



# Why do papers from international collaborations get more citations? A bibliometric analysis of Library and Information Science papers

A. Velez-Estevez<sup>1</sup> · P. García-Sánchez<sup>2</sup> · J. A. Moral-Munoz<sup>3</sup> · M. J. Cobo<sup>4</sup>

Received: 11 October 2021 / Accepted: 1 August 2022  
© The Author(s) 2022

## Abstract

Scientific activity has become increasingly complex in recent years. The need for international research collaboration has thus become a common pattern in science. In this current landscape, countries face the problem of maintaining their competitiveness while cooperating with other countries to achieve relevant research outputs. In this international context, publications from international collaborations tend to achieve greater scientific impact than those from domestic ones. To design policies that improve the competitiveness of countries and organizations, it thus becomes necessary to understand the factors and mechanisms that influence the benefits and impact of international research. In this regard, the aim of this study is to confirm whether the differences in impact between international and domestic collaborations are affected by their topics and structure. To perform this study, we examined the Library and Information Science category of the Web of Science database between 2015 and 2019. A science mapping analysis approach was used to extract the themes and their structure according to collaboration type and in the whole category (2015–2019). We also looked for differences in these thematic aspects in top countries and in communities of collaborating countries. The results showed that the thematic factor influences the impact of international research, as the themes in this type of collaboration lie at the forefront of the Library and Information Science category (e.g., technologies such as artificial intelligence and social media are found in the category), while domestic collaborations have focused on more well-consolidated themes (e.g., academic libraries and bibliometrics). Organizations, countries, and communities of countries must therefore consider this thematic factor when designing strategies to improve their competitiveness and collaborate.

**Keywords** International research collaboration · Impact analysis · Science mapping analysis · Coword analysis · Library and Information Science

---

✉ A. Velez-Estevez  
antonio.velez@uca.es

Extended author information available on the last page of the article

## Introduction

Over the years, literature on international research collaboration (IRC) has reported an increase in international scientific activity (Gazni et al., 2012). According to Adams (2013), this phenomenon of internationalization is a consequence of the increasing complexity of science, since difficult problems require multidisciplinary teams, as well as a large pool of funding resources (Larivière et al., 2015). Hence, IRC is considered a strategic approach to enhance the competitiveness and economic wealth and prosperity of nations and communities of countries (European-Commission, 2021). Accordingly, nations started to implement science policies to attract new talent, encouraging academics to move abroad and participate in international projects (Suresh, 2012).

The global science system has changed toward the internationalization of science, or what is known as the fourth age of research (Adams, 2013), and nations assess their competitiveness by measuring their scientific production and impact (Franzoni et al., 2011). However, they face the trade-off of cooperating while maintaining a high level of competitiveness (Chinchilla-Rodríguez et al., 2019). This dichotomy is increasing, mainly in two regards: the knowledge required to be competitive and the assets that the countries exclusively have (i.e., the knowledge needed to be competitive might not be fully covered by the country's assets), and the part of the knowledge that is produced domestically and that which is produced via international collaboration (Adams, 2013).

Mobility and migrant researchers are also increasing, as researchers who move abroad tend to achieve higher impact and enhance their careers (Robinson-Garcia et al., 2019; Sugimoto et al., 2017). Indeed, mobility is the pathway to IRC (Kato & Ando, 2017), and nations might lose this talent if returning policies are not implemented efficiently (Adams, 2013).

In this context, IRC has great benefits in terms of increasing impact in comparison with domestic collaboration practices (Chinchilla-Rodríguez et al., 2019; Gazni et al., 2012; Larivière et al., 2015; Narin & Whitlow, 1990; Persson, 2010; Sooryamoorthy, 2017; Sugimoto et al., 2017), which adds complexity to the problem of designing policies to improve the competitiveness of nations. In this sense, IRC has been studied during the last three decades to try to understand and measure its growth, impact, and causes of its increase, and to aid the design of better science policies (Chen et al., 2019). Numerous works have confirmed the increase in collaboration practices (Adams, 2012; Fortunato et al., 2018; Gazni et al., 2012; Larivière et al., 2015; Narin & Whitlow, 1990). Cultural, political, geographical, and linguistic factors have been shown to influence IRC strongly (Frame & Carpenter, 1979; van Raan, 1997). With regards to the causes of this increase, self-citations do not seem to be the cause of the difference in impact between international and domestic collaborations (Van Raan, 1998), and the aforementioned growth is not uniform between fields of study (Gazni et al., 2012). Also, the increased impact of IRC does not affect all countries equally, and indeed a strong dependence on IRC may point to a lack of resources to be independent; countries that do not benefit from IRC may neglect leadership as well as their own development, with negative consequences on linguistic and thematic diversity (Chinchilla-Rodríguez et al., 2019). Publishing in open access and with international collaborators increases impact (Gabrielle Breugelmans et al., 2018). Government funding does not, on average, have a significant effect on the citation impact achieved by IRC studies in the Organisation for Economic Co-operation and Development (OECD) (Leydesdorff et al., 2019). However, there are still unexplored factors that may influence the benefits that different countries and organizations obtain from IRC (Chen et al., 2019).

Moreover, it is worth mentioning that IRC is not the only factor affecting the impact of papers. Tahamtan et al. (2016) concluded that three categories of factors might affect the number of citations: paper-related factors (e.g., the quality of the paper, the novelty of the work, and the characteristics of each field), journal-related factors (e.g., the impact factor and scope of the journal, and the form of publication), and author(s)-related factors (e.g., the number of authors, the international and national collaborations of the authors, and their gender, age, and race).

Moreover, the competitiveness of countries can be related to their low dependence on IRC in some scientific fields (Chinchilla-Rodríguez et al., 2019), and in particular those that are more scientifically relevant (with high relative citation counts). Therefore, to be competitive, it is key to understand the strong themes (in terms of citation impact) that are at the forefront of the research, in which the nations and institutions have low dependencies, to potentiate them even more. In addition, it is also important identifying themes where the nations or institutions are depending too much on external collaboration, which is related to the lack of resources and low competitiveness, to increase their competitiveness in these themes by, for example, attracting foreign researchers or increasing the funding in these topics. Citation impact is partially related to scientific impact as well, as topics with high citation impact hold the attention or interest of the scientific community (Aksnes et al., 2019). Nonetheless, the use of citation impact has been criticized by other authors since it does not reflect all the dimensions of scientific impact, such as solidity/plausibility, originality, and societal value (Aksnes et al., 2019; Wilsdon, 2015). However, citation impact reflects some aspects related to scientific impact and relevance, and has been used as a proxy to carry out research evaluations in different scientific contexts, including institutions, nations Chinchilla-Rodríguez et al. (2019) and Moed (2005), and research proposals (Cabezas-Clavijo et al., 2013), among others. According to Robinson-Garcia et al. (2018), other indicators, such as altmetrics, could extend this coverage of impact to some of these other dimensions, e.g., societal impact.

In addition, owing to the increasing availability of digital data (Fortunato et al., 2018), new tools to analyze science and detect its patterns have emerged. Bibliographical networks and science mapping analysis tools have taken advantage of this opportunity to better understand science and its evolution (Batagelj & Cerinšek, 2013; Cobo et al., 2012; Moral-Muñoz et al., 2020).

Therefore, to define a strategy for international collaboration, there is first a need to understand the features of international collaboration and what factors are causing its increased citations compared with other types of collaboration. The aim of this paper is to provide insights specifically in this vein, focusing on the theme of collaboration types, as expertise in areas of knowledge is one of the exclusive assets that institutions and nations have in order to be competitive (Adams, 2013).

In this paper, we specifically examined the Library and Information Science (LIS) part of the Web of Science database during 2015–2019 through its bibliographical networks, as well as applying science mapping analysis methods. Some works have analyzed themes and research trends in the LIS category (e.g., Bauer et al. (2016), Galvez (2018), Han (2020), Han et al. (2014), Hsiao and Hua Chen (2020), Ma and Lund (2021), Mokhtarpour and Khasseh (2020), Olmeda-Gómez et al. (2017), Yan (2015), Yan et al. (2010)), and the benefits of IRC in LIS have also been confirmed in literature (Asubiaro, 2019; Sin, 2011). However, to the best of the authors' knowledge, no study has examined the themes and how they are structured in terms of types of collaboration in the LIS field, nor has it been shown whether they might be one of the factors that explain the gain in impact of IRC in LIS.

The main objective of this paper is thus to identify whether the asymmetry in impact between research collaboration types might be caused by thematic and structural factors, which is a key focus in IRC (Chen et al., 2019). These factors can be verified at multiple levels, and we focus herein on a comparison of the themes and structure between collaboration types as a whole, as well as at the levels of country and of communities of collaborating countries. Understanding how the knowledge base of a field is disseminated among local, national, and international research, as well as in countries and communities of cooperating countries, is fundamental to help the design of future policies. We thus seek to answer the following research questions:

- *RQ1* Are there any differences in citation impact between research collaboration types?
- *RQ2* Are there any topological differences within the conceptual network between the different collaboration types?
- *RQ3* Are there any topological differences within the social network between the different collaboration types?
- *RQ4* Are there any differences in themes and their impact on the different collaboration types?
- *RQ5* Are there any differences in the structure of themes of collaboration types according to the strategic diagram?
- *RQ6* Are there any differences in top countries according to themes for different collaboration types?
- *RQ7* Are there any differences in groups of collaborating countries at the international collaboration level according to international themes?

To answer these research questions, the papers in the LIS category published during 2015–2019 were retrieved from the Web of Science database, and the resulting corpus was analyzed by means of science mapping analysis, bibliographical networks, and performance measures.

The remainder of this paper is organized as follows: First, the methodology used to carry out the study is explained. Then, the results obtained are presented and described. Next, the discussion regarding all the results from a “whole” perspective and the findings of the study are presented. Finally, the conclusions and possibilities for future work are described.

## Methodology

In this section, the process used to retrieve, analyze, and visualize the data is detailed to address the stated research questions. Before explaining the specific methods used, the retrieval and cleaning of the data are specified. To perform this study, the required data could be retrieved from several bibliographical databases, such as Scopus, Web of Science, Dimensions, or ScholarMetrics, among others (Martín-Martín et al., 2018; Visser et al., 2021). For this paper, data were collected from the Web of Science using the query *WC=“Information Science & Library Science” AND PY = 2015–2019 AND DT = (ARTICLE OR REVIEW)* to search in the Science Citation Index Expanded (SCIE) and Social Sciences Citation Index (SSCI). The authors acknowledge that delimiting the field using a whole category could introduce limitations since some papers might not be related. However, it is important to note that limiting the field of LIS with a query on the basis of

keywords is very complex and may introduce biases. Moreover, selecting journals manually also increases the problem of biases. We thus chose to limit the field using a standard category, since this is the best way to ensure an objective and clear criterion to select all papers related to the LIS field. It should also be pointed out that our objective is to analyze the field of LIS rather than a subfield such as bibliometrics.

A cleaning process was then applied to the data by removing papers without affiliations and preprocessing the author's keywords. After that, the author's keywords representing the same concept were joined (e.g., artificial neural networks, ANN, neural networks; CRIS, CRIS System; current research information system; *h*-index, Hirsch index), and authors' keywords with a broad meaning, known as stop words, were removed (e.g., adolescent/s; algorithm/s; case study/studies; number/s). This preprocessing was performed using the SciMAT software (Cobo et al., 2011b, 2012; Moral-Muñoz et al., 2020) by joining singular with plural words automatically. After that, the authors of this paper manually revised the whole set of keywords to join them if and only if the words represented the same concept. Next, the "whole" dataset was created, and it was divided into three different collaboration types according to previous studies (Chinchilla-Rodríguez et al., 2019; Gazni et al., 2012):

- *Local type* Papers in this category must include only one organization and one country.
- *National type* Papers with national collaboration must include a number of organizations greater than one and include only a single country.
- *International type* Papers are classified under this category when two or more countries are collaborating.

It is important to clarify that, apart from the aforementioned types of collaboration, this study also utilized the whole dataset. After this common process was finished, the research questions could be addressed by using specific methods over the four datasets obtained (i.e., whole, local, national, and international).

Regarding the question of asymmetry in citation impact between the different collaboration types (RQ1), a variety of indicators related to citation impact, such as the *h*-index, the geometric mean of citations (Thelwall, 2016), and the standard deviation and median of citations, were measured. Moreover, and specifically focusing on highly cited papers (HCP), they were extracted utilizing two methods:

- *H-classics* for each year of the whole dataset, we measure the *h*-index (Hirsch, 2005), and papers belonging to the *h*-core were identified as HCP (Martínez et al., 2014). The *h*-core is the set of documents with *h* or more citations, with *h* being the Hirsch index (Rousseau, 2006).
- *Best 1%* for each year of the whole dataset, we sort the papers by number of citations and identify 1% of each year as HCP.

General measures were also computed to understand how papers are divided into the collaboration types, specifically the number of papers, percentage of papers, total citations, number of uncited papers, and percentage of uncited papers.

To address whether the differences in citations could be caused by topological differences in the conceptual (RQ2) or social networks (RQ3), we built two different networks, based on coworkers and coauthors. These networks were built using the author's keywords and author names from each document, which are the nodes in the networks. The edges joining two nodes are representative if two keywords or authors (depending on the network built) appeared together in the documents or not. The network has two attributes: the

frequency of a node (i.e., how many times this node appeared in the documents), and the co-occurrence frequency of two nodes (or the weight of an edge), which is the number of times that the two nodes appear together in a document. To precisely address the question of finding social and conceptual topological differences in collaboration types, we computed the following topological properties of the constructed networks: average degree, diameter, assortativity based on the degree, number of components, number of nodes of the biggest component (absolute and relative), average path length, and density. The number of nodes and number of edges were also computed to understand the size of the networks.

Furthermore, to reveal the themes and citation impact in the different collaboration types (RQ4), we detected the communities (themes) in the “whole” coword network, and extracted the citation impact for each of them. To do so, a normalization process was first performed over the frequency of co-occurrence (edge weight) by means of the equivalence index (Callon et al., 1991). After that, we extracted the “whole” themes by applying the Leiden community detection algorithm (Traag et al., 2019) to the normalized “whole” coword network. The Leiden algorithm was chosen owing to its advantages (e.g., guarantees regarding well-connected communities) over other algorithms.

After extracting the themes, the documents with keywords belonging to each theme were also analyzed. For this study, the documents of a theme are the union of the documents with any of the keywords of the theme. Using these, we measured the citation impact and general indicators of the themes found in the whole network. It is worth mentioning that, in these “whole” themes, the documents can have either local, national, or international collaboration. Therefore, to address RQ4, the global citation impact of each theme was measured by means of the following: citations geometric mean, the *h*-index, mean normalized citation score (MNCS) (Waltman et al., 2011), total and percentage of citations, percentage of uncited papers, and percentage of HCP. The percentage of HCP in each theme was measured in two ways: over the total number of HCP identified in the first research question (regardless of the type of collaboration), and over the number of papers related to the theme. The former facilitates the identification of themes with higher ratio of HCP, with the values being comparable between types of collaboration, while the latter helps to identify high-impact themes in the context of each collaboration type. We also measured the percentage of papers with each collaboration type.

The following citation impact indicators for the “whole” themes and collaboration types were also measured: percentage of uncited papers, citations’ geometric mean, and HCP (percentage of HCP over the total papers with the collaboration type related to the theme). The percentage of papers by collaboration type was also measured to obtain the degree of collaboration of each theme.

To provide a more complete picture of the collaboration types, we also carried out the same process but splitting the dataset before constructing the network. The themes extracted in this analysis thus provide a detailed overview of each collaboration type. For the specific themes extracted from each collaboration type, we measured the same indicators for the “whole” analysis without the specific measures for each collaboration type, as in these themes only papers with a specific collaboration type are found. Therefore, the HCP percentage measure for each theme over the “whole” HCP papers helps to compare the citation impact of particular themes according to collaboration type.

The structure of collaboration types (RQ5) was revealed by measuring the internal and external cohesion of the themes according to their density and centrality. With these two measures, the themes of each collaboration type (i.e., whole, local, national, international) were plotted in a strategic diagram (Cobo et al., 2011a). The strategic diagram classifies the themes into four categories: motor, basic and transversal, highly developed and isolated,

and emerging or declining. Motor themes (upper right) are those related externally to concepts that are applied to other themes that are conceptually close. Highly developed and isolated themes (upper left) have strong internal links, but not important external ties so have marginal importance for the field. In the bottom left part, themes are weakly developed and marginal, mainly representing emerging or disappearing themes. Themes in the bottom right part of the diagram are basic and transversal, since they have strong ties with the rest of the themes but are not developed. The themes were labeled by selecting the most central keyword of the cluster. Using this analysis and these visualizations, the differences and similarities of the structure between collaboration types were studied.

Moreover, to analyze the differences of top productive countries according to the themes (RQ6), the relative contribution of each one in relation to the themes in each type of collaboration was compared. To achieve this, we focused on the top ten countries globally (i.e., selecting the top ten countries with the most papers published in the whole dataset). For each type of collaboration, the top five countries not included in the global ranking were selected as well. In the case of the whole network, we focused on the top 15 countries. Consequently, we could observe how the research output of top countries was distributed throughout the themes and compared the “whole” ones concerning the types of collaboration. To complement the analysis of top countries by collaboration type, we also studied the coauthors’ networks of each top country, thereby showing how the knowledge base of countries varies according to collaboration practices.

Regarding the above-mentioned fourth age of research, as countries cooperate among themselves, it is interesting to study the communities of countries and their respective knowledge bases (themes) (RQ7). To do so, we built a co-country network using the international collaboration dataset. We then applied the Leiden algorithm over the whole co-country international network (Traag et al., 2019) to find communities of countries that collaborate together. Subsequently, we obtained the distribution of the papers in the international themes extracted when answering RQ4. Finally, once the results were obtained and combined with the performance measures of the detected themes, we studied how research in communities of countries varies, taking into consideration the impact of the themes.

## Results

According to the methodology described in “[Methodology](#)” section, a bibliometric analysis was performed, aiming to analyze whether one of the factors for the increase of IRC impact is the existence of a difference in the conceptual structure of collaboration types. A total of 22,127 papers were retrieved on 6 October 2020 using the advanced query explained above.

### Differences in the citation impact (RQ1)

To address RQ1, we carried out a quantitative analysis of the impact according to collaboration type as explained in “[Methodology](#)” section. The results (Table 1) showed that there exists a difference in impact, as noted in the geometric mean of citations. This benefit can also be seen in the number of HCP, either by 1% or the H-classics method, as they increase when the collaboration level is higher, whereas the proportion of uncited papers decreases

**Table 1** Impact measures according to collaboration type

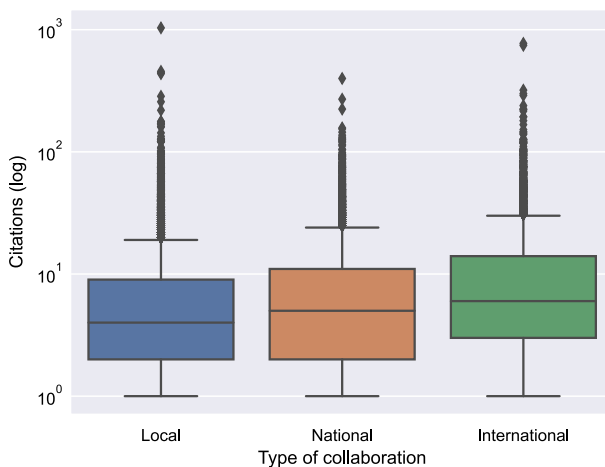
Measure	Local	National	International
<i>H</i> -index	70	61	75
Citations (geometric mean)	4.14	5.13	6.3
Citations (standard deviation)	17.77	15.05	23.39
Citations (median)	3	4	5
Uncited papers	2069	703	636
Uncited papers (%)	20.37%	14.78%	11.75%
HCP (H-classic)	92	55	125
HCP (H-classic) (%)	0.91%	1.16%	2.31%
HCP (1% whole)	65	43	96
HCP (1% whole) (%)	0.64%	0.9%	1.77%
Papers (%)	49.98%	23.39%	26.63%
Number of papers	10,158	4755	5413
Total citations	69,244	39,995	60,863

To compute the geometric mean, zeros were treated by adding one to every value

accordingly. Papers with local collaboration represent half of the total, while national and international papers accounted for the rest, with similar amounts.

Note that 1801 papers had no affiliation information (i.e., neither the country nor the organization) and were discarded (these 1801 papers represent 8.13% of the 22,127 papers). Therefore, the sum of the papers from the three types of collaboration was not equal to the total number of retrieved papers.

In addition, as shown in Fig. 1, there is a pattern in the citation distributions of the types of collaboration. As the collaboration level increases, the number of outliers decreases, and the citation distribution shifts to higher citations. This can be noticed in the maximum



**Fig. 1** Boxplot of local, national, and international collaboration citations. The number of citations is plotted on a log scale; zeros were treated by adding one to each value of the distribution



**Table 2** Topological properties of the conceptual networks

	Local	National	International	Whole
#nodes	19,776	11,366	13,954	35,578
#edges	100,438	49,975	63,089	201,939
Avg. degree	10.2	8.8	9	11.3
Diameter	11	10	10	10
Degree assortativity	- 0.03	0	- 0.01	- 0.04
#Components	279	225	228	380
#Nodes P. Component	18,498	10,351	12,978	33,939
%Nodes P. Component	93.5%	91.1%	93%	95.4%
Avg. path length	3.8	4.1	4	3.7
Density	0.05%	0.08%	0.06%	0.03%

**Table 3** Structural measures of the coauthor networks

	Local	National	International	Whole
#nodes	17,238	13,642	13,434	37,878
#edges	23,833	36,617	41,671	99,910
Avg. degree	2.8	5.4	6.2	5.3
Diameter	25	19	20	27
Degree assortativity	0.7	0.5	0.9	0.7
#Components	5810	2037	1620	6,59
#Nodes P. Component	1840	4722	6233	18,924
%Nodes P. Component	10.7%	34.6%	46.4%	49.9%
Avg. path length	9.1	7.2	7	7.3
Density	0.02%	0.04%	0.05%	0.01%

number of citations without outliers, as well as the median of citations, which is higher in each collaboration type than in the previous one (local < national < international).

In addition, as shown in Fig. 1, the maximum of citations without outliers is higher when the collaboration level increases, and the outliers are more reduced in national and international collaboration in comparison with the local collaboration type. The 25% of most cited papers increases when the academic collaboration is higher; this can be noticed by the third quartile being higher in each collaboration type.

**Topological analysis of conceptual networks (RQ2)**

Analyzing the conceptual and social structure of the LIS field is fundamental to discovering their effects on the scientific impact of collaboration levels. Aiming to understand how scientific impact of collaboration types may be influenced by social and conceptual structures, we studied the coauthor and cword networks. Therefore, Tables 2 and 3 present different measures to characterize the cword and coauthor collaboration networks: the size of each collaboration network (#nodes), the number of links between them (#edges), the average number of edges incident on nodes (Avg. degree), the largest distance between any pair of nodes (diameter), the tendency of a node to be linked to similar nodes on the basis of the degree (assortativity), the number of network-independent components, the number of nodes in the biggest component (#Nodes P. Component), the percentage of nodes of the

biggest component [Nodes. P. Component (%)], the average number of steps to reach any possible node (Avg. path length), and the density of the network (i.e., the percentage of edges of the possible ones in a complete network). We found that the conceptual networks do not follow any pattern on the basis of the degree of their nodes, as there is no assortativity based on the degree (i.e., there is no degree-based homophily), as presented in Table 2. In fact, this shows that specific concepts with low connections can be interrelated with each other and with other strongly connected concepts in the network following no pattern.

### Topological analysis of social networks (RQ3)

In the social network (Table 3) the number of authors is very similar, with the exception of the local collaboration type, which has around 3000 more authors compared with the rest. The average degree of national and international social networks is higher, which means that the social network of authors is bigger, on average, for these collaboration types. Also, the assortativity is remarkable in the case of the international social network, which points to strong homophily on the basis of the links between authors.

Components are very reduced in the international collaboration network compared with the rest, and the number of nodes in the largest component is also greater. The average path length between nodes in the national and international collaboration types is much shorter than in the local collaboration type. All of this means that authors are closer to other authors (i.e., the number of researchers between a researcher and the rest of the network is smaller), increasing the opportunity to collaborate.

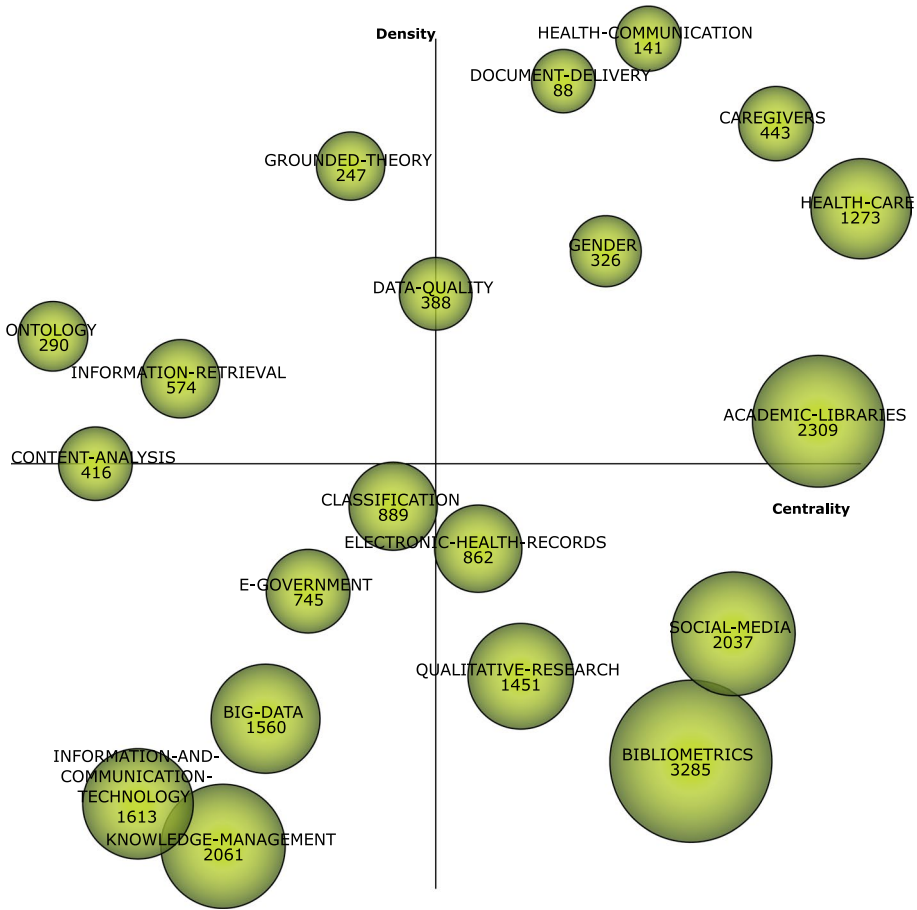
### Themes and structure of the collaboration types (RQ4 and RQ5)

In this section, we present the themes detected for each collaboration type and for the whole dataset in different subsections, and in each subsection, the themes and their performance measures are described (RQ4). In each subsection, the structure of the collaboration type is also revealed by plotting out the themes in a strategic diagram, which is then commented upon to address RQ5. On the basis of these results, RQ4 and RQ5 will be discussed globally.

Before presenting the results of the themes, it is important to clarify that, in the rest of the paper, the name of the themes will be shown in *italic*. Moreover, bold themes in the performance and impact measures tables are the ones with the highest impact (MNCS). It is also important to point out that a theme is composed of a sub-coword network; thus, the nodes enclosed in a theme could determine the concepts on which the theme is focused.

### LIS whole themes

The “whole” strategic diagram is shown in Fig. 2. For these themes, the general performance measures are presented in Table 4 while the collaboration-type-related measures are presented in Table 5. As Fig. 2 shows, the primary “whole” themes in the field were: *bibliometrics*, focused on citation, altmetrics, and social network analyses; *academic libraries*, focused on information literacy, collaboration, and higher education; and *social media*, focused on social networks such as Facebook and Twitter. Other themes are also distinctive owing to the mean normalized impact they achieved: *big data*, related to natural language processing, sentiment analysis, and text mining; *knowledge management*, related



**Fig. 2** Whole dataset strategic diagram. The number under the themes signifies the number of papers that belong to that theme

to knowledge sharing and information management; and *information retrieval*, related to search engines and meta-analysis.

Concerning the themes with the highest impact, the results show that most of the themes were related to new technologies. In fact, *big data*, *social media*, *information and communication technology*, *information retrieval*, or *e-government* are very important in the LIS category regarding the impact they achieved. In contrast, more traditional themes, such as *academic libraries*, *qualitative research* (i.e., information behavior, digital libraries), or *content analysis*, are below the world average citation in the field ( $MNCS < 1$ ).

Moreover, focusing on the high-impact themes and the uncited papers within them, papers with international collaboration contributed more in relative terms to impact, since there were fewer uncited papers with international collaboration. In almost all the themes, the HCP were also international, and it is important to note that the national and local HCP were usually found in high-impact themes.

As can be seen in Table 5, themes that achieved an above-average impact mostly have an intense degree of international collaboration. However, an important part of

**Table 4** Whole LIS themes, general performance, and impact measures

Theme	DOC ▼	CGM	HI	CIT	CIT%	UP%	HCHC%W	HCHC%T	HCI%W	HCI%T
Bibliometrics	3015	5.3	52	26,891	15.8	13.4	12.9	1.2	11.3	0.8
Academic-Libraries	2433	3.6	39	13,553	8	22.1	4.8	0.5	4.4	0.4
<b>Social-media</b>	<b>1945</b>	<b>6.9</b>	<b>62</b>	<b>24,758</b>	<b>14.6</b>	<b>11.6</b>	<b>22.8</b>	<b>3.2</b>	<b>21.6</b>	<b>2.3</b>
Knowledge-management	1937	6.6	61	23,963	14.1	11.7	22.1	3.1	20.6	2.2
Information-and-communication-technology	1644	6.1	53	18,106	10.6	11.1	15.8	2.6	16.2	2
<b>Big-data</b>	<b>1510</b>	<b>6.7</b>	<b>56</b>	<b>19,958</b>	<b>11.7</b>	<b>11.9</b>	<b>18</b>	<b>3.2</b>	<b>19.6</b>	<b>2.6</b>
Health-care	1124	5.5	35	9233	5.4	12.4	1.5	0.4	2	0.4
Electronic-health-records	819	6.5	40	9032	5.3	12.6	4	1.3	4.4	1.1
Classification	808	3.9	32	5587	3.3	23.9	3.7	1.2	3.9	1
E-government	782	6.1	44	9032	5.3	12.1	7.7	2.7	5.9	1.5
Qualitative-research	779	4.6	27	5261	3.1	14.5	0.7	0.3	1.5	0.4
Information-retrieval	542	5.8	36	6580	3.9	12.5	8.5	4.2	8.3	3.1
Content-analysis	400	4.6	24	2965	1.7	18.8	1.8	1.2	1.5	0.8
Gender	366	5	22	2783	1.6	12.6	0.7	0.5	0.5	0.3
Data-quality	365	6.2	29	4169	2.5	14	2.6	1.9	2	1.1
Caregivers	292	5.4	19	2154	1.3	8.6	0.7	0.7	1	0.7
Decision-making	288	5.7	25	2563	1.5	12.2	1.8	1.7	2	1.4
Grounded-theory	244	5	16	1776	1	9	0.4	0.4	0.5	0.4
Health-communication	131	3.8	13	767	0.5	22.9	0	0	0	0

DOC ▼ Number of papers (ordered), CGM Citations geometric mean, HI Hirsch index, CIT Citations, CIT% Percentage of citations over total citations of the dataset, UP% Percentage of uncited papers, HCHC%W Percentage of HCP (H-classes) over whole HCP, HCHC%T Percentage of HCP (H-classes) over theme papers, HCI%W Percentage of HCP (1%) over whole HCP (1%), HCI%T Percentage of HCP (1%) over theme papers  
 Themes with the highest MNCS are highlighted in bold

**Table 5** Performance measures and impact indicators for whole LIS themes by type of collaboration

Theme	MNCS ▼	1%	N%	L%	UI%I	UN%N	UL%L	CGMI	CGMN	CGML	HCHCI%I	HCHCN%N	HCHCL%L	HCI%I%I	HCI%N%N	HCI%L%L
<b>Big-data</b>	1.7	31.3	24.8	43.9	8	12	14.6	8.3	6.4	5.8	3.8	2.9	3	3.2	2.7	2.3
<b>Social-media</b>	1.6	32	21.6	46.4	8.5	11.2	13.9	8.5	7.1	5.8	4.3	2.9	2.5	2.9	1.9	2
Knowledge-management	1.6	35.4	20.3	44.3	9.3	10.7	14	8.5	6.3	5.4	4.8	3.1	1.7	3.6	2.5	0.8
Information-retrieval	1.5	21	25.1	53.9	5.3	14	14.7	7.9	5.8	5.1	5.3	4.4	3.8	2.6	4.4	2.7
E-government	1.4	27	19.3	53.7	11.8	12.6	12.1	8.8	5.6	5.3	5.2	1.3	1.9	4.3	0.7	0.5
Information-and-communication-technology	1.4	33.9	19.6	46.5	8.2	9.6	13.9	7.2	6.3	5.4	3	1.6	2.7	2.9	1.2	1.7
Electronic-health-records	1.4	17.9	42.5	39.6	11.6	9.8	16	7.1	7.4	5.3	0.7	1.7	1.2	1.4	1.1	0.9
Data-quality	1.3	29	21.4	49.6	12.3	11.5	16	8.1	6.9	5	0.9	3.8	1.7	0	3.8	0.6
Decision-making	1.2	29.5	27.4	43.1	12.9	11.4	12.1	6.5	6.2	4.9	1.2	1.3	2.4	0	1.3	2.4
Bibliometrics	1.1	28.4	20.1	51.5	10.8	12.9	15.1	6.4	5.1	4.9	1.9	0.2	1.2	0.9	0.2	0.9
Caregivers	1	26.4	32.5	41.1	7.8	10.5	7.5	6.2	5.3	5	2.6	0	0	2.6	0	0
Health-care	1	23.1	29.4	47.4	12.3	10	13.9	6.7	5.9	4.9	0.8	0.6	0	0.8	0.6	0
Grounded-theory	1	23.8	30.3	45.9	15.5	5.4	8	4.9	5.6	4.7	0	1.4	0	0	1.4	0
Qualitative-research	0.9	25.2	22.3	52.5	12.2	10.3	17.4	6.2	4.9	4	0.5	0	0.2	0.5	0.6	0.2
Content-analysis	0.9	17.2	22.8	60	14.5	11	22.9	8	5.6	3.6	4.3	2.2	0	2.9	1.1	0
Gender	0.9	26.2	24.9	48.9	13.5	12.1	12.3	5.4	5.3	4.6	1	1.1	0	1	0	0

**Table 5** (continued)

Theme	MNCS	I%	N%	L%	UI%	UN%	UL%	CGMI	CGMN	CGML	HCHCI%	HCHCN%	HCHCL%	HC1%I	HC1%N	HC1%L
Classification	0.9	23.1	18.6	58.3	15.5	28	25.9	5.6	3.6	3.5	2.7	0	1.1	1.6	0	1.1
Health-communication	0.9	11.5	24.4	64.1	40	12.5	23.8	3.7	4.7	3.6	0	0	0	0	0	0
Academic-libraries	0.7	19.6	19.4	61	14.2	21.2	24.9	4.6	3.7	3.3	1	0.4	0.4	0.8	0.4	0.2

*MNCS* ▼ Mean normalized citation score (ordered), *I, N, L* % Percentage of {international, national, local} papers, *U/I, N, L* % {*I, N, L*} papers, *U/I, N, L* % {*I, N, L*} Percentage of uncited {international, national, local} papers over {international, national, local} papers of the theme, *CGM/I, N, L* Citations geometric mean of {international, national, local} papers, *HCHC/I, N, L* % {*I, N, L*} Percentage of {international, national, local} HCP (H-classes) over {international, national, local} papers of the theme, *HC1/I, N, L* % {*I, N, L*} Percentage of {international, national, local} HCP (1 %) over {international, national, local} papers of the theme

the high-impact themes were papers with local collaboration. In themes with an impact above the world average ( $MNCS > 1$ ), HCP were mainly found in the international collaboration type, while the percentage of uncited papers was higher in the domestic collaboration types (i.e., local and national).

In addition, most of the international papers were found in high-impact themes (except for *bibliometrics*). Thus, international papers not only acquired a high impact in these themes, but national and local papers also increased it (see CGMI, CGMN, and CGML in Table 5) compared with other themes, such as *health communication* or *academic libraries*. In contrast, international papers in more traditional themes, such as those mentioned above, acquired very low impact compared with other themes.

From this perspective, international collaboration within a theme had, in most cases, more citations than local or national collaboration (except in the themes of *electronic health records*, *grounded theory*, and *health communication*). The impact of the theme will thus depend on the topic in question, meaning that papers on technological themes may have more impact than papers on traditional themes, and local papers on high-impact topics may have a greater impact than international papers on traditional topics. For instance, a local paper published on the *big data* theme may achieve a higher impact, on average, than an international paper on the *academic libraries* theme. This reinforces the idea that the themes are an important factor in the impact achieved by collaboration types, since a high proportion of papers with international collaboration belonged to high-impact themes (e.g., big data, social media).

Low-impact themes were less international, the percentage of uncited papers was higher, and the number of HCP was lower in them, independent of collaboration type. However, in the low-impact themes, the collaboration type with the most uncited papers was local collaboration, and HCP were mainly international. Although the international collaboration papers contributed to the impact of low-impact themes, it is noted how the international uncited papers, as well as the percentage of HCP, decreased when compared with modern themes such as *big data* or *social media*. This also suggests an influence of the themes on the scientific impact of collaboration types.

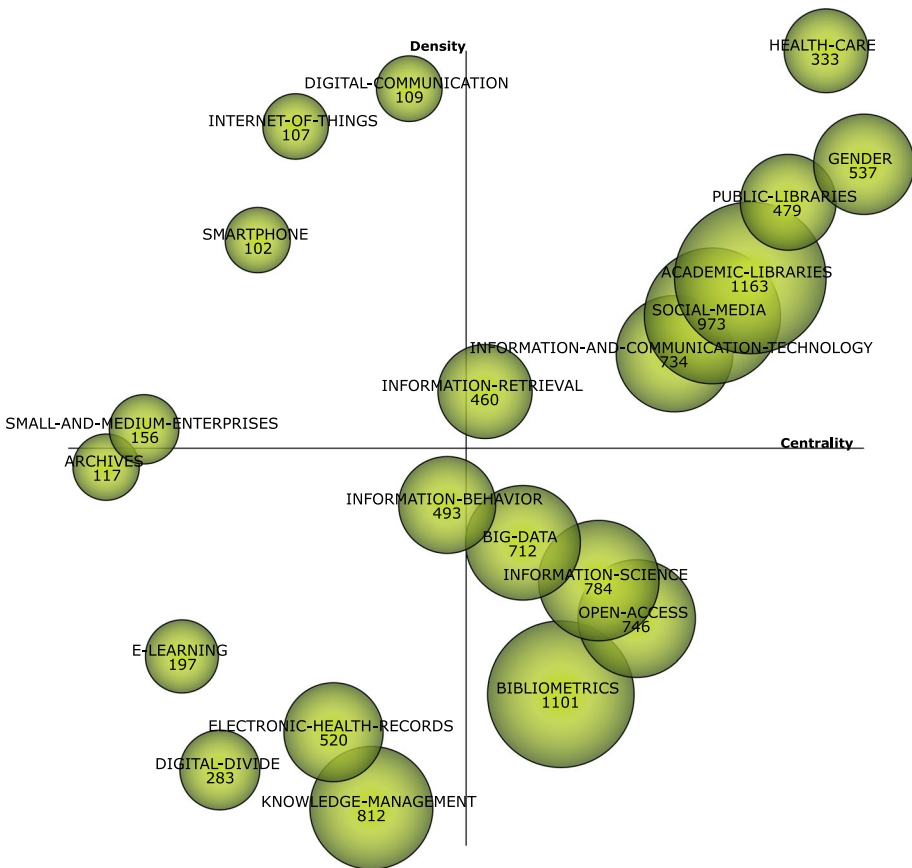
Nonetheless, as presented in Table 5, not all of the high-impact themes had a high rate of papers with international collaboration, although there were some noteworthy exceptions. For instance, the theme *information retrieval* had high normalized impact, but it was more present in the local collaboration type than in the international one. Additionally, this theme had a minimal number of uncited international papers while maintaining a high degree of HCP in all three types of collaboration. In contrast, *bibliometrics* is a theme with an appreciable degree of international collaboration, but remained in the middle- to high-impact themes. Moreover, it was the most important theme in our study considering the number of papers, but the percentage of HCP, either 1% or H-classics, was lower than in other themes, regardless of collaboration type.

Regarding the structure of the LIS “whole” themes, the category was focused on themes such as *bibliometrics*, *academic libraries*, or *health care*, which are motor and basic themes. However, it is observed that the themes with the highest number of citations and proportion of HCP were other themes that are neither motor nor basic, implying that the themes that achieved high scientific impact are novel themes and do not belong to the core of the conceptual structure. In the coming years, these novel themes will probably become more integrated into the category.

### Local collaboration

As shown in Fig. 3 and presented in Table 6, local collaboration focused on the following themes: *Academic libraries* related to information literacy, collaboration, and higher education; *bibliometrics* related to citation analysis, altmetrics, and social network analysis; and *social media* related to topics such as social networks, social networking sites, and political communication. In addition, new technologies such as *big data* have a place in this collaboration type, focusing on concepts such as e-government, transparency, open data, and developing countries. Yet another important theme was *information and communication technology*, which is related to content analysis, digital media, and journalism.

The themes with the strongest impact were *internet of things* and *digital divide*. The former is related to digitalization, academic librarians, and blockchain, which is a new technology, while the latter is related to cloud computing, user experience, and customer relationship management. These last two themes had a MNCS of 2.32 and 1.41, respectively, as presented in Table 6. In contrast, the themes that achieved less impact were *academic libraries* and *information science*, with MNCS of 0.63 and 0.68, respectively. HCP were



**Fig. 3** Local collaboration strategic diagram. The number under the themes signifies the number of papers that belong to that theme



**Table 6** Local collaboration: performance and impact measures of themes. Bold themes are the ones with the highest MNCS

Theme	DOC ▼	CGM	HI	MNCS	CIT	CIT%	UP%	HCHC%W	HCHC%T	HC1%%W	HC1%%T
Academic-Libraries	1163	3.6	29	0.6	6323	9.1	22.8	3.3	0.8	2	0.3
Bibliometrics	1101	5.7	41	1.1	11,047	16	12	6.6	1.6	6.4	1.2
Social-media	973	5.6	42	1.2	9958	14.4	15.3	8.5	2.4	7.8	1.6
Knowledge-management	812	5.5	35	1.1	7531	10.9	13.9	5.5	1.8	4.4	1.1
Information-science	784	3.3	25	0.7	4118	5.9	26.4	2.2	0.8	2.9	0.8
Open-access	746	4.2	27	0.8	4934	7.1	18.8	1.8	0.7	2.5	0.7
Information-and-communication-technology	734	3.9	29	0.8	4864	7	20.8	2.2	0.8	2.5	0.7
Big-data	712	5.3	37	1.1	6897	10	13.8	7	2.7	3.4	1
Gender	537	4.7	23	0.9	3584	5.2	13	1.1	0.6	1.5	0.6
Electronic-health-records	520	5.9	29	1.3	5134	7.4	13.5	2.6	1.3	2.9	1.2
Information-behavior	493	4.1	20	0.7	2642	3.8	16	0.7	0.4	1	0.4
Public-Libraries	479	4.2	24	0.8	3053	4.4	17.1	0.4	0.2	0.5	0.2
Information-retrieval	460	4	20	0.9	3660	5.3	18.5	2.2	1.3	2.9	1.3
Health-care	333	4.8	22	0.9	2382	3.4	15.9	0	0	0	0
<b>Digital-divide</b>	<b>283</b>	<b>5.8</b>	<b>27</b>	<b>1.4</b>	<b>3455</b>	<b>5</b>	<b>15.5</b>	<b>5.1</b>	<b>4.9</b>	<b>3.9</b>	<b>2.8</b>
E-LEARNING	197	5.5	19	1.1	1743	2.5	13.7	1.1	1.5	1	1
Small-and-medium-enterprises	156	6	17	1.1	1289	1.9	8.3	1.5	2.6	1	1.3
Archives	117	3.7	13	0.7	574	0.8	22.2	0	0	0	0
Digital-communication	109	4.5	13	1.2	1117	1.6	16.5	0.7	1.8	1	1.8
<b>Internet-of-things</b>	<b>107</b>	<b>5.8</b>	<b>18</b>	<b>2.3</b>	<b>1430</b>	<b>2.1</b>	<b>14</b>	<b>2.9</b>	<b>7.5</b>	<b>2.9</b>	<b>5.6</b>
Smartphone	102	4.8	14	0.9	767	1.1	17.6	0.4	1	0	0

mostly found in technology-related themes, such as *digital divide*, *social media*, and *big data*, but also in *bibliometrics*. This collaboration type also provides a high number of HCP with respect to the total of the whole dataset, although it must be pointed out that local collaboration contained 49.98% of the corpus. Interestingly, the most productive themes in the local collaboration type had a remarkable rate of uncited papers, which may indicate lower impact (MNCS).

Regarding the structure of local collaboration, high-impact themes (i.e., *digital divide*, *internet of things*) were found on the left-hand side of the strategic diagram (low centrality). As noted in the previous section, this implies that high-impact themes were novel themes that do not belong to the core of the LIS category.

### National collaboration

The following themes regarding national collaboration type were addressed in Fig. 4 and Table 7: *bibliometrics*, *academic libraries*, and *social media*, all focusing on the same topics as in local collaboration type; and *electronic health records*, which adds new technologies to the health context, such as machine learning and natural language processing.

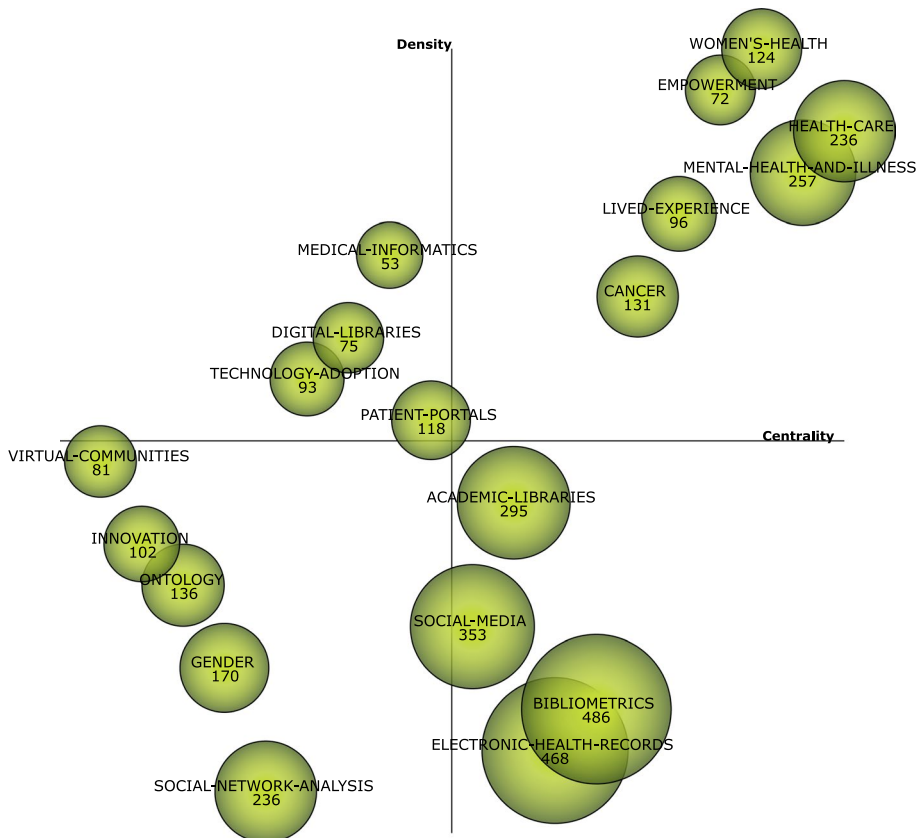


Fig. 4 National collaboration strategic diagram

**Table 7** National collaboration: performance and impact measures of themes. Bold themes are the ones with the highest MNCS

Theme	DOC ▼	CGM	HI	MNCS	CIT	CIT%	UP%	HCHC%W	HCHC%T	HCI%W	HCI%T
Bibliometrics	486	5.2	26	1	4077	10.2	14	1.1	0.6	1.5	0.6
<b>Electronic-health-records</b>	<b>468</b>	<b>6.7</b>	<b>33</b>	<b>1.6</b>	<b>5313</b>	<b>13.3</b>	<b>10</b>	<b>3.7</b>	<b>2.1</b>	<b>4.9</b>	<b>2.1</b>
Social-media	353	6.8	29	1.5	4158	10.4	10.5	3.3	2.5	3.4	2
Academic-Libraries	295	3.7	17	0.6	1554	3.9	20	0.4	0.3	0	0
Mental-health-and-illness	257	5.3	19	1	2082	5.2	12.8	0.4	0.4	0.5	0.4
Social-network-analysis	236	6.9	25	1.3	2395	6	7.2	0.4	0.4	0	0
Health-care	236	6.5	23	1.2	2247	5.6	9.7	0.7	0.8	0.5	0.4
Gender	170	4.8	19	1	1355	3.4	16.5	0.7	1.2	0.5	0.6
Ontology	136	3.9	17	0.8	903	2.3	25	0.4	0.7	0	0
Cancer	131	5.3	16	1	951	2.4	10.7	0	0	0.5	0.8
Women's-health	124	5.8	15	1.1	1096	2.7	10.5	0.4	0.8	0.5	0.8
<b>Patient-portals</b>	<b>118</b>	<b>7.5</b>	<b>21</b>	<b>1.6</b>	<b>1506</b>	<b>3.8</b>	<b>11</b>	<b>1.1</b>	<b>2.5</b>	<b>1.5</b>	<b>2.5</b>
Innovation	102	5.5	17	1.4	1192	3	16.7	1.5	3.9	1	2
Lived-experience	96	6.5	16	1.1	862	2.2	5.2	0	0	0	0
Technology-adoption	93	6.5	20	1.5	1136	2.8	11.8	0.4	1.1	0.5	1.1
Virtual-communities	81	6.7	17	1.2	959	2.4	11.1	0.7	2.5	0.5	1.2
Digital-Libraries	75	4.9	12	1	575	1.4	8	0.4	1.3	0.5	1.3
Empowerment	72	5.6	14	1.2	618	1.5	9.7	0	0	0.5	1.4
Medical-informatics	53	7.1	13	1.5	606	1.5	7.5	0.4	1.9	0.5	1.9

The strongest themes in terms of impact were *electronic health records* and *patient portals*, with MNCS of 1.56 and 1.63, respectively, as presented in Table 7. The *patient portals* theme focused on mobile health and electronic health. Otherwise, the themes with lower impact were *academic libraries* and *ontology* (e.g., information retrieval, data sharing, and interoperability topics). The HCP were mainly found in themes related to health and technology, such as *electronic health records*, *social media*, *patient portals*, and *innovation* (i.e., patents and information and communication technology). Uncited papers were also reduced in the same type of themes as HCP, which increased their impact. However, the contribution of national collaboration to the percentage of HCP over the whole dataset was very low, despite national papers representing 23.39% of the whole.

In the national collaboration structure, neither motor nor basic themes were found among the high-impact ones (e.g., *innovation* and *social network analysis*). Similar to local collaboration, high-impact themes in national collaboration were novel themes that were not a part of the LIS category's core.

### International collaboration

Finally, as shown in Fig. 5 and presented in Table 8, the themes at the international collaboration level have some similarities but also many differences from the other collaboration types: *academic libraries*, strongly related in this case to data mining, content analysis, and text mining; *developing countries*, focusing on topics such as information and communication technology and small and medium enterprises; *social media* and *bibliometrics*, focusing on the same concepts as for the other collaboration types; *knowledge management*, focusing on knowledge sharing, innovation, and knowledge transfer; and themes focused on artificial intelligence (AI), such as *sentiment analysis*, *big data*, *natural language processing*, and *machine learning*.

Also, the international collaboration type, as shown in Fig. 5 and presented in Table 8, was focused on some themes that also appeared in the other collaboration types, e.g., *academic libraries*, *social media*, and *bibliometrics*. Moreover, some themes that only appeared in this type of collaboration are *developing countries*, which focused on information and communication technology and small and medium enterprises; *sentiment analysis*, which is related to artificial intelligence; and the themes *natural language processing* and *machine learning*. The *bibliometrics* theme, which focuses on similar topics in all the considered types of collaboration (e.g., citation analysis and altmetrics, among other topics), is an example of a theme that has similarities to other themes. In addition, *social media* is shared between all the collaboration types, aggregating topics such as social networks and political communication.

However, looking at the differences, *academic libraries*, also related to artificial intelligence concepts such as data mining and text mining, had twice the normalized impact in this type of collaboration. Moreover, new themes centered on the development of artificial intelligence techniques, which did not appear in the other types, are: *sentiment analysis*, applied to concepts such as online reviews and electronic commerce; *machine learning*, related to ethnography and topic modeling; and, *big data*, which aggregates the internet of things and cloud computing, including resource-planning topics.

The theme with the strongest impact was *literature review*, which contained papers carrying out systematic literature reviews and meta-analyses on information systems and enterprise architecture, obtaining a MNCS of 3.13, as presented in Table 8. In addition, the artificial intelligence themes obtained a high impact; indeed *sentiment*

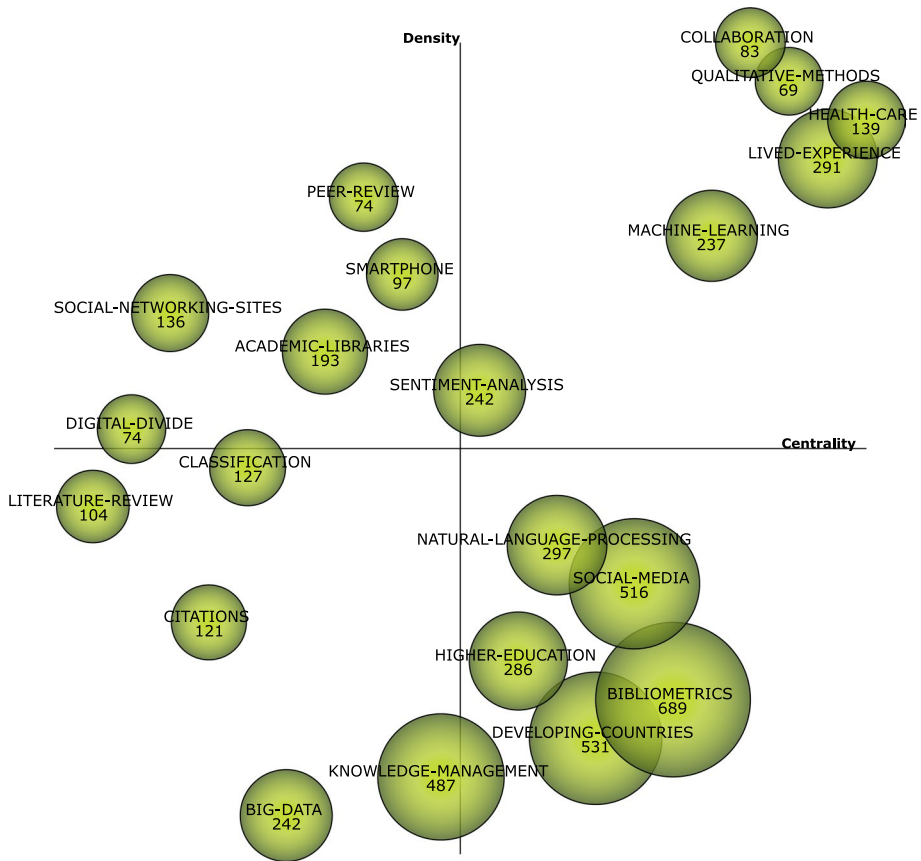


Fig. 5 International collaboration strategic diagram

analysis, big data, machine learning, and natural language processing had twice the expected average impact (MNCS > 2). In contrast, more traditional themes, such as higher education and health care, obtained a lower rate of 0.97 and 1.27 MNCS, respectively, and not all the themes had a remarkable number of HCP. Nonetheless, the number of uncited papers in the themes was very low compared with the rest of the collaboration types. HCP were mostly found in technology-related themes, such as social media, knowledge management, or sentiment analysis, and it is important to mention that themes of this type of collaboration contribute, in the same way as in local collaboration, to the global number of HCP, despite the former only accounting for 26.63% of the corpus.

The high-impact and technology-related themes were structured as motor and basic themes in the structure of international collaboration in LIS, in contrast to domestic collaboration, so these themes made up the core of international collaboration. Moreover, motor and basic themes had higher centrality (i.e., are strongly linked to the rest of the themes), which explains the rise in scientific impact of the rest of the themes.

**Table 8** International collaboration: performance and impact measures of themes. Bold themes are the ones with the highest MNCS

Theme	DOC ▼	CGM	HI	MNCS	CIT	CIT%	UP%	HCHC%W	HCHC%T	HC1%%W	HC1%%T
Bibliometrics	689	6.4	39	1.4	7830	12.9	10.7	6.2	2.5	4.4	1.3
Developing-countries	531	7	37	1.8	6565	10.8	9.2	5.5	2.8	7.8	3
Social-media	516	8.9	44	2	8567	14.1	8.9	8.8	4.7	9.3	3.7
Knowledge-management	487	7.6	37	1.9	6461	10.6	9.4	6.2	3.5	6.4	2.7
Natural-language-processing	297	8.1	35	2	4674	7.7	11.8	5.1	4.7	6.4	4.4
Lived-experience	291	6	25	1.5	3411	5.6	14.4	0.4	0.3	0.5	0.3
Higher-education	286	4.9	23	1	2212	3.6	14.3	1.5	1.4	1.5	1
<b>Sentiment-analysis</b>	<b>242</b>	<b>10.8</b>	<b>37</b>	<b>2.5</b>	<b>4970</b>	<b>8.2</b>	<b>5.8</b>	<b>4</b>	<b>4.5</b>	<b>4.9</b>	<b>4.1</b>
Big-data	242	8.5	33	2.3	3934	6.5	9.9	4.8	5.4	5.4	4.5
Machine-learning	237	5.6	21	1.6	2906	4.8	11.8	1.8	2.1	2	1.7
Academic-Libraries	193	5.8	23	1.2	1723	2.8	11.4	0.4	0.5	0.5	0.5
Health-care	139	6.7	18	1.3	1354	2.2	12.2	0.4	0.7	0.5	0.7
Social-networking-sites	136	9.7	27	2.1	2,359	3.9	5.1	2.2	4.4	1.5	2.2
Classification	127	5.7	20	1.3	1249	2.1	14.2	1.1	2.4	1	1.6
Citations	121	6.7	18	1.2	1124	1.8	9.9	0	0	0	0
<b>Literature-review</b>	<b>104</b>	<b>10.3</b>	<b>25</b>	<b>3.1</b>	<b>2351</b>	<b>3.9</b>	<b>7.7</b>	<b>4</b>	<b>10.6</b>	<b>5.4</b>	<b>10.6</b>
Smartphone	97	6.6	19	1.6	1138	1.9	17.5	0.4	1	0.5	1
Collaboration	83	7.4	17	2	1538	2.5	4.8	0.4	1.2	0.5	1.2
Digital-divide	74	6.3	14	1.1	668	1.1	5.4	0.4	1.4	0.5	1.4
Peer-review	74	7	16	1.6	995	1.6	6.8	0.4	1.4	0.5	1.4
Qualitative-methods	69	5.2	12	1.1	424	0.7	13	0	0	0	0

### Themes in top countries (RQ6)

The results of the analysis of themes in top countries for each type of collaboration, as well as for the whole dataset, are described in the next subsections to address RQ6.

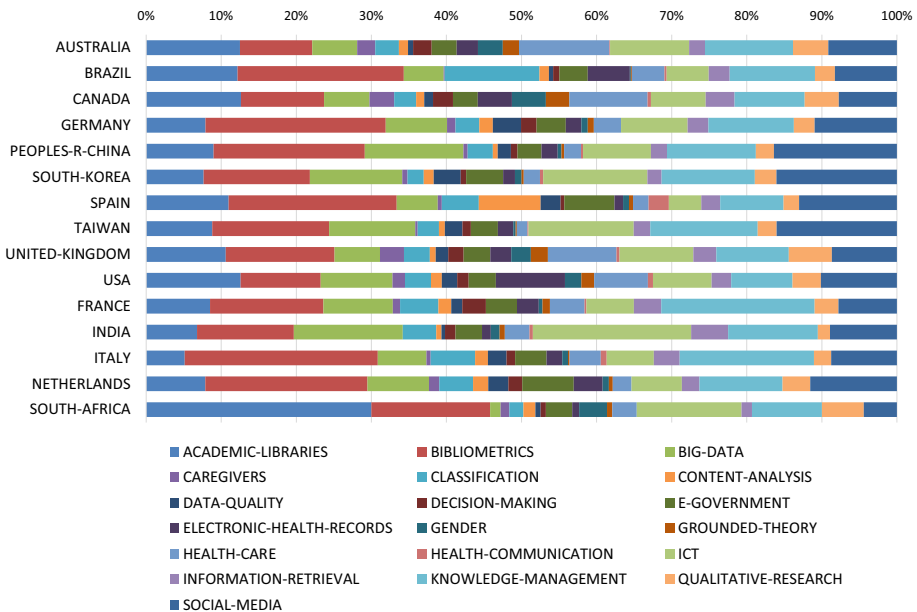
#### Top “whole” countries

Figure 6 shows the themes of the top 15 countries in the world by scientific output. The themes of *academic libraries*, *bibliometrics*, *knowledge management*, and *social media* were well represented in all countries. Other themes, such as *content analysis*, *gender*, *grounded theory*, and *health communication*, were weakly present in almost all countries.

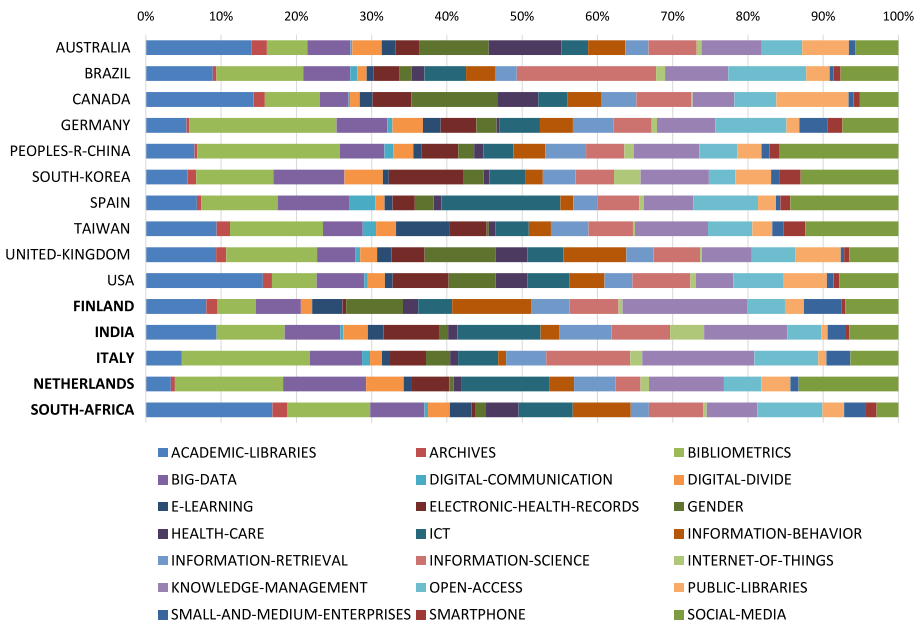
Some themes were relevant only to specific countries, such as in the case of *content analysis* in Spain, or *classification* in Brazil. Additionally, some countries were more concerned with high-impact themes than others; for example, India, South Korea, Taiwan, and China dedicated around half of their scientific output to *big data*, *e-government*, *information retrieval*, *knowledge management*, *social media*, and *information and communication technology*. South Africa also dedicated almost 50% to *academic libraries* and *bibliometrics*.

#### Top local collaboration countries

Most of the themes strongly appeared in all the countries in terms of local collaboration (Fig. 7), whereas some were marginal. In fact, themes such as *academic libraries*, *bibliometrics*, *social media*, and *knowledge management* were powerfully present in all



**Fig. 6** Themes in top countries (whole). *ICT* is an abbreviation for information-and-communication-technology. (Color figure online)



**Fig. 7** Themes of top countries (local collaboration). The top five countries of local collaboration are highlighted in bold. *ICT* is an abbreviation for information-and-communication-technology. (Color figure online)

the countries. In contrast, marginal themes such as *archives*, *digital communication*, and *smartphones* had a low representation in all the countries.

There were also particular countries in which some themes were not present, such as in the case of the Netherlands, in which neither the *digital communication* nor *smartphone* themes were present, or Australia, which also did not have studies in the *smartphone* theme. Regarding specific countries with a special interest in specific themes, the case of Spain is noteworthy, with a focus on the *digital communication* theme, in which Spanish authors published 50% of the total papers.

Concerning countries focused on high-impact themes, Table 6 shows that the two themes that achieved the highest impact were *internet of things* and *digital divide*. In Fig. 7, it is seen that the former acquired a high interest in relative terms from India and South Korea, while the latter was well represented by all the countries, although South Korea, the Netherlands, and Germany should be highlighted because they dedicated a large amount of attention to this theme.

On the other hand, taking into account the social structure of each country in the local collaboration, in Table 9 it is seen that the average degree was lower than 3, except for in China, where the value was fairly low, as will be introduced later. There were also some important values to highlight, such as the assortativity based on degree being very close to zero in Germany, China, South Korea, and Taiwan. This means that authors of these countries collaborated with others regardless of their importance in the network. In contrast, in other countries such as the United Kingdom or France, authors with a large number of connections collaborated, in most cases, with similar ones, while the



**Table 9** Coauthor networks: measures of top countries (local)

	#nodes	#edges	Avg. degree	Diameter	Degree assortativity	#Components	#Nodes P. Component	%Nodes P. Component	Avg. path length	Density
Australia	718	900	2.5	5	0.73	265	19	2.6	1.52	0.35
Brazil	714	873	2.4	5	0.76	236	22	3.1	1.58	0.34
Canada	734	873	2.4	5	0.81	302	12	1.6	1.22	0.32
Germany	456	479	2.1	4	0.66	186	16	3.5	1.40	0.46
Peoples-R-China	1308	2380	3.6	22	0.19	211	649	49.6	8.11	0.28
South-Korea	338	439	2.6	11	0.05	83	176	52.1	4.89	0.77
Spain	1317	1360	2.1	5	0.6	468	20	1.5	1.57	0.16
Taiwan	387	376	1.9	7	0.01	139	34	8.8	2.70	0.50
United-Kingdom	961	1291	2.7	5	0.85	353	22	2.3	1.69	0.28
USA	5468	9165	3.3	15	0.84	1915	317	5.8	5.86	0.06
Finland	256	309	2.4	5	0.64	89	24	9.4	1.73	0.95
India	376	390	2.1	4	0.44	127	16	4.3	1.50	0.55
Italy	271	299	2.2	4	0.83	100	8	3	1.18	0.82
Netherlands	300	352	2.3	6	0.38	99	23	7.7	1.78	0.78
South-Africa	192	162	1.7	5	0.22	77	17	8.8	1.82	0.88

same occurs with authors with fewer connections. This shows that some countries had different types of relationships when collaborating locally.

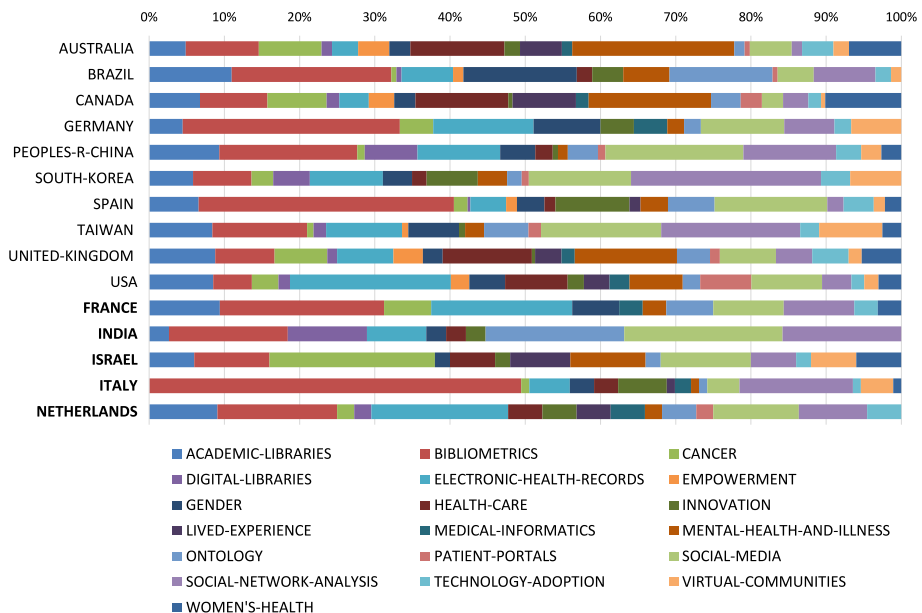
### Top national collaboration countries

Regarding the national collaboration type, Fig. 8 shows the relative number of papers in themes published by countries in the national collaboration type. Analyzing this figure, it is noted that some themes were strongly underrepresented, such as *empowerment*, *medical informatics*, *lived experience*, and *patient portals*. Nonetheless, several themes that were well structured among the countries, such as *academic libraries* (except for Italy), *bibliometrics*, *social media*, and *social network analysis*, were also identified.

In terms of how countries structured their research output, a minority were present in all the themes (i.e., the USA, Canada, Australia, and the United Kingdom), and the countries mentioned were also quite similar concerning their themes, whereas the rest varied (e.g., the USA is quite different from Spain in *bibliometrics*, *electronic health records* and *patient portals*).

Although all the countries had some similarities in this type of collaboration, it is worth noting that each country focused on its conceptual niche. The most remarkable cases were India and France, with the former mainly focusing on *social media*, *ontology*, *bibliometrics*, and *social network analysis*, and the latter focusing more on *bibliometrics*, *electronic health records*, *social network analysis*, and *social media*. It is also worth mentioning the case of Italy, which focused around 50% of national collaboration on *bibliometrics*.

The presence of some high-impact themes in most countries, such as *electronic health records* and *social media*, is remarkable. Nonetheless, not all countries were equally



**Fig. 8** Themes of top countries (national collaboration). The top five countries of national collaboration are highlighted in bold. (Color figure online)

interested in the previous themes (e.g., Australia or Canada were loosely interested, in comparison with the USA or France). There were also other high-impact themes, such as *patient portals* and *medical informatics*, which received minimal attention from most countries.

With respect to the coauthor network of each analyzed country, the average degree in national collaboration was generally higher than in local collaboration (Table 10). The assortativity also tended to increase, which reveals how the relationships between authors differ between types of collaboration. These measures may explain why countries focus much more on high-impact themes compared with local collaboration, owing to, for example, sharing of resources between organizations in the same country.

### Top international collaboration countries

Finally, Fig. 9 shows the conceptual structure of top countries in the international collaboration type. Themes of this type of collaboration were prevalent in almost every country. The themes that did not have universal presence were *collaboration*, *digital divide*, *health care*, and *peer review*, which were not present in Malaysia, Sweden, Taiwan, and South Korea, respectively.

The countries were very similar concerning the themes, although in the cases of Brazil and Spain both were very interested in *higher education*, while other countries had less interest. In addition, Brazil had a low focus on *sentiment analysis* in comparison with the rest.

High-impact themes (e.g., *literature review*, *sentiment analysis*, and *big data*) were present in every country. The relative number of papers published on these themes in each country was similar. However, there were some special cases where some themes did not appear in specific countries, such as *sentiment analysis* in Brazil or *big data* in the Netherlands.

Finally, the measures of the social structure of each country in the international collaboration (Table 11) are presented. In these networks, the average degree of the authors in the network (i.e., the average number of links they have with other authors) differed strongly from the rest of the collaboration types, with the exception of the largest countries (i.e., the USA or China). The assortativity was also much closer to 1, which reveals strong links between similar-degree authors, which may affect the visibility of the papers and increase their probability of being cited. Nonetheless, China differed from other countries in that there was no pattern of collaboration, with authors both related to similar and non-similar ones (based on the degree). Strong links between different countries tended to homogenize the thematic landscape owing to the sharing of resources and knowledge, enabling these more modern and citable topics to be addressed.

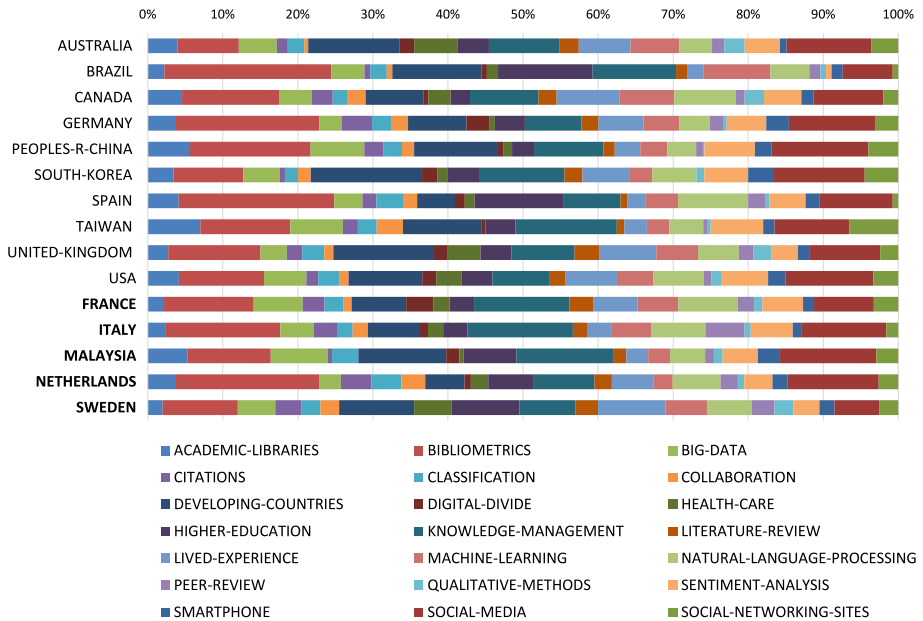
### Themes in country communities (RQ7)

As stated in the methodology, the Leiden algorithm was applied over the co-country network to identify communities of countries. The relative contribution of groups of collaborating countries to international themes is shown in Fig. 10.

Before describing the results of Fig. 10, it is vital to understand the nature of the eight groups of countries obtained. The ten most relevant countries making up the groups,

**Table 10** Coauthor networks: measures of top countries (national)

	#nodes	#edges	Avg. degree	Diameter	Degree assortativity	#Components	#Nodes P. Component	%Nodes P. Component	Avg. path length	Density (%)
Australia	438	817	3.7	4	0.66	96	16	3.65	1.4	0.85
Brazil	457	701	3.1	6	0.47	111	21	4.6	1.8	0.67
Canada	587	1522	5.2	5	0.75	113	40	6.8	1.63	0.88
Germany	210	343	3.3	4	0.69	50	16	7.6	1.43	1.56
Peoples-R-China	1229	3014	4.9	12	0.09	76	963	78.3	5.56	0.4
South-Korea	290	560	3.9	8	0.09	27	208	71.7	3.88	1.34
Spain	663	1011	3	12	0.45	154	89	13.4	4.17	0.46
Taiwan	391	644	3.3	11	0.12	57	75	19.2	3.91	0.84
United-Kingdom	693	1404	4	5	0.91	158	18	2.6	1.37	0.59
USA	6744	23,056	6.8	19	0.46	761	2999	44.5	7.16	0.1
France	164	334	4.1	2	0.96	41	10	6.1	1.07	2.5
India	147	229	3.1	4	0.64	36	14	9.5	1.61	2.13
Israel	120	222	3.7	4	0.68	24	16	13.3	1.60	3.11
Italy	228	521	4.6	4	0.91	53	32	14	1.65	2.01
Netherlands	184	282	3.1	2	0.48	46	8	4.3	1.17	1.67



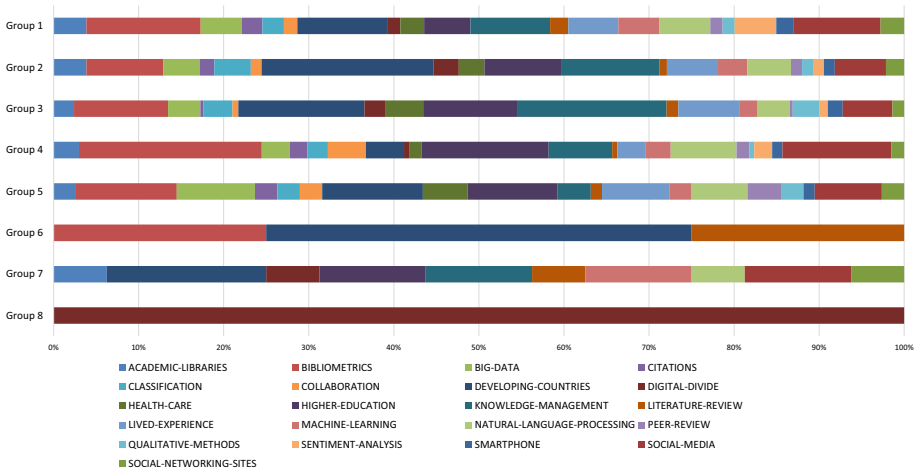
**Fig. 9** Themes of top countries (international collaboration). The top five countries of international collaboration are highlighted in bold. (Color figure online)

in terms of production, are presented in Table 12. In this table, countries in italics only appeared in the international collaboration; That is, in the LIS category during 2015–2019, they only had papers in collaboration with other countries but never alone. The eight groups obtained can be described as follows:

- *Group 1* Most European countries, as well as a significant portion of Asian and North American countries, were included in this group. It is also worth noting that 23 of the countries in this category were not involved in any local or national collaboration (e.g., Luxembourg, Liechtenstein, and Mozambique, among others). This group of countries was the most productive, as it is the largest and contains the top producing countries, such as the USA and China.
- *Group 2* This group consists primarily of Middle Eastern countries (i.e., Cyprus, Jordan, Kuwait, Oman, and Turkey), Eastern European countries (i.e., Latvia, Lithuania, Macedonia, Moldova, and Romania), and some African countries (i.e., Botswana, Morocco, Libya, Oman, and Ethiopia), among others. Turkey, Bangladesh, and Jordan were the most productive countries in this group. Furthermore, only five nations in this group (i.e., Bhutan, Palestine, Libya, Moldova, and Mongolia) had no presence in the other collaboration categories.
- *Group 3* Several African countries belong to this group, being geographically very close. This group also included five countries that only published papers by taking advantage of international collaboration (i.e., Cameroon, Cote d’Ivoire, Liberia, Senegal, and Swaziland).
- *Group 4* This group included only Latin American countries. This aggregation of countries has six countries that were closely linked to international collaboration

**Table 11** Coauthor networks: measures of top countries (international)

	#nodes	#edges	Avg. degree	Diameter	Degree assortativity	#Components	#Nodes P. Component	%Nodes P. Component	Avg. path length	Density (%)
Australia	1747	8751	10	19	0.99	213	662	37.89	6.2	0.57
Brazil	438	892	4.1	8	0.77	85	44	10.05	2.31	0.93
Canada	1814	10,787	11.9	15	0.99	241	560	30.87	6.33	0.66
Germany	1402	8770	12.5	11	0.99	211	140	9.99	3.01	0.89
Peoples-R-China	2898	8190	5.7	16	0.06	132	2395	82.64	5.48	0.2
South-Korea	814	6570	16.1	14	0.97	61	606	74.45	5	1.99
Spain	1470	9255	12.6	10	0.99	235	252	17.14	4.23	0.86
Taiwan	490	1136	4.6	16	0.68	62	133	27.14	6.09	0.95
United-Kingdom	3056	13,958	9.1	19	0.99	404	1048	34.29	7.36	0.30
USA	6482	24,818	7.7	18	0.95	642	3612	55.72	6.29	0.12
France	922	3205	6.9	8	0.99	162	83	9.00	2.69	0.75
Italy	867	6683	15.4	7	1	141	99	11.42	1.35	1.78
Malaysia	590	5923	20.1	12	0.99	71	145	24.58	3	3.41
Netherlands	1199	7724	12.9	13	0.99	170	348	29.02	5.37	1.08
Sweden	624	6130	19.7	10	0.99	90	99	15.87	2.08	3.15



**Fig. 10** Thematic breakdown of groups of collaborating countries. (Color figure online)

**Table 12** Groups of collaborating countries

ID	Top 10 countries in group
1	USA, P.R. China, United Kingdom, Australia, Canada, Spain, Germany, the Netherlands, South Korea, France
2	Turkey, Bangladesh, Jordan, Kuwait, Tanzania, Oman, Romania, Ethiopia, Lithuania, Kazakhstan
3	South Africa, Nigeria, Ghana, Kenya, Zimbabwe, Uganda, Malawi, Cameroon, <i>Liberia</i> , Zambia
4	Mexico, Chile, Cuba, Colombia, Ecuador, Argentina, Peru, <i>Venezuela</i> , Uruguay, <i>Guatemala</i>
5	Ireland, <i>Sudan</i> , <i>Myanmar</i>
6	<i>Niger</i> , <i>Burkina Faso</i> , <i>Nicaragua</i>
7	<i>Sri Lanka</i> , <i>Trinidad and Tobago</i>
8	<i>Fiji</i> , <i>Marshall Islands</i>

Countries in italics are present only in the international collaboration type

since they did not appear in the other collaboration types (i.e., Bolivia, Costa Rica, Guatemala, Panama, Paraguay, and Venezuela).

- *Groups 5–8* These groups were the smallest ones. Groups 6 and 7 comprised countries appearing only in the international collaboration. Group 8 was formed by only two countries. Furthermore, they were very isolated since the low production of their members results in their being slightly bound to the main component of the co-country network.

Having identified the groups, the presence of international themes in them can be analyzed. First, almost all themes were poorly present in groups with limited scientific production (i.e., groups 6–8). In contrast, it is noted that nearly all the groups in groups 1–5 were represented in each theme. The only exceptions include *digital divide* and *sentiment analysis*, in which the fifth group was not present.

Focusing then on the conceptual analogies between the communities of countries, it is observed that groups 1, 2, and 3 were extremely similar (except for the *social media* and *higher education* themes). Moreover, the group of Latin American countries (group 4) is close to the previous groups but with a low focus on *knowledge management* and more interest in *bibliometrics*.

The presence of high-impact themes (i.e., *literature review*, *big data*, *machine learning*, *natural language processing*, *sentiment analysis*, and *social media*) is thus very relevant. In fact, the relative number of papers on the aforementioned themes is above 20% in groups 1, 2, 4, 5, 6, and 7. The group with the highest presence of high-impact themes was the first, corresponding to nearly one-third of its relative production. In contrast, the lowest attention was received from the group of African countries (group 3), at approximately 18%. Moreover, it is worth highlighting the interest of group 5 (Niger, Nicaragua, and Burkina Faso) in *big data*. A special interest in low-impact themes (*digital divide* and *higher education*) via international collaboration was seen for groups 3–5 and 8.

## Discussion

In this paper, the LIS category during the period 2015–2019 was examined to analyze whether the themes and structure of collaboration types are important factors that may explain the increase in IRC citations. In this regard, and concerning RQ1, IRC yields a higher impact in terms of citations, which has already been stated in several works examining the LIS category of Web of Science (Asubiaro, 2019; Sin, 2011), and other works in different fields of science (Chinchilla-Rodríguez et al., 2019; Gazni et al., 2012; Larivière et al., 2015; Narin & Whitlow, 1990; Persson, 2010; Sooryamoorthy, 2017; Sugimoto et al., 2017). HCP were mainly found in themes with high IRC, which is in line with the results obtained by Persson (2010), however, according to our results, HCP are not found in the same proportion as IRC among the themes covered by the category, so the thematic dimension must also be taken into account when designing strategies for IRC.

The topology of the conceptual networks is similar, so they do not seem to be an important factor contributing to the asymmetrical impact of the collaboration types (RQ2). In contrast, the social network topology differs (RQ3), showing two main aspects: authors in IRC networks have a large social neighborhood, and authors relate to authors with a similar number of relationships. Important authors tend to work with other important authors, as opposed to domestic collaboration, where this tendency is less pronounced. Although authors have a very good network of collaborators, self-citations have been rejected as a sole factor explaining the increasing impact of IRC (Van Raan, 1998). Nonetheless, in our study, the probability of being cited by non-collaborators and/or researchers is higher, as papers in these IRC networks may have greater visibility, which advances the conclusions made by Yan et al. (2010), who found that central authors correlate with citations.

Regarding the thematic landscape of the LIS category (RQ4), there are some thematic similarities according to collaboration types. The similarities noted are: *bibliometrics*, which is present in the three collaboration types and has a similar impact on all of them; *social media*, which is quite similar and highly related to social networks, but in international collaboration is a stronger theme in terms of MNCS; and *health care*, which is very similar in the three collaboration categories as well, focusing on topics such as education, culture, and patient safety. In fact, there are two health-related themes in the three collaboration types: *empowerment*, which is about physical activity and mental illness; and



*gender*, which is about how gender relates to health, race, and information technology. All of them have about the same impact in all the collaboration types. Furthermore, *academic libraries*, which is related to collaboration and information literacy, is also similar. Nevertheless, in the international sphere, computer science (e.g., data mining and text mining, among others) plays an important role in this theme. In fact, the MNCS of *academic libraries* increased significantly in the international collaboration type and reached twice the impact acquired in the local and national collaboration types.

In this sense, it is important to highlight that the roles played by artificial intelligence (AI) and social media topics in the articles of the LIS category were also introduced in previous works (Hsiao & Hua Chen, 2020; Ma & Lund, 2021), but we found that these specialized themes were placed in the international collaboration type. In fact, our results show that there are multiple thematic differences that can explain the asymmetrical impact of collaboration types in LIS. Indeed, while in the local and national collaboration, AI topics are more spread throughout the themes and received much less attention, in the international collaboration type, they are made up of four specific themes: *natural language processing*, *big data*, *sentiment analysis* and *machine learning* (which has the highest MNCS), with almost all of them reaching twice the average citation rate of this category. Furthermore, the *bibliometrics* theme in international collaboration is related to AI topics such as data mining and text mining, and *literature review*, which is in the international collaboration type, focuses on the revision of previous works on enterprise architecture and information systems and has the highest impact. There are also other themes with a significant impact: *social networking sites* and *collaboration*. These novel and technological themes mostly come from international collaboration, where multidisciplinary, a mix of knowledge, and the skills of teams play an important role (Larivière et al., 2015; Suresh, 2012), which may explain the increase of citations, as they are at the vanguard of LIS scientific production.

Hence, technological themes may achieve greater impact than traditional themes. Therefore, a local paper in a high-impact theme may achieve a higher impact than an international paper in a low-impact theme. For instance, a local paper on *big data* may achieve more citations, on average, than a paper with international collaboration on *academic libraries* (see the measures of the whole LIS themes). In fact, a theme with remarkable international collaboration does not always mean higher impact (see *bibliometrics*, *grounded theory*, *caregivers*, or *gender* at the “whole” themes) and a low degree of international collaboration does not mean low impact either (see *information retrieval*). These may be influenced by other aspects, which are very interesting for the competitiveness of countries, such as funding-biased policies in some areas, as in the case of *information retrieval* from 1998 to 2008 as reported by Zhao (2010).

Moreover, regarding the structure of the LIS field according to the strategic diagram (RQ5), the high-impact themes were structured as non-motor themes when examining the whole category of LIS. From this perspective, the conceptual structure of LIS focuses on *bibliometrics* and *academic libraries*, which have been reported to be stable topics in the area (Hsiao & Hua Chen, 2020; Ma & Lund, 2021). However, themes with the highest impact are other themes that are structured as emerging (i.e., *big data*, *knowledge management*), or very specialized, such as *information retrieval*, according to their position in the strategic diagram. *Big data*- and *knowledge management*-related topics were reported as an emerging trend in LIS by Ma and Lund (2021). This implies that the themes achieving high scientific impact do not belong to the core of the LIS category.

Continuing with the results of the structure of the specific collaboration types, in international collaboration, the high-impact themes are structured as motor and basic themes

(*natural language processing, sentiment analysis, and collaboration*), and thus are salient in the rest of the international works (i.e., very central themes), and the internal development of these themes is high (i.e., very dense themes). Meanwhile, in domestic (i.e., local and national) collaboration, the high-impact themes are shown to be less central, which indicates that they are not used in the rest of the works at such levels of collaboration.

Therefore, there are many thematic and structural differences according to the strategic diagram that explain the asymmetrical impact of the collaboration types (RQ4 and RQ5). Domestic collaboration in LIS is focused on stable themes such as *bibliometrics* and *academic libraries*, and minor works are dedicated to new themes. However, for international collaboration, motor themes are new themes that lie at the forefront of the LIS category. It is important to highlight that these new themes have emerged in a relatively short period of time (Ma & Lund, 2021). This result is relevant for future studies on policy implications, as a key assumption might be that a variation in the collaboration patterns and/or in the themes may change the impact of an institution or a nation. Hence, future studies should examine the importance of being at the vanguard of scientific fields, taking into account the degree of international collaboration needed to develop those particular themes. Nations and institutions should also update their internal studies on the thematic and collaboration landscape from time to time to verify whether they are at the vanguard of the field, and how much dependence on collaboration they need Chinchilla-Rodríguez et al. (2019) to advance in these topics to make decisions aided by a good assessment of the situation.

Differences in the themes of top countries (RQ6) were found mainly concerning some countries, such as South Africa, which had a large amount of its research output dedicated to two traditional themes, while others were focused on high-impact and new themes, such as China, Taiwan, or India. This shows that there is room for competitiveness in the LIS area, as there is a variation in the knowledge base of countries; however, they must design strategies to ensure a path and clear goals for the specific themes. In contrast, and regarding the themes of specific collaboration types, the difference in themes between countries in domestic collaboration is much higher than in international collaboration. For international collaboration, the thematic landscape of the countries is mostly homogeneous, which might reveal a tendency toward a common international agenda. This is a problem, as noted by Adams (2012), because these strategies can end up being driven by a bland establishment consensus, which can ultimately lead to some researchers in underdeveloped nations abandoning project leadership, thus reducing thematic variety (Chinchilla-Rodríguez et al., 2019). Countries and organizations should examine their knowledge and expertise, and identify those that differentiate them from the rest where they can lead research to be in the forefront of the fields of study. This will lead to better understanding of their competitive assets, while contributing to the wealth of scientific development and thematic diversity.

Eight communities of countries were identified herein with a remarkable thematic difference between them (RQ7). In previous literature, Leydesdorff and Wagner (2008) found a core group that includes the most international collaborations, which also exists in our study, involving all the well-established economies and top countries, plus 23 countries that appear only when collaborating internationally. This core group is the one that focuses much more on high-impact themes. Moreover, there are seven other groups in which the formation seems to be guided by cultural, linguistic, and/or political factors (Frame & Carpenter, 1979; van Raan, 1997). There is a community of African countries, which were also previously reported in literature (Adams et al., 2014) and which is dedicated to stable and less impactful themes. It is also important to highlight some practices; for example, the community of countries of Niger, Nicaragua, and Burkina Faso is very interested in

*big data*, being at the forefront of the field. As can be seen, the core group leads the vanguard of research in LIS, while the rest of the communities of developing countries differ in their strategy. There are some countries in these communities that only appear at the international collaboration level, which seems to be caused by the lack of resources in their research systems, as pointed out by Chinchilla-Rodríguez et al. (2019).

Therefore, the themes and how they are structured in the LIS field are two crucial factors affecting the increasing impact of IRC in the category. These themes, which are structured as motor and basic themes in the international collaboration research, lie at the forefront of the LIS field and must be considered by organizations and nations to understand the impact of IRC. A suggestion emerging from this paper is that future studies should look further into including the thematic and collaborative factors, studying their possible implications for scientific policy design.

Finally, the results of this paper are only for LIS, while future studies should confirm whether these results are consistent in other fields/areas of knowledge. Therefore, the results are not generalizable, and other fields have to be studied specifically, since in some fields, such as computer science, there are less emergent technologies related to other fields of science or no major differences between collaboration levels (Velez-Estevéz et al., 2020) that contribute to the scientific knowledge of the area itself. With regards to the limitations, the methods employed in this paper are difficult to apply to the social sciences and humanities, since citation-based indicators are less representative of the research impact of these disciplines (Waltman, 2016).

## Conclusions

An analysis of the asymmetry in the impact of research collaboration types according to themes is presented herein. To achieve this, the LIS category was examined between the years of 2015 and 2019 (both included). The key finding of this paper is that the increased impact of IRC seems to be explained by a special focus on the forefront of research in the LIS category, which has been found to be an important factor. Therefore, we shed new light on the potential factors influencing the benefits of IRC. The forefront of research is mostly present in IRC of LIS, with emergent technologies applied in the field of study (i.e., artificial intelligence or social media, among others), while domestic research focuses on stable themes (i.e., academic libraries or bibliometrics).

Our findings also point out that differences in themes are present at the top country level (i.e., those with more scientific output), mostly in domestic collaboration. In contrast, IRC homogenizes the research, which could be a risk because it may lead to vague priorities around peer consensus research without a clear goal for organizations and nations. Communities of countries also exhibit a large difference in the thematic aspect, indicating that there is room for competitiveness and a need for future studies concerning policies at the community level on the basis of these results, which might enhance the research systems of developing communities of countries, and also might help the core group to maintain their research system.

Moreover, organizations, countries, and communities of countries may update the analysis of their situation from time to time to check their position in the thematic landscape and at the domestic and international collaboration levels, identifying whether themes are

at the vanguard of the field, to precisely design and revise their strategies for both domestic improvement and IRC strategies.

While thematic and structural factors are important in IRC, it can also be influenced by other factors, such as publication location, funding, research leadership, or publishing in open access. Therefore, future research on this topic should consider these other variables to demonstrate what was stated in this study.

Finally, although the results of this paper emphasize the thematic implications in different collaboration types, some limitations should also be noted: Only one database (Web of Science) was considered, and only the Library and Information Science category and the period of 2015–2019 were covered. Regarding the scope of our research, future studies must incorporate additional databases, different research categories, and a longer period to check whether these results are generalizable. New studies could also enrich these types of analyses with altmetrics or leadership indicators to reveal the possible societal impact of research carried out on the different collaboration levels and themes, and to observe the dominance of countries and/or communities of countries in the IRC landscape.

**Acknowledgements** This paper extends our contribution presented at ISSI 2021 Velez-Estevez et al. (2021). It has also been supported by the Spanish State Research Agency through project PID2019-105381GA-I00/AEI/10.13039/501100011033 (iScience).

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

## Declarations

**Conflict of interest** The authors have no conflicts of interest to declare that are relevant to the content of this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adams, J. (2012). The rise of research networks. *Nature*, 490(7420), 335–336. <https://doi.org/10.1038/490335a>
- Adams, J. (2013). The fourth age of research. *Nature*. <https://doi.org/10.1038/497557a>
- Adams, J., Gurney, K., Hook, D., & Leydesdorff, L. (2014). International collaboration clusters in Africa. *Scientometrics*, 98, 547–556. <https://doi.org/10.1007/s11192-013-1060-2>
- Aksnes, D. W., Langfeldt, L., & Wouters, P. (2019). Citations, citation indicators, and research quality: An overview of basic concepts and theories. *SAGE Open*, 9(1), 2158244019829575. <https://doi.org/10.1177/2158244019829575>
- Asubiaro, T. (2019). How collaboration type, publication place, funding and author's role affect citations received by publications from Africa: A bibliometric study of LIS research from 1996 to 2015. *Scientometrics*, 120(3), 1261–1287. <https://doi.org/10.1007/s11192-019-03157-1>
- Batagelj, V., & Cerinšek, M. (2013). On bibliographic networks. *Scientometrics*. <https://doi.org/10.1007/s11192-012-0940-1>

- Bauer, J., Leydesdorff, L., & Bornmann, L. (2016). Highly cited papers in Library and Information Science (LIS): Authors, institutions, and network structures. *Journal of the Association for Information Science and Technology*, 67, 3095–3100. <https://doi.org/10.1002/asi.23568>
- Cabezas-Clavijo, A., Robinson-García, N., Escabias, M., & Jiménez-Contreras, E. (2013). Reviewers' ratings and bibliometric indicators: Hand in hand when assessing over research proposals? *PLoS ONE*, 8(6), 68258. <https://doi.org/10.1371/journal.pone.0068258>
- Callon, M., Courtial, J., & 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. <https://doi.org/10.1007/BF02019280>
- Chen, K., Zhang, Y., & Fu, X. (2019). International research collaboration: An emerging domain of innovation studies? *Research Policy*, 48(1), 149–168. <https://doi.org/10.1016/j.respol.2018.08.005>
- Chinchilla-Rodríguez, Z., Sugimoto, C. R., & Larivière, V. (2019). Follow the leader: On the relationship between leadership and scholarly impact in international collaborations. *PLoS ONE*, 14(6), 1–18. <https://doi.org/10.1371/journal.pone.0218309>
- Cobo, M., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011a). An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field. *Journal of Informetrics*, 5(1), 146–166. <https://doi.org/10.1016/j.joi.2010.10.002>
- Cobo, M., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011b). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology*, 62(7), 1382–1402. <https://doi.org/10.1002/asi.21525>
- Cobo, M., López-Herrera, A., Herrera-Viedma, E., & Herrera, F. (2012). SciMAT: A new science mapping analysis software tool. *Journal of the American Society for Information Science and Technology*, 63(8), 1609–1630. <https://doi.org/10.1002/asi.22688>
- European-Commission. (2021). *The global approach to research and innovation: Europe's strategy for international cooperation in a changing world. COM/2021/252 final*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2021:252:FIN>. Accessed 4 July 2022.
- Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., Vespignani, A., Waltman, L., Wang, D., & Barabási, A. L. (2018). Science of science. *Science*, 359(6379), 185. <https://doi.org/10.1126/science.aao0185>
- Frame, J. D., & Carpenter, M. P. (1979). International research collaboration. *Social Studies of Science*, 9, 481.
- Franzoni, C., Scellato, G., & Stephan, P. (2011). Changing incentives to publish. *Science*, 333(6043), 702–703. <https://doi.org/10.1126/science.1197286>
- Gabrielle Breugelmanns, J., Roberge, G., Tippett, C., Durning, M., Struck, D. B., & Makanga, M. M. (2018). Scientific impact increases when researchers publish in open access and international collaboration: A bibliometric analysis on poverty-related disease papers. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0203156>
- Galvez, C. (2018). Análisis de co-palabras aplicado a los artículos muy citados en Biblioteconomía y Ciencias de la Información (2007–2017). *Transinformação*, 30(3), 277–286. <https://doi.org/10.1590/2318-08892018000300001>
- Gazni, A., Sugimoto, C. R., & Didegah, F. (2012). Mapping world scientific collaboration: Authors, institutions, and countries. *Journal of the American Society for Information Science and Technology*. <https://doi.org/10.1002/asi.21688>
- Han, P., Shi, J., Li, X., Wang, D., Shen, S., & Su, X. (2014). International collaboration in LIS: Global trends and networks at the country and institution level. *Scientometrics*, 98(1), 53–72. <https://doi.org/10.1007/s11192-013-1146-x>
- Han, X. (2020). *Evolution of research topics in LIS between 1996 and 2019: An analysis based on latent Dirichlet allocation topic model* (Vol. 125). Springer. <https://doi.org/10.1007/s11192-020-03721-0>
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569–16572. <https://doi.org/10.1073/pnas.0507655102>
- Hsiao, T. M., & Hua Chen, K. (2020). The dynamics of research subfields for library and information science: An investigation based on word bibliographic coupling. *Scientometrics*, 125(1), 717–737. <https://doi.org/10.1007/s11192-020-03645-9>
- Kato, M., & Ando, A. (2017). National ties of international scientific collaboration and researcher mobility found in Nature and Science. *Scientometrics*, 110(2), 673–694. <https://doi.org/10.1007/s11192-016-2183-z>
- Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*. <https://doi.org/10.1002/asi.23266>

- Leydesdorff, L., Bornmann, L., & Wagner, C. S. (2019). The relative influences of government funding and international collaboration on citation impact. *Journal of the Association for Information Science and Technology*. <https://doi.org/10.1002/asi.24109>
- Leydesdorff, L., & Wagner, C. S. (2008). International collaboration in science and the formation of a core group. *Journal of Informetrics*, 2, 317–325. <https://doi.org/10.1016/j.joi.2008.07.003>
- Ma, J., & Lund, B. (2021). The evolution and shift of research topics and methods in library and information science. *Journal of the Association for Information Science and Technology*, 72(8), 1059–1074. <https://doi.org/10.1002/asi.24474>
- Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & Delgado López-Cózar, E. (2018). Google scholar, web of science, and scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, 12(4), 1160–1177. <https://doi.org/10.1016/j.joi.2018.09.002>
- Martínez, M. A., Herrera, M., López-Gijón, J., & Herrera-Viedma, E. (2014). H-Classics: Characterizing the concept of citation classics through H-index. *Scientometrics*, 98(3), 1971–1983. <https://doi.org/10.1007/s11192-013-1155-9>
- Moed, H. F. (2005). Citation analysis in research. *Evaluation*. <https://doi.org/10.1007/1-4020-3714-7>
- Mokhtarpour, R., & Khasseh, A. A. (2020). Twenty-six years of LIS research focus and hot spots, 1990–2016: A co-word analysis. *Journal of Information Science*, 47, 15.
- Moral-Muñoz, J. A., Herrera-Viedma, E., Santisteban-Espejo, A., & Cobo, M. J. (2020). Software tools for conducting bibliometric analysis in science: An up-to-date review. *El Profesional de la Información*, 29(1), 3. <https://doi.org/10.3145/epi.2020.ene.03>
- Narin, F., & Whitlow, E. S. (1990). *Measurement of scientific cooperation and coauthorship in CEC-related areas of science* (Vol. 1). Office for Official Publications of the EC.
- Olmeda-Gómez, C., Ovalle-Perandones, M. A., & Perianes-Rodríguez, A. (2017). Co-word analysis and thematic landscapes in Spanish information science literature, 1985–2014. *Scientometrics*, 113(1), 195–217. <https://doi.org/10.1007/s11192-017-2486-8>
- Persson, O. (2010). Are highly cited papers more international? *Scientometrics*. <https://doi.org/10.1007/s11192-009-0007-0>
- Robinson-García, N., Sugimoto, C. R., Murray, D., Yegros-Yegros, A., Larivière, V., & Costas, R. (2019). The many faces of mobility: Using bibliometric data to measure the movement of scientists. *Journal of Informetrics*, 13(1), 50–63. <https://doi.org/10.1016/j.joi.2018.11.002>
- Robinson-García, N., van Leeuwen, T. N., & Råfols, I. (2018). Using altmetrics for contextualised mapping of societal impact: From hits to networks. *Science and Public Policy*, 45(6), 815–826. <https://doi.org/10.1093/scipol/scy024>
- Rousseau, R. (2006). New developments related to the hirsch index. *Science Focus*, 1, 23–25.
- Sin, S. C. J. (2011). International coauthorship and citation impact: A bibliometric study of six LIS journals, 1980–2008. *Journal of the American Society for Information Science and Technology*, 62(9), 1770–1783. <https://doi.org/10.1002/asi.21572>
- Sooryamoorthy, R. (2017). Do types of collaboration change citation? A scientometric analysis of social science publications in South Africa. *Scientometrics*, 111, 379. <https://doi.org/10.1007/s11192-017-2265-6>
- Sugimoto, C. R., Robinson-García, N., Murray, D. S., Yegros-Yegros, A., Costas, R., & Larivière, V. (2017). Scientists have most impact when they're free to move. *Nature*, 550(7674), 29–31. <https://doi.org/10.1038/550029a>
- Suresh, S. (2012). Global challenges need global solutions. *Nature*, 490(7420), 337–338. <https://doi.org/10.1038/490337a>
- Tahamtan, I., Safipour Afshar, A., & Ahamdzadeh, K. (2016). Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics*, 107(3), 1195–1225. <https://doi.org/10.1007/s11192-016-1889-2>
- Thelwall, M. (2016). The precision of the arithmetic mean, geometric mean and percentiles for citation data: An experimental simulation modelling approach. *Journal of Informetrics*, 10(1), 110–123. <https://doi.org/10.1016/j.joi.2015.12.001>
- Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, 9(1), 5233. <https://doi.org/10.1038/s41598-019-41695-z>
- van Raan, A. F. J. (1997). Science as an international enterprise. *Science and Public Policy*, 24(5), 290–300. <https://doi.org/10.1093/spp/24.5.290>
- Van Raan, A. F. J. (1998). The influence of international collaboration on the impact of research results: Some simple mathematical considerations concerning the role of self-citations. *Scientometrics*, 42(3), 423–428. <https://doi.org/10.1007/BF02458380>
- Velez-Estevéz, A., García-Sánchez, P., Moral-Munoz, J., & Cobo, M. (2020). Thematical and impact differences between national and international collaboration on artificial intelligence research. In 2020

- IEEE conference on evolving and adaptive intelligent systems (EAIS)* (pp 1–8). <https://doi.org/10.1109/EAIS48028.2020.9122769>
- Velez-Estevez, A., Garcia-Sanchez, P., Moral-Munoz, J., & Cobo, M. (2021). *Thematic influence on academic impact according to types of collaboration: An analysis of the library and information science field from 2015 to 2019* (pp. 1193–1204). International Society for Scientometrics and Informetrics.
- Visser, M., van Eck, N. J., & Waltman, L. (2021). Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. *Quantitative Science Studies*, 2(1), 20–41. [https://doi.org/10.1162/qss\\_a\\_00112](https://doi.org/10.1162/qss_a_00112)
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365–391. <https://doi.org/10.1016/j.joi.2016.02.007>
- Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, 5(1), 37–47. <https://doi.org/10.1016/j.joi.2010.08.001>
- Wilsdon, J. (2015). *The metric tide: Independent review of the role of metrics in research assessment and management*. SAGE.
- Yan, E. (2015). Research dynamics, impact, and dissemination: A topic-level analysis: Research dynamics. *Impact, and Dissemination*. <https://doi.org/10.1002/asi.23324>
- Yan, E., Ding, Y., & Zhu, Q. (2010). Mapping Library and Information Science in China: A coauthorship network analysis. *Scientometrics*, 83(1), 115–131. <https://doi.org/10.1007/s11192-009-0027-9>
- Zhao, D. (2010). Characteristics and impact of grant-funded research: A case study of the library and information science field. *Scientometrics*, 84(2), 293–306. <https://doi.org/10.1007/s11192-010-0191-y>

## Authors and Affiliations

A. Velez-Estevez<sup>1</sup>  · P. García-Sánchez<sup>2</sup>  · J. A. Moral-Munoz<sup>3</sup>  · M. J. Cobo<sup>4</sup> 

P. García-Sánchez  
pablogarcia@ugr.es

J. A. Moral-Munoz  
joseantonio.moral@uca.es

M. J. Cobo  
mjcobero@decsai.ugr.es

<sup>1</sup> Department of Computer Science and Engineering, University of Cadiz, Cádiz 11519, Spain

<sup>2</sup> Department of Computer Architecture and Technology, University of Granada, Granada 18071, Spain

<sup>3</sup> Department of Nursing and Physiotherapy, University of Cadiz, Cádiz 11009, Spain

<sup>4</sup> Department of Computer Science and Artificial Intelligence, Andalusian Research Institute in Data Science and Computational Intelligence (DaSCI), University of Granada, Granada 18071, Spain