identified nine research fronts: 'digital rights management', 'citation analysis', 'translation services', 'bibliometric analysis', 'co-authorship', 'electronic books', 'webometrics', 'information systems' and 'world wide web'. The most recent trends in the subject areas addressed in Spain were found to lie in metrics-related sub-specialities: the h-index, scientific collaboration, journal bibliometric indicators, rankings, universities and webometrics.

Keywords Co-word · Co-citation analysis · Citation networks · Information visualisation · Scientometrics · Research fronts

Introduction

Quantitative analysis of the bibliographic elements contained in formal scientific publications can be applied to determine the intellectual structure of research specialities (Small and Griffith 1974). This study explores the library and information science (LIS) output attributable to authors working at Spanish institutions from 1985 to 2014. The purpose is to define nationwide research subjects and thematic trends in this domain and identify any changes taking place across those years.

Library and information science was institutionalised in Spain in 1975 with the creation in the country's primary research body, the National Research Council, of three research

✉ Carlos Olmeda-Gómez
olmeda@bib.uc3m.es

[1] Department of Library & Information Science, Carlos III University, Madrid, Spain

institutes, one each for scientific information and documentation-related education, research and service provision (Plaza 2014). Library and information science was acknowledged as a university discipline in 1978. Education in the field began in 1981–1982 at 'university schools' (comparable to US junior colleges) at the Universities of Barcelona and Granada, with other universities following suit in the nineteen nineties. At this time, the discipline is taught in 17 public and private Spanish universities (Moneda-Corrochano 2016).

The importance attached to research on library and information science output in Spain has grown steadily from the outset. Spanish production from 1985 to 1994 was mapped in 1998 by Moya-Anegón et al. (1998), using co-citation techniques when the discipline was still in a stage of pre-development. The first 17 years of scientific production in Spain (1977–1994) were the object of bibliometric analysis by Cano (1999). Other quantitative studies (Jiménez-Contreras 2002) or papers on specific factors such as co-authorship in listed Spanish LIS research (Ardanuy 2012) have been published in the interim. The contribution and impact of Spanish institutions' output and their publication practices has been compared with those of other European Union countries (Olmeda-Gómez and Moya-Anegón 2016). A recent analysis of national production in a worldwide context found that authors affiliated with Spanish institutions ranked third in a total of 27 countries (Walters and Wilder 2016).

This article analyses the co-occurrence of entities such as keywords or jointly cited bibliographic references (i.e., co-citations) in scientific articles. The year 1985 was chosen as the start date for an earlier study (Arquero Aviles 2001) on author co-citation in Spanish library and information science output covered the period 1975–1984, although the literature analysed in that Ph.D. thesis was not indexed in the WoS.

The aim was to determine: (1) the core knowledge around which LIS research revolved in Spain in the years studied by creating co-keyword networks to ascertain frequency and proximity; (2) the temporal changes in the trends in research subjects; and (3) the structure of the intellectual base and research front identified on the grounds of co-citation networks.

The article is organised as follows. "Review of the literature" section describes the two most widely used approaches to identify knowledge structures in scientific fields. "Materials and methods"section explains how the data were gathered and the techniques and metrics used to analyse LIS articles attributed to Spanish institutions. The results of co-keyword and co-citation analysis are listed in "Results"section. The article concludes with a discussion of the results and pointing out some conclusions.

Review of the literature

To date, bibliometric studies have used primarily two methods to determine the subjects explored in a given research speciality: co-word analysis and document co-citation analysis (DCA). Both build on academic networks based on similarity and can be generated for different units of analysis (authors, journals, institutions or scientific fields).

In co-word analysis, content is explored through the co-occurrence of pairs of terms or lexemes (such as words or phrases) in a corpus of papers. Macro- or 'signal-words' (Rip and Courtial 1984) can be mathematically aggregated by building indexes that count the frequency of co-occurrence of terms in a given group of papers. The co-occurrence rates of the terms chosen are assumed to show the strength of the relationship between them. These pairs of terms may be grouped into clusters of the co-words that tend to be the most frequent in the subject area analysed and subsequently mapped. If the words appear together in many articles, the community of authors may in all likelihood be thought to see some manner of logical connection between them (Small 2003). The terms represented on these maps express and channel the conceits, interests and problems common to a given domain at a given time (hot topics), as well as areas of minority interest and the vocabulary associated with the subject (Callon et al. 1983, 1991).

Document co-citation analysis (Small 1973) is based on the citation in a given paper (X1) of two previously published 'co-cited' works (A and B). The strength of this co-citation bond is determined by the number of times that other citing articles (X2, X3, X4) jointly list the cited articles (A and B) in their references. The papers cited can be quantitatively pooled and represented on a graph, for instance, where they are shown as nodes and their inter-relationships as connecting arcs. Texts cited jointly are assumed to have some degree of thematic similarity or cognitive relationship (Garfield et al. 1978). Reference clusters with the highest co-citation values have been identified as the intellectual base for a given speciality (Persson 1994) and to represent and symbolise the conceptual structure of the corpus (Small 1980). The variety and validity of such methods have been examined (Gmür 2003), along with their most prominent variations, author (Rousseau and Zuccala 2004) and journal (Ding et al. 2000) co-citation analysis.

Many authors have analysed specialities using such traditional techniques either exclusively or in combination (Chang et al. 2015; Liu and Mei 2016) or by comparing them (Yan and Ding 2012) with others. Recent examples of co-word analysis can be found in research on nutrition (Blázquez-Ruiz et al. 2016), information metrics (Ravikumar et al. 2015; Sedighi 2016), the internet of things (Yan et al. 2015), recommendation systems (Hu and Zhang 2015), marketing (Wang et al. 2015), institutional repositories (Cho 2014), fuzzy set theory (López-Herrera et al. 2009), consumer behaviour (Muñoz-Leiva et al. 2012) and the economic transition in eastern and central European countries (Topalli and Ivanaj 2016), to mention but a few of the high diversity of subjects addressed.

Other recent studies have used document co-citation analysis to identify the knowledge base for specialities including transport geography (Liu and Gui 2016), organisational ambidexterity (García-Lillo et al. 2016), intercultural relations (Chi and Young 2013), innovation (Shafique 2013), substance abuse (González-Alcaide et al. 2016), corporate information systems (Shiau 2016), magnetic nanoparticles (Liu et al. 2016) and business and intra-company trust (Yang 2016), among others.

Both co-word and document co-citation analysis are characteristic of "knowledge domain analysis", a library and information science topic of research. "Knowledge domain analysis" is defined as an interdisciplinary area of research geared to mapping, content mining, classifying, analysing knowledge and enabling its presentation and navigation (Shiffrin and Börner 2004).

Materials and methods

Data

The search conducted in the Web of Science core collection yielded 2247 papers of all types published in journals classified under the WoS heading 'Information Science and Library Science' in which at least one of the authors was affiliated with a Spanish research institution, university or body. Of those, only the ones classified as articles (2209) were selected i.e., conference communications, reviews, notes and others were excluded. The citations analysed were the ones contained in the Web of Science's core collection when the papers were downloaded (2 September 2016).

The co-word and co-citation analyses of Spanish output were conducted with Java Citespace software (Chen 2004), which supports the construction and visualisation of bibliographic record networks (Chen 2006) and is suitable for visualising relative small ones. Each network, built for the number of years defined as the study period, was composed by the application to analyse the knowledge domain. Inasmuch as the fundamentals and details of these techniques are well known and fully explained in Chen et al. (2010), only the specific procedures deployed in this study are described below.

Co-keyword analysis

In this first procedure, the keywords listed in the set of 2209 bibliographic records were retrieved. No pre-processing of the keywords included in the fields DE e ID have been done. CiteSpace does not mix the terms in the keywords fields. We believe nuances might be lost if they were merged. Merging synonyms is potentially more computationally consuming than what you aim to achieve with your analysis in the first place. It is not cost-effective, you won't gain any more insight by merging them. In this particular study, analyzing the intellectual landscape of a specific scientific field, we decided that extra pre-processing was not worthwhile. The same approach is followed in similar works (Chen et al. 2014; Kim and Chen 2015).

The duration defined for all the networks, 1985–2014, was divided into 1 year intervals ("time slices" in Citespace jargon). Each network was formed from the 50 articles most frequently cited in each interval. The co-word frequencies in the matrix were converted using the value of the cosine as the similarity measure (Leydesdorff 2008). The networks were subsequently merged into a single entity in which the links were pruned where necessary using the Pathfinder network technique. Pathfinder network scaling simplifies the co-word matrix by retaining the strongest co-word links with reference to the triangle inequality condition (Schvaneveldt 1990). The Kamada and Kawai (1989) algorithm was applied to visualise the spatial organisation of the co-keyword node graph.

Burst

A second procedure to determine thematic trends consisted in identifying the noun phrases contained in the titles and abstracts of the articles in the record base created. The terms appeared at a given time in the period analysed in the wake of their intense use by the research community. These 'bursts of activity' were detected with Kleinberg's (2002) algorithm. Instead of simple frequencies, this algorithm deploys a probabilistic "automaton" whose states are conditioned by noun phrase frequencies. Inter-state transitions are the time points around which term frequencies change significantly. The algorithm generates a list of the strongest noun phrase bursts found in the set of records, as well as the years when they appeared.

Document co-citation analysis

Spanish LIS document co-citation analysis was conducted by building co-citation networks for the references contained in 2202 articles (the software excluded seven articles for unknown technical reasons). It detected the structural subjects informing the speciality. Co-citation counts were used to calculate the similarity between pairs of references. The interval, selection thresholds and link normalisation criterion were the same as used to build the co-keyword network. The largest connected or main component on the network built was selected for subsequent analysis.

Spectral clustering (Luxburg 2007) was applied to that main component to subdivide the network into clusters or modules, subsequently subjected to dual analysis. The references cited in each cluster with the highest co-citation values were identified as the intellectual base for the speciality and the subjects of the citing articles as the research fronts.

Two procedures were used to summarise research front subjects on the grounds of the descriptors for the keywords of the citing articles. The first consisted in listing subjects based on the keywords with the highest term frequency-inverse document frequency (tf*idf) value (Salton and McGill 1986). In the second, keywords were retrieved using the log likelihood ratio, comparing the observed to the expected keyword frequencies based on contingency tables generated from the corpus (Dunning 1993). The clusters on the network graphic created as a result of these operations were labelled with the keyword having the highest inverse document frequency.

Metrics

The analytical and representational procedures used to create networks included the calculations supported by the software to generate a series of metrics. (1) A betweenness centrality value, which is the proportion of all the geodesic distances between network pairs including the keyword at issue, was calculated for each network node (Freeman 1977). A keyword with a high value meant that it occupied a significant position as a network intermediary and was associated with a large number of other keywords. (2) Q modularity is a quality measure that expresses the goodness of the overall network breakdown into clusters (Newman 2006). A high value for this parameter means that the initial network was divided into clusters in which the constituent nodes were closely interconnected and any weak inter-cluster links were deleted. A low value means that a given network division failed to locate more node clusters than would have appeared randomly (Shibata et al. 2008). (3) The silhouette value indicates which terms fit well into the clusters and which are located between several, without assignment to any. Silhouette values denote the relative quality of the breakdown into clusters. The mean silhouette value assesses clustering validity, in which perfect network subdivision is 1 and the worst possible -1 (Rousseeuw 1987). (4) Sigma is a metric that combines the value of a structural measure, such as betweenness, with a time-related node property, burst strength. When used in co-keyword network analysis to identify the most novel co-keywords, it is defined as: $(\textit{betweenness centrality} + \text{sim1})^{\text{burstiness}}$ (Chen et al. 2009).

Results

The yearly output of authors affiliated with Spanish institutions publishing in journals listed under *Journal Citation Reports'* (JCR) category 'Information Science & Library Science' is shown in Fig. 1. Annual output never exceeded 35 papers in 1985–2004. That number was surpassed for the first time in 2005 and from 2007 onward, production never dipped below 130. This abrupt rise was largely due to the provisional inclusion in JCR of

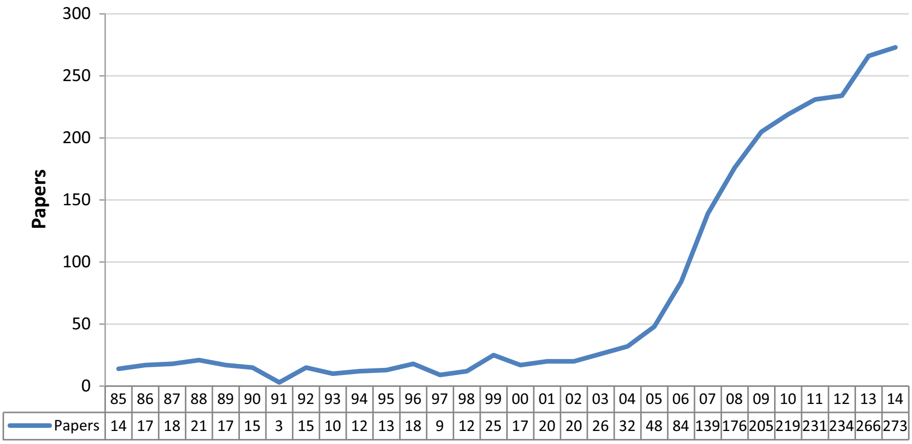

Fig. 1 Number of articles with Spanish affiliations published yearly under the WoS category ‘Information Science & Library Science’

two Spanish professional journals, *El Profesional de la Información* and *Revista Española de Documentación Científica*, prior to their permanent inclusion in 2008 and 2010. These publications are major vehicles for Spanish LIS research (Olmeda-Gómez and Moya-Anegón 2016).

Table 1 lists the universities and institutions publishing the largest number of papers in the journals analysed in this study. With the sole exception of Catalonia’s Universitat Oberta, a private institution, only public universities are listed. The universities in the highest positions (1–13) all had departments or faculties delivering library and information science courses. The top spot was held by the University of Granada, the first Spanish university to offer a degree in this area, beginning in the nineteen eighties. The only non-university on the list was the Spanish National Research Council (CSIC). Specific centres

Table 1 Number of Wos category ‘Information Science & Library Science’ papers published by Spanish institutions, 1985–2014

Rank	# Papers	Institution	Rank	# Papers	Institution
1	352	Univ Granada	12	63	Univ Alcalá Henares
2	246	Spanish National Research Council. CSIC	13a	55	Univ Salamanca
3	184	Univ Carlos III Madrid	13b	55	Univ Politécnica Madrid
4	133	Univ Complutense Madrid	14	53	Univ Navarra
5	127	Univ Barcelona	15a	37	Univ Sevilla
6	117	Univ Politécnica de Valencia	15b	37	Univ Autónoma Barcelona
7	104	Univ Extremadura	16	36	Univ Autónoma Madrid
8	86	Univ Valencia	17	34	Univ Oberta Catalunya
9	82	Univ Murcia	18	33	Univ Alicante
10	79	Univ Zaragoza			
11	65	Univ Pompeu Fabra			

and institutions under the aegis of the CSIC are responsible for gathering and managing information on science, technology and the humanities and employ staff whose research is geared to studies on science and technology.

The data on the two networks built are given in Table 2. For the co-keywords, the individual networks were merged into one with 407 co-keywords. The initial network had 1566 links. The proportion of actual to possible links was a very small 1.9%. The largest connected sub-graph grouped 356 co-keywords or 87% of the total identified. The silhouette values suggested good clustering and the modularity value denoted good but not exceptional interconnections among them.

The co-citation network was built from an analysis of 2202 records, which yielded 51,674 valid (97.5% of the total) and 1297 (2.4%) invalid (incomplete and therefore excluded) references. This single network had 4746 nodes and 10,506 links, for a network density of 0.0009. The main component had 931 nodes, or 19% of the total.

Thematic trends. Co-word analysis and burst detection (1985–2014)

The thematic trends informing the domain were defined from the frequencies of the co-keywords most used by researchers. As the data in Table 3 show, the main poles of attraction for researchers working in the domain in the period analysed revolved around the terms 'science', 'Spain', 'internet', 'citation analysis', 'impact', 'indicators' and 'university'. With such abstract terms, research subjects can only be identified on a macro-scale. The betweenness values for each co-term might reflect the power of the individual term to match subordinate subjects. 'Information' (0.44), 'internet' (0.36) and 'citation analysis' (0.29), for instance, exhibited the highest values.

An additional way to illustrate the subjects preferred in the Spanish domain is by representing the keywords as a network. Figure 2 shows part of the largest component in the Pathfinder-pruned co-keyword network, with the most frequent terms highlighted. Adjacent terms were the ones with a high joint assignment rate. Near the top of the graph, for instance, the terms 'absorptive capacity', 'information technology', 'firm' and 'knowledge management' are positioned close to one another. The terms most frequently used are circled, with purple outer circles identifying terms with structural betweenness value. The thicker the outer ring, the higher the value.

The presence of concurrent terms on the network denotes relative consensus about the macro-scale research issues addressed. Network density values are an indirect measure of the degree of diversity and cohesion. The low values found here were an indication that consensus in the use of common elements related to the subjects contained in the articles was also low. This is an initial sign that scientific discourse in this domain is only weakly integrated.

Table 2 Co-keyword and document co-citation networks (both based on 1 year intervals)

	# Nodes	# Links	Density	Largest component	Modularity	Mean silhouette	# Records
1985–2014 co-keyword network	407	1566	0.019	356 (87%)	0.585	0.587	2209
1985–2014 document co-citation network	4746	10,506	0.0009	931 (19%)	0.971	0.191	2202

Table 3 Most frequent keywords (frequency over 15), 1985–2014, in descending order

Frequency	Betweenness centrality	Keyword	Frequency	Betweenness centrality	Keyword
193	0.05	Science	39	0	Ranking
146	0.04	Spain	37	0.02	Bibliometric indicator
129	0.36	Internet	35	0.1	Database
126	0.3	Impact	34	0.03	Scopus
100	0.44	Information	33	0	Social network
99	0.2	Indicator	33	0.01	Research performance
97	0.07	University	31	0.12	Information technology
74	0.18	System	29	0.02	Evaluation
73	0	Web	28	0.03	*h* index
69	0.02	Innovation	27	0.01	Index
67	0.29	Citation analysis	26	0.12	Firm
59	0.2	Citation	25	0.01	Academic library
57	0.06	Bibliometrics	22	0.09	Behaviour
55	0.18	Journal	22	0	Social science
50	0.14	Collaboration	21	0.08	Productivity
48	0.15	Information retrieval	20	0.09	Webometrics
47	0.13	Library	20	0.04	Organization
47	0.04	Web 20	19	0	Scientific journal
45	0.05	Technology	18	0.08	Bibliometric analysis
45	0.08	Network	18	0.09	Knowledge management
44	0.03	Management	17	0.03	Trust
43	0.14	Communication	17	0	Web site
40	0.1	Knowledge	17	0.05	Absorptive capacity
39	0.01	Publication	16	0.03	Semantic web
39	0.01	Impact factor	16	0.01	E government

A more detailed breakdown of the noun-phrases or terms retrieved by CiteSpace from the title and abstract fields in 2202 records and derived from the burst analyses described above is given in Table 4. Citespace generates networks of co-occurring terms or noun phrases. This is undertaken using a Part-of-speech (POS) tagging algorithm. By means of this procedure, the informative content of the concepts is represented. Studies show that such n-grams preserve far more semantic content than individual term extraction (Ding and Chen 2014). By furnishing more specific information on single words, these findings facilitate interpretation, for such phrases can be more readily linked to conceits (Schneider 2006).

The period studied, 1985–2014, was divided into three sub-periods to determine the thematic trends in each, including the most recent years based on the publication trends observed in the data analysed (Fig. 1). The first period, 1985–1993, can be regarded as a preliminary stage, followed in 1994–2003 by development, in which papers began to be written in the university departments created in the early nineteen nineties. The last sub-period may be regarded as a growth or consolidation stage characterised by a sustained rise in output.

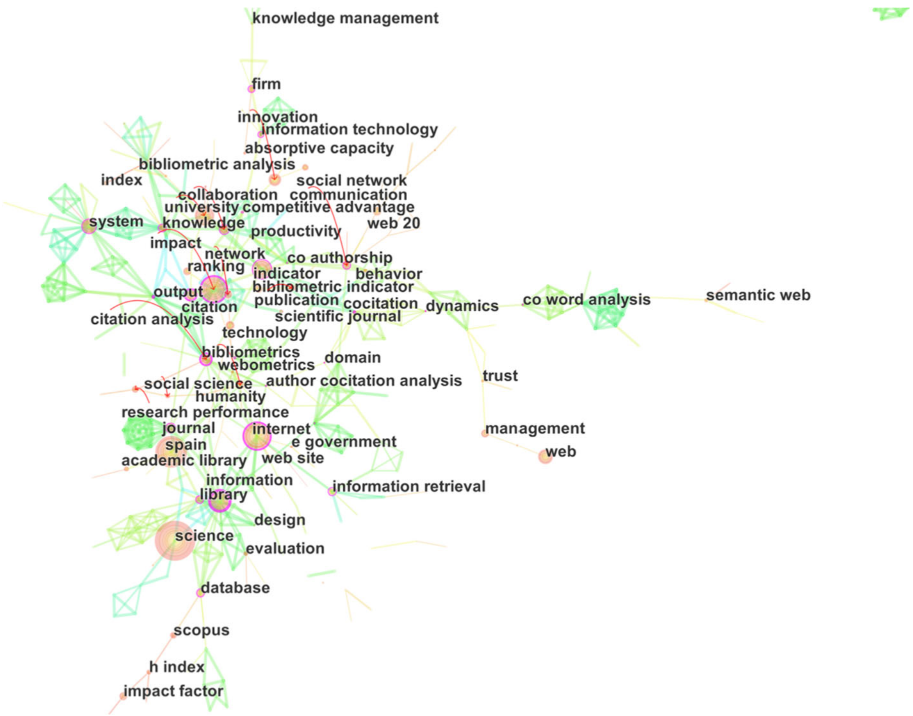

Fig. 2 Co-keyword network of Spanish library and information science papers, 1985–2014, containing 407 nodes and 1566 links in 2209 records. Pruning with Pathfinder. *Purple rings*: keywords with high betweenness. (Color figure online)

The earliest sub-period lists noun phrases found in a Spanish journal (*Revista de Informática y Automática*) classified under the subject category until 1989 when it ceased publication. The phrase 'information professionals' first appeared in 1988–1990. In the second (1994–2003), high burst strength values were found for 'electronic library' (1.87), 'scientific journal' (1.93) and 'information retrieval system' (1.74). The most recent sub-period confirmed consolidated subject trends, such as 'library science' (3.24) and others relating to science more generally: 'scientific research' (5.09), 'scientific production' (3.14), 'scientific output' (3.93), 'scientific collaboration' (2.88), 'scientific information' (4.8) and 'scientific field' (4.76). Note the high strength value for 'social media' (5.4).

Figure 3 shows the subjects with the highest structural values that also attracted researchers' attention for more time, according to their burst activity scores. They included 'information retrieval' (3.9), 'co-citation' (2.36), 'information technology' (1.73), 'database' (1.5), 'web' 2.0 (1.4), 'world wide web' (1.3), 'Scopus' (1.28), 'semantic web' (1.09), 'bibliometric indicator' (1.07), 'digital library' (1.12), 'research performance' (1.04), 'information literacy' (1.04), 'information science' (1.04), 'impact factor' (1.03) 'media' (1.03), 'web site' (1.03), 'map' (1.03), 'scientometrics' (1.01) and 'social network' (1.01).

Table 4 Noun phrases with the strongest bursts (in bold type, duration >3 years)

Noun phrase	Year	Strength	Begin*	End*
1985–1993				
Automatic speech recognition	1985	0.957	1987	1988
Isolated word recognition	1985	0.957	1987	1988
Path composition	1985	0.957	1987	1988
Information professionals	1985	0.834	1988	1990
1994–2003				
Electronic library	1994	1.875	1994	1995
Electronic book	1994	1.023	1994	1995
Electronic publishing	1994	1.023	1994	1995
Scientific journals	1994	1.935	1995	1996
Biomedical subfields	1994	0.977	1996	1997
Cardiovascular system	1994	0.770	1996	1997
Bibliometric indicators	1994	0.849	1996	1997
Collaboration pattern	1994	0.816	1997	1999
Bibliometric analysis	1994	0.802	1999	2000
Information retrieval system	1994	1.741	2001	2003
Cluster analysis	1994	0.827	2001	2003
2004–2014				
Library science	2004	3.243	2004	2009
Information retrieval system	2004	2.642	2005	2009
Web pages	2004	2.971	2005	2006
Science citation index	2004	3.334	2006	2009
Scientific research	2004	5.092	2006	2007
Information source	2004	3.457	2007	2009
Information literacy	2004	3.268	2007	2011
Semantic web	2004	3.190	2007	2008
Digital library	2004	3.448	2007	2009
Scientific production	2004	3.149	2008	2010
Information technology	2004	3.998	2008	2010
Public library	2004	3.513	2008	2009
Scientific output	2004	3.938	2009	2010
Scientific collaboration	2004	2.886	2009	2010
Search engine	2004	2.73	2009	2010
Information management	2004	2.574	2009	2011
Communication technology	2004	3.624	2009	2010
Scientific information	2004	4.825	2009	2010
Digital media	2004	2.525	2010	2012
Scientific community	2004	3.374	2011	2012
Spanish university system	2004	2.711	2011	2014
Scientific field	2004	4.766	2011	2014
European Union	2004	3.320	2011	2014
Knowledge organization	2004	3.166	2012	2014
Research output	2004	3.166	2012	2014

Table 4 continued

Noun phrase	Year	Strength	Begin*	End*
Scopus database	2004	4.991	2012	2014
Business model	2004	2.912	2012	2014
Social media	2004	5.449	2012	2014
Content analysis	2004	3.891	2012	2014

Terms excluded: electronic books; dynamic system; different periods; different field; positive effect; education area; main objective; empirical evidence (* beginning and end of period when activity was most intense)

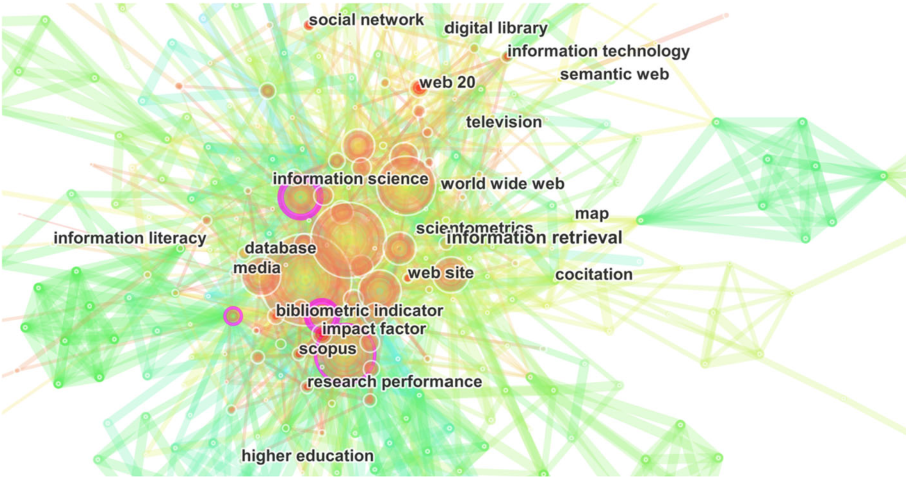

Fig. 3 Co-keyword network of Spanish library and information science papers with the highest sigma values. 1985–2014. *Purple rings*: keywords with high betweenness. (Color figure online)

Document co-citation analysis: research fronts and intellectual base (1985–2014)

The research fronts and intellectual base for the speciality were identified using document co-citation analysis for objective classification and independent and automatic generation. The boundaries were set on the grounds of the breadth of the period studied, the size of the clusters resulting from the breakdown, the validity of that process as measured by the silhouette values for each cluster and the consistency of the subject contents identified and the research front labels.

The breakdown of the co-citation network into nine sub-specialities is given in Table 5. The 771 references involved accounted for 82.8% of the total contained in the largest connected network. The silhouette values indicative of the relative validity of the clusters were very high, denoting a clearly separated structure. Values of 1 mean that the respective cluster was perfectly separated and uniform. The silhouette values for each cluster varied depending on the size of the sets of references from which it was derived. Somewhat lower values were found for clusters with a larger number of references, such as clusters #0 and #1, for their internal structure and contents were more diverse. The most recent clusters were identified on the grounds of the average year of publication.

Table 5 Number and size of co-cited reference clusters (minimum ≥59)

Cluster ID	Size	Distribution (% of 931)	Average year	Silhouette
0	112	12.03	1999	0.991
1	108	11.6	2007	0.915
2	93	9.99	1998	1
3	89	9.56	1996	0.99
4	88	9.45	1996	0.98
5	84	9.02	1991	1
6	69	7.41	2006	0.93
7	69	7.41	1998	0.99
8	59	6.34	1998	0.99
Total	771	82.81		

The clusters were described on the basis of the keywords in the articles citing the references classified (Table 6), in turn used to label the research fronts active during the period. Assuming the same *p* value, the identifiers with the highest log likelihood ratios (LLR) were the most unique, i.e., the most suitable labels.

The four sub-specialities labelled generically (#1, #3, #4, #6) (Figs. 5 and 6) exhibited research metrics that differed fairly slightly. Five other sub-specialities were also identified and labelled using the term frequency-inverse document frequency formula, as follows: #5, ‘electronic books’; #7, ‘information systems’, ‘structured document retrieval’, ‘information retrieval’, ‘indexing services’ and ‘hypertexts’; #2, ‘translation and abstracting services’;

Table 6 Major co-cited reference clusters (1985–2014)

# Cluster	Mean year	Tf-idf label[a]	LLR label ($p = 0.0001$)
0	1999	Digital rights management | copyright law	Intellectual property (859.98); copyright (627.32); trusted systems (442.1)
1	2007	Citation analysis | similarity measure	H index (340.55); hirsch index (208.5); ranking (189.65)
2	1998	Translation services | data representation	Structured abstracts (412.55); abstracts (307.33); abstracting (307.33)
3	1996	Bibliometric analysis |science citation index	International scientific cooperation (180.849)
4	1996	Co-authorship | international collaboration	Mathematics (246.83); world science (246.83); developing country (229.91)
5	1991	Electronic books | information retrieval	Electronic books (156.88)
6	2006	Webometrics | bibliometrics	Link analysis (295.68); site interlinking (287.7); webometrics (281.11)
7	1998	Information system | structured document retrieval	Information retrieval (532.09); indexing services (245.5); hypertext (187.51)
8	1998	World wide web	Maps (361.45); information retrieval (159.93); author co-citation analysis (139.9)

Stopwords size, single, hard and online excluded

LLR log likelihood ratio

[a] Term frequency—inverse document frequency

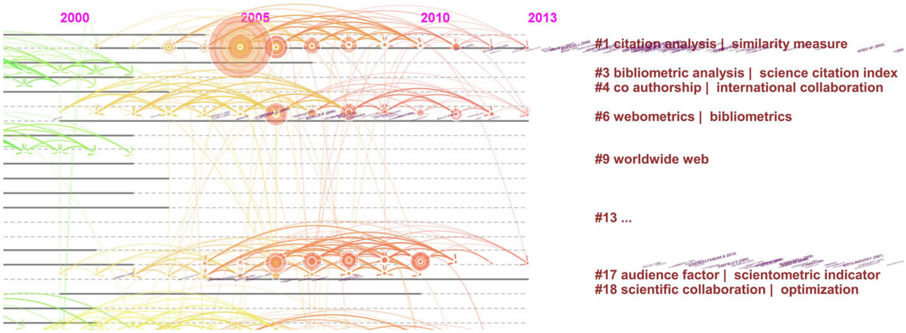

Fig. 4 Document co-citation cluster timeline

#8, ‘world wide web’, ‘maps’, ‘information retrieval’ and ‘author co-citation analysis’; and #0, ‘digital rights management’, ‘copyright law’ and ‘intellectual property’.

The four clusters with references related to scientific information metrics have prevailed to date. They also contain the most recent references and may have developed further after the end date for this study. The aforesaid references constitute the primary intellectual base for Spanish LIS. Of these four clusters, described in greater detail below, only #1 (citation analysis | similarity measure) and #6 (webometrics | bibliometrics) contained references with citation bursts (the nodes with red concentric rings in Fig. 6).

Cluster #1 was the second largest, with 108 references and mean year 2007. The five most cited papers are listed in Table 7. J. Hirsch’s well-known study on the *h*-index stood out above all the others, which dealt with bibliometric rankings, the *g*-index, assessment bibliometrics and the application of citation analysis to social science and humanities output. This was the cluster with the largest number of citation burst papers.

Cluster #3 (Table 8), with a mean year of 1996, contained 89 papers, including some by Spanish authors. The following terms appeared in the titles: ‘national journals’, ‘international collaboration’, ‘multi-disciplinary research’ and ‘journal internationalisation’.

Table 7 References with the highest citation frequency in cluster #1

Cluster	Citations	Author	Year	Source	Title
1	52	Hirsch JE	2005	*P Natl Acad Sci USA*	‘An index to quantify an individual’s scientific research output’. *Proceedings of the National Academy of Sciences of the United States of America* Volume 102, Issue 46, 15 November 2005, pp. 16,569–16,572
1	23	Van Raan AFJ	2005	*Scientometrics*	‘Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods’. *Scientometrics*. 62 (1), pp. 133–143
1	21	Moed HF	2005	*Citation Anal Res Ev*	*Citation Analysis in Research Evaluation*. Springer, 2005
1	20	Egghe L	2006	*Scientometrics*	Theory and practice of the g-index. *Scientometrics*. 69 (1), pp. 131–152
1	20	Nederhof AJ	2005	*Scientometrics*	Bibliometric monitoring of research performance in the social sciences and the humanities: A review. *Scientometrics* 66 (1), pp. 81–100

Table 8 References with the highest citation frequency in cluster: (a) #3, (b) #4

Cluster	Citations	Author	Year	Source	Title
3	3	Rey-Rocha J	1999	*Scientometrics*	'The role of domestic journals in geographically-oriented disciplines: The case of Spanish journals on Earth Sciences'. *Scientometrics* 45 (2), pp. 203–216
3	3	Luukkonen T	1992	*Sci Technol Hum Val*	'Understanding Patterns of International Scientific Collaboration'. *Science, Technology & Human Values* 17 (1), pp. 101–126
3	3	Mendez A	1993	*Dinamica Investigación*	*Dinámica de la investigación multidisciplinar sobre nuevos materiales en España. Un análisis bibliométrico*. Cindoc, 1993
3	2	Bordons M	1993	*Res Evaluat*	'Is collaboration improving research visibility? Spanish scientific output in pharmacology and pharmacy'. *Research Evaluation* 3 (1), pp. 19–24
3	2	Zitt M	1998	*Scientometrics*	'Internationalization of scientific journals: A measurement based on publication and citation scope'. *Scientometrics* 41 (1–2), pp. 255–271
4	3	Moya-Anegon F	1998	*Scientometrics*	'Research fronts in library and information science in Spain (1985–1994)'. *Scientometrics* 42 (2), pp. 229–246
4	3	MacRoberts MH	1996	*Scientometrics*	'Problems of citation analysis'. *Scientometrics*. 36 (3), pp. 435–444
4	2	Fernandez A	1998	*Rev Espanola Documen*	'Síntesis de estudios bibliométricos españoles en educación. Una dimensión evaluativa'. *Revista Española de Documentación Científica*. 21 (3), 269–285
4	2	Hjorland B	1995	*J Am Soc Inform Sci*	'Toward a new horizon in information science: Domain-analysis'. *Journal of the American Society for Information Science* 46 (6), pp. 400–425
4	2	Rennie D	1997	*Jama-J Am Med Assoc*	'When authorship fails: A proposal to make contributors accountable'. *Journal of the American Medical Association* 278 (7), pp. 579–585

The mean year for the 88 references in cluster #4 (Table 8) was 1996. They addressed subjects similar to the ones found for cluster #3: 'citation analysis problems', 'research fronts', 'education study bibliometrics', 'domain analysis' and 'authorship delimitation' (Figs. 5, 6).

Cluster #6, identified by the term 'webometrics', focused on quantitative studies of websites and traffic. The most frequently cited references, listed in Table 9, addressed 'link analysis', 'website interlinking', 'academic and university web analysis' and the 'Google Books impact measure'.

Table 10 shows the seven references with the strongest citation bursts listed in the 50 articles with the strongest bursts in the period studied (see Table 12). The Hirsch (2005) paper

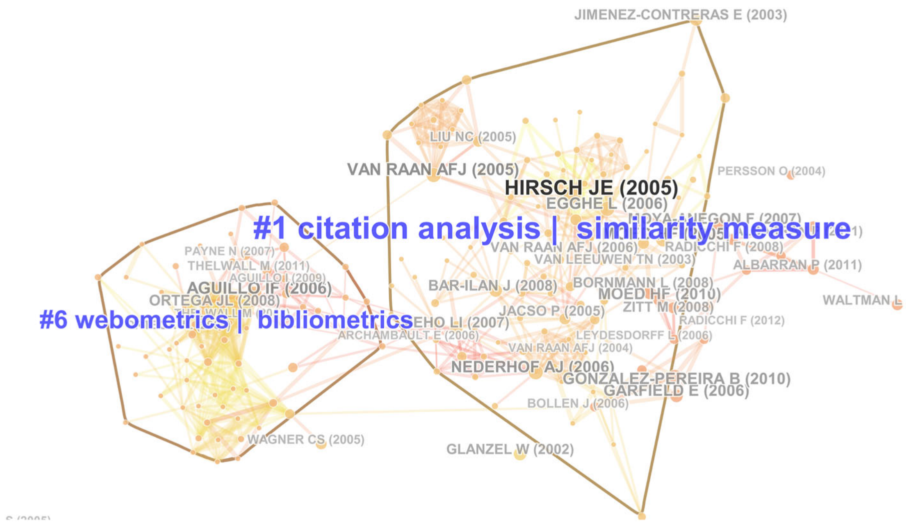

Fig. 5 Networks of co-cited references, with clusters #1 and #6 as the intellectual base

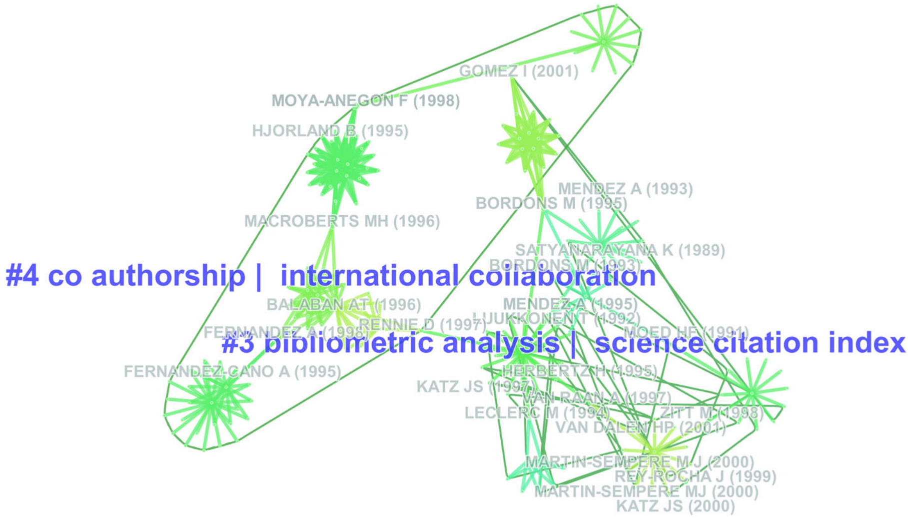

Fig. 6 Networks of co-cited references, with clusters #3 and #4 as the intellectual base

proposing the *h*-index as an indicator for individual researchers' scientific output, which drew much international attention, had the strongest citation burst, with a score of 18.3.

Table 11 shows the twelve references with the most recent citation bursts, from 2012 onward. In some cases, citation bursts began in 2012 for papers published earlier. The highest burst strength was recorded for a paper by Moed (2010), who proposed a new indicator denominated the source-normalised impact per paper (SNIP) to measure journal impact based on citations. Zitt and Small (2008), with the second highest score, introduced a new normalisation strategy known as the audience factor and Albarrán et al. (2011), with the third, published a study on the distribution of citations and references by categories and

Table 9 References with the highest citation frequency in cluster #6

Cluster	Citations	Author	Year	Source	Title
6	15	Ortega JL	2008	*Scientometrics*	'Maps of the academic web in the European Higher Education Area—An exploration of visual web indicators'. *Scientometrics* 74 (2), pp. 295–308
6	9	Thelwall M	2011	*J Am Soc Inf Sci Tec*	'Assessing the citation impact of books: The role of Google Books, Google Scholar, and Scopus'. *Journal of the American Society for Information Science and Technology* 62 (11), pp. 2147–2164
6	8	Thelwall M	2008	*Scientometrics*	'A university-centred European Union link analysis'. *Scientometrics* 75 (3), pp. 407–420
6	8	Payne N	2007	*Scientometrics*	'A longitudinal study of academic webs: Growth and stabilisation'. *Scientometrics* 71 (3), pp. 523–539
6	8	Aguillo I	2009	*Libr Hi Tech*	'Measuring the institution's footprint in the web'. *Library Hi Tech* 27 (4), pp. 540–556

Table 10 Seven references (only first authors' names shown) with the strongest citation bursts, 1985–2014 (bursts detected from an analysis of all the citations in the 50 articles most cited in each year in the period) (* beginning and end of period when activity was most intense)

References	Year	Strength	Begin*	End*
Hirsch JE, 2005, *P Natl Acad Sci USA*, V102, P16569	2005	18.35	2007	2014
Moed HF, 2010, *J Informetr*, V4, P265	2010	8.95	2012	2014
Moed HF, 2005, *Citation Anal Res Ev*, V, P	2005	8.43	2007	2012
Gonzalez-Pereira B et al., 2010, J *Informetr*, V4, P379	2010	8.32	2011	2014
Garfield E, 2006, *Jama-J Am Med Assoc*, V295, P90	2006	8.32	2011	2014
Aguillo IF et al., 2006, *J Am Soc Inf Sci Tec*, V57, P1296	2006	8.24	2009	2014
Van Raan AFJ, 2005, *Scientometrics*, V62, P133	2005	8.20	2009	2014

subject sub-fields. In the fourth ranked paper, Radicchi et al. (2008) proposed a method to compare the impacts of articles in different disciplines. Thelwall and Sud's (2011) article, with the fifth strongest burst, compared web link counts delivered by two alternative methods to measure organizations' online impact. The sixth, by Aguillo (2010), introduced website assessment metrics, and the seventh by Bordons et al. (2010) analysed the relationship between research findings based on Spanish universities' publications and structural factors and socio-economic data.

Discussion and conclusions

This paper identifies the subjects addressed in information science and library science literature authored by researchers working out of Spanish institutions by building co-keyword and document co-citation networks. The main subjects are represented graphically by discipline in a national context. The burst activity algorithm is deployed to detect the noun phrases attracting the greatest attention and ascertain the evolution over time of

Table 11 Twelve references (only first authors' names shown) with the most recent citation bursts, 1985–2014 (bursts detected from an analysis of all the citations in the 50 articles most cited in each year in the period) (* beginning and end of period when activity was most intense)

References	Year	Strength	Begin*	End*
Moed HF, 2010, J *Informetr*, V4, P265	2010	8.952	2012	2014
Zitt M, & Small, H. 2008, *J Am Soc Inf Sci Tec*, V59, P1856	2008	6.840	2012	2014
Albarrán P, et al. 2011, *Scientometrics*, V88, P385	2011	6.313	2012	2014
Radicchi F, et al. 2008, *P Natl Acad Sci USA*, V105, P17268	2008	5.785	2012	2014
Thelwall M, & Sud, P. 2011, *J Am Soc Inf Sci Tec*, V62, P1488	2011	4.732	2012	2014
Aguillo I, 2009, *Libr Hi Tech*, V27, P540	2009	4.205	2012	2014
Bordons M, et al. 2010, *Rev Esp Doc Cient*, V33, P9,	2010	3.974	2012	2014
Leydesdorff L, & Opthof, T. 2010, *J Informetr*, V4, P644,	2010	3.974	2012	2014
Torres-Salinas D, et al. 2011, *Prof Inform*, V20, P111	2011	3.97	2012	2014
Leydesdorff L, 2010, et al. *J Am Soc Inf Sci Tec*, V61, P352	2010	3.97	2012	2014
Torres-Salinas D, et al. 2011, *Prof Inform*, V20, P701	2011	3.974	2012	2014
Docampo D, 2011, *Scientometrics*, V86, P77	2011	3.974	2012	2014

the subjects of greatest interest. The articles with the largest and most recent citation bursts included in the intellectual base for the Spanish domain are also identified.

Globally speaking, library and information science rests on a combination of essentially three major subject categories: library science, information retrieval and bibliometrics, although the first is vanishing (Lariviere et al. 2012). The present analysis reveals that in Spain, all the sub-specialities and hot topics fall under one of those three categories, with no deviations relative to the worldwide domain.

The document co-citation analysis and subsequent characterisation conducted here detect nine sub-specialities. These fronts represent the degree of consensus around a given combination of conceits found in the corpus analysed. Four of these are inter-related and readily characterised. They concur with the research fronts detected in global analyses under the heading 'analysis of scientific literature' (Zhao and Strotmann 2008, 2014).

The references cited in the 'scientific literature analysis'-related fronts and the ones with the strongest bursts (Table 12) are the result of convergent interests among researchers working out of Spain's largest public scientific body, the National Research Council and, once the discipline was institutionalised, out of the country's universities. They also include papers by outstanding non-Spanish bibliometricians.

The intellectual content of the oldest 'scientific literature' research fronts (#3 and #4) has been updated and diversified. This is attested to by the intellectual base and the subjects of interest. The use of references is more intense and uniform. The fact that nearly all the references with the strongest burst factors in the Spanish domain fall under the two most recent fronts (#1 'citation analysis' and #6 'webometrics') denotes the existence of shared cognitive objects with their respective discourses and symbolisms. Such circumstances make it easier to assess the nature and significance of the findings (Whitley 2000), for they reveal a certain intellectual identity.

Based on the mean year of the references in the intellectual base, the remaining research fronts (Table 6, clusters #0, #2, #5, #7 and #8) are of an intermediate age. Here the interpretation of keyword retrieval and subsequent labelling is less clear-cut. 'Information retrieval', for instance, appears in three clusters, in the company of different mixes of other

Table 12 Top 50 references with the strongest citation bursts, 1985–2014

References	Year	Strength	Begin*	End*
Almind TC, 1997, *J Doc*, V53, P404	1997	3.705	2003	2005
Jimenez-Contreras E, 2003, *Res Policy*, V32, P123	2003	6.160	2005	2011
Bordons M, 2003, *Scientometrics*, V57, P159	2003	3.716	2005	2009
Glanzel W, 2002, *Scientometrics*, V53, P171	2002	7.259	2006	2010
Moya-Anegon F, 2004, *Scientometrics*, V61, P129	2004	6.389	2006	2010
Hirsch JE, 2005, *P Natl Acad Sci USA*, V102, P16569	2005	18.355	2007	2014
Moed HF, 2005, *Citation Anal Res Ev*, V, P	2005	8.435	2007	2012
Egghe L, 2006, *Scientometrics*, V69, P131	2006	7.017	2007	2014
Van Leeuwen TN, 2003, *Scientometrics*, V57, P257	2003	5.243	2007	2011
Van Raan AFJ, 2006, *Scientometrics*, V67, P491	2006	4.555	2007	2014
Maybee C, 2006, *J Acad Libr*, V32, P79	2006	3.767	2007	2008
Ruiz-Perez R, 2002, *J Med Libr Assoc*, V90, P411	2002	3.767	2007	2008
Wagner CS, 2005, *Res Policy*, V34, P1608	2005	4.274	2008	2012
Cothey V, 2005, P 10 *Int C Int Soc S*, V, P	2005	3.565	2008	2009
Aguillo IF, 2006, *J Am Soc Inf Sci Tec*, V57, P1296	2006	8.247	2009	2014
Van Raan AFJ, 2005, *Scientometrics*, V62, P133	2005	8.205	2009	2014
Nederhof AJ, 2006, *Scientometrics*, V66, P81	2006	7.494	2009	2014
Ortega JL, 2008, *Scientometrics*, V74, P295	2008	5.615	2009	2014
Jacso P, 2005, *Curr Sci India*, V89, P1537	2005	5.414	2009	2014
Bollen J, 2006, *Scientometrics*, V69, P669	2006	4.201	2009	2012
Ranganathan C, & Ganapathy, S. 2002, *Inform Manage-Amster*, V39, P457	2002	3.636	2009	2010
Bar-Ilan J, 2008, *Scientometrics*, V74, P257	2008	7.155	2010	2014
Moya-Anegon F, 2007, *Scientometrics*, V73, P53	2007	6.957	2010	2014
Meho LI, 2007, *J Am Soc Inf Sci Tec*, V58, P2105	2007	6.521	2010	2014
Guallar J, & Abadal, E. 2009, *Prof Inform*, V18, P255	2009	5.214	2010	2014
Alonso S, 2009, *J Informetr*, V3, P273	2009	4.218	2010	2011
Abadal E, & Guallar, J.2010, *Prensa Digital Bibli*, V, P	2010	3.688	2010	2012
Payne N, 2007, *Scientometrics*, V71, P523	2007	3.473	2010	2014
Gonzalez-Pereira B, 2010, *J Informetr*, V4, P379	2010	8.329	2011	2014
Garfield E, 2006, *Jama-J Am Med Assoc*, V295, P90	2006	8.329	2011	2014
Liu NC, 2005, *Higher Ed Europe*, V30, P127	2005	6.053	2011	2014
Albarran P, 2011, *J Am Soc Inf Sci Tec*, V62, P40	2011	5.546	2011	2014
Leydesdorff L, 2009, *J Am Soc Inf Sci Tec*, V60, P348	2009	4.157	2011	2014
Subramaniam M, Youndt, M. A. 2005, *Acad Manage J*, V48, P450	2005	3.882	2011	2012
Glanzel W, 2007, *J Informetr*, V1, P92	2007	3.882	2011	2012
Archambault E, 2006, *Scientometrics*, V68, P329	2006	3.694	2011	2014
Bergstrom C, 2007, *Coll Res Lib News*, V68, P314	2007	3.527	2011	2014
Opthof T, 2010, *J Informetr*, V4, P423	2010	3.527	2011	2014
Moed HF, 2010, *J Informetr*, V4, P265	2010	8.952	2012	2014
Zitt M, & Small, H. 2008, *J Am Soc Inf Sci Tec*, V59, P1856	2008	6.840	2012	2014
Albarran P, 2011, *Scientometrics*, V88, P385	2011	6.313	2012	2014
Radicchi F, et al. 2008, *P Natl Acad Sci USA*, V105, P17268	2008	5.785	2012	2014
Thelwall M, & Sud, P. 2011, *J Am Soc Inf Sci Tec*, V62, P1488	2011	4.732	2012	2014

Table 12 continued

References	Year	Strength	Begin*	End*
Aguillo I, 2009, *Libr Hi Tech*, V27, P540	2009	4.205	2012	2014
Bordons M, et al. 2010, *Rev Esp Doc Cient*, V33, P9,	2010	3.974	2012	2014
Leydesdorff L, 2010, *J Informetr*, V4, P644	2010	3.974	2012	2014
Torres-Salinas D, 2011, *Prof Inform*, V20, P111	2011	3.974	2012	2014
Leydesdorff L, 2010, *J Am Soc Inf Sci Tec*, V61, P352	2010	3.974	2012	2014
Torres-Salinas D, 2011, *Prof Inform*, V20, P701	2011	3.974	2012	2014
Docampo D, 2011, *Scientometrics*, V86, P77	2011	3.974	2012	2014

Burst detection based on the citations in the 50 articles most frequently cited per year, 1985–2014. Only first authors' names are used in the references

* Beginning and end of period when activity was most intense

phrases. In cluster #7, where its log likelihood ratio is highest and consequently the most representative, it is identified with terms such as 'indexing services' and 'hypertext', but not others related to the use of information and characteristic of LIS globally such as 'information behaviour' or 'information seeking'. The detection of two related fronts, world wide web (#8) and webometrics (#6), denotes the interest in the technical activities supported by this technology and in the activities, situations, people and social practice interconnected by these technical artefacts.

The duration of the burst periods for the noun phrases related to these clusters is brief, denoting low stability in the attention paid to the subjects formulated by the Spanish community of citing researchers. Only five of the 50 references in the intellectual base are not related to at least one of these biliometric fronts (Table 12: Maybee 2006; Ranganathan and Ganapathy 2002; Guallar and Abadal 2009; Abadal and Guallar 2010; Subramaniam and Youndt 2005). Moreover, the presence of several media—('social media', 'digital media', 'communication', 'television', 'web 2.0'), semantic web- and information literacy-related terms with high sigma values is an invitation to speculate on the possible configuration in the near future of one or several research fronts in these connections.

Overall, semantic differentiation is weak and the references fairly old in all the subspecialities. Network density values are indicative of low relationship density. From that standpoint, the domain was thematically 'open' in the period studied. That may be a sign of fragmented and weakly related cognitive structures in which mutual dependence is not particularly high. These are features associated by Whitley (2000) with specialities in which the degree of technical tasks performed is uncertain.

The main component of the co-citation network covers only 19% of the references. Therefore, although research fronts emerge in this study, most of the research does not lie within thematic structures. The inference may be that the research skills with which the findings are generated and the materials from which they are drawn are not standardised or generalised across the domain. That in turn may denote a prevalence of local research cultures.

One limitation to this study lies in the use of a single source of data, the Web of Science, which lists peer-reviewed journals meeting the citation level requirements in the respective disciplines. A second is related to the deployment of both traditional techniques and others which the authors intend to continue using in future research. That notwithstanding, the simultaneous use of a series of techniques, including structural and content analysis indicators, automatic generation of cluster labels and detection of the subjects appearing with highest intensity in a given period, helps understand the recent past and prevalent subject trends in Spanish library and information science research.

Acknowledgements The authors gratefully acknowledge the insightful comments and suggestions provided by anonymous reviewers. The authors wish to thank Margaret Clark, translator, for her linguistic support.

References

Abadal, E., & Guallar, J. (2010). *Prensa digital y bibliotecas*. Gijon: Trea.

Aguillo, I. (2010). Measuring the institution's footprint in the web. *Library Hi Tech, 27*(4), 540–556.

Albarrán, P., Crespo, J. A., Ortuño, I., & Ruiz-Castillo, J. (2011). The skewness of science in 219 sub-fields and a number of aggregates. *Scientometrics, 88*(2), 385–397.

Ardanuy, J. (2012). Scientific collaboration in library and information science viewed through the web of knowledge: The Spanish case. *Scientometrics, 90*(3), 877–890.

Arquero Aviles, R. (2001). *Análisis de la investigación Española en Biblioteconomía y Documentación: 1975–1984* (Tesis doctoral, Universidad Complutense) http://biblioteca.ucm.es/tesis/inf/ucm-t25424.pdf Accessed 20 May 2017.

Blázquez-Ruiz, J., Guerrero-Bote, V. P., & Moya-Anegón, F. (2016). New scientometric-based knowledge map of food science research (2003–2014). *Comprehensive Reviews in Food Science and Food Safety, 15*(6), 1040–1055.

Bordons, M., Sancho, R., Morillo, F., & Gómez, I. (2010). Scientific activity profile of Spanish universities in four thematic areas: A multifactor analysis. *Revista Espanola de Documentacion Cientifica, 33*(1), 9–33.

Callon, M., Courtial, J. P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Science Information, 22*(2), 191–235.

Callon, M., Courtial, J. P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: the case of polymer chemistry. *Scientometrics, 22*(1), 155–205.

Cano, V. (1999). Bibliometric overview of library and information science research in Spain. *Journal of the American Society for Information Science, 50*(8), 675–690.

Chang, Y.-W., Huang, M.-H., & Lin, C.-W. (2015). Evolution of research subjects in library and information science based on keyword, bibliographical coupling, and co-citation analyses. *Scientometrics, 105*(3), 2071–2087.

Chen, C. (2004). Searching for intellectual turning points: Progressive knowledge domain visualization. *Proceedings of the National Academy of Sciences of the United States of America, 101*(suppl. 1), 5303–5310.

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology, 57*(3), 359–377.

Chen, C., Chen, Y., Horowitz, M., Hou, H., Liu, Z., & Pellegrino, D. (2009). Towards an explanatory and computational theory of scientific discovery. *Journal of Informetrics, 3*(3), 191–209.

Chen, C., Ibekwe-SanJuan, F., & Hou, J. (2010). The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis. *Journal of the American Society for Information Science and Technology, 61*(7), 1386–1409.

Chen, C., Dubin, R., & Kim, M. C. (2014). Emerging trends and new developments in regenerative medicine: A scientometric update (2000–2014). *Expert Opinion on Biological Therapy, 14*(9), 1295–1317.

Chi, R., & Young, J. (2013). The interdisciplinary structure of research on intercultural relations: A co-citation network analysis study. *Scientometrics, 96*(1), 147–171.

Cho, J. (2014). Intellectual structure of the institutional repository field: A co-word analysis. *Journal of Information Science, 40*(3), 386–397.

Ding, W., & Chen, C. (2014). Dynamic topic detection and tracking: A comparison of HDP, C-word, and cocitation methods. *Journal of the American Society for Information Science and Technology, 65*(10), 2084–2097.

Ding, Y., Chowdhury, G. G., & Foo, S. (2000). Journal as markers of intellectual space: Journal co-citation analysis of information retrieval area, 1987–1997. *Scientometrics, 47*(1), 55–73.

Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics, 19*(1), 61–74.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry, 40*(1), 35–41.

García-Lillo, F., Úbeda-García, M., & Marco-Lajara, M. (2016). Organizationl ambidexterity: Exploring the knowledge base. *Scientometrics, 107*(3), 1021–1040.

Garfield, E., Malin, M. V., & Small, H. (1978). Citation data as science indicators. In Y. Elkana et al. (Eds.), *Toward a metric of science: The advent of science indicators* (pp. 179–208). New York, NY: Wiley.

Gmür, M. (2003). Co-citation analysis and the search for invisible colleges: A methodological evaluation. *Scientometrics, 57*(1), 27–57.

González-Alcaide, G., Calafat, A., Becoña, E., Thijs, B., & Glänzel, W. (2016). Co-citation analysis of articles published in substance abuse journals: Intellectual structure and research fields (2001–2012). *Journal of Studies on Alcohol and Drugs, 77*(5), 710–722.

Guallar, J., & Abadal, E. (2009). Evaluación de hemerotecas de prensa digital: Indicadores y ejemplos de buenas prácticas. *Profesional de la Informacion, 18*(3), 255–269.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America, 102*(46), 16569–16572.

Hu, J., & Zhang, Y. (2015). Research patterns and trends of recommendation system in China using co-word analysis. *Information Processing and Management, 51*(4), 329–339.

Jiménez-Contreras, E. (2002). La aportación española a la producción científica internacional en biblioteconomía y documentación: balance de diez años (1992–2001). *BiD*, 9 http://bid.ub.edu/09jimen2.htm. Accessed 11 January 2017.

Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters, 31*(1), 7–15.

Kim, M. C., & Chen, C. (2015). A scientometric review of emerging trends and new developments in recommendation systems. *Scientometrics, 104*(1), 239–263.

Kleinberg, J. (2002). Bursty and hierarchical structure in streams. In *Proceedings of the eighth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 91–101). NewYork, NY: ACM Press.

Lariviere, V., Sugimoto, C. S., & Cronin, B. (2012). A bibliometric chronicling of library and information science's first hundred years. *Journal of the American Society for Information Science and Technology, 63*(5), 997–1016.

Leydesdorff, L. (2008). On the normalization and visualization of author co-citation data: Salton's cosine versus the Jaccard index. *Journal of the American Society for Information Science and Technology, 59*(1), 77–85.

Liu, P., Chen, B.-L., Liu, K., & Xie, H. (2016). Magnetic nanoparticles research: A scientometric analysis of development trends and research fronts. *Scientometrics, 108*(3), 1591–1602.

Liu, Ch., & Gui, Q. (2016). Mapping intellectual structures and dynamics of transport geography research: A scientometrics overview from 1982 to 2014. *Scientometrics, 109*(1), 159–184.

Liu, L., & Mei, S. (2016). Document visualizing the GVC research: A co-occurrence network based bibliometric analysis. *Scientometrics, 109*(2), 953–977.

López-Herrera, A. G., Cobo, M. J., Herrera-Viedma, E., Herrera, F., Bailón-Moreno, R., Jiménez-Contreras, E. (2009). Visualization and evolution of the scientific structure of fuzzy sets research in Spain. *Information Research,* 14(4) http://www.informationr.net/ir/14-4/paper421.html. Accessed 11 January 2017.

Luxburg, U. V. (2007). A tutorial on spectral clustering. *Statistics and Computing, 17*(4), 395–416.

Maybee, C. (2006). Undergraduate perceptions of information use: The basis for creating user-centered student information literacy instruction. *Journal of Academic Librarianship, 32*(1), 79–85.

Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics, 4*(3), 265–277.

Moneda-Corrochano, M. de la. (2016). Las cifras de la enseñanza universitaria en documentación en España: 2014. *Anuario ThinkEPI, 1,* 29–47.

Moya-Anegón, F. de, Jiménez-Contreras, E., & Moneda Corrochano, M. de la. (1998). Research fronts in library and information science in Spain (1985–1994). *Scientometrics, 42*(2), 229–246.

Muñoz-Leiva, F., Viedma-del-Jesús, M. I., Sánchez-Fernández, J., & López-Herrera, A. G. (2012). An application of co-word analysis and bibliometric maps for detecting the most highlighting themes in the consumer behaviour research from a longitudinal perspective. *Quality & Quantity, 46*(4), 1077–1095.

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America, 103*(23), 8577–8582.

Olmeda-Gómez, C., & Moya-Anegón, F. de. (2016). Publishing trends in library and information sciences across European countries and institutions. *The Journal of Academic Librarianship, 42*(6), 27–37.
Persson, O. (1994). The intellectual base and research fronts of JASIS 1986–1990. *Journal of the American Society for Information Science, 45*(1), 31–38.
Plaza, L. M. (2014). La información y documentación científica en el CSIC. *Revista Española de Documentación Científica, 37*(2), e048.
Radicchi, F., Fortunato, S., & Castellano, C. (2008). Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences of the United States of America, 105*(45), 17268–17272.
Ranganathan, C., & Ganapathy, S. (2002). Key dimensions of business-to-consumer web sites. *Information and Management, 39*(6), 457–465.
Ravikumar, S., Agrahari, A., & Singh, S. N. (2015). Document mapping the intellectual structure of scientometrics: A co-word analysis of the journal scientometrics (2005–2010). *Scientometrics, 102*(1), 929–955.
Rip, A., & Courtial, J. P. (1984). Co-words of biotechnology: An example of cognitive scientometrics. *Scientometrics, 6*(6), 381–400.
Rousseau, R., & Zuccala, A. (2004). A classification of author co-citations: Definitions and search strategies. *Journal of the American Society for Information Science and Technology, 55*(6), 513–529.
Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20,* 53–65.
Salton, G., & McGill, M. J. (1986). *Introduction to modern information retrieval*. New York, NY: McGraw-Hill.
Schneider, J. W. (2006). Concept symbols revisited: Naming clusters by parsing and filtering of noun phrases from citation contexts of concept symbols. *Scientometrics, 68*(3), 573–593.
Schvaneveldt, R. W. (Ed.). (1990). *Pathfinder associative networks: Studies in knowledge organization.* Norwood, NJ: Ablex.
Sedighi, M. (2016). Application of word co-occurrence analysis method in mapping of the scientific fields (case study: The field of informetrics). *Library Review, 65*(1/2), 52–64.
Shafique, M. (2013). Thinking inside the box? Intellectual structure of the knowledge base of innovation research (1988–2008). *Strategic Management Journal, 34*(1), 62–93.
Shiau, W.-L. (2016). The intellectual core of enterprise information systems: A co-citation analysis. *Enterprise Information Systems, 10*(8), 815–844.
Shibata, N., Kajikawa, Y., Taked, Y., & Matsushima, K. (2008). Detecting emerging research fronts based on topological measures in citation networks of scientific publications. *Technovation, 28*(11), 758–775.
Shiffrin, R. M., & Börner, K. (2004). Mapping knowledge domains. *Proceedings of the National Academy of Sciences of the United States of America, 101*(suppl. 1), 5183–5185.
Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science, 24*(4), 265–269.
Small, H. (1980). Co-citation context analysis and the structure of paradigms. *Journal of Documentation, 36*(3), 183–196.
Small, H. (2003). Paradigms, citations, and maps of science. *Journal of the American Society for Information Science and Technology, 54*(5), 394–399.
Small, H., & Griffith, B. C. (1974). The structure of scientific literatures I: Identifying and graphing specialties. *Science Studies, 4*(1), 17–40.
Subramaniam, M., & Youndt, M. A. (2005). The influence of intellectual capital on the types of innovative capabilities. *Academy of Management Journal, 48*(3), 450–463.
Thelwall, M., & Sud, P. (2011). A comparison of methods for collecting web citation data for academic organizations. *Journal of the American Society for Information Science and Technology, 62*(8), 1488–1497.
Topalli, M., & Ivanaj, S. (2016). Mapping the evolution of the impact of economic transition on Central and Eastern European enterprises: A co-word analysis. *Journal of World Business, 51*(5), 744–759.
Walters, W. H., & Wilder, E. I. (2016). Disciplinary, national, and departamental contributions to the literature of library and information science, 2007–2012. *Journal of the Association for Information Science and Technology, 67*(6), 1487–1506.
Wang, Z., Zhao, H., & Wang, Y. (2015). Social networks in marketing research 2001–2014: A co-word analysis. *Scientometrics, 105*(1), 65–82.
Whitley, R. (2000). *The intellectual and social organization of the sciences*. New York, NY: Oxford University Press.
Yan, E., & Ding, Y. (2012). Scholarly network similarities: How bibliographic coupling networks, citation networks, cocitation networks, topical networks, coauthorship networks, and coword networks relate to

each other. *Journal of the American Society for Information Science and Technology, 63*(7), 1313–1326.

Yan, B.-N., Lee, T.-S., & Lee, T.-P. (2015). Mapping the intellectual structure of the Internet of Things (IoT) field (2000–2014): A co-word analysis. *Scientometrics, 105*(2), 1285–1300.

Yang, K.-C. (2016). Intellectual structure of trust in business and management: A co-citation analysis. *Electronic Library, 34*(3), 358–370.

Zhao, D., & Strotmann, A. (2008). Evolution of research activities and intellectual influences in information science 1996–2005: Introducing author bibliographic-coupling analysis. *Journal of the American Society for Information Science and Technology, 59*(13), 2070–2086.

Zhao, D., & Strotmann, A. (2014). The knowledge base and research front of information science 2006–2010: An author co-citation and bibliographic coupling analysis. *Journal of the Association for Information Science and Technology, 65*(5), 996–1006.

Zitt, M., & Small, H. (2008). Modifying the journal impact factor by fractional citation weighting: The audience factor. *Journal of the American Society for Information Science and Technology, 59*(11), 1856–1860.