


The Importance of Information Systems and Artificial Intelligence Research in Biomedical, Analysis of All Publications in the Scopus Database with R Software

İNCİLAY YILDIZ^{1*},

¹Department of Management Information Systems, Altınbaş University, İstanbul, Türkiye
incilay.yildiz@altinbas.edu.tr

 <https://orcid.org/0000-0001-5572-5058>

Received: 18.03.2024

Accepted: 29.03.2024

Published: 30.04.2024

*Corresponding author

Research Article

pp.33-48

Abstract

As technological advancements continue to exert influence across various facets of our daily existence, the significance of information systems and artificial intelligence is progressively escalating within the discipline of biomedical. The aim of the study is to analyze the studies biomedical and information systems and artificial intelligence. Informations were questioned by selecting article title, abstract, keywords using the Scopus database for analysis. As a result of the scanning, it was determined that there were 319 academic studies from 1976 to 2023. All studies were included in the bibliometric analysis. 4.3.1 of the R software for analysis. The bibliometric library of the version was used. The overview, sources, authors, documents, conceptual structure and social structure groups in the library were analyzed. The most published years are 2007 (20) 2008 (20) and 2006 (19). The most published journals Are Journal of Biomedical Informatics (15). The most published author is Demner-Fushman with 5 publications. The countries that publications the most are USA (59) The most cited countries are USA (1922). It was determined that 827 keywords were used. The most used keywords were determined as artificial intelligence (352).

Keywords: Artificial Intelligence, Information Systems, Biomedical, Data Analysis

1. Introduction

The term "biomedical" is a portmanteau of the words "biological" and "medical" and generally describes disciplines involving scientific research, technology, and engineering applications in the fields of biology and medicine. Biomedical sciences and engineering focus on human health and cover topics such as understanding biological processes, diagnosis and treatment of diseases, and design and development of medical devices. Biomedicine refers to the intersection of biological and medical engineering sciences. This discipline focuses on technological solutions developed for the understanding of biological systems and the diagnosis and treatment of diseases. Biomedical engineers play an important role in the design, development and use of medical devices by combining biological processes with engineering principles "This discipline exists at the intersection of scientific research, technological innovation, and clinical practice. (Enderle and Bronzino, 2012). Information systems encompass a framework that integrates individuals, information technologies, and operational procedures within organizations. These systems facilitate the acquisition, manipulation, retention, and dissemination of data. Moreover, they play a pivotal role in aiding decision-making processes, refining operational workflows, and facilitating the attainment of organizational objectives. Furthermore, information systems streamline the exchange of information across various functions inherent within the organization. (Laudon and Laudon, 2009). Artificial intelligence (AI) is a technology used to enable computer systems to gain human-like abilities and perform complex tasks. AI aims for computers to realize abilities similar to human intelligence, such as learning, reasoning, problem solving, language understanding, perception and decision making. In this field, there are many different approaches and techniques. For example, techniques such as machine learning and deep learning enable algorithms to learn from data to build models and

then make inferences based on new data. In addition, subfields such as expert systems, natural language processing, image processing, robotics and autonomous systems are among the application areas of artificial intelligence (Yılmaz, 2022).

The integration of these three fields presents opportunities for significant advancements in healthcare delivery, ultimately leading to better health outcomes for individuals. The aim of the study is to evaluate the academic results by examining the studies carried out with information systems and artificial intelligence technologies, which are an important part of technology in the field of biomedicine, which has become an important branch of health sciences, with bibliometric analysis.

In the health sciences, the confluence of biomedicine, information systems, and artificial intelligence is crucial for a number of reasons.

- 1. Information Management and Analysis:** The quality and effectiveness of healthcare services are improved by the efficient management and analysis of biomedical and health data. When it comes to gathering, storing, analyzing, and exchanging medical data, information systems are essential. Large-scale biomedical data may be analyzed using artificial intelligence and data analytics tools to yield insightful information about the diagnosis and management of disease.
- 2. illness Diagnosis and Treatment:** Biomedical systems and artificial intelligence can speed up and enhance the precision of illness diagnosis and treatment. Deep learning algorithms, for example, can help medical personnel diagnose pathological illnesses more accurately by extracting relevant information from medical pictures.
- 3. Personalized Treatment and Care:** Artificial intelligence can customize treatment and care plans to each patient's unique traits and medical needs by evaluating biological data. This individualized approach to care can reduce side effects and increase efficacy.
- 4. Design, Development, and Improvement of Medical equipment:** Artificial intelligence and information systems may be applied to the design, development, and improvement of medical equipment. This makes it possible to develop medical equipment that is more intelligent and effective, which eventually improves patient care.
- 5. Efficiency in Healthcare Services:** Artificial intelligence and information technologies help to improve the efficiency of healthcare operations and administration. AI-enabled planning and scheduling systems, for example, can enhance resource allocation, decrease wait times, and optimize patient visits

In the first part of the study, the 10 most cited studies were examined and their common results were evaluated. In the 3rd chapter, bibliometric analysis, which is the analysis method of the study, and scopus, which is a database, are mentioned. In the results and discussion section of the study, the findings of the study analysis are explained with the help of tables and graphs. The study ended with the conclusion part.

2. Literature Review

The literature research was evaluated based on the 10 most cited studies and is summarized in table 1. "Total Citations" shows the number of citations of a work over all time. This is an indicator of how effective or important the work is. The more citations a work receives, the more it has been referenced and interacted with by other works. "TC per year" refers to the annual number of citations. This refers to the number of citations the work received in a given year. The annual citation count can be used to show how long a work has been interacted with or referenced. It is an important metric, especially for measuring the impact and popularity of a new study. The study conducted by Bunescu R. et al in 2005 was the most cited study on this subject with 280 citations. The study conducted by Bunescu R. et al in 2005 was the most cited study on this subject with 280 citations. Da Costa C.A. The study conducted by et al. in 2018 appears to be the most popular study with a citation rate of 26.43 years.

Table 1. Most Grobal Cited Documents

No	Authors	Source Name	Total Citations	TC per year
1	Bunescu R. et al, 2005	Artificial Intelligence in Medicine	280	14,00
2	Da Costa C.A. et al, 2018	Artificial Intelligence in Medicine	185	26,43
3	Butte A. J. & Kohane, I. S. 2006,	Nature Biotechnology	153	8,05
4	Pratt W. et al, 2004	Journal of Biomedical Informatics	139	6,62
5	Isaksson A. et al, 2008	Pattern Recognition Letters	138	8,12
6	Mishra, R. et al 2014	Journal of Biomedical Informatics	136	12,36
7	Bhatti, U. A., et al, 2019	Enterprise Information Systems	131	21,83
8	Mbunge, E. 2020	Diabetes & Metabolic Syndrome: Clinical Research & Reviews	128	25,60
9	Virone G, Noury N, Demongeot J, 2002	IEEE Transactions on Biomedical Engineering	128	5.57
10	Alexander C. Yu., 2006	Journal of Biomedical Informatics	126	6.63

Study 1 states that the process of extracting information from biomedical texts holds the promise of easily consolidating biological information in large quantities and into a computer-accessible form. In particular, it is stated that it is a strategy that facilitates the extraction of data about genes in the human genome from 11 million abstracts in Medline. Authors have developed and evaluated several learned information extraction systems to identify human protein names in Medline abstracts and then extract information about interactions between these proteins. The authors used various machine learning methods to automatically extract gene/protein name, function, and interaction information from Medline abstracts to identify human proteins and their interactions. They present cross-validated results using datasets for training and testing from approximately 1000 hand-marked Medline abstracts discussing human genes/proteins (Bunescu et al, 2005).

Study 2 notes that the large amounts of patient data collected in hospitals through separate medical devices are often difficult to combine and analyze electronically, as they are not part of the electronic health record. As a way to overcome these limitations, connecting medical devices over the Internet through a distributed platform called Internet of Things (IoT) is proposed. This approach combines data from different sources to better diagnose patients' health conditions and identify possible preventive measures. The study introduces the concept of Internet of Health Things (IoHT), which examines different approaches that can be applied to collect and combine vital signs data in hospitals (Da Costa et al, 2018).

In study number 3, it is emphasized that gene and protein measurements alone do not fully characterize the phenotype of the samples. By examining the relationships between phenotype, genotype and environmental components, it is aimed to identify genes that may affect responses to the environment. To this end, relationships between genes associated with phenotypic, disease, environmental and experimental contexts were investigated using gene expression datasets obtained from the Gene Expression Omnibus. The study highlights that the comprehensive identification of genes associated with phenotype and environment is a step towards the Human Phenomenal Project (Buttle and Kohane, 2006).

Study 4 notes that many information systems fail when applied to complex healthcare environments, and one of the reasons for this is the difficulty of systematically accounting for the collaborative and exception-filled nature of medical work. It is emphasized that research from the field of CSCW can help biomedical information professionals recognize these challenges and work around them when designing their systems. Using an example such as electronic medical record systems, it is explained how three specific principles from CSCW – accounting for incentive structures, understanding workflow, and incorporating awareness – can improve the design and delivery of medical information systems (Pratt et al, 2004).

According to study number 5, there is a growing interest in statistical classification in significant applications, including crucial ones that involve using supervised learning to diagnose patient samples. A trustworthy performance estimate is a must for any decision support system to be approved, particularly in situations when judgments made later on might have grave repercussions, such in the case of cancer therapy. Over the past 20 years, practically all scientific fields have adopted cross-validation (CV) and bootstrap (BTS) as the most popular techniques for determining these predictions because to their adaptability to small sample difficulties. Here, the attention is focused on the fact that the uncertainty of a point estimate derived using CV and BTS is relatively

substantial and sometimes disregarded in small sample classification issues found in biomedical applications and other domains. It is advised that, until better alternatives are found, final classification performance should always be reported in the form of a Bayesian confidence interval obtained from a straightforward retention test or using another method that gives conservative measures of uncertainty in order to avoid this fundamental issue with using CV and BTS (Isaksson et al, 2008).

The aim of study 6 is to conduct a thorough evaluation of current literature on text document summarization in the biomedical domain. We searched the IEEE Digital Library, the ACM Digital Library, and MEDLINE (2000–October 2013). The researchers looked at text summary strategies in the biological sector and independently screened and summarized papers on the subject. Five aspects of information were extracted from the chosen articles: introduction, purpose, outcome, technique, and evaluation. 34 (0.3%) out of 10,786 studies satisfied the requirements for inclusion. The most popular summarizing methodologies were natural language processing (50%) and a hybrid technique like statistics, machine learning, and natural language processing (44%). 82% of the research carried out an internal evaluation. In the discipline of biomedicine, this is the first comprehensive review of text summarizing. In addition to identifying research gaps, the paper offers suggestions for future directions in biological text summarizing research. A hybrid approach that combines machine learning, statistics, and language processing methods has been the subject of recent study. The use and assessment of text summarizing in real research or patient care settings require more investigation (Mishra et al, 2014).

In article number 7, it is stated that medical information systems are becoming increasingly important by improving and facilitating the quality of healthcare services along with evidence-based medical treatment in areas such as health monitoring, disease trend modeling and early intervention through data mining and feature extraction methods. Such systems are part of the corporate information systems of healthcare organizations. A new algorithm on recommendation algorithms, fb-kNN, is proposed based on the analysis of disease patterns and patterns in the human body, and this algorithm has been implemented in Healthcare 4.0 for diagnosis and treatment recommendation. The developed tool is a complete package solution for the Corporate Management System (ERP) and contributes to improvement in health, reduction of chronic diseases and reduction of mortality rates (Bhatti et al, 2019).

Study 8 addresses the need for many healthcare systems and governments to use contact tracing as one of the prevention and control methods due to the lack of approved vaccines for COVID-19 as of August 2020. However, considering the difficulties in contact tracing and the inability of infected people to access the contact information of their contacts, this study analyzes the potential opportunities and challenges encountered in the integration of new technologies in COVID-19 contact tracing. While the results highlight the potential benefits of technology-based contact tracing applications, they also consider the various technical, social, security and privacy risks these technologies face. Therefore, the conclusion of the study calls for further research and policy development to increase the effectiveness and efficiency of the integration of new technologies into COVID-19 contact tracing (Mbunge, 2020).

An automated circulatory activity monitoring system has been created as part of telemedicine and is outlined in article number 9 of the "Health Integrated Smart Home Information System" (HIS²). HIS² is an experimental platform designed to test and develop solutions that protect people needing home-based medical monitoring and

improve their quality of life. Position sensors that are placed in every room enable the tracking of activity patterns in the patient's living space over time. We next proceeded to discover modest rhythmic fluctuations with an hourly sample. We calculated an average value and confidence limits based on a large number of measurements. Additionally, it enabled us to establish a boundary for the patient's behavior that would be seen as "predictable." If patient activity deviates from this zone, alarms are triggered (Virone, Noury, Demongeot, 2002).

Research on ontologies is becoming increasingly common in the biomedical informatics community, but the challenges associated with their construction and maintenance have proven more formidable than expected. As a result, identifying effective methods has become a primary focus for those looking to leverage ontologies. Article 10 provides an overview of contemporary approaches to creating, maintaining, aligning, and evaluating ontologies (Alexander, 2006).

The common focus of these studies is research conducted in various fields to improve the quality of health care by using developments in health services and new technologies. In the field of biomedical information, the focus is on topics such as text summarization, disease diagnosis and treatment, medical monitoring systems, COVID-19 contact tracing and ontology creation. Each study provides recommendations for improving or optimizing a specific area or technology in healthcare. A common theme is that with the increasing role of information technologies and artificial intelligence in healthcare, the importance of innovative solutions to monitor patients' health status, diagnose and treat diseases is emphasized.

3. Methods

A research technique called bibliometric analysis is used to assess and analyze scientific books quantitatively. By using statistical techniques on scientific papers, articles, citation data, and related data, this kind of study seeks to gather a variety of information (De Solla Price, 1963). Research trends, academic exchanges, author impact, and advancements in certain fields are all assessed using bibliometrics. Publication Numbers and Trends, Citation Analysis, Author Analysis, Institution Analysis, Journal Analysis, and Co-Authorship Analysis are all carried out as part of the analysis. It is employed to assess the influence of scholars, organizations, or publications as well as to gauge success in the pertinent subject. When it comes to making decisions on things like allocating research funds, choosing scientific methods and policies, and enhancing scientific communication, bibliometric analysis is frequently a valuable source of information.

Scopus is a multidisciplinary scientific database and contains a large collection of scientific articles in academic literature. Provides access to high-quality academic resources published by researchers, scientists and institutions. Scopus covers millions of articles across a variety of disciplines, including science, technology, medicine, social sciences, and arts (Scopus, 2023 <https://www.scopus.com/search>). In this study, the Scopus database was used for analysis. All publications until the end of 2023 were included in the study. Within the purview of this investigation, the Scopus database was used to look for works in the fields of biomedicine, information systems, and artificial intelligence. Article titles, abstracts, and keywords were searched in order to gather studies for the data collection. The R program version 4.3.1's bibliometric library was utilized for bibliometric analysis. The overview, resources,

authors, documents, conceptual structure, social structure groups, and documents were among the library's components that were examined.

4. Results and Discussion

In this section, the findings obtained as a result of bibliometric analysis are explained with tables and figures. Table 2 shows the number of studies and citation rates by year.

Table 2. Average Citations Per Year

Year	MeanTCperArt	N	MeanTCperYear	CitableYears
2007	25,35	20	1,41	18
2008	31,45	20	1,85	17
2006	30,68	19	1,61	19
2018	18,47	17	2,64	7
2020	15,71	14	3,14	5
2011	11,69	13	0,84	14
2021	7,08	12	1,77	4
2022	14,67	12	4,89	3
2005	46,09	11	2,30	20
2017	14,73	11	1,84	8

“meanTCperArt” represents the average number of citations per year obtained by dividing the total number of citations over the life of an article by the number of years it was published. This criterion is used to evaluate the impact and importance of an article. The expression "TC" means "Total Citations". The term "MeanTCperYear" indicates the average number of citations per year obtained by dividing the total number of citations of an article by the years it was published. A high "MeanTCperYear" value may indicate that an article has received long-term attention and is frequently cited. It can be used to evaluate how much attention an article has received since its publication or how much it has contributed to other studies in the field. The term "CitableYears" refers to the total number of years for which an article or a work is citable. This term is used to indicate how long an article has been available in scientific literature. The expression N indicates the number of articles.

According to Table 1, the most publications were made in 2007 (20), 2008 (20) and 2006 (19), respectively. It seems that 20 articles published in 2007 have been cited for a total of 18 years. In this case, it was observed that the articles

were citable for approximately 0.9 years on average. In general, the yearly citation rates appear to have increased in citable years to the number of articles. "The years with the highest MeanTCperArt value appear to be 2005 (46.09), 2008 (31.45) and 2006 (30.68). MeanTCperYear value, that is, the annual average number of citations, appears to be highest in 2022 (4.89). Although the number of articles is less than other years, such as 12 articles, the fact that there are so many citations emphasizes that the importance of the subject has increased in recent years.

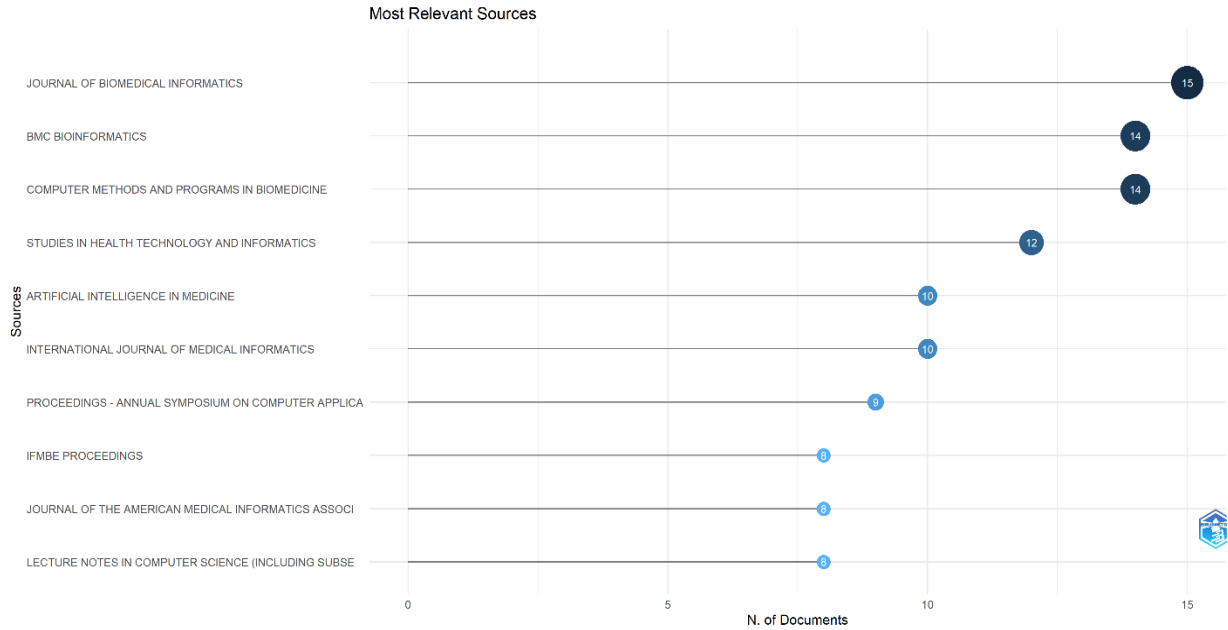


Figure 1. Most Relevant Sources

The 10 most published journals are shown in figure 10 under the title of most relevant sources. The most relevant sources are in order Journal of Biomedical Informatics (15), BMC Bioinformatics (14), Computer Methods and Programs In Biomedicine (14).

Figure 2 shows the number of publications of sources cumulatively over time. Figure 2 shows the number of publications of sources cumulatively over time. According to Figure 2, it can be seen that publications started to be made between 1996 and 2006. Table 3 shows the most published authors and their contributions in proportion to the number of authors in the article. (Fractionalized: divided into separate groups or parts.) According to Table 3 the most published author is Demner Fushman (5), while the authors who seem to have contributed the most to the field are Holzinger (1,48) and Rau (1,48). Figure 3 shows in which years and how much the authors published.

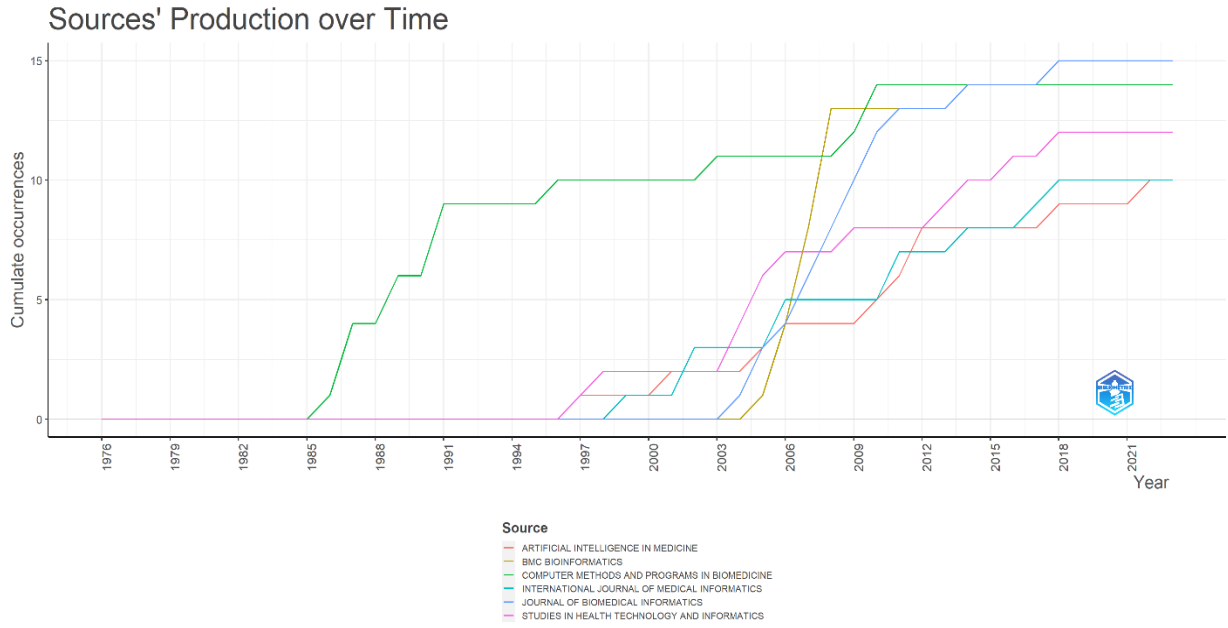


Figure 2. Sources Production Over Time

Table 3. Most Relevant Authors

Authors	Articles	Articles Fractionalized
DEMNER-FUSHMAN D	5	1,07
GROTH T	4	1,20
ANTANI S	3	0,70
BELLAZZI R	3	0,81
FERNÁNDEZ-BREIS JT	3	0,70
HOLMES JH	3	0,37
HOLZINGER A	3	1,48
MARTÍNEZ-BÉJAR R	3	0,70
RAU G	3	1,48
RINDFLESCH TC	3	0,62

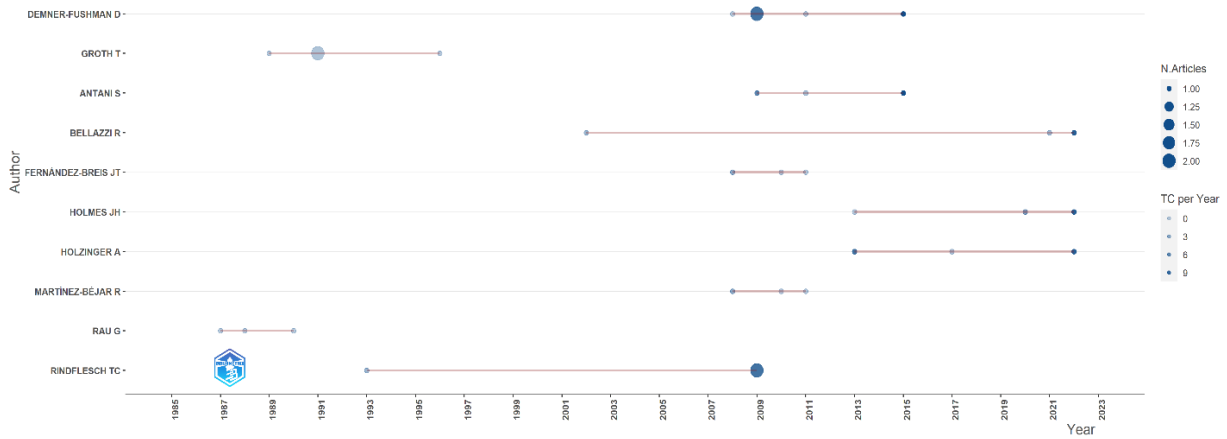


Figure 3. Author's Production Over Time

Figure 4 shows the countries of the corresponding authors. Countries with the most documents USA (59), CHINA (15) UK (12). According to Figure 4, The parts shown in red in the table represent publications made jointly with other countries, and the blue parts represent publications made only with authors from that country. USA (5) and UK (5) are the most common in documents made with partner countries.

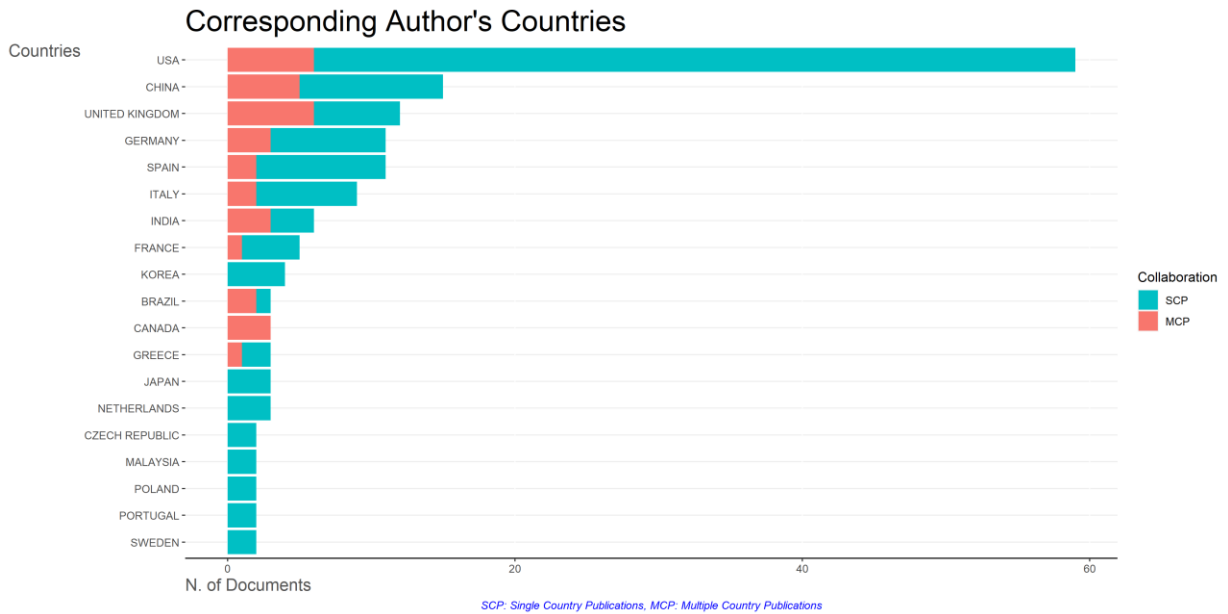


Figure 4. Corresponding Author's Countries

In figure 5, the countries with the most broadcasts are highlighted in dark blue. As the number of publications decreases, the blue color fades to gray. The countries with the most publications were USA (275), CHINA (62) and SPAIN (41).

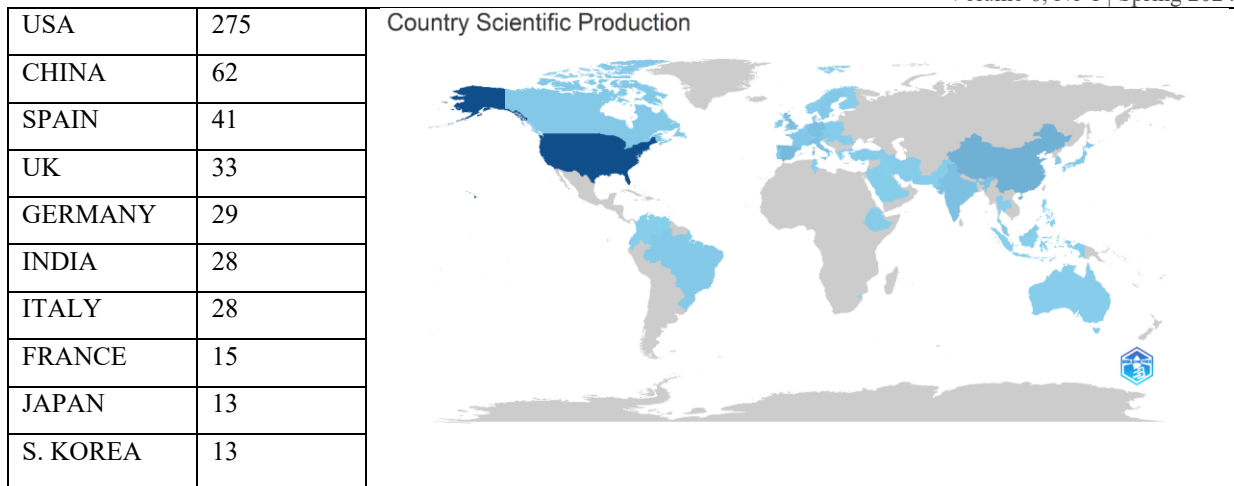


Figure 5. Country Scientific Production

Figure 6 shows the 10 most cited countries. USA 1922 seems to be clearly ahead with the reference. Followed by the UK 279 and China (276), which are the other countries with the most publications.

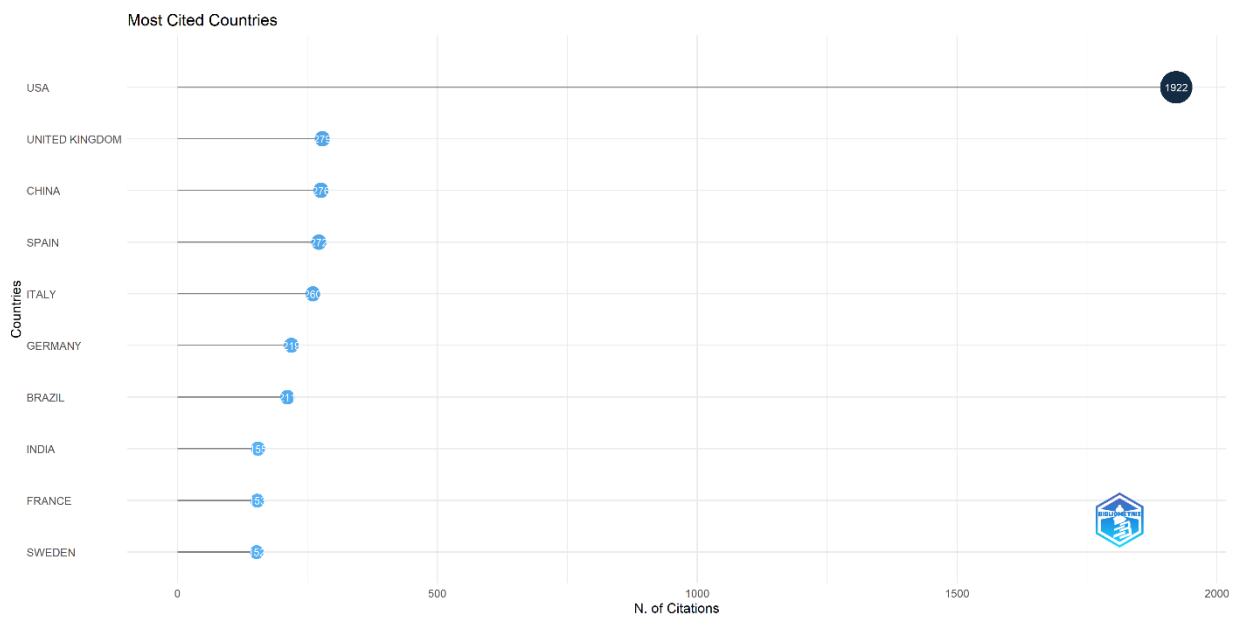


Figure 6. Most Cited Countries

Figure 7 shows the 10 most relevant affiliations. Figure 7 shows the 10 most published affiliations. Accordingly, Stanford University, Columbia University, University of Washington are seen as the affiliations that make the most publications.

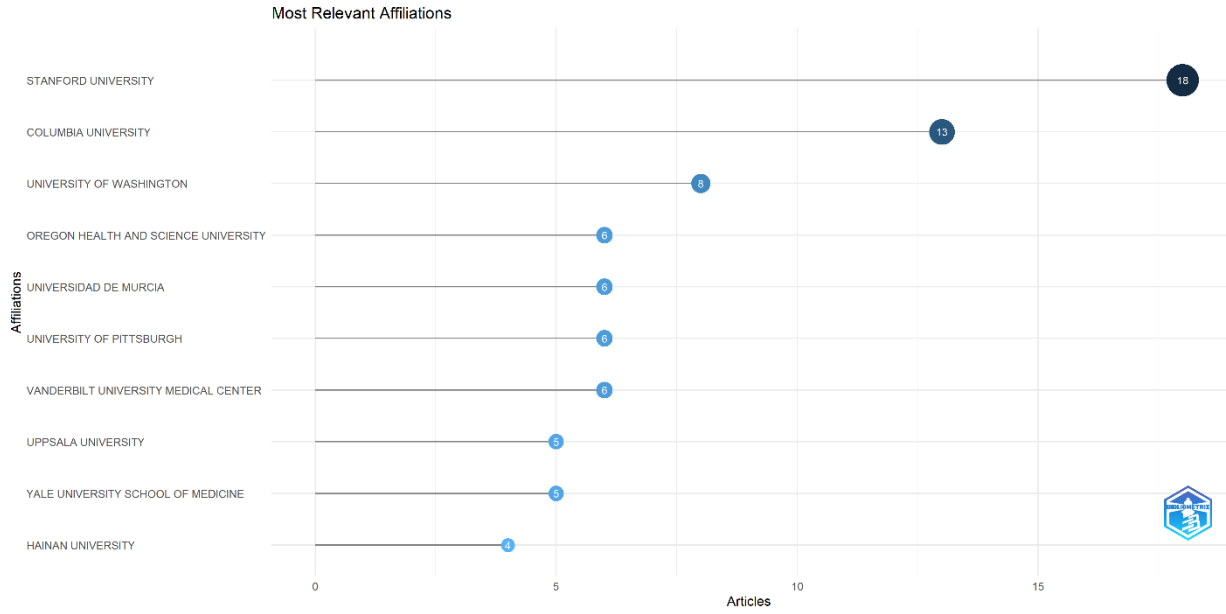


Figure 7. Most Relevant Affiliations

In figure 8, the most used keywords are shown as a word cloud.

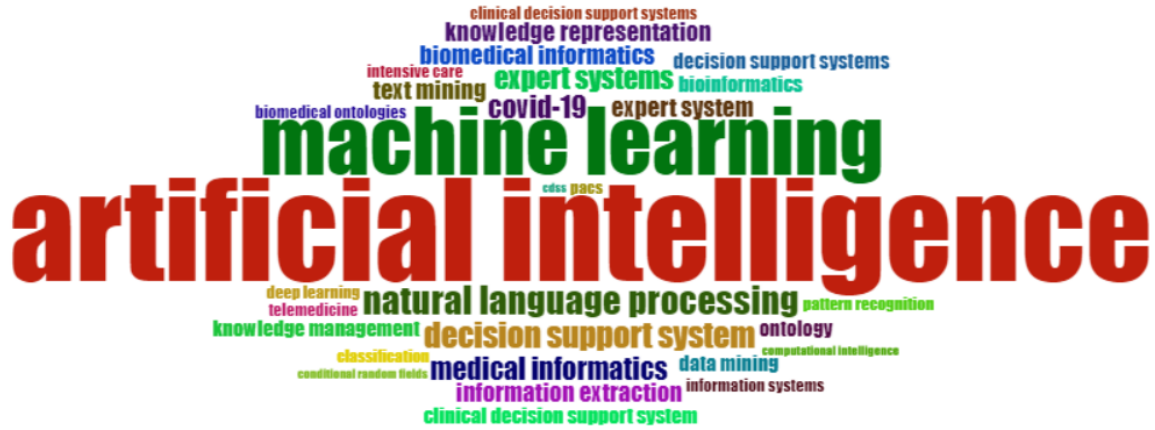


Figure 8. Word Cloud

Figure 9 shows the 10 most used keywords. Accordingly, artificial intelligence (32), machine learning (21), natural language processing (9), decision support system (8), covid-19 (7). We can conclude that in the field of biomedical, machine learning and natural language processing algorithms of artificial intelligence are used for decision support systems, especially during the Covid-19 period.

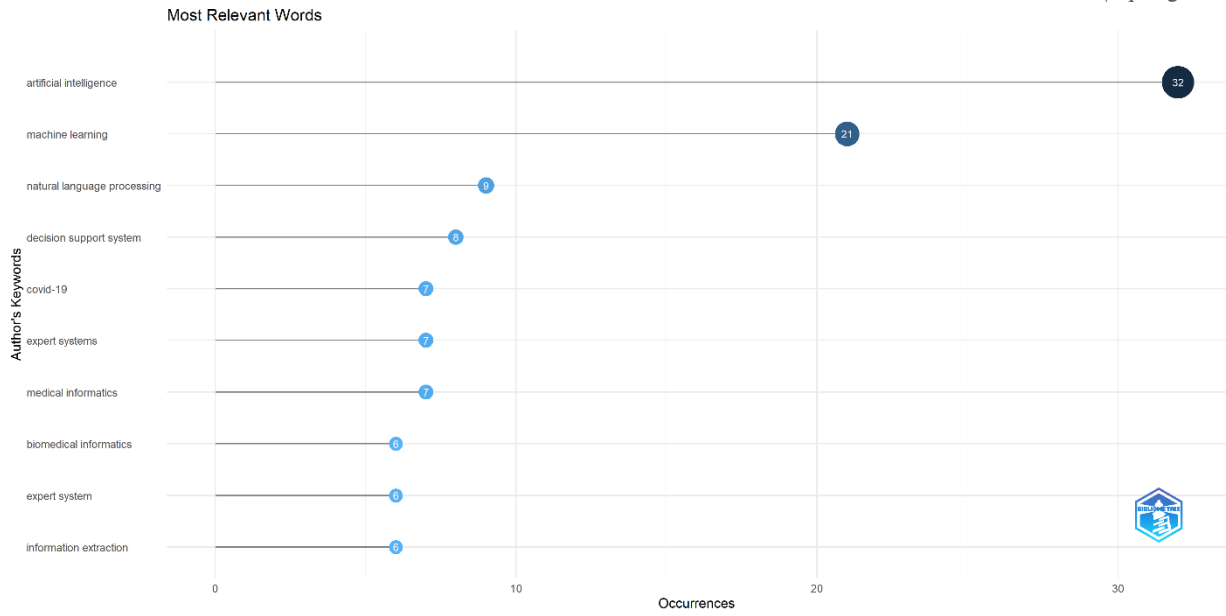


Figure 9. Most Relevant Words

Figure 10 shows the word tree and the percentage rates of the 30 most used keywords. Accordingly, artificial intelligence is seen as the most used keyword 32 times with a rate of 16%.

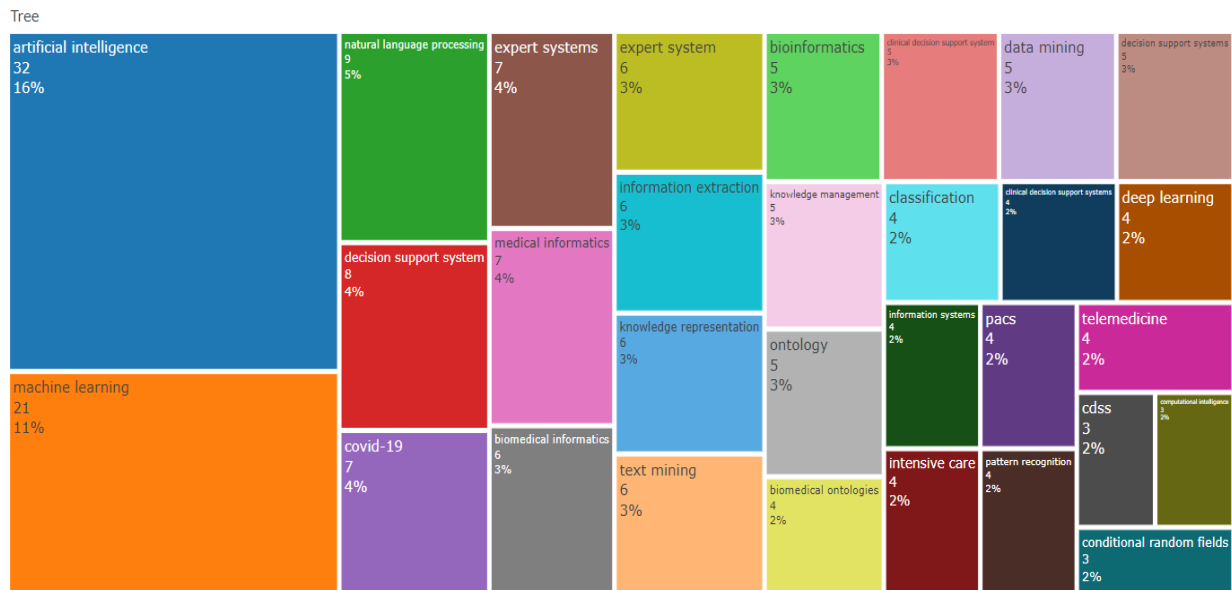


Figure 10. Word Tree

In figure 11, the periods in which the words are used most in the timeline are highlighted.

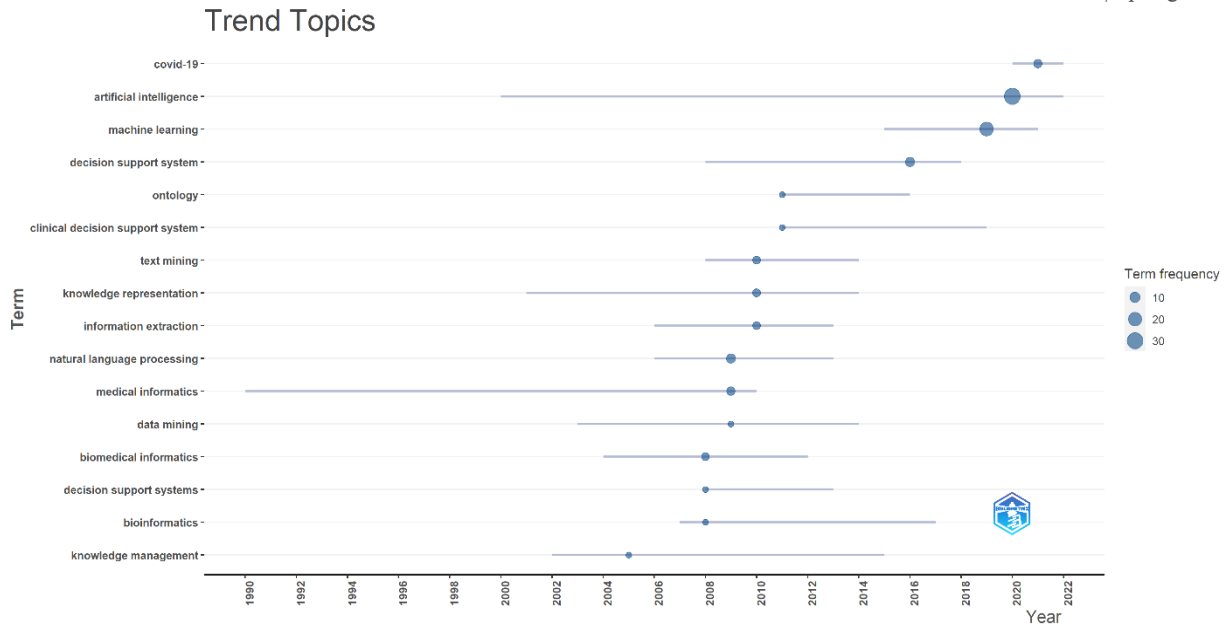


Figure 11. Trend Topics

5. Conclusion

This study conducts a bibliometric analysis focusing on the intersection of information systems and artificial intelligence within the field of biomedicine. The evolving integration of these domains holds significant promise for advancing medical research, diagnosis, and treatment. Employing bibliometric techniques, this research identifies key trends, influential authors, leading journals, and important research themes. The findings offer valuable insights into the current research landscape at this interdisciplinary junction, providing guidance for future research and strategic developments in biomedical informatics.

According to the results of the study, 319 publications were identified in 168 sources between 1976 and 2023. While the average number of citations per document was calculated as 19.05, the average age of the documents was 17,1. While it was determined that 1100 authors worked in this field, it was observed that only 59 authors worked alone. Co-authorship per document was calculated as 3.71. The international co-authorship percentage is 15,67 in publications with 827 keywords. The most common document types are articles (152) and academic symposium papers (115). The most published journals are Journal of Biomedical Informatics (15) The most published author is Demner Fushman with 6 publications. The countries that publications (59) and cites (1922) the most are USA. The year with the most publications was 2007. Since the Annual Growth Rate of the number of publications is 4.52%, it has been revealed that studies on this subject are increasing and therefore considered important.

The conclusion drawn from the 10 most cited studies in this field underscores the escalating significance of innovative solutions in monitoring patients' health status, diagnosing, and treating diseases, propelled by the expanding role of information technologies and artificial intelligence in healthcare services.

References

- Alexander, C. Y. (2006). Methods in biomedical ontology. *Journal of biomedical informatics*, 39(3), 252-266.
- Bhatti, U. A., Huang, M., Wu, D., Zhang, Y., Mehmood, A., & Han, H. (2019). Recommendation system using feature extraction and pattern recognition in clinical care systems. *Enterprise information systems*, 13(3), 329-351.
- Bunescu, R., Ge, R., Kate, R. J., Marcotte, E. M., Mooney, R. J., Ramani, A. K., & Wong, Y. W. (2005). Comparative experiments on learning information extractors for proteins and their interactions. *Artificial intelligence in medicine*, 33(2), 139-155.
- Butte, A. J., & Kohane, I. S. (2006). Creation and implications of a phenome-genome network. *Nature biotechnology*, 24(1), 55-62.
- Da Costa, C. A., Pasluosta, C. F., Eskofier, B., Da Silva, D. B., & da Rosa Righi, R. (2018). Internet of Health Things: Toward intelligent vital signs monitoring in hospital wards. *Artificial intelligence in medicine*, 89, 61-69.
- De Solla Price, D. J. (1963). Little science, big science--and beyond. Columbia University Press.
- Enderle, J., & Bronzino, J. (Eds.). (2012). *Introduction to biomedical engineering*. Academic press.
<https://www.scopus.com/search> available 01.01.2024
- Isaksson, A., Wallman, M., Göransson, H., & Gustafsson, M. G. (2008). Cross-validation and bootstrapping are unreliable in small sample classification. *Pattern Recognition Letters*, 29(14), 1960-1965.
- Laudon, K., & Laudon, J. (2009). *Management Information Systems: International Edition*, 11/E. London, UK: Pearson Higher Education.
- Mbunge, E. (2020). Integrating emerging technologies into COVID-19 contact tracing: Opportunities, challenges and pitfalls. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 14(6), 1631-1636.
- Mishra, R., Bian, J., Fiszman, M., Weir, C. R., Jonnalagadda, S., Mostafa, J., & Del Fiol, G. (2014). Text summarization in the biomedical domain: a systematic review of recent research. *Journal of biomedical informatics*, 52, 457-467.
- Pratt, W., Reddy, M. C., McDonald, D. W., Tarczy-Hornoch, P., & Gennari, J. H. (2004). Incorporating ideas from computer-supported cooperative work. *Journal of biomedical informatics*, 37(2), 128-137.
- Virone, G., Noury, N., & Demongeot, J. (2002). A system for automatic measurement of circadian activity deviations in telemedicine. *IEEE Transactions on Biomedical Engineering*, 49(12), 1463-1469.
- Yılmaz, A (2022). Yapay Zeka. Kodlab Yayın Dağıtım Yazılım Ltd. Şti. 10.Basım

aurum

AURUM SAĞLIK BİLİMLERİ DERGİSİ
AURUM JOURNAL OF HEALTH SCIENCES

Cilt 6, Sayı 1 | Bahar 2024
Volume 6, No 1 | Spring 2024