University of North Dakota
UND Scholarly Commons

Theses and Dissertations | Theses, Dissertations, and Senior Projects

1-1-2023

Application Of Blockchain Technology And Integration Of Differential Privacy: Issues In E-Health Domains

David Isie

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://commons.und.edu/theses

Recommended Citation

Isie, David, "Application Of Blockchain Technology And Integration Of Differential Privacy: Issues In E-Health Domains" (2023). *Theses and Dissertations*. 5676.
https://commons.und.edu/theses/5676

This Thesis is brought to you for free and open access by the Theses, Dissertations, and Senior Projects at UND Scholarly Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UND Scholarly Commons. For more information, please contact und.commons@library.und.edu.

APPLICATION OF BLOCKCHAIN TECHNOLOGY AND INTEGRATION OF DIFFERENTIAL PRIVACY: ISSUES IN E-HEALTH DOMAINS

by

David Friday Isie
Bachelor of Science, New Mexico Institute of Mining & Technology, 2018

A Thesis

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Science

Grand Forks, North Dakota
December
2023

Copyright 2023 David Isie

This document, submitted in partial fulfillment of the requirements for the degree from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

DocuSigned by:
Hassan Reza

DocuSigned by:
Eunjin Kim

DocuSigned by:
Wen-Chen Hu

This document is being submitted by the appointed advisory committee as having met all the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

DocuSigned by:
Chris Nelson

Chris Nelson
Dean of the School of Graduate Studies

11/27/2023

Date

PERMISSION

Title Application of Blockchain Technology and Integration of Differential Privacy: Issues in E-Health Domains

Department School of Electrical Engineering & Computer Science

Degree Master of Science

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

David Isie
04/14/2023

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	**Definitions**
EMR	Electronic Medical Record
BT	Blockchain Technology
DP	Differential Privacy
BC	Blockchain
SMS	Systematic Mapping Study
PHI	Protected Health Information
ePHI	Electronic Protected Health Information
HIPAA	Health Insurance Portability and Accountability Act
CIA	Confidentiality Integrity Availability
IoT	Internet of Thing
RQ	Research Question
HIT	Health Information Technology
PoW	Proof of Work
PoV	Proof of Vote
B2B	Business-to-Business
GDPR	General Data Protection Regulation
ROI	Return on Investment
FL	Federated Learning
SMS	Systematic Mapping Study
QAQ	Quality Assessment Question

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to the members of my advisory committee for their guidance and support during my time in the master's program at the University of North Dakota.

To my late eldest sister Christiana Isie and my late father Friday Isie

To my mom Adiaha Isie

The world’s best family

ABSTRACT

A systematic and comprehensive review of critical applications of Blockchain Technology with Differential Privacy integration lies within privacy and security enhancement. This paper aims to highlight the research issues in the e-Health domain (e.g., EMR) and to review the current research directions in Differential Privacy integration with Blockchain Technology.

Firstly, the current state of concerns in the e-Health domain are identified as follows: (a) healthcare information poses a high level of security and privacy concerns due to its sensitivity; (b) due to vulnerabilities surrounding the healthcare system, a data breach is common and poses a risk for attack by an adversary; and (c) the current privacy and security apparatus needs further fortification.

Secondly, Blockchain Technology (BT) is one of the approaches to address these privacy and security issues. The alternative solution is the integration of Differential Privacy (DP) with Blockchain Technology.

Thirdly, collections of scientific journals and research papers, published between 2015 and 2022, from IEEE, Science Direct, Google Scholar, ACM, and PubMed on the e-Health domain approach are summarized in terms of security and privacy. The methodology uses a systematic mapping study (SMS) to identify and select relevant research papers and academic journals regarding DP and BT.

With this understanding of the current privacy issues in EMR, this paper focuses on three categories: (a) e-Health Record Privacy, (b) Real-Time Health Data, and (c) Health Survey Data Protection. In this study, evidence exists to identify inherent issues and technical challenges associated with the integration of Differential Privacy and Blockchain Technology.

Keywords**:** e-Health domain, Differential Privacy, Blockchain, IoT, real-time data, health survey, electronic medical record

CHAPTER 1

INTRODUCTION

1.1 Overview

The evolving nature of the e-Health domain (e.g., EMR) in recent years has drawn government attention to address the issues surrounding the privacy and security of EMR. The Health Insurance Portability and Accountability Act (HIPAA) was introduced in 1996 as a federal law to regulate three significant components of healthcare data as follows [1]: (a) HIPAA Privacy Rules: Regulates the disclosure and use of Protected Health Information (PHI) by entities such as employer-sponsored health plans, health insurers, and transactions that involve medical services; (b) Security Rules: Specifically designed to address Electronic Protected Health Information (ePHI) and to safeguard three security compliances which are administrative, physical, and technical; and (c) Breach Notification Rules: Requires organizations to report an incident of PHI breach to patients. Confidentiality, Integrity, and Availability (CIA) strongly correlate with HIPAA compliance and must be implemented. Confidentiality means the privacy of PHI is ensured. Integrity means PHI is only changed or destroyed with due process. Availability means PHI remains accessible by keeping hardware and systems in good working condition [2].

This research focuses on privacy issues in e-Health domains (e.g., EMR) and the review of Blockchain Technology and Differential Privacy to address these vulnerabilities. Although Blockchain is still evolving, particularly in the e-Health system, its adoption has multiplied recently as more Internet of Things (IoT) uses electronic gadgets to manage and provide patient services [3]. Blockchain applications also apply in other industries like finance, supply chain,

insurance claim, clinical trial, and pharmaceutical counterfeit [4]. Therefore, this paper aims to review privacy issues in the e-Health domain using Blockchain Technology and the integration of Differential Privacy (DP).

1.2 Research Motivations

E-Health systems' privacy and security issues have triggered the need to explore the loopholes or vulnerabilities in handling, sharing, storing, and accessing patients' ePHI. The following are the current issues cited to back up the motivation of this research.

- Surveys have shown that many people are concerned about healthcare information privacy. Close to two-thirds of clients paid attention to the privacy of personal healthcare, and 39% of respondents assumed that their health data is safe [5].
- Some people are concerned that their healthcare data is not safe via the internet, and they are worried about security and privacy vulnerability [6].
- About half of the research participants believe that exchanging their medical records is not in their best interest to secure their privacy [7].
- In 2021, the Department of Health and Human Services Office for Civil Rights (OCR) implemented corrective action to settle potential violations of HIPAA, which included a privacy and security rules-related data breach that affected 9.3 million people [8].
- The existing EMR systems show that about 40% of physicians identified the design and interoperability as primary sources of dissatisfaction (sample size of 8,774 physicians) [9].

Blockchain and Differential Privacy are believed to provide solutions to mitigate these privacy issues. The benefit of Blockchain Technology spans healthcare systems to provide or reduce potential data breaches and unauthorized access or sharing of patients' PHI [10].

1.3 Problem Statements

This paper aims to evaluate the potential of using Differential Privacy as a complementary layer to enhance privacy protection in the e-Health domain, specifically in Electronic Medical Records (EMR) management systems. Despite the decentralized nature of Blockchain technology, it has been shown to have limitations in providing adequate privacy protection for users' sensitive personal health information. This is particularly important in today's digital age, where data breaches are increasingly common, and personal health information has become a commodity. The proposed integration of Blockchain and Differential Privacy aims to address these limitations by providing a more secure and private system for managing EMR. This study seeks to fill the current literature gap by evaluating this integration's effectiveness in terms of privacy and security and its potential for implementation in real-world e-Health systems.

1.4 Research Questions (RQs)

The Research Questions (RQs) are formulated based on the research motivations, problem statements, and the goal of this review. Table 1.1 below summarizes the research questions (RQs).

Table 1.1 Research Questions (RQs)

ID	Research Questions
RQ1	How can DP be integrated into BC to enhance privacy and security in the e-Health domain (e.g., EMR)?
RQ2	What factors contribute to the DP mechanisms integration in Blockchain Technology and associated issues?
RQ3	What types of datasets and programming languages are being considered for implementation?

RQ4	What are the limitations and inherent challenges of the BT and DP applications, and how can they be solved?

a. **Note that the above questions are narrowed to only e-Health domains**

1.5 Scope of Work

This research is focused on enhancing the privacy and security aspects within E-Health Domains, specifically Electronic Medical Records (EMRs). The proposed approach involves the integration of two distinct technologies, Blockchain and Differential Privacy, to bolster the security and privacy features of EMRs. The primary objective of this thesis is to assess how Blockchain Technology can effectively secure EMRs, fostering improved data sharing among healthcare providers, patients, and researchers. The study also delves into the potential of Differential Privacy techniques to safeguard patients' privacy in EMRs while enabling meaningful data analysis. Furthermore, the envisioned solution includes the development of a framework or proof-of-concept implementation that merges Blockchain and Differential Privacy, addressing specific challenges in EMR, such as data security, privacy, interoperability, data quality, and data governance. Consequently, the outcomes of this work aim to offer valuable insights for researchers and healthcare organizations striving to enhance the privacy and security of their EMR systems

1.6 Limitations of Work

The limitations of this work are inherent constraints that influence the outcome of this research. The inherent limitations are crucial for maintaining the integrity of the research as are stated bellow :

- This thesis explores data availability and quality issues, as EMR can be fragmented, incomplete, or inconsistent across different healthcare settings and systems.

- This thesis encounters regulatory and ethical barriers, as EMR contain sensitive personal health information subject to various privacy and security laws and regulations.

- This thesis may demand more generalizability, as the findings and recommendations may be specific to the healthcare context, Blockchain, and Differential Privacy tools and techniques used.

- This thesis may require multidisciplinary expertise and collaboration, combining computer science, statistics, healthcare, law, and ethics knowledge and skills.

1.7 Thesis Structure

This study's research organization and strategy consist of the following: (a) chapter 2 provides a comprehensive analysis of the existing literature related to the research question, research motivations, and problem statements to identify the gaps in the literature and the research questions that need further exploration to propose feasible solutions; (b) chapter 3 presents this study's methodology which consists of different research steps using the SMS and the selection processes of papers and publications related to the research objective and questions. Steps in this section include search items, literature sources, search process, selection, and study quality analysis. The research questions are framed according to three categories: Real-Time Health Data, Electronic Medical Records (EMR) Privacy, and Health Survey Data Protection; (c) chapter 4 is a presentation of the results and analysis where all the findings are listed and explained based on the three categories; (d) chapter 5 discusses the interpretation of the results, their implications, and relation to the literature review; and (e) chapter 6 includes the study's main findings and their significance, limitations, future research suggestions, and a conclusion.

1.8 Research Contributions

Based on the literature review regarding various research papers, the claims listed below summarize the contributions based on the findings.

- This paper presents three main categories of research work that streamlines inherent limitations and technical challenges in EMRs
- Twenty papers identify primary studies related to security and privacy of EMRs. Other healthcare organizations and researchers can use this list of studies to enhance their work and studies
- Additional selections from primary studies meeting the inclusion criteria are designated for quality assessment. These studies serve as a foundation for the integration of blockchain and differential privacy, involving intricate and advanced mathematical computations.
- The majority of literature could benefit from enhanced intuitiveness, elucidating the link between the academic foundation and the practical application of Differential Privacy. Additionally, a deeper understanding of researchers' and developers' expectations is necessary.

CHAPTER 2

BACKGROUND

2.1 Blockchain Technology: History and Concept

In 2008, Blockchain Technology emerged as an innovative tool designed for cryptocurrency management, with its foundational concept as a distributed ledger introduced by S. Nakamoto [11]. [12] further details that Blockchain relies on a hash-based proof-of-work chain. To comprehend the applications of Blockchain in e-Health, familiarity with Health Information Technology (HIT) is essential. Information Technology has progressively evolved, becoming an integral component of e-Health systems, which consist of various elements, with privacy and security as their defining characteristics. An illustrative example of e-Health is Electronic Medical Records (EMR), described in [13] as an electronic replica of the traditional paper copy containing comprehensive medical information, including the patient's treatment history.

The term "Blockchain Technology" carries a broad definition, contingent upon its intended purpose and application in specific contexts. This variability in definitions can pose challenges to a comprehensive understanding of the technology within a given context. The scope of this definition spans from Bitcoin Blockchain and smart contract Blockchain to distributed ledger Blockchain. Common characteristics across these definitions, serving as a means of data storage, typically encompass: (a) handling financial transactions, (b) replication across multiple systems, (c) establishment of a peer-to-peer network, (d) integration of cryptography and digital signature, (e) involvement of both writers and readers as active participants, and (f) a mechanism for maintaining historical records that is resistant to change [14]. Consequently, when applied in the

realm of Electronic Medical Records (EMR), Blockchain Technology functions as a distributed digital ledger, replicating across diverse nodes or Internet of Things (IoT) devices, such as computer systems not interlinked based on memory addresses [15]. Within this context, the functionalities of computer systems collectively operate as a full node, encompassing tasks such as storing, listing, validating transactions, and participating in the creation or mining of blocks

Blockchain can be envisioned as sequentially connected pages in a notebook. Each page follows the one before it, creating a continuous chain of information. In this analogy, each page represents a block, and these blocks are replicated across the entire system to form the blockchain [16]. Every block contains textual data and self-descriptive information. The connection between blocks is established through an algorithmic fingerprint, ensuring the security of data within each block. This algorithmic consistency fosters a consensus among all users regarding access to recorded data within each block. As mentioned earlier, links are employed between blocks, eliminating the need for memory addresses. Instead, these links constitute a network of cryptographic "hashings" that safeguard data integrity. The Blockchain Life Cycle initiates when a user initiates a transaction request, which is then transmitted to the network for validation. Upon acceptance, the transaction is incorporated into the existing transaction block. Conversely, if the request is rejected, the transaction is omitted. Upon acceptance, the block of transactions is linked to previous transaction blocks, confirming its validity. This cycle repeats with each subsequent transaction request, as depicted in Figure 2.1.

The process of the Blockchain Life Cycle has a unique feature such that the exchange of monetary or digital value units occur with no intermediaries involved. For example, in cryptocurrencies, as mentioned earlier, and in other applications, an authorized user can sell a digital asset on a

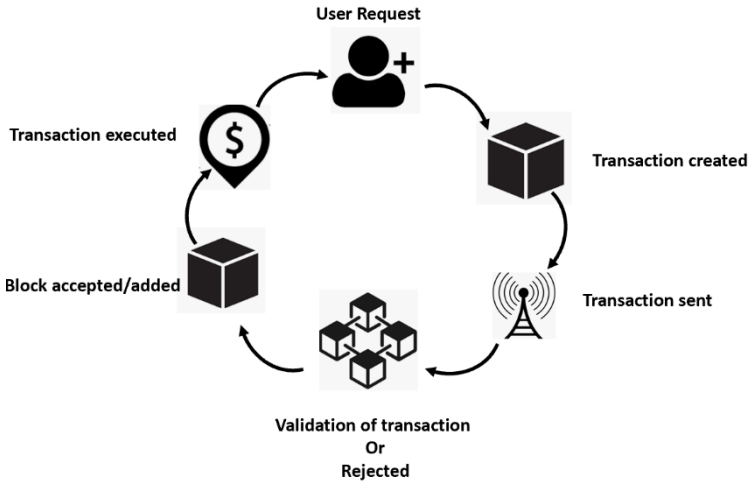

Figure 2.1 Blockchain Life Cycle

marketplace or transfer land properties without a notary [17]. In the healthcare industry, the exchange usually occurs through sharing or storing personal health information. Access to personal health information is shared, immutable, and transparent to all participants, creating a consensus without a centralized entity to manage these operations. Managing Blockchain applications in healthcare raises many concerns due to the susceptibility to privacy and security issues.

2.1.1 Types of Blockchain

Blockchain incorporates three authentication and control mechanisms: public, private, and consortium. Public authentication, exemplified by systems like Bitcoin and Ethereum, is decentralized and permissionless. For instance, Ethereum is implemented as a permissionless and programmable Blockchain, allowing any user to create and execute complex algorithms on the platform. In Ethereum, consensus is achieved through proof of work (PoW), where hashing is utilized to validate new blocks. However, PoW involves energy-intensive mining activities to meet

specific requirements, leading to a significant energy expenditure. This energy inefficiency makes the PoW consensus algorithm less environmentally sustainable and not a favorable approach [18][19].

On the flip side, private and consortium authentication mechanisms are constrained and regulated, necessitating permission. Private authentication stands out, especially in the context of the Hyperledger Fabric platform. Consequently, private authentication is deemed most suitable for ensuring Electronic Medical Record (EMR) security, as it mandates authorization for users to join the platform. The Hyperledger Fabric platform finds optimal use cases in Business-to-Business (B2B) data exchange, transaction settlement, and non-repudiation within Blockchain applications. In the healthcare sector, Blockchain's application for securing EMRs focuses on achieving non-repudiation, with the management of patients' EMRs holding significant growth potential. EMRs encapsulate a patient's comprehensive medical information, detailing their condition and clinical progress throughout treatment [20]. The advantages of employing a Blockchain-based network for EMRs include the decentralized storage of records, the absence of a centralized owner vulnerable to hacking, and the capacity for data updates [20]. In Hyperledger, chain-code services play a pivotal role in ensuring secure execution of smart contracts. These contracts, defined by logical rules, govern transactions executed with the associated World State. In this context, State refers to a database that stores data in arrays of arbitrarily assigned keys [27]. The paramount challenge encountered in Blockchain applications pertaining to EMRs is the preservation of security and privacy.

2.1.2 Characteristics of Blockchain

Blockchain Technology contains distinctive properties that make it suitable for EMR systems, which are listed below.

- **Decentralization:** This is a peer-to-peer transaction without a centralized validation or authorization system. The access is granted to each participant with the full right to verify transactions within the network [22]. To decentralize the network, technology such as cryptographic hash, digital signature, and distributed consensus mechanisms is required for security fortification. The consensus protocol is to ensure data integrity. Therefore, decentralization enhances protection against vulnerability in the network at risk of security attacks [23].

- **Immutability and Transparency:** This concept means that after creating and adding the block, it cannot be removed or changed [24].

- **Auditability:** Any transaction in the Blockchain network is traceable to its previous transaction. Therefore, the timestamp is incorporated into transaction validation and records [25].

- **Smart Contract:** This is based on certain conditions; when met, it is automatically filed and executed, such as control accesses and privileges [23].

- **Security:** By design, the Blockchain network uses a private or public key to access or make transactions. This is due to hashing that seals each block from a third party [25].

- **Computational Logic:** This is a feature in smart contracts in a Blockchain network that automatically triggers transactions [22].

2.1.3 Blockchain Benefits in EMR

The existing healthcare systems contain evidence suggesting that Blockchain presents inherent problems. The management and exchange of patients' data is a focal point for Blockchain applications. Other applications allow healthcare data to be distributed and immutable for greater security of patient records and data integrity [26]. Table 2.1 below shows a significant and brief summary of Blockchain Technology's benefits.

Table 2.1 Blockchain Benefits

Benefit		References
Transparency	Due to Blockchain immutability, data cannot be deleted or altered. Blockchain is a more transparent system that stores EMR.	[25][23]
Data Integrity	Blockchain ensures data integrity so that no centralized authority is at risk of security attacks.	[23][27]
Security	EMR is sensitive data and such Blockchain provides encryption capabilities that minimize attacks and protect vulnerably.	[25][27][23]
Interoperability	Decentralization helps to improve interoperability that facilitates the exchange of EMR and grants patients ownership and access control of their record.	[28][25]
Patient-Centered	The right of patients to access or grant access to authorized personnel in EMR systems is restored.	[23][29][30]

2.1.4 Limitations of Blockchain and Difficulties to Integrate into EMRs

The limitations of Blockchain in EMR are inherent and span through different categories. The limitations are grouped into the following: (a) privacy, (b) scalability, and (c) usability [123].

Privacy: The Blockchain system is encrypted, but it is possible for an adversary to interfere with patients' personal information. The primary limitation of Blockchain is that 51% of mining nodes on the EMR could result in the rewriting of the chain structure. Whereas, to achieve the advantage of a decentralized system, at least 50% of mining nodes trust is required from the participants to sustain the immutability of the Blockchain [124]. The second limitation occurs during Blockchain transaction analysis. For instance, permissioned Blockchain reduces unauthorized access to EMR, but it cannot mask the record of the transaction which allows unfavorable network analysis. As a result, an adversary may be able to determine a specific node visited by a physician or the provider [124]. The inability of Blockchain to erase patient EMR is another limitation [126]. As such, it is difficult for Blockchain to comply with regulations such as General Data Protection Regulation (GDPR) Article 17 of the GDPR.

Scalability: Blockchain in EMR is unable to store large file sizes (e.g., medical images), thereby slowing down the confirmation of the transaction process, especially when live streaming data is required from an Internet of Things device [125]. As a result, large data itself must be stored outside the Blockchain network which could be vulnerable to attack.

Usability: The affordability of Blockchain is another limitation as users are required to pay transaction fees [127]. As explained by Charanya et al. [128], a conventional EMR system contains password recovery mechanisms, but in a Blockchain EMR system, a patient would have difficulty accessing their record if their private key is lost.

2.2 Differential Privacy: History and Concept

Differential Privacy (DP) made its debut in 2006 through the paper titled "Calibrating Noise to Sensitivity in Data Analysis" [31]. This technique serves to quantify and anonymize personal data

within a network [32]. The efficacy of Differential Privacy relies on the parameter epsilon (ϵ-value), a determinant that gauges the compromise between privacy loss and the introduction or removal of noise from a specific data account. Balancing the addition or removal of noise within a dataset inherently impacts the utility of the original data [33]. Consequently, a range of ϵ-values, as explored in [34][35], has been tested to ascertain the suitable noise levels for diverse applications. To safeguard real-time data, an optimal amount of noise is strategically added, preserving a judicious balance between privacy and accuracy [36]. DP serves as a protective measure for statistical, database, and real-time data, achieving a reasonable equilibrium between privacy preservation and accuracy [36]. The core objective of Differential Privacy is to obscure the output result of any query, effectively concealing the identity of sensitive information. The mechanism of noise addition in Differential Privacy is visually represented in Figure 2.2, illustrating the interposition of Differential Privacy between the original data and the query data transmitted to analysts.

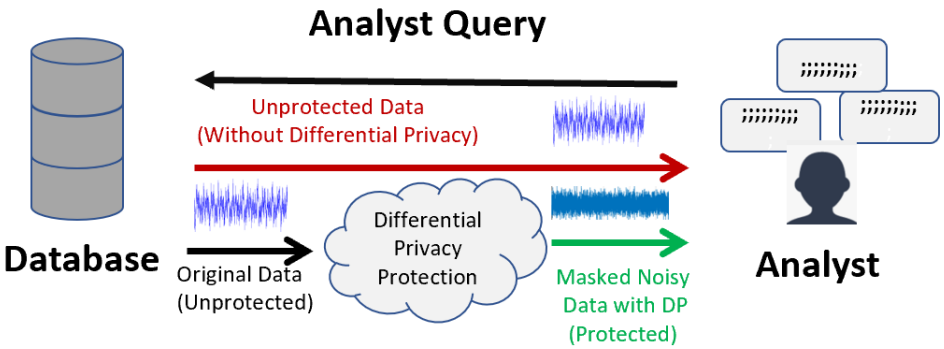

Figure 2.2 Query from E-Health Database (DP implementation)

To illustrate Differential Privacy, for instance, a randomized mechanism *M* gives (ε, δ)-DP for every set of output S, for any dataset that differs in only one value *D*, *D*' (where *D*, and *D'* are two

database neighbors if *D* can be obtained by adding or removing one data from *D'*) if *M* satisfies equation 1 below [37].

$$\left(\frac{\Pr\left[M(D)\in S\right]}{\Pr\left[M(D')\in S\right]}\right) \leq e^{\varepsilon} + \frac{\delta}{\Pr\left[M(D')\in S\right]} \quad (1)$$

For approximate Differential Privacy, if $\delta = 0$, equation 1 shows the ratio between the probability of the query output being dataset *D*, *D*' become $\leq e^{\varepsilon}$. This is called Differential Privacy when two datasets differ with values c, then the Differential Privacy is $\leq e^{\varepsilon c}$ which is called group privacy. The main actors are two mechanisms ε and *M*, where ε is the balance of e-privacy and utility, that is, a trade-off between privacy and accuracy. For the following conditions:

- $\varepsilon = 0$ leads to complete privacy but zero utility, and
- $\varepsilon \leq 1$ leads to less privacy but higher utility.

M determines how much noise will be added and the query output, and utility determines the degree of accuracy, where the ε privacy parameter is called Epsilon.

2.2.1 Mechanisms of Differential Privacy

Differential Privacy existing methods and noise addition mechanisms are the two branches of Differential Privacy [38]. For the sake of this study, the method was narrowed down to noise addition mechanisms (that is, data perturbation mechanisms). These mechanisms are:

- **Laplace Mechanism:** This mechanism is for numeric queries, which is a procedure of adding Laplace noise to query results [39]. The noise is a sample from Laplace distribution. Equation 2 shows the probability density function for Laplace distribution:

$$\text{Lap}(x|b)=\frac{1}{2b}\exp\left(-\frac{|x|}{b}\right) \quad (2)$$

In this equation, b > 0 is the scale parameter of the variable *x*, which is determined by sensitivity Δf of the Laplace function and the privacy parameter ε, b = Δf/ε [40]. The Laplace mechanism uses l_1 sensitivity (that is magnitude by which a single individual's data can change) and the variance of distribution is $\sigma^2 = 2b^2$.

- **Gaussian Mechanism:** In this mechanism, numeric queries are also used to add noise to given data. Rather than scaling to l_1 sensitivity, the curator scales the noise to the l_2 - sensitivity. Equation 3 shows the mechanism of adding Gaussian noise to the results:

$$M(D)=f(D)+N(0,\sigma^2) \quad (3)$$

In this equation, $\sigma = \Delta_2 f \sqrt{(2 \ln(2/\delta)/\varepsilon)}$. Differential Privacy is used for protection of statistical, database, or real time data by adding a desirable amount of noise to maintain a reasonable trade-off between privacy and accuracy [36]. $N(0, \sigma^2)$ is the added Gaussian noise. The Gaussian mechanism is calculated using normal (Gaussian) distribution [41]. The value of ε is between 0 and 1 in a query function.

- **Exponential Mechanism:** Exponential mechanism is used to implement Differential σPrivacy in case of non-numerical output. In this case, query output is measured using a score function [37]. The score function $q(D, \varphi)$ represents the query output φ and how good the output is for the database *D*. As shown below in equation 4, it represents exponential mechanism:

$$M(D) = \{\text{return } \phi \text{ with the probability } \} \ \alpha \exp\left(\frac{\varepsilon q(D,\phi)}{2\Delta q}\right) \quad (4)$$

In this equation, Δ_q represents the sensitivity of score function q and value of Δ_q varies according to the requirement of the user.

CHAPTER 3

LITERATURE REVIEW

3.1 Related Work: Overview of Privacy and Security in Blockchain

Emerging challenges concerning Electronic Medical Records (EMR) include the potential for misuse by third-party entities providing cloud services for EMR storage and sharing [81]. Health Information Technology (HIT) encompasses all systems utilized for storing, accessing, sharing, and transmitting sensitive information, including diagnoses, treatments, and patient tests [82]. Given the dynamic nature of HIT, ensuring the security of clinical data becomes a paramount concern. Privacy, security, and confidentiality stand out as prevalent issues associated with EMR

While privacy and security are closely interconnected, they represent distinct concepts in the context of Electronic Medical Records (EMR). Privacy refers to an individual's entitlement to determine who, how, and to what extent personal information is shared. On the other hand, security involves restricting access to personal information solely based on authorized authorization. Privacy breaches can occur in various scenarios, including systemic identification within the entire electronic Health Information Technology (HIT). Even with legitimate access, EMRs are susceptible to accidental or intentional abuse [83]. Security in HIT necessitates ensuring confidentiality, integrity, and availability. Confidentiality involves restricting information access exclusively to authorized parties. The subsequent sections elaborate on these aspects

Numerous studies indicate that 66% of clients prioritize the privacy of their personal health records, with 39% expressing confidence in the safety and security of their clinical data [84]. Further research demonstrates that half of the respondents harbor concerns about the necessity for their personal health data to traverse the internet [85]. Healthcare organizations commonly grapple with challenges related to handling vast amounts of data, commonly referred to as big data. These challenges encompass issues of data governance, the application of big data technology, and concerns regarding security and privacy [86].

Health Information Technology (HIT) grapples with several challenges posed by the distinct characteristics of Internet of Things (IoT) networks, encompassing heterogeneity, an uncontrolled environment, constrained resources, and the imperative for scalability. The security requirements specific to IoT systems, with their unique attributes, are categorized into key settings: (a) identity management, (b) network security, (c) resilience and trust, and (d) privacy [87]. The advent of IoT has rendered the HIT model acceptable, grounded in the abbreviated CIA principles: confidentiality (to prevent unauthorized access), integrity (ensuring no data alterations), and availability (ensuring accessibility when needed).

3.1.1 Blockchain Issues and Considerations in EMR Systems

This section highlights literature that researches Blockchain application as it relates to EMR and encompasses a deep exploration and comprehensive approach from different perspectives and factors.

- Clohessy et al. highlighted Blockchain applications from different technological, organizational, and environmental perspectives [43]. According to the research, the following organizational factors are ranked in order of their importance.

- Management factors as it relates to support, such as acquiring all the necessary resources ranging from new skills and equipment to integrating regulatory guidelines.
- Readiness is the second organizational factor. This concerns the availability of resources to enhance EMR systems. The innovation in these areas includes human resources, financial, and infrastructure facets in order to ensure cooperation and acceptance.
- Size is another organizational predictor of Blockchain application. Large organizations are more likely to adopt Blockchain. The research also shows a small or medium organization's likelihood of adopting Blockchain applications.

- McGhin et al. identify different and unique requirements that impact EMR systems as it relates to Blockchain application [44].
- Zyskind et al. highlight Blockchain applications for access control, management, and secure data storage [45].
- Asaph Azariaet et al. show Blockchain application based on a data-sharing system that decentralizes record management that handles EMR systems [51].
- Schatshy et al.'s report, as part of Deloitte Insight, highlights five barriers that need to be addressed by companies before pursuing Blockchain application [52].
- Batubara et al.'s literature explores the e-government application of Blockchain [29].

- The Deloitte survey presents responses from more than 1,000 senior executives from seven countries [26][52]. The respondents cited barriers such as regulatory issues (39% of respondents), implementation challenges (37%), security threats (35%), the uncertainty of Return on Investment (ROI) (33%), and lack of a skillful workforce (28%).

- IBM conducted a survey and published a report showing technical challenges restricting Blockchain applications. The major challenge is scalability [28].

- IBM Institute for Business surveyed executives from 200 healthcare entities in 16 nations. The surveys show that more than half of the executives cited these significant barriers to Blockchain application in EMR systems: early state of Blockchain, lack of skillful workforce, and regulatory constraints [48].

- Deloitte conducted a survey addressing Blockchain Technology challenges in life science and EMR systems [49]. Stakeholders engaged in multiple efforts such as healthcare organizations, health plans, scalability standardization, cost, and regulations to ensure commitment to Blockchain applications.

The major limitation of Blockchain is the difficulty in maintaining privacy and security [42]. Users with false identities can breach the security and privacy of EMR systems. Therefore, ensuring anonymity is one of Blockchain's ultimate challenges. Table 3.1 summarizes the literature on Blockchain challenges and considerations in EMR systems.

Table 3.1 Blockchain Issues and Consideration in EMR Systems

Literature	Challenges/Considerations	References
Blockchain Application: Technological, organizational, and environmental considerations	The top factors are management support, organizational, readiness, and organizational size.	[43]
Blockchain Application in EMR Systems	Requirements that impact EMR Systems as it relates to Blockchain application, such as non-standardized system, decentralized storage and privacy, key management and scalability, and IoT.	[44]
Blockchain application for access control management; secure data storage	The encrypted information is stored in a third party that hub services on the Blockchain.	[45]
A Blockchain that is based on data sharing system	Miners are provided with access to aggregate and reward the data bookkeeper.	[50]
Deloitte survey: 1,000 senior executives from seven countries	The respondents cited barriers such as regulatory issues (39% of respondents), implementation challenges (37%), security threats (35%), uncertain Return on Investment (ROI) (33%), and lack of skillful workforce (28%).	[26][52]
IBM report: Technical challenges that restrict Blockchain application	The major challenge is scalability. Blockchain ecosystems within corporate legacy and systems of record are challenging operations.	[47]
IBM Institute for Business Value survey: Executives from 200 healthcare entities in 16 nations	Studies show that more than half cited the early/immature state of Blockchain as an issue.	[48]
Deloitte Blockchain Technology challenges in life science and EMR Systems	Stakeholders engaging in multiple efforts such as healthcare organizations, health plans, scalability standardization, cost, and regulations to ensure commitment to Blockchain adoption.	[49]

3.1.2 Technical Challenges in the Application of Differential Privacy in EMR

Differential Privacy, a concept geared towards preserving privacy, seeks to offer strong protection for individual data points within a dataset, all while enabling the extraction of valuable insights. Nevertheless, the implementation of Differential Privacy comes with its share of technical challenges in a specific application as stated below.

- **Granularity of Data:** Due to highly sensitive information from individual patient records, DP should be applied at fine-grained results in insignificant noise addition, thereby reducing data utility. To address this challenge, data aggregation at higher levels such as hospitals or regions help to balance privacy risk and data utility [120].

- **Sensitivity (Data Utility):** The absence or presence of individual records in the dataset is indistinguishable and maintained. Introducing Differential Privacy in practical datasets requires statistical query and low-sensitivity evaluation because random noise can be excessive, thereby impacting the usefulness of the patient data for analysis and decision-making [53]. A trade-off exists between accuracy (utility) and privacy, which is a challenge that emerges in services and applications using different sensitivities [54][55]. To address this challenge, adaptive privacy budgets, query optimization, and advanced noise generation algorithms can be used to improve data while preserving privacy.

- **Data Heterogeneity and Interoperability:** EMR data have various sources and different formats, models, and standards which complicates the assurance of privacy. To address interoperability concerns, standardization efforts such as adopting a common data model, which facilitates integration while ensuring consistent privacy-preserving mechanisms across data sources [122].

- **Scalability:** It is challenging to manage a large volume of data from different sources in an EMR system when applying Differential Privacy techniques. To address scalability issues, distributed processing frameworks such as MapReduce or Spark can be utilized. These frameworks enhance parallelization techniques and efficient processing of data while preserving privacy [104].

- **Choosing Epsilon Value (ϵ-Privacy Loss):** Choosing the privacy parameter ε is a practice that users of Differential Privacy cannot avoid [3]. The strength of guaranteed privacy is controlled by ε, and it is not clear how to choose an appropriate value in a given situation, as shown in [56][57]. In [58], the smaller ε is, the higher the increase in security and vice versa.

- **Data Correlation:** In a real-world dataset, there is a correlation in certain records that leads to the disclosure of information. Many researchers have developed model-based and transformation-based approaches, such as sensitivity weights, correlation degree, and correlated sensitivity, that have overcome these challenges [59].

- **Other Challenges:** Other challenges include a lack of computing environment, an appropriate system that aligns with users' needs, and a lack of trained personnel to verify implementation and correctness [60]. Consent and governance are other challenges. EMR data is highly sensitive and subject to legal and ethical considerations. Robust governance frameworks, informed consent mechanisms, and data-sharing agreements are essential to address these challenges and maintain patient trust [121].

3.1.3 Integration of Differential Privacy and Blockchain

Blockchain and Differential Privacy are revolutionizing and altering the concept of data storage. The decentralized property of Blockchain is considered a secure system. However, there are issues in Blockchain that require solutions before implementation in a real-world situation. One of these issues is preserving data while maintaining privacy for Blockchain applications. The integration of Differential Privacy in each layer of Blockchain Technology is classified into six different layers according to [61]. These layers are: (a) the data layer, (b) the network layer, (c) the consensus layer, (d) the incentive layer, (e) the contract layer, and (f) the application layer. Each layer has functionality and privacy requirements. For instance, users' requirements differ from privacy requirements while creating blocks in the data or the consensus layer.

Researchers are actively investigating the effort to integrate Differential Privacy with a Blockchain-based healthcare system. In [62], the author proposed a proof of votes consensus that operates on a Blockchain-based healthcare network whereby data is mutually shared to create transaction blocks. As such, a third-party team is assigned to work and forward the blocks to companies within the network for verification through voting, thereby ensuring the decentralized characteristics of the Blockchain. Furthermore, the author discussed adding noise to their data to ensure privacy using decentralized Differential Privacy protection.

Establishing remote connections is pivotal for both doctors and patients, facilitating routine monitoring and fitness programs, especially in the context of elderly care [63]. As the demands of the modern healthcare system evolve, traditional service administration methods prove inadequate due to the requisite transparency and trust. Traditional healthcare systems are vulnerable to attacks and data tampering by adversaries. Hence, the integration of Blockchain becomes imperative to bolster the security of the contemporary healthcare infrastructure. While this trend presents

significant advantages, privacy concerns persist as Blockchain stores data in a decentralized distributed ledger, where each node holds a copy of the ledger. The decentralized nature introduces a vulnerability wherein a malicious node could potentially initiate an attack on the private information of a Blockchain node.

3.1.4 Integration of Differential Privacy in E-Health Domains

The implementation of Differential Privacy in e-Health domains is subdivided into real-time health data, electronic medical records (EMR) privacy, and health survey data protection.

- **Real-Time Health Data:** Real-time data is reported to the database or medical personnel such as doctors and lab technicians. The purpose is to keep track of users' activities, but disclosing this data can lead to privacy concerns [3]. Most real-time health data is derived from wearable devices, also known as IoTs. In [3], the scheme Re-DPoctor was proposed to provide budget allocation and adaptive sampling using Differential Privacy. The proposed strategy meets all conditions of Differential Privacy and reduces the mean relative error and mean absolute error of the transmitted. The strategy also uses proportional-integral-plus (PIP) to control and compare the trade-off between privacy and accuracy (utility) by applying Differential Privacy. The mathematical models of Differential Privacy are used for data protection by adding the value of noise in order to include wearable devices. Therefore, data perturbation using Differential Privacy is suitable for data preservation.

- **Electronic Medical Records (EMR) Privacy:** Health systems have adopted an electronic way of storing, sharing, and transmitting patient data and integrated mechanisms [64][65]. Differential Privacy is also applied over an end-to-end based deep learning approach to

enhance training accuracy and efficiency by integrating cryptographic encryption while preserving privacy with machine learning [66][67]. The preservation of privacy mechanism is accomplished by adding Laplace noise over perturbated data for optimization of error rate for statistical queries [90]. In [91], the Laplace noise addition mechanism of Differential Privacy provides an enhancement to data privacy by performing an experiment using cancer patients' data. The strategy's aim is to reduce computation using a framework that is compatible with data mining tasks and different SQL queries. Other fields, such as genomic data record protection and distributed clinical data by encrypting data, are discussed in [91] using Differential Privacy noise mechanism. The mathematical models of Differential Privacy best fit the e-Health domain database, and the data can be secured easily using Differential Privacy perturbation.

- **Health Survey Data Protection:** Before publishing data for a survey or statistical query to learn more about a particular disease, the sensitive data needs to be protected so that the adversary cannot compromise its integrity. For instance, mobile recommender systems can be used to suggest medication with fewer side effects [68], or a therapist can use recommender systems to suggest depositions for a patient [69][70]. The benefits come with a trade-off of privacy, especially the perspective of a common person toward Differential Privacy. The users' perspectives toward their confidential data are hindered for commercial purposes, while some users show reluctance for their data to be used for scientific purposes. However, Differential Privacy is the preferred data perturbation mechanism and provides a trade-off for privacy decisions.

The summary of the literature on Differential Privacy in e-Health is shown in Table 3.2 below. Table 3.3 presents a review of the pros and cons of approaches implemented in e-Health systems regarding real-time health data, electronic medical records (EMR), and survey data recorded.

Table 3.2 Summary of Related Work of Differential Privacy in E-Health Domains

Main Category	**Privacy Mechanism**	**Technique of DP Used**	**Enhancement due to DP**	**Privacy Criterion**	**Scenario**	**Ref**
Real-time Health Data	Real-time health data releasing scheme Re-Dpoctor	Data perturbation is used along with adaptive sampling and filtering	Mean absolute error and mean relative errors are enhanced	(ε, δ)-Differential Privacy	Real-time	[3]
Electronic Medical Record (EMR) Privacy	Efficient E-health data release	Heuristic hierarchical query method and private partition algorithm proposed for DP	Enhanced time, overhead, and query error	(ε, Δ)-Differential Privacy	Statistical Database	[117]
	Private and secure management of databases of health care database	Used Laplace mechanism for data privacy	Reduced computational overhead	(ε, Δ)-Differential Privacy	Statistical Database	[99]
	Health data Differential Privacy algorithm for range queries	Partitioning by data and workload are implemented with use of Laplacian noise	Optimized error rate of queries	ε-Differential Privacy	Statistical Database	[90]
	MedCo (Privacy preservation of genomic and	Encryption in combination with Differential Privacy is	• Enhanced i2b2 database privacy • Optimized runtime and	ε-Differential Privacy	Statistical Database	[91]

	distributed clinical data)	used to secure and preserve sensitive data	network overhead			
	Genomic data privacy protection	Protecting encrypted data using Differential Privacy and two step decryption	• Enhanced execution time • preserved secret keys leakage in dual decryption	ε Differential Privacy	Statistical Database	[101]
	End-to-end differentially private deep learning health record protection	Differentially private stochastic gradient descent based deep learning method	• Enhanced training accuracy • Improved computational cost	(ε, δ)-Differential Privacy	Statistical Database	[90]
	Differentially private data clustering (EDPDCS) framework for medical data	K-means clustering based differentially private machine learning over MapReduce	• Optimized privacy allocation budget • Improved learning accuracy	(ε, δ)-Differential Privacy	Statistical Database	[66]
	Secure e-Health data aggregation with fair incentives	Combined local Differential Privacy with Boneh-GohNissim crypto system and Shamir's secret sharing	• Improved key generation overhead • Aggregation privacy	ε - Differential Privacy	Real-time	[100]
Health Survey Data Protection	Privacy-Utility trade-off in health record systems	K-Anonymity and random data perturbation discussed	• Discussed and improved survey data according to users' perspectives		Statistical Database	[107]

Table 3.3 A Review of Pros and Cons of Each Approach

Approaches	Pros	Cons	Reference
Electronic Medical Records (EMR) Privacy	Overall improvement in the accuracy of perturbated data and eliminate background knowledge of attack	Risk of privacy attack is a major concern	[98][118]
Real-time Health Data	Improve the utility with excellent performance under small privacy cost	Real time data creates temporal correlation problem	[116][119]
Health Survey Data Protection	Reduces communication overhead as well as cloud burden	-----	[112]

3.1.5 Key Technical Issues with Integration of DP and BC in E-Health Domains

Generally, there are technical issues encountered during the implementation of Differential Privacy. These issues are listed below.

- **Decision of ϵ-Value (Privacy Loss):** choosing small ϵ versus large ϵ can be challenging [60].
- **Decision of Sensitivity Value:** There is a lack of guidelines for choosing the optimal value of sensitivity to balance a reasonable trade-off between sensitivity and data utility; as such, researchers tend to use low sensitivity value on a statistical database [53][60].
- **Overcoming Data Coupling:** Data correlation is one of the challenging issues during Differential Privacy implementation [4]. In a real-world operation, the dataset correlates

with different attributes, allowing attackers to make references to obtain personal information [71].

3.1.6 Other Approach to Enhance Privacy in EMR – Federated Learning (FL)

Federated Learning (FL) is another learning paradigm designed to address the problem of data sharing and privacy [72]. The FL approach was initially developed for different domains, such as mobile and edge devices, but in recent years, FL has gained traction in EMR [73]. In collaboration with the consensus model, FL enables and gains insight into data without sharing patient information beyond the firewalls of the institutions where it resides [74]. In this case, the FL process is positioned locally at each institution, and only the model characteristics, such as parameters and gradient, are transferred [74]. Therefore, in the context of EMR, for instance, FL helps in the following areas.

- Finding patients with similar clinical results [75]
- Prediction of hospitalization due to cardiac [76]
- Medical imaging for whole brain segmentation in MRI [77]

The advantages of FL only solve some inherent challenges in EMR. Some factors, such as data quality, bias, and standardizations, depend on the successful model training [78]. Data heterogeneity is challenging in FL since collaborative learning strategies are not uniformly distributed across the institution [79]. Other considerations are privacy and security, the trade-off strategies, and risk regarding the privacy-preserving potential of FL performance and techniques [73]. Differential Privacy can also enhance privacy in an FL setting [80]. Developing

countermeasures, such as limiting the granularity of the updates and adding appropriate noise, may be needed [59]. In effect, the discussion regarding FL is still open for further research.

3.2 Research Gap Analysis

The Research Questions (RQs) from Table 1.1 are used as the basis for the research gap analysis. From the previously reviewed literature regarding Blockchain and EMR systems, the following sections highlight what has been identified.

3.2.1 Gap 1: Lack of Assessment from Multiple Perspectives

The application of Blockchain needs a multi-criteria model to evaluate its capabilities to enhance the security and privacy of EMR systems. The in-depth studies of EMR systems have shown that the issues of security and privacy lack full integration of all aspects of PHI, which includes the readiness of healthcare management to deploy or adopt Blockchain for the interoperability of EMR. The factors that impact the consideration of Blockchain adoption stem from multiple perspectives, such as technological, organizational, and environmental. Therefore, there is a need to explore various influential factors in order to assess and then propose a platform that ensures the security and privacy of EMR systems. The study of the assessment of Blockchain application from different perspectives is deployed for clarification, understanding of the healthcare industry, and factors that would influence and enhance the application of Blockchain for EMR systems.

3.2.2 Gap 2: Lack of a Comprehensive Chronological Model: Lack of Approach

There is no chronological order of the model that highlights the application of Blockchain in EMR systems. Most of the literature centers around other fields, such as finance and cryptocurrency, and there is a well-proposed technical platform for implementation that shows the capability of Blockchain application. The studies indicated that Blockchain application maintains various

benefits, but the salient issues that require resolution include scalability, interoperability, security, and privacy. Therefore, the need for a holistic platform and model to enhance healthcare organizations' management of EMR systems is geared toward Blockchain application.

3.2.3 Gap 3: Highlight Inherent Issues

There is a lack of studies designed to enlighten inherent issues in Blockchain applications for the management of EMR systems. The current healthcare system is rather dynamic regarding sharing, storing, and exchanging PHI. Due to the increase in the usage of IoTs, EMR systems possess inherent limitations that technology cannot solve. These limitations revolve around ethics and moral obligations, which are hard to detect. Most of the literature lack this aspect to investigate, identify, and evaluate the inability of Blockchain technology to solve these inherent issues.

3.2.4 Gap 4: Lack of Experts' Assessments and Quantifications

Current studies are based on the characteristics of Blockchain applications related to EMR systems, and there are no collaborations from different experts' assessments. The essence of these collaborations is to quantify the important factors that utilize different sectors' perspectives, which could include government officials, healthcare administrators, Blockchain experts, cybersecurity specialists, and legal experts with different experiences in implementing and applying Blockchain applications to manage EMR systems.

3.2.5 Gap 5: Lack of Legal Framework for EMR System

Most literature focused on Blockchain applications without a legal framework defining their operations. A framework that combines technicality and legality to improve the management of the EMR system is needed. The current guideline on PHI is based on HIPAA, which needs improvement in the existing regulatory matrix to meet these issues. The foundation provided by

this legal framework is the future of health information with an understanding of its design, deployment, and Blockchain application.

3.2.6 Gap 6: Leveraging Differential Privacy for Privacy Protection

Most of the user's information is collected through IoT, which contains sensitive data such as lab results about a diagnosis. Differential mechanisms are used to protect data to avoid privacy leakage. In [39], the paper proposes the use of the Laplace mechanism, which is a procedure of adding Laplace noise to query results. The exponential mechanism is used to implement Differential Privacy in case of non-numerical output. In this case, query output is measured using a score function [53] to leverage Differential Privacy in order to develop a framework that will prevent an attack. However, the authors fail to discuss how to reuse the Differential Privacy budget.

3.2.7 Gap 7: Fundamental and Applied Research Approaches of Differential Privacy

Regarding the most viable protection technique for real-world data analysis, there is a discrepancy between these two approaches. The literature regarding Differential Privacy is less intuitive as it relates to a concept that does not match the industry. Therefore, there is no connection between the academic Differential Privacy platform and the practical application of Differential Privacy [88]. That is, there is a lack of knowledge as to what software tools can be leveraged in application and development with Differential Privacy principles. This also includes the degree of performance researchers and developers expect from the current platforms [89].

3.3 Research Output

The output from gap analysis measures the current state versus the desired state of the reviewed literature. It identifies missing collaborations, technologies, and processes. The research output is summarized below.

- Identification of the inherent limitation in EMR systems as it relates to privacy and security
- Identification of challenges in Blockchain applications concerning scalability and interoperability
- Highlighting the Blockchain application from different perspectives, such as technological, organizational, and environmental factors
- Highlighting Blockchain application for access control management and secure data storage in a third party that hub services on the Blockchain
- Examination regarding limited credibility that a third party can reduce data disclosure or a breach in the EMR system
- Identification of significant barriers to Blockchain application in EMR Systems, such as the early state of Blockchain, lack of skillful workforce, and regulatory constraints
- Identification of the existing technical challenges for leveraging Differential Privacy to secure sensitive data
- Highlighting the lack of literature revealing that Differential Privacy needs to be more intuitive as it relates to a concept that does not match the healthcare industry

Figure 3.1 summarizes the research gaps, goal, and research output.

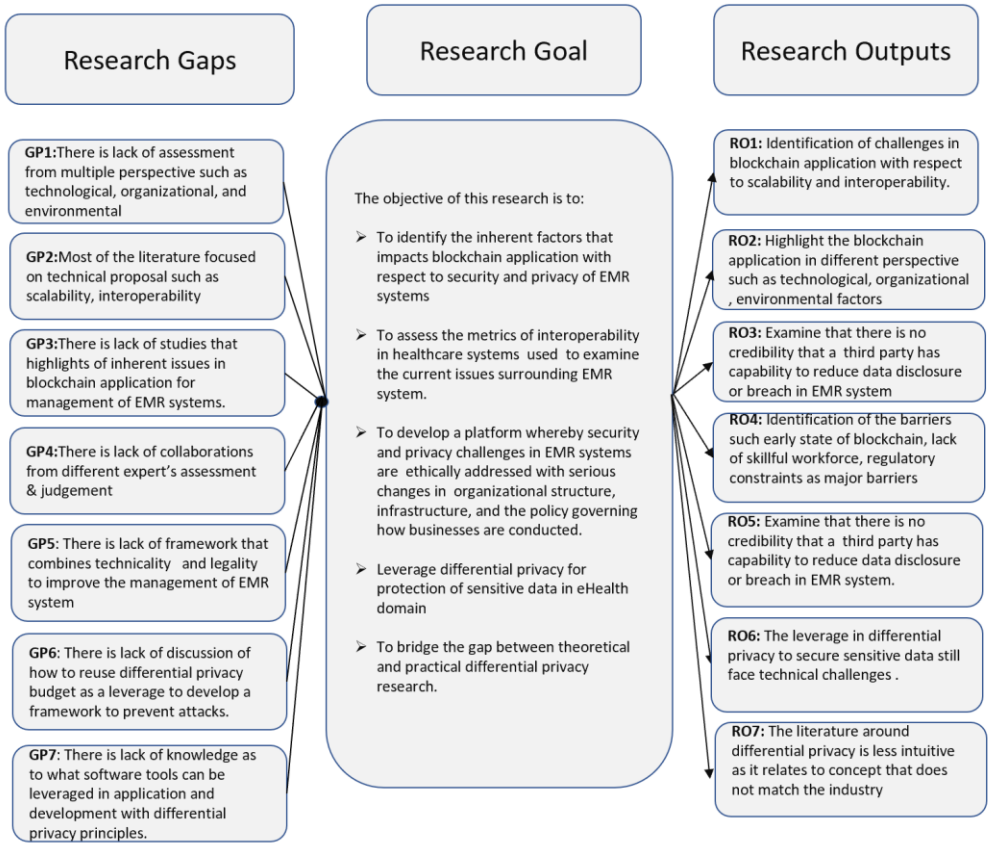

Figure 3.1 Research Gaps, Goals, and Output

CHAPTER 4

METHODOLOGY

The research methodology employed in this study is a systematic mapping study (SMS), a widely utilized approach in scientific surveys. SMS involves various research steps and the systematic selection of papers and publications to address formulated research questions (RQs) [92]. The primary objective of utilizing SMS in this research was to achieve a comprehensive overview of relevant research papers through an unbiased assessment, aiming to identify research gaps and collect evidence for future proposals [93]. The research method followed the guidelines proposed by [94], encompassing the following steps: (a) defining research goals; (b) formulating research questions (RQs); (c) outlining the research strategy, which includes search terms, literature sources, search processes, and study selection; (d) assessing the quality of selected studies; and (e) analyzing the results. The framework directions and phases of this research are illustrated in Figure 4.1 shows the framework directions and phases of this research.

4.1 Research Goal

Electronic Medical Records (EMR) contain patients' medical history. The issues concerning privacy and security have exponentially widened because of the era of IoT. In EMR systems, the management is ineffective without a proper system to share, store, and transmit these records in a server in a secure manner. The goals of this research are listed below.

Figure 4.1 Research Design

- To identify the inherent factors that impact Blockchain applications concerning the security and privacy of EMR systems and to investigate the supporting platform that permits integration of Differential Privacy as a covering layer

- To categorize this investigation into three areas that address (a) Real-Time Health Data, (b) Electronic Medical Records (EMR) Privacy, and (c) Health Survey Data Protection

This research aimed to facilitate the trade-off between security and privacy during the application of Blockchain in the management of EMR systems and to formulate a proposal for future research in an area that needs more attention where inherent security and privacy challenges exist in EMR systems.

4.2 Research Questions (RQs)

The Research Questions (RQs) were formulated based on the research motivations, problem statements, and the goal of this review. Table 4.1 below summarizes the research questions (RQs).

Table 4.1 Research Questions (RQs)

ID	Research Questions
RQ1	How can DP be integrated into BC to enhance privacy and security in the e-Health domain (e.g., EMR)?
RQ2	What factors contribute to the DP mechanisms integration in Blockchain Technology and associated issues?
RQ3	What types of datasets and programming languages are being considered for implementation?
RQ4	What are the limitations and inherent challenges of the BT and DP applications, and how can they be solved?
b.	**Note that the above questions are narrowed to only e-Health domains**

4.3 Research Strategy

The sources of information in the literature are academic publications, including conference papers, journal articles, Google scholarly books, and reports. Sources also include government agencies and reputable organizations such as IBM and AMIA. The research strategy intended to identify relevant works and applications of Blockchain and Differential Privacy and mechanisms in the e-Health domain, including the cost of privacy and challenges of the proposed solutions. As recommended by [95], two research strategies were primary and secondary. The primary strategy includes search terms, literature resources, and the search process, as explained below.

4.3.1 Search Terms

The search keywords used in this research are shown in Table 4.2. Online libraries, various journals, and papers were considered during the keyword search. The date filter was used to screen for current literature.

Table 4.2 Search Terms and Keywords

Numbers	Keywords
1	Review, survey, literature review, background
2	Electronic medical records, e-Health domain*, electronic health record, health information technology, patient health information
3	Blockchain Technology, Differential Privacy*, privacy, data
4	Data perturbation, Differential Privacy mechanisms

a. ***the keyword noted while searching**

4.3.2 Literature Sources

The search was conducted for papers on four different electronic databases from online libraries. During the collection process, the title, the year of publication, the journal name, the number of citations, and the link were considered. The search terms with keywords for collecting conference papers and reviewing academic journals were used to formulate conceptual building blocks. The search also covered keywords in the title and abstract. The summary of the collected search is shown in Table 4.3.

Table 4.3 Numbers of Literature Retrieved from Online Libraries

Online Libraries	Numbers of Retrieved Literature
IEEE	32
ACM	8
ScienceDirect	6
AMIA	1
Others	58
Total	**105**

4.3.3 Search Process

A Systematic Mapping Study (SMS) examined the resources' maturity and comprehensibility during the search. The systematic review process can be divided into two main phases.

- **Phase One:** Initial searching phase consisted of the four online library databases. Each paper was searched separately with keywords, as shown in Table 4.3.

- **Phase Two:** In this phase, the search was conducted based on the references of a particular paper. By scanning the list of references for relevant papers, they were added if there was a relation to the keywords.

The search results were stored and managed in Microsoft Excel. From the phase-one search, 300 papers were gathered. Eight-five papers were gathered from phase two of the reference search, as shown in Figure 4.2.

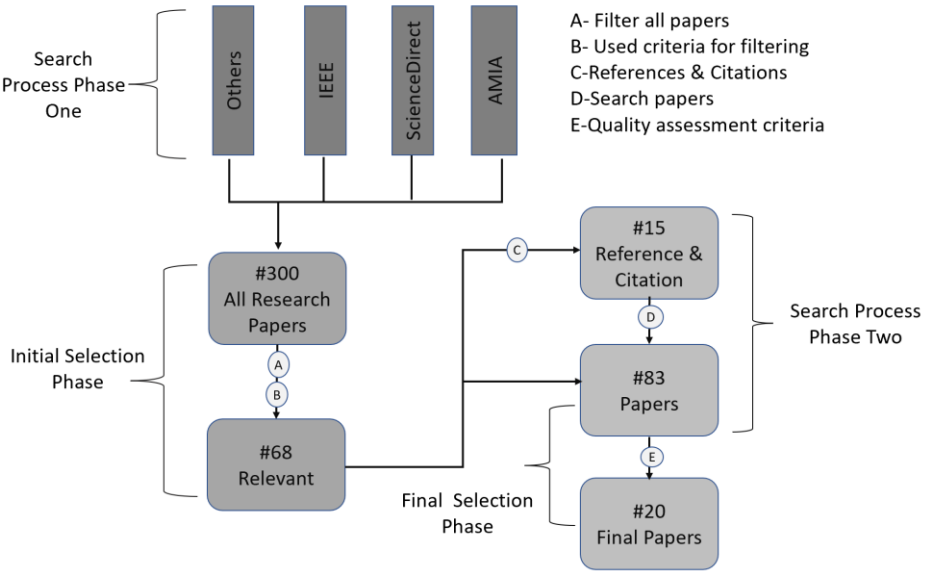

Figure 4.2 Search and Selection Process

4.3.4 Study Selection

Research papers were selected from different websites. Some of the papers did not offer helpful inside knowledge concerning this research, so an extensive filtering process was performed. The selection process consisted of two phases.

- **Initial Selection Phase:** The aim was to obtain papers that offer sufficient background regarding this research. The researcher applied inclusion criteria (IC) and exclusion criteria (EC) to filter any related papers that answer the research questions. IC and EC are defined below.
    - The inclusion criteria (IC) are listed below.
        - Papers published from 2008 (only a few papers published in 2005 and 2006)
        - Papers published until 2022
        - Papers that describe Blockchain and Differential Privacy
        - Papers that describe EMR, e-Health domain
        - Academic papers and journals
        - Review or survey papers
        - Check for duplicate publications – completed or newly released of the same study
    - The exclusion criteria (EC) are presented below.
        - Papers in digital libraries that are duplicated

- News, correspondences, comments, summaries of presentations, posters, and workshops
- Abstracts of papers that are not written in the English language

- **Final Selection Phase:** This phase selected papers with the acceptable quality needed to extract information. The selection in the final phase used study quality assessment, as explained in section 4.4 below.

The citations and references from the above papers were also reviewed, and the last step included quality assessment criteria for data extraction.

4.4 Study Quality Assessment

This section addresses how quality assessment questions (QAQs) give credit to the reviewed paper. These questions are shown in Table 4.4. The questions were used for the quality assessment of the paper and the criteria. QAQ1 evaluated how the e-Health domain uses Blockchain and Differential Privacy to enhance protection for sensitive health information. Noticeably, the researchers used the DP method to address the security concerns in the e-Health domain. QAQ2 attempted to discover if the papers provide a more comprehensive perspective other than EMR systems since the privacy of personal information is cut across all fields. QAQ3 explored whether the research results can be deployed to real-world applications. QAQ4 evaluated common limitations in the papers that are inherent. QAQ5 identified similarities in research questions, while QAQ6 defined different methods to provide solutions. Finally, 20 papers were selected, as shown in Table 4.5.

Table 4.4 Quality Assessment Questions

ID	Quality Assessment Questions
QAQ1	Are the review papers related to e-Health domain under Blockchain and Differential Privacy?
QAQ2	Do the papers cover other Differential Privacy applications under different fields?
QAQ3	Do the papers use theoretical or practical based methods to answer research questions?
QAQ4	Are there common or inherent limitations in their studies?
QAQ5	Is the research question similar or different from other papers?
QAQ6	Do the proposed methods provide solutions that are different from the existing papers?

Table 4.5 List of Papers for Methodology

Category	Papers Selection*
EMR Privacy	Roehrs et al. [22], ElSalamouny et al. [54], Saleheen et al. [96], Raisaro et al. [91], Lin et al. [97], Guan et al. [66], Machanavajjhala et al. [55], Alnemari et al. [90], Hadian et al. [98], Mohammed et al. [99], Tang et al. [100], Raisaro et al. [101]
Real-Time Health Data	Geo et al. [102], McSherry et al. [103], Machanavajjhala et al. [56], Zhang et al. [3]
Health Survey Data Protection	Luo et al. [104], Narayanan et al. [105], Valdezet al. [107], Narayanan et al. [106]
***This selection is for the research framework and methodology**	

CHAPTER 5

RESULTS ANALYSIS

The analysis of research results is based on the research questions (RQs) in chapter 1. After an extensive literature review and rigorous investigation into different papers, the Differential Privacy mechanisms used to enhance Blockchain Technology in e-Health domains have been organized into three main categories namely Real-Time Health Data, EMRs Privacy, and Health Survey Data Protection.

Figure 5.1 portrays the taxonomy diagram for Differential Privacy in e-Health domain, as well as health systems and approaches implemented in e-Health systems. The figure shows each category: real-time health data, electronic medical records (EMR) privacy, and health survey data protection.

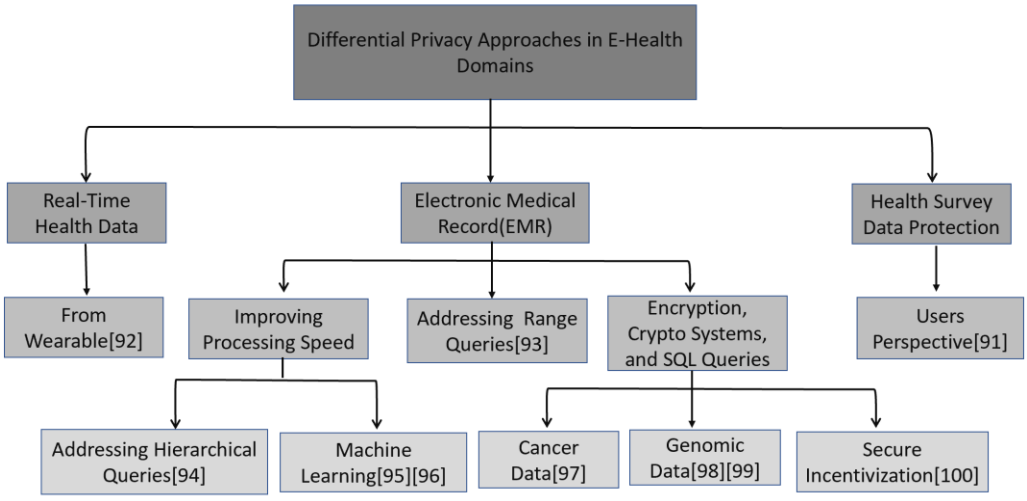

Figure 5.1 The Taxonomy for Differential Privacy in Approaches in E-Health Domains

5.1 RQ1: How can DP be Integrated into BC to Enhance Privacy and Security in the E-Health Domain (e.g., EMR)?

Integrating Differential Privacy in decentralized healthcare is considered part of modern smart cities. Every patient, doctor, and hospital are connected to provide services such as remote health monitoring, fitness programmers, and elderly care [62]. The integration trend has potential benefits; however, it raises privacy concerns as data over the Blockchain is stored in a decentralized ledger. Therefore, the authors in [62] proposed a healthcare system whereby a secure Blockchain-based system is used as a proof of vote (PoV) consensus mechanism. The use of real-time health data is considered as a solution to RQ1. Real-time health data used in e-Health domains mostly comes from IoT devices, which is different from conventional health data [108]. The mechanisms are also called data perturbation, including the Laplace, Exponential, and Gaussian mechanisms. This research question explores which mechanisms researchers have used to protect the privacy of sensitive health data in real-time.

5.2 RQ2: What Factors Contribute to the DP Mechanisms Integration in Blockchain Technology and Associated Issues?

This research question explores the factors contributing to Blockchain Technology integration with DP based on reliability, utility-privacy trade-off, and risk minimization. The data over the Blockchain is stored in a decentralized distributed Hyperledger. Furthermore, the node contains a copy of that ledger [39]. The researchers suggest privacy preservation strategies based on e-Health systems, and one of these strategies is called Differential Privacy in decentralized healthcare [62]. A report of diagnosis of disease falls under this EMR, and the technical work is kept secure by using centralized Differential Privacy and pseudo-identity mechanisms [109]. The researchers introduced a risk minimization strategy using test errors to overcome adversaries in a public Blockchain environment. The associated issues concern the navigation between utility (accuracy)

and privacy, the trade-off. For example, adding noise to the data may reduce the accuracy of the information in the e-Health domain [54]. Furthermore, this may put the safety and welfare of the patient at risk. Therefore, an adequate trade-off between privacy and utility (accuracy) must be maintained. The proposal in [110] involves a Differential Privacy-based solution and optimization of privacy parameters to obtain a helpful utility (accuracy) and privacy trade-off.

5.3 RQ3: What Types of Datasets and Programming Languages are being Considered for Implementation?

The implementation of any proposed solution depends on the quality of the datasets. As reviewed in section 2.2.1, data perturbation mechanisms for Laplace and Gaussian use numerical datasets, while Exponential uses a non-numerical dataset. The dataset used in data perturbation mechanisms is a structured dataset. However, further research reveals that most proposal applications grouped datasets into public and private datasets based on availability.

Real-Time Health Data: This category utilizes the flu dataset that Geo et al. [102] harnesses. Wearables are also used to record and share real-time health datasets. In [40], the heart rate dataset was recorded in order to research real-time data. The summary is shown in Table 5.1 below.

EMR Privacy: According to Saleheen et al. [96], EMR privacy shows a dataset with 660 hours of ECG (electrocardiogram) from participants whose private dataset was collected. Lin et al. [97] collected private datasets from wearable sensors, [79] collected heart disease datasets, and Hadian et al. [98] collected datasets from wearable devices that users attached to their bodies to monitor heart rate. A blood bank dataset containing individual information utilized a research record dataset [66]. In addition, datasets were obtained during activities such as walking, running, and

sleeping. Kim et al. [111] obtained a dataset from daily step counts using a Gear S3 smartwatch. Table 5.1 below shows a summary of dataset utilization in the e-Health domain.

Health Survey Data Protection: This category discusses and provides inside surveys according to users' perspectives. Most of these datasets from a database are statistically queried. Luo et al. [104] surveyed two real-world case studies. One of the cases used a health survey based on students' heart rates to find the average and distribution statistically. The second case focused on collaboration to classify models based on emotions. Yang et al. [104] also used real-world public datasets with one million health datasets. The summary is shown in Table 5.1.

Table 5.1 Different Types of Datasets

Data Type	**EMR Privacy**	**Real-Time Health Data**	**Health Survey Data Protection**
Private (Heart-related)	[97, 98]	[97]	[104]
Public	[66]	_	[112]
Private	[97]	_	_
Public (Activities, e.g., running, walking)	[66]	_	_
Private (Wearable sensors)	[111, 98]	[102]	_

Programming Laanguages: A thorough review of programming Languages used for the implementation of blockchain and differential privacy application. The review shows that Solidity, JavaScript(Node.js), Python, Go(Golang), Java, C++, and Swift are frequently used to implement the algorithms in research work. Among this programming language, python has emerged as the most used programing language[129].

5.4 RQ4: What are the Limitations and Inherent Challenges of the BT and DP Applications, and How can They be Solved?

The limitations of the existing methodology are visible, and researchers have conducted several experiments to evaluate different approaches.

Real-Time Health Data: Proposed solutions for real-time data in Differential Privacy applications suffer data perturbation errors [115] because of relative and absolute errors [116]. The strength of guaranteed privacy is controlled by ε, and it is not clear how to choose an appropriate value in a given situation, as shown in [56][57], where algorithms have chosen ε from the range of 0.01 to 7. For example, in [90], a large budget ($\varepsilon > 1$) shows no corresponding advantages. Similarly, in [102], there is evidence that increasing the epsilon value weakens the algorithm. Therefore, choosing an appropriate epsilon value is challenging for a threshold application.

EMR Privacy: As discussed in [22], Blockchain Technology has scalability issues. Most of the proposed solutions for Differential Privacy are for static database information as it confines to a single dimension [113]. Another issue is that most of the privacy protection approach needs a practical roadmap for implementation, and some models suffer from degradation in performance as the number of cloud resources increases [114]. Zhang et al. [113] proposed a more complex algorithm than existing works. The methods are also vulnerable to information leakage, giving adversaries more knowledge about sensitive data.

Health Survey Data Protection: According to [107], survey participants potentially revealed sensitive information which is just one example of the challenges of complete privacy protection.

CHAPTER 6

DISCUSSIONS, RECOMMENDATIONS, AND CONCLUSION

6.1 Discussion

The e-Health domain requires mechanisms for flexible solutions to preserve the privacy of health information. As IoT devices and their application have grown exponentially, Blockchain and Differential Privacy have emerged as viable routes to enhance data security. Researchers have shown inherent limitations on Blockchain and Differential Privacy and the concern regarding these challenges [24, 67, 91]. In this section, the discussions are grouped into the three categories of real-time health data, electronic medical records (EMR) privacy, and health survey data protection.

- **Real-Time Health Data:** This represents papers that were investigated based on the real-time health data releasing scheme. Most of this data comes from IoT devices, such as wearables, for real-time data collection and sharing. Therefore, all papers that discuss Differential Privacy and Blockchain were placed under this category.

- **Electronic Medical Records (EMR) Privacy:** This represents papers that were covered by EMR systems. The EMR consists of all clinical data, laboratory tests, and diagnosis results that come in different numeric and non-numeric queries. These papers discussed how to protect sensitive health data from the database using Differential Privacy mechanisms.

- **Health Survey Data Protection:** This represents papers that were discussed in the statistical database, such as how health survey data improves based on users' perspectives and the Differential Privacy mechanisms used to enhance privacy-utility trade-off in e-Health domains.

6.2 Challenges and Limitations

Firstly, Blockchain presents scalability and interoperability issues that create unreasonable constraints on exchanging patient data [22]. Secondly, Differential Privacy maintains challenges when choosing epsilon (ϵ) values [35, 55, 56]. Sensitivity is another challenge while navigating the trade-off between privacy and accuracy (utility) [53, 54, 55]. Data correlation-dataset used in a real-world situation is strongly correlated, which gives an adversary a chance to combine obfuscated data to obtain sensitive health information [59, 10]. Mechanism implementation of Differential Privacy, such as Laplace noise, is vulnerable to being tracked or attacked [104].

6.3 Recommendations and Future Work

Recommendations for implementing a privacy-preserving Blockchain-based solution are as follows. One approach could be to use homomorphic encryption to encrypt the sensitive data stored on the Blockchain, allowing for computations to be performed on the encrypted data without exposing it. The data could then be decrypted only by authorized parties.

Additionally, Differential Privacy techniques could be used to add random noise to the data before it is stored on the Blockchain to protect the privacy of individual patients further. By implementing these privacy-enhancing technologies, a secure and private system for EMR storage and management could be established, maintaining the confidentiality of sensitive medical information while allowing for the benefits of a decentralized, tamper-proof system.

6.4 Conclusion

The three categories, real-time health data, EMR privacy, and health survey data protection, are significant concerns in e-Health domains as they relate to privacy. Blockchain Technology and Differential Privacy have emerged as suitable mechanisms. This project aims to understand Blockchain and Differential Privacy in e-Health domains for privacy protection, as well as limitations and future direction to enhance integration and implementation of Blockchain and Differential Privacy in e-Health domains.

The literature review and related works reveals gaps, necessitating the incorporation of additional mechanisms to enhance privacy and security in e-Health domains. These gaps are:

- Lack of Assessment from Multiple Perspectives
- Lack of a Comprehensive Chronological Model: Lack of Approach
- Highlight Inherent Issues
- Lack of Experts' Assessments and Quantifications
- Lack of Legal Framework for EMR System
- Leveraging Differential Privacy for Privacy Protection
- Fundamental and Applied Research Approaches of Differential Privacy

In addition, the trade-off between privacy and utility (accuracy) in Differential Privacy and the integration of Blockchain with Differential Privacy is a complex computational problem. Recently, most companies and establishments have experienced a rapid increase in cybersecurity attacks from adversaries, compromising the privacy of sensitive information. The attackers exploit weaknesses such as correlated data, despite the use of Differential Privacy, to breach the security mechanisms.

This review thoroughly surveyed and summarized Differential Privacy mechanisms in real-time health data, EMR privacy, and health survey data protection while highlighting limitations and challenges, as well as exploring future research areas in Blockchain and Differential Privacy.

REFERENCES

[1] Centers for Disease Control and Prevention. (2018, September 14). Health Insurance Portability and accountability act of 1996 (HIPAA). Centers for Disease Control and Prevention. Retrieved January 31, 2022, from https://www.cdc.gov/phlp/publications/topic/hipaa.html

[2] Compliancy Group. (2022, July 22). The CIA triad: Confidentiality, integrity, availability for HIPAA. Compliancy Group. Retrieved December 2, 2022, from https://compliancy-group.com/the-cia-triad-confidentiality-integrity-availability-for-hipaa/

[3] J. Zhang, X. Liang, Z. Zhang, S. He, and Z. Shi, "Re-dpoctor: Real-time health data releasing with w-day differential privacy," arXiv preprint arXiv:1711.00232, 2017. https://arxiv.org/pdf/1812.02282v1.pdf

[4] M. U. Hassan, M. H. Rehmani and J. Chen, "Differential Privacy Techniques for Cyber Physical Systems: A Survey," in IEEE Communications Surveys & Tutorials, vol. 22, no. 1, pp. 746-789, Firstquarter 2020, doi: 10.1109/COMST.2019.2944748

[5] Keshta, I., & Odeh, A. Security and Privacy of Electronic Health Records: Concerns and Challenges, Egyptian Informatics Journal, Volume 22, Issue 2, 2021, Pages 177-183, ISSN 1110-8665, https://doi.org/10.1016/j.eij.2020.07.003 (https://www.sciencedirect.com/science/article/pii/S1110866520301365)

[6] G. Perera, A. Holbrook, L. Thabane, G. Foster, D. J. Willison Views on Health Information Sharing and Privacy from Primary care Practices using Electronic Medical Records Int J Med Informatics, 80 (2) (2011), pp. 94-101 https://doi.org/10.1016/j.ijmedinf.2010.11.005

[7] J. Ancker, M. Silver, M. Miller, R. Kaushal Consumer Experience with and Attitude Toward Health Information Technology: A Nationwide Survey Am Medical Informatics Assoc, 1 (2012), pp. 152-156 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3555333/

[8] (OCR), O. for C. R. (2021, June 28). Health insurer pays $5.1 million to settle data breach affecting over 9.3 million people. HHS.gov. Retrieved February 1, 2022, from https://www.hhs.gov/hipaa/for-professionals/compliance-enforcement/agreements/excellus/index.html

[9] EHRIntelligence, K. M. (2018, September 18). Ehr Design, Interoperability Top List of physician pain points. EHRIntelligence. Retrieved February 1, 2022, from https://ehrintelligence.com/news/ehr-design-interoperability-top-list-of-physician-pain-points

[10] H. Wang, Z. Zheng, S. Xie, H. N. Dai, and X. Chen, "Blockchain Challenges and Opportunities: A Survey," International Journal of Web and Grid Services, vol. 14, no. 4, p. 352, 2018, https://www.inderscienceonline.com/doi/abs/10.1504/IJWGS.2018.095647

[11] S. Sarmah, (2018). Understanding Blockchain Technology. Computer Science and Engineering, 8(2), 23-29.

[12] Y. Zou, T. Meng, P. Zhang, W. Zhang and H. Li, "Focus on Blockchain: A Comprehensive Survey on Academic and Application," in IEEE Access, vol. 8, pp. 187182-187201, 2020, https://ieeexplore.ieee.org/abstract/document/9220919

[13] P. Garret, & J. Seidman. (2011, August 26). EMR vs EHR – What is the Difference? Health IT Buzz. Retrieved January 31, 2022, from https://www.healthit.gov/buzz-blog/electronic-health-and-medical-records/emr-vs-ehr-difference

[14] Lewis, A. (n.d.). A gentle introduction to blockchain technology web. https://miethereum.com/wp-content/uploads/2017/11/A.-A-Gentle-Introduction-To-Blockchain-Technology.pdf. Retrieved January 22, 2022, from https://miethereum.com/wp-content/uploads/2017/11/A.-A-Gentle-Introduction-To-Blockchain-Technology.pdf

[15] N. R. Pradhan, A. P. Singh, S. Verma. et al. A blockchain based lightweight peer-to-peer energy trading framework for secured high throughput micro-transactions. Sci Rep 12, 14523 (2022). https://doi.org/10.1038/s41598-022-18603-z

[16] D. Guru, S. Perumal, V. Varadarajan. Approaches towards Blockchain Innovation: A Survey and Future Directions. Electronics 2021, 10, 1219. https://doi.org/ 10.3390/electronics10101219

[17] Che-Ming Yang, Herng-Ching Lin, Polun Chang, Wen-Shan Jian, Taiwan's perspective on electronic medical records' security and privacy protection: Lessons learned from HIPAA, Computer Methods and Programs in Biomedicine, Volume 82, Issue 3, 2006, Pages 277-282, ISSN 0169-2607, https://doi.org/10.1016/j.cmpb.2006.04.002

[18] B. Cao, Z. Zhang, D. Feng, S. Zhang, L Zhang, M. Peng, Y. Li, Performance analysis and comparison of PoW, PoS and DAG based blockchains, Digital Communications and Networks, Volume 6, Issue 4, 2020, Pages 480-485, ISSN 2352-8648, https://doi.org/10.1016/j.dcan.2019.12.001.

(https://www.sciencedirect.com/science/article/pii/S2352864819301476)

[19] D. Guru, S. Perumal, S. Varadarajan. Approaches towards Blockchain Innovation: A Survey and Future Directions. Electronics 2021, 10, 1219. https://doi.org/ 10.3390/electronics10101219

[20] F. Casino, T. K. Dasaklis, Constantinos Patsakis, A systematic literature review of blockchain-based applications: Current status, classification and open issues, Telematics and Informatics, Volume 36, 2019, Pages 55-81, https://doi.org/10.1016/j.tele.2018.11.006

[21] Dubovitskaya, A., Xu, Z., Ryu, S., Schumacher, M., & Wang, F. (2018). Secure and Trustable Electronic Medical Records Sharing using Blockchain. AMIA ... Annual Symposium proceedings. AMIA Symposium, 2017, 650-659.

[22] A. Roehrs, C. A. da Costa, and R. da Rosa Righi, "OmniPHR: A distributed architecture model to integrate personal health records," Journal of Biomedical Informatics, vol. 71, pp. 70-81, Jul. 2017, doi: 10.1016/j.jbi.2017.05.012.

[23] S. Ølnes, J. Ubacht, M. Janssen, Blockchain in government: Benefits and implications of distributed ledger technology for information sharing, Government Information Quarterly, Volume 34, Issue 3, 2017, Pages 355-364, ISSN 0740-624X, https://doi.org/10.1016/j.giq.2017.09.007

[24] M. Iansiti and K. R. Lakhani, "The Truth About Blockchain," Harvard Business Review, no. January-February 2017, Jan. 01, 2017.

[25] H. Wang, Z. Zheng, S. Xie, H. N. Dai, and X. Chen, "Blockchain Challenges and Opportunities: A Survey," International Journal of Web and Grid Services, vol. 14, no. 4, p. 352, 2018, doi: 10.1504/IJWGS.2018.10016848

[26] L. Pawczuk, R. Massey, and D. Schatsky, "Breaking blockchain open Deloitte's 2018 global blockchain survey," Deloitte United States, 2018. https://www2.deloitte.com/us/en/pages/consulting/articles/innovation-blockchainsurvey.html (accessed Feb. 09, 2022).

[27] M. C. Wong, K. C. Yee, and C. Nohr, "Socio-technical consideration for blockchain technology in healthcare: the technological innovation needs clinical transformation to achieve the outcome of improving quality and safety of patient care," Studies in Health Technology and Informatics, vol. 247, pp. 636-640, 2018, doi: 10.3233/978-1-61499-852-5-636

[28] IBM, “Blockchain: The Chain of Trust and its Potential to Transform Healthcare – Our Point of View,” Aug. 2016. Accessed: Feb. 11, 2022. [Online]. Available: https://www.healthit.gov/sites/default/files/8-31-blockchain-ibm_ideation-challenge_aug8.pdf

[29] A. W. Peters, B. M. Till, J. G. Meara, and S. Afshar, “Blockchain technology in health care: A primer for surgeons,” The Bulletin, Dec. 06, 2017

[30] D. V. Dimitrov, “Blockchain Applications for Healthcare Data Management,” Healthcare Informatics Research, vol. 25, no. 1, p. 51, 2019, doi: 10.4258/hir.2019.25.1.51

[31] C. Dwork, F. McSherry, K. Nissim, & A. Smith. (2006). Calibrating Noise to Sensitivity in Private Data Analysis. In Proceedings of the 3rd Conference on Theory of Cryptography (pp. 265-284).

[32] C. Dwork. (2008). Differential Privacy: A Survey of Results. Lecture Notes in Computer Science Theory and Applications of Models of Computation, 1-19. doi:10.1007/978-3-540-79228-4

[33] C. Dwork, A. Roth, “The Algorithmic Foundations of Differential Privacy,” in The Algorithmic Foundations of Differential Privacy, now, 2014. https://ieeexplore.ieee.org/document/8187424

[34] J. Lee, and C. Clifton. (2011) How Much Is Enough? Choosing ϵ for Differential Privacy. In: Lai X., Zhou J., Li H. (eds) Information Security. ISC 2011. Lecture Notes in Computer Science, vol 7001. Springer, Berlin, Heidelberg. https://doi.org/10.1007/ 978-3-642-24861-0_22

[35] J. Hsu et al., “Differential Privacy: An Economic Method for Choosing Epsilon,” 2014 IEEE 27th Computer Security Foundations Symposium, Vienna, 2014, pp. 398-410, doi: 10.1109/CSF.2014.35

[36] C. Dwork, “A firm foundation for private data analysis,” Communications of the ACM, vol. 54, no. 1, pp. 86-95, 2011.

[37] T. Zhu, G. Li, W. Zhou, S.Y. Philip, Differential privacy and applications, vol. 69 (Springer International Publishing)

[38] U. M. Hassan. M. H. Rehmani, & J. Chen. (2019). Differential Privacy Techniques for Cyber Physical Systems: a survey. IEEE Communications Surveys & Tutorials, 22(1), 746-789. Firstquarter 2020, doi: 10.1109/COMST.2019.2944748

[39] C. Dwork, A. Roth, et al.: An adaptive approach to real-time aggregate monitoring with differential privacy. Foundations and Trends in Theoretical Computer Science 9(3-4), 211 (2014)

[40] Y. Tian, B. Song, T. Ma, A. Al-Dhelaan and M. Al-Dhelaan, "Bi-Tier Differential Privacy for Precise Auction-Based People-Centric IoT Service," in IEEE Access, vol. 9, pp. 55036-55044, 2021, doi: 10.1109/ACCESS.2021.3067138

[41] F. Liu, "Generalized Gaussian Mechanism for Differential Privacy," IEEE Transactions on Knowledge and Data Engineering, vol. 31, no. 4, pp. 747-756, Apr. 2018.

[42] I. C. Lin, T. C. Liao. A Survey of Blockchain Security Issues and Challenges. Int. J. Netw. Secur. 2017, 19, 653-659

[43] T. Clohessy, T. Acton, and N. Rogers. (2019). "Blockchain Adoption: Technological, Organisational and Environmental Considerations", in Treiblmaier, H. and Beck, R. "Business Transformation through Blockchain", Volume 1, Cham, Switzerland: Palgrave Macmillan, pp. 47-76. DOI: 10.1007/978-3-319-98911-2

[44] T. McGhin, K.-K. R. Choo, C. Z. Liu, and D. He, "Blockchain in healthcare applications: Research challenges and opportunities," Journal of Network and Computer Applications, vol. 135, pp. 62-75, Jun. 2019, doi: 10.1016/j.jnca.2019.02.027

[45] G. Zyskind, O. Nathan, and A. Pentland, Decentralizing privacy: Using blockchain to protect personal data. Proceedings of IEEE Security and Privacy Workshops: 180-184, 2015. https://doi.org/ 10.1109/SPW.2015.27

[46] A. Lippman, T. Vieira, A. Ekblaw, A. Azaria, et al., MedRec: Using blockchain for medical data. Presented at International Conference on Open & Big Data. 2016. Available: http://ieeexplore.ieee.org/document/7573685/

[47] IBM, "Blockchain: The Chain of Trust and its Potential to Transform Healthcare – Our Point of View," Aug. 2016. Accessed: Feb. 11, 2022. [Online]. Available: https://www.healthit.gov/sites/default/files/8-31-blockchain-ibm_ideation-challenge_aug8.pdf

[48] S. Hogan, H. Fraser, P. Korsten, V. Pureswaran, and R. Gopinath, "Healthcare rallies for blockchains Keeping patients at the center," IBM Institute for Business Value, Dec. 2016. https://www.ibm.com/downloads/cas/BBRQK3WY

[49] Deloitte, "Blockchain to Blockchains in Life Sciences and Health Care," Deloitte, 2018. https://www2.deloitte.com/uk/en/insights/topics/understanding-blockchain-potential/global-blockchain-survey-2019/2019-adoption-by-industry.html

[50] A. Lippman, T. Vieira, A. Ekblaw, A. Azaria, et al., MedRec: Using blockchain for medical data. Presented at International Conference on Open & Big Data. 2016. Available: http://ieeexplore.ieee.org/document/7573685/

[51] D. Schatsky, A. Arora, and A. Dongre, "Blockchain and the Five Vectors of Progress," Deloitte, 2018. Accessed: Feb. 13, 2022. [Online]. Available: https://www2.deloitte.com/insights/us/en/focus/signals-for-strategists/value-ofblockchain-applications-interoperability.html

[52] R. Krawiec and M. White, "Blockchain: Opportunities for health care," Deloitte United States, Aug. 2016. https://www2.deloitte.com/us/en/pages/publicsector/articles/blockchain-opportunities-for-health-care.html (accessed Feb. 13, 2022).

[53] J. Soria-Comas, J. Domingo-Ferrer, D. Sanchez, and D. Megias, "Individual differential privacy: A utility-preserving formulation of differential privacy guarantees," IEEE Transactions on Information Forensics and Security, vol. 12, no. 6, pp. 1418-1429, 2017.

[54] E. ElSalamouny and S. Gambs, “Differential privacy models for location-based services,” Transactions on Data Privacy, vol. 9, no. 1, pp. 15-48, 2016.

[55] X. He, G. Cormode, A. Machanavajjhala, C. M. Procopiuc, and D. Srivastava, “Dpt: differentially private trajectory synthesis using hierarchical reference systems,” Proceedings of the VLDB Endowment, vol. 8, no. 11, pp. 1154-1165, 2015

[56] A. Machanavajjhala, A. Korolova, and A. D. Sarma. Personalized social recommendations accurate or private? PVLDB, 4(7):440-450, 2011. https://arxiv.org/ftp/arxiv /papers/1105/1105.4254.pdf

[57] F. McSherry and R. Mahajan. Differentially-private network trace analysis. In Proc. SIGCOMM, pages 123-134, 2010. https://ratul.org/papers/sigcomm2010-privacy.pdf

[58] P. Jain, M. Gyanchandani, & N. Khare. Differential privacy: its technological prescriptive using big data. J Big Data 5, 15 (2018). https://link.springer.com/article/10.1186/s40537-018-0124-9

[59] C. Dwork, K. Talwar, A. Thakurta, and L. Zhang, “Analyze gauss: optimal bounds for privacy-preserving principal component analysis,” in Proceedings of the forty-sixth annual ACM symposium on Theory of computing, 2014, pp. 11-20.

[60] S. Garfinkel, J. Abowd, and S. Powazek. 2018. Issues Encountered Deploying Differential Privacy. In Proceedings of the 2018 Workshop on Privacy in the Electronic Society (WPES’18). Association for Computing Machinery, New York, NY, USA, 133-137. https://doi.org/10.1145/3267323.3268949

[61] P. Wang, J. Huang, Z. Cui, L. Xie, J. Chen, A Gaussian error correction multi-objective positioning model with nsga-ii, Concurr. Comput.: Pract. Exper. 32 (5) (2020) e5464.

[62] J. Vora, A. Nayyar, S. Tanwar, S. Tyagi, N. Kumar, M. S. Obaidat, J. J. Rodrigues, BHEEM: A blockchain-based framework for securing electronic health records, in: IEEE Globecom Workshops (GC Wkshps), 2018, pp. 1-6. https://ieeexplore.ieee.org/ abstract/document/8644088

[63] J. Xie, H. Tang, T. Huang, F.R. Yu, R. Xie, J. Liu, Y. Liu, A survey of Blockchain Technology applied to smart cities: Research issues and challenges, IEEE Commun. Surv. Tutor. (2019)

[64] B. Shickel, P. Tighe, A. Bihorac, and P. Rashidi, "Deep ehr: a survey of recent advances in deep learning techniques for electronic health record (ehr) analysis," arXiv preprint arXiv:1706.03446, 2017.

[65] T. Sahama, L. Simpson, and B. Lane, "Security and privacy in ehealth: Is it possible?" in IEEE 15th International Conference on e-Health Networking, Applications & Services (Healthcom), 2013, pp. 249-253.

[66] Z. Guan, Z. Lv, X. Du, L. Wu, and M. Guizani, "Achieving data utility privacy tradeoff in internet of medical things: A machine learning approach," Future Generation Computer Systems, in Print, 2019.

[67] B. K. Beaulieu-Jones, W. Yuan, S. G. Finlayson, and Z. S. Wu, "Privacy-preserving distributed deep learning for clinical data," arXiv preprint arXiv:1812.01484, 2018.

[68] W. Zhang, H. Zou, L. Luo, Q. Liu, W. Wu, and W. Xiao, "Predicting potential side effects of drugs by recommender methods and ensemble learning," Neurocomputing, vol. 173, pp. 979-987, 2016.

[69] Q. Zhang, G. Zhang, J. Lu, and D. Wu, "A framework of hybrid recommender system for personalized clinical prescription," in IEEE 10th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), 2015, pp. 189-195.

[70] B. Esteban, A. Tejeda-Lorente, C. Porcel, M. Arroyo, and E. Herrera-Viedma, "Tplufib-web: A fuzzy linguistic web system to help in the treatment of low back pain problems," Knowledge-Based Systems, vol. 67, pp. 429-438, 2014.

[71] C. Han and K. Wang, "Sensitive disclosures under differential privacy guarantees," in Proc. IEEE Int. Congr. Big Data, Jun. 2015, pp. 110-117.

[72] T. Li, A. K. Sahu, A. Talwalkar. & V. Smith. Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine 37, 50-60 (IEEE, 2020).

[73] P. Kairouz, et al. Advances and open problems in federated learning. arXiv preprint arXiv:1912.04977 (2019).

[74] N. Rieke, J. Hancox, W. Li. et al. The future of digital health with federated learning. npj Digit. Med. 3, 119 (2020). https://doi.org/10.1038/s41746-020-00323-1

[75] J. Lee. et al. Privacy-preserving patient similarity learning in a federated environment: development and analysis. JMIR Med. Inform. 6, e20 (2018).

[76] T. S. Brisimi. et al. Federated learning of predictive models from federated electronic health records. Int. J. Med. Inform. 112, 59-67 (2018).

[77] A. G. Roy, S. Siddiqui, S. Pölsterl, N. Navab, & C. Wachinger Braintorrent: a peer-to-peer environment for decentralized federated learning. arXiv preprint arXiv:1905.06731 (2019).

[78] F. Wang, L. P. Casalino. & D. Khullar. Deep learning in medicine–promise, progress, and challenges. JAMA Intern. Med. 179, 293-294 (2019).

[79] M. J. Sheller, G. A. Reina, B. Edwards, J. Martin. & Bakas, S. Multi-institutional deep learning modeling without sharing patient data: a feasibility study on brain tumor segmentation. In International MICCAI Brainlesion Workshop, 92-104 (Springer, 2018)

[80] M. Abadi. et al. Deep learning with differential privacy. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, 308-318 (ACM, 2016)

[81] O. Dorgham, B. Al-Rahamneh, A. Almomani and K.F. KhatatnehInt. J. Cloud Appl. Computing (IJCAC), 8 (1) (2018), pp. 154-172

[82] C.-L. Chen, P.-T. Huang, Y.-Y. Deng, H.-C. Chen and Y.-C. WangHuman-Centric Computing Information Sci., 10 (2020), pp. 1-31

[83] M. Cifuentes, M. Davis, D. Fernald, R. Gunn, P. Dickinson and D.J. Cohen J Am Board Fam Med, 28 (Supplement 1) (2015), pp. S63-S72

[84] J. Ancker, M. Silver, M. Miller and R. Kaushal Am Medical Informatics Assoc, 1 (2012), pp. 152-156

[85] Win. K. T. (2005), A review of Security of electronic health records, Health Information Management Journal, vol. 34, is. 1, pp. 13-18

[86] D. Lafky and T. Horan Health Informatics J, 17 (1) (2011), pp. 63-71

[87] S.B. Wikina Perspect Health Inf Mana, 2014 (2014), pp. 1-16

[88] K. L. van der Veen, R. Seggers, P. Bloem, and G. Patrini, "Three tools for practical differential privacy," CoRR, vol. abs/1812.02890, 2018.

[89] W. Zhang, "Privacy-preserving statistical tools: Differential privacy and beyond," Ph.D. dissertation, Georgia Institute of Technology, Atlanta, GA, USA, 2021.

[90] A. Alnemari, C. J. Romanowski, and R. K. Raj, "An adaptive differential privacy algorithm for range queries over healthcare data," in IEEE International Conference on Healthcare Informatics (ICHI), 2017, pp. 397-402

[91] J. L. Raisaro, J. Troncoso-Pastoriza, M. Misbach, J. S. Sousa, S. Pradervand, E. Missiaglia, O. Michielin, B. Ford, and J.-P. Hubaux, "Medco: Enabling secure and privacy-preserving exploration of distributed clinical and genomic data," IEEE/ACM transactions on computational biology and bioinformatics, in Print, 2018.

[92] A. Dewey, & A. Drahota (2016) Introduction to systematic reviews: online learning module Cochrane Training https://training.cochrane.org/interactivelearning/module-1-introduction-conducting-systematic-reviews

[93] M. Salama, R. Bahsoon, N. Bencomo, Chapter 11 - Managing Trade-offs in Self-Adaptive Software Architectures: A Systematic Mapping Study, Editor(s): Ivan Mistrik, Nour Ali, Rick Kazman, John Grundy, Bradley Schmerl, Managing Trade-Offs in Adaptable Software Architectures, Morgan Kaufmann, 2017, Pages 249-297, ISBN 9780128028551, (https://www.sciencedirect.com/science/article/pii/B9780128028551000113)

[94] https://arxiv.org/ftp/arxiv/papers/1702/1702.02653.pdf

[95] A. Fink: Conducting research literature reviews: from the internet to paper. Thousand Oaks, Thousand Oaks (2019)

[96] N. Saleheen, S. Chakraborty, N. Ali, M. M. Rahman, S. M. Hossain, R. Bari, E. Buder, M Srivastava, S. Kumar. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (2016), pp. 706-717

[97] C. Lin, Z. Song, H. Song, Y. Zhou, Y. Wang, G. Wu.: Differential privacy preserving in big data analytics for connected health. J. Med. Syst. 40(4), 97 (2016)

[98] M. Hadian, X. Liang, T. Altuwaiyan, M. M. Mahmoud: In: 2016 IEEE Global Communications Conference (GLOBECOM) (IEEE, 2016), pp. 1-6

[99] N. Mohammed, S. Barouti, D. Alhadidi, and R. Chen, “Secure and private management of healthcare databases for data mining,” in IEEE 28th International Symposium on Computer-Based Medical Systems (CBMS), 2015, pp. 191-196

[100] W. Tang, J. Ren, K. Deng, and Y. Zhang, “Secure Data Aggregation of Lightweight E-healthcare IoT Devices with Fair Incentives,” IEEE Internet of Things Journal, in Print, 2019 https://ieeexplore.ieee.org/abstract/document/8737719

[101] J. L. Raisaro, G. Choi, S. Pradervand, R. Colsenet, N. Jacquemont, N. Rosat, V. Mooser, and J.-P. Hubaux, “Protecting privacy and security of genomic data in i2b2 with homomorphic encryption and differential privacy,” IEEE/ACM transactions on computational biology and bioinformatics, vol. 15, no. 5, pp. 1413-1426, Sep. 2018

[102] R. Gao, X. Ma: In: 2018 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom) (IEEE, 2018), pp. 737-743

[103] F. McSherry and R. Mahajan. Differentially-private network trace analysis. In Proc. SIGCOMM, pages 123-134, 2010. https://ratul.org/papers/sigcomm2010-privacy.pdf

[104] C. Luo, X. Liu, W. Xue, Y. Shen, J. Li, W. Hu, Liu, A.X.: Predictable privacy-preserving mobile crowd sensing: a tale of two roles. IEEE/ACM Trans. Network. 27(1), 361 (2019)

[105] A. Narayanan and V. Shmatikov. Robust de-anonymization of large sparse datasets. In Proc. S&P. IEEE, May 2008 https://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf

[106] A. Narayanan and V. Shmatikov. De-anonymizing social networks. In IEEE Symposium on Security and Privacy (S&P), Oakland, California, pages 173-187, 2009 https://ieeexplore.ieee.org/abstract/document/5207644

[107] A. C. Valdez and M. Ziefle, "The users' perspective on the privacy utility trade-offs in health recommender systems," International Journal of Human-Computer Studies, in Print, vol. 121, pp. 108-121, Jan. 2019 https://moam.info/arxiv181202282v1-cscr-6-dec-2018_5c362c14097c478d538b456d.html

[108] M. S. Ali, M. Vecchio, M. Pincheira, K. Dolui, F. Antonelli, M. H. Rehmani, Applications of blockchains in the internet of things: A comprehensive survey, IEEE Commun. Surv. Tutor. 21 (2) (2018) 1676-1717

[109] X. Wang, K. Yang, Asynchronous blockchain-based privacy-preserving training framework for disease diagnosis, in: IEEE International Conference on Big Data (Big Data), 2019, pp. 5469-5473.

[110] Q. Hu, R. Chen, H. Yang, and S. Kumara, ''Privacy-preserving data mining for smart manufacturing,'' Smart Sustain. Manuf. Syst., vol. 4, no. 2, Jul. 2020, Art. no. 20190043

[111] J. W. Kim, J. H. Lim, S. M. Moon, H. Yoo, B. Jang: In: 2019 IEEE International Conference on Consumer Electronics (ICCE) (IEEE, 2019), pp. 1-4

[112] M. Yang: Improving privacy preserving in modern applications. Deakin University, Tech. rep. (2019)

[113] Z. Zhang, B. Han, H. C. Chao, F. Sun, L. Uden, Tang, D.: A new weight and sensitivity based variable maximum distance to average vector algorithm for wearable sensor data privacy protection. IEEE Access 7, 104045 (2019)

[114] Y. Zhang, Y. Qu, L. Gao, T. H. Luan, X. Zheng, S. Chen, Y. Xiang: APDP: Attack-Proof Personalized Differential Privacy Model for a Smart Home. IEEE Access 7, 166593 (2019)

[115] V. Rastogi, S. Nath: In: Proceedings of the 2010 ACM SIGMOD International Conference on Management of data (2010), pp. 735-746

[116] L. Fan, L. Xiong: An adaptive approach to real-time aggregate monitoring with differential privacy. IEEE Trans. Knowl. Data Eng. 26(9), 2094 (2013)

[117] H. Li, Y. Dai, and X. Lin, “Efficient e-health data release with consistency guarantee under differential privacy,” in IEEE 17th International Conference on

[118] M. Saifuzzaman, T. N. Ananna, M. J. M. Chowdhury. et al. A systematic literature review on wearable health data publishing under differential privacy. Int. J. Inf. Secur. 21, 847-872 (2022). https://doi.org/10.1007/s10207-021-00576-1

[119] Y. Cao, L. Xiong, M. Yoshikawa, Y. Xiao, S. Zhang: ConTPL: controlling temporal privacy leakage in differentially private continuous data release. Proc. VLDB Endowm. 11(12), 2090 (2018)

[120] J. Raisaro, J. Scheibner, J. Troncoso-Pastoriza, M. Ienca, J. Fellay, E. Vayena, J. Hubaux Revolutionizing Medical Data Sharing Using Advanced Privacy-Enhancing Technologies: Technical, Legal, and Ethical Synthesis J Med Internet Res 2021;23(2):e25120 URL: https://www.jmir.org/2021/2/e25120DOI: 10.2196/25120

[121] M. Lenca, A. Ferretti, S. Hurst, M. Puhan, C. Lovis, E. Vayena (2018) Considerations for ethics review of big data health research: A scoping review. PLoS ONE 13(10): e0204937. https://doi.org/10.1371/journal.pone.0204937

[122] G. R. Signorelli, F. Lehocki, M. M. Fernadez, G. O'Neil, D. O'Connor, & J.Garicia-Gpmez (2019). A research roadmap: connected health as an enabler of cancer patient support. *Journal of Medical Internet Research*, *21*(10), e14360. https://www.jmir.org/2019/10/e14360/

[123] H. S. A. Fang, T. H. Tan, Y. F. C. Tan, & C. J. M. Tan (2021). Blockchain Personal Health Records: Systematic Review. *Journal of medical Internet research*, *23*(4), e25094. https://doi.org/10.2196/25094

[124] G. G. Dagher, J. Mohler, M. Milojkovic, P. B. Marella, Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology, Sustainable Cities and Society, Volume 39, 2018, Pages 283-297, ISSN 2210-6707, https://doi.org/10.1016/j.scs.2018.02.014. (https://www.sciencedirect.com/science/article/pii/S2210670717310685)

[125] A. A. Mamun, S. Azam and C. Gritti, "Blockchain-Based Electronic Health Records Management: A Comprehensive Review and Future Research Direction," in *IEEE Access*, vol. 10, pp. 5768-5789, 2022, doi: 10.1109/ACCESS.2022.3141079.

[126] Y. R. Park, E. Lee, W. Na, S. Park, Y Lee, J. Lee. Is Blockchain Technology Suitable for Managing Personal Health Records? Mixed-Methods Study to Test Feasibility. *J Med Internet Res.* 2019 Feb 08;21(2):e12533. doi: 10.2196/12533.

[127] R. Sangeetha, B. Harshini, T. K. P. Rajagopa. Electronic Health Record System using Blockchain. *Int Res J Multidiscip Technovation.* 2019 Mar 25;1(2):57–61. doi: 10.34256/irjmt1927

[128] R. R. Charanya, SeFra: A Secure Framework to Manage eHealth Records Using Blockchain Technology. *International Journal of E-Health and Medical Communications (IJEHMC)* 2020:1–5. doi: 10.4018/ijehmc.2020010101.

[129] https://www.simplilearn.com/blockchain-programming-languages-article