Emerging Trends in Retail Analytics: A Bibliometric Analysis of the Last Decade

Tendencias Emergentes en la Analítica del Retail: Un Análisis Bibliométrico de la Última Década

Juan David Velásquez-Henao

*Facultad de Minas, Universidad Nacional de Colombia, Medellín, Colombia. jdvelasq@unal.edu.co*

Received: June 16th, 2023. Received in revised form: November 10th, 2023. Accepted: November 10th, 2023.

**Abstract**

.

*Keywords*:.

**Resumen**

.

*Palabras clave*:.

# Introduction

.

.

The rest of this paper is organized as follows: Section 2 discusses the methodology used. Section 3 presents the results. Section 4 discusses the findings. Finally, Section 5 presents the conclusions.

# Literature Review

# Materials and Methods

This section outlines and examines the standard workflow commonly employed in literature analysis and tech-mining studies, which also serves as the methodological foundation of this paper. This process has been widely discussed in the research literature, including by Aria and Cuccurullo (2017), Donthu et al. (2021), and Page et al. (2021). The adopted methodology comprises four key stages:

1. Design of the study.

2. Collection and preparation of data.

3. Analysis of the data.

4. Interpretation of the findings.

## Design of the Study

The study's parameters are presented below, following the literature review criteria established by Kitchenham (2004). Scopus was used as the bibliographic database to retrieve information on engineering journals indexed in SciELO, aligning with the study's objective of analyzing SciELO-indexed engineering journals. The search strategy consisted of listing the ISSNs of the relevant journals from SciELO and using them to perform targeted searches within Scopus. The study parameters are:

* Database: Scopus.
* Years of analysis: From Jan 2015 to Dec 2024.
* Data retrieval: Feb 3, 2025.
* Search string: The search operator ISSN() was used to retrieve the information of each engineering journal indexed in SciELO from Scopus.
* Inclusion criteria: None.
* Exclusion criteria: Documents whose abstracts are not written in English.

The exclusion criterion was established based on two primary considerations. First, most abstracts in the dataset are written in English, making it the dominant language of scholarly communication in the selected journals. Second, the text processing tools employed in this study are configured to analyze English-language content, ensuring consistency and accuracy in the analysis. As a result, documents with abstracts written in other languages were excluded from the dataset.

**Fig. 1** presents the PRISMA flow chart. During the identification phase, the search string returned 29,499 documents. A time restriction was applied in the screening phase, excluding 12,849 papers published before 2015 or in 2025. The eligibility phase involved reviewing the titles and abstracts of the remaining 16,603 documents published between 2015 and 2024, excluding 1,307 papers. Consequently, the final dataset consists of 15,286 documents.

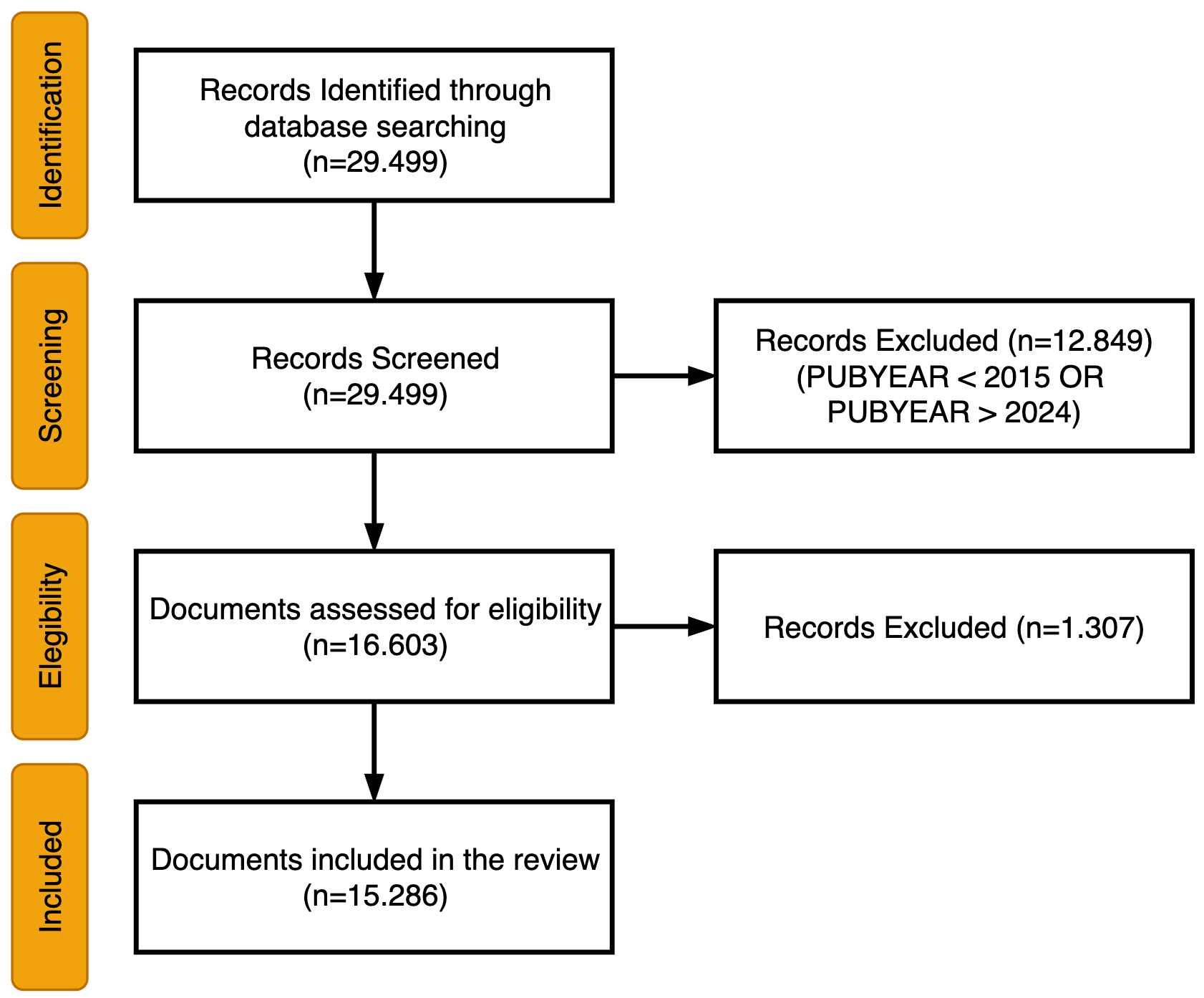


Figure 1. The PRISMA flow chart

Source: The authors.

## Collection and preparation of data

Following widely accepted practices in the literature, all bibliographic data were downloaded from Scopus in CSV format for analysis (Donthu et al., 2021; Page et al., 2021). The downloaded dataset includes document titles, abstracts, author keywords, index keywords, authors, affiliations, source titles, and bibliographic references.

Data processing combined automated procedures with manual refinements to ensure consistency and quality. The cleaning process applied to the raw author keywords included:

* Converting all text to uppercase.
* Translating British spelling to American variants.
* Removing multiple consecutive spaces within strings.
* Standardizing the formatting of hyphenated words.

This preprocessing step was essential to ensure accurate grouping and analysis of keywords in the subsequent stages. Author keywords were used directly for thematic exploration, reflecting the terminology and concepts prioritized by the authors themselves. Identifying dominant themes based on cleaned author keywords supports a multifaceted analysis of the engineering papers published in SciELO, as detailed in the following section.

## Analysis of the data

Various analyses were conducted to understand better the papers published in engineering journals indexed in SciELO. The following sections detail each analysis's specific procedures and findings.

## Interpretation of the findings

The interpretation phase focused on reviewing and synthesizing the results obtained from the various analyses. Patterns, trends, and thematic clusters identified during the data analysis phase were examined to derive meaningful conclusions aligned with the study's objectives. This step involved contextualizing the findings within the broader research landscape and ensuring consistency between the results and the study’s research questions.

# Results and Discussion

This section presents the fundamental bibliometric indicators of the analyzed dataset.

## General Metrics

The dataset includes scientific publications from January 2015 to December 2024, totaling 15,286 documents and reflecting an annual growth rate of 28.05%. The average document age is 5.77 years, with each work receiving approximately 4.11 citations overall, or 0.41 citations per year. Publications originate from 37 different sources, averaging 413.14 documents per source. The dataset comprises 14,822 journal articles, 35 conference papers, 21 editorials, two errata, one letter to the editor, and 402 review papers. A total of 40,773 authors (1,609 unique) contributed to the publications, with a strong tendency toward collaboration—averaging 3.61 authors and 3.92 co-authors per document. International collaborations account for 18.84% of the dataset. The contributing authors are affiliated with 10,687 organizations across 127 countries.

## Leading Scopus Subject Areas

This section analyzes the subject areas assigned by Scopus, which are attributed at the journal level rather than to individual articles. These classifications help identify the disciplinary focus of the research published in SciELO-indexed engineering journals. The dataset used in this study comprises 15,286 documents, which are distributed across 18 of Scopus's 27 subject areas. Notably, eight subject areas are associated with 1,000 or more documents. The most prominent areas are:

* Engineering: 7,257 documents (as anticipated).
* Materials Science: 2,664 documents.
* Computer Science: 1,884 papers.
* Chemical Engineering: 1,629 documents.
* General: 1,609 documents.
* Agricultural and Biological Sciences: 1,401 papers.
* Environmental Science: 1,263 documents.
* Business, Management, and Accounting: 1,017 papers.

**Fig. 2** displays a cross-correlation map between Scopus subject areas and SciELO journals. This visualization provides insight into the interconnections between journals based on their associated subject areas. The numbers alongside each subject area indicate the corresponding documents and citations. Links between nodes reflect the degree of similarity, while node size is proportional to the number of documents associated with each subject area.

The correlation map reveals a strong central cluster of subject areas closely related to Engineering and its subfields (including “Material Science,” “Computer Science,” and “Energy”). In addition, several other groups of nodes emerge that exhibit strong internal correlations but relatively weak connections to the engineering core. These include:

* “Social Sciences,” which appears as an isolated node.
* “Physics and Astronomy” and “Mathematics” form a distinct cluster.
* A group consisting of “Medicine,” “Immunology and Microbiology,” “Biochemistry, Genetics and Molecular Biology,” and “Environmental Science,” which are closely interrelated but show limited overlap with engineering disciplines.

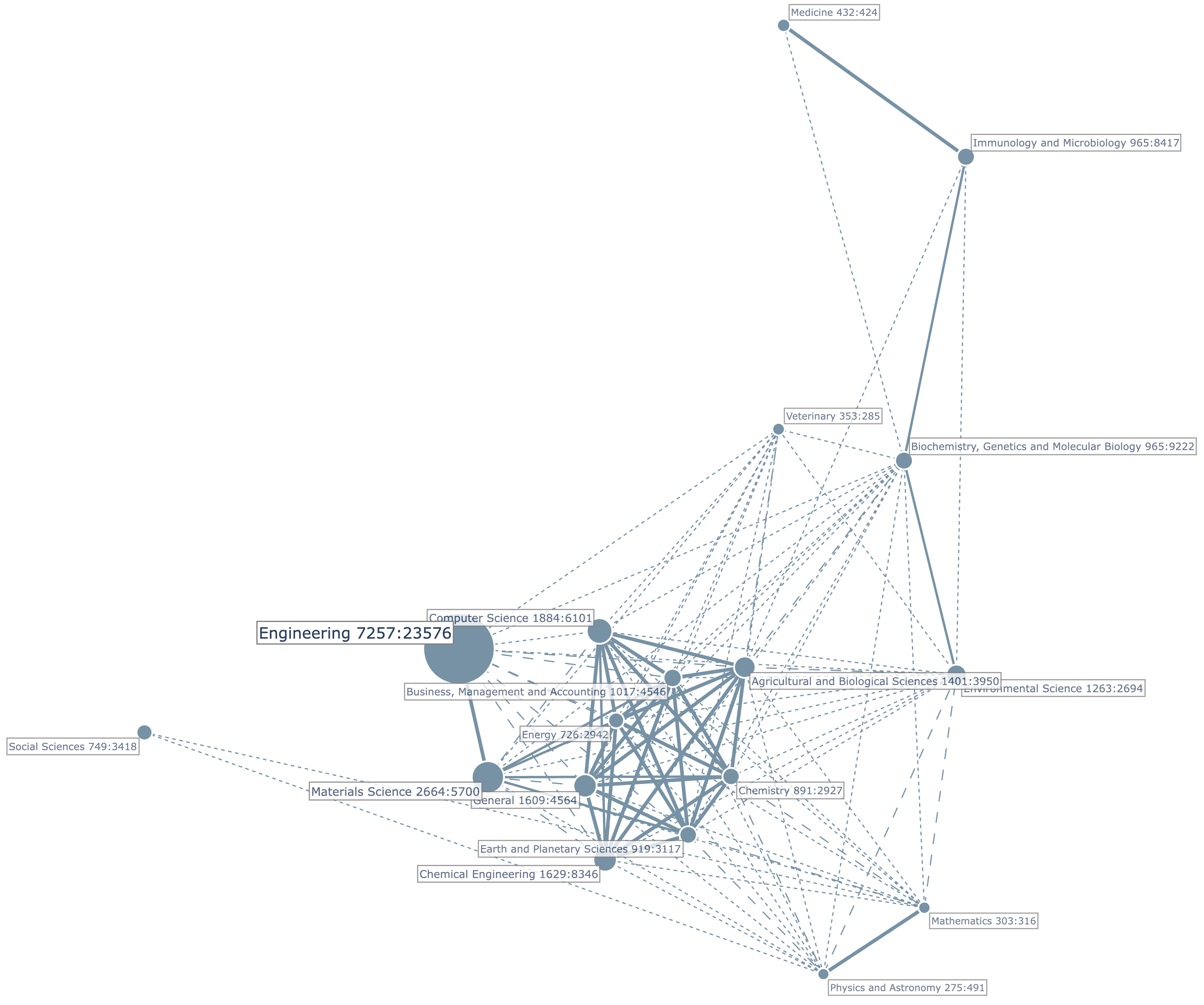


Figure 2. Correlation map of Scopus Subject Areas crossed with engineering SciELO journals.

Source: The authors.

## Basic Performance Metrics

### Authors

**Table 1** presents the 20 most frequent authors publishing in engineering journals indexed in SciELO between 2015 and 2024, including the number of papers, citations received within SciELO, total publications indexed in Scopus, and the ratio of SciELO to Scopus publications. This ratio provides insight into each author's publication preferences and engagement with the SciELO platform relative to the broader Scopus-indexed literature.

Among the authors, Ganga-Contreras F.A. leads with 39 publications in SciELO, accounting for 25.4% of their total output in Scopus, followed closely by Pedraja-Rejas, L.M. and Campos-Aranda D.F., both with 37 papers. Notably, Campos-Aranda D.F. shows a very high SciELO/Scopus ratio of 86.0%, suggesting a strong focus on SciELO as their primary publication venue. Similarly, Severino-González P. and Acevedo-Correa D. also exhibit high ratios (59.0% and 57.9%, respectively), indicating a substantial portion of their research is disseminated through SciELO journals.

In contrast, authors such as Gelbukh A. and Valencia-Arias A. have a much lower ratio (11.5% and 11.8%, respectively), suggesting a broader or more international publication strategy favoring Scopus journals outside SciELO. Despite this, they maintain a notable presence in SciELO and receive many citations.

Authors like Bustamante-Ubilla M.A. (52.4%) and Castrillón O.D. (60.6%) also strongly engage with SciELO relative to their total output. In contrast, others such as Preciado-Rangel, P. and Escobar D.A. have lower ratios but remain among the most frequent contributors.

Table 1.

Most Frequent Authors.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Author Name | Papers  SciELO | Citations  Scielo | Papers  Scopus | Ratio |
| Ganga-Contreras F.A. | 39 | 182 | 153 | 25.4 % |
| Pedraja-Rejas, L.M. | 37 | 240 | 108 | 34.3 % |
| Campos-Aranda D.F. | 37 | 64 | 43 | 86.0 % |
| Severino-González P. | 36 | 221 | 61 | 59.0 % |
| Gelbukh A. | 33 | 219 | 287 | 11.5 % |
| Acevedo-Correa D. | 33 | 164 | 57 | 57.9 % |
| López-Lezama J.M. | 30 | 177 | 131 | 22.9 % |
| Vidal-Silva C.L. | 29 | 152 | 88 | 33.0 % |
| Valencia-Arias A. | 28 | 178 | 237 | 11.8 % |
| Muñoz-Galeano, N. | 27 | 148 | 85 | 31.8 % |
| Rodríguez-Ponce E.R. | 26 | 185 | 87 | 29.9 % |
| Ortega-Toro R. | 26 | 157 | 96 | 27.1 % |
| Dávila-Morán, R.C. | 25 | 21 | 84 | 29.8 % |
| Preciado-Rangel, P. | 24 | 63 | 118 | 20.3 % |
| Escobar D.A. | 23 | 86 | 88 | 26.1 % |
| Paz-Pellat F. | 23 | 58 | 44 | 52.3 % |
| Bustamante-Ubilla M.A. | 22 | 117 | 42 | 52.4 % |
| Fontalvo-Herrera, T. | 21 | 167 | 46 | 45.7 % |
| Castrillón O.D. | 20 | 110 | 33 | 60.6 % |
| Zapata J.E. | 20 | 109 | 48 | 41.7 % |

Source: The authors.

### Organizations

**Table 2** lists the ten most frequent institutions contributing to engineering journals indexed in SciELO between 2015 and 2024. These institutions are ranked based on their total number of publications in the dataset. Two citation metrics are included: Global citations, which refer to the total number of citations recorded in Scopus, and Local citations, which count citations made between documents within the analyzed database.

Out of 10,687 unique organizations represented in the dataset (and 4,956 as first-author affiliations), these top 10 institutions account for a significant proportion of the overall scientific output. The Universidad Nacional de Colombia (COL) leads with 829 papers, followed by Instituto Politécnico Nacional (MEX) with 602, and Universidad de Antioquia (COL) with 353 publications. These three institutions alone contribute over 11% of all documents analyzed, highlighting their central role in SciELO’s engineering research landscape.

Regarding global impact, measured by Scopus citations, the Universidad Nacional de Colombia again stands out with 3,754 citations, the highest among all institutions, suggesting strong visibility and influence beyond the local database. Instituto Politécnico Nacional also shows a high global citation count (2,342), while other institutions, such as Universidad Nacional Autónoma de México and Universidad del Valle, maintain solid performance with over 900 citations each.

As expected in a diverse and broad dataset, local citations (a proxy for intra-database scholarly exchange) are generally lower. However, Instituto Politécnico Nacional shows a relatively high local citation count (39), suggesting a strong internal connection within the SciELO engineering network. Other institutions, such as Universidad Nacional Autónoma de México (13) and CONICET (ARG) (13), also reflect moderate levels of local scholarly interaction.

The top institutions are geographically concentrated in Colombia and Mexico, with seven out of ten based in these two countries, underscoring their prominent role in regional engineering research. Notably, institutions like Universidad de Tarapacá (CHL) and CONICET (ARG) also contribute significantly, reflecting the broader participation of Latin American countries in SciELO’s engineering publications.

Table 2.

Most Frequent Institutions.

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Papers | Global  Citations | Local  Citations |
| Univ Nac de Colombia (COL) | 829 | 3754 | 28 |
| Inst Politec Nac (MEX) | 602 | 2342 | 39 |
| Univ de Antioquia (COL) | 353 | 1406 | 6 |
| Univ Nac Autonoma de Mexico (MEX) | 254 | 1171 | 13 |
| Univ Autonoma Metropolitana (MEX) | 248 | 910 | 18 |
| Univ de Tarapaca (CHL) | 237 | 761 | 4 |
| Tecnol Nac de Mexico (MEX) | 234 | 695 | 7 |
| Univ del Valle (COL) | 232 | 983 | 9 |
| CONICET (ARG) | 192 | 528 | 13 |
| Univ Ind de Santander (COL) | 179 | 656 | 4 |

Source: The authors.

### Countries

**Table 3** highlights the top ten countries contributing to engineering publications in SciELO between 2015 and 2024. Colombia and Mexico lead with 3,524 and 3,452 papers, respectively, accounting for nearly 46% of the total 15,286 documents. While Colombia has the highest output, Mexico shows more vigorous intra-database citation activity, indicating greater local engagement. China ranks third in productivity (2,148 papers) and demonstrates high global visibility through citations but limited local interaction. Other Latin American countries (Chile, Brazil, Ecuador, Peru, Argentina, and Cuba) further emphasize SciELO’s strong regional identity, with Latin America and the Caribbean contributing 11,094 documents, 39,040 global citations, and 506 local citations. Spain also stands out for its citation impact despite fewer publications. Regionally, the Americas dominate with 11,235 documents and 40,220 global citations, followed by Asia (3,237 papers) and Europe (1,922). Subregional analysis confirms that Latin America and the Caribbean are the core of SciELO’s engineering output, while Eastern and Southern Asia show notable participation. Southern Asia, in particular, demonstrates high local citation rates, suggesting active engagement with the SciELO research network. Overall, SciELO serves as a key platform for engineering scholarship, particularly within Latin America, with growing international participation.

Table 3.

Most Frequent Countries.

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Papers | Global  Citations | Local  Citations |
| Colombia | 3,524 | 13,447 | 102 |
| Mexico | 3,452 | 12,408 | 154 |
| China | 2,148 | 6,326 | 59 |
| Chile | 1,265 | 6,100 | 85 |
| Spain | 1,226 | 5,341 | 42 |
| Brazil | 951 | 4,351 | 133 |
| Ecuador | 859 | 1,697 | 8 |
| Peru | 661 | 1,474 | 25 |
| Argentina | 543 | 1,162 | 15 |
| Cuba | 415 | 889 | 8 |

Source: The authors.

### Sources

Table 4 presents the ten most frequent journals in the dataset, drawn from 37 sources, with an average of 413 documents per journal. Boletín Técnico/Technical Bulletin leads with 1,820 papers but shows relatively low citation counts (2,810 global and 12 local), suggesting high productivity but limited impact. In contrast, “DYNA (Colombia)”, with 1,341 papers, has a much higher global citation count (5,321), indicating greater visibility. “Computación y Sistemas” and “Revista Mexicana de Ingeniería Química” also show strong performance, with over 1,000 and 900 papers, respectively, and significant local citations—particularly the latter, which leads in local citations (197), suggesting strong engagement within the SciELO community.

Electronic Journal of Biotechnology stands out with the highest global citation count (11,229) despite ranking lower in document volume (593 papers), reflecting a high citation impact per paper. Similarly, “Journal of Applied Research and Technology” shows strong citation metrics relative to its output. These results indicate that while some journals focus on volume, others achieve more significant influence through citation impact. Overall, the data reflect a diverse set of publication strategies across journals in SciELO’s engineering collection, with varying balances between productivity and scholarly impact.

Table 4.

Most Frequent Countries.

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Papers | Global  Citations | Local  Citations |
| Boletin Tecnico/Technical Bulletin | 1,820 | 2,810 | 12 |
| DYNA (Colombia) | 1,341 | 5,321 | 45 |
| Computacion y Sistemas | 1,037 | 3,310 | 100 |
| Revista Mexicana de Ingeniera Quimica | 903 | 5,404 | 197 |
| Informacion Tecnologica | 882 | 3,698 | 7 |
| Interciencia | 871 | 2,342 | 8 |
| Formacion Universitaria | 701 | 4,205 | 7 |
| Tecnologia y Ciencias del Agua | 625 | 1,552 | 2 |
| Electronic Journal of Biotechnology | 593 | 11,229 | 85 |
| Journal of Applied Research and Technology | 580 | 5,889 | 65 |

Source: The authors.

## Author Keywords Analysis

The analyzed dataset contains 42,141 unique text strings representing the author-assigned keywords. As described in Section 3.2, these raw keywords were subjected to a cleaning procedure to unify different text strings representing the same concept or idea. As a result of this process, 39,450 cleaned author keywords were obtained.

**Fig. 3** presents the six most frequent author keywords per year from 2015 to 2024, based on their occurrence and citation impact, and serves as the basis for analyzing temporal keyword dominance. Early years (2015–2016) featured specialized terms like “social networking,” “bi-level programming,” and “asphaltenes,” reflecting narrower research focuses. From 2017 onward, broader and high-impact terms such as “big data,” “data mining,” and “genetic algorithm” emerged, marking a shift toward digital technologies and computational methods. Keywords like “simulations,” “artificial neural network,” and “optimization” gained prominence with sustained relevance and high global citations. Education-related terms— “teaching,” “learning,” and “higher education”—also became more visible, peaking around 2018–2020. More recently, terms like “machine learning,” “deep learning,” and “natural language processing” reflect current AI-driven trends. Post-2020 entries such as “covid19”, “circular economy,” and “digital transformation” indicate responses to global disruptions. In 2023–2024, emerging technologies like the “large language model,” “blockchain,” and “cybersecurity” suggest a forward-looking research agenda. The diversity and evolution of keywords underscore the dynamic nature of engineering research indexed in SciELO, spanning traditional disciplines, applied technologies, and societal challenges.

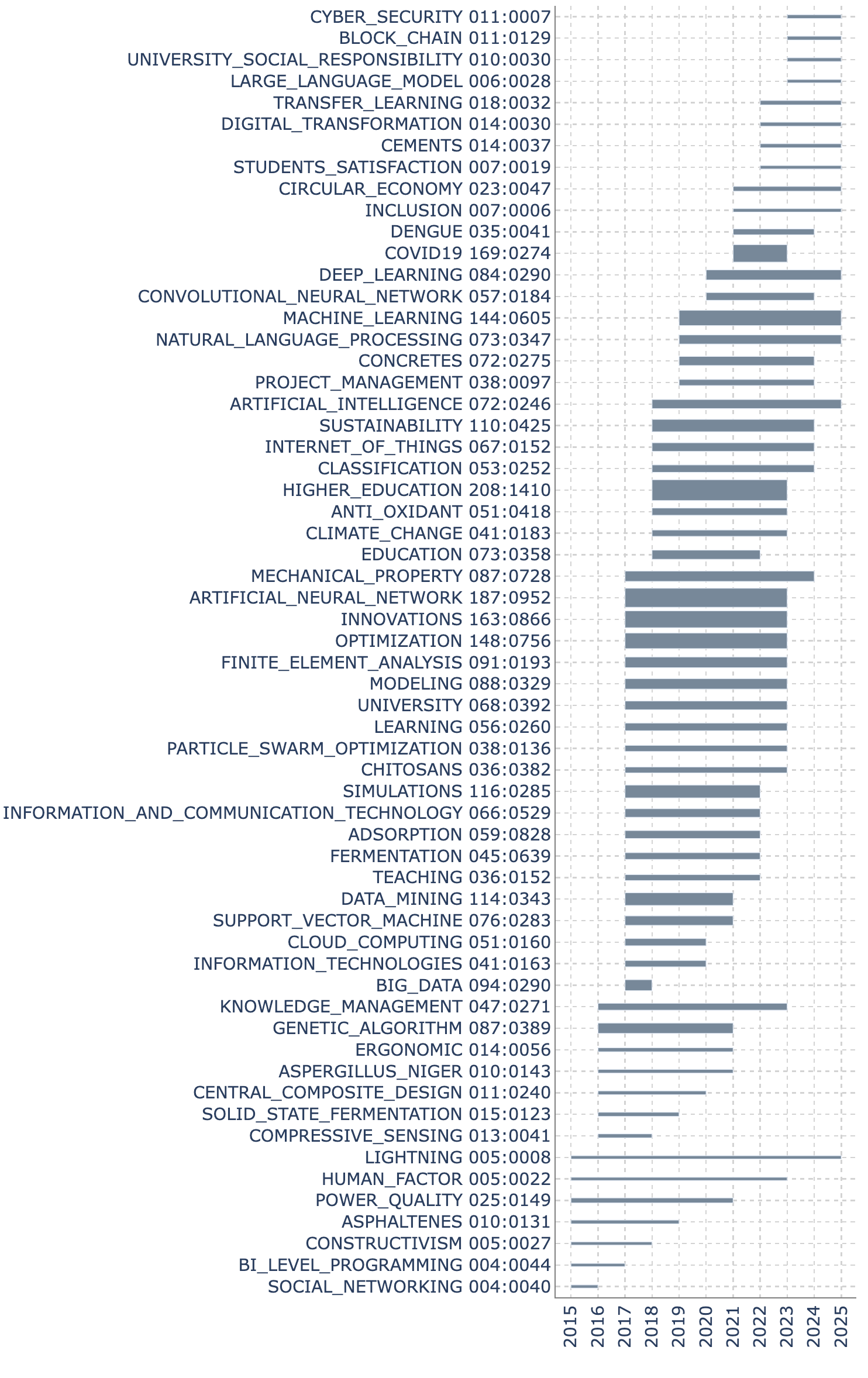


Figure 3. Dominant author keywords per year

Source: The authors.

## Correlation Analysis

**Fig. 4** presents a correlation map of SciELO engineering journals based on the cleaned author keywords used as the cross-variable. The size of each node is proportional to the number of documents published by the corresponding journal. At the same time, the width and intensity of the connecting lines reflect the strength of the cross-correlations between journals. The map reveals a well-defined core of closely interrelated journals actively published in engineering fields. This core includes prominent journals such as “DYNA (Colombia),” “Boletín Técnico,” and “Computación y Sistemas.” In contrast, journals like “Terra Latinoamericana” and “Papers in Physics” appear on the periphery of the correlation map, indicating that their thematic focus diverges from the central topics shared by the core journals. The map highlights a strong correlation between “Anales de Investigación en Arquitectura” and “Estoa,” suggesting a shared thematic scope or overlapping research interests within a specific subfield.

## Network Analysis

### Citation Network

**Fig. 5** presents the citation network of engineering journals indexed in SciELO, clustered using the Louvain algorithm. In bibliometric analysis, a citation network is a graphical representation where nodes correspond to documents (e.g., journals or articles) and edges represent citation links between them. These networks help identify thematic clusters, influential sources, and patterns of knowledge dissemination across disciplines. The Louvain clustering revealed 14 distinct clusters, each grouping journals with stronger internal citation ties. The luster, most densely populated by grouping 13 sources, includes highly cited journals such as “DYNA (Colombia)”, “Información Tecnológica”, and “Interciencia”, suggesting this cluster forms the core of SciELO’s engineering literature. These journals exhibit high publication volume and citation impact, indicating centrality in the network. The second major cluster is conformed by the journals BOLETIN DE malarologia y salud Ambiental, journal

The third major cluster groups the journals related to architecture and construction and is comprised by:

* Revista Ingeniería de la Construcción
* Ingeniería (Colombia)
* Revista Alconpat.

There are no more essential relationships to highlight among the journals.

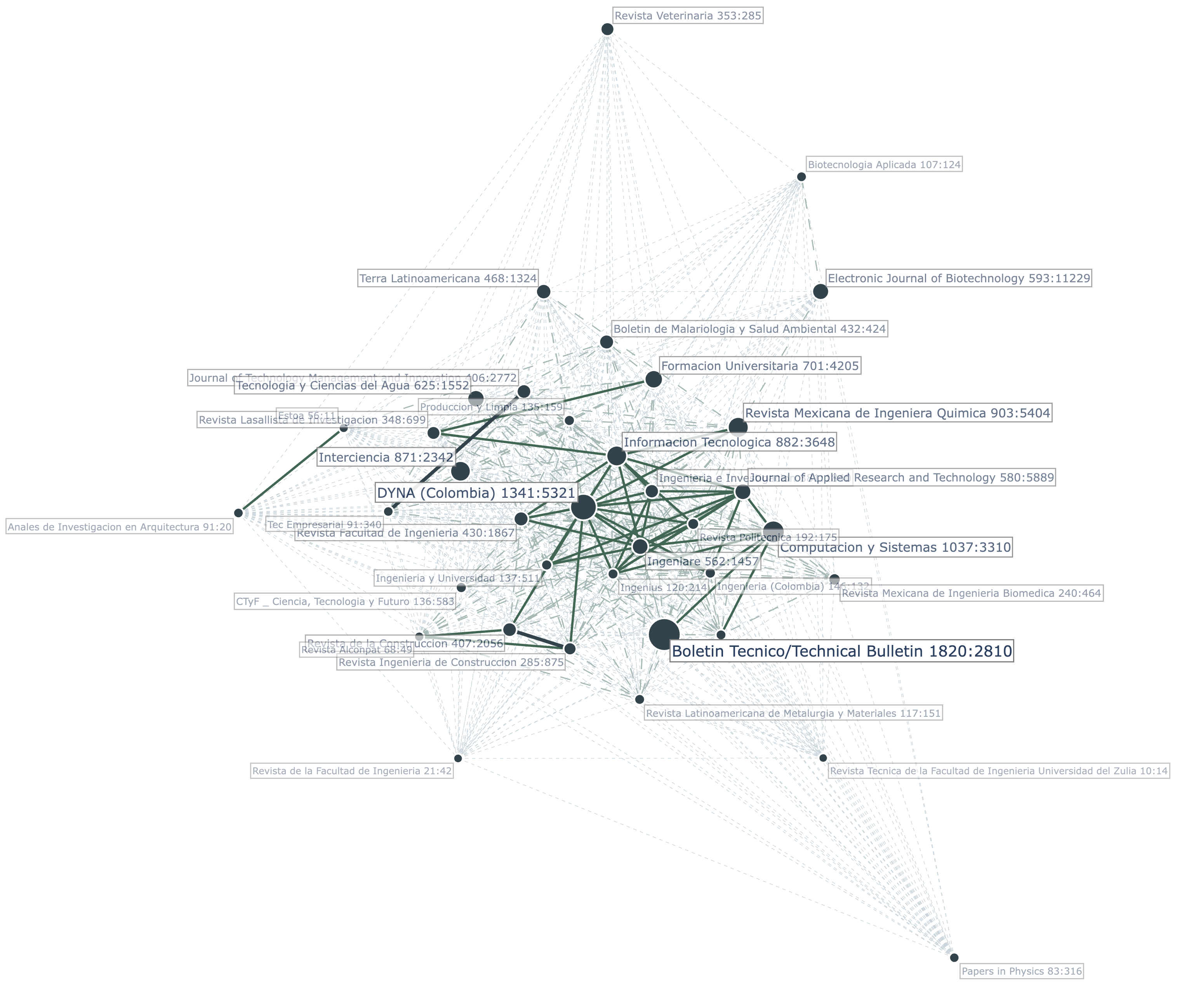


Figure 4. Correlation map of the engineering SciELO journals based on the author keywords.

Source: The authors.

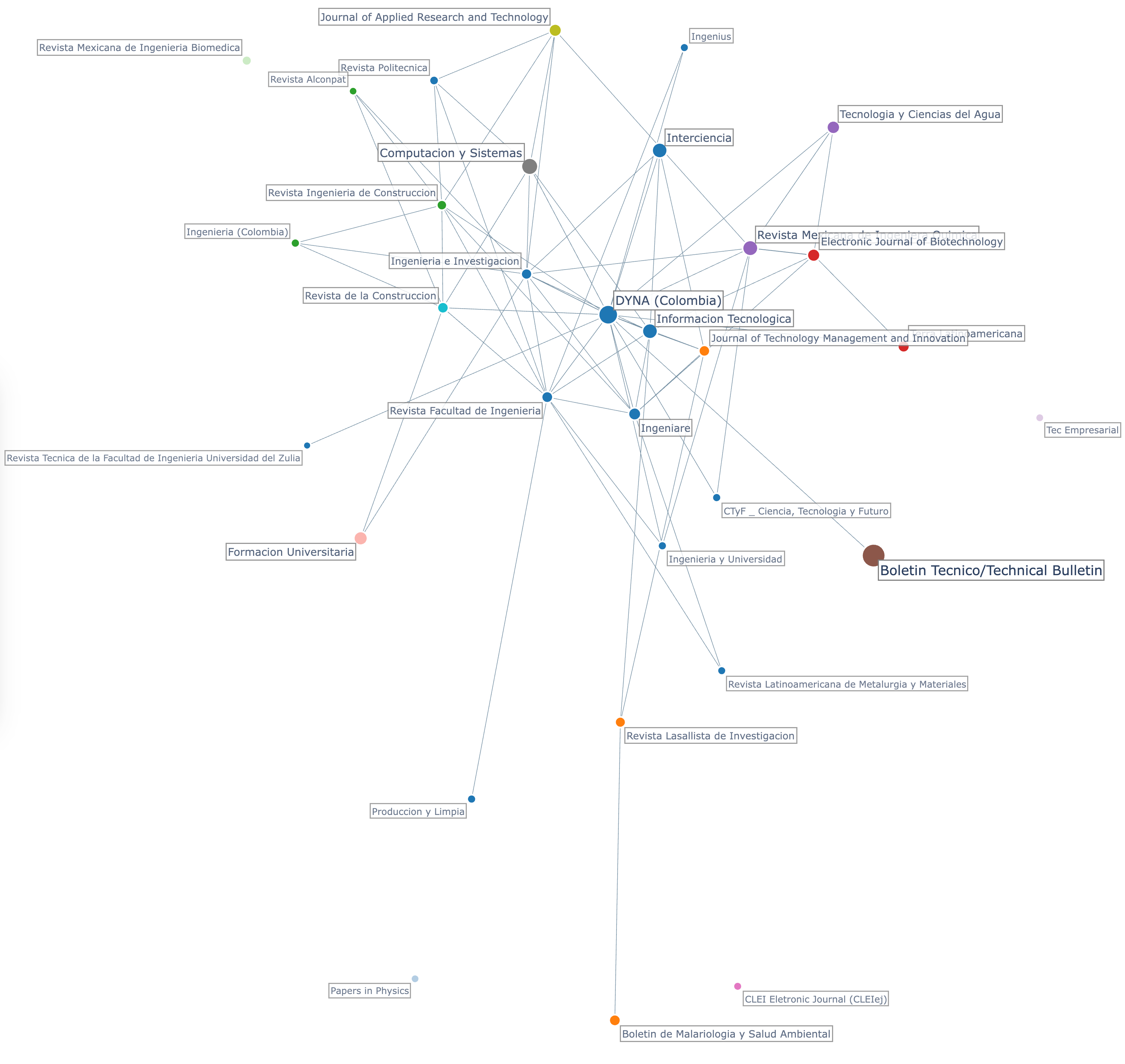


Figure 5. Citation network.

Source: The authors.

### Co-citation Network

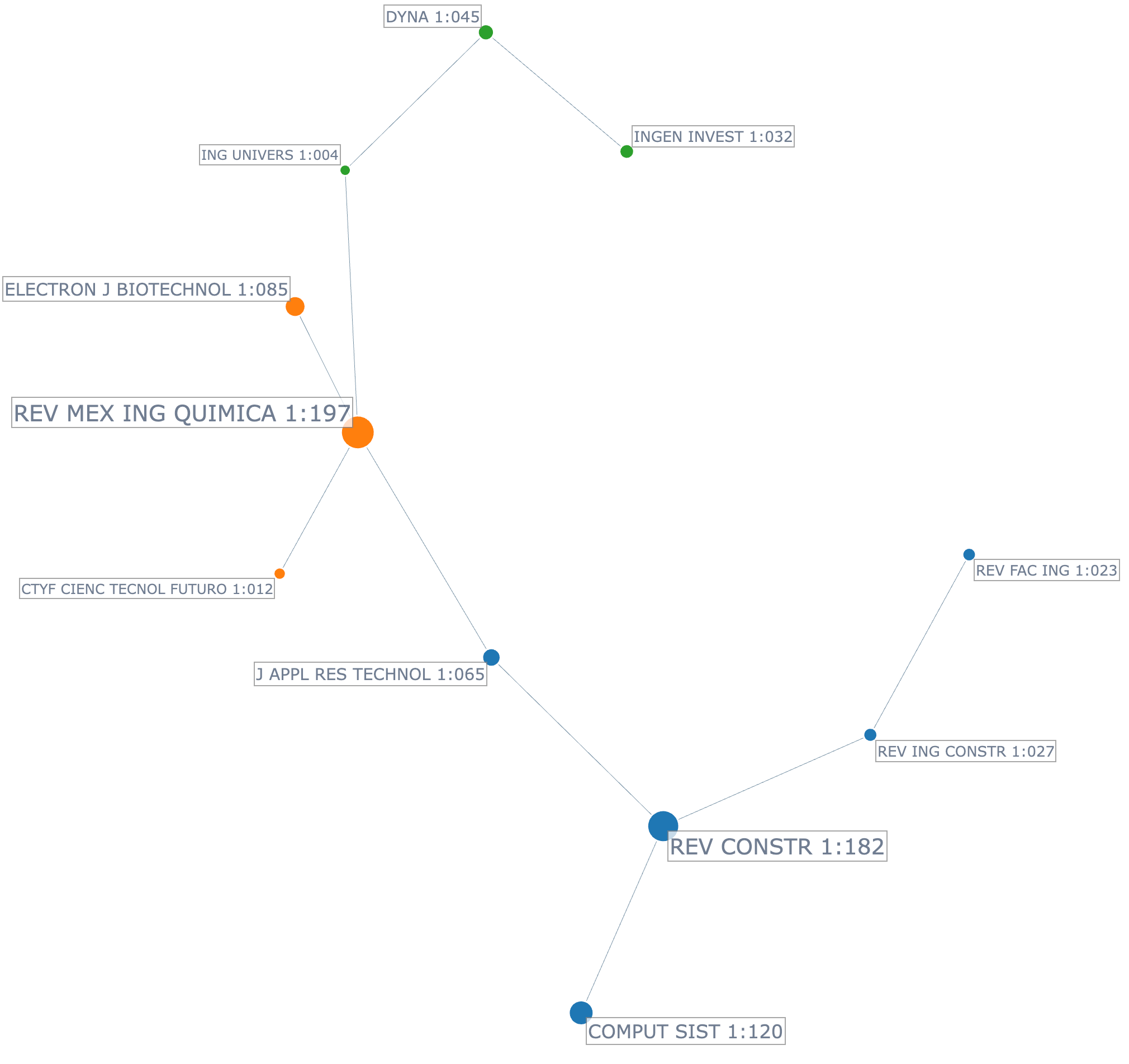


Figure 6. Co-citation network

Source: The authors.

### Coupling

### Co-occurrence

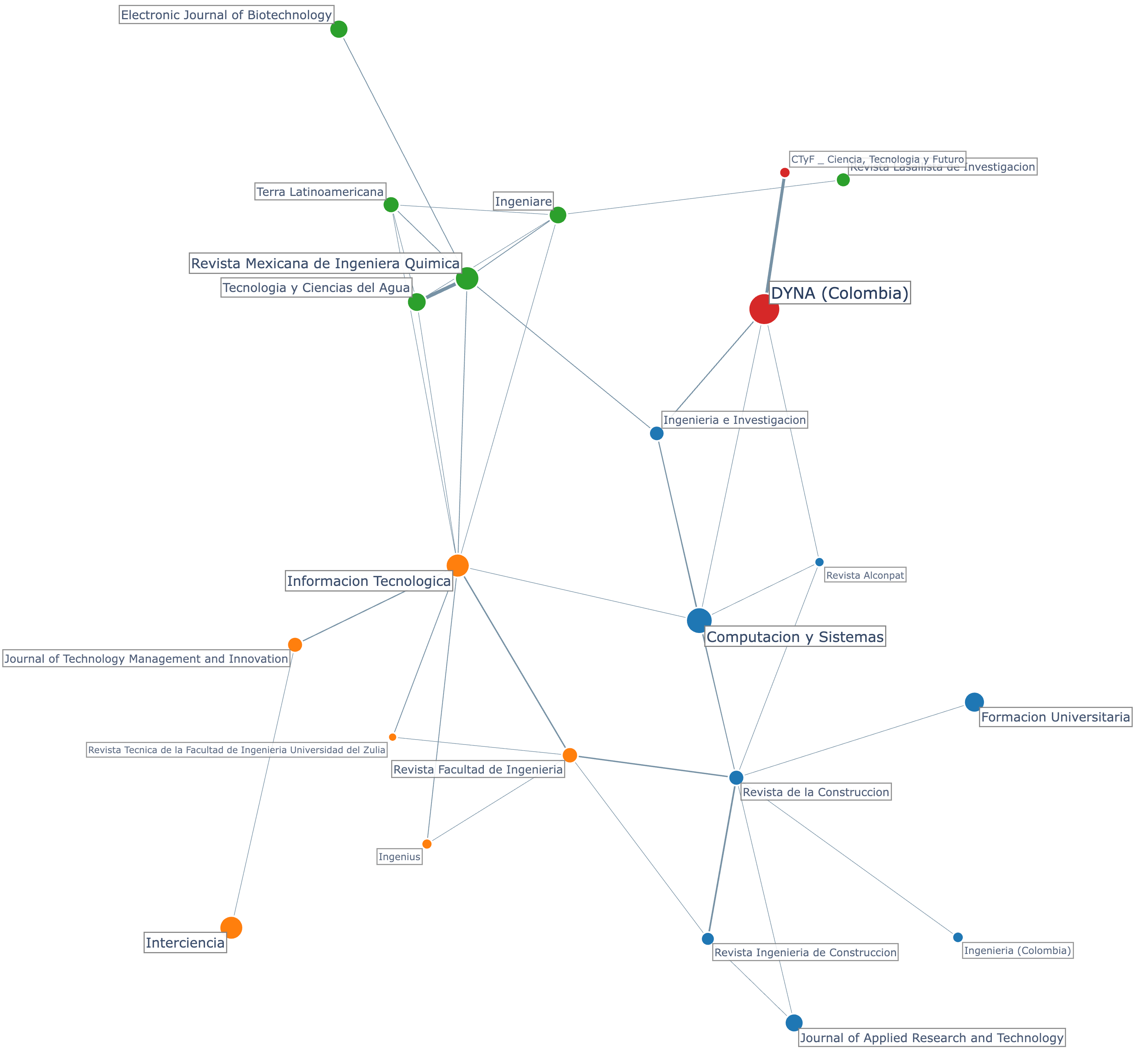


Figure 7. Coupling network.

Source: The authors.

## Most Cited Documents

.

Table 1.

Most Cited Documents in Retail Analytics.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Title | Authors | Document  Type | Rank  Global  Citations | Global  Citations | Rank  Local  Citations | Local  Citations |
| Shopping intention at AI-powered automated retail stores (AIPARS) | Pillai et al. [17] | Article | 1 | 201 | 2 | 13 |
| A study on investments in the big data-driven supply chain, performance measures Stuckenschmidt [\*142\*] and organizational performance in Indian retail 4.0 context | Gawankar et al. [19] | Article | 2 | 130 | 7 | 6 |
| Shopping with a robotic companion | Bertacchini et al. [20] | Article | 3 | 127 | 4 | 9 |
| Big data analytics and demand forecasting in supply chains: a conceptual analysis | Hofmann and Rutschmann [22] | Article | 4 | 111 | 16 | 4 |
| Consumer engagement via interactive artificial intelligence and mixed reality | Sung et al. [23] | Article | 5 | 103 | 52 | 1 |
| Drivers and impact of big data analytic adoption in the retail industry: A quantitative investigation applying structural equation modeling | Lutfi et al. [24] | Article | 6 | 95 | 27 | 3 |
| Retail business analytics: Customer visit segmentation using market basket data | Griva et al. [25] | Article | 7 | 91 | 11 | 5 |
| Agent-Based Modeling of Retail Electrical Energy Markets with Demand Response | Dehghanpour et al. [26] | Article | 8 | 91 | 117 | 0 |
| Indian shopper motivation to use artificial intelligence: Generating Vroom’s expectancy theory of motivation using grounded theory approach | Chopra [21] | Article | 9 | 87 | 5 | 8 |
| Retail sales forecasting with meta-learning | Ma and Fildes [27] | Article | 10 | 86 | 12 | 5 |
| Daily retail demand forecasting using machine learning with emphasis on calendric special days | Huber and Stuckenschmidt [16] | Article | 11 | 80 | 1 | 16 |
| State-of-the-art and adoption of artificial intelligence in retailing | Weber and Schütte [18] | Article | 18 | 70 | 3 | 12 |
| Revolution of Retail Industry: From Perspective of Retail 1.0 to 4.0 | Har et al. [28] | Article | 32 | 49 | 8 | 6 |
| Artificial intelligence in retail: applications and value creation logics | Cao [29] |  | 40 | 43 | 6 | 8 |
| Low cost embedded system for increasing retail environment intelligence | Pierdicca et al. [30] |  | 49 | 37 | 9 | 6 |
| Incorporating big data within retail organizations: A case study approach | Aversa et al. [31] | Article | 58 | 31 | 10 | 6 |

Source: The authors.

.

Table 3.

Dimensions for analysis of emerging topics.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Rank  Occurrences | Occurrences | Rank  Citations | Citations |
| JOURNALS |  |  |  |  |
| Journal of Retailing and Consumer Services | 1 | 18 | 1 | 632 |
| Lecture Notes in Networks and Systems | 2 | 14 | 56 | 25 |
| Lecture Notes in Computer Science | 3 | 10 | 16 | 79 |
| ACM International Conference Proceeding Series | 4 | 9 | 42 | 37 |
| International Journal of Retail and Distribution Management | 5 | 8 | 3 | 232 |
| Advances in Intelligent Systems and Computing | 6 | 8 | 74 | 18 |
| Sustainability (Switzerland) | 7 | 5 | 12 | 95 |
| Lecture Notes in Electrical Engineering | 8 | 5 | 92 | 12 |
| International Journal of Production Research | 9 | 4 | 2 | 246 |
| Procedia Computer Science | 10 | 4 | 13 | 93 |
| Annals of Operations Research | 19 | 3 | 9 | 116 |
| International Journal of Information Management | 29 | 2 | 4 | 182 |
| European Journal of Operational Research | 30 | 2 | 5 | 162 |
| Computers in Human Behavior | 31 | 2 | 6 | 148 |
| International Journal of Logistics Management | 32 | 2 | 7 | 131 |
| International Journal of Physical Distribution and Logistics Management | 33 | 2 | 8 | 118 |
| Expert Systems with Applications | 34 | 2 | 10 | 99 |
| COUNTRIES |  |  |  |  |
| India | 1 | 124 | 2 | 1131 |
| United States | 2 | 69 | 1 | 1190 |
| China | 3 | 56 | 5 | 429 |
| United Kingdom | 4 | 51 | 3 | 1020 |
| Germany | 5 | 35 | 4 | 497 |
| Italy | 6 | 20 | 6 | 412 |
| Canada | 7 | 20 | 8 | 188 |
| Russia | 8 | 18 | 25 | 63 |
| Brazil | 9 | 14 | 7 | 215 |
| Ireland | 10 | 13 | 9 | 188 |
| Hong Kong | 21 | 7 | 10 | 177 |
| ORGANIZATIONS |  |  |  |  |
| Maynooth Univ. (IRL) | 1 | 8 | 9 | 129 |
| Univ. of Appl. Sciences Upper Austria (AUT) | 2 | 6 | 5 | 158 |
| Amity Univ. (IND) | 3 | 6 | 36 | 76 |
| Univ. of Bristol (GBR) | 4 | 5 | 14 | 107 |
| Dublin City Univ. (IRL) | 5 | 5 | 37 | 76 |
| Massachusetts Inst. of Technol. (USA) | 6 | 5 | 58 | 55 |
| Univ. of Moratuwa (LKA) | 7 | 5 | 204 | 14 |
| Univ. of Duisburg-Essen (DEU) | 8 | 4 | 20 | 100 |
| Univ. of Tennessee (USA) | 9 | 4 | 33 | 81 |
| Univ. of Bologna (ITA) | 10 | 4 | 59 | 55 |
| Swansea Univ. (GBR) | 37 | 2 | 1 | 242 |
| Montana State Univ. (USA) | 38 | 2 | 4 | 194 |
| Univ. of Mannheim (DEU) | 39 | 2 | 6 | 156 |
| Pune Inst. of Bus. Manag. (IND) | 116 | 1 | 2 | 201 |
| Sri Balaji Univ. (IND) | 117 | 1 | 3 | 201 |
| California State Univ. (USA) | 118 | 1 | 7 | 130 |
| Nac. Inst. of Ind. Eng. (NITIE) (IND) | 119 | 1 | 8 | 130 |
| Università della Calabria (ITA) | 120 | 1 | 10 | 127 |
| AUTHORS |  |  |  |  |
| Razmochaeva N.V. | 1 | 7 | 117 | 46 |
| Bezbradica M. | 2 | 5 | 57 | 76 |
| Cirqueira D. | 3 | 5 | 58 | 76 |
| Helfert M. | 4 | 5 | 59 | 76 |
| Klionskiy D.M. | 5 | 5 | 240 | 21 |
| Griva A. | 6 | 4 | 12 | 122 |
| Frontoni E. | 7 | 4 | 17 | 103 |
| Pantano E. | 8 | 4 | 20 | 103 |
| Frazzon E.M. | 9 | 3 | 15 | 106 |
| Pereira M.M. | 10 | 3 | 16 | 106 |
| Huber J. | 33 | 2 | 4 | 156 |
| Stuckenschmidt H. | 34 | 2 | 5 | 156 |
| Dwivedi Y.K. | 137 | 1 | 1 | 201 |
| Pillai R. | 138 | 1 | 2 | 201 |
| Sivathanu B. | 139 | 1 | 3 | 201 |
| Gawankar S.A. | 140 | 1 | 6 | 130 |
| Gunasekaran A. | 141 | 1 | 7 | 130 |
| Kamble S. | 142 | 1 | 8 | 130 |
| Bertacchini F. | 143 | 1 | 9 | 127 |
| Bilotta E. | 144 | 1 | 10 | 127 |

Source: The authors.

Table 4.

Emergent topics clusters.

|  |  |  |  |
| --- | --- | --- | --- |
| Cluster Name | Num Terms | Percentage | Main Terms |
| Retail Sales prediction | 9 | 12.5 | Prediction; Retail Sales; Sales Forecasts; Sales Data; Time Series; Retail Trade; Strategic Decisions; Sales Promotions; Sales Prediction |
| AI-Driven Customer Insights | 9 | 12.5 | Artificial Intelligence; Retail Industry; Retail Organizations; Customer Satisfaction; Artificial Intelligence Technology; Computer Vision; Experience; Information Technology; Business Performance |
| Consumer Behavior and Price Dynamics | 8 | 11.1 | Consumer Behavior; Insights; Retail Operators; Data Sets; Consumption Behaviors; Customer Engagement; Price Dynamics; Pricing |
| ML for Predictive Modeling | 7 | 9.7 | Random Forest; Decision Trees; Retail Location; Logistic Regression; Predictive Models; Support Vector Machine; Boosting |
| AI-driven Retail Performance | 7 | 9.7 | Learning Systems; Neural Network; Deep Learning; Convolutional Neural Networks; Radio Frequency Identification; Performance Metrics; Supervised Learning |
| Data-Driven Social and Consumer Dynamics | 6 | 8.3 | Social Media; Consumer; Retail Banks; Finance; Robots; Retail Data |
| Customer-Centric Data-Driven Strategies | 6 | 8.3 | Customers; Data Mining; Customer Relationship Management; Business Analytics; Electronic Commerce Websites; Mobile Devices |
| Consumer-Centric Experience | 5 | 6.9 | Customer Experience; Management; Customer Service; Customer Data; Consumer Data |
| Predictive Customer Behavior Systems | 5 | 6.9 | Customer Behavior; Information System; Recommender Systems; Transaction Data; Customer Demands |
| Predictive Customer Behavior Systems | 4 | 5.6 | Decision Making; Decision Support Systems; Decisions; Efficiency |
| Human-Centered Business Process Strategy | 3 | 4.2 | Strategy; Business Processes; Human Resource Managers |
| Fashion Analytics | 3 | 4.2 | Data Science; Data Analysis; Fashion |

Source: The authors.

.

# Conclusions

.

# References

[1]

**J.D. Velásquez-Henao** earned his BS in Civil Engineering in 1994, an MS in Systems Engineering in 1997, and a PhD in Energy Systems in 2009, all from the National University of Colombia in Medellin, Colombia. From 1994 to 1999, he worked in electricity utilities and consulting companies in the power sector. In 2000, he joined the National University of Colombia in Medellin and was appointed a full professor of computer science in 2012. Between 2004 and 2006, he served as an Associate Dean for Research, and from 2009 to 2018, he led the Computing and Decision Science Department at the Facultad de Minas, National University of Colombia, Medellin. His research and publications span simulation, modeling, optimization, and forecasting in energy markets. He specializes in nonlinear time-series analysis and forecasting using statistical and computational intelligence techniques, numerical optimization with metaheuristics, and analytics and data science. He currently instructs postgraduate courses in data science, machine learning, and big data in the Analytics program, emphasizing Python programming.

ORCID: 0000-0003-3043-3037