Rev. Cient. Sist. Inform. 5(1), e911, doi: 10.51252/rcsi.v5i1.911

UNIVERSIDAD NACIONAL DE SAN MARTÍN

RCSI

https://revistas.unsm.edu.pe/index.php/rcsi
e-ISSN: 2709-992X

Original article
Artículo original
Jan-Jun, 2025

Latest research trends on open-source research

Últimas tendencias de investigación en investigación de código abierto

La-Cruz-Orbe, Sergio [1]
Ramos-y-Yovera, Santiago Ernesto [1]
Luperdi-Ríos, Fiorella Victoria [1]
Neri Ayala, Abrahán César [1*]

[1]Universidad Nacional José Faustino Sánchez Carrión. Huacho, Perú

Received: 18 Oct. 2024 | **Accepted:** 30 Dic. 2024 | **Published:** 20 Jan. 2025

Corresponding author*: aneri@unjfsc.edu.pe

How to cite this article: La-Cruz-Orbe, S., Ramos-y-Yovera, S.E., Luperdi-Ríos, F.V., & Neri Ayala, A. C. (2025). Latest research trends on open-source research. *Revista Científica de Sistemas e Informática*, *5*(1), e911. https://doi.org/10.51252/rcsi.v5i1.911

ABSTRACT

We analyzed the latest trends in open-source research from 2020 to 2024. We used bibliometric techniques such as productivity and author collaboration, productivity and institutional collaboration, and co-occurrence of terms. We used Scopus to collect the sample we analyzed. Regarding the productivity by authors, those who stand out the most have exceeded the threshold of 9 articles in the last five years. Among the organizations, the most productive are the University of the Chinese Academy of Sciences (China), Carnegie Mellon University (United States), Zhejiang University (China), Massachusetts Institute of Technology (MIT, United States), and Stanford University (United States). The co-word analysis revealed five topical clusters: open-source software and artificial intelligence, medical research and scientific methodologies, systematic reviews and health impact studies, simulation, quantum physics and complex systems, and medical and demographic research.

Keywords: open-source; open systems; bibliometric analysis; research trends

RESUMEN

Analizamos las últimas tendencias en investigación sobre código abierto desde el 2020 hasta el 2024. Utilizamos técnicas bibliométricas como productividad y colaboración de autores, productividad y colaboración institucional, y coocurrencia de términos. Se empleó Scopus para recopilar la muestra analizada. En cuanto a la productividad por autores, los que más destacan han superado el umbral de nueve artículos en los últimos cinco años. Entre las organizaciones, las más productivas son la University of the Chinese Academy of Sciences (China), la Carnegie Mellon University (Estados Unidos), la Zhejiang University (China), el Massachusetts Institute of Technology (MIT, Estados Unidos) y la Stanford University (Estados Unidos). El análisis de palabras asociadas reveló cinco grupos temáticos: software de código abierto e inteligencia artificial, investigación médica y metodologías científicas, revisiones sistemáticas y estudios de impacto en la salud, simulación, física cuántica y sistemas complejos, e investigación médica y demográfica.

Palabras clave: código abierto; sistemas abiertos; análisis bibliométrico; tendencias de investigación

1. INTRODUCTION

The term “open source” emerged in 1998 when Netscape decided to release the code for its browser. It was later formalized with the Open-Source Initiative to distinguish it from free software and its associated connotations (Aksulu & Wade, 2010). Open-source software is typically developed through collaborative projects, often initiated by individuals or groups to create a technological solution that meets their needs (Von Krogh & Von Hippel, 2006). According to Von Krogh & Spaeth (2007), since the start of the 21st century, open-source software has spurred significant research across various social science disciplines. Five key characteristics explain its academic appeal: (1) impact, as it significantly influences the economy and society; (2) theoretical tension, as it challenges the predictions and explanations of existing theories in different fields; (3) transparency, because it offers unprecedented access to data for research; (4) community reflexivity, as developers actively analyze and discuss the functioning of the software, even forming their own research community; and (5) proximity to science, given that its innovation process resembles the production of knowledge in academic research.

Furthermore, since the early 21st century, a considerable body of research on open-source software has emerged, spanning various academic disciplines and employing multiple methodological approaches (Von Krogh & Von Hippel, 2006). Consequently, several authors utilize bibliometric techniques to analyze the literature on scientific topics (Ayala et al., 2024). For instance, Raasch et al. (2013) investigate research development in open-source innovation, illustrating a shift from interdisciplinarity to multidisciplinarity during its first decade. Based on 306 publications and 10,000 references, the authors observe that academic collaboration shifts from joint problem-solving to parallel approaches, accompanied by a rapid decline in interdisciplinary exchange. The findings suggest that these patterns are prevalent and influenced by uncertainty and the modularization of research, with significant implications for academics and science policy.

Murray et al. (2022) analyze the use of open-source information in human rights investigations, focusing on its application in United Nations missions and commissions of inquiry. Through interviews and source reviews, they highlight its value in planning, evidence gathering, and overcoming access barriers. However, they also note challenges such as the exclusion of marginalized communities, disinformation, and the significant consumption of resources. They conclude that, given the increasing role of this information, it is crucial to provide these investigations with institutional support, funding, and methodological rigor. Toro et al. (2020) also examine the evolution and trends of the subject through a bibliometric study of 289 documents extracted from Scopus. The results indicate a growing interest in the topic, with high academic productivity in countries like the United States. The barriers to its adoption in specific sectors are identified, and the main areas of study include innovation, Linux, FLOSS, engineering requirements, risk management, open innovation, the public sector, social network analysis, and total cost of ownership.

Using a longitudinal approach, Duarte & Teodoro (2021) examine the evolution of free software application development in geographic information systems (GIS) from 2010 to 2020, focusing on two main objectives: (i) providing an overview of the creation of open-source GIS software in desktop, web, and mobile environments, and (ii) evaluating open-source GIS plugins used in environmental sciences. Through a bibliometric analysis based on the Web of Science database, the authors highlight the rise of mobile applications in the context of big data and the Internet of Things (IoT). The study concludes that these advancements have enhanced the use and applicability of GIS software across various areas of environmental research. Newby et al. (2003) apply Lotka's Law to analyze metadata on open-source software development, aiming to evaluate the productivity of authors in this field. The research relies on data extracted from the Linux Software Map (LSM) and SourceForge, one of the largest platforms for free software developers. The findings indicate that this law aids in understanding productivity trends in open-source software development and has the potential to predict developers' aggregate behavior in this ecosystem.

Based on data from IEEE Xplore and ACM Digital Library, Silva et al. (2019) examine the key research trends related to free and open-source software. By employing keyword co-occurrence and co-authorship techniques, the findings indicate an almost exponential increase in publications on this topic, despite academic output being dominated by small groups of researchers with limited collaboration. Using data from Emerald, Khode and Thakkar (2012) carried out a bibliometric analysis of open-source software publications in journals the company published between April 1998 and May 2010. The results reveal that most articles (63.49%) were authored by a single writer, with the journals Library Hi Tech and Program: Electronic Library and Information Systems being the primary sources of publication on the subject.

Other research precedents include the study by Kanwal (2017), which examines the scientific production of free and open-source software in Scopus. The author indicates that medicine generates the highest production (31.72%), that PLOS ONE is the most influential journal (20.73%), and that the University of Washington is the most productive institution. Silva et al. (2024) analyze 897 articles on free and open-source software, highlighting key topics such as health, costs, information systems, and emerging areas like IoT, educational computing, and digital health. Lastly, using bibliometrics and visual analysis, Shang et al. (2022) explore open-source intelligence research's status, trends, and key points. The findings reveal accelerated growth in the field, with increasing interdisciplinarity and technological dependence. It is emphasized that this technological advancement improves the accuracy of open-source intelligence, supports sustainable development, and facilitates its industrialization and future research. There are several bibliometric studies that have been carried out on this subject (Evangelatos et al., 2018; Bansal & Bansal, 2021; Velmurugan, C., & Radhakrishnan, 2016; Zhao & Wei, 2017).

In this study, we will analyze the latest trends in open-source research. We will use bibliometric techniques to identify productivity patterns and collaboration among authors and institutions to achieve this. We will also employ the co-word technique to generate maps that help us determine the topics most explored in recent years.

2. MATERIALS AND METHODS

Given the importance of Scopus in conducting bibliometric studies (Vizalote-Rodriguez, 2022; Ausejo Sánchez et al., 2024; Patralekh, Vaishya & Vaish, 2024), we used the database to collect the sample we analyzed. We chose the terms "open source" and "open system" in the title to extract the most relevant literature. To examine recent trends, we focused on the last five years (2020-2024). The final search strategy was as follows: (TITLE (open AND source*) OR TITLE (open AND system*)) AND PUBYEAR > 2019 AND PUBYEAR < 2025 AND (LIMIT-TO (DOCTYPE, "cp") OR LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "ch") OR LIMIT-TO (DOCTYPE, "re") OR LIMIT-TO (DOCTYPE, "cr")). Through this method, we retrieved 14,152 documents, which were processed for further analysis.

Given that we worked with three indicator groups—productivity and author collaboration, productivity and institutional collaboration, and co-occurrence of terms—we standardized each variable related to these indicators in an ad hoc database, including author names, institution names, and author keywords and keyword plus. We then counted the frequency distribution to determine productivity, while VOSviewer was employed to generate collaboration and co-word maps. We applied a threshold of five for the author and institution collaboration maps as the minimum productivity. In contrast, the co-occurrence map was created from keywords with a frequency greater than 20. A description and qualitative analysis of each map were conducted, considering the frequency and grouping of the nodes, as well as the link weight between nodes and its grouping structure.

3. RESULTS AND DISCUSSION

The first of our results refers to productivity by authors. Those who stand out the most have published many studies, exceeding the threshold of 9 articles in the last five years (Table 1). Among them, Pearce, Joshua M. (57), Wang, Wei (28), Liu, Yang (24), Tamura, Yoshinobu (18), Yamada, Shigeru (18), Benini, Luca (15), Riehle, Dirk (15), Steinmacher, Igor (15), Vasilescu, Bogdan (15), among others, stand out. In the case of the most productive, Pearce, Joshua M., we are in the presence of an author who researches open source applied to distributed manufacturing, 3D printing, renewable energy, and medical devices. His work encompasses the development of open-source hardware and software for accessible manufacturing, automation and optimizing sustainable energy systems. It also studies technological resilience through open solutions for health crises and supply shortages, promoting the democratization of access to technology. Another case is Wang, Wei, who researches the control of permanent magnet linear motors with a focus on traction and drive systems with open windings. His main areas of interest include fault-tolerant control, zero-sequence current suppression, predictive voltage modeling, and fault detection in sensors and open circuits.

On the other hand, there is a group of small producers, who have published less than five articles, such as Alam, Ekky Novriza (5), Alenezi, Mamdouh (5), Auer, Hans (5), Balasubramanian, Karthikeyan (5), Barcomb, Ann (5), Boudaoud, Hakim (5), and many others. It should be emphasized that, although their scientific output is less, they are authors who may specialize in specific niches within the field, contributing valuable knowledge in a more focused way.

Table 1. Author scientific output on open source

Author	Documents	Citations
Pearce, Joshua M.	57	744
Wang, Wei	28	116
Liu, Yang	24	119
Tamura, Yoshinobu	18	20
Yamada, Shigeru	18	20
Benini, Luca	15	269
Riehle, Dirk	15	75
Steinmacher, Igor	15	224
Vasilescu, Bogdan	15	162
Braune, Katarina	14	235
De Maria, Carmelo	13	60
Carollo, Federico	12	223
Díaz Lantada, Andrés	12	23
Lewis, Dana M.	12	159
Wang, Tao	12	70
Gao, Wei	11	42
Hua, Wei	11	142
Sun, Dan	11	245
Wang, Zheng	11	68
Zhang, Wei	11	37
Zhou, Minghui	11	70
Ahluwalia, Arti	10	12
Cao, Jian	10	15
Cheng, Ming	10	54
Gamalielsson, Jonas	10	73
Liu, Xin	10	77
Nian, Heng	10	223
Pan, David Z.	10	130
Saligane, Mehdi	10	59
Wang, Yang	10	40
Wang, Yu	10	67
Zhang, Bin	10	43

Some authors stand out for their productivity and the high number of citations they receive. Those who have accumulated over 159 citations significantly impact the scientific community (Table 1). Among them are Pearce, Joshua M. (744), Benini, Luca (269), Steinmacher, Igor (224), Vasilescu, Bogdan (162), Braune, Katarina (235), Carollo, Federico (223), Lewis, Dana M. (159), Sun, Dan (245), Nian, Heng (223), Wiese, Igor (180), Curri, Vittorio (298), Lovett, Brendon W. (243), Raile, Klemens (213), Wang, Yan (377), Zhang, Yu (217), Eliceiri, Kevin W. (231), Grammel, Gert (254), Chen, Chen (522), Galimberti, Gabriele (295), Hussain, Sufyan (183), Li, Ying (212), Liu, Zhiyuan (211), Ouni, Ali (185), Sun, Maosong (234), Balasubramanian, Karthikeyan (251), Chen, Xin (323), Ferrari, Alessio (251), Filer, Mark (251), Keeling, Jonathan (159), Kundrat, Jan (251), Mkaouer, Mohamed Wiem (182), Plate, Henrik (243), and Yin, Yawei (251). Their work has been widely referenced, indicating the relevance of their research.

We also highlight authors who exhibit both high productivity and substantial citation rates. This group includes Pearce, Joshua M. (57 documents, 744 citations), Benini, Luca (15 documents, 269 citations), Steinmacher, Igor (15 documents, 224 citations), Vasilescu, Bogdan (15 documents, 162 citations), Braune, Katarina (14 documents, 235 citations), Carollo, Federico (12 documents, 223 citations), Lewis, Dana M. (12 documents, 159 citations), Sun, Dan (11 documents, 245 citations), and Nian, Heng (10 documents, 223 citations). They not only produce numerous publications but also significantly impact the academic community.

A cluster analysis of the co-authorship network reveals distinct collaborative communities among researchers (Figure 1). The largest cluster includes influential authors such as *Chen, Chen*, *Chen, Lei*, *Chen, Liang*, *Jiang, Xuefeng*, *Li, Feng*, *Li, Qiang*, *Li, Yang*, *Li, Yun*, *Liu, Qian,* and *Lu, Yao*, who have a significant number of publications and citations. They play a central role in disseminating knowledge. Their collaboration is characterized by joint efforts across multiple open-source fields, covering technical aspects and applications in various industries. The work published by the authors in cluster one primarily addresses developing and applying open models and diagnostic systems for advanced electrical and thermal systems. This research focuses on the simulation and optimization of energy systems, emphasizing Modelica, an open modeling language, to analyze the efficiency and control of air conditioning and refrigeration systems, particularly in chiller plants with water economizers and steam-based district heating systems. Additionally, open models have been developed for thermal management in data centers, facilitating the optimization of energy consumption within these infrastructures.

Another significant cluster consists of *Bao, Yungang*, *Chen, Hao*, *Li, Huawei*, *Li, Shuai*, *Li, Wei*, *Lin, Yibo*, *Pan, David Z.*, *Wang, Hao*, *Yu, Bei*, and *Zhang, Peng*, who have conducted research focused on specific approaches to open-source software, emphasizing a technical and applied perspective. This group is marked by constant collaboration, reflecting strong thematic cohesion in their studies. The work published by the authors of cluster 2 centers on developing open-source tools for the design and automation of integrated circuits, hardware synthesis, and the optimization of computational architectures. A central theme involves creating analog and digital circuit design systems using automated open-source methodologies, highlighting tools such as MAGICAL, an automatic layout system for analog integrated circuits that transitions from the netlist to the final design generation in GDSII format. Also featured is DREAMPlace 2.0, a GPU-accelerated system for placing components in Very-Large-Scale Integration (VLSI) designs, optimizing performance and scalability for large-scale projects. Another area of research within this cluster focuses on synthesizing and automating the design of reconfigurable hardware and open-source processors. Studies like DREAMPlaceFPGA propose solutions for analytical placement of elements in heterogeneous FPGAs utilizing deep learning tools.

Additionally, the design and implementation of XiangShan, a high-performance processor based on the RISC-V architecture, stands out, aiming to establish an open and efficient alternative in CPU development. Besides circuit and hardware design, researchers in cluster 2 also address the impact and evolution of the open-source model in the semiconductor industry and technological innovation in China. One paper examines the construction of an open-source-based CPU ecosystem, exploring both its opportunities and challenges regarding adoption and development in the global technology landscape.

A third prominent cluster consists of *Gao, Wei*, *Li, Hui*, *Sun, Yu*, *Wang, Qian*, *Wang, Tao*, *Wang, Zhe*, *Yang, Yang*, *Yu, Yue*, *Zhang, Wei,* and *Zhang, Yang. This group* illustrates a community of researchers who frequently collaborate in open-source software development and software engineering. Their co-authorship network has expanded over time, incorporating new collaborators and strengthening connections across various knowledge areas. Additionally, smaller but highly specialized clusters have been identified, such as one that includes *Chen, Yu*, *Li, Chao*, *Liu, Xin*, *Rabault, Jean*, *Wang, Hui*, *Wang, Yu*, *Wu, Jie*, *Xie, Lei*, and *Zhang, Jing Jing*, which focuses on studies regarding the economic impact and adoption of open-source software. Another cluster involves *Chen, Bihuan*, *Chen, Yang*, *Li, Hao*, *Liu, Yang*, *Wang, Yang*, *Yang, Chao*, *Zhang, Jian,* and *Zhao, Bin*, whose work concentrates on process automation and optimizing open-source software.

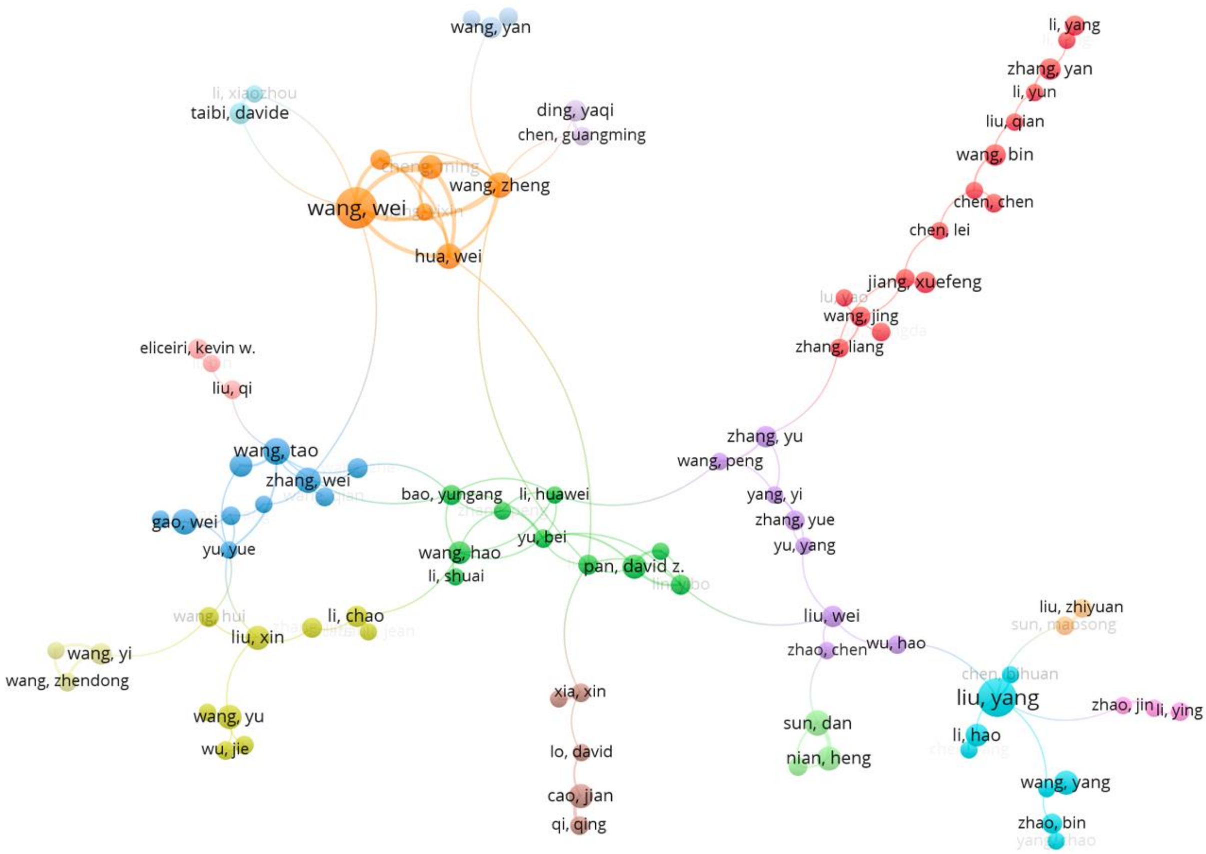

Figure 1. Co-authorship network on open-source research based on Scopus

Among the organizations, the most productive are those that have published a substantial number of documents, exceeding the threshold of 14 articles during the analyzed period (Table 2). Notable among them are the University of the Chinese Academy of Sciences (China), Carnegie Mellon University (United States), Zhejiang University (China), Massachusetts Institute of Technology (MIT, United States), Stanford University (United States), National Renewable Energy Laboratory (United States), Western University (Canada), National University of Singapore (Singapore), Politecnico di Torino (Italy), Southeast University (China), and Tampere University (Finland). These institutions have made significant contributions to the academic literature in this field through consistent, high-impact output, establishing themselves as leaders in research on open-source technologies.

The organizations with the lowest productivity have produced fewer than 5 documents. While their output is lower, these institutions might specialize in specific areas. Although these contributions are fewer in number, they can significantly influence particular developments or research niches that impact smaller academic or industrial communities.

Table 2. Research output per institution

Organization	Documents	Citations
University Of Chinese Academy of Sciences, China	39	339
Carnegie Mellon University, United States	34	247
Zhejiang University, China	19	305
Massachusetts Institute of Technology, United States	18	264
Stanford University, United States	16	371
National Renewable Energy Laboratory, United States	15	214
Western University, Canada	15	92
National University of Singapore, Singapore	14	100
Politecnico Di Torino, Italy	14	95
Southeast University, China	14	83
Tampere University, Tampere, Finland	14	112
Georgia Institute of Technology, United States	13	76
Tsinghua University, Beijing, China	13	178
University of Toronto, Canada	13	71
Harvard University, United States	12	101
University College London, United Kingdom	12	46
Duke University, United States	11	136
Peng Cheng Laboratory, China	11	251
Universität Tübingen, Germany	10	202
University Of Science and Technology of China, China	10	134

Some organizations stand out for their productivity and the high number of citations they receive. Those that have accumulated over 178 citations significantly impact the scientific community. Among these are the University of Chinese Academy of Sciences (China), Carnegie Mellon University (United States), Zhejiang University (China), Massachusetts Institute of Technology (MIT, United States), Stanford University (United States), National Renewable Energy Laboratory (United States), Tsinghua University (China), Peng Cheng Laboratory (China), and the Institut für Theoretische Physik, Universität Tübingen (Germany). Their work has been widely referenced, indicating its relevance and applicability in academic and technological fields. Organizations that combine high productivity with high citation rates include those that have published over 14 articles and received more than 178 citations. The combination of consistent output and high visibility in citations reflects their influence on knowledge generation and consolidating trends within the open-source field.

The institutional open-source collaboration network highlights a series of key interconnections among the most influential academic institutions (Figure 2). Notable organizations with the highest number of collaborations include Carnegie Mellon University (USA), Harvard University (USA), the National University of Singapore (Singapore), Georgia Institute of Technology (USA), Massachusetts Institute of Technology (MIT, USA), Singapore Management University (Singapore), Stanford University (USA), Harvard Medical School (USA), Nanyang Technological University (Singapore), and the University of Toronto (Canada). These institutions have established connections with various entities within the global open-source research network, acting as central nodes in academic cooperation.

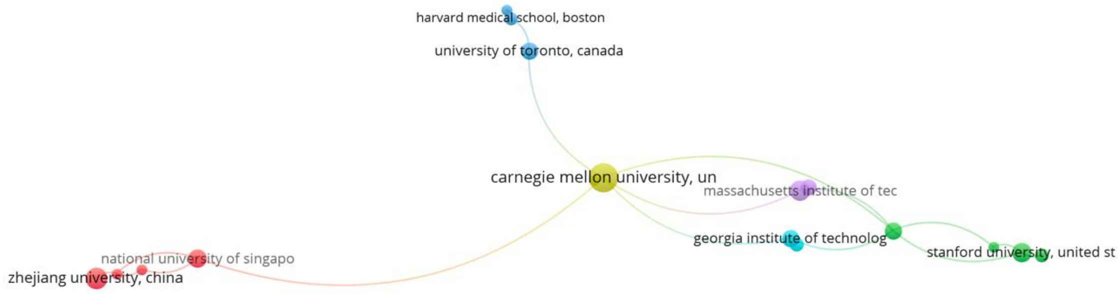

Figure 2. Institutional collaboration network on open-source research based on Scopus

The most interconnected institutions promote academic cooperation and the sharing of knowledge within the scientific community. For instance, Harvard University and Carnegie Mellon University sustain active collaboration networks with research centers in North America, Europe, and Asia, forging strategic partnerships. Stanford University and MIT have enhanced their cooperation with Asian universities like Zhejiang University and the National University of Singapore, reflecting the increasing significance of open source in technological innovation throughout the Asian region.

Another noteworthy aspect of the network is the strong connection among Singaporean universities, such as Nanyang Technological University and Singapore Management University, which have formed strategic partnerships with both Chinese and American institutions for projects in artificial intelligence, cybersecurity, and open-source hardware optimization. In Europe, institutions like Politecnico di Torino and the University of Tübingen have established collaborations that strengthen their positions within the global research ecosystem in this field.

When it comes to influence, universities like Stanford and MIT, which have the highest number of citations, have significantly impacted the academic community. This impact stems not only from their high-quality publications but also from their roles as centers of collaboration, facilitating the integration of knowledge from various regions around the globe. These connections illustrate how open-source research has transcended borders, promoting a model of international cooperation that fosters innovation and free access to knowledge.

The analysis of keywords in the co-occurrence map of open-source research terms shows a highly concentrated distribution of certain key concepts (Figure 3). A total of 1,278 different terms were identified, averaging 69.81 occurrences per keyword, with notable dispersion in the distribution (Table 3). The most frequent term in the dataset is “open source,” appearing 3,493 times, followed by “open-source software” with 3,418 instances, and “open systems,” which occurs 3,366 times. These terms highlight the main focus of the co-occurrence map on open-source software.

Beyond these dominant terms, other keywords suggest complementary areas of study within the field of research. For example, “human” appears 2,081 times, indicating a possible connection between the study of open source and its implications for human interaction, ethics, or impact on society. Other relevant keywords include “software”, “procedures”, “systematic review” and “controlled study”, suggesting a strong presence of rigorous methodological approaches within open source-related studies. In addition, terms such as “male” and “female” suggest studies that analyze open-source software's impact on different population groups or from a gender perspective. Although “open source” is still a term that is highly connected to others, the term with the strongest link in the dataset is “human”, with a value of 27,891, indicating its interconnection with a wide range of concepts within the field of study. It is followed by “open-source software” with 24,416, “open systems” with 23,613, and “open source” with 22,493. Other terms with high link strength also stand out, such as “procedures”, “male”, “female”, “systematic review,” and “controlled study”.

The distribution of keywords in this co-occurrence map shows a marked concentration in a few key terms, with a long tail of less frequent words. This indicates that, although there are multiple approaches in research on open source, studies tend to converge on certain essential concepts, such as open systems, research methodology, and human interaction with software. In addition, the link strength analysis suggests that some terms, although less frequent, have a significant impact within the knowledge network, as they are highly connected with other concepts.

Table 3. Top co-occurring terms per cluster

#	Keyword	Occurrences	#	Keyword	Occurrences
1	open source	3493	51	benchmarking	218
2	open source software	3418	52	open source hardware	217
3	open systems	3366	53	optimization	217
4	human	2081	54	length of stay	212
5	software	944	55	middle aged	210
6	procedures	762	56	reproducibility	210
7	systematic review	693	57	outcome assessment	204
8	controlled study	676	58	risk assessment	204
9	male	676	59	software testing	203
10	female	664	60	data handling	202
11	adult	579	61	image analysis	201
12	meta-analysis	553	62	laparoscopy	196
13	machine learning	547	63	digital storage	195
14	deep learning	417	64	students	195
15	open data	411	65	diagnostic imaging	194
16	algorithms	407	66	priority journal	193
17	python	393	67	3d printing	191
18	open quantum systems	391	68	open source frameworks	190
19	quantum optics	383	69	remote-sensing	190
20	treatment outcome	361	70	information management	189
21	open surgery	350	71	workflow	188
22	internet of things	345	72	physiology	183
23	application programs	342	73	data visualization	180
24	software design	335	74	adverse event	179
25	nonhuman	318	75	codes (symbols)	177
26	performance	318	76	matlab	173
27	costs	315	77	cost effectiveness	170
28	open sources	309	78	open source platforms	169
29	open source tools	306	79	clinical outcome	168
30	postoperative complications	303	80	intermethod comparison	165
31	automation	300	81	geographic information systems	161
32	learning systems	298	82	open source software projects	155
33	randomized controlled trial	287	83	robotics	152
34	simulation	274	84	visualization	152
35	open source projects	268	85	e-learning	151
36	artificial intelligence	259	86	learning algorithms	149
37	aged	258	87	forecasting	147
38	decision-making	257	88	network security	147
39	comparative study	255	89	fault detection	146
40	major clinical study	248	90	graphical user interfaces	146
41	high level languages	246	91	c++ (programming language)	144

42	quality control	239	92	clinical article	143
43	animals	237	93	computational linguistics	143
44	low-costs	234	94	covid-19	143
45	retrospective study	233	95	ecosystems	141
46	follow up	231	96	dataset	139
47	animal	227	97	data acquisition	138
48	operation duration	227	98	sensitivity analysis	137
49	image processing	223	99	computational fluid dynamics	135
50	timing circuits	223	100	computer simulation	135

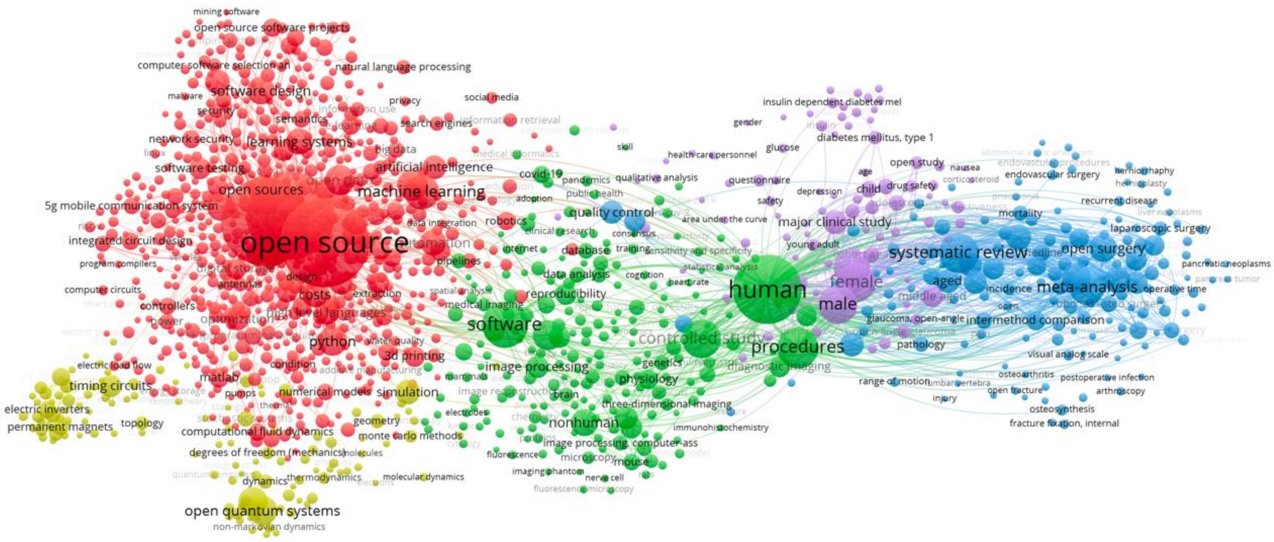

Figure 3. Co-word map on open-source research based on Scopus

In the first cluster on the map in Figure 3, which highlights open-source software and artificial intelligence, the most frequently co-occurring terms are “open-source software” (3,418), “open systems” (3,366), “open source” (3,493), “machine learning” (547), and “deep learning” (417). The strong relationship among “open-source software," “open systems," and “open source” suggests that studying open source is central to this cluster, linking it to emerging fields such as machine learning and deep learning. The co-occurrence of “machine learning” and “deep learning” indicates that artificial intelligence technologies significantly depend on open-source tools, possibly due to the availability of open libraries and frameworks.

In the second cluster focused on medical research and scientific methodologies, the most frequently co-occurring terms include “human” (2,081), “procedures” (762), “controlled study” (676), “software” (944), and “algorithms” (407). The relationship between “human” and “procedures” indicates that the studies in this cluster are oriented toward scientific methodologies applied to health and social sciences. The co-occurrence of “controlled study” and “procedures” suggests that open-source research is utilized to enhance the reproducibility of scientific studies and clinical experiments through open-source software that facilitates data analysis in a transparent and verifiable manner. The term “algorithms” within this group indicates that the development of new analytical methodologies in medical studies also relies on open tools, providing access to advanced computational models for interpreting clinical data.

The third cluster, focusing on systematic reviews and health impact studies, has its most commonly co-occurring terms: “systematic review” (693), “meta-analysis” (553), “open surgery” (350), “treatment outcome” (361), and “postoperative complications” (303). The relationship between “systematic review” and “meta-analysis” is crucial in this group, as both terms are strongly linked to studies that synthesize scientific evidence. The co-occurrence of “open surgery” with “treatment outcome” and “postoperative complications” indicates an interest in evaluating surgical outcomes through open tools, allowing researchers to collaboratively analyze clinical data and improve the efficiency of medical procedures based on empirical evidence.

In the fourth cluster, which covers simulation, quantum physics, and complex systems, the most frequently co-occurring terms are “simulation” (274), “quantum optics” (383), “timing circuits” (223), “open quantum systems” (391), and “fault detection” (146). The co-occurrence of “simulation” and “quantum optics” suggests that open-source tools are being employed to simulate quantum phenomena, an essential area in the advancement of quantum computing. The relationship between “open quantum systems” and “fault detection” implies that research in this field is also exploring ways to improve the stability and accuracy of quantum systems through open methods, which may be crucial for progress in the detection and correction of errors in emerging quantum technologies.

Finally, the fifth cluster, which focuses on medical and demographic research, includes terms such as “male” (676), “female” (664), “adult” (579), “major clinical study” (248), and “middle-aged” (210). The strong correlation between “male” and “female” indicates that gender segmentation is a vital component in studies using open-source software, possibly related to biomedical data analysis and demographic population studies. The co-occurrence of “adult” and “middle-aged” with “major clinical study” suggests that these studies involve large-scale clinical research, where open tools facilitate more accessible and replicable analysis of epidemiological data.

CONCLUSIONS

The findings demonstrate a rapid expansion of open-source research driven by highly productive authors, collaborative institutional networks, and the integration of AI, energy systems, quantum computing, and healthcare technologies. Leading institutions are crucial in establishing global connections that foster knowledge exchange. At the same time, thematic research trends show a continued evolution of open-source applications across various scientific and technological domains. The study also emphasizes the convergence of AI, quantum computing, medical research, and open systems. It suggests a future where open-source solutions will drive innovation and democratized access to technology.

FINANCING

The authors did not receive sponsorship to carry out this study-article.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest related to the development of the study.

AUTHORSHIP CONTRIBUTION

Conceptualization; Data curation; Formal analysis; Research; Methodology; Visualization; Writing - original draft; Writing - review and editing: La-Cruz-Orbe, S., Ramos-y-Yovera, S.E., Luperdi-Ríos, F.V. and Neri Ayala, A. C.

REFERENCES

Aksulu, A., & Wade, M. R. (2010). A comprehensive review and synthesis of open source research. *Journal of the Association for Information Systems*, *11*(11), 6. 576-656. https://doi.org/10.17705/1jais.00245

Ausejo Sánchez, J. L., Soto, F. G. C., Rosa, P. E. R. L., Palma, D. F. M., Campos, G. A. C., & Cadillo, A. J. R. (2024). Big data research in the business, management and accounting field: Revealing the thematic structure based on co-word analysis. *Iberoamerican Journal of Science Measurement and Communication*, *4*(1), 1–8. https://doi.org/10.47909/ijsmc.116

Ayala, A. C. N., Valenzuela Narváez, D. A. O., Manes Cangana, G. A., Orbe, S. L. C., Torres Anaya, A. J., & Ayala Huaynatte, E. (2024). Blended learning research: Identifying research output patterns in Scopus (2000-2023). *Iberoamerican Journal of Science Measurement and Communication*, *4*(2), 1–9. https://doi.org/10.47909/ijsmc.124.4

Bansal, J., & Bansal, M. (2021). Open Source Software (OSS): A Scientometric Study of Global Publications during 1999-18. *Journal of Indian Library Association*, 56(2), 84-92. https://www.ilaindia.net/jila/index.php/jila/article/view/707

Duarte, L., & Teodoro, A. C. (2021). GIS open-source plugins development: A 10-year bibliometric analysis on scientific literature. *Geomatics*, *1*(2), 206-245. https://doi.org/10.3390/geomatics1020013

Evangelatos, N., Satyamourthy, K., Levidou, G., Brand, H., Bauer, P., Kouskouti, C., & Brand, A. (2018). Use of free/libre open source software in sepsis "-omics" research: A bibliometric, comparative analysis among the United States, EU-28 Member States, and China. OMICS: A Journal of Integrative Biology, 22(5), 365-372. https://doi.org/10.1089/omi.2018.0032

Kanwal, P. (2017). A Bibliometric Study of World Research Output on Free and Open Source Software Literature during 1960-2016. *International Journal of Advanced Research in Computer Science*, *8*(3), 1067-1072. https://doi.org/10.26483/ijarcs.v8i3.3157

Khode, S., & Thakkar, S. (2012). Bibliometric analysis of literature published in emerald journals on open source. *Gyankosh-The Journal of Library and Information Management*, *3*(2), 49-54.

Murray, D., McDermott, Y., & Koenig, K. A. (2022). Mapping the use of open source research in UN Human Rights investigations. *Journal of Human Rights Practice*, *14*(2), 554-581. https://doi.org/10.1093/jhuman/huab059

Newby, G. B., Greenberg, J., & Jones, P. (2003). Open source software development and Lotka's law: bibliometric patterns in programming. *Journal of the American Society for information science and technology*, *54*(2), 169-178. https://doi.org/10.1002/asi.10177

Patralekh, M., Vaishya, R., & Vaish, A. (2024). Impact of COVID-19 on Indian biomedical research: A bibliometric analysis using online data from 2017 to 2022. *AWARI*, 5, 1–13. https://doi.org/10.47909/awari.65

Raasch, C., Lee, V., Spaeth, S., & Herstatt, C. (2013). The rise and fall of interdisciplinary research: The case of open source innovation. *Research policy*, *42*(5), 1138-1151. https://doi.org/10.1016/j.respol.2013.01.010

Shang, R.; Zhang, B.; Ma, H.; & Zhang, T. (2022). Research Status, Hotspots and Trends of Open Source Intelligence Abroad. *Journal of Modern Information*, 42(1), 5-13. https://doi.org/10.3969/j.issn.1008-0821.2022.01.001

Silva, D. G., Coutinho, C., & Costa, C. J. (2024). Exploration of Open Source in Low-Income Countries: Review and Future Research Directions. In *2024 International Symposium on Sensing and Instrumentation in 5G and IoT Era (ISSI)* (Vol. 1, pp. 1-6). IEEE. https://doi.org/10.1109/ISSI63632.2024.10720490

Silva, D., Aparicio, M., & Costa, C. J. (2019,). Free and open source bibliometric study. In *2019 14th Iberian Conference on Information Systems and Technologies (CISTI)* (pp. 1-6). IEEE. https://doi.org/10.23919/CISTI.2019.8760695

Toro, O. P., Correa, Y. A., Valencia-Arias, A., & Benjumea-Arias, M. (2020). A bibliometric analysis of the use of open source software in educational contexts. *Problems of Education in the 21st Century*, *78*(1), 114. https://doi.org/10.33225/pec/20.78.114

Velmurugan, C., & Radhakrishnan, N. (2016). Literature Research Trends and Authorship Pattern of Open Software Engineering Journal: a Scientometric Study. In A. Kishore et al (Eds.), Modern Trends in Library & Information Science in Digital Environment Festschrift Volume in Honour of Dr. U.M. Thakur (pp. 263 -274). Uttar Pradesh, India: University Prakashan.

Vizalote-Rodriguez, G. A. (2022). Contribución de las TI en la mejora de la productividad de las PYME. *Revista Amazonía Digital*, 1(1), e164. https://doi.org/10.55873/rad.v1i1.164

Von Krogh, G., & Spaeth, S. (2007). The open source software phenomenon: Characteristics that promote research. *The Journal of Strategic Information Systems*, *16*(3), 236-253. https://doi.org/10.1016/j.jsis.2007.06.001

Von Krogh, G., & Von Hippel, E. (2006). The promise of research on open source software. *Management science*, *52*(7), 975-983. https://doi.org/10.1287/mnsc.1060.0560

Zhao, R., & Wei, M. (2017). Impact evaluation of open source software: An altmetrics perspective. *Scientometrics*, 110, 1017-1033. https://doi.org/10.1007/s11192-016-2204-y