



Bibliometric Characterization of Electronic Health Records in Privacy and Security


Chaimae Moumouh

(SIGL Laboratory, ENSATE, Abdelmalek Essaadi University, Tetouan, Morocco
 <https://orcid.org/0000-0001-9999-847X>, chaimae.moumouh@etu.uae.ac.ma)


José A. García-Berná

(Department of Computer Science and Systems, University of Murcia, Spain
 <https://orcid.org/0000-0002-9526-8565>, josealberto.garcia1@um.es)


Begoña Moros

(Department of Computer Science and Systems, University of Murcia, Spain
 <https://orcid.org/0000-0002-3092-7654>, bmoros@um.es)


Juan M. Carrillo de Gea

(Department of Computer Science and Systems, University of Murcia, Spain
 <https://orcid.org/0000-0002-3320-622X>, jmcdg1@um.es)

Mohamed Yassin Chkouri

(SIGL Laboratory, ENSATE, Abdelmalek Essaadi University, Tetouan, Morocco
 <https://orcid.org/0000-0001-7256-329X>, mychkouri@uae.ac.ma)

José L. Fernández-Alemán

(Department of Computer Science and Systems, University of Murcia, Spain
 <https://orcid.org/0000-0002-0176-450X>, aleman@um.es)

Abstract: The impact of technology on improving health and well-being of individuals is remarkable. EHealth boosts the transition from paper-based health records to Electronic Health Records (EHRs). The use of EHRs can lead to improve quality of care, costs and time. In eHealth systems the health data is stored in digital form, and can be exchanged or accessed securely by authorised users. It is worth noting that medical data is considered very confidential information. However, the privacy and security of medical data remains a critical issue. Any leak or breach in security can lead to serious privacy damages for patients. Despite the safeguards, training courses and the consciousness on keeping data safe, the human error continues to be a problem. The main purpose of this paper is to present a bibliometric overview on the academic research related to privacy and security in EHRs. For this purpose, the papers of this study were searched in the Scopus. A period of 24 years was considered for selecting the papers. The information gathered in the database identified a total of 3,077 publications. Some key findings revealed that in the year 2015 the highest number of publications was produced. The Harvard Medical School was the most prolific institution with 2.44% papers from the total number of publications. A total of 97.21% of the documents were written in English. Finally, the results provided in this manuscript allowed us to make a picture on the current relevance in academic literature on privacy and security in EHRs.

Keywords: Electronic Health Records; Privacy; Security; eHealth, bibliometric analysis
Categories: H.2; H.3

1 Introduction

In recent years, the adoption of eHealth technologies has grown significantly, with Electronic Health Records (EHRs) becoming a central component in the digital transformation of healthcare. EHRs store health data in a structured digital format, enabling secure information exchange among patients and healthcare providers [Shahnaz et al., 2019]. These systems have been promoted across both public and private sectors to improve care quality and operational efficiency [Plantier et al., 2017]. Closely related are Electronic Medical Records (EMRs), which focus on clinical data within specific institutions [Heart et al., 2017], and Personal Health Records (PHRs), which are managed directly by patients and allow controlled access to medical information [Fernández-Alemán et al., 2013].

The integration of these systems supports enhanced data sharing, real-time monitoring, and personalized care. Notably, mobile personal health records (mPHRs) have gained prominence by enabling users to collect and visualize health data via smartphones [Dohan et al., 2014], with applications such as pregnancy tracking [Alqahtani et al., 2021]. However, the increasing connectivity and mobility of health information systems raise serious concerns about privacy and security. Ensuring confidentiality, integrity, and availability of medical data is essential in preventing unauthorized access, data manipulation, or service disruption [Ma et al., 2013].

Despite the use of encryption and administrative safeguards, vulnerabilities persist due to weak physical security, insufficient training, and user negligence [McLeod and Dolezel, 2018], [Chernyshev et al., 2019]. Studies have shown that a significant proportion of healthcare professionals lack awareness of proper data handling protocols, further exacerbating the risk of breaches [Flaumenhaft and Ben-Assuli, 2018], [Fernández-Alemán et al., 2015a]. Regulatory frameworks such as the General Data Protection Regulation (GDPR) aim to address these risks, yet the complexity of eHealth environments continues to pose challenges [European Union, 2016].

While the technical and regulatory aspects of EHR privacy and security are well explored, there is a lack of comprehensive bibliometric studies that systematically map the scientific discourse in this domain. A bibliometric approach enables the identification of research trends, influential authors and institutions, collaborative patterns, and emerging topics—offering valuable insights for academics, practitioners, and policymakers.

Objective and Scope

This study aims to analyze academic publications related to privacy and security in Electronic Health Records through bibliometric and network analysis techniques. Using data extracted from the Scopus database, we examine key indicators such as publication growth, geographic distribution, author collaboration, citation dynamics, and keyword co-occurrence.

Research Questions

The following research questions guide our analysis:

- RQ1: What is the growth of literature on privacy and security in EHRs?
- RQ2: What are the hot topics in EHRs' privacy and security papers?
- RQ3: What is the collaborative profile of authors in this field of research?
- RQ4: What are the most relevant research networks among the authors?

The remainder of the paper is organized as follows. Section 2 presents the search process and tools used to analyze the data collected from the Scopus database. This section also describes how the search string was validated and how the experiment was designed for replication. Section 3 introduces a range of bibliometric indicators along with brief descriptions. A more in-depth analysis of the results is presented in Section 4, where the key findings are discussed in relation to the existing literature. Finally, Section 5 outlines the study's limitations and suggests directions for future work.

2 Related work

The integration of digital technologies into healthcare systems has led to the emergence of a variety of eHealth platforms, among which Electronic Health Records (EHRs) play a central role. EHRs enable structured and secure access to medical data, facilitating communication among healthcare professionals and enhancing the continuity of care [Shahnaz et al., 2019]. Their implementation has been actively promoted across both public and private sectors, reflecting their potential to improve clinical outcomes and reduce costs [Plantier et al., 2017].

Closely related technologies include Electronic Medical Records (EMRs)—typically used within single institutions—and Personal Health Records (PHRs), which allow patients to manage and share their own health information [Heart et al., 2017]. The increasing interconnection between these systems supports data sharing and patient empowerment, but also introduces new vulnerabilities. Recent developments such as mobile PHRs (mPHRs) have further expanded accessibility by enabling real-time health monitoring via smartphones, particularly in use cases like pregnancy tracking [‘Proceedings of the 39th Hawaii International Conference on System Sciences - 2006’, 2006], [Bachiri et al., 2018].

A common thread across these systems is the critical importance of information security. Three foundational principles—confidentiality, integrity, and availability—guide the protection of medical data [Ma et al., 2013]. Breaches in these areas can result in identity theft, medical fraud, or compromised patient safety. The risks are amplified in networked environments where data is transmitted to external devices and cloud platforms, often lacking robust endpoint security [Collins et al., 2011]. Studies report that even with encryption protocols in place, physical security and user behavior remain weak links in the security chain [Li, 2015], [McLeod and Dolezel, 2018]

Several surveys underscore the role of human error and limited staff awareness in privacy violations. For instance, a clinical setting study revealed that 31.7% of professionals were unaware of their organization's data deletion protocols, and over 50% inadvertently exposed patient data on screens visible to unauthorized viewers

[Flaumenhaft and Ben-Assuli, 2018], [Fernández-Alemán et al., 2015a]. These findings highlight the need for continuous training and institutional policies to address behavioral and procedural vulnerabilities.

Research has extensively addressed privacy and security challenges in Electronic Health Records (EHRs), adopting various methodological approaches. Fernández-Alemán et al. [Fernández-Alemán et al., 2013a] conducted a comprehensive systematic literature review highlighting common technical safeguards, such as encryption, access control, and authentication, but also noted a lack of standardization across implementations [Kruse et al., 2017], and McLeod and Dolezel [McLeod and Dolezel, 2018] reviewed modern cybersecurity threats in healthcare, identifying phishing, insider misuse, and ransomware as persistent vulnerabilities. From a technical standpoint, several studies have proposed frameworks for secure EHR sharing using cloud computing [Zhang and Liu, 2010] and blockchain technologies [Azaria et al., 2016], though these are often theoretical and lack real-world validation.

Behavioral and user-centered research has underscored the impact of privacy concerns on EHR adoption. Policy analyses by Ben-Assuli [Ben-Assuli, 2015] and Menachemi and Collum [Menachemi and Collum, 2011] further pointed out that legal frameworks such as HIPAA and GDPR play critical roles in shaping privacy practices, yet inconsistencies in enforcement and compliance remain problematic. Despite this body of work, most existing studies focus on isolated aspects—technical, behavioral, or regulatory—without offering integrated models that address EHR privacy and security holistically. This gap reinforces the need for comprehensive, evidence-based mappings of the field, such as the bibliometric analysis presented in this study.

Governments and regulatory bodies have responded with legal frameworks aimed at strengthening data governance. The General Data Protection Regulation (GDPR), enacted in the European Union in 2018, has introduced stringent requirements for handling personal data, including health information [European Union, 2016]. While it has improved transparency and accountability, practical implementation across healthcare systems remains uneven.

Although extensive research exists on the technical, legal, and organizational aspects of EHR privacy and security, no bibliometric studies have yet synthesized this knowledge to reveal the structure and evolution of research in this area. Bibliometric analyses offer a systematic way to uncover publication trends, citation dynamics, and collaboration patterns. Such an approach is essential for identifying knowledge gaps and guiding future investigations.

3 Methodology

The bibliometrics term is defined as “the application of mathematical and statistical methods to books and other communication medium” [Patra et al., 2006]. Bibliometric studies are used to analyse literature related to a certain topic. Moreover, they are used to examine the influence of a research field. The bibliometric methods evaluate the impact of researchers and papers, assess research performance of institutions on a particular knowledge area. In this paper, an evaluation of the data was carried out by means of statistical methods for classification, providing with a representative summary of the contributions [Broadus, 1987]. The Scopus database was employed in this study

to collect the data to be analysed. Scopus is one of the academic databases with the largest number of abstracts online [Burnham, 2006].

Although no restriction on years was imposed on the search in Scopus, a period of 24 years was considered in this bibliometric study, from 1996 to 2019. For clarifying purposes data were shown since 2009 in some indicators. The search string was built including the words that define the critical issues considered in this paper concerning EHRs, which are privacy and security. The terms were the following ones : (privacy OR security) AND "electronic health record". Moreover, they should appear in the abstract section, title or keywords of the manuscripts.

The search in the database was carried out on March 11th, 2020. The search results included a total of 3,077 publications that were exported to a CSV file and analysed in a bibliometric manner. Excel was used to manipulate the CSV file downloaded. The library Python Data Analysis (Pandas) was employed, allowing to work with the CSV file and manage the data to calculate the parameters presented in Section 3. Finally, the VOSviewer tool was utilized to show clustered data in the same section [Sweileh et al., 2017].

A sensibility analysis of the results allowed to study the accuracy of the collection of papers reported by Scopus concerning the field under study. Thus, a sample of publications was randomly selected from the whole list. Then, the suitability of each publication for the purpose of the study was analysed. The amount of documents randomly selected to perform the validation of the search string was obtained with Cochran's sample size formula (1), which is as follows [Cochran and William, 1977].

$$n = \frac{N \cdot Z^2 \cdot p \cdot (1 - p)}{(N - 1) \cdot e^2 + Z^2 \cdot p \cdot (1 - p)} \quad (1)$$

In this formula (1) n is the number of randomly selected papers. N is the total number of items found in Scopus. Z is the deviation from the mean value accepted for a confidence level ($Z=1.645$ implies that the level of confidence is of 90%, $Z=1.96$ means a level of confidence of 95%, and $Z =2.575$ signifies a level of 99%). The parameter e is the margin of error, and p is the ratio of items that are expected to be invalid for this study. By taking the values of $N=3,077$ and $Z=1.96$ (level of confidence 95%), $e=0.05$ and $p=0.08$ (a low value is expected from the randomly chosen sample), the number of papers to check was obtained, $n = 109.12$. Consequently, a subset of 110 random papers was picked to carry out the validation. The suitability analysis revealed that only 12 publications had nothing to do with privacy and security in EHR, which means a 10.9% of invalid results (8% expected).

Together with the validation of the set of articles, the verification of the search string was also tackled. To do that, the most frequent keywords in the papers obtained in Scopus were checked. The terms most frequently found in the publications were electronic health record/s, privacy, security, EHR, cloud computing, access control and confidentially. These concepts were narrowly related to the topic under study, and confirmed the appropriateness of the search string.

4 Results

In this bibliometric analysis, the results were organised into a set of sections. The sections are related to the analysis of parameters of the current situation of publication growth, the number of publications by authors and journals, collaboration between authors, the study of citations between authors and the evaluation of citations in journals [Andrés, 2009].

4.1 Descriptive analysis

The descriptive analysis presented in this section involves the discussion of different parameters classified into four categories: 1) temporal evolution, 2) institutions and countries, 3) language and 4) type of documents.

4.1.1. Temporal evolution

- RQ1: What is the growth of literature on privacy and security in EHRs?

In order to respond RQ1, the production of literature related to EHR privacy and security is showed graphically in Figure 1. The total number of papers published is depicted together with the percentage of publications per year with respect to the total of publications found. With this parameter, the upward trend in the number of publications can be observed, especially in the period between 2010 and 2015, in which there is a notable growth.

In this section, the trend regarding the numbers of publications along the years was studied. By observing the results, the number of publications increased from the year 2010 to the year 2015. Since 2010 a high amount of studies have been published regarding privacy and security in EHRs, which lead to a noticeable growth in the production of literature. The total number of publications from 2010 to 2015 was 1,495 (48.58%), being 2015 the most prolific year with 320 publications (21.40% of the total production over these five years). After a decline in 2016, the upward trend started from 2017 onwards. In 2019, when this study was conducted, the number of publications almost matches the number of 2015 with 311 publications (10.11%).

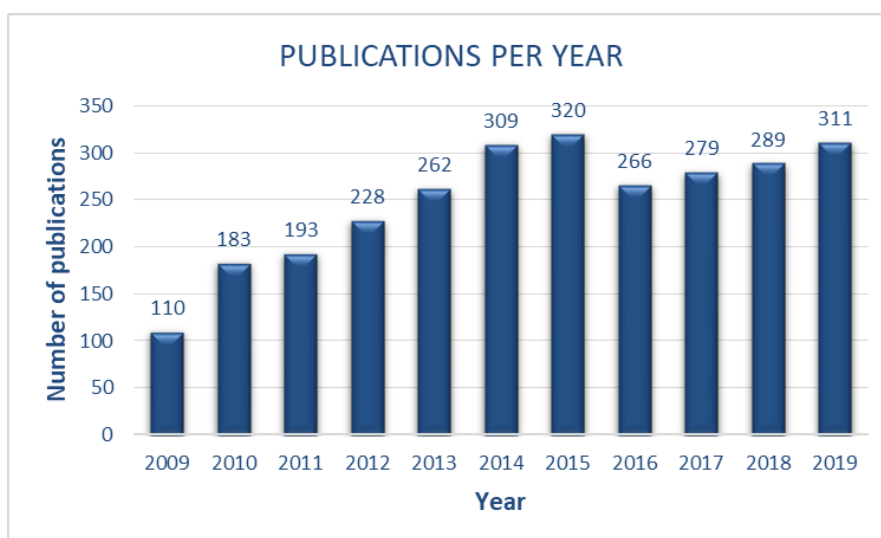


Figure 1: Growth of publications

4.1.2. Institutions and countries

Most prolific institutions

Table 1 presents the most productive institutions when it comes to literature related to privacy and security in EHRs. According to the data gathered from the search process, the most productive and prolific institution was the Harvard Medical School with 75 publications, which means the 2.44% of the total amount of publications in the aforementioned topic. This institution was followed by the University of California with 46 publications (1.49%). The Vanderbilt University, Brigham and Women's Hospital, University of Washington and Queensland University of Technology obtained similar values with a percentage around 1%. University of Toronto, University of Victoria, University College London and University of Ottawa completed the top ten institutions with a number of publications between 27 and 24, which was a percentage also close to 1%.

Pos.	Institution	Number of publications	%
1	Harvard Medical School	75	2.44
2	University of California	46	1.49
3	Vanderbilt University	44	1.43
4	Brigham and Women's Hospital	33	1.07
5	University of Washington	31	1.01
6	Queensland University of Technology	31	1.01
7	University of Toronto	27	0.88
8	University of Victoria	25	0.81

9	University College London	24	0.78
10	University of Ottawa	24	0.78

Table 1: Most prolific institutions

Geographical distribution of the publications

The data provided in Table 2 shows a collection of countries that achieved the highest number of publications related to security and safety in EHR. The countries were collected based on the country affiliations of the authors as indexed in the Scopus database. The percentages shown in the table come from dividing the number of publications from each country by the total number of publications appearing in Scopus (3,077 according to Figure 1). The country with the highest output was United States (US) with a total of 1,138 publications (36.98%). The next country in the list, Australia, was quite far from US, with a difference nothing less than 923 publications. The total number of publications from Australia was 215, which corresponds to a percentage of 6.99%. This country was followed by United Kingdom with 208 publications (6.76%). In the same range (between 100 and 200 publications) Canada (5.49%), Germany (5.07%), India (4.84%) and China (4.39%) were found. Finally, Italy, Taiwan, Austria, Spain and France contributed with a percentage around 2%. In Figure 2 the world map has been coloured to depict the current situation.

Pos.	Country	Number of publications	%
1	United States	1138	36.98
2	Australia	215	6.99
3	United Kingdom	208	6.76
4	Canada	169	5.49
5	Germany	156	5.07
6	India	149	4.84
7	China	135	4.39
8	Italy	66	2.14
9	Taiwan	62	2.01
10	Austria	61	1.98
11	Spain	59	1.92
12	France	57	1.85

Table 2: Geographical distribution of publications

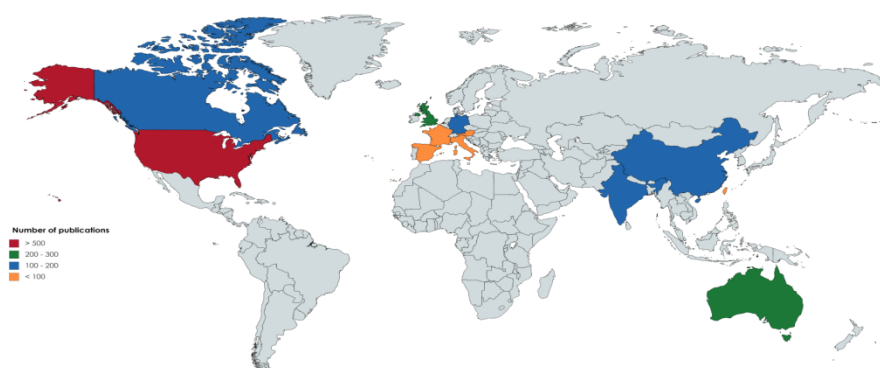


Figure 2: Most productive countries

4.1.3. Language used in publications

The list of ten languages most used for the writing of the publications is showed in Table 3. English was on the top of the list, with a total of 2,991 publications (97.21% of the total). The remaining 2.89% is distributed mainly among German (1.33%), French (0.52%) and Spanish (0.36%). This group of languages was followed by Chinese, Swedish, Portuguese, Japanese, Turkish and Italian with 6 publications or less. It is worth noting that the total amount of papers in Table 3 was greater than the total amount of publications retrieved from the database. This is because in some cases the abstract was written in two different languages.

Pos.	Country	Number of publications	%
1	English	2991	97.21
2	German	41	1.33
3	French	16	0.52
4	Spanish	11	0.36
5	Chinese	6	0.19
6	Swedish	5	0.16
7	Portuguese	3	0.10
8	Japanese	3	0.10
9	Turkish	2	0.06
10	Italian	2	0.06

Table 3: Distribution as regards language used

4.1.4. Type of document

In Table 4 the distribution of publications is presented according to the means used. The list revealed that more than a half of the publications regarding privacy and security in EHRs can be found in the form of Articles. The total number of articles were 1,660,

which means 53.95% over the total amount of publications in the field. This kind of publication was followed by Conference Papers (25.90%), Reviews (7.44%) and Book Chapter (3.57%). The rest of the publications means were less significant since the number of publications found were less than 100, reaching the number of 7 in the form of Books (0.23%).

Pos.	Forms of publication	Publications	%
1	Article	1660	53.95
2	Conference Paper	797	25.90
3	Review	229	7.44
4	Book Chapter	110	3.57
5	Note	89	2.89
6	Editorial	54	1.75
7	Conference Review	51	1.66
8	Short Survey	37	1.20
9	Letter	36	1.17
10	Article in Press	7	0.23
11	Book	7	0.23

Table 4: Forms of publications

- RQ2: What are the hot topics in EHRs’ privacy and security papers?

The most commonly terms used by the authors in publications related to privacy and security in EHRs were used to answer the research question. The keywords that would allow the Internet users to find these publications can be seen in the data collected, analysed and presented in the next subsection "Common keywords in the publications".

4.1.5. Common keywords in the publications

Table 5 exposes the most common keywords that have been identified in the papers. These terms are Electronic health record(s) and its abbreviation EHR, privacy, security, cloud computing, access control, confidentiality, healthcare and eHealth. All these keywords correspond to concepts closely related to the field analyzed in the present manuscript. Among them, the most frequent keywords are “**Electronic Health Records / EHR,**” “**Privacy,**” and “**Security**”, which clearly reflect the central research focus on protecting patient information and ensuring data confidentiality in healthcare systems.

Pos.	Keywords	Number of publications
1	Electronic health records	376
2	Privacy	341
3	Electronic health record	278
4	Security	243
5	EHR	119

$$PPA = \frac{\text{ProductivityPerAuthors}}{\text{AuthorsPerPaper}} = \frac{1}{\text{AuthorsPerPaper}} \quad (10)$$

Year	Papers	Authors	AAPP	PPA
2019	311	1235	3.97	0.25
2018	289	1257	4.35	0.23
2017	279	1133	4.06	0.25
2016	266	1010	3.80	0.26
2015	320	1116	3.49	0.29
2014	309	1166	3.77	0.27
2013	262	867	3.31	0.30
2012	228	806	3.54	0.28
2011	193	565	2.93	0.34
2010	183	518	2.83	0.35
<2009	437	1222	2.80	0.36

Table 6: Author productivity

The results in 6 revealed that was common to find publications written by 2 or 3 authors (AAPP) from the first years to 2011. From 2012 to 2016 the value of AAPP is near 4. However, it is not until 2017 when this value is reached. Thus, a more collaborative trend starts from 2017. On the other hand, the mean value of productivity per author (PPA) has varied from 0.36 to 0.25 over the period under study.

4.2.2. Authorship trend analysis

Table 7 depicts the total amount of publications per year broken down by the number of authors collaborating in them. Anonymous publications were not taken into account in this classification. The total amount of publications considered for the calculation of the percentages have been obtained by including the papers from 2010 to 2019 (2,640 in Table 7 instead of the total number of papers). The figures in Table 7 revealed that publications written by multiple authors contributed more to literature production throughout the decade under study. Multiple authors' publications have trended upward from 2010 to 2015, and were particularly interrupted in 2016.

Year	Single author		Multiple authors		0 to N authors
	Total output	%	Total output	%	Total output
2019	311	1235	3.97	0.25	311
2018	289	1257	4.35	0.23	289
2017	279	1133	4.06	0.25	279
2016	266	1010	3.80	0.26	266
2015	320	1116	3.49	0.29	320
2014	309	1166	3.77	0.27	309
2013	262	867	3.31	0.30	262
2012	228	806	3.54	0.28	228
2011	193	565	2.93	0.34	193

2010	183	518	2.83	0.35	183
<2009	437	1222	2.80	0.36	2640

Table 7: Authorship trend analysis

4.2.3. Most prolific authors

The most prolific authors concerning privacy and security in EHRs had been collected, and the results are showed below. Table 8 reveals that Bradley Malin was the most prolific author with 22 publications; followed by Tony Sahama with 21 publications and Bernd Blobel with 19 publications. The percentages were calculated taking into account the total number of publications obtained from the search in Scopus (see in Table 8).

Pos.	Author	Number of publications	%
1	Malin B.	22	0.71
2	Sahama T.	21	0.68
3	Blobel B.	19	0.62
4	Ohno-MachadoL	16	0.52
5	Jiang X.	13	0.42
6	Kalra D.	13	0.42
7	Wang X.	13	0.42
8	Bates D.W.	12	0.39
9	Wang H.	12	0.39
10	Chen Y.	11	0.36
11	Pharow P.	11	0.36
12	Susilo W.	11	0.36
13	Wang S.	11	0.36
14	Win K.T.	11	0.36
16	Zhang X.	11	0.36

Table 8: Most prolific authors

4.2.4. Lotka's law

According to the literature, the Lotka's law [Bailón-Moreno et al., 2005] is commonly used for bibliometric analysis. In this law the number of authors is related to their productivity. Lotka's law focuses on the premise that most authors publish a single manuscript, while a small number of authors generate a large number of articles [Andrés, 2009].

As is shown in Table 9, 8,257 authors ($\sum yx$) contributed to the publications in the field of study. Most of them published just one publication (6,883) as expected according to the premise of the law. Column 7 ($yx/(\sum yx)$) in Table 9 represents the observed frequency. That is to say, the frequency of authors with x number of publications decreases as the number of publications increases. Once the coefficient n in Lotka's equation (11) is known, the constant c can also be calculated as stated in the equation (13). With the collected data the result of the value was $c = 0.834$. So, the

expected frequency of authors publishing x number of publications (fe in Table 9) was calculated by applying Lotka's equation (11).

$$c = \frac{1}{\sum 1/x^n} \quad (13)$$

Finally, a Kolmogorov-Smirnov goodness-of-fit test was done to study if the collected data fit the expected distributions according to the Lotka's law. To do so, the null hypothesis (i.e.: the data follow the Lotka's law) is rejected if the maximum difference (D in table 9) between the observed cumulative distribution (column 8 in Table 9) and the expected cumulative distribution (column 10 in Table 9) is greater than the critical value obtained in the equation (14). The coefficient 1.63 in the equation (14) corresponds to a level of significance of 1%. Since the critical value was 0.017, greater than the maximum value of D, which was 0.003, the null hypothesis could not be rejected. Consequently, the author productivity in the literature related to privacy and security in EHRs do not fits Lotka's law.

x	yx	X=log x	Y=log yx	X ²	XY	yx/(Σyx)	Σ(yx/(Σ yx))	fe=c x-n	Σfe	D
1	6883	0.000	3.838	0.000	0.000	0.834	0.834	0.835	0.835	0.001
2	864	0.301	2.937	0.091	0.884	0.105	0.939	0.104	0.939	0.000
3	238	0.477	2.377	0.228	1.134	0.029	0.967	0.031	0.969	0.002
4	98	0.602	1.991	0.362	1.199	0.012	0.979	0.013	0.982	0.003
5	71	0.699	1.851	0.489	1.294	0.009	0.988	0.007	0.989	0.001
6	44	0.778	1.643	0.606	1.279	0.005	0.993	0.004	0.993	-0.001
7	17	0.845	1.230	0.714	1.040	0.002	0.995	0.002	0.995	0.000
8	16	0.903	1.204	0.816	1.087	0.002	0.997	0.002	0.997	-0.001
9	5	0.954	0.699	0.911	0.667	0.001	0.998	0.001	0.998	0.000
10	6	1.000	0.778	1.000	0.778	0.001	0.999	0.001	0.999	0.000
11	6	1.041	0.778	1.084	0.810	0.001	0.999	0.001	0.999	0.000
12	2	1.079	0.301	1.165	0.325	0.000	1.000	0.000	1.000	0.000
13	3	1.114	0.477	1.241	0.531	0.000	1.000	0.000	1.000	0.000
16	1	1.204	0.000	1.450	0.000	0.000	1.000	0.000	1.000	0.000
19	1	1.279	0.000	1.635	0.000	0.000	1.000	0.000	1.000	0.000
21	1	1.322	0.000	1.748	0.000	0.000	1.000	0.000	1.000	0.000
22	1	1.342	0.000	1.802	0.000	0.000	1.000	0.000	1.000	0.000
Σ	8257	14.94 2	20.105	15.341	11.029					

Table 9: Data needed for Lotka's law calculation

$$cv = \frac{1.63}{\left(\sum y + \left(\sum y/10\right)^{1/2}\right)^{1/2}} \quad (14)$$

Figure 4 shows a high coefficient of determination (R2 = 0.9844), which determines a high linearity between the number of authors and papers published.

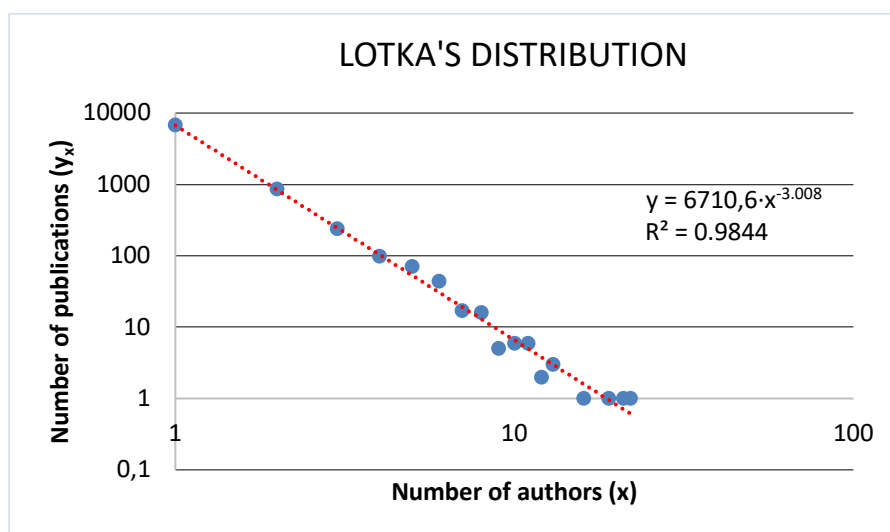


Figure 4: Relationship between the number of authors and publications

4.3. Journal and conference ranks

Both journals and conferences have been considered as sources for publication in the area of privacy and security in EHRs. A total of 3,077 publications spread out between 1,205 different sources were retrieved from Scopus. As Table 10 shows, the most preferred source was the book series Studies in Health Technology and Informatics where 224 articles from the total have been published (7.28%). This book series was followed by the Journal of Medical Systems (3.15%) and the International Journal of Medical Informatics (2.83%). Next, the Journal of the American Medical Informatics Association and the Journal of Biomedical Informatics, both with 67 publications, together with the Lecture Notes in Computer Science including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics and Journal of Medical Internet Research, all of them with a percentage of publications around 2%. The last 3 sources of publications in the top ten ranking in the subject area, with a percentage of publications close to 1%, were Health Management Technology, Journal of the American Medical Association and BMC Medical Informatics and Decision Making.

Pos.	Sources	Number of publications	%
1	Studies in Health Technology and Informatics	224	7.28
2	Journal of Medical Systems	97	3.15
3	International Journal of Medical Informatics	87	2.83
4	Journal of the American Medical Informatics Association	67	2.18

5	Journal of Biomedical Informatics	67	2.18
6	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	58	1.88
7	Journal of Medical Internet Research	52	1.69
8	Health Management Technology	26	0.84
9	JAMA - Journal of the American Medical Association	24	0.78
10	BMC Medical Informatics and Decision Making	24	0.78

Table 10: Most preferred sources

4.4. Scientific collaboration

Following the recommendations given in the literature [Andrés, 2009], the collaboration between authors and institution has been analysed in this subsection. Some of the indicators studied were the co-authorship index (CAI), collaboration index (CI) and degree of collaboration. Finally, a research network based on co-authorships was depicted.

4.4.1. Collaboration index

Around the 42% of the publications were carried out by 1 or 2 authors as depicted in Table 11. Each group represented almost a 20% of the total amount of publications separately. These figures were almost similar to the ones concerning 3 and 4 authors, which reached the 17% and 13% respectively.

Number of authors	Number of publications	%
Anonymous	83	2.70
1	624	20.28
2	649	21.09
3	542	17.61
4	414	13.45
5	271	8.81
6	164	5.33
7	113	3.67
8	68	2.21
9	45	1.46
10	26	0.84
>10	78	2.53

Table 11: Authorship pattern of publications

Equation (17) and data from Table 11 were used for the degree of collaboration calculation. This parameter revealed the cooperation trend among the authors. By discarding the publications for which no author was reported, the value of the index

was equal to 0.7915. The result revealed that the authors tend to cooperate in the field of privacy and security in EHRs. A value near to 1 showed a general willingness to cooperate on the part of the authors.

$$C = \frac{\text{NumberOfPublications}_{\text{multipleAuthors}}}{\text{NumberOfPublications}_{\text{multipleAuthors}} + \text{NumberOfPublications}_{\text{singleAuthor}}} \quad (17)$$

The number of authors involved in writing the publications was then studied to find the trend over time. The collaboration index (CI) represented this information and was calculated as shown in equation (18). Table 12 contains the resulting values.

$$CI = \frac{\text{NumberOfSignatoriesInMultiauthoredPublications}}{\text{NumberOfMultiauthoredPublications}} \quad (18)$$

The value of CI remained around 4 in the years of the data studied and reached the value of 5 in 2018. Consequently, a greater collaboration was produced in that year. This pattern was also revealed from the growth in both total authors in multi-authored publications and multi-authored publications parameters.

- *RQ3: What is the collaborative profile of authors in this field of research?*

The number of authors in each publication and the country of the authors were collected to address RQ3. An analysis was done in order to detect if they belong to the same institutions or countries. The results were presented in the subsection “National and international collaborations”.

Year	Multi-authored publications	Total signatories in multi-authored publications	CI
2019	277	1235	4.46
2018	243	1257	5.17
2017	235	1133	4.82
2016	223	1010	4.53
2015	247	1116	4.52
2014	244	1166	4.78
2013	189	867	4.59
2012	170	806	4.74
2011	130	565	4.35
2010	122	518	4.25
2009	77	345	4.48
2008	59	251	4.25

Table 12: Collaboration index

4.4.2. National and international collaborations

The number of papers that were published by multiple authors from different institutions or countries allowed to obtain a collaboration profile. The types of collaborations were classified into (1) international, when authors have different

nationalities between them; (2) national, when authors came from the same country but they belonged to different institutions; and (3) no collaboration, in all other cases. Figure depicts these collaborations. National collaborations were preferred rather than international ones.

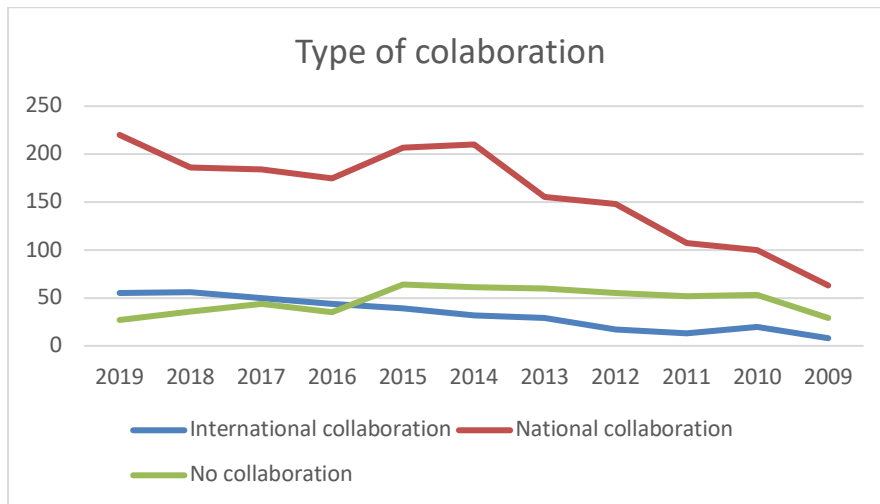


Figure 5: Collaboration pattern in international and national publications

4.4.3. Co-authorship index

The co-authorship index (CAI) allowed to show that the number of publications in a country was in the average of a pattern of co-authorship [Andrés, 2009]. This index was calculated as stated below.

$$CAI = ((N_{ca}/N_{ct}) / (N_{ta}/N_{tt})) \times 100 \quad (19)$$

In equation (19), N_{ca} represented the papers in the c th country co-authored by a authors, N_{ct} had the value of the papers in the c th country, N_{ta} was the papers in the total number of countries co-authored by a authors and N_{tt} valued the total amount of papers in all the countries.

Figure shows the CAI values for co-authored papers (from two to eight authors) in the 10 countries with more publications). The highest CAI value corresponded to Germany for eight authored papers, followed by Italy for the same numbers of authors. The next country was China for six authored papers and Taiwan for four authored papers. CAI index took the value of 0 for eight co-authored papers in Austria and India.

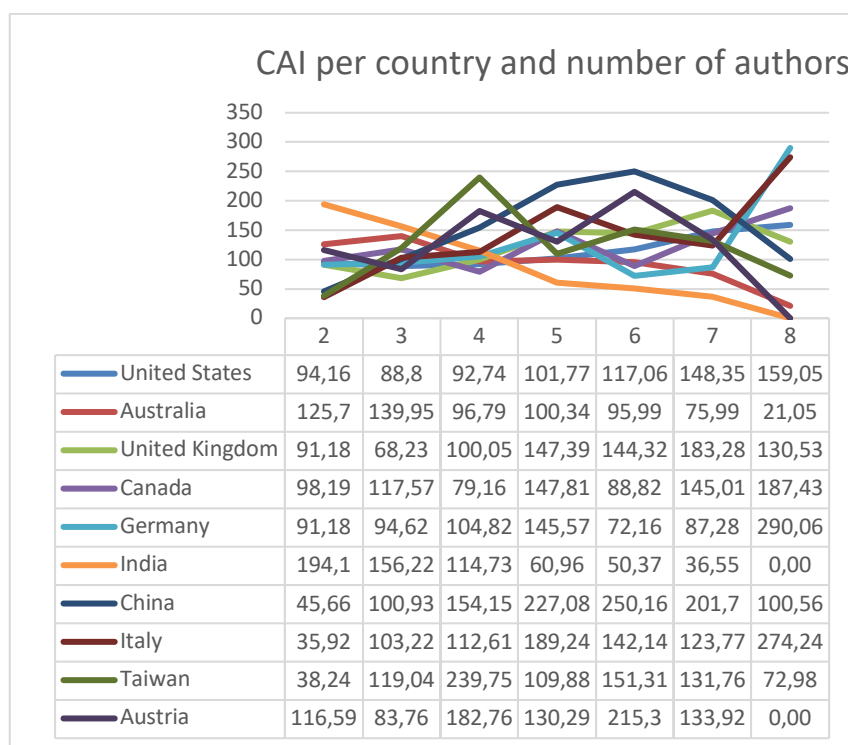


Figure 6: The co-authorship index plot

- RQ4: What are the most relevant research networks among the authors?

In the multi-authored publications, data on the most relevant groups of authors along with their names and relationships between them were gathered. Figure presents the results aiming at answering the fourth research question.

4.4.4. Research networks

The collaboration patterns among authors belonging to the same or different institution or country gave raise to research networks. Network assessment allowed to identify the number of members in the network, the intensity of the relationships and the most relevant members. By focusing on authors, the requirement for belonging to the network was to have published 6 papers in co-authorship. Figure 7 depicts the research network for the authors regarding privacy and security in EHRs obtained after conducting a clustering technique [Sweileh et al., 2017]. The diagram shows the main authors in a subject area. Moreover, the intensity of the relationships between them depends on the numbers of articles s in co-authorship.

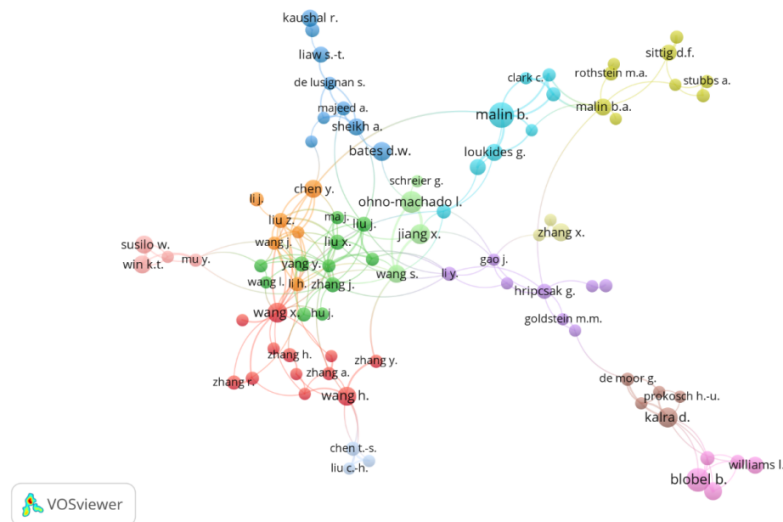


Figure 7: The most relevant authors separated into each cluster depending on co-authorship

4.5. Author citation analysis

Together with the analysis of research networks, citation analysis was a way of identifying relationships between authors or sources such as journals. In this section, the authors together with their most frequently cited publications were depicted. Moreover, a co-citation analysis was carried out to identify intellectually related authors.

4.5.1. Most cited publications

Table 13 shows the manuscripts ordered by number of citations in Scopus. In this study the most cited papers were “Mining electronic health records: Towards better research applications and clinical care”, with 634 citations. The second most cited paper was “Adoption of electronic health records in the presence of privacy concerns: The elaboration likelihood model and individual persuasion”, with 486 citations.

Pos.	Title	Authors	Source	Citations
1	Mining electronic health records: Towards better research applications and clinical care	Jensen P.B., Jensen L.J., Brunak S.	Nature Reviews Genetics	634
2	Adoption of electronic health records in the presence of privacy concerns: The elaboration likelihood	Angst C.M., Agarwal R.	MIS Quarterly: Management Information Systems	486

	modeland individual persuasion			
3	The eMERGE Network: A consortium of biorepositories linked to electronic medical records data for conducting genomic studies	McCarty C.A., Chisholm R.L., Chute C.G., Kullo I.J., Jarvik G.P. et alt.	BMC Medical Genomics	365
4	Big data in health care: Using analytics to identify and manage high-risk and high-cost patients	Bates D.W., Saria S., Ohno-Machado L., Shah A., Escobar G.	Health Affairs	353
5	Healthcare Data Gateways: Found Healthcare Intelligence on Blockchain with Novel Privacy Risk Control	Yue X., Wang H., Jin D., Li M., Jiang W.	Journal of Medical Systems	255
6	Security and privacy in electronic health records: A systematic literature review	Fernández-Alemán J.L., Señor I.C., Lozoya P.T.O., Toval A.	Journal of Biomedical Informatics	253
7	Social, ethical and legal barriers to E-health	Anderson J.G.	International Journal of Medical Informatics	240
8	Benefits and drawbacks of electronic health record systems	Menachemi N., Collum T.H.	Risk Management and Healthcare Policy	232
9	Policy: Achieving a nationwide learning health system	Friedman C.P., Wong A.K., Blumenthal D.	Science Translational Medicine	220
10	Opportunities and obstacles for deep learning in biology and medicine	Ching T., Himmelstein D.S., Beaulieu-Jones B.K., Kalinin A.A., Do B.T. et alt.	Journal of the Royal Society Interface	213

Table 13: Most cited publications

4.5.2. Co-citation analysis

Author co-citation appears when two authors are cited in the same publication. Moreover, an academic relationship between them might be assumed. Table 14 shows the list of the top ten most cited authors. In addition, depicts the author co-citation network obtained by applying a mapping and clustering approach [Sweileh et al., 2017]. In the aforementioned diagram, a total of 5 different groups of authors were identified as the most co-cited.

Author	Number of citations
Bates, D.W.	467
Waters, B.	279
Li, J.	262
Blumenthal, D.	256
Malin, B.	247
Uzuner, O.	220
Sahai, A.	211
Sweeney, L.	210
Blobel, B.	204
Jha, A.K.	196

Table 14: Ten most cited authors

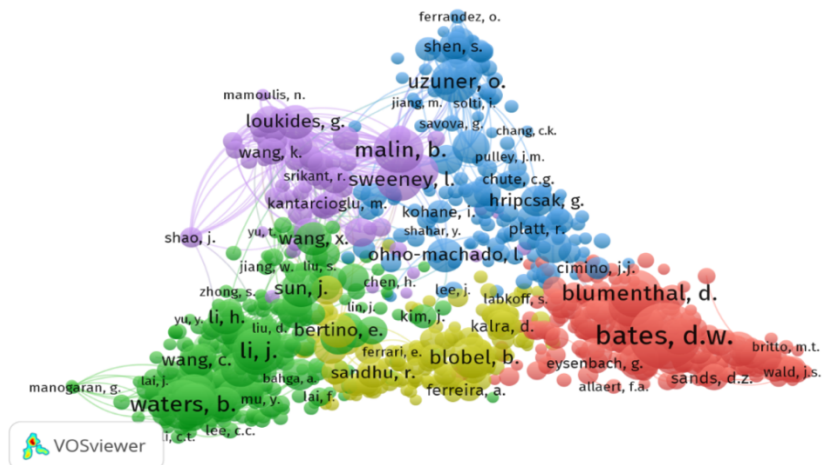


Figure 8: Network map revealing co-citation of authors

4.6. Journal citation analysis

In this subsection, the impact factor of the ten sources with the highest number of citations was studied. Moreover, a co-citation analysis allowed discovering connections between them.

The most relevant journals in privacy and security in EHRs were depicted in this section. The 5-year impact factor of the top ten journals where more publications were published were showed in Table 15. The impact factors were collected from the 2019 Journal Citation Report (JCR) and Scimago Journal Rank (SJR). It can be observed that 7 out of 10 journals were included in the 2019 JCR report and almost the 70% of them were in the first 2 quartiles.

Quartile SJR JCR	Source	2019 SJR	5-year JCT IF
Q3 -	Studies in Health Technology and Informatics	0.27	-
Q2	Journal of Medical Systems	0.69	3.058
Q1 Q2	International Journal of Medical Informatics	0.95	3.025
Q1	Journal of the American Medical Informatics Association	1.83	4.112
Q1 Q2	Journal of Biomedical Informatics	1.14	3.526
-	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	-	-
Q1	Journal of Medical Internet Research	1.19	5.034
Q4 -	Health management technology	0.1	-
Q1	JAMA - Journal of the American Medical Association	5.91	45.54
Q1 Q3	BMC Medical Informatics and Decision Making	0.91	2.317

Table 15: 2019 Impact Factors

5 Discussion

In this section, an analysis of the results has been carried out. For a better organisation of the information, subsections were created for each one of the research questions. The discussion of the results allowed for a detailed study of the research topic and a deeper understanding of its relevance.

- RQ1: What is the growth of literature on privacy and security in EHRs?

Concerning the results of this bibliometric study, a clear upward trend appeared in the publications from 2010 to 2015. In particular, USA was ranked as the first country in academic literature production. More than 30% of the documents were published in this country. Moreover, the period between 2010 and 2015 led to a major increase in the number of publications. Several factors may explain these results. EHRs have received particular attention since 2009. In that year, the Congress of the United States enacted the Health Information Technology for Economic and Clinical Health Act (HITECH).

Finally, the act was passed in the same year. The goals of HITECH were to promote the health information exchange by adopting the use of EHRs [Blumenthal, 2011]. As a result, the investments increase the adoption of EHR. From 2010 to 2012 the use of EHR systems escalated in ambulatory care in USA [Sun et al., 2018]. It is worth noting that the world's best-equipped hospitals can be found in this country. Since 2000, the budget spent in health is double compared with the median of the industrialised countries. Moreover, the grow of the healthcare services in USA has been growing more rapidly in the last decades [Schoen et al., 2006]. The HITECH Act aims at widening the scope of privacy and security protections. As an example, this act requires that patients must be notified when any unsecured breach occurs [Clark and Bilimoria, 2013].

According to the data gathered in this study, the most prolific institutions were mostly American. The Harvard Medical School and the University of California were the two institutions that stood out between the others in number of publications. Harvard owes its solid reputation in part to its prestigious medical school. More than 11,000 faculty members conduct research on diseases such as Parkinson. Moreover, a total of 15 researchers have been involved in work carried out during their time at Harvard Medical School for which a total of 9 Nobel Prizes have been awarded [QS World University Rankings, 2020].

The number of papers published in European countries was also prominent. An important number of contributions arose from the European Union (EU) after the broad reform of personal data protection legislation that came into force in May 2018 [VIORESCU, 2017]. The most prolific author was Bradley Malin, Professor of Biomedical Informatics at the Vanderbilt University. The Vanderbilt University was the third most productive institution according to the results of this study. Professor Malin's research focuses on privacy-enhancing technologies. They allow to enable analysis of organisational, political and health information architectures in the context of the real world. He is also the director of the Health Information Privacy Lab (HIPLab). The aforementioned institution aims to face the increasing need for research in data privacy concerning eHealth technologies [Bradley Malin, 2021].

English remained the most used language. More than 90% of the manuscripts were written in this language. This result is in line with a previous study. According to the International Federation for Documentation, almost 85% of all technological and scientific information in the world is presented in English [Almuhaisen et al., 2020, Sherwood, 2002].

Pattern of authorship found in our study revealed an upward trend in the number of authors per paper. Privacy and security matters in EHR involve mainly three disciplines: medicine, technology and law [Hiller et al., 2011, Schumacher, 2010]. Therefore, multidisciplinary collaborations are required. This type of cooperation was observed in our study, positively reflecting a prevalence with several authors from different academic fields.

- RQ2: What are the hot topics in EHRs' privacy and security papers?

The most common keywords found in this study were electronic health records accompanied by different words, privacy, security, cloud computing and access control. Privacy and security in EHRs is an important matter to face in research since medical

data are highly sensitive information [Mayer et al., 2020]. Human error is one of the main factors causing privacy breaches [Evans et al., 2019]. Both corrective and preventive actions need to be taken to reduce the number of incidents caused by healthcare workers. A survey filled by health staff revealed that 31.7% had ever shared their password with someone else [Fernández-Alemán et al., 2015b]. In a more recent study, similar results were obtained. A total of 73% of medical professionals shared passwords to allow others to get access to EHRs [Hassidim et al., 2017]. Privacy and security requirements are specially critical, and limits the attainment of a good acceptance among the users [Price et al., 2015, Showell, 2017]. Privacy and security concerns, such as unauthorised access to electronic patient records, has prevented these eHealth systems from providing credible and authentic information [Idoga et al., 2016]. In addition, healthcare organisations must recognise the most sensitive information to be protected, impose sanctions, promote security training and establish what constitutes inappropriate behaviour in the use of EHRs. Well-trained staff is as important as technical data protection mechanisms, preventing data breaches and potential threats, moreover, ensuring privacy and security in EHRs [Churi et al., 2021].

The most cited article, with 634 citations, addressed the use of data from the EHRs and the challenges that must be overcome, since the potential to advance medical research and clinical care are relevant [Jensen et al., 2012]. Another high-cited paper introduced the adoption of EHRs by individuals. This study looked at how to persuade users to change their attitudes towards the use of EHRs, and allow the digitisation of their medical information even in the case of significant privacy gaps [Angst and Agarwal, 2009]. In the fourth most cited article the use of big data techniques in EHRs was highlighted [Bates et al., 2014]. Different sources such as clinical, genetic and social data were employed in predictive analytics. New challenges concerning privacy and ethics were involved, requiring ways to preserve privacy unknown until now. As a result, controversy is growing in this area. Some authors argue that patients have a collective responsibility to improve the clinical care [Faden et al., 2013].

There are unprecedented challenges in using big data to reduce the costs of healthcare services. Legislation on privacy and security issues in healthcare, such as the Health Information Portability and Accountability Act (HIPAA), does not address these matters related to linked data sources. By way of example, a systematic literature review on privacy and security in EHRs was carried out [Fernández-Alemán et al., 2013a]. The main goal of the research was to identify and analyse the privacy and security propositions found in the academic literature. The main findings of this study distinguish what may or may not be true privacy and security propositions in EHRs.

- RQ3: What is the collaborative profile of authors in this field of research?

The main international collaboration was between authors from the United Kingdom and the United States. Similarities in legislation and regulations imposed in the past by the EU and the United States may explain this situation. In both countries, the security and confidentiality of health information is of great importance. In the UK, the Data Protection Act 2018 points out how personal information must be employed by organisations, businesses or the government. On the other hand, different laws protect the privacy and security of the data to restrict the exchange of medical information in the USA. The US Congress enacted the Health Insurance Portability and Accountability Act (HIPAA) of 1996. The main goal of this law was to modernise the flow of health

information by stipulating how personal information has to be maintained in the healthcare environment. This law addresses the growing need to safeguard the privacy of the health data. It was created primarily to update the circulation of healthcare-related information, stipulating how personally identifiable information has to be preserved by its users. A set of requirements were proposed in HIPAA that ensure the integrity, confidentiality and availability of electronic medical data. In this act, privacy is the choice of a group of individuals or a single individual to exclude others from accessing their own personal information. In contrast, confidentiality prevents the dissemination of personal or identity information to other unauthorised users, unless there is prior approval [Walsh et al., 2010]. Another law enacted by the 114th United States Congress in December 2016 is the 21st Century Cures Act, which promotes research in the medical area and explains HIPAA privacy guidelines. The Cures Act defines interoperability as the ability to exchange, and use electronic medical data in the patient's record or manage them without special effort and without information blocking [U.S. Department of Health & Human Services, 2012].

In the EU, the GDPR (EU) 2016/679 on data protection and privacy for citizens in the EU and the European Economic Area (EEA) was enacted. The main objective was to allow individuals to have control of the personal data.

- RQ4: What are the most relevant research networks among the authors?

The majority of collaborations found in this study were between authors from the United States. From 7, which contains the authors' research network on privacy and security in EHRs, patterns of collaboration between authors belonging to the same or different institutions or countries were detected. Observation of the diagram allowed to identify the main authors by subject area and the level of collaboration between them. The most relevant groups of authors shown in Figure 7 are grouped by colour and the main topics covered in their work are described below.

Professor Bradley Malin stood out in the group of the most prolific author found in our study. They were focused earlier in the decade on security challenges in HCEs. They worked on Role-Based Access Control (RBAC) [Zhang et al., 2011] and the natural language de-identification framework [Li et al., 2016] as proposed strategies to address these problems. Different methods were analysed to improve the performance of de-identification. Several reasons may explain why the authors paid a lot of interest to this topic. First, EHRs are considered a relevant source of patient information. This information is accessible by authorised parties. However, unless properly managed, these systems may also provide information to those who should not have access. Addressing these weaknesses is a challenge to ensure the privacy and security of EHRs and the personal data stored. In addition, when the access to health data is not controlled, security breaches can happen, compromising privacy and causing harm. Therefore, according to the authors, it is worth further exploring other approaches that could be used in eHealth to access control.

Another group of authors was interested in introducing a method to obscure information on the date of clinical events and patient characteristics aiming at reduce the risk to find individuals through publicly available data [Hripcsak et al., 2016]. In another study, the validity of using EHRs for the research of genetics and heritability of diseases was proved [Polubriaginof et al., 2018]. For a better implementation of precision medicine, it is important to understand the risk factors for each disease, and

one of the most relevant risk factor is family genetic history. Since EHRs contain relatives' information collected through emergency contact forms of patients, it would be a waste of information if it were not used in research within the aforementioned scope. As a result, it could be possible to provide patients with potentially quick and in-depth studies of the heritability of diseases.

On the other hand, another group of collaborators chose to study the impact of the use of EHRs on the quality of health care [Liyanage et al., 2018]. The results of this work showed the barriers, facilitators and outcomes when patients are offered online access to EHR [De Lusignan et al., 2014].

6 Conclusion and Future Work

EHRs adoption has been increasing these past few years. Moreover, several aspects of health services are expected to change accordingly due to the progress in eHealth technologies. Investing in EHR and reassuring the privacy and security of user data should be a priority for governments and organisations. EHR systems should be promoted with the aim of improving health services and helping people and countries to cope with major public health crises.

Encrypting the information is not enough to prevent possible data breaches. Important precautions should be taken as cited in Section 164.530 of the HIPAA privacy rule. This section indicates that the health institutions must offer training for employees, such as students, volunteers and contract employees on privacy, aiming at securing and ensuring the privacy of the patient's data. Since the HIPAA of 1996 went into effect (i.e.: Public Law 104-191), volunteers, hospital staff and providers have put so much effort to guarantee that the national standards in health information protection are met [Carrión et al., 2011b, 2011a, Señor et al., 2012].

The storage of patient medical data is crucial in the healthcare sector. Medical information is critical and therefore the main target of cyber-attacks. Blockchain offer promising solutions for improving the protection of medical data and has attracted interest as a technology to enhance the transparency and authenticity of the medical data across many cases, like for example managing the permissions in EHRs to streamline claims processing. Blockchain-based healthcare systems could also help strengthen the information of the user, allowing to share the health data between medical facilities. These systems could enhance the security and accuracy of health data, as users would control the medical records. Blockchain technology is strong against cyber-attacks and failures, and offers several access control alternatives. Therefore, blockchain can be a suitable solution for health data protection.

The data presented in this paper makes a picture on the current situation of the academic literature on privacy and security in EHRs. Furthermore, the results will serve to compare and document the impact of health technologies in future research. As an example, mHealth is considered an appropriate technology that presents several research lines to study its benefits, such as the continuous evaluation of the patients's health state, detection of abnormal situations or changes in the health conditions and the early detection of emergency situations between others. Thus, the evolution of mHealth provides ubiquity in health services due to the mobile devices that are very common in societies. These devices could expand the current healthcare services, allowing the global monitoring, wide availability and immediacy in healthcare.

These technologies are already widely used in society and have the potential to improve the reach and efficiency of healthcare. However, they create additional privacy issues that must be properly addressed.

The fact that only employing the Scopus database to perform the search could be a limitation. Future work could be performed with other databases, such as Web of Science or PubMed, to widen the study. Differences between databases could be detected if this study is extended. Further study might also focus on specific subtopics (for example, blockchain in healthcare, behavioral risks, and cross-border data protection) or compare research output across regions to identify particular difficulties and improvements.

Current research includes a limited number of empirical studies on real-world data breaches. By identifying this gap, our study provides a solid foundation for future research to explore practical solutions in real healthcare settings.

In conclusion, medical data protection remains a shared responsibility between technology developers, governments and healthcare institutions. Strong technical safeguards must come along with effective and useful training, clear policies, and patient-centered frameworks. Researchers can help construct safer, more trustworthy digital health systems by continuing to explore and improve these areas.

References

- [Almuhaisen et al., 2020] Almuhaisen, O., Habes, M., Alghizzawi, M.: ‘an Empirical Investigation the Use of Information, Communication Technologies To English Language Acquisition: a Case Study From the’; *Novateur Publications International Journal of Innovations in Engineering Research and Technology*, 7, May (2020), 2394–3696.
- [Alqahtani et al., 2021] Alqahtani, F., Winn, A., Orji, R.: ‘Co-designing a mobile app to improve mental health and well-being: Focus group study’; *JMIR Formative Research*, 5, 2 (2021). <https://doi.org/10.2196/18172>
- [Andrés, 2009] Andrés, A.: ‘Measuring Academic Research: How to Undertake a Bibliometric Study’; *Measuring Academic Research: How to Undertake a Bibliometric Study* (2009). <https://doi.org/10.1533/9781780630182>
- [Angst and Agarwal, 2009] Angst, C. M., Agarwal, R.: ‘Adoption of electronic health records in the presence of privacy concerns: The elaboration likelihood model and individual persuasion’; *MIS Quarterly: Management Information Systems*, 33, 2 (2009), 339–370. <https://doi.org/10.2307/20650295>
- [Azaria et al., 2016] Azaria, A., Ekblaw, A., Vieira, T., Lippman, A.: ‘MedRec: Using blockchain for medical data access and permission management’; In *Proceedings - 2016 2nd International Conference on Open and Big Data, OBD 2016* (2016), 25–30. <https://doi.org/10.1109/OBD.2016.11>
- [Bachiri et al., 2018] Bachiri, M., Idri, A., Fernández-Alemán, J. L., Toval, A.: ‘Evaluating the Privacy Policies of Mobile Personal Health Records for Pregnancy Monitoring’; *Journal of Medical Systems*, 42, 8 (2018), 144. <https://doi.org/10.1007/s10916-018-1002-x>
- [Bailón-Moreno et al., 2005] Bailón-Moreno, R., Jurado-Alameda, E., Ruiz-Baños, R., Courtial, J. P.: ‘Bibliometric laws: Empirical flaws of fit’; *Scientometrics*, 63, 2 (2005), 209–229. <https://doi.org/10.1007/s11192-005-0211-5>
- [Bates et al., 2014] Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., Escobar, G.: ‘Big data in health care: Using analytics to identify and manage high-risk and high-cost patients’; *Health*

Affairs, 33, 7 (2014), 1123–1131. <https://doi.org/10.1377/hlthaff.2014.0041>

[Ben-Assuli, 2015] Ben-Assuli, O.: ‘Electronic health records, adoption, quality of care, legal and privacy issues and their implementation in emergency departments’; *Health Policy* (2015), 287–297. <https://doi.org/10.1016/j.healthpol.2014.11.014>

[Blumenthal, 2011] Blumenthal, D.: ‘Implementation of the Federal Health Information Technology Initiative’; *New England Journal of Medicine*, 365, 25 (2011), 2426–2431. <https://doi.org/10.1056/nejmsr1112158>

[Bradley Malin, 2021] Bradley Malin: ‘CV’; (2021).

[Broadus, 1987] Broadus, R. N.: ‘Toward a definition of ‘bibliometrics’’; *Scientometrics*, 12, 5–6 (1987), 373–379. <https://doi.org/10.1007/BF02016680>

[Burnham, 2006] Burnham, J. F.: ‘Scopus database: A review’; *Biomedical Digital Libraries*, 3 (2006), 1–8. <https://doi.org/10.1186/1742-5581-3-1>

[Carrión et al., 2011a] Carrión, I., Alemán, J. L. F., Toval, A.: ‘Assessing the HIPAA standard in practice: PHR privacy policies’; *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (2011a), 2380–2383. <https://doi.org/10.1109/IEMBS.2011.6090664>

[Carrión et al., 2011b] Carrión, I., Fernández-Alemán, J. L., Toval, A.: ‘Usable privacy and security in personal health records’; *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6949 LNCS, PART 4 (2011b), 36–43. https://doi.org/10.1007/978-3-642-23768-3_3

[Chernyshev et al., 2019] Chernyshev, M., Zeadally, S., Baig, Z.: ‘Healthcare Data Breaches: Implications for Digital Forensic Readiness’; *Journal of Medical Systems*, 43, 1 (2019). <https://doi.org/10.1007/s10916-018-1123-2>

[Churi et al., 2021] Churi, P., Pawar, A., Moreno-Guerrero, A. J.: ‘A comprehensive survey on data utility and privacy: Taking indian healthcare system as a potential case study’; *Inventions*, 6, 3 (2021), 45. <https://doi.org/10.3390/inventions6030045>

[Clark and Bilimoria, 2013] Clark, L. W., Bilimoria, N. M.: ‘How HIPAA final rules affect health information technology vendors.’; *The Journal of Medical Practice Management : MPM*, 29, 1 (2013), 56–58.

[Cochran and William, 1977] Cochran, W. G., William, G.: ‘Sampling Techniques. New York: John Wiley & Sons’; Inc (1977).

[Collins et al., 2011] Collins, J. D., Sainato, V., Khey, D. N.: ‘Organizational Data Breaches 2005-2010: Applying SCP to the Healthcare and Education Sectors’; *International Journal of Cyber Criminology*, 5, July (2011), 794–810.

[De Lusignan et al., 2014] De Lusignan, S., Mold, F., Sheikh, A., Majeed, A., Wyatt, J. C., Quinn, T., et al.: ‘Patients’ online access to their electronic health records and linked online services: a systematic interpretative review’; *BMJ Open*, 4, 9 (2014), e006021.

[Dohan et al., 2014] Dohan, M. S., Abouzahra, M., Tan, J.: ‘Mobile personal health records: Research agenda for applications in global health’; *Proceedings of the Annual Hawaii International Conference on System Sciences* (2014), 2576–2585. <https://doi.org/10.1109/HICSS.2014.325>

[European Union, 2016] European Union: ‘European Union General Data Protection Regulation & Directive’; (2016), 1–5.

[Evans et al., 2019] Evans, M., He, Y., Maglaras, L., Janicke, H.: ‘HEART-IS: A novel technique

for evaluating human error-related information security incidents'; *Computers and Security*, 80 (2019), 74–89. <https://doi.org/10.1016/j.cose.2018.09.002>

[Faden et al., 2013] Faden, R. R., Kass, N. E., Goodman, S. N., Pronovost, P., Tunis, S., Beauchamp, T. L.: 'An Ethics Framework for a Learning Health Care System: A Departure from Traditional Research Ethics and Clinical Ethics'; *Hastings Center Report*, 43, SUPPL. 1 (2013). <https://doi.org/10.1002/hast.134>

[Fernández-Alemán et al., 2015a] Fernández-Alemán, J. L., Sánchez-Henarejos, A., García-Amicis, V. M., Toval, A., Sánchez-García, A. B., Hernández-Hernández, I.: 'Estudio sobre la importancia y la seguridad de uso de las contraseñas en el ámbito laboral sanitario'; *Gaceta Sanitaria*, 29, 1 (2015a), 72–79. <https://doi.org/10.1016/j.gaceta.2014.07.003>

[Fernández-Alemán et al., 2015b] Fernández-Alemán, J. L., Sánchez-Henarejos, A., Toval, A., Sánchez-García, A. B., Hernández-Hernández, I., Fernandez-Luque, L.: 'Analysis of health professional security behaviors in a real clinical setting: An empirical study'; *International Journal of Medical Informatics*, 84, 6 (2015b), 454–467. <https://doi.org/10.1016/j.ijmedinf.2015.01.010>

[Fernández-Alemán et al., 2013a] Fernández-Alemán, J. L., Señor, I. C., Lozoya, P. ángel O., Toval, A.: 'Security and privacy in electronic health records: A systematic literature review'; *Journal of Biomedical Informatics*, 46, 3 (2013a), 541–562. <https://doi.org/10.1016/j.jbi.2012.12.003>

[Fernández-Alemán et al., 2013b] Fernández-Alemán, J. L., Seva-Llor, C. L., Toval, A., Ouhbi, S., Fernández-Luque, L.: 'Free web-based personal health records: An analysis of functionality'; *Journal of Medical Systems*, 37, 6 (2013b). <https://doi.org/10.1007/s10916-013-9990-z>

[Flaumenhaft and Ben-Assuli, 2018] Flaumenhaft, Y., Ben-Assuli, O.: 'Personal health records, global policy and regulation review'; *Health Policy* (2018), 815–826. <https://doi.org/10.1016/j.healthpol.2018.05.002>

[Hassidim et al., 2017] Hassidim, A., Korach, T., Shreberk-Hassidim, R., Thomaidou, E., Uzefovsky, F., Ayal, S., Ariely, D.: 'Prevalence of sharing access credentials in electronic medical records'; *Healthcare Informatics Research*, 23, 3 (2017), 176–182. <https://doi.org/10.4258/hir.2017.23.3.176>

[Heart et al., 2017] Heart, T., Ben-Assuli, O., Shabtai, I.: 'A review of PHR, EMR and EHR integration: A more personalized healthcare and public health policy'; *Health Policy and Technology*, 6, 1 (2017), 20–25. <https://doi.org/10.1016/j.hlpt.2016.08.002>

[Hiller et al., 2011] Hiller, J., McMullen, M. S., Chumney, W. M., Baumer, D. L.: 'Privacy and Security in the Implementation of Health Information Technology'; *B.U. J. Sci. & Tech.*, 1, 1 (2011), 40.

[Hripcsak et al., 2016] Hripcsak, G., Mirhaji, P., Low, A. F. H., Malin, B. A.: 'Preserving temporal relations in clinical data while maintaining privacy'; *Journal of the American Medical Informatics Association*, 23, 6 (2016), 1040–1045. <https://doi.org/10.1093/jamia/ocw001>

[Idoga et al., 2016] Idoga, P. E., Agoyi, M., Coker-Farrell, E. Y., Ekeoma, O. L.: 'Review of security issues in e-Healthcare and solutions'; In *13th HONET-ICT International Symposium on Smart MicroGrids for Sustainable Energy Sources Enabled by Photonics and IoT Sensors, HONET-ICT 2016* (2016), 118–121. <https://doi.org/10.1109/HONET.2016.7753433>

[Jensen et al., 2012] Jensen, P. B., Jensen, L. J., Brunak, S.: 'Mining electronic health records: Towards better research applications and clinical care'; *Nature Reviews Genetics*, 13, 6 (2012), 395–405. <https://doi.org/10.1038/nrg3208>

- [Kruse et al., 2017] Kruse, C. S., Frederick, B., Jacobson, T., Monticone, D. K.: 'Cybersecurity in healthcare: A systematic review of modern threats and trends'; *Technology and Health Care* (2017), 1–10. <https://doi.org/10.3233/THC-161263>
- [Li, 2015] Li, J.: 'Ensuring privacy in a personal health record system'; *Computer*, 48, 2 (2015), 24–31. <https://doi.org/10.1109/MC.2015.43>
- [Li et al., 2016] Li, M., Carrell, D., Aberdeen, J., Hirschman, L., Kirby, J., Li, B., et al.: 'Optimizing annotation resources for natural language de-identification via a game theoretic framework'; *Journal of Biomedical Informatics*, 61 (2016), 97–109. <https://doi.org/10.1016/j.jbi.2016.03.019>
- [Liyanage et al., 2018] Liyanage, H., Liaw, S. T., Konstantara, E., Mold, F., Schreiber, R., Kuziemy, C., et al.: 'Benefit-risk of Patients' Online Access to their Medical Records: Consensus Exercise of an International Expert Group'; *Yearbook of Medical Informatics*, 27, 1 (2018), 156–162. <https://doi.org/10.1055/s-0038-1641202>
- [Ma et al., 2013] Ma, Y., Liu, J., Liu, W.: 'Security and privacy issues in electronic health network'; *Wuhan University Journal of Natural Sciences*, 18, 6 (2013), 523–529. <https://doi.org/10.1007/s11859-013-0967-z>
- [Mayer et al., 2020] Mayer, A. H., da Costa, C. A., Righi, R. da R.: 'Electronic health records in a Blockchain: A systematic review'; *Health Informatics Journal*, 26, 2 (2020), 1273–1288. <https://doi.org/10.1177/1460458219866350>
- [McLeod and Dolezel, 2018] McLeod, A., Dolezel, D.: 'Cyber-analytics: Modeling factors associated with healthcare data breaches'; *Decision Support Systems*, 108 (2018), 57–68. <https://doi.org/10.1016/j.dss.2018.02.007>
- [Menachemi and Collum, 2011] Menachemi, N., Collum, T. H.: 'Benefits and drawbacks of electronic health record systems'; *Risk Management and Healthcare Policy*, 4 (2011), 47–55. <https://doi.org/10.2147/RMHP.S12985>
- [Patra et al., 2006] Patra, S. K., Bhattacharya, P., Verma, N.: 'Bibliometric Study of Literature on Bibliometrics'; *DESIDOC Bulletin of Information Technology*, 26, 1 (2006), 27–32. <https://doi.org/10.14429/dbit.26.1.3672>
- [Plantier et al., 2017] Plantier, M., Havet, N., Durand, T., Caquot, N., Amaz, C., Biron, P., et al.: 'Does adoption of electronic health records improve the quality of care management in France? Results from the French e-SI (PREPS-SIPS) study'; *International Journal of Medical Informatics*, 102 (2017), 156–165. <https://doi.org/10.1016/j.ijmedinf.2017.04.002>
- [Polubriaginof et al., 2018] Polubriaginof, F. C. G., Vanguri, R., Quinnes, K., Belbin, G. M., Yahi, A., Salmasian, H., et al.: 'Disease Heritability Inferred from Familial Relationships Reported in Medical Records'; *Cell*, 173, 7 (2018), 1692–1704.e11. <https://doi.org/10.1016/j.cell.2018.04.032>
- [Price et al., 2015] Price, M., Bellwood, P., Kitson, N., Davies, I., Weber, J., Lau, F.: 'Conditions potentially sensitive to a Personal Health Record (PHR) intervention, a systematic review Healthcare Information Systems'; *BMC Medical Informatics and Decision Making*, 15, 1 (2015), 1–12. <https://doi.org/10.1186/s12911-015-0159-1>
- ['Proceedings of the 39th Hawaii International Conference on System Sciences - 2006', 2006] In *Proceedings of the Annual Hawaii International Conference on System Sciences (Vol. 7)*. Kauai, HI (2006).
- [QS World University Rankings, 2020] QS World University Rankings: 'Top Universities in Life Sciences and Medicine 2020'; (2020).

[Schoen et al., 2006] Schoen, C., Davis, K., How, S. K. H., Schoenbaum, S. C.: 'U.S. health system performance: A national scorecard'; *Health Affairs*, 25, 6 (2006), 457–475. <https://doi.org/10.1377/hlthaff.25.w457>

[Schumacher, 2010] Schumacher, R. M.: 'Commentary: Electronic health records and human performance'; *Journal of Oncology Practice*, 6, 3 (2010), 125–126. <https://doi.org/10.1200/JOP.091098>

[Señor et al., 2012] Señor, I. C., Alemán, J. L. F., Toval, A.: 'Personal health records: New means to safely handle health data?'; *Computer*, 45, 11 (2012), 27–33. <https://doi.org/10.1109/MC.2012.285>

[Shahnaz et al., 2019] Shahnaz, A., Qamar, U., Khalid, A.: 'Using Blockchain for Electronic Health Records'; *IEEE Access*, 7 (2019), 147782–147795. <https://doi.org/10.1109/ACCESS.2019.2946373>

[Sherwood, 2002] Sherwood, B.: 'Review of Ammon (2001): The Dominance of English as a Language of Science: Effects on Other Languages and Language Communities'; *Language Problems and Language Planning* (Vol. 26). Mouton de Gruyter (2002). <https://doi.org/10.1075/lplp.26.2.08she>

[Showell, 2017] Showell, C.: 'Barriers to the use of personal health records by patients: A structured review'; *PeerJ*, 2017, 4 (2017), e3268. <https://doi.org/10.7717/peerj.3268>

[Sun et al., 2018] Sun, J., Garcia, J. A., Wang, Y.: 'Ambulatory EMR adoption in the USA: A longitudinal study'; *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2018-Janua (2018), 2855–2864. <https://doi.org/10.24251/hicss.2018.361>

[Sweileh et al., 2017] Sweileh, W. M., Al-Jabi, S. W., AbuTaha, A. S., Zyoud, S. H., Anayah, F. M. A., Sawalha, A. F.: 'Bibliometric analysis of worldwide scientific literature in mobile - health: 2006-2016'; *BMC Medical Informatics and Decision Making*, 17, 1 (2017), 1–12. <https://doi.org/10.1186/s12911-017-0476-7>

[U.S. Department of Health & Human Services, 2012] U.S. Department of Health & Human Services: 'The Office of the National Coordinator for Health Information Technology'; (2012), 1–22.

[VIORESCU, 2017] VIORESCU, R.: '2018 Reform Of Eu Data Protection Rules'; *European Journal of Law and Public Administration* (2017), 27–39. <https://doi.org/10.18662/eljpa/11>

[Walsh et al., 2010] Walsh, D., Passerini, K., Varshney, U., Fjermestad, J.: 'Legal Issues in the Transition to Electronic Records in Health Care'; *Information Systems: People, Organizations, Institutions, and Technologies* (2010), 321–326. <https://doi.org/10.1007/978-3-7908-2148-2>

[Zhang and Liu, 2010] Zhang, R., Liu, L.: 'Security models and requirements for healthcare application clouds'; In *Proceedings - 2010 IEEE 3rd International Conference on Cloud Computing*, CLOUD 2010 (2010), 268–275. <https://doi.org/10.1109/CLOUD.2010.62>

[Zhang et al., 2011] Zhang, W., Gunter, C. A., Liebovitz, D., Tian, J., Malin, B.: 'Role prediction using Electronic Medical Record system audits.'; In *AMIA ... Annual Symposium proceedings / AMIA Symposium*. AMIA Symposium (Vol. 2011) (2011), 858–867.