Revista Cubana de Salud y Trabajo. 2026;27:e1033

Estudio bibliométrico

Thematic and prospective analysis of open science studies within Web of Science

Análisis temático y prospectivo de estudios de ciencia abierta en *Web of Science*

Luis Ernesto Paz Enrique[1,2*] https://orcid.org/0000-0001-9214-3057

Wileidys Artigas Morales[3,4] https://orcid.org/0000-0001-6169-5297

Eduardo Alejandro Hernández Alfonso[1,2] https://orcid.org/0000-0002-6446-1653

[1]Universidad Nacional Autónoma de México. México.

[2]Universidad de Granada. España.

[3]Universidad del Zulia. Venezuela.

[4]Universidade Óscar Ribas. Angola.

[*]Corresponding author: luisernestopazenrique@gmail.com / pazenrique@correo.ugr.es / luispaz@comunidad.unam.mx

ABSTRACT

Introduction: Open science has gained increasing relevance in contemporary scientific communication due to its emphasis on transparency, accessibility, and social engagement. The growing adoption of open access policies and collaborative research practices has transformed knowledge production, making it necessary to analyze its thematic evolution, actors, and structural dynamics.

Objective: To analyze the scientific production on open science through a bibliometric approach, identifying publication trends, leading contributors, thematic structures, and emerging research lines.

Methods: An observational, descriptive, and bibliometric study was conducted. Data were extracted, normalized, and analyzed using RStudio and the Bibliometrix package. Indicators related to publication trends, citations, authorship, countries, collaboration networks, keywords, and thematic evolution were examined.

Results: The findings reveal a sustained increase in scientific production, driven mainly by international collaboration and open science policies. The United States, the United Kingdom, Germany, and Spain stand out as the most productive countries. Citizen science, open data, and open access emerged as central themes, while technological applications and responses to global challenges, such as climate change and COVID-19, appear as emerging topics. Thematic mapping shows a dual structure that combines normative, technological and applied-participatory axes, highlighting strong links between transparency, data accessibility, and social engagement.

Conclusions: Open science exhibits a growing interdisciplinary and international character. Its development shows strong potential to guide future research agendas, strengthen collaborative networks, and promote methodological innovation and social impact within the scientific ecosystem.

Keywords: open science; citizen science; open data; open access; bibliometrics

RESUMEN

Introducción: La ciencia abierta ha adquirido una relevancia creciente en la comunicación científica contemporánea debido a su énfasis en la transparencia, la accesibilidad y la participación social. La expansión de las políticas de acceso abierto y de las prácticas colaborativas ha transformado la producción del conocimiento, lo que hace necesario analizar su evolución temática, actores y dinámicas estructurales.

Objetivo: Analizar la producción científica sobre ciencia abierta mediante un enfoque bibliométrico, identificando tendencias de publicación, principales contribuyentes, estructuras temáticas y líneas emergentes de investigación.

Métodos: Se realizó un estudio observacional, descriptivo y bibliométrico. Los datos fueron extraídos, normalizados y analizados mediante *RStudio* y el paquete *Bibliometrix*. Se examinaron indicadores relacionados con tendencias de publicación, citas, autoría, países, colaboración, palabras clave y evolución temática.

Resultados: Los resultados evidencian un crecimiento sostenido de la producción científica, impulsado principalmente por la colaboración internacional y las políticas de ciencia abierta. Estados Unidos, Reino Unido, Alemania y España destacan como los principales países productores. La ciencia ciudadana, los datos abiertos y el acceso abierto se consolidan como temas centrales, mientras que las aplicaciones tecnológicas y las respuestas a desafíos globales, como el cambio climático y la COVID-19, emergen como líneas en expansión. El mapeo temático revela una estructura dual que integra los ejes normativo-tecnológico y aplicado-participativo, destacando la interrelación entre transparencia, accesibilidad a los datos y participación social.

Conclusiones: La ciencia abierta presenta un marcado carácter interdisciplinario e internacional, con un alto potencial para orientar futuras agendas de investigación, fortalecer redes de colaboración y promover la innovación metodológica y el impacto social.

Palabras clave: ciencia abierta; ciencia ciudadana; datos abiertos; acceso abierto; bibliometría

Received: January 21st of 2026
Accepted: February 21st of 2026
Published: February 22nd of 2026
Editor in charge: MSc. Jesús Salvador Hernández Romero
ARTICLE WITH INTERNATIONAL COLABORATION

Introduction

Thematic studies in scientific production hold great importance both individually and globally. This type of analysis has gained popularity in recent years and has been prioritized by editorial policies.[1] They allow for the identification of the most relevant research areas, being useful for institutions and researchers to direct their efforts towards topics that have greater scientific and social impact.

By conducting thematic studies, it is possible to detect areas of science that are underexplored or that present knowledge gaps. These gaps can represent opportunities for the formulation of new hypotheses, the development of new experiments, and the generation of new knowledge. They also serve as a tool to assess the quality of scientific production in a given area.[2] Thus, analyzing the number of publications, the citation of works, and other indicators, it is possible to identify the most relevant research and the most influential scientists in a specific field.

Thematic studies can also foster collaboration among researchers working on similar topics. By having an overview of the advances in a particular area, it is possible to establish links between scientists who share common interests and to promote interdisciplinary collaboration. Likewise, they also provide relevant information for decision-making in scientific policy and resource allocation. A thorough understanding of high-potential research areas and impactful topics enables decision-makers to optimize resource allocation and strategically advance priority research initiative.

Furthermore, prospective analyses enable the forecasting and evaluation of potential future scenarios related to a specific topic or phenomenon. These studies provide an anticipated view of possible changes and trends that may arise in the future. This benefits organizations by allowing them to make strategic decisions based on the information gathered and analyzed, thus preparing for potential challenges and capitalizing on future opportunities.

Through prospective studies, emerging trends in different fields can be identified and analyzed.(3) This enables the detection of business or development opportunities in specific areas, anticipating change and enhancing growth. Additionally, they help to identify and assess potential risks that may arise in the future, thus allowing for preventive measures to be taken to minimize their impact. Moreover, by having a clearer vision of what may happen, uncertainty can be reduced and appropriate contingency strategies can be established.

Prospective studies allow for the development of long-term plans and strategies, both individually and organizationally. This provides a clear vision of the objectives to be achieved in the future and how to achieve them, facilitating data-driven and reliable information-based decision-making. An analysis of this type fosters innovation and the adaptability of organizations and individuals.(4) This is particularly important in a globalized and highly competitive environment, where the ability to anticipate and adapt quickly is key to success.

Open science as a perspective and research approach has been expanding in recent years. It refers to the practice of freely and accessibly sharing the results of scientific research, as well as the data, tools, and methodologies used in that process. This implies eliminating the traditional barriers that exist in scientific communication and making the results available to everyone, without restrictions.

This approach is based on the principles of open access, open data, open peer review, and open collaboration. These principles seek to eliminate the traditional barriers that have limited access to research. One of the most notable benefits of open science is the democratization of knowledge. By allowing free and open access to research results, the gap between those who have the resources to access scientific publications and those who do not is reduced. This is especially important in institutions with fewer resources, where access to quality scientific information may be limited.

One of the most prominent aspects of open science is open access publishing, which implies that scientific articles are published in online journals or repositories free of charge and without access restrictions. This allows anyone, regardless of their geographical location or academic institution, to access scientific information and use it to advance their own research. Additionally, this perspective fosters transparency in the research process.(5) This implies that the data and materials used in the study, as well as the protocols and methodologies employed, are shared to promote the reproducibility of research results. Other categories associated with this movement are open access, open data, citizen science, and participatory science.

Open science also promotes collaboration. Researchers can share their ideas, findings, and experiences through online platforms, which fosters knowledge exchange and the generation of new ideas. This collaboration can lead to the creation of multidisciplinary projects and the achievement of more robust and effective results.

Consequently, it has the potential to drive innovation and technological development. Open science allows "freely sharing research results, facilitating the creation of new technologies and applications that can have a significant impact on society.[6]

The involvement of different actors such as companies, non-governmental organizations, and independent professionals facilitates the acquisition of a greater diversity of perspectives and innovative solutions. However, some more traditional researchers have reservations about sharing their data and results, whether due to ethical concerns or academic competition. Furthermore, it is also necessary to establish standards and protocols for data management and curation, as well as incentives and recognition for those who participate in open science practices.

This revolutionary approach seeks to expand access and participation in the production of scientific knowledge. With the increasing demand for transparency and accountability in research, the need to implement open science in institutions with research activity has become increasingly evident. The goal is to ensure research reproducibility by making data and the research process publicly available. This contributes to reducing scientific fraud and improving the quality of discoveries.

The implementation of open science requires significant changes in research culture and institutional policies.[7,8] It is necessary to promote the adoption of open licenses such as Creative Commons, which allow for the sharing and reuse of research. Likewise, it is essential to develop technological infrastructures that facilitate access and management of open data. Incentive and recognition policies that value the practice of open science must be established.

Prospective studies allow for the anticipation and analysis of future trends in the scientific production of open science. This is essential to adapt to the changes and challenges that may arise in the field. Furthermore, such anticipation will allow for the identification of opportunities and the development of strategies to promote open science effectively. Through prospective and thematic analysis, the existing scientific production of open science can be evaluated. This involves analyzing the quality and efficiency indicators in open scientific production, such as data reuse, transparency in results, and collaboration among researchers. Knowing these metrics will allow for the establishment of clear objectives and strategies to improve open scientific production.

It is necessary to conduct a prospective and thematic study on the scientific production of open science to identify the challenges and barriers that researchers face in adopting these practices. These challenges include the lack of incentives, legal and ethical restrictions, as well as the lack of training and awareness about open science. By understanding these obstacles, policies and programs can be implemented to overcome them and promote greater adoption of open practices.

The prospective and thematic study of open science's scientific production allows for the evaluation of the quality and validity of research ascribed to this perspective. This involves analyzing the transparency of the data used, the methodology employed, and the results obtained. A rigorous quality assessment ensures that open scientific production is reliable and can be used by other researchers and society in general. Consequently, this is fundamental to promoting the reproducibility and replicability of scientific results.

Conducting a study of these characteristics will allow for the evaluation of the impact of existing policies and programs in promoting open science. This includes analyzing how promotion policies, such as open access mandates and open data declarations, have influenced open scientific production. This evaluation will allow for the adjustment and improvement of existing policies to promote greater adoption of open practices.

In the digital and information age, open access and transparency have become crucial aspects in the scientific community.(9) Open science has gained increasing attention in recent years. Despite the growing interest in open science, there is no evidence in the published scientific literature of thematic and prospective studies of this field in its entirety.(10) In particular, there is a lack of understanding of the specific research trends and approaches that have been carried out in relation to open science. This field of study is complex and multidisciplinary, which hinders its analysis and understanding.

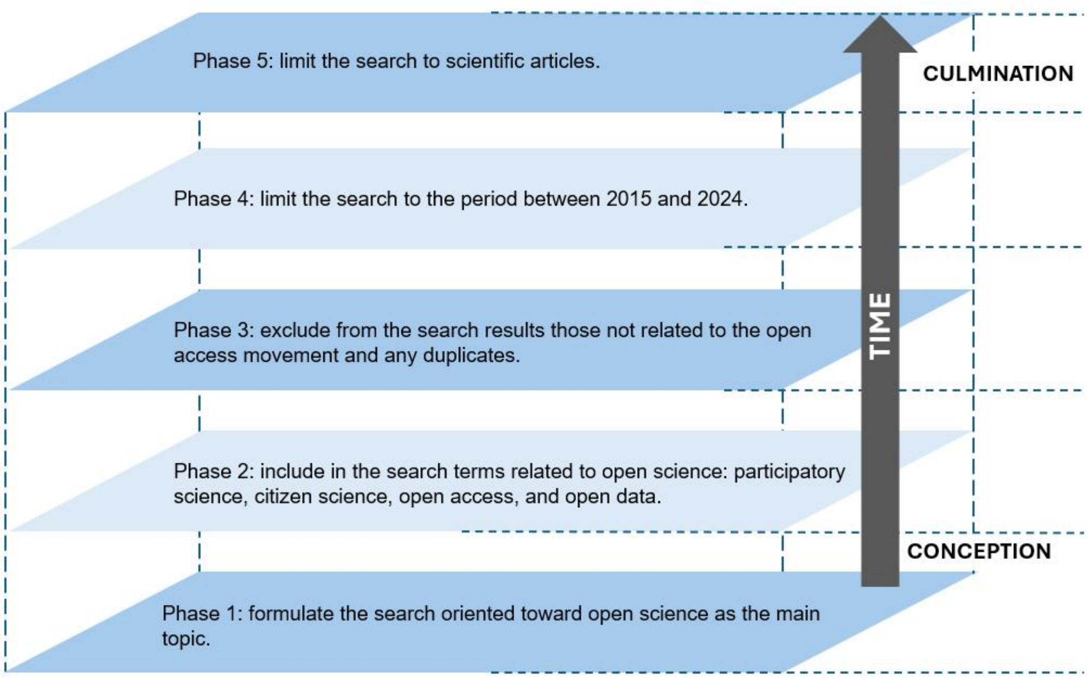
Therefore, it is necessary to conduct research that thematically and prospectively analyzes the scientific field of open science studies published in Web of Science (WOS). This scientific database collects and provides access to a wide range of academic publications across various disciplines and fields. By focusing on studies published in WOS, it will be possible to obtain a comprehensive view of the research conducted in relation to open science.

The thematic analysis will identify the main research áreas(11) in the field of open science. This integrates the implementation of open access practices in scientific publishing, the use of open data in research, and the benefits and challenges of collaboration and public participation in science.(12) Furthermore, the prospective analysis will allow for the identification of emerging trends and potential future research directions in open science. The results of this analysis can be used by researchers, academic institutions, and regulatory bodies to make informed decisions and promote best practices in open science. The present study can contribute to identifying underrepresented or emerging research areas in open science, which could drive new research and advances in the field. The objective of the study was set as follows: to thematically and prospectively analyze the scientific field of open science studies published in WOS in the period between 2015 and 2024.

Methods

An observational, descriptive, bibliometric study was conducted as a specialized method for the analysis of scientific production. The study was also classified as a retrospective longitudinal study, analyzing events that occurred over a past period of time, but with a prospective focus to visualize trend topics within open science. For the present study, the following steps within the bibliometric method were used:

1. Selection of Information Sources: Scientific output on open science and its associated categories, contained in the Web of Science (WOS) database, was selected.
2. Spatial and Temporal Dimensions: The study focused on scientific output from 2015 to 2024, representing the last decade in terms of knowledge dissemination in the field. This time period was selected because the open science movement has proliferated and gained many supporters in recent years.
3. Search Strategy: The search was performed using various iterations, as shown in figure 1.



Source: Own elaboration.

Fig 1. Iteration in the information search.

4. Data Normalization and Processing: The data underwent a process of duplicate removal and keyword standardization. The data was processed using RStudio and the Bibliometrix application. Using these two tools, the Biblioshiny Report was obtained in Excel, which allowed for the processing and in-depth analysis of some of the proposed indicators.
5. Bibliometric indicators used: 1) scientific production by years, 2) number of citations by years, 3) mean total citations per article (Mean TC per Art), 4) mean total citations per year (Mean TC per Year), 5) most productive authors, 6) most productive sources, 7) most productive countries, 8) type of collaboration (single country publication: SCP, multiple country publication: MCP), 9) most cited research, 10) most frequent terms in keywords and abstract, 11) co-occurrence of keywords, 12) trend topics, 13) thematic map.

Results

The scientific output on open science in Web of Science during the 2015-2024 period totals 10,294 documents, published in 2,823 different sources. The average age of the documents is 4.27 years, with an average of 19.34 citations per publication. The temporal dynamics reveal a sustained growth in the number of publications from 2015 (406 articles) until reaching a peak in 2021 with 1,508 records, followed by a stabilization in the last three years with values close to 1,480 annual documents. Despite this absolute increase, the average annual growth rate indicator shows a negative value (-11.04%), which is explained by the recent slowdown and the high concentration of output in recent years.

In terms of content, the corpus includes 12,280 terms in Keywords Plus and 26,848 author keywords, which demonstrates a wide thematic diversity. Individual authorship is a minority, with only 1,307 documents by a single author, while the majority of the research is the result of collaborative efforts, consistent with the interdisciplinary and global nature of open science. This pattern reinforces the orientation towards international networks and the broad circulation of knowledge in the field. The scientific output by year and citation indicators can be seen in table 1.

Table 1. Scientific output by year and citation indicators

Year	Articles	Mean TC per Art	Citation	Mean TC per Year	Citable Years
2015	406	52.30	424	4.75	11

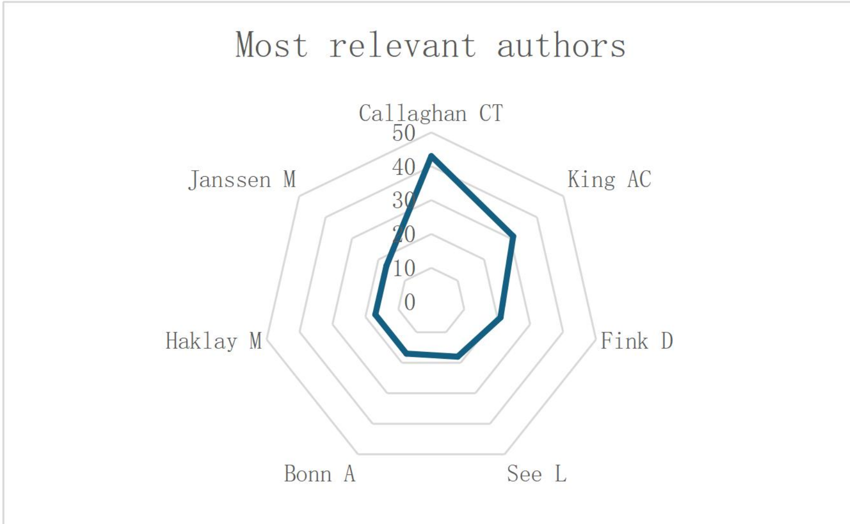
2016	461	54.09	1872	5.41	10
2017	586	42.25	4406	4.69	9
2018	721	36.91	7344	4.61	8
2019	884	32.06	12496	4.58	7
2020	1246	20.37	18017	3.40	6
2021	1508	15.09	27302	3.02	5
2022	1486	10.20	31952	2.55	4
2023	1482	5.51	34013	1.84	3
2024	1488	2.38	37963	1.19	2

Source: Research database.

The evolution of scientific output between 2015 and 2024 shows a pattern of steady growth in the number of articles, accompanied by a progressive decrease in the average citations per document (Mean TC per Art). In 2015, with 406 articles, the average citations per document reached 52.3. In 2024, with 1,488 articles, this value dropped to just 2.38. This behavior is linked to the younger age of the most recently published studies and increased competition in the field.

In terms of total citations, the trend is upward and shows a shift towards more recent years. This is a result of the push to publish research findings in open and free-access formats. While articles from 2015 have accumulated only 424 citations, those from 2024 already have 37,963, despite their short academic life.[13] This indicates growing visibility and interest in the topic of open science, as well as a faster circulation of knowledge over the last decade. However, the average citations per year (MeanTCperYear) systematically decreases from 5.41 in 2016 to 1.19 in 2024, with a decline that is adjusted for the time elapsed.

The analysis of authors reveals that this body of scientific output was produced by 43,473 authors, with an average of 5.43 co-authors per document and an international collaboration rate of 32.32%. This shows that open science fosters collaboration at both the local and global levels, facilitating the exchange of knowledge and the creation of research networks. The most productive authors are shown in figure 2.



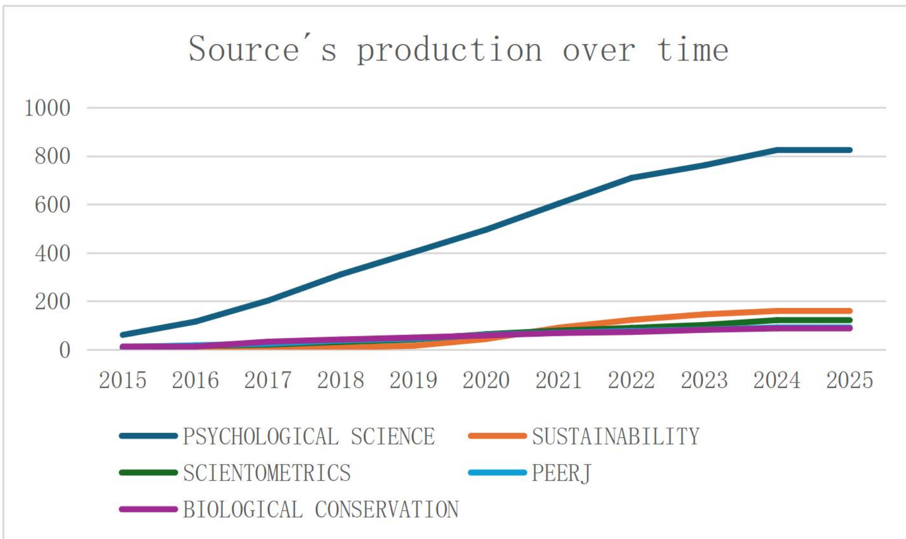
Source: Research database.
Fig. 2. Most productive authors.

Within the analysis of scientific output on open science, Corey T. Callaghan (Callaghan CT) stands out as a highly productive author with a strong focus on environmental sciences and biodiversity conservation. His affiliation with highly relevant international institutions like the University of Florida, the German Centre for Integrative Biodiversity Research, and the Martin Luther University Halle Wittenberg reinforces his global position in research networks. His career is directly linked to sub-fields of open science, such as participatory and citizen science, particularly in biodiversity monitoring and the use of open data for conservation.

Abby C. King (King AC) is recognized as a Highly Cited Researcher. This author combines public health, psychology, and gerontology with participatory methodologies that naturally fit within the open science paradigm. Her leadership in the Our Voice Global Citizen Science Research Initiative is an example of the practical implementation of citizen science to promote community participation in research. The impact of her work, supported by multiple international affiliations such as Stanford University and the University of Haifa, demonstrates how open science transcends disciplinary boundaries and is positioned as a tool to address social and health issues collaboratively and globally.

Daniel Fink (Fink D) presents an interdisciplinary profile, with affiliations ranging from medical institutions like the Shaare Zedek Medical Center to European and North American universities. His output is distributed across areas including environmental sciences, oncology, and biodiversity, suggesting a hybrid approach that combines biomedicine with sustainability issues. This thematic breadth, along with his connection to high-prestige research centers like the University of Zurich and Cornell University, positions him as an author who contributes to the expansion of the open science field into less traditional domains such as medicine and biomedical engineering.[14]

Scientific articles are the predominant document type, with 10,059 records, followed by smaller contributions such as book chapters, data papers, and conference proceedings. The field of open science and its associated categories is a multidisciplinary one. This encourages the publication of studies on this topic in journals from various areas of knowledge. The most productive sources by year are shown in figure 3.



Source: Research database.

Fig. 3. Most productive sources per year.

Psychological Science shows sustained growth in the number of articles related to open science, increasing from 60 in 2015 to 825 in 2025. This trend reflects an editorial policy focused on data and method transparency, as well as the reproducibility of results, which are central elements of the open science movement. Its focus on experimental and cognitive psychology has benefited from initiatives like Open Practices Badges, which facilitates the adoption of best practices in the exchange of data and protocols.

Sustainability has shown a notable upward trajectory, from a marginal presence in 2016 (1 article) to reaching 159 in 2024 and 2025. Its editorial policy encourages the inclusion of citizen science and open data studies. This strategy broadens the geographical scope of contributions and integrates inter- and transdisciplinary approaches, a key aspect for understanding the thematic evolution of open science.

Scientometrics reflects a constant growth in output, increasing from seven articles in 2015 to 122 in 2025. Its editorial policy is centered on the evaluation of science, academic communication, and bibliometrics, making it a natural platform for publishing studies that analyze open science from a meta-scientific perspective.[15] The journal directly contributes to the objective of this study, as its content offers theoretical and methodological frameworks for measuring and tracking the impact of scientific openness policies. Scientific output by country and author correspondence based on collaboration is shown in table 2.

Table 2. Scientific output by country and authors' correspondence based on collaboration.

Country	Articles	Articles %	Citations	SCP	MCP	MCP %
Usa	2576	25.08	68467	2015	561	21.7779503
United Kingdom	914	8.90	32965	553	361	39.4967177
Germany	614	5.97	11376	384	230	37.4592834
Spain	579	5.63	6064	416	163	28.1519862
Canada	485	4.72	11178	301	184	37.9381443
Italy	456	4.44	5847	304	152	33.3333333
Australia	433	4.21	8399	285	148	34.1801386
Netherlands	343	3.34	9614	184	159	46.3556851
China	341	3.32	4963	231	110	32.2580645
France	271	2.63	4305	166	105	38.7453875
Brazil	255	2.48	1405	197	58	22.7450980
Japan	183	1.78	2026	141	42	22.9508197
India	174	1.69	1238	142	32	18.3908046
Switzerland	141	1.37	2648	67	74	52.4822695
Finland	129	1.26	2374	87	42	32.5581395

Source: Research database.

Scientific output in the field of open science between 2015 and 2024 has shown significant collaboration among different countries. The United States leads scientific production with 2,576 articles, accounting for 25.08% of the total, and accumulating 68,467 citations. Its high proportion of single-country publications (SCP) at 78.22%, compared to 21.78% for multiple-country collaborations (MCP), reflects a strong internal capacity to generate research on open science. This pattern indicates a consolidated infrastructure and robust national institutional networks, though with a relatively lower degree of internationalization compared to other leading countries.

The United Kingdom, with 914 articles (8.90%) and 32,965 citations, presents a more balanced profile between SCP (553) and MCP (361), reaching 39.50% MCP. This figure reveals intense international collaboration, consistent with its scientific policy oriented towards openness and integration into global networks. The citations per article support the idea that these alliances contribute to greater visibility and reach of publications. Germany and Spain show similar production volumes (5.97% and 5.63%, respectively) but with different collaboration profiles. Germany has 37.46% MCP, while Spain records 28.15%, which shows a greater German emphasis on international alliances. In both cases, their presence in European open science projects and thematic consortia explains their relevance in the global structure of the field.(16) The most cited investigations are shown in table 3.

Table 3. Most cited investigations

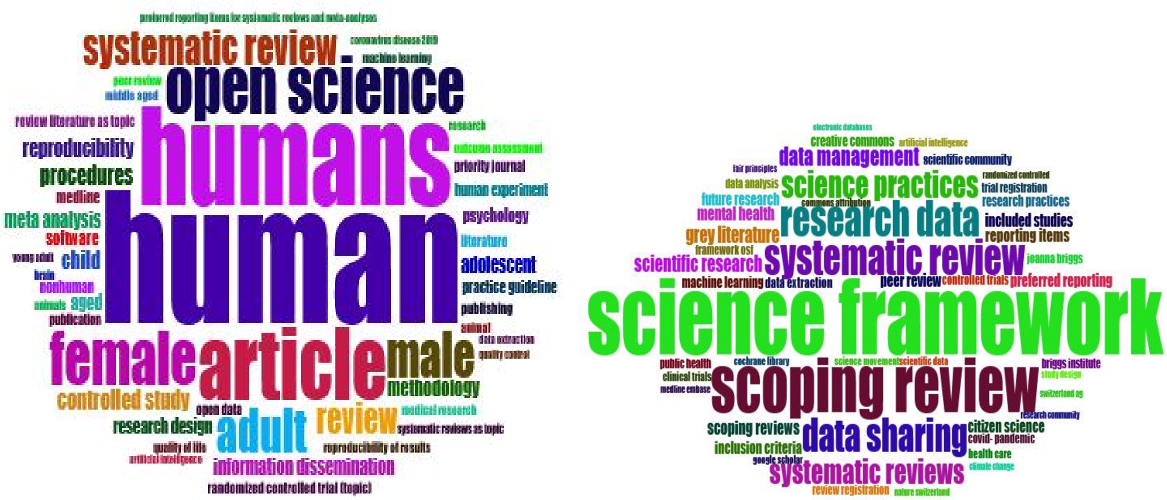
References	Citations
Groom CR, Bruno IJ, Lightfoot MP, Ward SC. The Cambridge Structural Database. Acta Crystallogr B Struct Sci. 2016;72(2):171–79. DOI: https://doi.org/10.1107/S2052520616003954	8379
Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H, et al. PsychoPy2: Experiments in behavior made easy. Behav Res Methods. 2019;51(2):195–203. DOI: https://doi.org/10.3758/s13428-018-01193-y	3090
Nosek BA, Ebersole CR, DeHaven AC, Mellor DT. The preregistration revolution. Proc Natl Acad Sci U S A. 2018;115(11):2600–6. DOI: https://doi.org/10.1073/pnas.1708274114	1080

Source: Research database.

The first study on the Cambridge Structural Database is a fundamental contribution to open science by consolidating one of the most complete and accessible databases for crystallographic structures. Its contribution lies in the standardization and curation of over a million chemical structures, accompanied by high-quality metadata and interoperable analysis tools. This resource not only facilitates the reuse and verification of data but also boosts the reproducibility of research in chemistry, structural biology, and materials science. Furthermore, the work promotes a culture of sharing structural data under open access policies, strengthening global collaboration networks and accelerating scientific innovation across multiple disciplines.(17)

The study on PsychoPy2 represents a significant advance in the field of open science. It provides an accessible and flexible platform for the creation of experiments in psychology and behavioral sciences. PsychoPy2 allows researchers of various experience levels to design, execute, and analyze experiments without the need for extensive programming knowledge, which democratizes access to advanced scientific tools. By facilitating the creation of experimental protocols that can be easily shared and replicated by other researchers, PsychoPy2 contributes to the creation of a more solid and reliable knowledge base with an emphasis on open-source software.[18]

The last study represents a significant advance for open science. It addresses the need to improve transparency and reproducibility in scientific research. One of the main contributions of the work is the promotion of study preregistration, a process that involves documenting and registering the hypotheses, methods, and planned analyses before carrying out the research. This practice helps to mitigate publication bias, data manipulation, and fosters a culture of responsibility among researchers by forcing them to define and communicate their intentions before observing the results. Consequently, a practical approach to improve research quality and a cultural change in the scientific community is proposed.[19] The frequency of terms in keywords and abstracts are shown in figure 4.



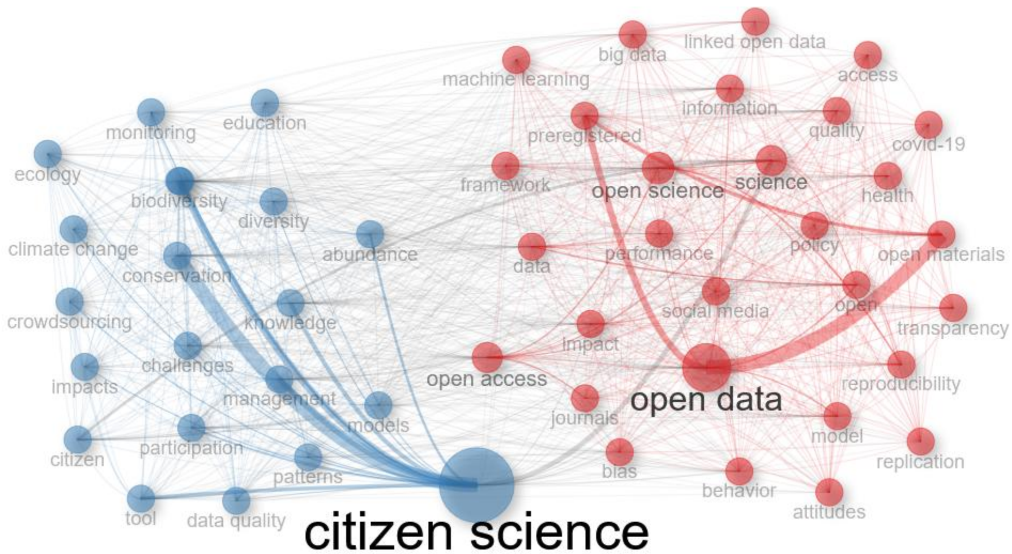
a) b)

Source: Research database.

Fig. 4. Most frequent terms in key words (a) and abstracts (b).

In the keyword analysis, citizen science leads with 3,752 occurrences, followed by open data (2,275) and open access (1,370), highlighting the importance of citizen participation and data openness as central themes in open science output. Other terms like open science (1,347), open materials (572), and linked open data (279) reinforce the presence of open practices at different stages of the research cycle. Applied concepts such as conservation (569), biodiversity (442), and climate change (243) also appear, showing a focus on global environmental challenges. The presence of methodological terms like preregistered (377), reproducibility (243), and transparency (236) confirms the adoption of scientific integrity and quality principles.

In the abstracts, the most frequent term is also citizen science (5,170), but with a greater presence of related expressions like citizen scientists (938) and community science (236), which reflects a narrative emphasis on the active role of non-specialist individuals in the research process. Environmental topics such as climate change (574), water quality (301), and air quality (280) appear strongly, as do terms related to data management and circulation (data sharing, 495; research data, 365; government data, 265). Expressions related to open access and its infrastructure (oa journals, 293; access oa, 353) and the use of emerging technologies (machine learning, 304) are also prominent. The co-occurrence of keywords in scientific output is shown in figure 5.

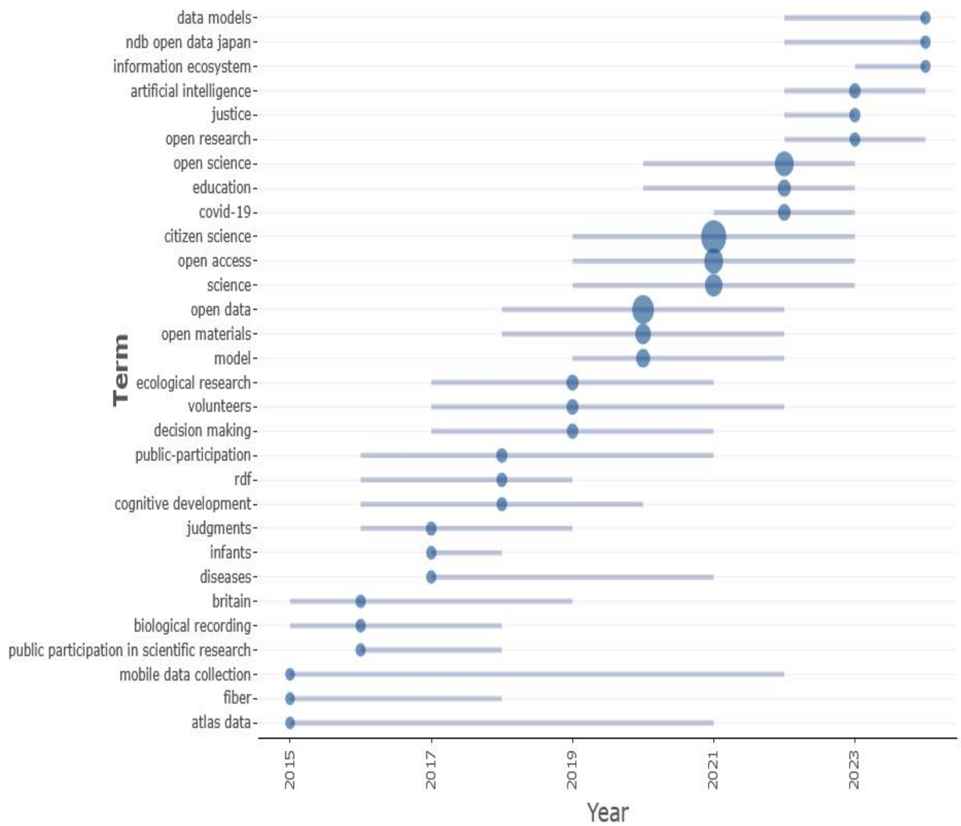


Source: Research database.

Fig. 5. Co-occurrence of keywords.

The analysis of the co-occurrence matrix reveals two distinct thematic clusters. The first group contains core open science concepts such as open data, open access, open science, and science, which have high betweenness centrality and closeness values. This indicates their role as bridge nodes and their conceptual proximity to the other terms. In particular, open data (52.43) and science (28.65) stand out as articulating axes that connect different sub-themes, reinforcing their central role in the field's semantic structure. This thematic core also integrates notions related to transparency, reproducibility, big data, and access policies, thus solidifying a discursive framework focused on the infrastructure and governance of scientific information.

The second cluster concentrates terms linked to citizen science. This cluster has the highest betweenness centrality (63.62) in the entire network and a very high PageRank (0.1154). This indicates its high relevance and ability to connect with other concepts. Topics related to biodiversity management and conservation, citizen participation, and the collaborative generation of knowledge emerge in this cluster, as well as global environmental challenges like climate change. The presence of terms like knowledge, management, and conservation shows that open science is not limited to data availability but extends to its application in specific socio-environmental contexts. Trending topics are shown in figure 6.



Source: Research database.

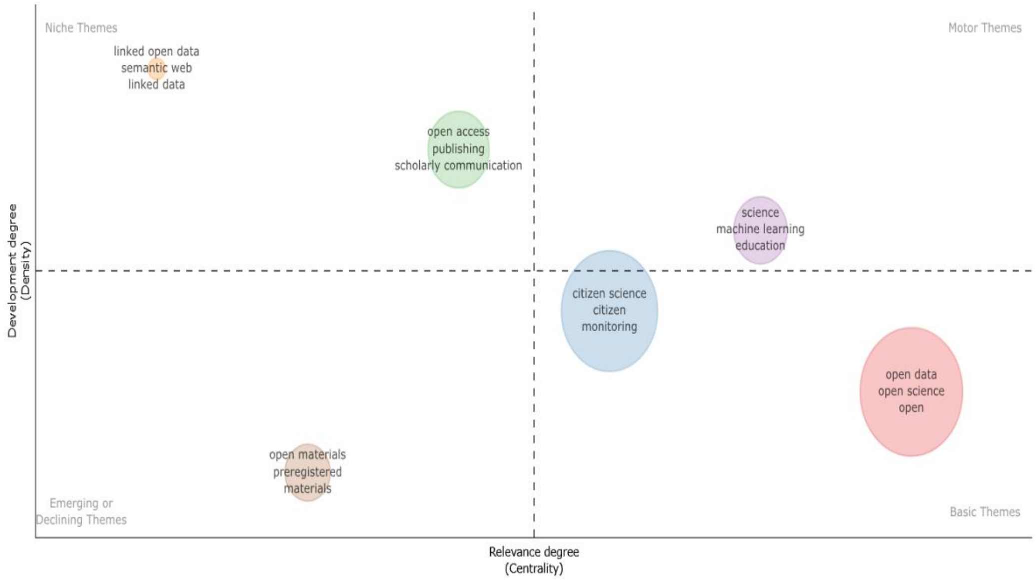
Fig. 6. Trending Topics.

The analysis of trending topics reveals a clear evolution in the field of open science during the 2015-2024 period. The early years show an emphasis on terms related to traditional data collection and management, such as atlas data, fiber, and biological recording, as well as citizen participation in research projects, reflected in public participation in scientific research and public-participation. These terms indicate the initial consolidation of collaborative practices and the recording of biological information, with a focus on local actors and established scientific communities.

From 2018 onwards, there is a shift toward concepts associated with the infrastructure and principles of open science, highlighting open data, open materials, open access, and open science. The increasing frequency of these terms and their appearance in the most recent years of the period (median 2020-2022) indicate that the openness of data, materials, and publications has become a central axis of scientific discourse. This trend reflects the progressive adoption of institutional policies and international mandates that promote information availability and methodological transparency.

In parallel, applied and contextual topics emerge, such as citizen science, ecological research, volunteers, and decision making, which show a combination of citizen participation, ecological research, and the analysis of decision-making processes. The persistence of these terms until 2023 reflects the continuity of collaborative projects and the integration of data from multiple sources, linking open science with concrete social and environmental issues. Likewise, terms like education and cognitive development indicate an expansion into the educational and social spheres, solidifying the transversal nature of open science.

The most recent years show the emergence of topics linked to new technologies and global events, such as covid-19, artificial intelligence, open research, data models, and information ecosystem. The appearance of these concepts between 2021 and 2024 shows that open science is rapidly adapting to emerging contexts, integrating artificial intelligence, the management of large volumes of data, and responses to health crises. This dynamic reflects the field's ability to incorporate innovative tools and its function as a strategic framework for applied and prospective research. The thematic map of scientific output is shown in figure 7.



Source: Own elaboration.
Fig. 7. Thematic Map.

The thematic map analysis shows that the scientific output on open science is organized into six main clusters, each reflecting a specific conceptual axis of the field. The "open data" cluster emerges as the most frequent and central, with high betweenness and PageRank, indicating that it acts as the articulating axis for multiple sub-topics. These include open science, transparency, reproducibility, and covid-19. This pattern highlights the relevance of open data and materials as a structural basis for open research and its integration with access policies and information management practices.

The second most prominent cluster corresponds to citizen science, characterized by the presence of terms associated with public participation, environmental monitoring, biodiversity, and conservation. The centrality of nodes such as citizen, monitoring, and climate change reflects that open science is not only focused on access to information but also on the co-production of knowledge with non-academic actors. This demonstrates the interaction between open data and the implementation of collaborative projects, reinforcing the applied and social dimension of the field.

The open access cluster groups terms related to scientific publishing, academic communication, peer review, and bibliometrics. Its high density shows a consolidated and specialized development, where journals, evaluation metrics, and knowledge dissemination are strategic elements to guarantee the accessibility and visibility of research. This cluster connects with open data and science, showing that information availability and editorial processes are complementary components within open science.

The science and linked open data clusters reflect methodological and technological dimensions. The first includes terms such as machine learning, education, data quality, and management, highlighting the integration of advanced tools and analytical methodologies in the study of open science. The second, with nodes like semantic web and ontology, indicates the construction of digital infrastructures and interoperable data models, reinforcing the connectivity and reuse of scientific information.

Finally, the open materials cluster emphasizes the availability of experimental and methodological resources. This includes preregistration, study materials, and decision making. Its position on the map indicates an emerging development, with high density but lower centrality. Consequently, this axis is in the process of consolidation and may become stronger in future research.

Discussion

The open science movement has proliferated with the development of online platforms and open-access repositories. These have enabled the quick and efficient dissemination of research. They also allow researchers from different parts of the world to access work that would otherwise be restricted. This digital environment has fostered a culture of collaboration and exchange, leading to an increase in publications and the development of new ways to share knowledge.

Furthermore, funding policies and the requirements of funding agencies have also influenced this increase. Many institutions and funding bodies have started to demand that the results of the research they fund be published in open access. Authors are also required to manage their research data and make it available in open data repositories. This has led researchers to adapt their publication practices. This trend promotes transparency and ensures that knowledge generated with public funds is available for the benefit of society as a whole.

Finally, the COVID-19 pandemic accelerated the adoption of open science by highlighting the urgent need to share information and data quickly and accessibly. During this period, there was a significant increase in the publication of preprints and in interdisciplinary collaboration. This underscores the importance of a more open and collaborative approach to scientific research. This context has created momentum that will continue to influence scientific output in the years to come.

The Citable Years indicator decreases from 11 in 2015 to only 2 in 2024, which is an expected effect due to the shorter time window for receiving citations. The longer scientific output is exposed, the more likely it is to be cited. This explains the low number of citations in the last two years, because, it takes two years for a published document to start receiving its first citations. However, the sharp drop in the average citations per document, even in intermediate periods such as 2019 (32.06) or 2020 (20.37), suggests that recent output faces an environment of greater literature saturation. This phenomenon is consistent with the increase in articles and the dispersion of the scientific community's attention due to access to a greater diversity of sources.

The inflection point in average impact occurred between 2019 and 2020, when the Mean TC per Art dropped from 32.06 to 20.37, while the output increased from 884 to 1,246 articles. This phenomenon is common in emerging fields, where the accumulation of literature can make newer publications less visible. Likewise, it can be related to an increase in publications motivated by open science policies and the greater availability of data and collaborative resources. It also coincides with an increase in local or applied studies, which tend to have less international citation projection. The COVID-19 pandemic accelerated this trend by concentrating efforts and citations on specific areas, temporarily relegating other sub-topics.

The presence of the authors Corey T. Callaghan (Callaghan CT), Abby C. King (King AC), Daniel Fink (Fink D); illustrates the disciplinary and geographical diversity of the most productive researchers in the field. Their careers reflect that open science is not limited to a single domain but acts as a transversal framework that spans from ecosystem conservation to public health and precision medicine. This feature strengthens the idea that the expansion of the field requires multilevel collaboration networks and the integration of data from multiple sources and contexts.

Prospectively, the leadership of researchers like Callaghan, King, and Fink helps to consolidate a research agenda that combines high-volume scientific output with sustained international impact. Their multidisciplinary and multi-centric profiles anticipate that future work in open science will continue to diversify and broaden its applications. This aligns with the central hypothesis of this study: open science is configured as a space of convergence for different disciplines, fostering both methodological innovation and broad participation in the generation and use of knowledge.

The results correspond to the most productive thematic categories in WOS. The Information Science & Library Science category tops the list with 1,411 records (13.54%), showing that open science has become a priority topic for disciplines related to information management, repositories, and bibliometrics. Editorial policies in this field often promote open access, system interoperability, and metadata standardization—essential aspects for data visibility and reuse—directly aligning with the objective of this study.

Environmental Sciences and Ecology, with 1,234 (11.84%) and 1,113 (10.68%) records, respectively, demonstrate the strong foothold of open science in the study of natural systems. These areas have massively incorporated the use of open data, remote sensors, and participatory monitoring, allowing for the integration of international observation networks. The connection with citizen science and resource conservation reinforces their role as strategic fields in the thematic expansion of the movement.

Psychology Multidisciplinary (1,073; 10.29%) stands out for its early adoption of transparency practices, such as preregistration of studies and the publication of datasets. The Biodiversity Conservation category (721; 6.92%) reaffirms the importance of open science in the management and protection of ecosystems, while Computer Science Information Systems (597; 5.73%) shows the key role of digital platforms and technologies in data infrastructure and technical support for scientific openness.

Canada, Italy, and Australia, with contributions between 4.72% and 4.21%, stand out for a significant balance between SCP and MCP, with MCP percentages over 33%. This pattern reflects scientific systems that are open to cooperation and have editorial policies favorable to open access, which reinforces their role as strategic nodes in international networks. The case of the Netherlands is particularly notable: with only 3.34% of the total production, it achieves 46.36% MCP, demonstrating a highly internationalized scientific model.

In the context of countries with a lower absolute volume of production, Switzerland stands out with 52.48% MCP, the highest value in the group, confirming its strong orientation towards transnational collaborative research. Meanwhile, Japan, India, and Brazil show lower MCP percentages (22-23%), which indicates production structures more focused on national or regional networks. Finland, with 1.25% of the production and 32.56% MCP, maintains a balanced profile and is aligned with European policies on open data and participatory science.

In the context of the European Union (EU), several countries with high production and collaboration in open science directly reflect the impact of policies like Plan S, promoted by cOAlition S, which requires that publicly funded research be published in immediate open access. Countries such as the United Kingdom, the Netherlands, Germany, France, Italy, and Finland have adopted or adapted their national policies to align with this framework. Although the UK is not currently an EU member, it pioneered open access mandates through UK Research and Innovation (UKRI) and maintains a high percentage of international collaboration (39.50% MCP), consistent with the philosophy of openness and global knowledge circulation.

The Netherlands is an exemplary case: with 46.36% MCP, it demonstrates a strongly internationalized scientific system, supported by transformative agreements negotiated by VSNU (now Universities of the Netherlands) that guarantee full open access in hybrid and gold journals. Germany and France also show high levels of MCP (37.46% and 38.75%, respectively), benefiting from national policies that incentivize open access publishing and integration into European infrastructures like OpenAIRE. Italy and Finland have implemented national plans that promote institutional repositories and compliance with open access mandates for publicly funded research. These policies have favored the international visibility of their publications and participation in collaborative European projects.

This is proportional to institutional-level scientific output. The University of California System leads the list with 467 publications, equivalent to 4.48% of the total, reflecting the strength of an institutional network that integrates multiple campuses with active open access policies and robust institutional repositories. This university system has been a pioneer in transformative agreements with publishers and in the adoption of mandates for depositing data and articles, which favor the high visibility and impact of its research in open science, in line with the objective of this study.

The University of London, with 251 records (2.41%), brings together various prestigious academic institutions, including University College London and King's College London. Both have explicit open access policies and participate in European and international open science projects. Their contribution shows how university federations can consolidate relevant scientific production and foster the adoption of international standards for open data and methodological transparency.

The Centre National de la Recherche Scientifique (CNRS), with 224 publications (2.15%), is a key player in open science in France and Europe. The CNRS has implemented a specific roadmap for open science (Plan national pour la science ouverte), which includes the mandatory deposit of publications and data in open access repositories. Its output reflects a combination of basic and applied research, reinforcing this study's perspective on the thematic diversification of open science.

Harvard University and the State University System of Florida, with 181 (1.74%) and 177 (1.70%) publications respectively, represent models of high productivity with their own open access dissemination strategies. Harvard has been a benchmark in implementing institutional open access mandates since 2008, while the Florida university system has focused on integrating open data into environmental and biodiversity research.

The University of Oxford, with 156 publications (1.50%), combines academic tradition with leadership in European and global projects focused on open science, particularly in public health, social sciences, and digital humanities. Finally, the Consejo Superior de Investigaciones Científicas (CSIC) in Spain, with 151 records (1.45%), has adopted an institutional open access policy and actively participates in the OpenAIRE network, which has favored its international projection. The role of these institutions confirms that significant production in open science is closely linked to institutional and national policies that prioritize openness, interoperability, and collaboration—central elements for the prospective analysis proposed in this study.

The comparison between most frequent terms in key words and abstracts shows overlap in the main thematic areas: citizen science, open data, open access, and environmental sustainability, though with differences in the level of detail. Keywords tend to reflect more standardized and strategic concepts, designed for indexing and retrieval in databases, whereas abstracts include a greater variety of expressions, specific actors (e.g., citizen scientists), applied contexts, and technological mentions. Abstracts also incorporate terms related to specific events like COVID-pandemic (225), which are absent from the keywords. According to Price-Whelan et al. (2018), this textual space is used to contextualize research in specific temporal and social scenarios.[20]

From a prospective standpoint, the comparison between the two lists reveals that terms from abstracts capture a broader thematic diversity and depth. They incorporate contextual nuances, involved actors, emerging issues, and concrete applications. This shows that analyzing abstracts allows for a more precise anticipation of future development lines in open science, by highlighting consolidated areas, trends, and intersections with fields like public health, climate change, or machine learning. These areas could gain greater relevance in the global research agenda.

The comparison between clusters on co-occurrence of keywords shows that while the first is oriented toward the conceptual and normative foundations of open science, the second prioritizes its implementation in participatory projects and specific issues. This duality reflects a thematic structure that combines normative and technological frameworks with citizen collaboration practices, broadening the scope and social impact of research. Additionally, the citizen science node acts as a bridge between the two spheres, connecting openness policies with on-the-ground actions. This illustrates the interdependence between theory and practice.

In terms of network metrics, the relatively homogeneous closeness values across most nodes indicate that the structure is not dominated by a single, centralized concept. It is characterized by multiple points of access to knowledge. However, the weight of open data and citizen science in the mediation measures reveals that these terms have a special ability to link disparate areas. This positions them as catalysts for interdisciplinary convergence. This configuration favors the circulation of ideas and the integration of approaches, which are key aspects for the future expansion of open science.

The analysis of trending topics reveals a clear evolution in open science research between 2015 and 2024. In the early years, the field was characterized by an emphasis on data collection and management, with recurring terms related to biological records, citizen participation, and local scientific practices, reflecting the initial consolidation of collaborative research models. Overall, the evolution of topics demonstrates the increasing maturity, interdisciplinarity, and strategic relevance of open science as a framework for collaborative, data-driven, and socially oriented research. The thematic analysis shows that scientific production on open science is structured around six interconnected clusters that define the conceptual organization of the field. Open data emerges as the central axis, linking transparency, reproducibility, COVID-19 research, and broader open science practices, and highlighting its key role in research accessibility and information management. Citizen science represents the second most relevant cluster, emphasizing public participation, environmental monitoring, and climate-related studies. Overall, the thematic structure reflects an increasingly interconnected, data-driven, and socially oriented model of open science, characterized by strong links between openness, collaboration, and technological innovation.

Conclusions

Open science is a necessary and transformative approach for institutions engaged in research. By promoting open access, open data, open peer review, and open collaboration, it democratizes knowledge, increases transparency and reproducibility, and encourages collaboration among various actors. Implementing open science requires significant changes in institutional culture and policies. This perspective fosters collaboration, networking, and the promotion of innovation.

Overall, the data reveals an expanding field with a high capacity for production but facing challenges in consolidating its impact per document. The quantitative growth is accompanied by a change in citation patterns, with a faster initial accumulation of citations but a lower average adjusted per year and per article. This highlights the need for more effective positioning and dissemination strategies to maintain the visibility and relevance of contributions in an increasingly competitive and diverse scientific communication ecosystem. The upward trend in publishing, along with an international collaborative approach, underscores the current importance of open science. It also reinforces its potential to transform how knowledge is created and shared in academy.

The period from 2015 to 2024 has seen a significant increase in scientific output related to open science, characterized by greater collaboration between institutions from different countries. This phenomenon reflects the globalization of science and the evolution of the open access philosophy, where the democratization of knowledge becomes a central goal. The trend suggests that interest in open science will continue to grow, promoting a more collaborative and accessible environment for global research.

The analysis of the most frequent terminology in keywords and abstracts confirms that open science is structured around three strategic axes: citizen participation, data openness and management, and responding to global challenges, especially environmental ones. Keywords show a stable conceptual core oriented toward bibliographic retrieval, while abstracts broaden the scope to include actors, applications, and situational contexts. This complementarity reinforces the vision of open science as a transversal, dynamic, and socially connected field, which supports its future projection as a reference framework for the production and circulation of scientific knowledge.

The co-occurrence of keywords in scientific output within the open science field reveals an interconnected network of approaches and methodologies. These address the quality and validity of research and its social relevance. The co-occurrence of keywords shows that open science is structured around a normative-technological axis (open data, open science, open access) and an applied-participatory axis (citizen science, conservation, climate change). This configuration consolidates the principles of access and transparency and projects a scenario where citizen participation and the application of knowledge to global challenges will play an increasingly decisive role.

The evolution of trending topics shows that open science has transitioned from initial practices of citizen participation and data collection to a comprehensive approach that combines openness, collaboration, and technological application. The emerging terms in recent years demonstrate the field's ability to adapt to contemporary challenges and project new lines of research, thus solidifying its relevance both methodologically and socially. This prospective pattern underscores that open science will continue to expand into interdisciplinary, technological, and strategic domains.

The thematic map shows that open science is structured around the axes of data access and transparency, citizen participation, scientific publishing, and advanced methodologies. The centrality of open data and citizen science shows that openness and collaboration are the main drivers of the field, while the emerging technology and materials clusters project areas for future consolidation. These results can guide research toward integrating access policies, social participation, and the development of digital infrastructures to strengthen global open science.

Funding

No funding was received for the conduct of this study.

Data set

This article has an associated dataset published in the Scientific Editing Seminar community hosted by Zenodo. The DOI is: https://doi.org/10.5281/zenodo.17454437

Bibliographic references

1. Jiménez P. Tendiendo puentes entre la divulgación y la comunicación de las ciencias a nivel mundial. Estoa. 2023;12(24):6. DOI: https://doi.org/10.18537/est.v012.n024.e01

2. Yuni JA. Dilemas de la investigación como práctica institucional y experiencia biográfica. Praxis Educ. 2021;25(1):40–60. DOI: https://doi.org/10.19137/praxiseducativa-2021-250105

3. Rodríguez S, Trejo C, Hernández EL. Análisis prospectivo del diseño gráfico en México. Zincografía. 2020;4(8):109–20. DOI: https://doi.org/10.32870/zcr.v0i8.78

4. Moreno GA, Gutiérrez RE. Estudio prospectivo de la tecnología en la educación superior en Colombia al 2050. Rev Univ Empresa. 2020;22(38):160–82. DOI: https://doi.org/10.12804/revistas.urosario.edu.co/empresa7a.7583

5. Fernández M. ¿Ciencia abierta para intereses privados? La lógica de la ciencia abierta y la comercialización de la investigación. Rev Econ Inst. 2022;24(47):179–201. DOI: https://doi.org/10.18601/01245996.v24n47.08

6. Prieto D. Ciencia abierta: desafíos y oportunidades para Uruguay y el Sur Global. Informatio. 2022;27(1):253–83. DOI: https://doi.org/10.35643/info.27.1.5

7. Seroubian M. Acceso abierto y ciencia abierta. Experiencia desde la gestión del repositorio institucional COLIBRI de la Universidad de la República. Informatio. 2022;27(1):284–308. DOI: https://doi.org/10.35643/info.27.1.6

8. Viñar D. Software libre, ciencia libre: explorando algunas enseñanzas que el software libre puede aportar a la ciencia abierta. Informatio. 2022;27(1):336–71. DOI: https://doi.org/10.35643/info.27.1.9

9. Rodés V, Motz R. Construcciones interdisciplinarias en educación y ciencia abiertas. Informatio. 2022;27(1):142–66. DOI: https://doi.org/10.35643/info.27.1.10

10. De Giusti MR. Ciencia abierta: el corazón del problema. Informatio. 2022;27(1):309–35. DOI: https://doi.org/10.35643/info.27.1.3

11. Imbernó AL, Souto L. Innovación y economía circular, un binomio perfecto. Econ Desarro. 2023 [access 01/10/2026];167(2):1–21. Disponible en: https://revistas.uh.cu/econdesarrollo/article/view/7562

12. Rojas L. Robustez de los índices de ajuste del análisis factorial confirmatorio a los valores extremos. Rev Mat Teor Apl. 2020;27(2):383–404. DOI: https://doi.org/10.15517/rmta.v27i2.33677

13. Hosseini M, Horbach SP, Holmes K, Ross-Hellauer T. Open science at the generative AI turn: An exploratory analysis of challenges and opportunities. Quant Sci Stud. 2025;6:22–45. DOI: https://doi.org/10.1162/qss_a_00337

14. Viancos P, Valdés M, Alarcón N, Abello J. Producción científica en el ámbito de la gobernanza universitaria. J Acad. 2024;11(2):175–94. DOI: https://doi.org/10.47058/joa11.10

15. Yuan HY, Cao Y, Kamra A, Duine RA, Yan P. Quantum magnonics: When magnon spintronics meets quantum information science. Phys Rep. 2022;965:1–74. DOI: https://doi.org/10.1016/j.physrep.2022.03.002

16. Bautista J, Badillo R, Lobato C. Ciencia abierta como una nueva forma de hacer investigación. Dialogos Educ. 2023;14(26):1–16. DOI: https://doi.org/10.32870/dse.v0i26.1230

17. Groom CR, Bruno IJ, Lightfoot MP, Ward SC. The Cambridge Structural Database. Acta Crystallogr B Struct Sci. 2016;72(2):171–9. DOI: https://doi.org/10.1107/S2052520616003954

18. Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H, et al. PsychoPy2: Experiments in behavior made easy. Behav Res Methods. 2019;51(2):195–203. DOI: https://doi.org/10.3758/s13428-018-01193-y

19. Nosek BA, Ebersole CR, DeHaven AC, Mellor DT. The preregistration revolution. Proc Natl Acad Sci U S A. 2018;115(11):2600–6. DOI: https://doi.org/10.1073/pnas.1708274114

20. Price-Whelan AM, Sipöcz BM, Günther HM, Lim PL, Crawford SM, Conseil S, et al. The Astropy Prfoject: Buildingan Open-science Project and status of the v2.0 Core Package. Astronomial Journal.2028;156(3):e123. DOI: https://doi.org/10.3847/1538-3881/aabc4f

Conflict of interest

The authors declare that they have no conflict of interest.

Author contributions

Conceptualization: Luis Ernesto Paz Enrique, Wileidys Artigas Morales, Eduardo Alejandro Hernández Alfonso.

Data curation: Luis Ernesto Paz Enrique, Wileidys Artigas Morales, Eduardo Alejandro Hernández Alfonso.

Investigation: Luis Ernesto Paz Enrique, Wileidys Artigas Morales, Eduardo Alejandro Hernández Alfonso.

Methodology: Luis Ernesto Paz Enrique.

Formal analysis: Luis Ernesto Paz Enrique, Wileidys Artigas Morales, Eduardo Alejandro Hernández Alfonso.

Writing – original draft: Luis Ernesto Paz Enrique, Wileidys Artigas Morales.

Writing – review and editing: Eduardo Alejandro Hernández Alfonso.