informatics

MDPI

Article

From Innovation to Regulation: Insights from a Bibliometric Analysis of Research Patterns in Medical Data Governance

Iulian V. Nastasa [1,2], Andrada-Raluca Artamonov [2,*], Ștefan Sebastian Busnatu [2,3], Dana Galieta Mincă [1] and Octavian Andronic [2]

[1] Discipline of Public Health and Management, Faculty of Medicine, Carol Davila University of Medicine and Pharmacy, 050474 Bucharest, Romania
[2] Center for Innovation and e-Health, Carol Davila University of Medicine and Pharmacy, 030167 Bucharest, Romania
[3] Cardiology Department, Carol Davila University of Medicine and Pharmacy, Bagdasar-Arseni Emergency Clinical Hospital, 041915 Bucharest, Romania
* Correspondence: andrada.artamonov@umfcd.ro

Abstract

This study presents a comprehensive bibliometric analysis of the evolving landscape of data protection in medicine, examining research trends, thematic developments, and scholarly contributions from the 1960s to 2024. By analyzing 2159 publications indexed in the Scopus database using the Bibliometrix R package (v.4.3.2), based on R (v.4.4.3), this paper maps key research areas, leading journals, and international collaboration patterns. Our findings reveal a significant shift in focus over time, from early concerns centered on data privacy and management to contemporary themes involving advanced technologies such as artificial intelligence, blockchain, and big data analytics. This transition reflects the increasing complexity of balancing data accessibility with security, ethical, and regulatory requirements in healthcare. This analysis also highlights persistent challenges, including fragmented research efforts, disparities in global contributions, and the ongoing need for interdisciplinary collaboration. These insights offer a valuable foundation for future investigations into medical data governance and emphasize the importance of ethical and responsible innovation in an increasingly digital healthcare environment.

Keywords: data protection; data governance; medicine; patient confidentiality; cybersecurity

check for updates

Academic Editor: Antony Bryant

Received: 21 May 2025
Revised: 30 June 2025
Accepted: 4 July 2025
Published: 8 July 2025

Citation: Nastasa, I.V.; Artamonov, A.-R.; Busnatu, Ș.S.; Mincă, D.G.; Andronic, O. From Innovation to Regulation: Insights from a Bibliometric Analysis of Research Patterns in Medical Data Governance. *Informatics* **2025**, *12*, 66. https://doi.org/10.3390/informatics12030066

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The management of medical data has undergone significant transformations, driven by the need for enhanced privacy, security, and patient empowerment. The evolution of privacy and security frameworks in this domain reflects a shift from traditional, centralized approaches to more decentralized, patient-centric models, increasingly leveraging emerging technologies [1,2].

The increasing implementation of electronic health records (EHRs) and the demand for interoperability have heightened concerns regarding data control, as patients express greater interest in managing their personal information. Despite the transition from a predominantly isolated, paper-based health record system to an integrated, electronic framework, minimal policy development has initially addressed the substantial privacy concerns arising from this shift. Furthermore, advancements in Information Technology have exposed patient health data to novel security and privacy risks, outpacing the capacity of existing legislation to adapt to these technological developments [1]. In the digital age,

cyberattacks pose significant risks to individuals, organizations, and national infrastructure, leading to financial loss, operational disruption, data breaches, and threats to public safety and security [3]. The healthcare sector is continually exposed to risks of compromise, as hackers seek to exploit vulnerabilities in operating systems, browsers, or hardware, potentially leading to loss of sensitive patient information. Regular security risk assessments are essential for identifying these threats, implementing proactive safeguards, and reducing the chances of successful attacks. Importantly, cybersecurity risk assessment for healthcare workstations must be an ongoing process to ensure robust protection of critical data [4].

Contemporary research indicates that numerous patients are unwilling to permit unrestricted access to their complete health data, even in de-identified formats; rather, they prefer to govern the parameters of data access [2]. Personal information protection relies primarily on firewalls and encryption. Firewalls monitor and control network traffic to prevent unauthorized access, while encryption secures data by making it accessible only to authorized users through decryption methods. Both approaches play crucial roles in ensuring data security [5].

The regulatory landscape for medical data management is shaped by the global influence of key regulations. Initial regulations emerged in the mid-1990s, beginning with the European Data Protection Directive 95/46/EC [1]. In the United States, the principle of patient confidentiality emerged as a central tenet of healthcare information technology, leading to the enactment of the Health Insurance Portability and Accountability Act (HIPAA) in 1996 [6]. In 2018, the General Data Protection Regulation (GDPR) took effect, superseding the Data Protection Directive 95/46/EC—the foundation of the Data Protection Act 1998. The GDPR introduced new obligations for organizations handling the data of European Union citizens [7]. Historically, healthcare privacy and security frameworks were largely centered on adhering to these main regulatory frameworks. However, disparities exist between developed and developing countries in the implementation of these standards, which highlights the difficulties in establishing consistent global standards [8].

Recent studies suggest a transition toward more comprehensive strategies. Some authors propose a methodology that incorporates adaptive security measures throughout the lifecycle of medical devices, underscoring the necessity for dynamic security frameworks in contemporary healthcare environments [9]. Furthermore, others emphasize the potential of blockchain technology, particularly its decentralization and immutability features, to bolster data integrity and patient control within healthcare systems [10,11].

Despite the theoretical advantages of these evolving frameworks, their effectiveness is still being assessed. While the reviewed literature extensively discusses potential benefits, there are increasingly more studies pointing out the practical challenges in implementation, especially concerning scalability and regulatory compliance [12]. A significant trend involves the creation of holistic models that tackle technical, governance, and ethical considerations in a unified manner. For instance, Mishra et al. propose a global framework aimed at standardizing security and privacy rules for medical data, utilizing advanced analytical methods like K-means clustering to categorize concepts and prioritize their implementation [13]. Additionally, Lea et al. discuss the concept of data safe havens, emphasizing the crucial role of public engagement and involvement in governance to foster trust [14].

The development of these frameworks also highlights a growing emphasis on patient empowerment and data ownership, which is reflected in the increasing interest in patient-generated health data and the associated privacy and security challenges [15]. Moreover, recent regulatory changes, such as GDPR, have transferred data ownership to patients, thus presenting new challenges in managing access and stewardship of health data [16]. The integration of emerging technologies, such as blockchain, also necessitates compliance with privacy-related regulations like GDPR and HIPAA [10]. However, this is the Western-

centric approach. China's rapid advancement in the digital sphere, coupled with its data localization policies rooted in the concept of "data sovereignty", is fundamentally different, emphasizing state control over data generated within national borders, in a centralized manner [17]. India's data governance model also seeks to increase national government oversight of cybersecurity. However, it adopts a more balanced approach, as reflected in the Digital Personal Data Protection Bill, which shares several features with the GDPR. Despite these similarities, it has received mixed reviews from legal experts. Some critics argue that it grants the government excessive control over consent and privacy matters, potentially undermining the effectiveness of the country's data protection framework [18].

The challenges of ensuring compliance with these regulations in Artificial Intelligence (AI) applications for healthcare are also becoming more and more prominent, stressing the importance of maintaining transparency in AI decision-making processes. Compliance strategies are evolving, with a growing emphasis on privacy-by-design principles and the development of comprehensive governance frameworks [19]. The integration of consent policies, ethics policies, and privacy policies to ensure compliance and trustworthiness in precision health data handling also became of utmost importance [20]. Technological solutions for compliance are also emerging, with Elluri et al. proposing a framework using machine learning and knowledge graphs to extract and manage COVID-19 data in compliance with HIPAA guidelines [21]. However, regulatory frameworks often lag behind technological advancements, as the proliferation of Internet of Medical Things (IoMT) technologies has outpaced policy reforms in some regions, leading to ethical and security concerns [22].

In response to these challenges, a diverse array of emerging technological solutions is being developed. Machine learning methods are valuable in complex situations because they can quickly learn from new information and adjust to unfamiliar problems [23]. This adaptability allows them to handle challenges that traditional approaches might struggle with, making them useful tools for solving difficult and changing issues [24]. Blockchain technology offers the potential for secure and efficient health data sharing through decentralization, trustlessness, immutability, traceability, and transparency [10,11], and it can be integrated with edge computing and machine learning to enhance IoMT security and efficiency [22]. Blockchain is a system where data is grouped into blocks, each marked with a timestamp and linked to the previous block using a unique code, creating a continuous chain of blocks, allowing for the transactions to be verified without relying on a central authority. Scalability is a key issue for blockchain, and resolving it is necessary for broader data management and security improvements [5].

Privacy-preserving technologies, like federated learning and differential privacy, are also gaining traction as promising approaches for maintaining data privacy while enabling collaborative research and AI model training [19]. Federated learning enables multiple institutions to collaboratively train a shared model without exchanging raw data. Each institution downloads a global model, updates it using local data, and transmits only encrypted model updates to a central server. Differential privacy allows for useful data sharing by adding noise to query results, ensuring that personal information remains protected, rather than extracting detailed insights from the data provided. It often focuses on enhancing the privacy features of other technologies, rather than serving as a standalone solution. Since individual technologies like federated learning and blockchain have inherent privacy limitations (vulnerability to central server attacks in federated learning and malicious participant risks in blockchain), combining two or more privacy-preserving technologies is a common practice to mitigate the risk of personal information leakage [5].

Thapa and Çamtepe (2021) propose a "no-peek learning" approach that combines federated learning, split learning, and differential privacy to ensure privacy in precision health

data analytics [20]. Advanced cryptographic techniques, such as homomorphic encryption and secure multi-party computation, have become methods for enabling computations on encrypted data, particularly relevant for genomic data analysis and collaborative research projects [2,25].

IoMT also presents unique security challenges, as many devices are not perceived by patients to be designed with modern security standards in mind. Therefore, concerns about discrimination and misuse by governmental or insurance agencies discourage people from disclosing sensitive health and financial information. Authors advocate for a comprehensive, multilayered security approach based on security-by-design principles [26] or for integrating it with blockchain and AI to enhance data security and enable real-time health monitoring [27]. Similarly, digital twins, which are virtual models that replicate physical systems or processes, are revolutionizing technology by enabling improved optimization, predictive maintenance, and real-time monitoring. However, their use within interconnected environments like the IoMT introduces notable security concerns [28]. Artificial intelligence and machine learning offer potential for enhancing diagnostic accuracy and treatment planning, but also necessitate privacy-preserving AI systems. Machine learning techniques are being applied to improve data de-identification processes and enhance the security of IoT applications in healthcare [2,19,27]. Integrated frameworks, such as the methodology proposed by Almazyad et al. for ensuring data security throughout the medical device lifecycle and the HiGHmed Platform described by Haarbrandt et al., which combines various standards and technologies to enhance interoperability and data security, are also emerging [9,29].

Finally, governance and ethical considerations are receiving increasing attention in medical data management. Stakeholder management involves public engagement and involvement in governance structures for data safe havens [14], while some papers advocate for a patient-centric approach that prioritizes patient empowerment and data ownership [16]. Ethical frameworks for emerging technologies are also crucial, with further research into AI and blockchain applications that respect privacy and ethical principles still needed. The integration of ethical considerations into technological solutions is seen as essential for maintaining public trust and ensuring responsible use of healthcare data [30]. Evolving best practices, such as dynamic consent mechanisms and an "ethics-by-design" approach, are also gaining prominence [20]. Data governance best practices, including the concept of data stewardship, are also gaining prominence, with an emphasis on responsible data management throughout the data lifecycle [16]. However, challenges in implementation persist, and the rapid pace of technological advancement often outpaces the development of appropriate governance structures, necessitating ongoing research and development in these areas [22].

Therefore, the field of medical data protection is characterized by rapid technological innovation and frequent regulatory updates, resulting in a research landscape that is both highly dynamic and fragmented. However, the pace of technological change often outstrips the development of regulatory standards, leading to uncertainties and inconsistencies in both research and practice. This fragmented evolution makes it difficult for stakeholders to gain a comprehensive understanding of the field's trajectory, influential contributors, and emerging gaps. The objectives of this study are to map the academic landscape of data protection in medicine by identifying the main research areas, topics, and trends within the field and analyzing their evolution over time, including key milestones, turning points, and emerging trends driven by technological advancements, regulatory changes, or societal concerns. Additionally, this study aims to identify key contributors, influential journals, and impactful articles that have shaped the field, as well as to map collaborations among researchers, institutions, and countries. Furthermore, it seeks to explore thematic areas and

emerging topics, such as the role of legal frameworks, cybersecurity measures, and ethical considerations in data governance, while uncovering gaps and suggesting future directions for research.

Related Works

In recent years, several bibliometric studies have mapped the evolving landscape of health informatics and data protection research. For example, Qiu and Hu (2025) conducted a comprehensive analysis of 2529 publications from 2000 onward, focusing on data governance and open sharing in the life sciences and medicine, with particular attention to research frontiers such as the FAIR principles, public trust, and data sharing attitudes [31]. Similarly, Abeykoon and Sirisena (2023) restricted their analysis to the decade 2012–2022, examining 757 articles to elucidate trends and collaborative networks in data governance, while highlighting emergent themes like artificial intelligence, blockchain, and data management [32]. Costa et al. (2024) further narrowed their scope, analyzing only 63 articles on system interoperability and data linkage within health information management [33]. Our study advances and refines the bibliometric methodology by specifically targeting data protection within the medical domain, leveraging a significantly broader historical perspective (1960–2024), and systematically interrogating the evolution and impact of technological innovations (such as AI, blockchain, and encryption) alongside major regulatory milestones (including HIPAA and GDPR). This expanded scope enables a more nuanced and longitudinal understanding of the interplay between technological, regulatory, and collaborative dynamics shaping the field of medical data protection.

2. Materials and Methods

This bibliometric analysis was conducted using the Scopus database, aligning with the standard practice of bibliometric research, as each database uses specific metadata formats and indexing points, and combining records from multiple databases can introduce inconsistencies and differences, such as duplicate entries and variations in author names.

The search was performed on 10 March 2025 and included all publications up to 2024. A comprehensive and targeted search formula was designed to capture studies addressing the privacy, confidentiality, protection, and security of medical data (including the most widely used synonyms), and also incorporated topics related to legal frameworks, cybersecurity, patient consent, data governance, and ethical considerations. Bibliometric analysis is a quantitative and descriptive research methodology designed to systematically map the structure, trends, and evolution of the scientific literature within a given field. Unlike hypothesis-driven empirical research, bibliometric studies do not typically require theoretical problematization or narrowly defined research questions. Instead, their primary aim is to provide an objective overview of research activity, identify influential works and contributors, and reveal patterns, gaps, and emerging trends in the literature [34].

Nevertheless, to ensure analytical relevance and interpretability, best practices in bibliometric research recommend articulating a clear rationale and well-defined objectives. In this study, we address the fragmented and rapidly evolving landscape of medical data protection research by mapping its academic development over time, with particular attention to the influence of technological innovations and regulatory milestones.

Our research question is:

How have technological advancements and regulatory milestones influenced the evolution, thematic focus, and collaboration patterns in medical data protection research from 1960 to 2024?

The search formula used was as follows:

TITLE-ABS-KEY ((“medical data” W/15 (“privacy” OR “confidentiality” OR “protection” OR “security*”) OR “healt* data” W/15 (“privacy” OR “confidentiality” OR “protection” OR “security*”) OR “clinical data” W/15 (“privacy” OR “confidentiality” OR “protection” OR “security*”) OR “patient data” W/15 (“privacy” OR “confidentiality” OR “protection” OR “security*”) OR “medical information” W/15 (“privacy” OR “confidentiality” OR “protection” OR “security*”) OR “healt* information” W/15 (“privacy” OR “confidentiality” OR “protection” OR “security*”) OR “clinical information” W/15 (“privacy” OR “confidentiality” OR “protection” OR “security*”) OR “patient information” W/15 (“privacy” OR “confidentiality” OR “protection” OR “security*”) OR “medical recor*” W/15 (“privacy” OR “confidentiality” OR “protection” OR “security*”) OR “healt* recor*” W/15 (“privacy” OR “confidentiality” OR “protection” OR “security*”) OR “clinical recor*” W/15 (“privacy” OR “confidentiality” OR “protection” OR “security*”) OR “patient recor*” W/15 (“privacy” OR “confidentiality” OR “protection” OR “security*”)) AND ((“GDPR” OR “HITECH” OR “HIPAA” OR “ISO 27001” OR “ISO/IEC 27001” OR “cybersecurity”) OR (“patient consent” OR “patient autonomy” OR “informed consent”) OR (“data” W/15 “sharing” OR “access” OR “control” OR “management” OR “law” OR “laws” OR “legislation” OR “legal” OR “regulatio*” OR “governance” OR “polic*” OR “standar*” OR “act” OR “righ*” OR “anonym*” OR “pseudoanonym*” OR “breac*” OR “lea*” OR “de-identification” OR “deidentification” OR “ethic*” OR “bioethics”) OR (“information” W/15 “sharing” OR “access” OR “control” OR “management” OR “law” OR “laws” OR “legislation” OR “legal” OR “regulatio*” OR “governance” OR “polic*” OR “standar*” OR “act” OR “righ*” OR “anonym*” OR “pseudoanonym*” OR “breac*” OR “lea*” OR “de-identification” OR “deidentification” OR “ethic*” OR “bioethics”) OR (“recor*” W/15 “sharing” OR “access” OR “control” OR “management” OR “law” OR “laws” OR “legislation” OR “legal” OR “regulatio*” OR “governance” OR “polic*” OR “standar*” OR “act” OR “righ*” OR “anonym*” OR “pseudoanonym*” OR “breac*” OR “lea*” OR “de-identification” OR “deidentification” OR “ethic*” OR “bioethics”))).

Articles were included if they contained the specified keywords in their title, abstract, or keywords sections (both authors’ keywords and index keywords). Publications without sufficient metadata (e.g., missing author or affiliation information) were excluded. This search yielded 2196 articles, of which 2159 articles contained sufficient metadata for analysis.

The bibliometric analysis was conducted using the Bibliometrix R package (v.4.3.2), based on R (v.4.4.3).

The following bibliometric indicators and analyses were performed: publication and citation trends, keyword analysis, thematic evolution, co-occurrence maps, journal and citation analysis, geographic analysis, and collaboration networks.

The results were visualized using trend graphs, treemaps, thematic maps, co-occurrence networks, bar charts, and tables generated by Bibliometrix. These visualizations provided insights into the intellectual structure of the field, key research themes, and collaboration patterns.

3. Results

The dataset spans the period from 1968 to 2024 and includes 2159 documents from 1175 sources (1.83 papers per journal). The annual growth rate of publications is 11.69%. The dataset contains 4358 unique keywords provided by the authors.

A total of 8044 authors contributed to the dataset, with 150 single-authored documents. The average number of co-authors per document is 4.2, and 22.05% of the publications involve international co-authorship. The average age of the documents is 6.12 years, and the average number of citations per document is 16.82.

Our initial analysis focused on the annual publication rate, as illustrated in Figure 1. The data reveals an exponential growth pattern, characterized by phases of stagnation or minimal growth until the early 2000s. A period of sustained acceleration is then observed, succeeded by a marked increase in the number of publications from 2016 onward. Some particular peaks are noticed in 1996 and 2006, which will be approached in the Discussion Section.

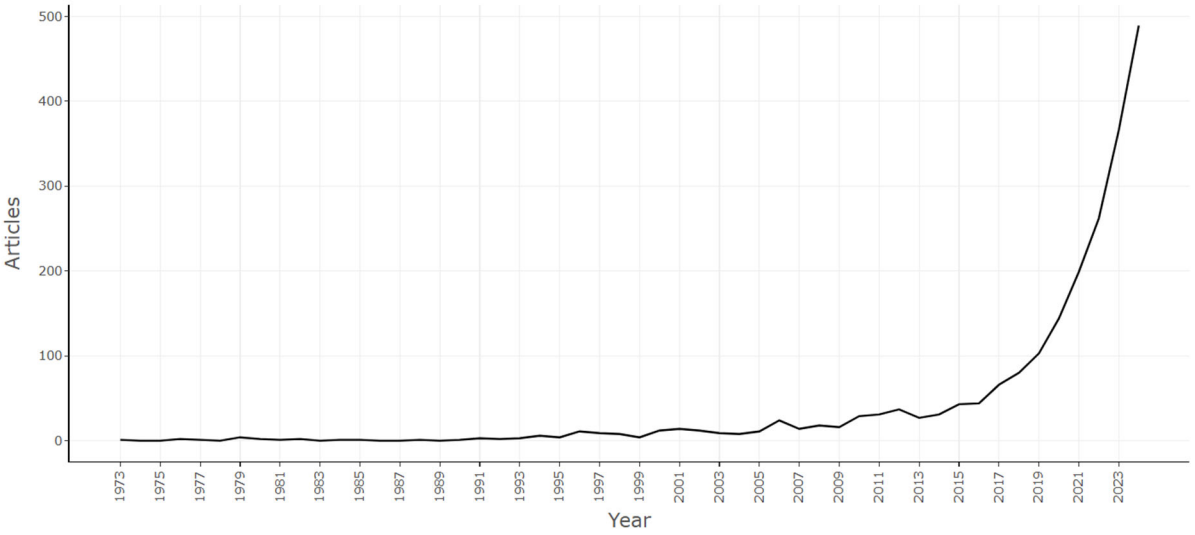


Figure 1. Annual publication rate.

This is intrinsically linked to the total no. of citations (Figure 2), which reflects the interest in and the relevance of a given research output within its field. The trend in citations appears to correlate with the annual publication rate, with notable peaks in citations occurring in the years following peaks in publication rate (as seen in Figure 1), suggesting that impactful publication years have subsequent reverberations. An initial, significant peak is observed in 2002, after which the interest appears to follow a cyclical pattern.

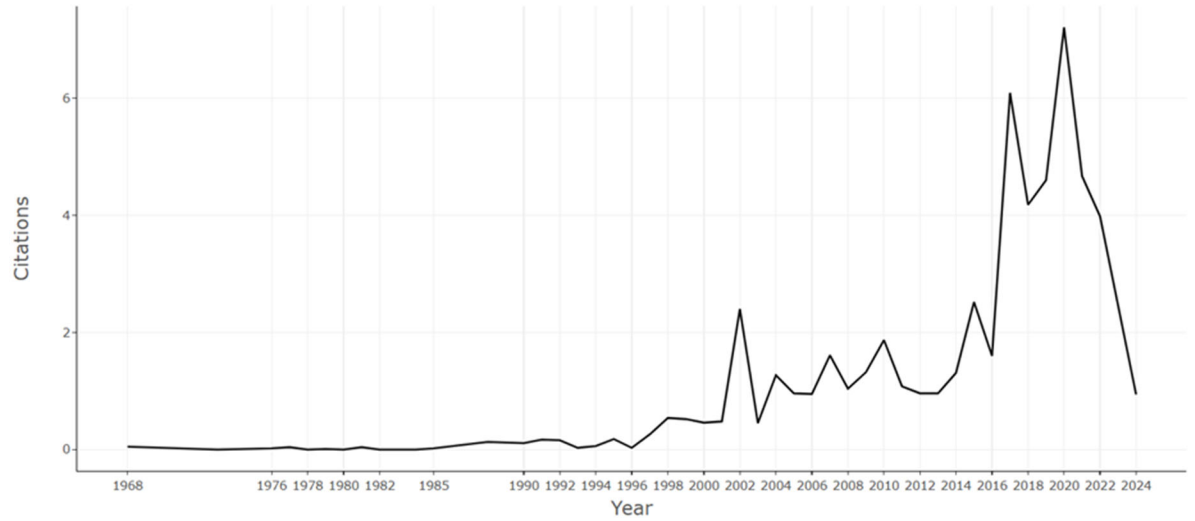


Figure 2. Average no. of citations per year.

Based on the information extracted from Figures 1 and 2, and correlated with historical advancements in healthcare technologies (which will be further discussed in the following section), we analyzed the main trends related to specific time periods: 1960s–1970s,

1980s–1990s, 2000–2015, and 2016–present, by using two-word treemaps sequences extracted from abstracts (Figures 3–6).

Overall, the wordtree data from 1968 to 1979 provides insight into the early stages of medical informatics, highlighting the importance of data management, privacy, and emerging technologies in the medical field. The primary focus during this period was on the collection, management, and application of medical data, as well as the provision of healthcare services.

Compared to the previous period, the wordtree data from 1980 to 1999 shows a significant increase in the frequency of terms related to medical data, healthcare, various "security" concerns, and "risk analysis". This suggests a growing recognition of the importance of information technology in the medical field and an increasing focus on the development of systems for managing and analyzing data, with growing attention on maintaining patient confidentiality and protecting sensitive information.



Figure 3. Treemap of the most widely used 2-word sequences in paper abstracts during 1968–1979.

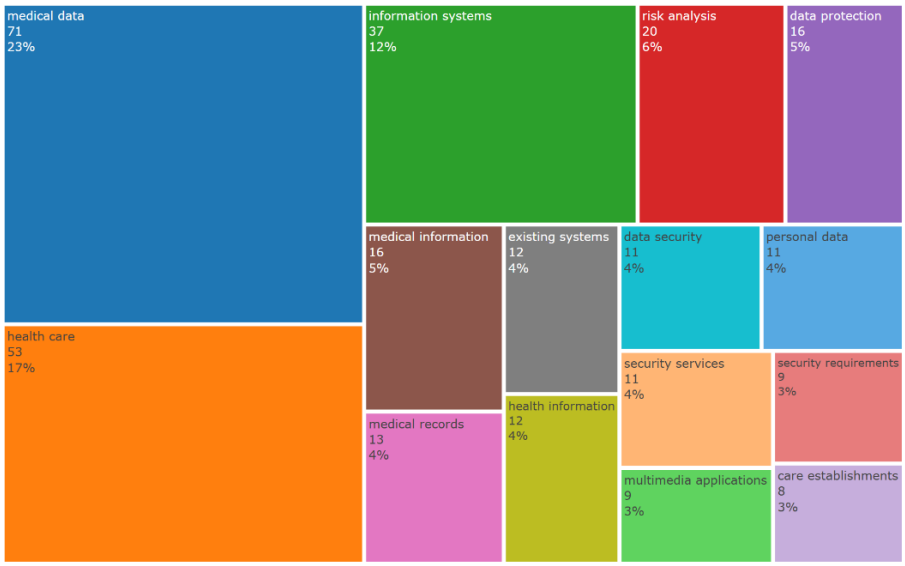


Figure 4. Treemap of the most widely used 2-word sequences in paper abstracts during 1980–1999.

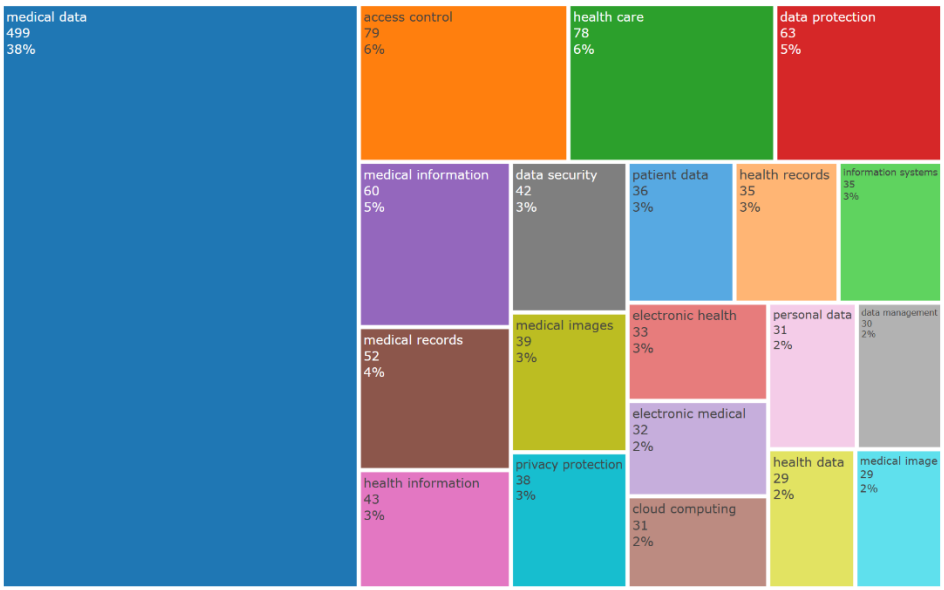


Figure 5. Treemap of the most widely used 2-word sequences in paper abstracts during 2000–2015.



Figure 6. Treemap of the most widely used 2-word sequences in paper abstracts during 2016–2024.

Compared to the previous periods, the wordtree data from 2000 to 2015 shows an even more significant increase in the frequency and diversity of terms related to medical data and data protection, a more extensive representation of electronic health records, and an increasing focus on the development of digital systems.

Finally, the wordtree data from 2016 to 2024 shows a significant increase in the frequency of terms related to data sharing, security, blockchain, and artificial intelligence. This suggests a growing recognition of the importance of secure and decentralized data management, as well as the potential of innovative technologies to transform the healthcare industry.

To provide a comprehensive overview of the landscape across the entire period, we conducted a thematic analysis, incorporating trigrams extracted from abstracts to capture nuanced conceptual relationships (Figure 7). A thematic map of this sort reveals the intellectual structure of the field, highlighting key areas of research and their interrelationships. "Motor themes" represent the driving forces of the field; in our case, they suggest a primary

focus on the practical challenges of making medical data accessible while maintaining security standards. "Basic themes" form the foundational knowledge and are centered on "electronic health records," confirming them to be the fundamental building block of the domain. "Niche themes" indicate specialized areas of intense investigation, and are centered around artificial intelligence; "emerging or declining themes" suggest areas of evolving interest. The themes clustered around the center of the map act as a bridge between the different quadrants. They represent core concerns that are relevant to all areas of the field and will be commented on in the Discussion Section.

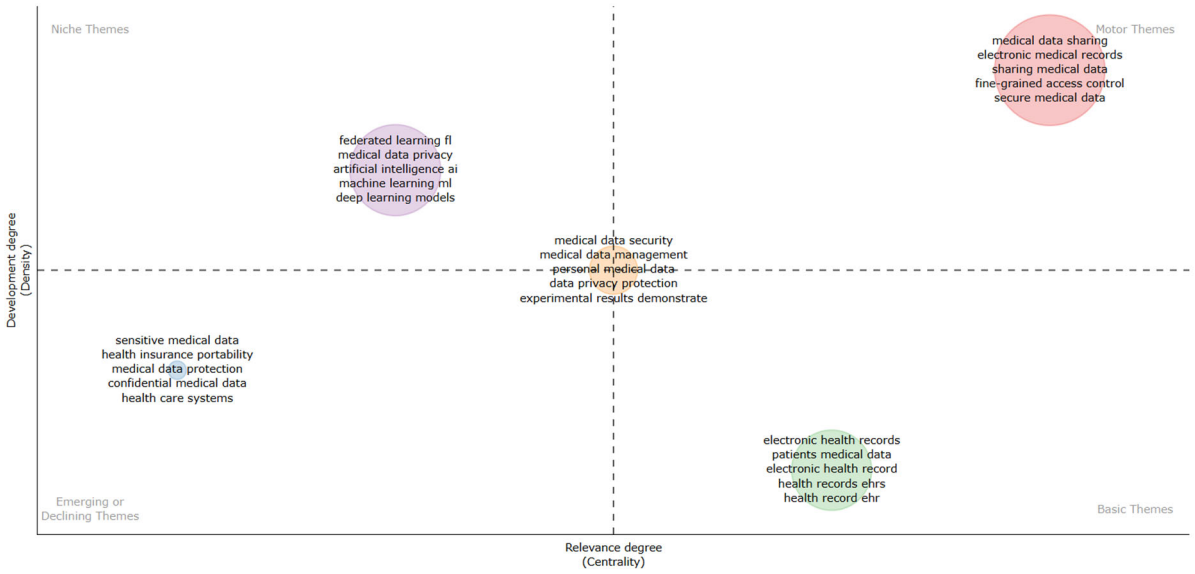


Figure 7. Thematic map on the most prominent topics.

A comparative analysis of the author-assigned keywords (Figure 8) and indexed keywords (Figure 9) reveals a nuanced landscape of thematic priorities.

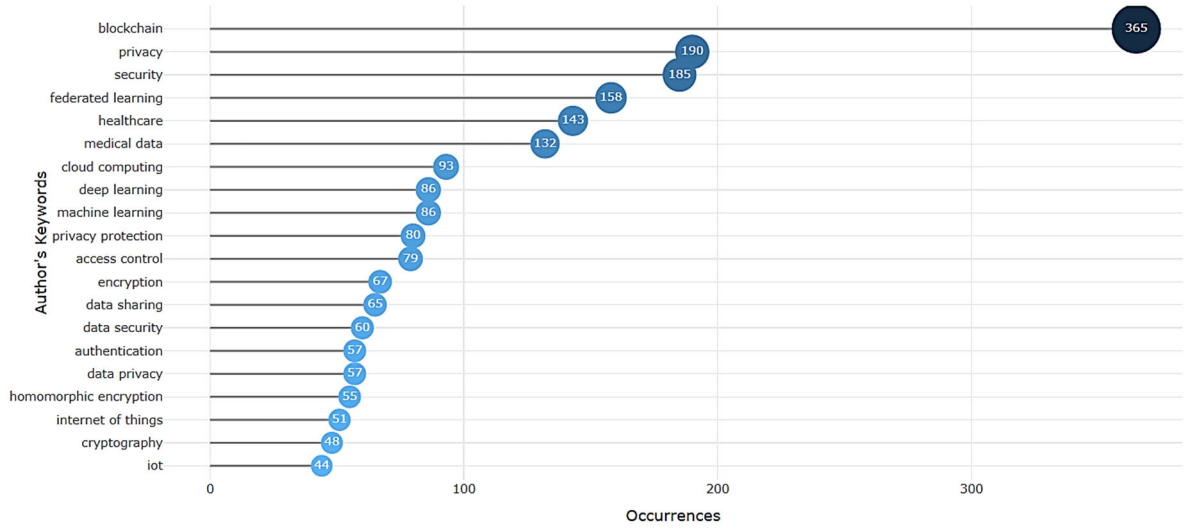


Figure 8. Most frequent author keywords.

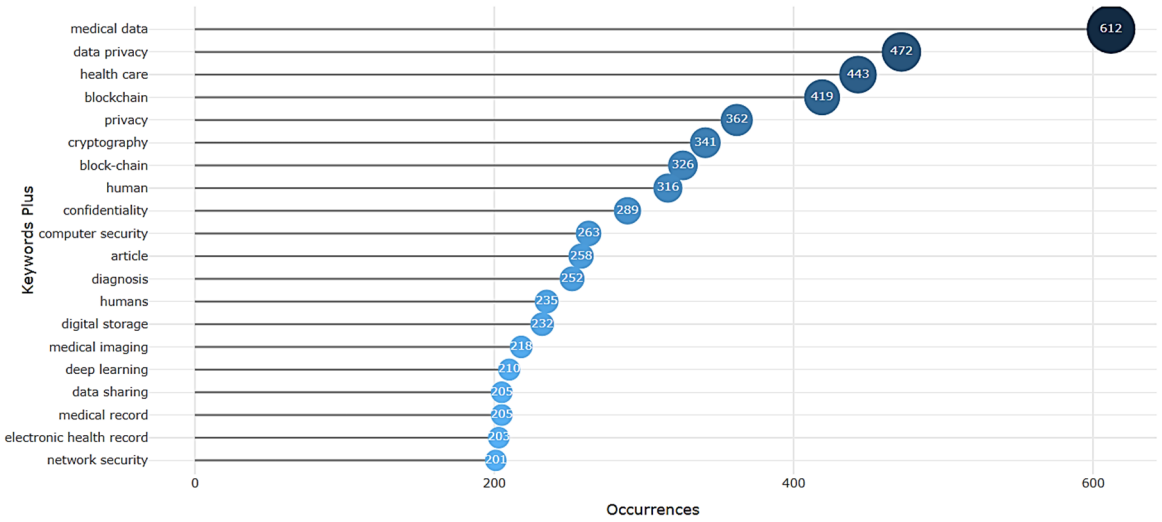


Figure 9. Most frequent indexed keywords.

While both keyword sets exhibit a pronounced focus on "blockchain", "privacy", and "security", the indexed keywords demonstrate a more pronounced emphasis on medical domain-specific terminology, such as "medical data", "medical imaging", and "electronic health record". In contrast, the author-assigned keywords display a broader thematic scope, encompassing "federated learning", "deep learning", and "cloud computing", which suggests a more expansive consideration of technological applications in the medical domain.

Regarding trend topics over time, we took into consideration three authors' keywords/year, appearing at least 10 times/year (Figure 10). Given the sparse literature of the previous century, trends in the authors' interests could only be identified after 2005. The analysis revealed a chronological progression of the main themes in research surrounding data protection in medicine. The earliest and most longstanding topic is that of "confidentiality", spanning the entire timeline; yet, "privacy" and "security" became prominent after 2020. "Telemedicine", then "e-health", and now "internet of medical things" became some of the most discussed topics of the last decade, and so are "pseudoanonymization", then "cloud computing", and then both "blockchain" and "blockchain technology". There is a clear evolution from basic data handling and security concerns to the application of advanced technologies like blockchain, federated learning, and IoMT, all while maintaining a strong focus on privacy and security. The graph illustrates the increasing complexity and sophistication of research in this field over time.

The network appears to be divided into three main clusters, indicated by the red, blue, and green. This suggests three relatively distinct, yet interconnected, areas of research. The red cluster is heavily centered around "blockchain," "medical data," and "healthcare." This indicates a strong research focus on applying blockchain technology to manage and secure medical data within the healthcare sector. Related keywords include "smart contracts," "privacy protection," "Internet of Medical Things (IoMT)," and "edge computing," suggesting exploration of decentralized, secure, and connected healthcare solutions. The blue cluster revolves around "privacy" and "security," which are strongly interconnected. This cluster includes terms like "data security," "encryption," "confidentiality," "e-health," and "telemedicine." This suggests a focus on the fundamental aspects of data protection and secure communication in healthcare. The green cluster is centered around "federated learning," "artificial intelligence," and "privacy-preserving" techniques. This indicates a research area focused on advanced analytics and machine learning methods that prioritize data privacy. Connections to "medical imaging" and "COVID-19" suggest applications in specific healthcare domains.

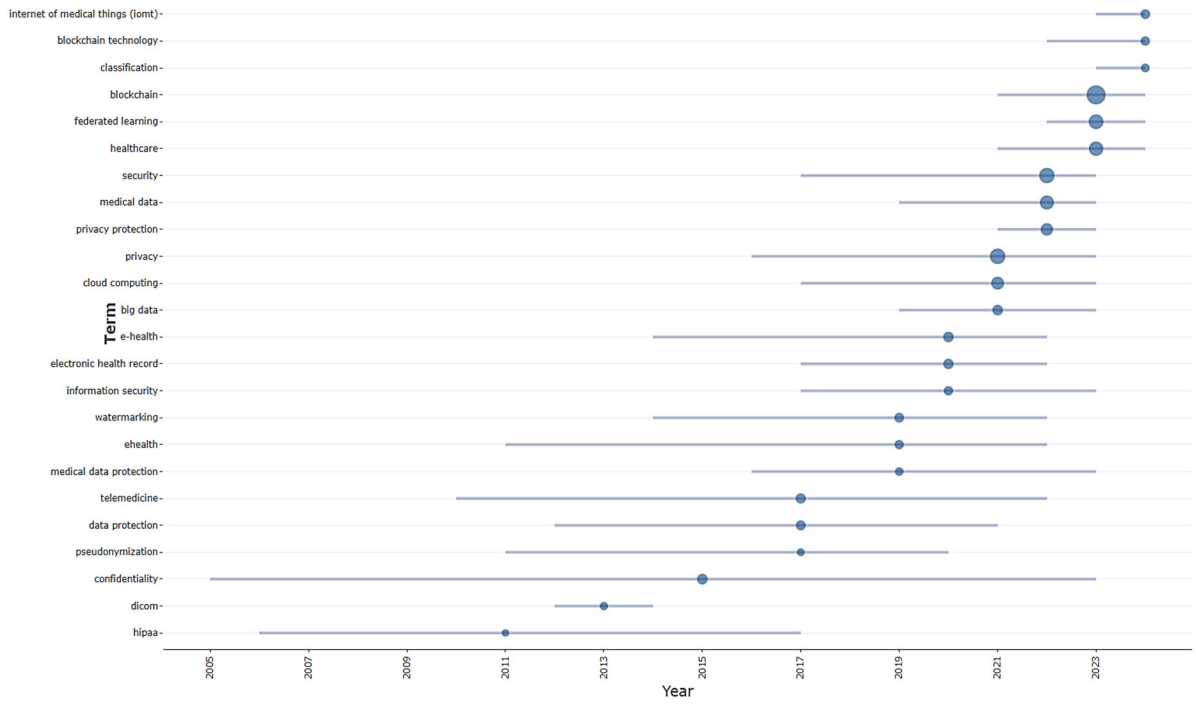


Figure 10. Trend topics in authors' keywords.

These observations are supported by the co-occurrence network of author keywords (Figure 11).

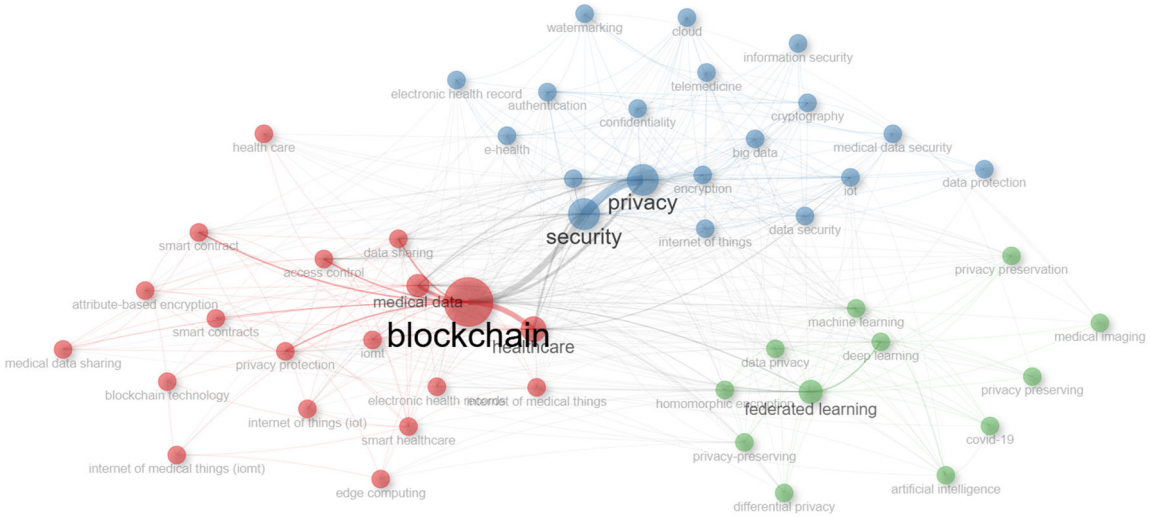


Figure 11. Co-occurrence network of author keywords.

To evaluate the influence of the research within the field, we investigated the institutions and journals (Table 1 and Figures 12 and 13) that have been the most prolific. However, owing to variations in reporting practices, we chose not to include the institutions, as the second most reported was identified as the broad category of "Department of Computer Science and Engineering," and the top 15 also encompassed "Not reported." This highlights the critical need for enhanced standardization in the reporting of research findings. In exchange, we analyzed the scientific output of countries (Figure 14), as well as the collaboration network among them (Figure 15).

Table 1. The most relevant journals in the field, by the total no. of publications.

Sources	Articles
Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	80
IEEE Access	47
Studies in Health Technology and Informatics	43
ACM International Conference Proceeding Series	35
Communications in Computer and Information Science	23
IEEE Internet of Things Journal	22
Lecture Notes in Networking and Systems	21
Electronics (Switzerland)	20
Journal of Medical Systems	20
Future Generation Computer Systems	17

Table 1 displays the journals that have published the most articles related to data protection and governance in medicine within the dataset. The presence of both technical and medical journals underscores the interdisciplinary nature of data protection and governance in healthcare. "Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)" is the most prolific source, with 80 articles, which is a significant portion but still only accounts for approximately 3.7% of the total 2159 papers in the dataset. Conference publications are a significant outlet for research in this area, meaning that the field is rapidly evolving, with researchers using scientific gatherings to disseminate and discuss their latest findings quickly, before they are formalized in journal articles.

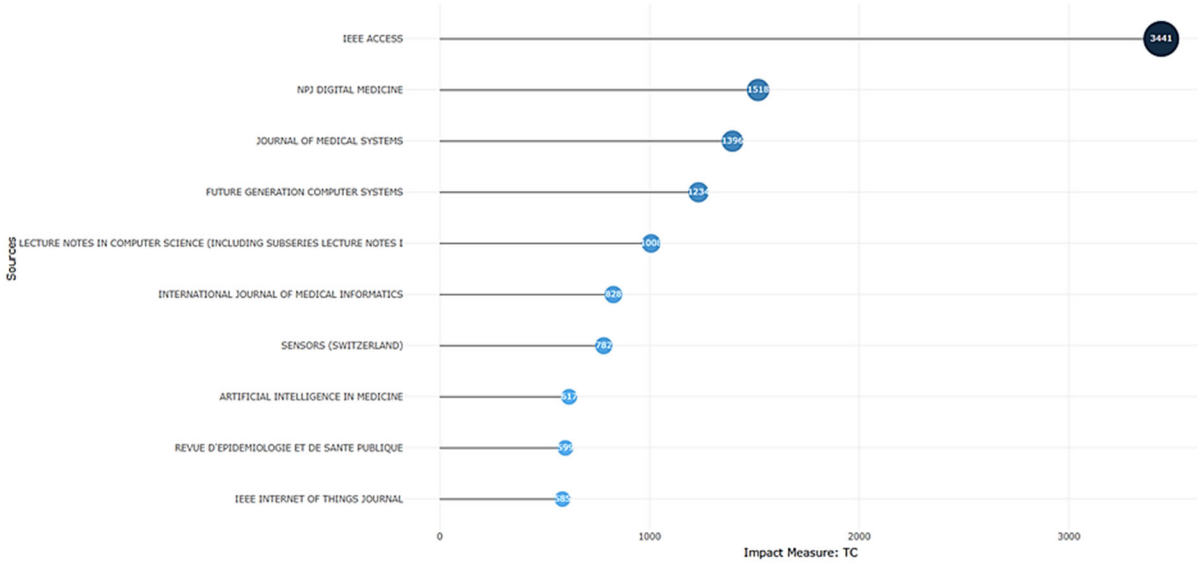


Figure 12. The most relevant journals in the field, by the total no. of citations.

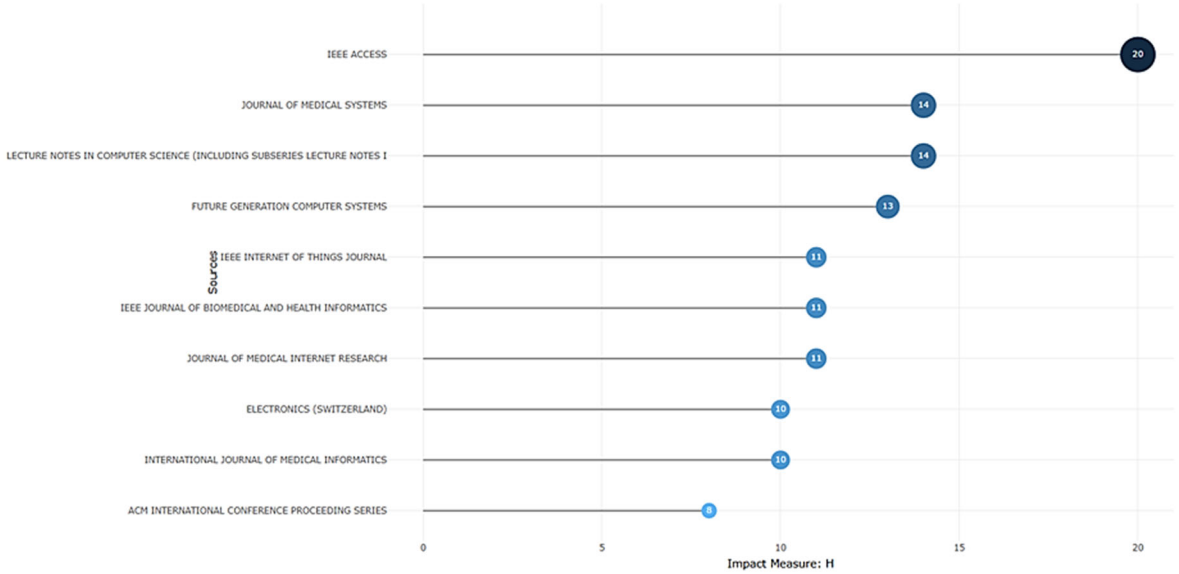


Figure 13. The most relevant journals in the field, by the H-index.

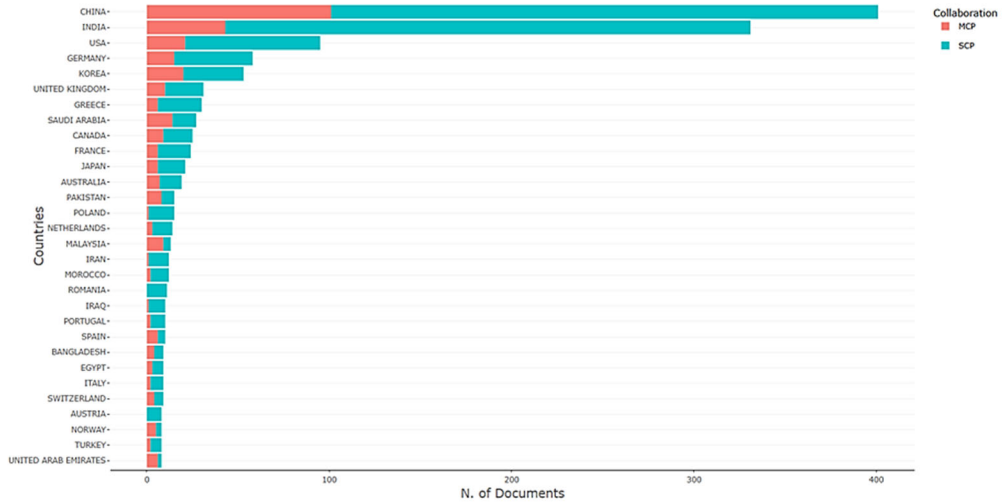


Figure 14. Countries with the highest no. of publications. MCP: Multiple-Country Publications; SCP: Single-Country Publications.

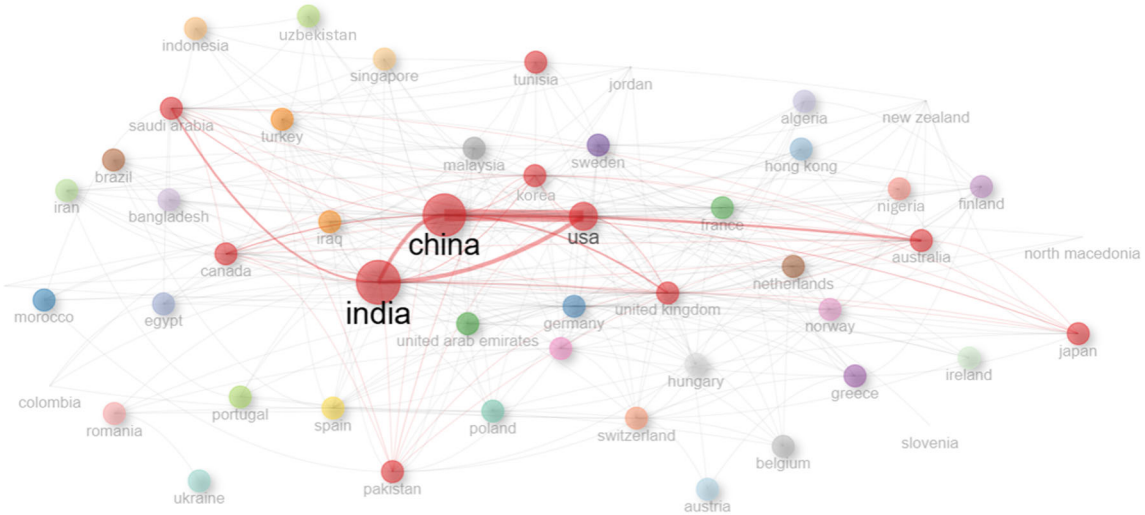


Figure 15. Cross-country collaboration network.

Both Figures 12 and 13 show IEEE Access as the leading journal in terms of total citations and H-index. This suggests that it is a highly influential and impactful publication venue in the field of data protection and governance in medicine. Total citation numbers reflect the overall impact of a journal's publications, while the H-index measures both the productivity and citation impact of the journal. A high TC but lower h-index might indicate that a journal has published a few highly cited articles, but its overall body of work is not as consistently impactful. A high H-index suggests a more sustained and consistent level of influence. Some journals appear high on one list but not the other, suggesting different strengths, further discussed in the following section.

In terms of total numbers, China is by far the most prolific country, followed by India and the USA (Figure 14). For most countries, the majority of publications are single-country publications (SCP). This suggests that a significant portion of the research is conducted within national boundaries, without international collaboration. The proportion of multiple-country publications (MCP) varies across countries. Some countries, like the USA and the UK, have a relatively higher proportion of MCP compared to others like China and India. Analysis of the co-authorship network (Figure 15) provides a nuanced understanding of international collaborations in data protection and governance in medicine. The network reveals a triad of influence, with China, India, and the USA forming a highly interconnected core. Beyond this central group, the network also includes more peripheral clusters, indicating that some countries are operating with a greater degree of autonomy in their research endeavors.

We included an overview of the most cited studies, by the total number of citations (Figure 16) and by the articles' average citations (Figure 17), which was conducted in relation to the countries of origin, with the aim of pinpointing key periods in the field's progression and the most influential research hubs globally.

China leads by a wide margin in total citations, indicating a massive overall research output in this field (Figure 16). When considering average citations (Figure 17), countries like Egypt, Canada, and Sweden rise to the top, while China falls out of the top rankings. Several less influential countries, such as Tunisia and Portugal, have high average citation rates, suggesting that they are producing high-quality research that is having a significant impact, even if their overall publication volume is limited. The USA performs well in both metrics.

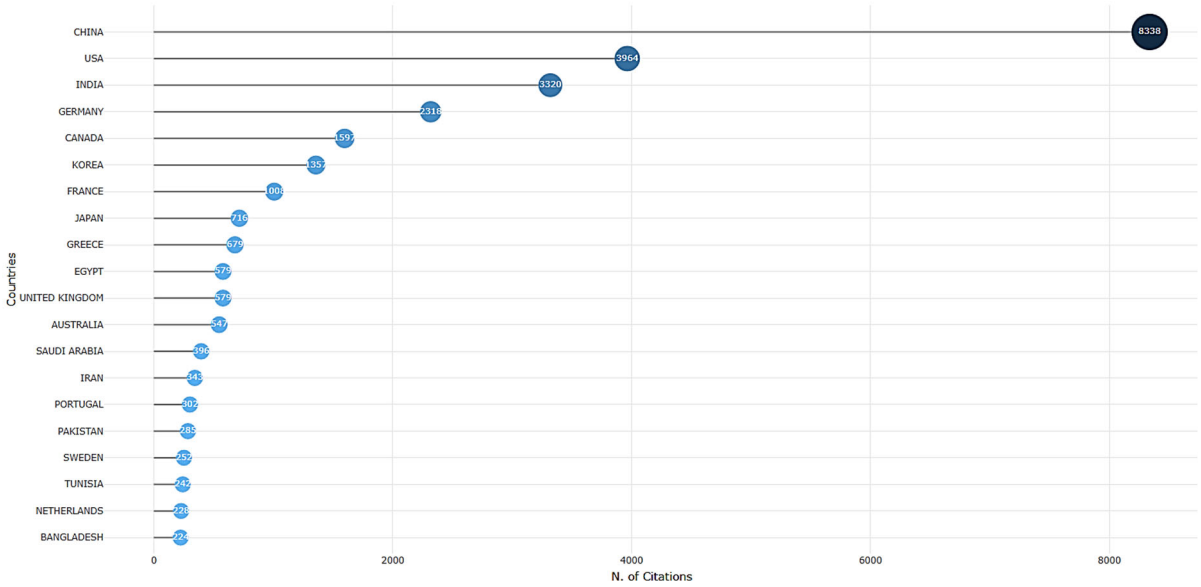


Figure 16. Most cited studies, by the total no. of citations and country of origin.

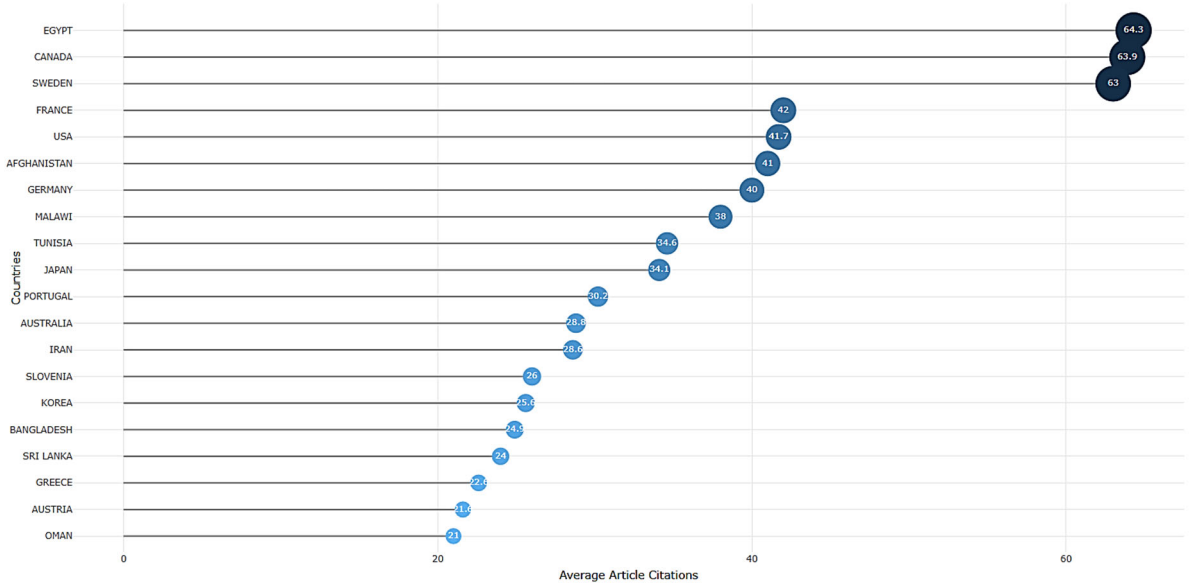


Figure 17. Most cited studies, by the average no. of citations and country of origin.

4. Discussion

The potential of computers to improve healthcare began to be discussed in the 1960s. It was hypothesized that physicians could use computers to quickly retrieve test results or to search the medical literature. Furthermore, these systems had the potential to decrease medical errors by providing alerts and reminders and aiding in clinical decision-making. Despite the potential advantages, doctors did not embrace this innovation due to several pragmatic concerns: the technology of the time was costly, slow, cumbersome, and unreliable. Moreover, medical professionals at the time prioritized their independence and were not very interested in structured decision support systems. Additionally, healthcare administrators were wary of investing in technologies that lacked a clear financial return [6]. This is reflected by the development of the scientific literature regarding data protection in medicine, with no publications on the subject predating the 1960s. The first paper on the topic is titled "A Legal Structure for a National Medical Data Center" and is authored by Roy N. Freed. It was published in the proceedings of the 1968 Fall Joint Computer Conference (AFIPS '68) [35]. This appears to be one of the early works discussing legal frameworks and proposed structures for handling medical data at a national level, published at a time when computerization of healthcare information was still in its early stages, which makes it a pioneering piece of work in the field of medical data protection and governance.

After several years of silence, the next decade came with additional topics. The following studies started to focus on the development and implementation of the data bank systems in healthcare in different countries. Karlsson's work emphasized the importance of secrecy and confidentiality in a Swedish medical data bank, addressing the ethical and technical measures required to protect sensitive medical information [36]. On the other hand, Goldberg and Doyon described SARI, a user-oriented data bank system designed to enhance the accessibility and management of medical data for healthcare professionals [37]. Both studies underscored the critical role of databases in improving healthcare information systems, with Karlsson focusing on security and ethical considerations, while Goldberg and Doyon prioritized usability and functionality.

During 1968–1979, only eight papers were published. The most widely used two-word sequences in the abstracts from the 1960s–1970s (Figure 3) likely reflect the emerging themes and priorities of that era in medical informatics. They highlight the early focus on

developing centralized systems for storing and managing medical information ("medical data", "database"), driven by the advent of computer technology, and highlighted by the phrase "medical applications" (presumably, of computers). However, concerns about "medical privacy" have been noted from the very beginning, yet the legal and bioethical terminology surrounding it had not yet been coined, thus having been referred to under the umbrella of "citizen rights".

Moving into 1980–1999, the emphasis shifted toward "information systems", alongside "risk analysis", "data protection" and "data security", "security services" and "security requirements", indicating both technological advancements and the growing importance of securing sensitive information (Figure 4). Historically, this period marked the transition from paper-based records to digital systems, as healthcare institutions began exploring the potential of computers to improve data storage, retrieval, and analysis. The emergence of terms like "medical records" reflects the increasing adoption of such electronic systems. This created a demand for a healthcare data interchange protocol, resulting in the creation of Health Level Seven (HL7). HL7 provides standards for the electronic exchange of clinical, financial, and administrative information between healthcare computer systems, improving interoperability. Despite evidence that information systems could improve results and lower costs, medical practices had limited funds to invest in them [6]. As hardware became more affordable, powerful, and compact in the late 1980s and early 1990s, the growing demand for personal computers, local area networks, and the internet streamlined access to medical information, leading to the early implementation of web-based Electronic Health Records (EHRs) in academic medical institutions. However, several factors impeded the widespread adoption of EHRs, including substantial costs, data entry errors, initial resistance from physicians, and a lack of compelling incentives. Consequently, EHRs were initially intended to supplement, rather than supplant, traditional paper records [38]. Moreover, the first peak in the annual publication rate (Figure 1) is noticed in 1996, which probably reflects the implementation of the European Data Protection Directive 95/46/EC in 1995 and HIPAA in 1996 [1,6].

From 2000 to 2015, the abstracts demonstrate an ever more significant focus on "medical data" (Figure 5). This could be related, in part, to the steady development of Big Data, starting from the 1990s but gaining significant traction from 2012 onward [39], as it plays an important part in research and medicine. Emerging themes such as "access control", or the combinations of "electronic health" or "electronic medical" and "health records" and "medical records" signaled the widespread adoption of EHRs and the challenges of managing access and privacy in digital environments. Technologies such as "cloud computing" started gaining attention in the scientific community. In addition, this time period demonstrated a noticeable focus on medical images, which probably corresponds to the integration of picture archiving and communication system (PACS) with radiology information systems (RIS), a type of imaging-specific EHR that greatly improved the workflow of radiologists [40].

In addition, an isolated citation peak is observed in 2002 (Figure 2), before research in the field became more prominent. This could be explained by discussions over HIPAA regulatory changes. After proposing revisions in March 2002, the Department of Health and Human Services published the modified "Privacy Rules" in August 2002, which introduced significant changes compared to the original regulations [41]. Particular peaks in the annual publication rate are also noticed in 2006 (Figure 1). Although no singular trigger could be identified, the peak could indicate a responsive adjustment to growing anxieties regarding data security and privacy in healthcare. It is plausible that researchers and policymakers were becoming more cognizant of the potential vulnerabilities to data breaches and the necessity for enhanced data protection protocols. Another perspective suggests that this

increased activity may have been prompted by a sense of urgency and a motivation to be perceived as proactively addressing the problem, although the sustained commitment to these solutions warrants further examination.

Finally, in the period from 2016 to 2024, the focus expanded further with "medical data" remaining central, while new terms like "data sharing", "federated learning", "blockchain technology", and "machine learning" reflect the rise of advanced technologies in healthcare (Figure 6). This period highlights the shift toward leveraging artificial intelligence, distributed systems, and blockchain for secure data sharing and analysis, alongside continued attention to "data privacy" and "privacy protection". The increased frequency of these terms reflects not only technological innovation but also a research community responding to new challenges and opportunities in data security, privacy, and data sharing. Thus, the evolution of publication patterns is closely intertwined with the adoption and development of key health information technologies. Consequently, the research landscape becomes increasingly diversified, with emerging topics and specialized subfields reflecting the complexity of these innovations. This diversification leads to less homogeneous publication trends, as the focus of research shifts to address a broader array of technical, ethical, and practical issues introduced by these technologies.

Overall, the evolution of abstract wording across these periods demonstrates a progression from foundational computerization and privacy concerns to advanced, technology-driven solutions for managing and securing medical data in increasingly interconnected healthcare systems. The trajectory of the medical data domain reflects a complex interplay between technological advancements, changing societal values, and the evolving needs of healthcare systems, highlighting the need for ongoing innovation, collaboration, and vigilance in ensuring the responsible and effective use of medical data.

This has also been highlighted by the thematic map analysis (Figure 6), painting a picture of a field that is actively working to balance the competing goals of data accessibility, security, and privacy. Interestingly, the presence of "experimental results demonstrate" in the center of the graphic suggests an emphasis on empirical validation and evidence-based approaches. The "motor themes" quadrant, dominated by "medical data sharing," "electronic medical records," and "fine-grained access control", indicates that the field is actively grappling with the tension between data utility and data protection. The emphasis on "fine-grained access control" implies a move beyond simple all-or-nothing approaches to data access, toward more sophisticated mechanisms for controlling who can access what data and under what conditions.

The "basic themes" are dominated by "electronic health records", synonyms and eponyms, and "Big Data" does not appear yet in the graphic. However, the "niche themes" quadrant features include "federated learning" and "artificial intelligence" as emerging areas of intense research activity (which are intimately linked to Big Data). The combination of AI and privacy suggests a growing interest in developing privacy-preserving AI techniques for healthcare applications.

Interestingly, the presence of "experimental results demonstrate" suggests an emphasis on empirical validation and evidence-based approaches. Also, the limited representation of patient empowerment, data control, and perspectives on data sharing suggests a potential overemphasis on technical aspects of data management. Further consideration is needed to ensure patients are viewed as active participants, not merely subjects, within the healthcare ecosystem.

On the other hand, a critical examination of the disparity between the author-assigned keywords (Figure 6) and indexed keywords (Figure 7) reveals a potential disconnect between the explicit intentions of researchers and the implicit biases embedded in their work. The pronounced emphasis on medical data and electronic health records in the

indexed keywords may suggest a surreptitious prioritization of data commodification and surveillance over the more altruistic goals of privacy and security. The fact that "human" and "humans" appear as indexed keywords, yet are absent from the author-assigned keywords, may also imply a troubling lack of consideration for the social and ethical implications of these technologies on vulnerable populations. Ultimately, this comparison highlights the need for a more nuanced and critically informed approach to the development and implementation of emerging technologies in healthcare, one that acknowledges the complex interplay between technological innovation, economic interests, and social justice. The differential ranking and frequency of keywords, such as blockchain and privacy, between the two sets may indicate a divergence in priorities between the authors' explicit categorizations and the implicit thematic structures embedded in the indexed keywords. This discordance underscores the complexity of keyword extraction and the need for a multifaceted approach to understanding the thematic topography of medical data research, particularly in the context of emerging technologies like blockchain and artificial intelligence.

The trend analysis of author-provided keywords over time (Figure 8) offers a unique glimpse into the evolving priorities and concerns of researchers in the healthcare technology domain. The fact that these keywords are provided by authors themselves lends significance to this analysis, as it reflects their self-identified areas of focus and expertise. Notably, cohesive trends before 2005 could not be identified, as research in this area was too limited, and the field was too broad. "Confidentiality" represents a foundational and enduring concern, consistently present throughout the timeline, and completed by "privacy" and "security" as prominent themes after 2020, reflecting increasing awareness and discussions surrounding data protection challenges in modern healthcare, possibly a reactive response to escalating data breaches and privacy controversies.

The evolution of technology is mirrored in the rise of "telemedicine," followed by "e-health," and now the burgeoning interest in the "Internet of Medical Things," indicating a shift toward remote and interconnected healthcare solutions. However, it might also highlight a tech-driven hype cycle, potentially outpacing practical implementation and security considerations. Similarly, "pseudoanonymization" gained traction, succeeded by "cloud computing" and subsequently, the dual emergence of "blockchain" and "blockchain technology," highlighting the exploration of advanced techniques for data security and decentralized systems in healthcare. These trends collectively illustrate a dynamic field adapting to technological advancements while grappling with persistent and emerging data governance challenges. This trend raises concerns about whether research is genuinely leading the way in responsible innovation or simply chasing the latest buzzwords while core vulnerabilities persist. Furthermore, the inclusion of HIPAA in all legislative frameworks suggests either an American-centric approach to the topic or a lack of concern from other world regions.

The significance of these author-provided keywords lies in their ability to reveal the collective mindset and priorities of the research community, providing a nuanced understanding of the field's evolution and future directions. By examining these keywords, we gain insight into the areas where researchers are focusing their efforts and where gaps in knowledge and expertise may exist, ultimately informing strategies for advancing healthcare technology research and development. In the co-occurrence network of author keywords (Figure 9), the connections between the clusters indicate relationships between these research areas. However, the distinct clustering of the network could indicate a lack of true interdisciplinary collaboration. The network may reflect echo chambers rather than genuine integration of diverse perspectives. For example, the connections between "blockchain" and "privacy/security" suggest an interest in using blockchain to enhance

data protection. On the other hand, the prominence of "blockchain" might indicate a rush to apply the technology to healthcare without fully considering its practical limitations or whether it truly solves existing problems better than established methods. The strong connection to "medical data" could suggest an overemphasis on technological novelty rather than addressing fundamental data governance issues. While "privacy" and "security" form a central cluster, their separation from the core technological innovations (like blockchain) could imply that they are often treated as secondary considerations, addressed only after the technology is developed rather than being integrated from the outset. This raises concerns about the ethical implications of deploying potentially vulnerable systems. The "federated learning" cluster, while promising, might represent an idealized vision of AI in healthcare that overlooks the practical challenges of data heterogeneity, algorithmic bias, and the difficulty of ensuring true privacy in real-world deployments. The connection to "COVID-19" could suggest a reactive application of AI to address immediate crises, potentially overshadowing long-term strategic considerations. It is worth considering the perspectives that may be underrepresented in the network. Specifically, do the keywords adequately reflect the concerns of key stakeholders such as patients, clinicians, and policymakers? The absence of certain terms could indicate areas where the research agenda might benefit from broader input and a more comprehensive consideration of diverse needs.

The fact that the dataset encompasses over five decades and comprises 2159 documents from 1175 sources, thus yielding an average of 1.83 papers per journal, indicates that the field is still relatively fragmented, lacking a small set of core journals. The most prolific journal ("*Lecture Notes in Computer Science*") accounts for only 80 out of 2159 papers (approximately 3.7%), suggesting that research in data protection and governance in medicine is highly dispersed across a wide range of publications (Table 1). This could indicate that the field is interdisciplinary, drawing contributions from various areas such as computer science, engineering, health informatics, and medical systems. The diversity of sources could also reflect the emerging nature of the field, where research is not yet concentrated in a few core journals. However, it could be interpreted as a failure of the academic publishing system to adequately recognize and support this specific area or a lack of coordinated effort. Security breaches primarily impact individuals rather than the entire system, and so do the bioethical dilemmas. The focus on technical journals and the lack of representation of humanistic approaches in the field could indicate a potential overemphasis on technological solutions without sufficient consideration of the ethical, legal, and social implications.

The consistent dominance of IEEE Access in both total citations and H-index, despite publishing approximately half the number of articles compared to "*Lecture Notes in Computer Science*", could be attributed to its open-access model (Figures 12 and 13). This could be a significant factor in its high impact, suggesting that open access publishing may be particularly beneficial in this interdisciplinary field, as it makes research more readily available and accessible to a global audience. A concern persists regarding the potential for researchers to prioritize the quantity of conference papers over the quality of research, possibly driven by a desire to inflate publication metrics. This raises questions about whether the peer-review processes associated with these proceedings are sufficiently stringent to guarantee the validity and influence of the published findings.

The presence of journals like "*Artificial Intelligence in Medicine*" and "*International Journal of Medical Informatics*" suggests that specialized journals play a role in disseminating research on specific aspects of data protection and governance in medicine. However, their relatively lower rankings compared to more general journals like "*IEEE Access*" could indicate that research in this area is often integrated into broader discussions of technology and healthcare, rather than being confined to niche publications.

Regarding publication metrics (Figure 14), China and India have emerged as significant contributors to research concerning data protection and governance in medicine. Research approaches vary across nations, with some, such as China and India, prioritizing domestic initiatives, while others, including the USA and UK, emphasize international collaborative efforts. This could be driven by national policies and funding initiatives aimed at promoting scientific advancement within China. It might also reflect a degree of self-sufficiency in terms of research infrastructure and expertise. Different governance models influence not only domestic research priorities but also the extent and nature of international collaboration. Furthermore, global imbalances in knowledge production are shaped by disparities in funding, facilities, access to journals, and the dominance of Western-centric citation metrics. Furthermore, potential biases introduced by language barriers and the prevalence of regional databases may skew the results, as exemplified by the underrepresentation of Russia. Addressing these contextual factors is essential for a nuanced interpretation of publication and collaboration patterns across countries.

The USA and the UK, while having fewer publications, show a higher proportion of MCP. This suggests that these countries are active hubs for international research collaboration, attracting researchers from around the world and participating in joint projects. Some smaller countries show a relatively higher proportion of MCP compared to their overall publication count. This could indicate that these countries rely on international collaboration to conduct research in this field, potentially due to population and geographic factors, along with strong research networks between institutions and high mobility of researchers in the Western sphere. Interestingly, Romania appears to have an SCP-only publishing model, which could be attributed to a smaller scientific infrastructure and funding.

Despite China and India primarily functioning as independent research hubs, the international collaboration network demonstrates their increasing leadership in the field (Figure 15). The strong connection between China and India suggests a significant level of collaboration between these two research powerhouses. This could be driven by shared interests, geographical proximity, or policy initiatives aimed at fostering collaboration between developing countries. The USA's strong connections to both China/India and Europe underscore its role as a bridge between these different research communities. The cluster of European countries suggests a strong regional collaboration network. This could be driven by EU funding programs, shared research priorities, and cultural and linguistic proximity.

However, China's focus may be on quantity over quality (Figures 16 and 17), with a strategy of producing a large volume of research to establish itself as a global leader in this field. The lower average citation rate suggests that there may be room for improvement in the quality and impact of its research. The high average citation rates in some smaller countries might be due to a focus on niche areas of expertise, where they have a competitive advantage, and the lower citation rates in some countries could be a direct result of funding disparities and brain drain. Citation metrics may inherently favor research deemed "relevant" by the dominant scientific community, which is often Western-centric, potentially introducing bias.

The ability to conduct and publish high-quality research is not equally distributed across countries. Researchers in developing countries often face significant challenges in terms of funding, infrastructure, access to journals, and opportunities for collaboration. The lower average citation rates in some countries may reflect these systemic inequalities, rather than a lack of talent or expertise. Both total and average citations are proxies for "impact," and "impact" itself is a complex and multifaceted concept. Citations do not tell us why a paper is cited. Is it cited because it is groundbreaking, or because it is flawed but influential? Is it cited because it is widely read or because it confirms existing biases?

Several key directions for future research and development emerge:

- Integration of ethical, social, and human perspectives: there is a clear need for future studies to more centrally address these issues in medical data protection, including patient empowerment, equity, and the societal implications of emerging technologies;
- Comparative analysis of data governance models: further research should examine how different national and regional data governance frameworks, particularly in rapidly growing research hubs like China and India, influence research priorities, collaboration patterns, and the practical implementation of data protection measures;
- Addressing global imbalances: efforts should be made to mitigate global imbalances in knowledge production and citation impact, fostering greater equity and inclusivity in the international research landscape.

5. Conclusions

This paper provides a comprehensive overview of the intellectual landscape of data protection in medicine, highlighting key themes, influential research hubs, and the ongoing evolution of the field. This analysis sets the stage for further discussions on the implications of these trends for future research and practice in healthcare data management.

The evolution of computer technology in healthcare has undergone significant transformations since the 1960s, reflecting a complex interplay between technological advancements, ethical considerations, and the evolving needs of healthcare systems. Initially met with skepticism due to concerns over cost, reliability, and the desire for physician autonomy, the integration of computers into healthcare has gradually progressed, particularly with the advent of electronic health records (EHRs) and data protection regulations like HIPAA. The focus has shifted from foundational issues of data management and privacy to advanced technologies such as artificial intelligence, blockchain, and big data analytics, which promise to enhance data sharing and security.

However, this trajectory also reveals persistent challenges, including the need for improved data governance, ethical considerations surrounding patient privacy, and the importance of interdisciplinary collaboration. The research landscape remains fragmented, with a diverse array of publications and varying contributions from different countries, indicating both the global nature of the field and the necessity for cohesive efforts to address emerging issues. As healthcare continues to evolve in an increasingly digital world, ongoing innovation, collaboration, and a critical examination of the ethical implications of technology will be essential to ensure that advancements in medical data management serve the best interests of patients and healthcare providers alike.

Author Contributions: Conceptualization, I.V.N.; methodology, I.V.N.; validation, Ș.S.B. and O.A.; formal analysis, I.V.N.; investigation: I.V.N.; resources, Ș.S.B. and O.A.; data curation, A.-R.A.; writing—original draft preparation, I.V.N. and A.-R.A.; writing—review and editing, A.-R.A.; visualization, A.-R.A.; supervision, D.G.M.; project administration: O.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data obtained from the Scopus database, using the provided search methodology.

Acknowledgments: Publication of this paper was supported by the University of Medicine and Pharmacy Carol Davila, through the institutional program Publish not Perish.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fernández-Alemán, J.L.; Señor, I.C.; Lozoya, P.Á.O.; Toval, A. Security and privacy in electronic health records: A systematic literature review. *J. Biomed. Inform.* **2013**, *46*, 541–562. [CrossRef]
2. Arellano, A.M.; Dai, W.; Wang, S.; Jiang, X.; Ohno-Machado, L. Privacy Policy and Technology in Biomedical Data Science. *Annu. Rev. Biomed. Data Sci.* **2018**, *1*, 115–129. [CrossRef] [PubMed]
3. Aldossary, A.; Algirim, T.; Almubarak, I.; Almuhish, K. Cyber Security in Data Breaches. *J. Cyber Secur. Risk Audit.* **2024**, *2024*, 14–22. [CrossRef]
4. Mousa, R.S.; Shehab, R. Applying risk analysis for determining threats and countermeasures in workstation domain. *J. Cyber Secur. Risk Audit.* **2025**, *2025*, 12–21. [CrossRef]
5. Shin, H.; Ryu, K.; Kim, J.-Y.; Lee, S. Application of privacy protection technology to healthcare big data. *Digit. Health* **2024**, *10*, 20552076241282242. [CrossRef]
6. Ambinder, E.P. A History of the Shift Toward Full Computerization of Medicine. *J. Oncol. Pract.* **2005**, *1*, 54–56. [CrossRef] [PubMed]
7. Chico, V. The impact of the General Data Protection Regulation on health research. *Br. Med. Bull.* **2018**, *128*, 109–118. [CrossRef] [PubMed]
8. Idoko, B.; Alakwe, J.A.; Ugwu, O.J.; Idoko, J.E.; Idoko, F.O.; Ayoola, V.B.; Ejembi, E.V.; Adeyinka, T. Enhancing healthcare data privacy and security: A comparative study of regulations and best practices in the US and Nigeria. *Magna Sci. Adv. Res. Rev.* **2024**, *11*, 151–167. [CrossRef]
9. Ibrahim, A.; Aakarsh, R.; Rozenblit, J. A Framework for Secure Data Management for Medical Devices. In Proceedings of the Spring Simulation Conference (SpringSim 2020), Fairfax, VA, USA, 18–21 May 2020; Society for Modeling and Simulation International (SCS): San Diego, CA, USA, 2020.
10. Arbabi, M.S.; Lal, C.; Veeraragavan, N.R.; Marijan, D.; Nygård, J.F.; Vitenberg, R. A Survey on Blockchain for Healthcare: Challenges, Benefits, and Future Directions. *IEEE Commun. Surv. Amp Tutor.* **2023**, *25*, 386–424. [CrossRef]
11. Shaikh, Z.A.; Memon, A.A.; Shaikh, A.M.; Soomro, S.; Sayed, M. BLOCKCHAIN IN HEALTHCARE: UNLOCKING THE POTENTIAL OF BLOCKCHAIN FOR SECURE AND EFFICIENT APPLICATIONS FOR MEDICAL DATA MANAGEMENT- A PRESENTATION OF BASIC CONCEPTS. *Liaquat Med. Res. J.* **2023**, *5*, 81–86. [CrossRef]
12. Osamor, V.C.; Edosomwan, I.B.; Damilola, O.O. Application of Blockchain Technology for Data Privacy and Secured Sharing in Electronic Medical Records: A Systematic Literature Review. In Proceedings of the 2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG), Omu-Aran, Nigeria, 2–4 April 2024; pp. 1–12.
13. Mishra, V.; Gupta, K.; Saxena, D.; Singh, A.K. A Global Medical Data Security and Privacy Preserving Standards Identification Framework for Electronic Healthcare Consumers. *IEEE Trans. Consum. Electron.* **2024**, *70*, 4379–4387. [CrossRef]
14. Lea, N.C.; Nicholls, J.; Dobbs, C.; Sethi, N.; Cunningham, J.; Ainsworth, J.; Heaven, M.; Peacock, T.; Peacock, A.; Jones, K.; et al. Data Safe Havens and Trust: Toward a Common Understanding of Trusted Research Platforms for Governing Secure and Ethical Health Research. *JMIR Med. Inform.* **2016**, *4*, e22. [CrossRef]
15. Khatiwada, P.; Yang, B.; Lin, J.-C.; Blobel, B. Patient-Generated Health Data (PGHD): Understanding, Requirements, Challenges, and Existing Techniques for Data Security and Privacy. *J. Pers. Med.* **2024**, *14*, 282. [CrossRef] [PubMed]
16. Maher, M.; Khan, I. From Sharing to Selling. *Blockchain Healthc. Today* **2022**, *5*, 184. [CrossRef]
17. Yun, H. China's Data Sovereignty and Security: Implications for Global Digital Borders and Governance. *Chin. Polit. Sci. Rev.* **2025**, *10*, 178–203. [CrossRef]
18. Jain, D. Regulation of Digital Healthcare in India: Ethical and Legal Challenges. *Healthcare* **2023**, *11*, 911. [CrossRef] [PubMed]
19. Yekaterina, K. Challenges and Opportunities for AI in Healthcare. *Int. J. Law Policy* **2024**, *2*, 11–15. [CrossRef]
20. Thapa, C.; Camtepe, S. Precision health data: Requirements, challenges and existing techniques for data security and privacy. *Comput. Biol. Med.* **2021**, *129*, 104130. [CrossRef]
21. Elluri, L.; Piplai, A.; Kotal, A.; Joshi, A.; Joshi, K.P. A Policy-Driven Approach to Secure Extraction of COVID-19 Data from Research Papers. *Front. Big Data* **2021**, *4*, 701966. [CrossRef]
22. Rajab, R.M.; Abuhmida, M.; Wilson, I.; Ward, R.P. A Review of IoMT Security and Privacy related Frameworks. *Eur. Conf. Cyber Warf. Secur.* **2024**, *23*, 733–743. [CrossRef]
23. Davarasan, A.; Samual, J.; Palansundram, K.; Ali, A. A Comprehensive Review of Machine Learning Approaches for Android Malware Detection. *J. Cyber Secur. Risk Audit.* **2024**, *2024*, 38–60. [CrossRef]
24. Alshuaibi, A.; Almaayah, M.; Ali, A. Machine Learning for Cybersecurity Issues: A systematic Review. *J. Cyber Secur. Risk Audit.* **2025**, *2025*, 36–46. [CrossRef]
25. Malin, B.; Goodman, K. Between Access and Privacy: Challenges in Sharing Health Data. *Yearb. Med. Inform.* **2018**, *27*, 055–059. [CrossRef]

26. Madanian, S.; Nakarada-Kordic, I.; Reay, S.; Chetty, T. Patients' perspectives on digital health tools. *PEC Innov.* **2023**, *2*, 100171. [CrossRef] [PubMed]
27. Naili, Y.T.; Afrilies, M.H.; Garunja, E.; Purwono, P. Protection of patient data privacy on IoT devices for healthcare in the era of smart cities: A health law perspective. *J. Huk. Nov.* **2024**, *15*, 87. [CrossRef]
28. Otoom, S. Risk auditing for Digital Twins in cyber physical systems: A systematic review. *J. Cyber Secur. Risk Audit.* **2025**, *2025*, 22–35. [CrossRef]
29. Haarbrandt, B.; Schreiweis, B.; Rey, S.; Sax, U.; Scheithauer, S.; Rienhoff, O.; Knaup-Gregori, P.; Bavendiek, U.; Dieterich, C.; Brors, B.; et al. HiGHmed–An Open Platform Approach to Enhance Care and Research across Institutional Boundaries. *Methods Inf. Med.* **2018**, *57*, e66–e81. [CrossRef] [PubMed]
30. Mirchev, M.; Mircheva, I. Digital health and some thematic shifts in bioethics in academic publications after the pandemic. *Eur. J. Public Health* **2024**, *34*, ckae144.1186. [CrossRef]
31. Qiu, Y.; Hu, Z. Data governance and open sharing in the fields of life sciences and medicine: A bibliometric analysis. *Digit. Health* **2025**, *11*, 20552076251320302. [CrossRef]
32. Abeykoon, B.B.D.S.; Sirisena, A.B. A Bibliometric Analysis of Data Governance Research: Trends, Collaborations, and Future Directions. *South Asian J. Bus. Insights* **2023**, *3*, 70–92. [CrossRef]
33. Costa, T.; Borges-Tiago, T.; Martins, F.; Tiago, F. System interoperability and data linkage in the era of health information management: A bibliometric analysis. *Health Inf. Manag. J. Health Inf. Manag. Assoc. Aust.* **2024**, 18333583241277952. [CrossRef] [PubMed]
34. Manoj Kumar, L.; George, R.J.; Anisha, P.S. Bibliometric Analysis for Medical Research. *Indian J. Psychol. Med.* **2023**, *45*, 277–282. [CrossRef] [PubMed]
35. Freed, R. Legal Structure for National Medical Data Center. In Proceedings of the International Workshop on Managing Requirements Knowledge, San Fransisco, CA, USA, 9–11 December 1968; IEEE Computer Society: San Fransisco, CA, USA, 1968; Volume 33, pp. 387–394.
36. Karlsson, Y. Advanced system of secrecy for a data bank in the medical services. *Nord. Med.* **1973**, *88*, 41–42. [PubMed]
37. Goldberg, M.; Doyon, B. SARI: A user oriented data bank system for medical applications. *Methods Inf. Med.* **1976**, *15*, 69–74. [CrossRef]
38. Evans, R.S. Electronic Health Records: Then, Now, and in the Future. *Yearb. Med. Inform.* **2016**, *25*, S48–S61. [CrossRef]
39. Balazka, D.; Rodighiero, D. Big Data and the Little Big Bang: An Epistemological (R)evolution. *Front. Big Data* **2020**, *3*, 31. [CrossRef]
40. Andriole, K.P. Picture archiving and communication systems: Past, present, and future. *J. Med. Imaging* **2023**, *10*, 061405. [CrossRef]
41. Cole, L.J.; Fleisher, L.D. Update on HIPAA privacy: Are you ready? *Genet. Med. Off. J. Am. Coll. Med. Genet.* **2003**, *5*, 183–186. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.