West Science Information System and Technology
Vol. 4, No. 01, April 2026, pp. 123~134

Database Management Systems (2000–2026): A Scopus-Based Bibliometric Review of Research Streams and Key Contributor

Loso Judijanto
IPOSS Jakarta, Indonesia

Article Info

Article history:

Received Apr, 2026
Revised Apr, 2026
Accepted Apr, 2026

Keywords:

Database Management Systems (DBMS)
Bibliometric Analysis
Scopus
VOSviewer
Key Contributor

ABSTRACT

The current research intends to analyze the knowledge structure and research trends of the DBMS by using a bibliometric approach to study articles published in the Scopus database between 2000 and 2026. This will be achieved using the VOSviewer as the main software for analyzing authors' networks, citations and keywords' co-occurrences. As expected, the findings reveal that research within the DBMS area has become increasingly interdisciplinary with notable influences from fields such as health care and bioinformatics and data science. In terms of authorship patterns, it is evident that collaboration within DBMS research happens in clusters. This is an indication of the presence of well-coordinated intra-cluster collaboration and a lack of inter-cluster collaboration. The findings from the citation analysis show that there are key publications within DBMS research that have been highly cited due to their contribution to developing the discipline. Finally, from the keyword co-occurrences, it is clear that the trend within DBMS research has shifted from more traditional subjects to newer areas such as artificial intelligence and machine learning. The findings highlight the growing complexity and integration of DBMS within modern digital ecosystems, while also emphasizing the need for enhanced global collaboration and innovation-driven research. This study provides valuable insights for researchers and practitioners in understanding the trajectory and future directions of DBMS research.

Corresponding Author:

Name: Loso Judijanto
Institution: IPOSS Jakarta, Indonesia
Email: losojudijantobumn@gmail.com

1. INTRODUCTION

The fast pace of development of the latest technologies in the last twenty years has greatly impacted how data can be collected, stored, analyzed, and used in many spheres. The key element of this technological revolution is the sphere of Database Management Systems (DBMS). DBMS refers to all activities related to collecting, processing, and retrieval of structured and unstructured data in an efficient manner. From basic relations' databases to complex cloud-based and distributed systems as well as NoSQL systems, DBMS has always developed to fit the latest demands of the market [1], [2].

The domain of DBMS experienced significant growth in terms of its research scope between 2000 and 2026. Thanks to progress in hardware technologies, network infrastructure, and software engineering, researchers have been able to create efficient

and high-performance DBMS architectures. Additionally, novel trends like cloud computing, IoT, and machine learning have raised new questions about the nature of data storage and processing, which have fueled interest in DBMS-related issues. Distributed databases, data privacy, query optimization, and data governance are among the key themes that have been explored in numerous studies [3], [4].

In this regard, bibliometric analysis has proved to be a helpful tool in understanding the structural and evolutionary aspects of scientific research. The use of bibliometric methods allows scholars to measure vast amounts of literature published on different subjects, detect research tendencies, and trace the collaboration between authors, institutes, and countries. With the help of bibliometric analysis, scientists can receive an objective picture of a subject area that is more extensive than any other conventional review.

The employment of big and credible academic databases, especially Scopus, has contributed immensely to enhancing the credibility of bibliometrics research. Scopus is renowned for its broad inclusion of peer-reviewed literature sources such as journals, conferences, and other scholarly output sources. It provides uniform metadata and citation tracing tools, making it a great database for conducting comprehensive bibliometric research. Unlike other databases, Scopus includes a broader range of subject matters in areas like computer science and information systems, which play a vital role in DBMS studies. Due to the benefits that Scopus has over other databases, many recent studies have made use of its data to study research advancements in different disciplines [5], [6].

It has been established through previous bibliometric studies undertaken in relevant fields that research domains progress through certain stages. These stages include emergence, growth, and expansion. For example, from the analysis of the publications indexed on Scopus in various scientific disciplines, it has been observed that there is a considerable increase in the production of research output together with the emergence of topic clusters and the development of international collaboration networks. Likewise, bibliometric mapping within the realm of information systems and technologies suggests that some important contributors, publications, and topics can be identified through such analysis.

In spite of the existing number of bibliometric studies conducted in relation to Information Systems and associated areas, it should be noted that there is a dearth of thorough investigations regarding the research dynamics of Database Management Systems over a long period of time. Taking into account the significance of this topic in view of today's world of information technologies, it is crucial to conduct an in-depth analysis of the evolution, prominent themes, and main researchers within this sphere. Thus, conducting a bibliometric investigation using Scopus data for 2000 to 2026 will give us an idea about current trends, themes, and contributors.

Despite an abundance of academic research on Database Management Systems (DBMS), there is no systematic review of existing studies that analyzes the whole research landscape, detects main research lines, and uncovers important authors in this research domain over a prolonged period of time. Most of previous studies have examined particular subareas, analyzed the state of research within a limited timeframe, or used small databases that do not allow drawing a complete picture of DBMS development. The lack of bibliometric study, which could be based on an all-encompassing database like Scopus, makes it difficult for scholars to get acquainted with trends in global scientific development, detect collaboration networks, and identify new research directions in the field.

The main objective of this study is to provide an in-depth analysis of the scientific output generated on the topic of Database Management Systems from 2000 to 2026 by employing the Scopus database. In particular, the study will attempt to (1) assess publication trends; (2) delineate key research streams and themes in DBMS; (3) explore collaboration trends amongst researchers from different

organizations across countries, and (4) identify key influential individuals and literature. By conducting a systematic analysis of the subject, this paper hopes to contribute towards gaining insight into the intellectual structure and development of DBMS research.

2. METHODS

This study uses a quantitative bibliometric methodology to examine the development of research in Database Management Systems (DBMS) from 2000 to 2026. The use of bibliometric methods for studying scientific activities and developments is widespread. The methodology provides a means for extracting and analyzing large amounts of publication information in a scientific and reproducible manner. In the current study, publication results, citations, keywords, and collaborative networks are analyzed. Through this analysis, the goal is to discover how research streams have developed and what contributions specific authors, organizations, and countries have made within the DBMS research field.

For this research paper, data was collected from the Scopus database due to the fact that Scopus provides wide coverage to peer-reviewed high-quality journals, especially in the field of computer science and information systems. A search strategy was formulated on the basis of appropriate keywords such as "database management system," "DBMS," "data management," and many others. The search was restricted to articles, conference papers, and reviews between 2000 and 2026 that have been written in English. Following the retrieval of information, it was filtered out to eliminate any duplicates, irrelevant data, and partial data. The final set of data thus obtained was exported to be used in bibliometric analysis.

In order to analyze the information, this study applies the bibliometric mapping and visualization methodology through the use of VOSviewer. This will allow the creation of authorship networks, co-citations, as well as keyword co-occurrences in order to define thematic and clustering structures. At the same time, descriptive statistics will be used for the analysis of bibliographic databases that concern the DBMS field in terms of publication, citation patterns, and geographical distribution of research outputs.

3. RESULTS AND DISCUSSION

3.1 Co-Authorship Analysis

The co-authorship network is used to analyze the collaborations that exist between researchers in the discipline of Database Management Systems (DBMS). Co-authorship analysis serves to identify the network of relationships among researchers, determine influential individuals and institutions, and the degree of collaboration between various countries. This method of analysis enables us to understand how knowledge is created and shared in this research domain.

1. Author-level Visualization

The visual representation of the co-authorship network is shown above for the years 2000–2026 within the area of research called Database Management Systems (DBMS). Authors are considered as nodes while the connections between them can be seen as links, and thus, one can see how scholarly interactions are established and how they are spread across the network. The different colors used to depict various clusters allow us to identify different research groups contributing to DBMS research.

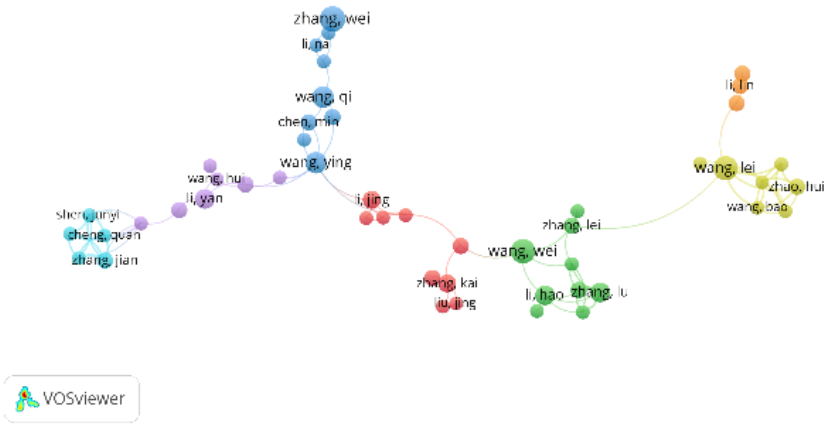

Figure 1. Author-level Visualization

Source: Data Analysis

From the above figure, we can observe a number of well-defined clusters. Each of these clusters denotes a group of authors who work with each other on a regular basis. For example, the blue cluster can be said to form a very centralized cluster because it demonstrates a relatively robust collaboration network among the various individuals, which include Wang Qi and his peers. Likewise, the red and green clusters also show signs of close collaboration, and it can therefore be concluded that they might conduct their work on particular aspects of DBMS systems.

The other side is that the distance among the different clusters implies a lack of interaction among researchers belonging to different groups. For instance, the yellow and purple clusters seem quite distant from each other, and this means that researchers of these colors tend to be active in research communities with a higher degree of specialization. As one can see, although there is much interaction within different groups, interaction among different groups is still not highly prevalent in the DBMS sphere.

2. Institution-level Visualization

The citation network graph shown here highlights the associations among institutions involved in the study of Database Management Systems (DBMS). Here, each node denotes an institution, whereas the connecting lines signify citations that denote the movement of information and influence. The arrangement of nodes gives some indication of how institutions interact via their intellectual output and how some institutions serve as pivotal centers in the citation network.

Figure 2. Institusion-level Visualization
Source: Data Analysis

As shown in the figure above, there are some organizations that hold central positions in the citation graph, reflecting their major impact on the research of DBMS. In particular, universities like University College London and its research institutes play vital roles in the citation graph through forming a dense group that has considerable citations and is often cited by other universities. Likewise, the School of Medical Sciences plays a bridging role in connecting various organizations. It helps to transfer knowledge between the fields.

On the other hand, the existence of smaller and less important nodes represents an institution that has a lower impact on citations or a narrower involvement in research in the DBMS area. This can be seen in the case of the Baylor College of Medicine that is present in another cluster, indicating a more specific contribution to the field or an application of DBMS into some other disciplines. Overall, the network configuration is characterized by a semi-centralized citation distribution model, where only several institutions are the core of the intellectual activity in the area, and others provide their contribution to the research process in a more specific way.

3. Country-level Visualization

The research framework of DBMS is portrayed by Figure 3. The nodes represent countries that have participated in the research area, while the links represent the frequency of occurrence or collaboration among countries in the literature. The bigger the nodes, the more they contributed or produced literature in the area. As seen from the figure, there are several different clusters based on themes or collaboration groups.

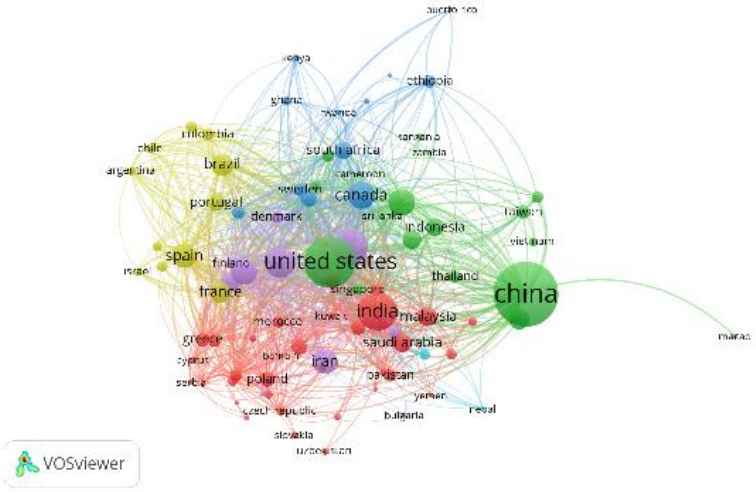

Figure 3. Country-level Visualization
Source: Data Analysis

As is evident from the figure, the US and China are clearly the dominant players in the field of DBMS research, given the fact that both are shown to have a significant number of connections and hence are positioned towards the center of the graph. The US, in particular, can be seen to serve as a prominent hub in terms of linking different clusters, given the fact that the country has been engaged in collaborations with countries from other regions like Europe, Asia, and even Africa. Likewise, China too serves as a prominent producer of research and makes strong links with Asian and developing nations.

Moreover, the clustering shows that there is a trend of collaboration within the region and themes. Countries in Europe such as Spain, France, and Germany are found to be part of a highly clustered group, implying that there is considerable collaboration within the region. At the same time, countries like Canada, Australia, and South Africa serve as bridges between the collaborating developed and developing regions.

3.2 Citation Analysis

Citation analysis is conducted to evaluate the academic impact and intellectual influence of publications within the DBMS domain. This approach identifies highly cited authors and documents that have significantly contributed to the development of the field.

Table 1. The Most Impactful Literatures

Citations	Authors and year	Title
17265	[7]	Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy
13993	[8]	Primer3 on the WWW for general users and for biologist programmers.
13757	[9]	The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data
12890	[10]	Comment: The FAIR Guiding Principles for scientific data management and stewardship
12631	[11]	2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer

Citations	Authors and year	Title
9854	[12]	Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research
9624	[13]	The behaviour change wheel: A new method for characterising and designing behaviour change interventions
8454	[14]	The graph neural network model
6174	[15]	A design science research methodology for information systems research
5407	[16]	Epidemiology of multimorbidity and implications for health care, research, and medical education: A cross-sectional study

Source: Scopus, 2026

The influential literatures identified in Table 1 are distinguished by having a very high number of citations. Such literatures represent the most influential in the dataset based on how many times they have been used in different research contexts. For instance, the most cited literature in the literature set is the work by Wang et al. (2007). This literature emphasizes the importance of computational classification techniques for the study of biological information. Another influential literature in the literature set is by Rozen and Skaletsky (2000). This work underlines the importance of bioinformatics tools developed early in time. On the other hand, there is another group of influential literatures whose high number of citations indicates their relevance to scientific data handling. These literatures include Cerami et al. (2012) and Wilkinson et al. (2016).

3.3 Keyword Co-Occurrence Analysis

Co-occurrence analysis of keywords can be employed to examine the underlying conceptual framework of DBMS studies as well as its thematic development through time. With the help of co-occurrence analysis, one can discover key topics that have attracted attention in DBMS research, upcoming trends in research, and thematic clusters that are emerging in DBMS research.

1. Network Visualization

This figure shows the co-occurrence network of the keyword terms in DBMS studies, which represents the conceptual framework and relationship of themes found in DBMS research. The nodes show the individual keyword terms, while the connections between the keywords show the association between them. The grouping of the keywords into different colors signifies the main themes that exist in the study of DBMS, which enables one to identify major trends and interrelationships of various disciplines in the study.

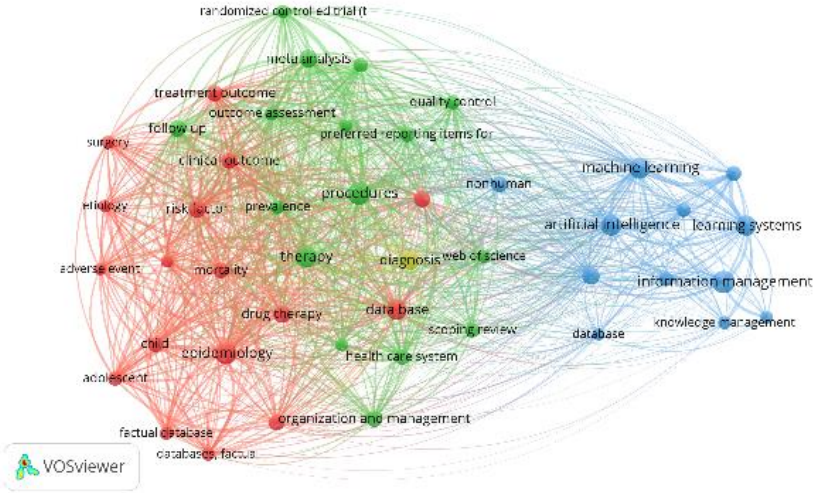

Figure 4. Network Visualization
Source: Data Analysis

Red cluster is assumed to reflect the presence of the dominant theme of research related to medicine and clinical studies, characterized by such key terms as “epidemiology,” “clinical outcome,” “therapy,” and “risk factor.” It implies that the DBMSs are widely utilized in working with medical information. High density and connections in the red cluster demonstrate the maturity and development of the corresponding research field, where database systems serve an essential function.

The green cluster represents those research topics that relate to database management and research approaches, including keywords like “meta-analysis,” “procedure,” “diagnosis,” and “quality control.” The fact that there is an abundance of those keywords in the green cluster means that researchers pay great attention to the application of DBMS in organizing, validating, and analyzing data sets, especially within the scope of experimental science. The keywords “randomized controlled trial” and “reporting items” demonstrate a connection between DBMS and research approaches.

The blue cluster shows a more tech-driven research line, one that emphasizes modern techniques related to computing and information systems. The use of key words like “machine learning,” “artificial intelligence,” “information management,” and “intelligent systems” indicates that DBMS research is gradually moving towards integration with new technology trends. The blue cluster clearly denotes the transformation of traditional database systems into more intelligent systems.

The keywords "database," "data base," and "procedure" function as bridging nodes within the network at its center. It is clear that the central terms have a high connectivity because they highlight the key importance of the DBMS as an integrative tool for different spheres. The terms denote that database management systems can be used beyond one field; hence, their application ranges from healthcare analytics to AI development.

Through the graph, it is evident that there exists an increasingly interdisciplinary field with the research on DBMS involving different sectors such as health care, data science, as well as sophisticated computing technologies. Even though some of the clusters seem to be more specific to certain domains, it can also be noted that the high degree of connectivity among them shows a clear sign of the increasing need for collaboration between them.

2. Overlay Visualization

The keyword co-occurrence visualization overlay shown in Figure 5 gives a historical view into the development of research topics in Database Management Systems (DBMS). While conventional clustering plots do not incorporate information about the time at which publications occurred, this particular visualization incorporates a color coding scheme that represents the average publication date of keywords. This enables one to determine the research topics that are already prevalent as well as those that are new.

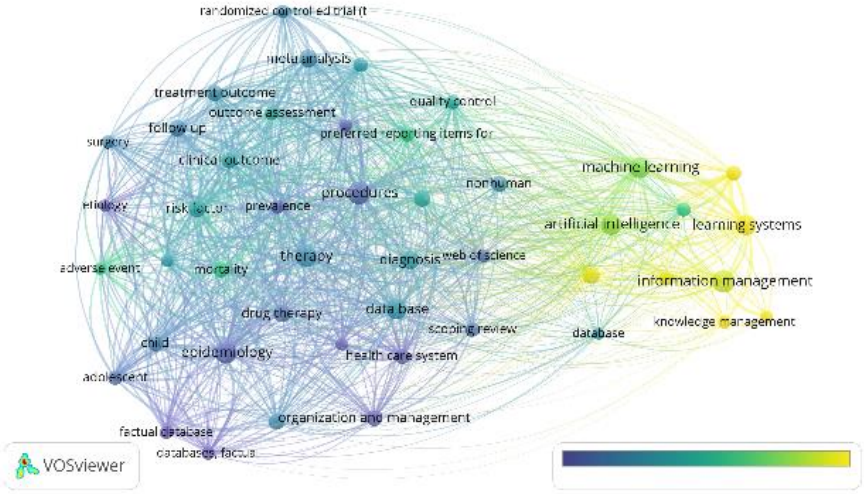

Figure 5. Overlay Visualization

Source: Data Analysis

The figure demonstrates that previous research areas, identified by dark blue and purple shades, are mostly centered around topics that can be regarded as fundamental or topical, including “epidemiology,” “clinical outcome,” “drug therapy,” and “database.” The presence of these keywords indicates that previous research was largely oriented toward practical application in fields where there were large amounts of structured data, for instance, healthcare. Moreover, numerous connections between keywords imply that previous research was mainly concerned with applying database management systems for analyzing and processing structured data.

When the color moves to green, the network highlights the emergence of more systematic and methodical approaches using the keywords such as "procedure," "diagnosis," "quality control," and "meta-analysis." This is the stage when there has been consolidation wherein the research about the DBMS went beyond data management to encompass analysis and data quality and reliability checks. The existence of bridge keywords between this period shows that the role of DBMS was increasingly significant in bridging other fields of research and conducting further research through data analysis.

The latest trends are indicated by yellow, including terms like "machine learning," "artificial intelligence," "information management," and "intelligent systems." These new trends indicate a definite move towards the combination of DBMS with intelligent technologies. The importance of these terms points out that the current research in DBMS emphasizes processing unstructured data, automatization, and intelligent decision-making systems. This development shows the change in the entire field because the database system is now not only the storage place for information but an important element in a sophisticated computer environment.

3. Density Visualization

This density map for co-occurrences of the keywords as shown below is indicative of the strength and density of research being carried out in the DBMS discipline. The brighter the color in the density map (from yellow to green), the higher will be the density of keywords and more active will be the research topic in that region, whereas the darker regions would be those that have less frequent research activity going on.

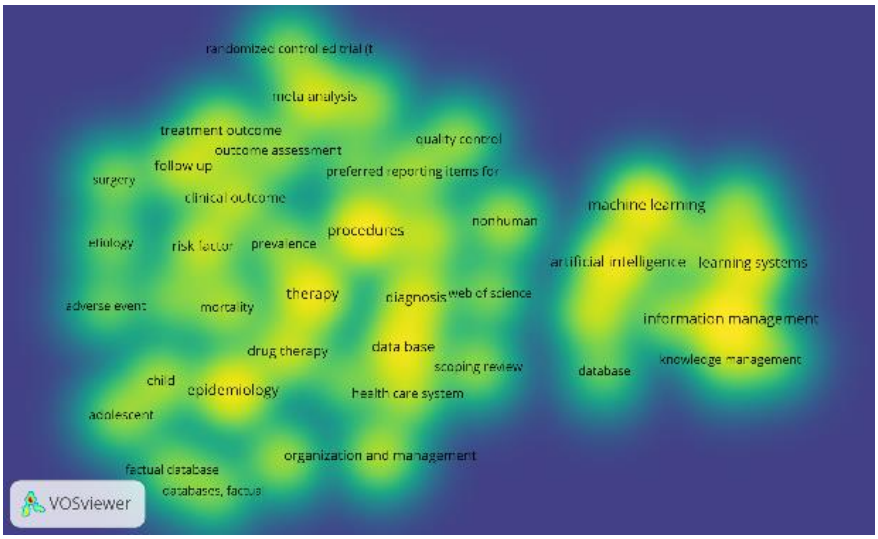

Figure 6. Density Visualization

Source: Data Analysis

This visualization shows many dense zones clustered around phrases like "procedures," "meta-analysis," "quality control," and "diagnosis," implying that these are important research domains. Clearly, these are strongly linked to procedures involved in data management and analytics, especially in scientific and medical fields. The density of these clusters also implies that DBMS is important for data-based decision-making processes where thorough data analysis, validation, and standardization is essential.

Another region characterized by high density is associated with terms like "epidemiology," "therapy," "drug therapy," and "clinical outcome." This demonstrates the significance of the application of health care services within the DBMS field of study. From the clusters, it is clear that database technology plays a significant role in conducting research due to the management of clinical and population data for analysis. The wide dispersion of the above-mentioned terms suggests that there is an established and developing research stream in this area.

On the other hand, innovative themes such as “machine learning,” “artificial intelligence,” “learning systems,” and “information management” are represented by relatively dense but clearly defined clusters. While being less concentrated than the primary medical care themes, the presence of these themes in the map demonstrates that more studies have begun exploring the relationship between DBMS and new computational techniques. Thus, there is a potential for creating new types of database systems that are more intelligent and adaptable.

Discussion

According to the findings from the bibliometric study, DBMS studies have turned out to be an extremely multidisciplinary research field, going well beyond the scope of traditional computer science. The dominance of health care concepts, as shown by keyword clustering and density maps, suggests that the DBMS system has been extensively implemented as the basic backbone technology to handle massive volumes of structured information. This can be attributed to the growing need for evidence-based decision-making practices in both clinical and epidemiologic studies, where precision, scalability, and reliability are vital considerations. Thus, DBMS has transcended beyond being just a technological application tool and has emerged as a crucial innovation catalyst for applied sciences.

Additionally, the results of the co-authorship analysis show that collaborations within the domain of DBMS research continue to be largely clustered, with high levels of cooperation within groups but low levels of

connection between groups. This implies that research teams usually form based on their interest in a particular sub-domain or region, which might constrain knowledge transfer between various clusters. Although specialization allows for a deeper understanding of certain areas, it might prevent a comprehensive solution from being achieved. Thus, there is a need for cross-institutional and cross-country collaboration in DBMS research.

Citation analysis reveals that the contributions of seminal works as well as methodology are greatly valued. Citation analysis also shows that the cited articles are often related to interdisciplinary fields, for example, bioinformatics and health sciences. The conclusion we can make from this is that valuable studies in DBMS are often application-oriented, thus emphasizing the importance of applicability of the DBMS systems. This proves that the applicability and relevance of the DBMS systems are of great value when determining the academic significance of DBMS research.

Furthermore, the keyword co-occurrence and overlay analysis indicates a definite trend towards a new direction in DBMS research; this trend is represented by the movement from classical applications of databases to new directions that include artificial intelligence, machine learning, and intelligent information systems. Thus, it appears that modern data environments require more complex approaches in their handling, which involves advanced data technologies. The appearance of the mentioned themes indicates that DBMS research enters a new era in which integration of new data technologies into DBMS becomes a crucial task. Therefore, future studies will be directed at the creation of hybrid information systems that unite all types of processes involved.

This study reveals that the nature of DBMS studies is such that it has firm foundational principles, greater interdisciplinary applicability, and an increasingly changing technology framework. Although much progress has been made within this field, especially in terms of its application within certain domains, further improvement can be made with respect to fostering international cooperation, expanding its research environment, and embracing new technologies. These perspectives provide useful areas in which to explore future avenues of research.

4. CONCLUSION

In conclusion, this bibliometric review shows that the area of Database Management Systems (DBMS) studies has become a rapidly growing field with deep roots, wide-ranging applications, and new technologies. It has been observed that although in earlier times, research conducted was mainly related to structured data management and its applications in fields such as health care and scientific studies, recent trends have indicated a clear trend towards the use of new technologies like artificial intelligence and machine learning. In addition to this, while there were many collaborative clusters present along with impactful literature, there is a need for greater collaboration between regions and more integrative research practices.

REFERENCES

[1] T.-T. T. Phan, C.-T. Vu, P.-T. T. Doan, D.-H. Luong, and T.-P. Bui, "Two decades of studies on learning management system in higher education: A bibliometric analysis with Scopus database 2000-2020," *J. Univ. Teach. Learn. Pract.*, vol. 19, no. 3, pp. 1–21, 2022.

[2] R. Farooq, "A review of knowledge management research in the past three decades: a bibliometric analysis," *Vine J. Inf. Knowl. Manag. Syst.*, vol. 54, no. 2, pp. 339–378, 2024.

[3] P. T. Chountalas and A. G. Lagodimos, "Integrated management systems: a content and bibliometric analysis," *TQM J.*, vol. 37, no. 7, pp. 1827–1873, 2025.

[4] Y. Gu, "Global knowledge management research: A bibliometric analysis," *Scientometrics*, vol. 61, no. 2, pp. 171–190, 2004.

[5] P.-T. Pham *et al.*, "Learning management system in developing countries: A bibliometric analysis between 2005 and 2020," *Eur. J. Educ. Res.*, vol. 11, no. 3, pp. 1363–1377, 2022.

[6] H. A. Schildt, "Sitkis: software for bibliometric data management and analysis," *Helsinki Inst. Strateg. Int. Bus.*, vol. 6, no. 1, 2002.

[7] Q. Wang, G. M. Garrity, J. M. Tiedje, and J. R. Cole, "Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy," *Appl. Environ. Microbiol.*, vol. 73, no. 16, pp. 5261–5267, 2007.

[8] S. Rozen and H. Skaletsky, "Primer3 on the WWW for general users and for biologist programmers," in *Bioinformatics methods and protocols*, Springer, 2000, pp. 365–386.

[9] E. Cerami *et al.*, "The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data," *Cancer Discov.*, vol. 2, no. 5, pp. 401–404, 2012.

[10] M. D. Wilkinson *et al.*, "Comment: The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3." 2016.

[11] B. R. Haugen *et al.*, "2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid Association guidelines task force on thyroid nodules and differentiated thyroid cancer," *thyroid*, vol. 26, no. 1, pp. 1–133, 2016.

[12] A. Conesa, S. Götz, J. M. García-Gómez, J. Terol, M. Talón, and M. Robles, "Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research," *Bioinformatics*, vol. 21, no. 18, pp. 3674–3676, 2005.

[13] S. Michie, M. M. Van Stralen, and R. West, "The behaviour change wheel: a new method for characterising and designing behaviour change interventions," *Implement. Sci.*, vol. 6, no. 1, p. 42, 2011.

[14] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Trans. neural networks*, vol. 20, no. 1, pp. 61–80, 2008.

[15] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *J. Manag. Inf. Syst.*, vol. 24, no. 3, pp. 45–77, 2007.

[16] K. Barnett, S. W. Mercer, M. Norbury, G. Watt, S. Wyke, and B. Guthrie, "Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study," *Lancet*, vol. 380, no. 9836, pp. 37–43, 2012.