

AI-BLOCKCHAIN BASED HEALTHCARE RECORDS
MANAGEMENT SYSTEM

BY

ALAA HADDAD

A dissertation submitted in fulfillment of the requirement for
the degree of Doctor of Philosophy (Engineering).

Kulliyyah of Engineering
International Islamic University Malaysia

OCTOBER 2023

ABSTRACT

Accessing healthcare services by several stakeholders for diagnosis and treatment has become quite prevalent owing to the improvement in the industry and high levels of patient mobility. Due to the confidentiality and high sensitivity of electronic healthcare records (EHR), the majority of EHR data sharing is still conducted via fax or mail because of the lack of systematic infrastructure support for secure and reliable health data transfer, delaying the process of patient care. As a result, it is critically essential to provide a framework that allows for the efficient exchange and storage of large amounts of medical data in a secure setting, where the storing the data over the cloud do not remain secure all the time. Since the data are accessible to the end user only by using the interference of a third party, it is prone to breach of authentication and integrity of the data. This thesis introduces the development of a Patient-Centered Blockchain-Based EHR Management (PCBEHRM) system that allows patients to manage their healthcare records across multiple stakeholders and to facilitate patient privacy and control without the need for a centralized infrastructure. In addition, the proposed system ensures a secure and optimized scheme for sharing data while maintaining data security and integrity over the Inter Planetary File System (IPFS). Further, the proposed system introduces a sophisticated End to End Encryption (E2EE) functionality by combining the ECC (Elliptic Curve Cryptography) method and the Advanced Encryption Standard (AES) method. This is to enhance the security of system, reduce the computational power for memory optimization, and ensure authentication and data integrity. We have also demonstrated how the proposed system design enables stakeholders such as patients, labs, researchers, etc., to obtain patient-centric data in a distributed and secure manner that is integrated using a web-based interface for the patient and all users to initiate the EHR sharing transactions. Finally, the thesis enhances the proposed PCBEHRM system with deep learning artificial intelligence capabilities to revolutionize the management of the EHR and offer an add-on diagnostic tool based on the captured EHR metadata. Deep learning in healthcare now had become incredibly powerful for supporting clinics and in transforming patient care in general and is increasingly applied for the detection of clinically important features in the images beyond what can be perceived by the naked human eye. Chest X-ray images are one of the most common clinical methods for diagnosing several diseases. The proposed enhancement integrated deep learning feature is a developed lightweight solution that can detect 14 different chest conditions from an X-ray image. Given an X-ray image as input, our classifier outputs a label vector indicating which of 14 disease classes does the image fall into. The proposed diagnostic add-on tool focuses on predicting the 14 diseases to provide insight for future chest radiography research. Finally, the proposed system was tested in Microsoft Windows@ environment by compiling a smart contract prototype using Truffle and deploying it on Ethereum using Web3. The proposed system was evaluated in terms of the projected medical data storage costs for the IPFS on blockchain, and the execution time for a different number of peers and document sizes. The results show that the proposed system achieves a reduced storage cost of 73.4172% and a 76% in execution time in

comparison to other proposed systems in the open literature. The Results of the study conclude that the proposed strategy is both efficient and practicable. The add-on deep learning diagnostic feature flags any present diseases predicted from the health records and assists doctors and radiologists in making a well-informed decision during the detection and diagnosis of the disease.



ملخص البحث

أصبح الوصول إلى خدمات الرعاية الصحية من قبل العديد من أصحاب المصلحة للتشخيص والعلاج سائداً إلى حد كبير بسبب التحسن في الصناعة والمستويات العالية من تنقل المرضى. نظراً للسرية والحساسية العالية لسجلات الرعاية الصحية الإلكترونية (EHR) ، لا تزال غالبية مشاركة بيانات السجلات الصحية الإلكترونية تتم عبر الفاكس أو البريد بسبب عدم وجود دعم منهجي للبنية التحتية لنقل البيانات الصحية بشكل آمن وموثوق ، مما يؤدي إلى تأخير عملية رعاية المرضى . نتيجة لذلك ، من الضروري للغاية توفير إطار عمل يسمح بتبادل وتخزين كميات كبيرة من البيانات الطبية بكفاءة في بيئة آمنة ، حيث لا يظل تخزين البيانات عبر السحابة آمناً طوال الوقت. نظراً لأن البيانات لا يمكن الوصول إليها إلا من خلال استخدام تدخل طرف ثالث، فهي عرضة لخرق المصادقة وسلامة البيانات.

تقدم هذه الأطروحة تطوير نظام إدارة السجلات الطبية الإلكترونية (PCBEHRM) القائم على Blockchain والمتمحور حول المريض والذي يسمح للمرضى بإدارة سجلات الرعاية الصحية الخاصة بهم عبر العديد من أصحاب المصلحة وتسهيل خصوصية المريض والتحكم فيه دون الحاجة إلى بنية تحتية مركزية. بالإضافة إلى ذلك، يضمن النظام المقترح مخططاً آمناً ومحسناً لمشاركة البيانات مع الحفاظ على أمن البيانات وسلامتها عبر نظام الملفات الكوكبي (IPFS). علاوة على ذلك، يقدم النظام المقترح وظيفة تشفير من طرف إلى طرف (E2EE) متطورة من خلال الجمع بين طريقة (ECC تشفير منحني إهليلجي) وطريقة معيار التشفير المتقدم (AES) هذا لتعزيز أمان النظام، وتقليل القوة الحسابية لتحسين الذاكرة ، وضمان المصادقة وتكامل البيانات.

تم استخدام Ethereum blockchain و IPFS للتنفيذ لتخزين السجلات نظراً لمزايا توزيعها، وضمان ثبات السجلات، والسماح بالتخزين اللامركزي للبيانات الوصفية الطبية (على سبيل المثال ، التقارير الطبية). لضمان وجود سياسة تحكم في الوصول آمنة وموزعة وجديرة بالثقة، اقترحت الأطروحة عقد Ethereum ذكي يسمى بروتوكول التحكم في الوصول المرتكز على المريض (PCAC). لقد أوضحنا أيضاً كيف يمكن تصميم النظام المقترح أصحاب المصلحة مثل المرضى والمختبرات والباحثين، وما إلى ذلك، من الحصول على بيانات تتمحور حول المريض بطريقة موزعة وآمنة تتكامل باستخدام واجهة على شبكة الإنترنت للمريض وجميع المستخدمين بدء معاملات مشاركة السجلات الصحية الإلكترونية.

أخيراً، تعزز الأطروحة نظام PCBEHRM المقترح بقدرات الذكاء الاصطناعي للتعلم العميق لإحداث ثورة في إدارة السجلات الصحية الإلكترونية وتقديم أداة تشخيص إضافية تعتمد على البيانات الوصفية للسجلات الصحية الإلكترونية التي تم التقاطها. أصبح التعلم العميق في مجال الرعاية الصحية الآن قوياً بشكل لا يصدق لدعم العيادات وفي تحويل رعاية المرضى بشكل عام ويتم تطبيقه بشكل متزايد للكشف عن الميزات المهمة سريرياً في الصور بما يتجاوز ما يمكن أن تراه العين المجردة. تعد صور الصدر بالأشعة السينية واحدة من أكثر الطرق السريرية شيوعاً لتشخيص العديد من الأمراض. تعد ميزة التعلم العميق المدمجة والمعززة المقترحة حلاً خفيف الوزن مطوراً يمكنه اكتشاف 14 حالة صدر مختلفة من صورة الأشعة السينية. بالنظر إلى صورة الأشعة السينية كمدخلات، يقوم المصنف لدينا بإخراج متجه تسمية يشير إلى أي فئة من فئات المرض الأربعة عشر التي تقع فيها الصورة. تركز الأداة الإضافية التشخيصية المقترحة على التنبؤ بـ 14 مرضاً لتوفير نظرة ثاقبة لأبحاث التصوير الشعاعي للصدر في المستقبل.

أخيراً، تم اختبار النظام المقترح في بيئة @ Microsoft Windows عن طريق تجميع نموذج أولي ذكي للعقد باستخدام Truffle ونشره على Ethereum باستخدام Web3. تم تقييم النظام المقترح من حيث تكاليف تخزين البيانات الطبية المتوقعة لـ IPFS على blockchain ، ووقت التنفيذ لعدد مختلف من النظراء وأحجام المستندات.

أظهرت النتائج أن النظام المقترح يحقق تكلفة تخزين مخفضة بنسبة 73.4172% و76% في وقت التنفيذ مقارنة بالأنظمة الأخرى المقترحة في الأدبيات المفتوحة. خلصت نتائج الدراسة إلى أن الاستراتيجية المقترحة فعالة وعملية. تعمل ميزة

التشخيص الإضافية للتعلم العميق على تمييز أي أمراض حالية متوقعة من السجلات الصحية وتساعد الأطباء وأخصائيي الأشعة في اتخاذ قرار مستنير أثناء اكتشاف المرض وتشخيصه.



APPROVAL PAGE

The thesis of Alaa Haddad has been approved by the following:



Mohamed Hadi Habaebi
Supervisor



Md. Rafiqul Islam
Co-supervisor

Suriza Ahmad Zabidi
Co-Supervisor

Mashkuri Yaacob
Internal Examiner

Musse Mohamud Ahmed
External Examiner

Akram M Z M Khedher
Chairman

DECLARATION

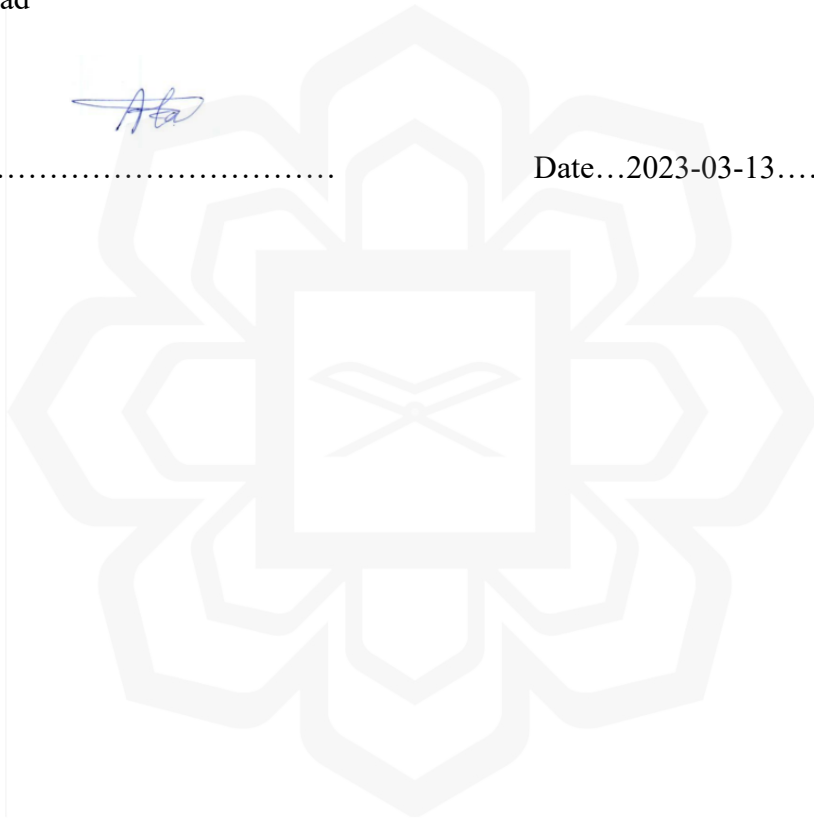
I hereby declare that this thesis is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Alaa Haddad



Signature.....

Date...2023-03-13.....



INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

**DECLARATION OF COPYRIGHT AND AFFIRMATION OF
FAIR USE OF UNPUBLISHED RESEARCH**

**AI-BLOCKCHAIN BASED HEALTHCARE MANAGEMENT
SYSTEM**

I declare that the copyright holder of this thesis are jointly owned by the student and IIUM.

Copyright © 2023 Alaa Haddad and International Islamic University Malaysia. All rights reserved.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

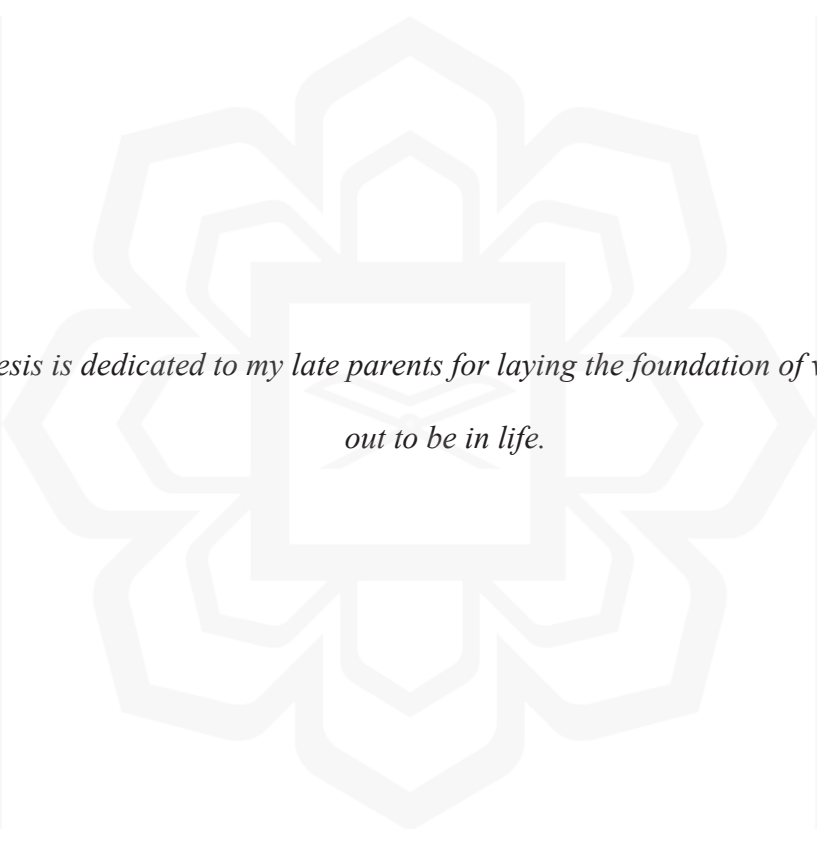
1. Any material contained in or derived from this unpublished research may only be used by others in their writing with due acknowledgement.
2. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purpose.
3. The IIUM library will have the right to make, store in a retrieval system and supply copies of this unpublished research if requested by other universities and research libraries.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Alaa Haddad

.....
Signature

.....2023-03-13.....
Date



*This thesis is dedicated to my late parents for laying the foundation of what I turned
out to be in life.*

ACKNOWLEDGEMENTS

Indeed, all praise is due to Allah, we praise Him, we seek His aid, and we ask for His forgiveness. We seek refuge in Allah from the evil of our actions and from the evil consequences of our actions. Whom Allah guides, no one can misguide, and whom Allah misguides, no one can guide. I bear witness that there is no god worthy of worship except Allah, and I bear witness that Muhammad is the servant and messenger of Allah.

I would like to express my sincere appreciation to my supervisor. This thesis would not have been possible without the kind support of my supervisor: Professor Mohamed Hadi Habaebi. Thank you for your constant advice, patience, and energy until the completion of this thesis. Despite the difficult conditions created by the global pandemic of Covid 19, I would like to thank you for great support, which definitely helped me to complete this work.

Many thanks to my co-supervisors Prof. Dr. Md. Rafiqul Islam and Assoc. Prof. Dr. Suriza Ahmad Zabidi for their guidance and support.

To the members of my supervisory committee, thank you for the guidance offered.

I am forever indebted to my late father for his prayer, and his doa'a always even in last his life day to continue my Ph.D. which is his ambition. I am forever indebted especially to my small family, my husband and my daughter for their endless support, encouragement and patience in all time that always have my back and give motivation in finishing this project. especially my dear mother and my brothers and for their love and encouragement.

Once again, we glorify Allah for His endless mercy on us one of which is enabling us to successfully round off the efforts of writing this thesis. Alhamdulillah

TABLE OF CONTENTS

Abstract	i
Abstract in Arabic	iii
Approval Page	v
Declaration	vi
Copyright	vii
Dedication	viii
Acknowledgements	ix
List of Tables	xiv
List of Figures	xv
CHAPTER ONE: INTRODUCTION	1
1.1 Background of study	1
1.2 Research Questions	3
1.2 Problem Statement	4
1.3 Objectives	5
1.4 Motivation	5
1.5 Research Scope	6
1.6 Research Philosophy	6
1.7 Research Methodology	7
1.8 Thesis Breakdown	7
CHAPTER TWO: LITERATURE REVIEW	9
2.1 Introduction	9
2.2 Background of Blockchain	14
2.2.1 Blockchain	14
2.2.2 Digital Signature	16
2.2.3 Algorithms for Building Consensus	17
2.2.4 Smart Contract	19
2.3 Blockchain applications in health records system	19
2.3.1 Data Management in Electronic Medical Record	19
2.3.2 Blockchain and Data Protection in Healthcare	19
2.3.3 Personal Health Record (PHR) Data Management on The Blockchain	21
2.4 Overview of Artificial Intelligence Health Records Management System	23
2.4.1 The Challenges of Using AI in Health Records	24
2.4.1.1 Transparency	24
2.4.1.2 Transportability	25
2.4.1.3 Training Size	26
2.4.2 AI Algorithm in Healthcare Systems	27

2.4.2.1 Supervised Algorithm	27
2.4.2.1.1 Artificial Neural Network (ANN)	27
2.4.2.1.2 Support Vector Machine (SVM).....	28
2.4.2.1.3 Decision Tree Random Forest	29
2.4.2.2 Unsupervised Algorithm	30
2.4.2.2.1 Association Analysis	30
2.4.2.3 Network Approach	30
2.5 Managing EHR Using AI and Blockchain	36
2.5.1 Methodology of the Literature Review	38
2.5.2 Research Questions	39
2.5.3 Filtering the literature of the study	39
2.5.4 Inclusion and Exclusion Criterion	46
2.5.5 Privacy and Security Issues	46
2.6 Taxonomy of AI-Blockchain	52
2.6.1 Decentralized Applications	52
2.6.1.1 Autonomous Computing	53
2.6.1.2 Optimization	55
2.6.1.3 Planning	55
2.6.1.4 Learning	56
2.6.2 Decentralized Operation	56
2.6.2.1 Storage	57
2.6.2.2 Data Management	57
2.6.2.3 Deployment	58
2.6.3 Blockchain Types for AI Application	58
2.6.3.1 Public	59
2.6.3.2 Private	59
2.6.3.3 Blockchain-As-A-Service	60
2.6.4 Decentralized Infrastructure	60
2.6.5 The Role of Consensus Protocol	62
2.6.5.1 Proof-of-Work (Pow)	62
2.6.5.2 Proof-of-Stake (Pos)	62
2.6.5.3 Proof-of-Activity (Poa)	63
2.6.5.4 Proof-of-Burn	63
2.6.5.5 Proof-of-Capacity (Poc)	64
2.6.5.6 Proof-of-Authority (Poa)	64
2.7 E2e Encryption	65
2.7.1 How Does E2EE Differ From Other Types Of Encryption	65
2.7.2 Benefits Of Using E2EE	66
2.8 Discussion	67
2.9 Open Challenges and Future Research Opportunities	69
CHAPTER THREE: RESEARCH METHODOLOGY	73
3.1 Introduction	73
3.2 Overview Of The Proposed System	73

3.2.1 Overview Blockchain and EHR management based on patient centered control.....	73
3.2.2 A Background of the Proposed System	74
3.2.2.1 Actors	79
3.2.2.2 Electronic Health Record (EHR)	80
3.2.2.3 Blockchain	81
3.2.2.4 Website Portal	81
3.2.2.5 Smart Contract	82
3.2.2.6 Access control	82
3.2.2.7 Hybrid Encryption	84
3.2.2.7.1 Definition of ECC and AES	84
3.2.2.7.2 A Proposed Hybrid Encryption Approach ECC-AES	87
3.2.2.7.3 Implement the Hybrid encryption algorithm	89
3.2.2.8 IPFS for storing the health data	93
3.2.2 AI Model	104
3.2.2.1 Dataset	105
3.2.2.2 Pre-Processing	107
3.2.2.3 Classification	108
3.2.2.3.1 DenseNet	109
3.2.3 Tools Used to Implement The Proposed System	112
3.2.3.1 Blockchain Platform	112
3.2.3.2 AngulaJs	114
3.2.3.3 Truffle	114
3.2.3.4 Ganache	115
3.2.3.5 MetaMask	116
3.2.3.6 IPFS	116
3.2.3.7 Web3	117
CHAPTER FOUR: RESULT AND ANALYSIS	118
4.1 Introduction	118
4.2 Result of The Proposed System	118
4.2.1 Results of the proposed system	118
4.3 Analysis	130
4.3.1 Encryption and decryption time	131
4.3.2 IPFS uploading and downloading time	136
4.3.3 Classification Results of The Diseases Prediction	138
4.4 Discussion	139
CHAPTER FIVE: CONCLUSION AND RECOMMENDATION	144
5.1 Conclusion	144
5.2 Novelty of The Work	145
5.3 Thesis Contribution	146
5.4 Future Work	147

5.4.1 Big Data	147
5.4.2 Artificial Intelligence	147
5.4.3 Edge Computing	148
5.4.4 Internet of Medical Things (IOMT)	149
APPENDIX X: REFERENCES	150
APPENDIX Y: LIST OF PUBLISHED PAPERS	170



LIST OF TABLES

Table 2.1	Comparison of Different Types of Risk Prediction Models with Study Goals For Various Diseases	32
Table 2.2	The Advantage and Disadvantages of Different Types of AI Algorithms in The Risk Prediction Model.	34
Table 2.3	Contributions in the publications	41
Table 2.4	Contributions in the publications	41
Table 2.5	Research Comparison Used Blockchain and Ai Based Approaches To Secure EHR Systems.	43
Table 2.6	Inclusion And Exclusion Criterion	46
Table 3.1	Patient roles	79
Table 4.1	Execution time and cost without hybrid encryption ECC-AES.	130
Table 4.2	Execution time and cost with hybrid encryption ECC-AES	131
Table 4.3	shows key generation time for ECC-AES In different size	131
Table 4.4	Compare encryption time with other methods	132
Table 4.5	Compare the decryption time with other methods	132
Table 4.6	compare encryption and decryption time with different EHR size.	135
Table 4.7	Comparison of uploading and downloading time for encrypt/decrypt files.	138
Table 4.8	Compare our proposed system with the related on the literatures.	143

LIST OF FIGURES

Figure 1.1	Overview of the current system.	6
Figure 1.2	System Flowchart	7
Figure 2.1	Standard Block Structure	16
Figure 2.2	Diagram of the transaction flow in the blockchain	17
Figure 2.3	Structure of blockchain technology for hospitals	21
Figure 2.4	Blockchain service for PHR data	22
Figure 2.5	Number of articles according to publishers	39
Figure 2.6	PRISMA Chart	40
Figure 2.7	Taxonomy AI- Blockchain	54
Figure 3.1	Overview of the traditional and the proposed solution	75
Figure 3.2	Actors and rule-based access in the proposed system	80
Figure 3.3	Interactions in the system	83
Figure 3.4	Hybrid Encryption steps	87
Figure 3.5	Hybrid Decryption steps	88
Figure 3.6	ECC and AES algorithm	89
Figure 3.7	Sample Images in NIH ChestX-Ray8 Dataset	105
Figure 3.8	sample of images with bad quality in NIH ChestX-Ray8 dataset.	107
Figure 3.9	DenseNet architecture (Huang G, et al.2017)	111
Figure 3.10	Ethereum Logo	112
Figure 3.11	Front-end software	114
Figure 3.12	Truffle Logo	115
Figure 3.13	Ganache Logo	115
Figure 3.14	MetaMask Logo	116
Figure 3.15	IPFS Logo	117
Figure 4.1	Book appointment interface.	119
Figure 4.2	Doctor didn't have edit access.	120

Figure 4.3	Sharing record with view access.	120
Figure 4.4	Sharing record with edit access.	121
Figure 4.5	Manage shared EHR.	121
Figure 4.6	Update Shared record	122
Figure 4.7	Access and received signed EHR.	123
Figure 4.8	illustrated signed EHR.	124
Figure 4.9	uploaded x ray image and predict disease.	124
Figure 4.10	confirm estimated gas consumed to update EHR.	125
Figure 4.11	EHR stored in IPFS.	126
Figure 4.12	show how encrypt SK.	127
Figure 4.13	show how encrypt EHR.	127
Figure 4.14	show how Sign EHR.	128
Figure 4.15	show how decrypt SK.	128
Figure 4.16	show how verify sign.	129
Figure 4.17	show how decrypt EHR.	129
Figure 4.18	Compare time with different key size for encryption process with other methods.	133
Figure 4.19	Compare time with different key size for decryption process with other methods.	133
Figure 4.20	Encryption and decryption time consumption with different EHR size.	134
Figure 4.21	Compare encryption time with other works.	136
Figure 4.22	Compare decryption time with other works.	136
Figure 4.23	Time to upload and download different EHR size.	137
Figure 4.24	ROC for the performance of classification with 14 diseases.	139

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND OF STUDY

Medical and healthcare researchers emphasize the importance of their ability to collect and analyze multi-source data in order to identify potential community health hazards, provide case-specific therapies, and deliver focused medicine (Kumari A et al. 2018), which could promote informed clinical decision making and lead to improved patient care quality. This information can help to improve personal health information systems such as patient health records (PHR) and patient portals. Patients frequently do not have easy access to their historical data, while clinicians retain primary ownership.

Incorporating blockchain, AI, and other readily available technologies into a business's DNA is the key to success (Tanwar S et al. 2020). To enhance medical research and attain patient-centricity, the industry needs to use technology to produce user- and customer-centric interfaces and data-driven decisions for creative ways to data processing and improved outcomes (Campanella P et al. 2016, Siyal AA et al. 2019). For example, artificial intelligence (AI) could assist in identifying and prioritizing patients for drug monitoring and development, which is essential for regulated drug production and accelerated timeframes (Tanwar S et al. 2020). Using numerical drug design methodologies and AI, clinical trial data was evaluated for repurposing marketed pharmaceuticals, exploring the efficacy of medication formulations, and dose measurement (Tagde P et al. 2021). Blockchain facilitates the development of a system that creates and manages content blocks known as ledgers, incorporating secure and automated data analysis. All health-related information will be recorded and analyzed securely, allowing physicians, healthcare providers, and payers to receive rapid updates. However, storing massive records on the blockchain, such as complete electronic medical records or genetic data records, would be expensively inefficient due to the large computational resources required. This is a major drawback of blockchain technology, as it makes data queries within a blockchain difficult. Implementing AI algorithms into the blockchain, however, can help overcome this drawback (Tagde P et al.

2021). To comprehend health trends and patterns, artificial intelligence began to learn and reason like a clinician. It collects unstructured data from a variety of sources, including the patient, the radiologist, and the pictures. AI is also capable of conducting complex computational processes and evaluating enormous quantities of patient information fast. However, some doctors are still hesitant to use AI in healthcare, particularly in positions that may affect a patient's health, due to the significant capabilities that AI may bring, which have proved that it can execute numerous dynamic and cognitive processes faster than a person. The automobile sector has already demonstrated its capacity to utilize AI to produce autonomous automobiles. However, some businesses have already identified machine learning-based methods for detecting fraud and identifying financial dangers and demonstrating AI's maturity level (Shahnaz A et al. 2019).

The following section discusses the main terms and principles of intelligent technology in healthcare. We look at how intelligent technologies evolve and the security criteria for their implementation in the healthcare industry sector. In addition, the advent of modular IT systems has been observed since the implementation of healthcare provisions in the 1970s.

Healthcare 1.0 is the name given to this period. Because of a lack of funding, healthcare services were limited and not coordinated with digital systems during this period. On the other hand, bio-medical machines had not yet been built and did not integrate with networked electronic devices. Paper-based medications and reports were commonly used in healthcare institutions during this period, resulting in increased costs and time.

From 1991 to 2005, the Healthcare 2.0 period was observed. During this time, health and information technology were merged to form the foundations of today's healthcare systems. This process saw the introduction of automated monitoring, which provided doctors with imaging systems for assessing patients' health. Simultaneously, new user-enabled innovations in the healthcare sector started to evolve, coinciding with the advent of social media. Healthcare services began to build online communities to exchange information and expertise, store data on cloud servers, and provide mobile access to documentation and patient records, allowing both the provider and the patient to have constant access. During this time, critics shared their dissatisfaction with the misleading facts and the invasion of patients' privacy. Healthcare systems used networked electronic health management

practices combined with clinical imaging systems to help doctors get more reliable, accurate, and timely access to patient's data.

Healthcare 3.0 debuted simultaneously as Web, allowing users to customize how patient healthcare records were distributed. User interfaces became simpler and more tailored, allowing for more customized and optimized experiences. Electronic Healthcare Records (EHRs) and wearable and implantable devices were also introduced, allowing for real-time, ubiquitous monitoring of patients' healthcare. Similarly, EHR systems (Vora J et al. 2018) emerged that incorporated stand-alone non-networked systems, such as social media networks, to store patient's data.

Finally, the care period proliferated, inspired by the idea of Industry 4.0, in which Hi-tech and Hi-touch systems are implemented, using cloud computing, fog, and edge computing, big data analytics, AI, and machine learning to create blockchains that allow for real-time access to patient's clinical data (Tanwar S et al. 2020). The fundamental goal of this period is to improve virtualization, allowing for real-time personalized healthcare. The emphasis is now on teamwork, coherence, and integration, using AI technology to make healthcare more predictive and personalized.

By considering the above scenario, this paper aims to identify the potentiality of AI-blockchain to manage EHRs and show the challenges and future scopes. This systematic review explores research that offers conceptual solutions, experimental results, prototypes, and blockchain implementations for managing EHRs.

1.2 RESEARCH QUESTIONS

1. How can we design a blockchain framework that ensures the security and integrity of healthcare medical records while maintaining time efficiency?
2. What cryptographic techniques and consensus mechanisms can be employed to enhance the security aspects of the proposed blockchain framework?
3. How can artificial intelligence models effectively filter and mine metadata from big healthcare datasets for the purpose of diagnosis sharing and decision-making?

4. What are the key challenges and opportunities in integrating AI models into the healthcare metadata filtration process?
5. How can end-to-end encryption be implemented to provide patients with centralized control over their medical records while maintaining security and accessibility?
6. What are the usability and acceptance factors associated with a patient-centered approach to medical records management?
7. How does the performance of the developed system compare to existing systems proposed in the open literature in terms of security, efficiency, and usability?
8. What are the key performance metrics and evaluation criteria that should be considered when benchmarking the system?

1.3 PROBLEM STATEMENT

The problem statement can be summarized in the following points.

1. Lack of management and distributed data, where anyone can access the medical data, because it is readable by anyone without authorization.
2. Medical records in database are vulnerable and can be easily tampered with, altered, modified or deleted completely.
3. Processing, accessing and retrieving are time-consuming because it is based on a centralized database for saving medical data from patient medical records to diagnostics reports and doctor's prescriptions.
4. Medical data need to end2end encryption to ensure the security ,integrity and confidentiality.
5. The numbers of medical records are heterogeneous massive Bigdata and it has proven to be a challenge so far to have a one solution fits all to secure them.
6. The handling of metadata is another challenge that calls for emerging AI technologies to be applied together with blockchain solutions to secure the data and reduce cost.
7. all recent standards require decentralization, distributed access and metadata maximum use without patient rights infringements.

1.4 OBJECTIVES

1. To develop a secure time-efficient blockchain framework for healthcare medical record management system
2. To utilize AI models for bigdata metadata filtration, mining and diagnosis sharing decision-making process
3. To enhance distributive accessibility and security of patient's medical records using e2e encrypted patient-centered control of medical records management plans
4. To evaluate and benchmark the performance of the developed system against other systems proposed in the open literature.

1.5 MOTIVATION

Content organizations traditionally utilize cloud databases to consolidate various types of health information, such as electronic health records (EHRs), electronic medical records (EMRs), clinical images, patient health records (PHRs), and personal data such as body measurements and home-checking gadget information. It is important to note, however, that a centralized database presents a vulnerability to cyberattacks, which can compromise the security and privacy of EHRs (Madine et al., 2020). Additionally, stakeholders and healthcare providers encounter challenges in sharing health information due to differences in standards and formats.

Furthermore, if a patient's EHR is deleted from a hospital's database, the record is permanently lost, exacerbating the problem. Therefore, any proposed system must be tamper-proof to prevent unauthorized parties from accessing the information (Saidi et al., 2022). Another issue with current healthcare systems is that patients have limited control over their health records as they are managed by service providers (Makridakis et al., 2019). As the amount of healthcare data continues to increase, security and scalability have become major concerns. Figure 1.1 illustrates the current system architecture for managing health records.

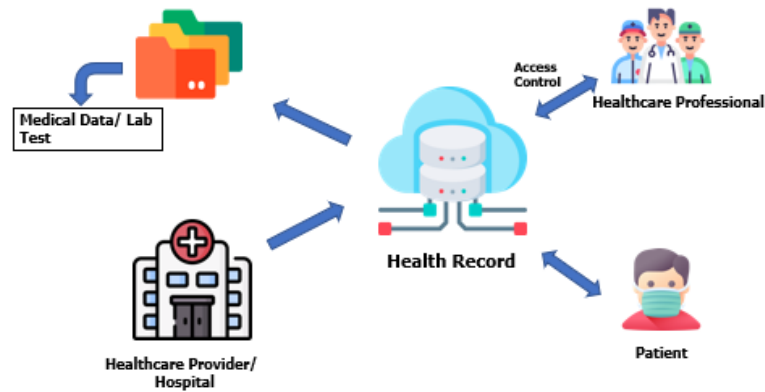


Figure 1.1. Overview of the current system.

1.6 RESEARCH SCOPE

The study in this thesis will involve the design and implementation of the proposed system in an actual testbed. No simulation studies will be considered. The system will be developed by using Ethereum blockchain platform with IPFS and Ganache. Truffle is a framework of this DApp, AngularJs as a front-end, and executed it on web3, back-end executed using Python. Furthermore, AI algorithms executed using Python and import all the required libraries to achieve our proposed system. However, it will be benchmarked against other systems reported in the open literature. The system will be evaluated in terms of its security, user-friendliness, distributive accessibility, time efficiency, and data analytics capabilities.

1.7 RESEARCH PHILSOPHY

The finding of this study will assume to provide a AI-Blockchain solution that can manage healthcare medical records from different heterogeneous sources, like IoT devices, ambulance records, EHR records, out-patient records, in-patient records, etc, with a high-security level to data and allow for better confidentiality. furthermore, the system aims by being designed in a distributive manner, to give more freedom to the patients themselves to control the level of accessibility and record management needed.

1.8 RESEARCH METHODOLOGY

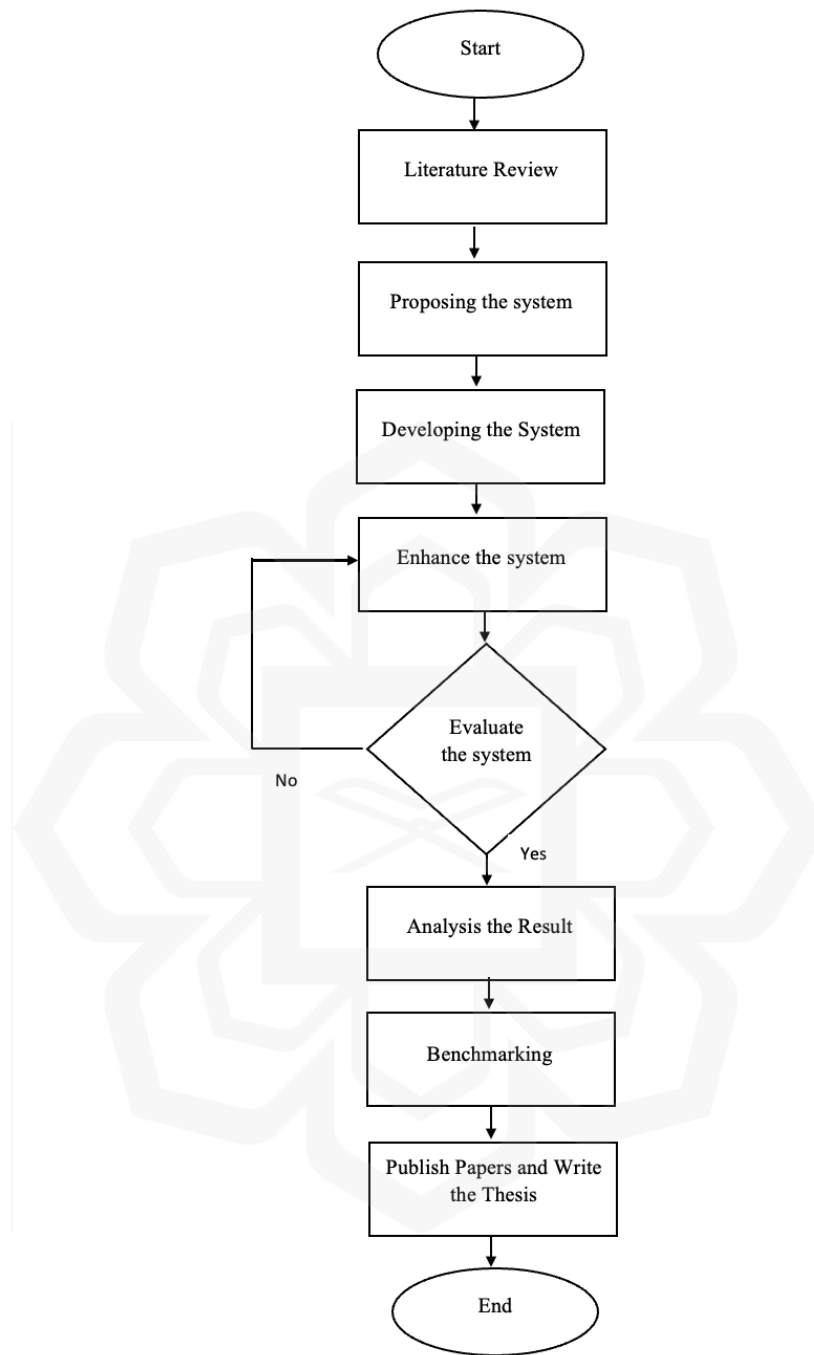
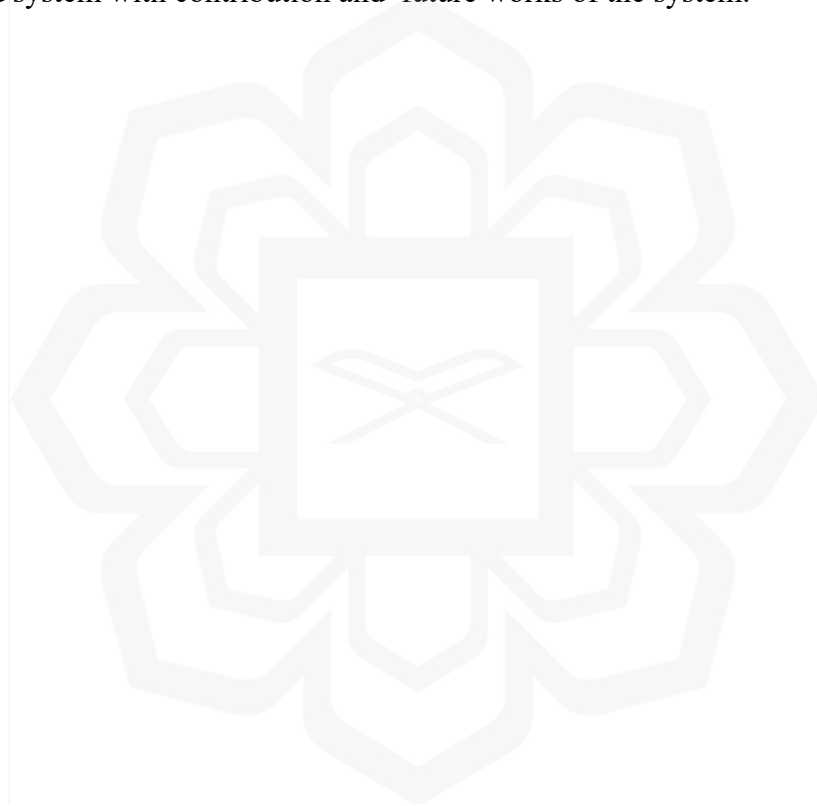


Figure 1.2 System Flowchart.

1.9 THESIS BREAKDOWN

In Chapter 1, the general idea of the project is demonstrated. The essential components such as background, problem statement, methodology, scope and organisation of report are

discussed under this chapter. Chapter 2 discusses the literature review that related to the proposed system which comprise of the compilation of published articles related to Blockchain and Artificial intelligence and E2E encryption. In Chapter 3, the methodology of the proposed system is described with methods and procedures that will be used to achieve the stated objectives of this system. For Chapter 4, this chapter contains the results and analysis the results by comparing with the recent related studies of this archiving blockchain and AI or encryption methods application. The concluding section, Chapter 5 summarises the findings of the system and also contains the summarize of novelty of the proposed system with contribution and future works of the system.



CHAPTER TWO

LITERATURE REVIEW

2.1 INTRODUCTION

Electronic health records (EHRs) typically contain sensitive information, including medical history, personal details such as age and weight, and laboratory test results. Therefore, it is crucial to ensure that this information remains secure and private. Hospitals in some countries, such as the United States, are subject to strict government scrutiny to ensure that patient data is protected (Ma S et al. 2021). However, deploying and implementing healthcare systems presents several challenges. As previously mentioned, centralized server models are vulnerable to single-point attack constraints and malicious insider assaults. Patients who store their data in these EHR systems lose control over their information, as they cannot determine who accesses it or for what purposes, leading to a potential violation of personal privacy. Malicious insiders could also leak the information to other organizations, resulting in consequences such as insurance coverage being denied based on leaked medical records. Meanwhile, sharing data is becoming increasingly important, especially as the population becomes more mobile. Shared data can improve medical service delivery by taking advantage of the interconnectedness between different healthcare organizations. However, overcoming the "Information and Resource Island" (information silo) due to privacy concerns and restrictions is challenging. Furthermore, information silos lead to data redundancy and bureaucracy.

To address these issues, the United States Congress enacted the Health Insurance Portability and Accountability Act (HIPAA) in 1996 (Nosowsky R and Giordano TJ. 2006). HIPAA established standards to protect the privacy and security of personal health information and implemented several programs to combat fraud and abuse in the healthcare system. The act includes five rules that help to safeguard patient privacy and improve healthcare delivery:

- The privacy rule. Regulations governing the use and dissemination of patient health information in the treatment and operations of healthcare organizations.

- The rule of Transactions and Code Sets. All health plans must uniformly engage in healthcare transactions.
- The rule of security. The security rule supplements privacy by limiting access to computer systems and preventing interception of communications via open networks.
- The Rule of Unique Identifiers. To secure patient personal information, only the National Provider Identifier (NPI) is used to identify covered entities in standard transactions.
- The Rule of Enforcement. For breaking HIPAA rules, there will be an investigation and fines.

ISO 27789 (Kubo et al. 2019) is another typical audit trail for EHRs that keeps personal health information auditable across systems and domains. A secure audit record must be created every time an operation is triggered by a system that complies with ISO 27789. As a result, a collaborative and open data-sharing system is essential, as it simplifies auditing and post-incident inquiry or forensics in the event of alleged misbehavior (e.g., data leakage). Forensic scholars also do highlight this concept (forensic-by-design) (Kubo et al. 2019, Davenport et al. 2019).

When the next generation of secure EHR systems has been generated, we should follow the next requirements based on the relevant standards listed above:

- Data accuracy and integrity: e.g., unauthorized data modification is not allowed and can be detected.
- Data security and privacy.
- Efficient data sharing mechanism (Feng Q et al. 2019).
- The patient control mechanism allows the patient control mechanism of EHRs (e.g., the patients will have control over their records and can get a notification if there is unauthorized access or loss of their data).

Data auditing and accountability (e.g., forensic by design) (Kubo et al. 2019, Davenport et al. 2019).

- The decentralization of power. In contrast to the centralized approach, blockchain does not require a semi-trusted third party.

- Safety and security. The blockchain-based decentralized system is resistant to a single point of failure and insider attacks.
- The use of a pseudonym. Each node is assigned a pseudonymous public address to safeguard its true identity.
- Impermanence. using the cryptographic hash function in one way, it will make the computationally hard to delete or change any records of any record of any block included in the chain.
- Independence. Patients have control over their data and can share it in a variety of ways thanks to the settings of special items in the smart contract
- Mechanism of motivation. Blockchain's incentive structure can encourage competitive institutions to collaborate and share information to advance medical services and research.
- Transparency. Can track every operation in the blockchain because every previous transaction is recorded in the chain.

Based on the following explanation, blockchain technology can be used to achieve the previously mentioned requirements.

A. BENEFITS OF USING BLOCKCHAIN AND ARTIFICIAL INTELLIGENCE.

The coronavirus epidemic can be dealt with in a variety of ways using blockchain and AI. There are many real-world applications for the blockchain that can be put to good use in the fight against the coronavirus outbreak. Blockchain could be used to monitor the spread of coronavirus infections around the world by installing blockchain network client software on users' mobile devices. One of the most important aspects of blockchain is its ability to protect user privacy, allowing early identification of epidemics while prohibiting the publication of user information. It also helps with epidemic and treatment management by making vaccine trials more efficient and transparent, as well as keeping track of all fundraising activities and donations. When it comes to combating the Coronavirus, AI has a range of approaches to help. AI may be used to identify viruses and anticipate how they will

spread by analyzing the accumulated knowledge of environmental factors (Tanwar et al. 2020), healthcare access, and the transmission method. By classifying coronavirus inside localized outbreaks of sickness, AI can help determine whether or not it is indeed there. Pneumonia, severe acute respiratory syndrome, and renal failure are all possible outcomes of coronavirus infections. For example, a genome-based neural network that has already been developed for personalized care can be very useful in managing these adverse events or symptoms caused by a coronavirus, particularly when virus impact is dependent on individual immunity and genome structure and no single treatment can effectively treat all symptoms at this time. AI may also be useful in speeding up the development of a new vaccination for novel coronaviruses (Tagde P et al. 2021). As a final application of AI, it may be possible to develop an automated model or correlation between medical records and results. Clinical protocols for coronavirus-like outbreaks could benefit from these models' quick identification of diagnostic and therapeutic options. A recent White House request to deploy AI to assist the US government in responding to the coronavirus pandemic (Vora J et al. 2018) is based on these prospective advantages.

Disintermediation is defined as the absence of a centralized authority that collects, processes, and validates data & models designed and shared. It enables a reduction in the time, error, and cost of process performance aimed at building and updating a predictive model that supports clinical practice and risk management. Transactions certified by the blockchain, and the data included within them are irreversible, in the sense that they cannot be changed or erased, ensuring their legitimacy while also strengthening the security of the system in which the activities take place (Kubo et al. 2019). Furthermore, the cryptographic system, the immutability of the data communicated across the network, and the lack of a centralized authority foster greater trust in the system, as the need to maintain this confidence among the parties involved in the process fades (Davenport et al. 2019).

B. THE HEALTHCARE SYSTEMS' SHORTCOMINGS

In the wake of the COVID-19 pandemic, current healthcare systems have come under scrutiny. Currently some existing healthcare systems may be overburdened by the COVID-19 outbreak. As of right now, there is no trustworthy data monitoring system in place

(Nguyen DC et al. 2021) to give key healthcare organizations the information they require about potential epidemics in real-time. In fact, most of the current coronavirus information comes from separate sources such as the public, hospitals, clinical labs with a large amount of inaccurate data without being monitored thoroughly. The use of unreliable information makes it challenging for potential outbreak identification and quarantine. Another limitation is the current time-consuming and in-accuracy coronavirus detection procedure that often takes several hours to complete the virus tests. This is unacceptable in light of the rapid spread of the coronavirus. It is critical to learn how to swiftly and accurately identify coronaviruses. Coronavirus data processing utilizing human-dependent medicinal software is exceedingly tough, especially when dealing with complex patterns and enormous volumes. Blockchain technology offers promising security solutions to aid in the fight against pandemics. Indeed, the blockchain creates immutable transaction ledgers for medical data sharing systems. More importantly, the combination of blockchain and smart contract technology eliminates the need for central servers to ensure fairness among transaction parties. Traceability and decentralization are two key characteristics of blockchain that are not found in other traditional security techniques. Furthermore, blockchain can provide reliable data analytics. Data collection is an important step in disease analytics. How to ensure the reliability of collected data during data collection is important for ensuring the high quality of disease data analytics (Pham QV et al. 2020). The use of incorrect data or untrustworthy database sources can lead to biased analytical results, which can have fatal consequences, such as incorrect diagnosis. Furthermore, in an emergency epidemic situation, many sources of contagious disease data are collected without protection from hospitals, the public, or the media, which can result in data modifications. These issues would undoubtedly affect the accuracy of the collected data, reducing the reliability of the analysis process. Because of its security, blockchain is in high demand in such contexts to ensure the reliability of collected data. Due to consensus mechanisms, blockchain also ensures the correct ordering of data records from data sources to destinations (e.g., hospitals or clinical labs), ensuring the high quality of data collection. These blockchain features would ensure accurate data collection and thus reliable disease analysis.

As the last point, there are privacy issues over the mass monitoring of the population to monitor the coronavirus. Healthcare organizations can monitor individuals' cell phones without a court order to prevent the spread of the COVID-19 coronavirus, for example (Mistry et al. 2020). However, human rights and privacy advocates have objected to the plan since it might potentially disclose citizens' private information, which could lead to major civil liberties abuses. To combat the spread of the coronavirus, real-time monitoring systems that protect user privacy are needed. As privacy become more of a concern, secret blockchain networks, that uses Privacy by Blockchain Design (PbBD) technologies to customize the level of privacy, are now gaining attention.

2.2 BACKGROUND OF BLOCKCHAIN

We briefly outline blockchain technology to assist readers in comprehending the remainder of the article. In the following subsections, we will cover the fundamental structure of blockchain technology to facilitate better grasping of the survey and the notion of blockchain.

2.2.1 Blockchain

A blockchain may be thought of as a decentralized public ledger that is accessible to all peers in a network where all committed, valid, and completed transactions are stored in a list or chain of blocks. The chain grows as new blocks are appended to it continuously. Blockchain technology employs a combination of two technologies: asymmetric cryptography and P2P distributed consensus to guarantee ledger consistency and user security. Hence, these time stamped blocks are linked together by a cryptographic hash (Feng Q et al. 2019). Typically, each block contains transaction records that have been verified by peers, often known as miners. The chain is continually lengthened, with each new block being added to the end. Each new block, on the other hand, contains a reference to the preceding block's header, which is essentially a cryptographic hash (e.g., SHA-256). the creation of each block has been with pseudonymity, transparency, and immutability (Lin C et al. 2020, Ma S et al. 2021)

A block is made up of the block header and the block body, defined below, as seen in Figure 1.

- Version: the cryptocurrency version number that indicates which set of block validation rules should be followed.
- Previous block hash: the hash value of the block before it.
- Time stamp: the current block's creation time is the timestamp.
- Nonce: to solve a PoW problem, miners alter a four-byte random field each time they hash the code.
- Hash target: new block's hash value must fall within a certain range before it is considered valid. Target hash is used in determining the difficulty of the input and can be adjusted in order to ensure that blocks are processed efficiently.
- Merkle Root: transactions in the block's body generate the Merkle tree root's hash value.

Transactions regularly are included in the block's body. Each leaf node of the Merkle tree represents a transaction, and every nonleafy node represents the hash value of the two concatenated child nodes that make up the leaf node. To validate the presence and integrity of a transaction, every node only needs to check the hash value of the two concatenated child nodes that make up the leaf node rather than the entire Merkle tree. There will be a new hash value generated in the top layer for any changes made to a transaction, which will result in one root hash. In addition to the block size, the maximum number of transactions per block is determined by the size of each transaction. Once the hash function is used, all blocks will be linked. Because data that has been validated cannot be modified or deleted in the blockchain, as new data comes in, it will be added to the linked blocks. Every change to the block will result in a new hash value (a new block) and a new link relationship based on this state. Immutability and security are fundamental features of blockchain technology.

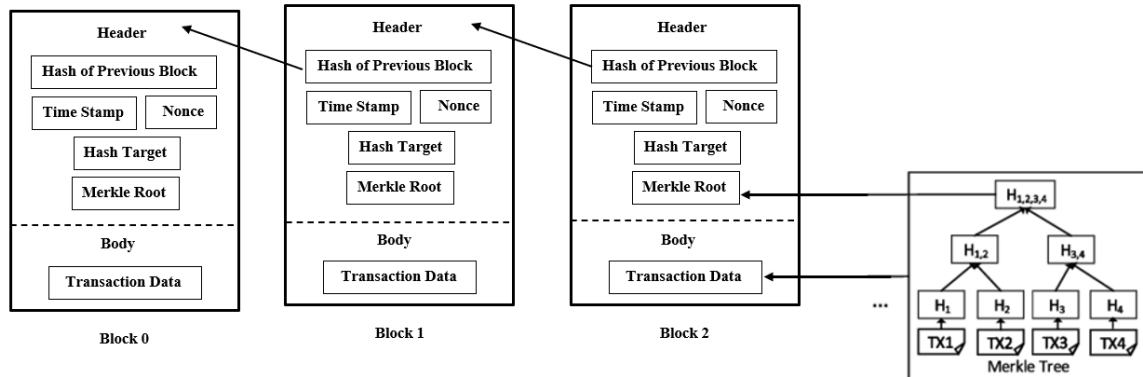


Figure 2.1 Standard Block Structure.

2.2.2 Digital Signature

For transaction authentication in an untrustworthy environment, asymmetric cryptography is often utilized (Feng Q et al. 2019). To send and verify the legitimacy of transactions, asymmetric cryptography is a key component of the Blockchain. In a P2P network, transactions are signed with the transaction initiator's private key before they are received. Most current blockchains use the elliptic curve digital signature technique (ECDSA) (Wang W et al. 2019).

When a transaction is requested or initiated, a block representing that transaction is generated and broadcasted to all adjacent nodes via the peer-to-peer (P2P) network, in which peers have equal Privileges. This block will be received by other nodes. The sender's public key is used to validate the legitimacy of the received block using specified block validation rules. If the block is genuine, it will be transmitted to other nodes until they have all verified it. If not, it will be discarded during the procedure. Only valid blocks can be added and stored in the blockchain network.

Figure 2.2 illustrates the process using coins, where Bob receives from Alice a specific number of coins. She initiates a transaction using her private key, which is then confirmed by the network. Anyone with access to Alice's public key can easily verify the transaction. In the second step, the P2P network disseminates the transaction to other nodes. In the third step, the transaction is verified by each node according to predetermined rules. Each validated

transaction will be grouped chronologically and added to a new block in step 4 after the miner solves the problem. Then, each node will update and back up the new block.

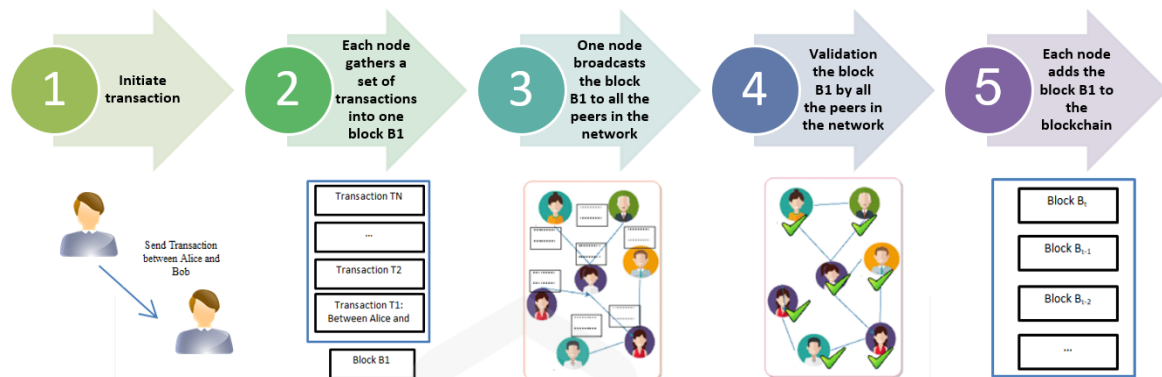


Figure 2.2 Diagram of the transaction flow in the blockchain.

2.2.3 Algorithms for Building Consensus

There is no one point of authority in the blockchain network. As a result, a fundamental issue is the Byzantine Generals Problem (Shostak R et al. 1982), a variant of which was created in the context of distributed networks in 1982. A gang of Byzantine generals is surrounding the city, and they have little chance to win the fight unless they all attack at the same time, the Byzantine Generals claims. There is a question as to whether or not there will be any traitors in a dispersed context. So they must make a choice: attack or retreat. It is the same challenge for the blockchain network.

To obtain a consensus protocol among all the distributed nodes before a new block can be attached to the blockchain, different protocols have been developed (Wang W et al 2019).

- **PoW (Proof of Work):** PoW is the name of Bitcoin's consensus algorithm (Proof of Work). Before receiving any rewards, a miner node with a certain level of computing (hashing) power must perform laborious task of mining to prove that he is not malicious (NN-A at S et al. 2017, GW-E project 2014). To find an eligible nonce value that is smaller than (or equal to) the target hash value, the node must continually perform hash calculations. It is difficult to generate a nonce, yet it is trivial for other nodes to check its validity. The task is

costly as a result of the numerous computations required (computational resources). If the blockchain network were to be attacked by a 51 percent attack (Li X et al. 2020), this would be an extreme case. A miner or a group of miners having more than 51% of the processing power can delay the generation of new blocks and create fraudulent records of transactions that benefit the attackers.

- Proof of Stake (PoS) Compared to PoW, PoS uses less power. It is widely believed that nodes with the highest stakes (such as cash) are less likely to attack the network (Bentov I et al. 2014). It's unfair to decide based on account balance because the wealthiest node is more likely to take over the network, making it a centralized one.

- Delegated Proof of Stake (DPoS) Similar to PoS, DPoS can also be used. The key distinction between DPoS and PoS is that the DPoS is democratically representative (Li C et al. 2019), whereas the PoS selection is based on all nodes. Stakeholders can elect delegates to decide who generates and validates new blocks. The fewer nodes that validate a block, the faster the transactions are confirmed by other nodes. In addition, dishonest representatives could be easily removed from office, making network maintenance simpler.

- Proof of Authority (PoA) is an efficient algorithm for achieving consensus network (Bentov I et al. 2014). Nodes with the ability to build new blocks are permitted. Each node must first undergo a pre-authentication process. On the other hand, this method produces a design that is centered by nature.

- Proof of Capacity (PoC) is a consensus mechanism that achieves consensus by utilizing available hard disc space rather than computational resources (Tschorsch F et al. 2017). With additional storage capacity, you may store more solutions, increasing the likelihood that a new block will be generated.

Rather than depending on a single consensus algorithm, an increasing trend is to combine many consensus algorithms to improve performance in a variety of applications.

2.2.4 Smart Contract

Smart contracts are self-executing programs that are implemented on the blockchain. They have been employed in a variety of areas, including finance, healthcare, and government.

Such a system can achieve complex programmable functionality by delivering a contract-invoking transaction to the appropriate contract address. The smart contract will execute the secure container's predefined terms automatically. Ethereum is the first open-source blockchain platform that includes Turing-complete smart contract languages, enabling developers to create any decentralized application (Dapps) they desire. Dapps, or decentralized applications, refer to programs built on the blockchain technology that facilitate communication between patients and doctors without relying on third-party intermediaries, except for the Ethereum network. Through Dapps, patients can exercise greater control over their medical records, as stated by Houtan et al. (2020).

2.3 BLOCKCHAIN APPLICATIONS IN HEALTH RECORDS SYSTEM

2.3.1 Data Management in Electronic Medical Records

Blockchain technology has gained interest in healthcare and pilot programs have been launched globally. Booz Allen Hamilton Consulting developed and launched a blockchain-based pilot platform in the United States last year, which is now being implemented at four large hospitals. They have also been tasked with advising the Food and Drug Administration's Office of Translational Sciences on the application of blockchain in healthcare data management (Figure 2.3). The pilot project uses Ethereum to regulate data access via virtual private networks and employs IPFS to decrease data replication by utilizing off-chain cloud components and cryptographic techniques to facilitate user sharing. This ensures encryption and data privacy for users (Cyran MA et al. 2018).

2.3.2 Blockchain and Data Protection In Healthcare

A connection exists between blockchain technology, and the General Data Protection Regulation (GDPR) implemented in the European Union. GDPR, on the other hand, places a high value on the inclusion of blockchain technology (when the data can be portable, for

data traceability, legal access auditing). Based on the information previously provided, a variety of issues can be experienced (the actual control may be weakened when the technical implementation of the smart contract over data). Dynamic consent management is a solution that is fully compliant with GDPR consent requirements. Enterprise blockchains, also known as private blockchains, are also believed to be suitable for GDPR compliance because they allow transactions involving digital records to be modified and removed by network owners or authorities using a particular type of consensus algorithm. These private blockchains are typically controlled by a single entity or organization, and access is limited to individuals or companies who meet certain predetermined criteria or restrictions. (Cédric Villani et al. 2021, Cyran MA et al. 2015, Lima C, 2018).

The way a firm handles its private web apps will be comparable to the way it handles its public web applications.

Their technology can cater to various use cases such as government agencies, owners of public health data, and healthcare reimbursement companies. In particular, private blockchains are expected to have a significant impact on healthcare policy and management in the future. Moreover, Novartis is leading the IMI (Innovative Medicine Initiative) Pilot project "Blockchain-Enabled Healthcare" under the European Commission's Research & Innovation Program to explore blockchain possibilities in healthcare. It hopes to capitalize on established standards like Ethereum while simultaneously developing supplementary standards as needed. The emphasis is on those who can facilitate programs that would directly benefit patients (Dimitrov D V. 2019).

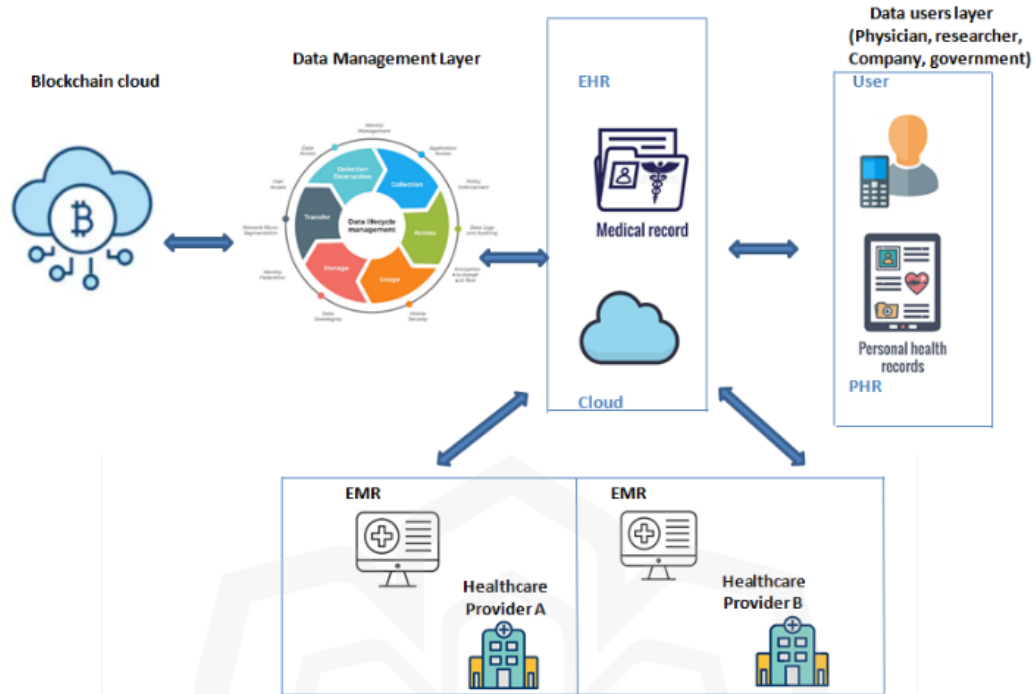


Figure 2.3 Structure of blockchain technology for hospitals.

2.3.3 Personal Health Record (PHR) Data Management on The Blockchain

Personal health records (PHR) have lately begun to be built utilizing data from sensors, which can be wearable or medical Internet of Things devices. A variety of stakeholders, including patients, doctors, pharmaceutical specialists, and payers will benefit from real-time AI-powered healthcare analytics (P. Zhang et al. 2018, Salah K et al. 2019). A key data source for blockchain service providers is the complete PHR service trajectory, which is becoming increasingly important. (See Figure 2.4).

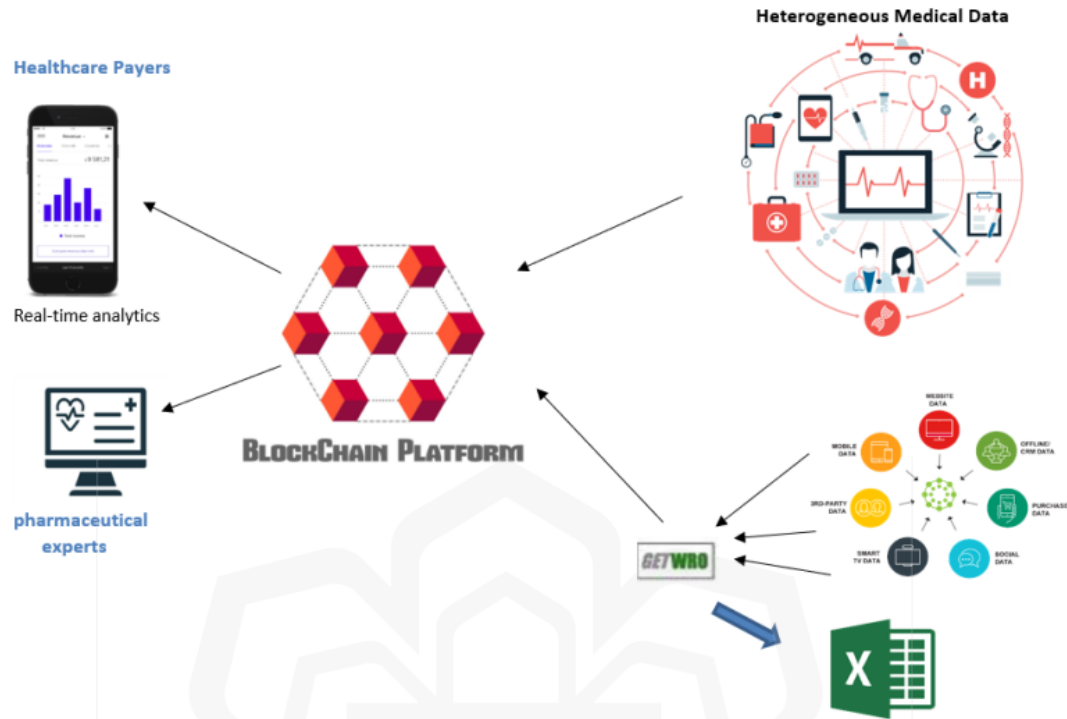


Figure 2.4 Blockchain service for PHR data.

Blockchain technology is also a feasible solution for managing personal electronic health records. Patients may be reimbursed with tokens for providing health data with physicians and research collaborators through the use of so-called "smart contracts," which are electronic contracts that exchange data between parties. Using blockchains to tokenize data, Health Wizz, for example, is experimenting with a blockchain- and Fast Healthcare Interoperability Resources (FHIR)-enabled EHR aggregator mobile app that will allow patient groups to aggregate and organize their medical records in a safe manner, as well as exchange, donate and/or swap their medical records (P. Zhang et al. 2018). To facilitate improved coordination between healthcare institutions and caregivers for a higher level of care, the goal is to make it as simple as managing online bank accounts to manage one's health information.

In the context of an EHR blockchain business (P. Zhang et al. 2018), medical chain allows a variety of healthcare agents to apply for and obtain authorization to view and communicate with patients' medical records. These agents include physicians, hospitals, laboratories, pharmacies, and insurers. In the medical chain's distributed ledger, each interaction is recorded as a transaction, and the ledger is auditable, open, and stable at all times.

2.4 OVERVIEW OF ARTIFICIAL INTELLIGENCE IN HEALTH RECORDS MANAGEMENT SYSTEM

AI systems in health care are often built upon supervised or unsupervised methods. In supervised learning, labeled data with regard to output or reaction of interest is used to train machines to predict these classifications using a set of predictors or inputs (Gareth J et al. 2021). The unsupervised method, on the contrary, does not use labeled data nor does it anticipate a result or reaction. Instead, it finds patterns and correlations in the data to classify variables or observations into related categories (Gareth J et al. 2021). The majority of existing machine learning structures in the health care industry, some of which build electronic phenotyping algorithms, employ supervised learning methods (Jiang F et al. 2021). In this part, we provide a quick review of a few machine learning approaches widely utilized to categorize clinical results from electronic health records, including random forests and support vector machines, as well as supervised and unsupervised models for deep learning and neural networks (Jiang F et al. 2021, LeCun Y et al. 2015, Resta M et al. 2018).

Support vector machines determine the optimum disconnected hyperplane in the covariate space between observations of different outcome groups for the identification of variables (Jiang F et al. 2021). The best hyperplane is defined as the one with the greatest margin or distance separation from the closest observation to either of its sides from distinct outcome groups, which essentially is referred to as ‘support vectors’ (Kaye J et al. 2015). On the other hand, random forests are regulation-based batch classifiers that identify inputs by averaging estimates through a group of decision trees models (Dimitrov D V, 2019).

Every tree in the random forest classifier would be trained with a sample of bootstrap data points, with the sample split at each node on the most descriptive among a randomized subset of the potential predictors (Jiang F et al. 2021). In recent years, deep learning models, as well as neural networks have grown in popularity, particularly for their application in diagnostic imaging and forecasting activities (Jiang F et al. 2021). When they are used for the supervised learning approach, the models may be regarded as progressively complicated extensions of the classic regression paradigm (Jiang F et al. 2021). A conventional logistic regression model consists of an output and input layer, with a node in the input layer for every parameter

and a collection of relation weights linking the input nodes to the coefficients or the output node. To generate the final output of the model, the output node undertakes the total from each parameter multiplied by the matching relation weight, which is known as input nodes' weighted sum, and runs it through the activation function, or the logistic function in this example. Neural networks extend this structure by including a concealed layer between the output and input layers, in which the nodes would allow the neural network to simulate more complicated and non-linear relationships between the response and predictor factors (Gareth J et al. 2021).

Following that, deep learning models improve this approach by incorporating several hidden layers between the output and input layers to detect even more subtleties in the data (Resta M et al. 2018).

All of these algorithms are designed in such a way that they can autonomously simulate sophisticated interactions and relationships in datapoints with priori constraints from the investigator and with little assumptions. These algorithms, however, can be harder to decipher, are prone to overfitting, and frequently need a large quantity of training data to provide appropriate results (Beam AL et al. 2018). To both justify and optimize the application of machine learning for the categorization of health responses from EHR data, researchers ought to consider when these techniques are most appropriate and for which tasks, they should be employed (Raghupathi W et al. 2021).

2.4.1 The Challenges of Using AI in Health Records System

In this section, we discuss the issues surrounding the use of AI for health record systems such as the portability and transparency of these algorithms, as well as the requirement of the training sizes necessary for satisfactory productivity.

2.4.1.1 Transparency

The issue of inadequate transparency related to elaborate algorithms of machine learning such as deep learning creates hurdles to their application in phenotyping tasks, especially when the stakes are significant and end-user confidence is essential. Clinicians, for example,

would prefer the algorithms to supplement or enhance their expertise as opposed to merely dictating their decision-making process (Rundo et al. 2020). Regulatory authorities, on the other hand, require algorithms to be decipherable for transparency reasons because their classification system may have substantial legal or financial ramifications (Gehrmann S et al. 2018). As a result, improving the interpretability of such "black box" models is crucial.

The results from previous research that employed a recently established approach in interpreting deep learning model predictions were outstanding (Gehrmann S et al. 2018, 33]. Researchers in (Gehrmann S et al. 2018) applied a modified variant of saliency, which is dubbed 'saliency' (Omar IA et al. 2021), to classify the most appropriate terms from clinical material and were subsequently utilized for prediction purposes by convolutional neural networks. Based on the authors, clinicians would assess these terminologies as more descriptive and applicable to the desired trait than the most crucial characteristics determined using a more standard definition of extraction-based NLP method. (Rajpurkar P et al. 2017) created heatmaps by using mappings of class activation in agreement with radiologists' assessments, representing the most significant portions of chest X-ray images applied by their deep neural network for the prediction of chest diseases (Zhou B et al. 2021). Such initiatives to improve the transparency and interpretability of complicated machine learning models strengthen the trust and confidence of physicians and other end users in these technologies, hence encouraging the number of uses.

2.4.1.2 Transportability

Due to privacy concerns and administrative constraints, a lot of electronic phenotyping research was conducted in a single-site environment (Ford E et al. 2021, Shivade C et al. 2021, Carrell DS et al. 2017, Kirby JC, et al. 2017). It was worth noting that there is a rising interest in exchanging the algorithms among researchers and healthcare bodies to improve their versatility, provided ample time and resources for the development of phenotyping algorithms were given (Kirby JC et al. 2016). Initiatives such as the Phenotype Knowledgebase (PheKB), an online platform meant to help researchers build, share, and validate electronic phenotyping algorithms, demonstrate that progress is being made in this area (Kirby JC et al. 2016). However, few phenotyping algorithms also have been customized

for applications in various settings, especially those involving machine learning (Wang W et al. 2019). To build scalable phenotyping algorithms, they should be externally verified to determine their portability, and then modified, if needed, to account for idiosyncrasies of site-specific. This "validation-adaptation" strategy is especially useful for phenotyping algorithms that use NLP systems even though it could be extremely laborious and work-intensive (Carrell DS et al. 2017). Since these versatile methods are vulnerable to overfitting, it is especially essential to validate phenotyping algorithms that implement machine learning externally before implementing them in different settings (Foster KR et al. 2014). If additional fine-tuning is required, for example, to model relationships differently or to detect new acronyms in clinical documentation, machine learning algorithms may take up less human work to be retrained than manual-engineered algorithms. In another instance, deep learning NLP systems would seldomly utilize manually supplied feedback and maybe quickly retrained to new datasets (Gehrmann S et al. 2018). Deep learning models that are usually employed to classify health responses from imaging procedures might theoretically be considerably more portable than those utilized for NLP tasks because of the smaller degree of between-site heterogeneity in medical images compared to clinical narratives.

2.4.1.3 Training Size

To achieve optimal efficiency, machine learning algorithms, particularly deep learning frameworks, would necessitate a significant quantity of labeled training data (Asperti A et al. 2018). In this regard, (Rajpurkar P et al. 2017) used the CheXNet algorithm and was trained on over 100,000 labeled images, which subsequently produced expert-level results. Many researchers, on the other hand, do not have such privilege given due to the limitation in time or resources for data annotation (Asperti A et al. 2018), or probably due to the constraint in the number of cases or rare health ramifications in the EHR system. In addition, as previously mentioned, pooling labeled data across different locations may not be a viable option owing to privacy concerns and administrative hurdles (Shivade C et al. 2021). However, innovative solutions such as image data enhancement and active learning by successively picking the most insightful instances for training can assist in minimizing the portion of training data required to obtain satisfactory performance (Wong SC et al. 2016, Kemp R et al. 2017). In

(Asperti A et al. 2018), for example, the annotated samples required to obtain an AUC of 0.95 was lowered by 68% when active learning was paired with support vector machines to create an electronic phenotyping approach for rheumatoid arthritis.

2.4.2 AI Algorithm in Healthcare Systems

This article brings machine learning and data mining together for a joint discussion because both disciplines are based on data science and frequently cross (Lorbieski R et al. 2018). However, there are a few fundamental differences between data mining and machine learning. The study of methods that can extract information automatically is known as machine learning (Lorbieski R et al. 2018). Forecasting future events requires two sets of data (training data and test data). On the other hand, data mining is an iterative process of uncovering various types of novel and useful patterns in data.

Data mining can employ machine learning, but it can also use other techniques besides or in addition to machine learning to identify new patterns. Machine learning and data mining technologies are employed mainly in the healthcare industry to extract knowledge from vast amounts of electronic health data. Machine learning and data mining approaches were included in the analysis in (Kavakiotis I et al. 2017) because they use similar mechanisms for disease prediction and are frequently discussed together in the literature.

2.4.2.1 Supervised Algorithm

2.4.2.1.1 Artificial Neural Network (ANN)

Artificial neural networks (ANNs) were first proposed by McCulloch and Pitts (McCulloch WS et al. 1943) and popularized in the 1980s by (Rumelhart DE et al. 1986). They can handle a range of categorization issues. The word "neural" in their name implies brain-inspired systems designed to mimic how human brains learn categories. ANNs were created to mimic the way the human brain works, in which a vast number of neurons are coupled to one another via many axon junctions. Neuron connections can be strengthened or decreased by reinforcing labeled training data, just as they can be in biological learning. A weighted matrix can be used to represent these neuronal connections. This matrix is referred to as a layer,

similar to the cortical layers in the brain. The training data used in ANNs serves as a form of 'biological learning' for people. In an ANN framework, there can be one or more hidden layers in addition to the input and output layers. ANNs are taught to generate an output from a set of input variables.

Several ANN research focused on the survival prediction problem were found in the literature. However, a few research relying on electronic health data were found.

Deep learning is a subfield of machine learning that deals with ANN-inspired algorithms (Schmidhuber J. et al. 2015). These algorithms have been utilized to model illness symptoms and hazards in recent years. Liu et al. (2014) created a deep learning-based technique for early identification of Alzheimer's disease and mild cognitive impairment in 2014. Neuroimaging data from the Alzheimer's Disease Neuroimaging Initiative database was used.

To get around the bottleneck, they used stacked auto-encoders. Cheng et al. (2016) suggested a method for phenotyping patient electronic health records (EHRs) using deep learning. Each patient's EHR was initially represented as a temporal matrix, with time on one axis and events on the other. The researchers built a four-layer convolution neural network (CNN) to extract phenotypes and forecast risk. (Zhang J et al. 2017) recently presented Heterogeneous Convolution Neural Network (HCNN), a new predictive learning model representing EHRs as graphs with heterogeneous properties such as diagnosis, procedures, and medicines. They used this information to create a new risk prediction model for numerous comorbid conditions.

2.4.2.1.2 Support Vector Machine (SVM)

SVM is a popular supervised learning approach for classifying linear and non-linear data. SVMs transform the input vector into a higher-dimensional feature space and find the hyperplane that divides the data points into two groups. An SVM may perform classification tasks by increasing the marginal distance between two classes while reducing classification errors. To determine the marginal distance for a given class, we must find the distance between the decision hyperplane and the closest instance of that class (Hossain ME et al. 2021). In order to accomplish this, each data point is initially represented as a coordinate in

an n-dimensional space, where n is the number of features. The hyperplane that maximizes the distance between the two classes is then located to complete the classification process. This technique, which has been applied in bioinformatics and healthcare, involves identifying the optimal decision boundary for separating two classes of data points (Hossain ME et al. 2021).

2.4.2.1.3 Decision Tree Random Forest

A decision tree (DT) is a sophisticated and deterministic data structure that looks like a tree, with internal nodes representing input variables or attributes and leaves representing decision outcomes. All nodes and their accompanying leaves are used to create a plan to attain a categorization goal. The leaves of a DT tree are on the last level of the relevant branch, and the nodes can be organized in more than one level. The root node is the tree's first node. It's similar to a flowchart in which each non-leaf symbolizes a test on a single property, each branch denotes the test's outcome, and each leaf indicates the class label. Many academics in the healthcare sector use decision trees extensively. For example, a decision tree-based prognostic approach was suggested to quantify disease recurrence and predict survival in breast cancer patients (Hossain ME et al. 2021). The model was developed for predicting breast cancer survival using two machine learning techniques (ANN and decision tree) and one statistical approach (logistic regression). They used the SEER breast cancer database, regarded as one of the few population-based data repositories for evaluating cancer care quality.

Random forest (Hossain ME et al. 2021) is an ensemble classifier made up of many decision trees. Individual trees represent the output of the classes. Among the machine learning-based algorithms, it is one of the most accurate. The method in (Hossain ME et al. 2021) combines Breiman's "Bagging" idea and the random selection of features to create a collection of decision trees with a controlled variation. In (Kavakiotis I et al. 2017), researchers suggested a classifier based on the random forest algorithm to estimate illness risk among individuals. The Healthcare Cost and Utilization Project (HCUP) dataset was used in their research. The work in (Rallapalli S et al. 2017) have developed a diabetes risk prediction model using a scalable random forest classification algorithm based on administrative data.

2.4.2.2 Unsupervised Algorithm

2.4.2.2.1 Association Analysis

Association analysis has been frequently utilized in data mining and machine learning literature for prediction because it can extract hidden and relevant information from huge datasets (Hossain ME et al. 2021). This function generates a collection of dataset item association rules (Rallapalli S et al. 2017). Power of association is an implication statement with $X \rightarrow Y$, where X and Y are disjoint item sets (i.e., $X \cap Y = \emptyset$). It means that the existence of X things in current transactions may result in one or more Y items appearing in future transactions. As a result, association analysis has been widely utilized with market basket data to forecast retail sales behavior, where each object reflects a customer's purchase (Hossain ME et al. 2021).

If an item is related to disease and the item set is specified as the patient's set of conditions until now, this method can be applied to the medical context to predict future disease risk.

The Hierarchical Association Rule Model (HARM) was introduced in (McCormick T et al. 2011) to predict illness risk from medical data using association analysis and a Bayesian estimate. First, a set of association rules is developed utilizing association analysis methods in this modeling technique. Then, using Bayesian estimation, these association rules are ranked. HARM can anticipate a patient's likely future medical issues based on her previous and present history of reported ailments, assuming that each patient regularly consults healthcare professional.

2.4.2.3 Network Approach

A network can be represented as a graph, which is made up of nodes (also known as vertices or actors) and edges (also known as ties or links). Edges represent the relationships between things, while nodes represent the entities themselves. Many scientific problems can be represented as graphs and modeled as networks. Many graphs theory approaches and algorithms for analyzing various problems, including disease prediction in the healthcare area, can be found in the literature.

Many statistical and data mining methods for predicting disease risk from healthcare data do not explicitly take into account the link between diseases and symptoms. Chronic and non-communicable diseases, on the other hand, do not arise in isolation (Hossain ME et al. 2021). They frequently share a risk factor, which might be genetic, environmental, or behavioral in nature.

These risk factors have a synergistic influence on health outcomes, which makes it difficult to forecast if they are studied separately. A network method may be more applicable in this scenario. Statistical methods are also used in a network-based approach. Another comparable approach is Social Network Analysis (SNA), which is built on a solid theoretical foundation drawn from network and graph theories. SNA is the study of the pattern of relationships among network entities, such as a group of people, departments, or organizations, as the name suggests. If the elements in the dataset have a lot of relationships between them, SNA can be especially useful. In a healthcare setting, for example, clinicians frequently need to confer among themselves about a patient's illness diagnosis. Patients are additionally cared for by pharmacists, nurses, and medical technicians. As a result, the recordings of these dialogues are bound by a network structure.

Each sort of entity participating in the healthcare data is represented as a node to describe the health care infrastructure as a social network. Edges linking the corresponding node pairs represent relationships between entities. SNA has been utilized to better analyze physician-patient partnerships as well as collaborations throughout a hospital network. Uddin et al. (2015) suggested an SNA framework to analyze the process of collaboration (amongst physicians) and coordination (between hospitals).

SNA was created with the intent of being utilized in the social sciences, but it is now frequently employed in medicine and public health. Each sort of entity participating in the healthcare data is represented as a node in the health care infrastructure's social network representation. Edges linking the corresponding node pairs are used to represent relationships between entities. SNA has been used to better understand physician-patient partnerships as well as hospital-to-hospital collaborations. Uddin et al. (2015) presented an SNA framework to describe the process of physician collaboration and coordination, for example (between hospitals). Their suggested framework looked at a patient-centric care coordination network,

a hospital-rehab coordination network, and a physician collaboration network, all using centrality theories. In the healthcare domain, for example, in obesity research, SNA is utilized to understand research trends and map knowledge structures (Uddin S et al. 2015, Khan A et al. 2016).

Large population-level studies aimed at understanding the nature of comorbidities (Khan A et al. 2018) and forecasting the likelihood of comorbid chronic diseases have a lot of potential with electronic health data. (Khan A et al. 2016) established a novel strategy in which they used graph theory and social network analysis methodologies to analyze and comprehend chronic disease progression using electronic health data. Their main goal was to forecast the likelihood of developing a chronic disease in new patients by modeling the health trajectory of chronic disease patients. The data was gathered from hospital admission and discharge records. The diagnoses of the patients (in ICD-10 Australian Modification format), as well as several socio-demographic characteristics, were taken into account. They created a baseline network based on the diagnosis data to better comprehend and reflect the health trajectory of chronic disease patients. Later, to better understand the comorbidities associated with type 2 diabetes, this approach was expanded and used. They proposed the concept of a 'comorbidity network,' which may be utilized to construct a model for predicting chronic illness risk (Khan A et al. 2019, Kang E et al. 2020).

Table 2.1 Comparison of Different Types of Risk Prediction Models with Study Goals For Various Diseases.

Risk prediction model	Diseases Name	Goals	Reference
Artificial neural network (Supervised)	Multiple cancer diseases	Using administrative and registry data, propose a machine learning model for cancer survival prediction.	(Gupta S et al. 2014)
	Pancreatic cancer	Using a boosting method and healthcare administrative data, propose a model for	(Velez-Serrano JF, 2017)

		predicting in-hospital mortality after pancreatic resection in pancreatic cancer patients.	
	Acute coronary syndrome	A significant volume of EHR data was used to stratify clinical risk and death for individuals with acute coronary syndrome.	(Huang Z et al. 2018)
	Heart failure	To offer an EHR-based architecture for heart failure prediction that is both effective and reliable.	(Jin B et al. 2018)
	Alzheimer's disease	Develop a deep learning-based approach for early detection of Alzheimer's disease and Mild Cognitive Impairment.	(Liu S et al. 2014)
	Generic	Propose a deep learning method for phenotyping patients' electronic health records (EHR)	(Cheng Y et al. 2016)
	Multiple chronic diseases	The goal is to create a new risk prediction model for comorbid disorders.	(Zhang J et al. 2017)
Support vector machine (Supervised)	Breast cancer	Using a hybrid SVM method, propose a predictive model for breast cancer diagnosis.	(Hossain ME et al. 2021)
	Cardiovascular	To create a system that analyses heart valve disease using a genetic SVM classifier.	(Rallapalli S et al. 2017)
	Cardiovascular	To create a model for predicting heart failure patients' 30-day readmission.	(Uddin S et al. 2015)
	Diabetes	Using a scalable random forest classification technique, create a model for predicting diabetes risk.	(Xu W et al. 2017)
	Coronary artery disease	Implement and analyze a set of supervised learning	(Forsen H et al. 2017)

		approaches for coronary artery disease prediction in a systematic way.	
Association analysis (Unsupervised)	Multiple diseases	Using electronic healthcare data, offer a method for forecasting disease risk in healthcare research.	(Uddin S et al. 2021)
Network Approach	Generic	To offer a SNA framework for analyzing the performance of physician collaboration and coordination (between hospitals).	(Uddin S et al. 2015)
	Generic	To determine the health trajectory of chronic disease patients and estimate the probability of new disease development.	(Khan A et al. 2016)
	Diabetes	Using graph theory and social network analysis methodologies provides a research framework for understanding and visualizing the evolution of type 2 diabetes.	(Wong SC et al. 2016, Xu W et al. 2017)

Table 2.2 The Advantage and Disadvantages of Different Types of AI Algorithms in The Risk Prediction Model.

Risk Prediction Model	Advantage	Disadvantage
Artificial Neural Network (ANN)	<ul style="list-style-type: none"> - When the relationships between variables are nonlinear and complicated, it is appropriate for predicting outcomes. - Requires less formal statistical training, and many training techniques for this methodology are available in the literature. - Can be used to solve both classification and regression issues. 	<ul style="list-style-type: none"> - It is referred regarded as a "black box" technology because the user is unable to see the exact decision-making process. - Training the network for a difficult classification task takes a long time with this technique. - Pre-processing of predictor variables is required.

<p>Support Vector Machine (SVM)</p>	<ul style="list-style-type: none"> - It introduces the kernel, which allows for non-linear transformation. - The ability to manage a large number of feature spaces. - In SVM, the risk of overfitting is lower. - Even unstructured and semi-structured data, such as words and photos, works well. 	<ul style="list-style-type: none"> - SVMs will not work as a classifier if the points on the boundaries are not informative owing to noise. - Larger, more complicated datasets will take longer to train. - The final model, variable weights, and individual impact are difficult to understand and interpret.
<p>Decision Tree (DT)</p>	<ul style="list-style-type: none"> - Easy to comprehend and interpret. - Requires minimal data preparation and can handle a variety of data formats, including numeric, nominal, and categorical information. - It is capable of producing robust classifiers that can be validated using statistical tests. 	<ul style="list-style-type: none"> -Classes must mutually exclude one another. - The final decision tree is determined by the order in which variables or attributes are chosen. - They don't perform as well as other classifiers (e.g., Artificial Neural Networks) (Gupta S et al. 2014) - When the needed value for the ancestor variable or attribute is absent, it is impossible to select which branch to choose.
<p>Random Forest</p>	<ul style="list-style-type: none"> - When compared to decision trees, random forest has a lesser likelihood of overfitting training data. - Produce less variance than decision trees since a random forest takes the average value of its constituent decision trees' results. - Random forests are almost always more accurate than decision trees. - It works well with huge datasets. - It can estimate which factors or attributes are most essential in classification. 	<ul style="list-style-type: none"> - The number of decision trees in the random forest must be defined. - When estimating variable importance, it favors variables or qualities that can take a large number of alternative values. - Overfitting is a common occurrence.

Association Analysis	<ul style="list-style-type: none"> - When diseases have a lot of comorbidities, it can forecast risk. - It can mine massive databases for interesting hidden relationships. 	<ul style="list-style-type: none"> - The methods utilized contain an excessive number of parameters. - The derived rules may be excessively complex and difficult to comprehend.
Network Approach	<ul style="list-style-type: none"> - It can make clinical decision-making more efficient and effective. - It can disclose the intricate relationships that exist between diseases, patients, and clinicians. 	<ul style="list-style-type: none"> - Traditional network models lack the longitudinal and spatial dimensions necessary to predict illness risks. - When compared to single-attribute networks, healthcare networks are far more complex.

2.5 MANAGING EHR USING AI AND BLOCKCHAIN

Machine learning can aid in the optimization of healthcare systems and the provision of intelligent services. How to safely store, exchange, and train sensitive datasets is a major difficulty for practical machine learning systems. Machine learning and blockchain are increasingly being combined to improve the security and privacy of datasets (Zheng X et al. 2018 , Lee SH et al. 2018).

Federated learning is a machine learning technique that is carried out over numerous computing nodes with the confidentiality and privacy of sensitive data protected throughout data sharing as a precondition. By exchanging encrypted datasets, different medical organizations can collaborate to create high-accuracy prediction models. To establish accountability and reliable cooperation, blockchain as a regulator can record associated training transactions in an immutable and transparent manner. Medical organizations and researchers will be more ready to share encrypted datasets to advance medical treatment and public health in this circumstance.

The security of data input is ensured by blockchain as a dependable backbone for machine learning algorithms. The first challenge raised by (Yaji S et al. 2018) is the sharing of huge datasets across different applications and domains. In reality, however, homo-morphic encryption has a substantial computational expense. Perhaps sensitive data can be encrypted in the future without affecting machine learning for intelligent services.

If the rate of erroneous predictions is high, blockchain can also be used to store rollback models. The pointers to essential data of retrained models are stored in a safe and immutable

manner on the blockchain. In the context of erratic arrhythmia alarm rate, (Juneja A and Marefat M, 1018) argued that retraining models indexed by pointers in the blockchain can improve accuracies for continuous remote systems.

The application of AI in automating the creation of secure and adaptable smart contracts is another potential use case. The healthcare sector has seen a growing interest in blockchain technology in both academic research and industry, including startups (Ekblaw A et al. 2016, Yue X et al. 2016, Gem 2021, Beninger P and Ibara MA, 2016, Dubovitskaya A et al. 2017, Randall D 2017). In their paper, Wang Z and O'Boyle M claim the originality of using blockchain technology for managing electronic health records (EHRs), while presenting a proposal for securely exchanging EHRs with a user-friendly approach. However, the proposed system is still in its conceptual phase and has not been implemented or evaluated for its anticipated advancements. Ekblaw et al. (2016) developed a decentralized EHR management framework called MedRec using blockchain technology. The framework used a modular architecture and an existing data storage system for ease of use and flexibility. They enticed the medical community and EHR stakeholders to participate as miners in the network's Proof of Work (Kaye J et al. 2015) verification. Permission to view aggregated and anonymized data will be granted in exchange. In collaboration with the Harvard Medical School Teaching Hospital, they developed and tested the first working prototype. They suggested that future research focus on areas where miners can rank their preferences for data attributes (demographic, gender, age group, and so on) to allow precision medicine and targeted research.

In (Dubovitskaya A et al. 2017), researchers built a prototype that differed greatly from the MedRec framework's permissionless mining. From a medical standpoint, they decided to create a closed, access-controlled blockchain EHR system.

To store patient data, MedRec utilized local node storage, while cloud storage and access key transfers for encryption were implemented. The latest research has not fully explored the benefits of incorporating AI into blockchain-based EHR management systems in both permissionless and permissioned prototypes, as indicated by Randall D. (2017) and Rifi et al. (2017).

(Wang Z and O'Boyle M. 2018) provides an overview of how blockchain technology can be used to monitor health records and obtain meaningful results in drug tracking and development, treatment effectiveness, safe patient management, and enhanced clinical results when combined with healthcare and big data. On the other hand, (LeCun Y et al. 2015) discusses other critical aspects of an EHR such as complete reporting, quality assurance, monitoring of patient health-related expenses and billing details, and confidentiality. It highlights the current systems as slow, inflexible, and insecure.

The work in (Khan A et al. 2018) highlights the importance of patient records availability. Due to a lack of time and patience, important aspects of a patient's medical history are often overlooked. A patient's medical history can be extremely useful during care. Doctors, on the other hand, are largely unable to access this information because they lack the expertise, time, or desire to retrieve what they need from a patient's medical data repository.

2.5.1 Methodology of the Literature Review

In conducting this review, we follow SLR guidelines in (S. Keele et al.,2007) as well as the Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines in (S. Keele et al. 2007). An SLR is a methodology for discovering, analyzing, and evaluating all recent literature on a research topic or subject field.

In December 2021, all review papers were chosen by searching for relevant and reliable academic repositories such as Google Scholar, IEEE, ACM, Science Open, Science Direct, Springer, Hindawi, Wiley Online Library, and MDPI.

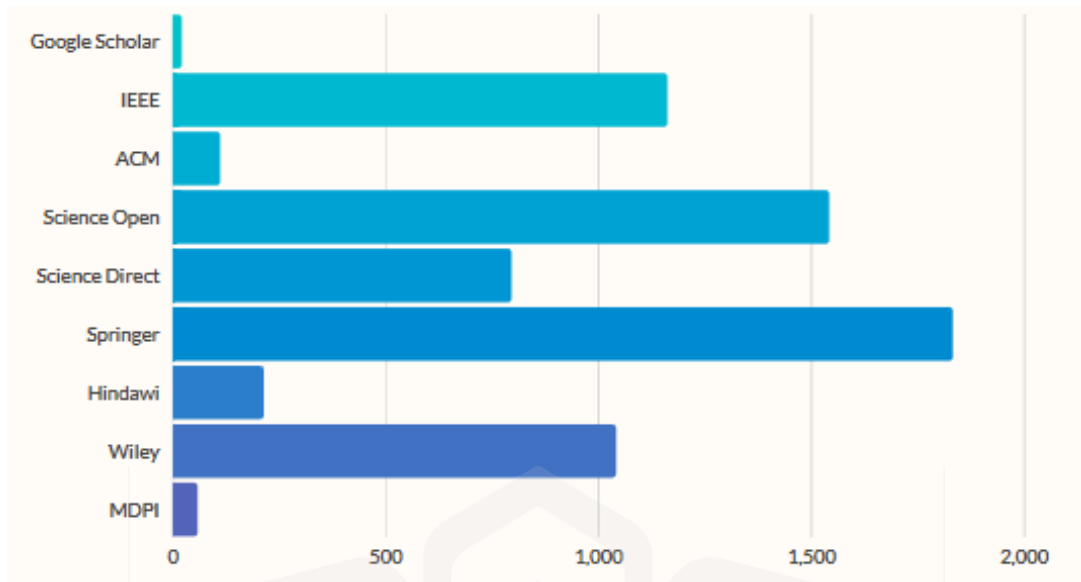


Figure 2.5 Number of articles according to publishers.

2.5.2 Research Questions

The goal of this systematic review paper was to provide answers to the following research questions:

- 1) To what extent has the blockchain been developed for the management of EHRs, and how has it evolved over time?
- 2) What standards are used to store EHRs in the blockchain?
- 3) How large amounts of EHR data are handled?
- 4) What blockchain platforms/mechanisms are used to manage EHRs?

2.5.3 Filtering the literature of the study

After reviewing papers from various categories, selected papers are presented in this portion. As indicated in Section 2.5.2, the article selection query was intentionally extensive to evaluate as many research issues as feasible.

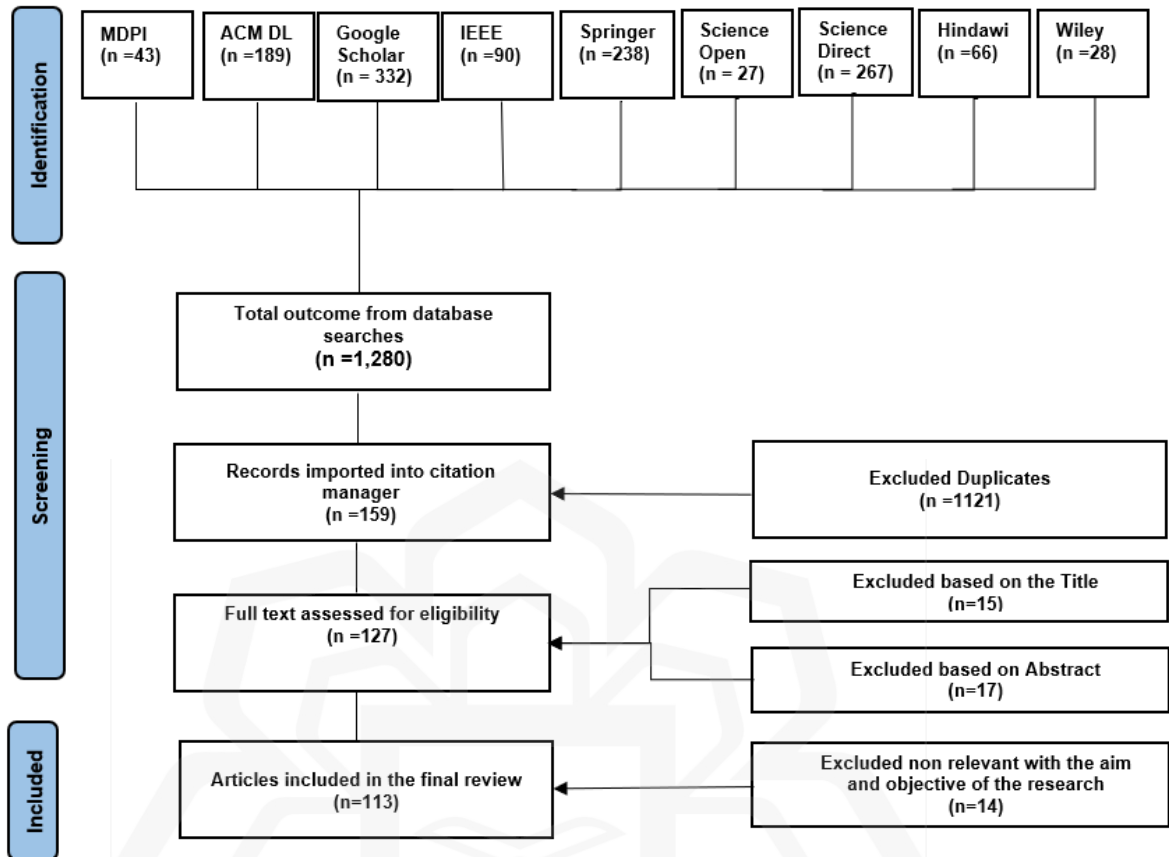


Figure 2.6 PRISMA Chart.

Selected papers are presented in this segment after screening from various categories. The selection query for the articles was purposely long enough to consider as many research questions as possible, as described in 2.5.2. Using the searching mechanism, we were able to retrieve 1280 research articles from the scientific repositories, as shown in Figure. 6. After the first screening step, we removed duplicates and retrieved 159 papers. Using the second and third screening methods (here, exclusion was based on title and abstract), a total of 32 articles were deleted accordingly, leaving 127 papers for further processing. We uploaded the remaining papers to the Mendeley software for thorough reading. Finally, all articles that did not serve the purpose of the SLR were deducted, and a total of 113 articles were there.

The second analysis we ran, as part of our systematic investigation, was to determine the purpose or field of blockchain application in the healthcare industry. As indicated in Table III, the majority of publications in the field of healthcare use blockchain for data interchange, health data records, and access control. A large number of applications of blockchain

technology in healthcare (for example, data sharing and access control) are frequently mentioned by authors, which is understandable given that the blockchain technology itself implies specific applications—for example. Essentially, distributed technologies such as blockchain technology are to be used for data sharing, so, understandably, this field of research would be frequently mentioned.

Table 2.3 Contributions in the publications

Contributions	Number of Publication
distributive mechanism	22
Access control	5
decision-making process	3
Increase interoperability	2

Table 2.4 Contributions in the publications

Field	Number of Publications
Data sharing	20
Health record	18
AI methods	3
Data Security	4
Data Analytics	3
Other	10

Table 2.4 shows the additional analysis for the selected papers. The table compares papers based on five key characteristics.

These characteristics are critical for EHRs. The following properties are discussed further below:

1) Privacy

The concept of privacy refers to a person's right to select when, how, and to what extent they can access, change, and share their own EHRs. (I. Keshta and A. Odeh, 2021). A healthcare provider may purposefully or unintentionally misuse electronic health records (EHRs) to violate patients' privacy, for example (M. Cifuentes et al. 2015). Many patients are concerned about their electronic health records (EHRs), according to the study article (K. T. Win, 2005).

About half of those polled (J. S. Ancker et al. 2013) thought that sharing health data would make it more difficult to protect their personal information. As a result, when comparing blockchain-based solutions that claim to protect EHRs' privacy, privacy is an important consideration.

2) Security

EHR security, on the other hand, refers to the extent to which an individual's electronic health records (EHRs) are confined to authorized individuals. According to (G. Perera et al. 2011), about half of patients are concerned about the security of their EHRs because they must travel via the Internet.

EHR security is more important to a doctor than to patients, according to (S. B. Wikina , 2014), and the majority of doctors prefer paper records over EHRs because they believe the former are safer. Because doctors use digital health records, they are more vulnerable to security breaches than paper-based records (C. S. Kruse et al. 2017). Liu et al. in (V. Liu et al. 2015) advised that ways of securing EHRs should be thoroughly studied first. These aspects indicate that the security of EHRs should be seriously considered.

3) Accessibility

Controlling and managing access to crucial or sensitive data is an essential part of accessibility. Access to data can be restricted using this method. (B. Yüksel et al. 2017) Role-based, attribute-based and identity-based access control are some of the most common strategies for healthcare systems. Because EHRs deal with sensitive patient health data, access management is a critical consideration.

4) Storability

In recent years, the scalability of blockchain technology has become a challenge. Bitcoin's first block had a storage limit of just 1MB when (S. Nakamoto, 2008) first began mining the cryptocurrency's network. But since then, the popularity of the blockchain has increased, as well as the number of participants and blocks. To understand and validate a transaction, a participant must download all of the chains, which consumes a significant amount of memory

and time. On the other hand, conventional blockchain applications have two ways to deal with scalability issues: on-chain and off-chain. Every piece of data that a user uploads will be stored directly on the blockchain. However, off-chain storage means that the true data is held someplace else, but is still linked to the main chain. Off-chain storage, on the other hand, provides less robust security. To store EHRs on-chain, a substantial amount of storage capacity is needed. To keep data safe and secure, it is important to consider storing information outside of the blockchain.

Table 2.5 Research Comparison Used Blockchain and Ai Based Approaches To Secure EHR Systems.

Ref	Objective	Pros	Cons
(Yue X et al. 2016)	Discovering healthcare intelligence focused on the blockchain with privacy	Patient-controlled documents.	Illustration for concept only.
(Zhang J et al. 2016)	For an extensive network to establish a safe health system.	Sharing the network load.	No schemes mature
(Xia Q et al. 2017)	To design health sharing based on blockchain with cloud-based services	Mechanism for Access management	Scalability, core leadership
(Liang X et al. 2018)	Usage of blockchain to exchange health information and communicate with mobile health users	Secure Merkle root tree for collaboration on transactions, data sharing, and healthcare.	The interoperability
(Jiang S et al. 2018)	To build a medical data exchange blockchain-based framework.	Joins the approach to safety and authenticity of off-chain storage and on-chain verification.	Performance and fairness of the system, and

			dynamic regulation of access.
(Li H et al. 2018)	Examination of data security systems in relation to health data	Immutable, memory management and cryptographic algorithms help to handle leaked information	Easily lose paper-based records, slow pace, low memory.
(Fan K et al. 2018)	To strengthen the exchange of effective and safe health information with a blockchain network.	Management and exchanging records from EMR systems, and method of access.	The greater computing capacity of miners contributes to the downstream method.
(Wang H et al. 2018)	To provide the cloud-based support of attribute-based cryptosystem and blockchain to a protected EHR system.	Identity-based encryption guarantees confidentiality and traceability to encrypt databases.	Deployment is not complete yet.
(Guo R et al. 2018)	To propose a stable ABE scheme with multiple blockchain authorities in EHRs	Immutability of the ledger of information	Interoperability, confidentiality
(Uddin MA et al. 2018)	To discuss continuous monitoring of patients with a patient-centered agent.	Lightweight encryption and authentication, tamper-proof, and single point of failure defense.	Delay End-to-End.
(Sun Y et al. 2018)	To suggest a blockchain-based decentralized	Large-scale and distributed EHR,	Certificates attribute, storage space

	attribute-based signature for healthcare.	anonymity, and stable verifiable sharing Non-rebatement.	
(Zhang X and Poslad S., 2018)	To propose access policies for blockchain-based EMR-based systems.	Finer regulation of granular	Proven theoretically.
(Yang G and Li C, 2018)	To build a blockchain-based secure EHR architecture.	Model for Safe records.	Implementation.
(Thakkar P et al. 2018)	To assess the efficiency and optimization of blockchain platforms	Ability to simulate network efficiency	Scalability
(Sukhwani H et al. 2017)	For the creation of a blockchain network dependent on permission.	Defined blockchain integrity permission.	Scalability
(Thakkar P and Natarajan S., 2020)	Using fabric to scale a blockchain network	Demonstrable network blockchain functionality.	It needs increased computing power.
(Chen L et al. 2019)	Using blockchain to design searchable encryption for EHR.	Analysis of protection with a searchable algorithm for encryption	Scalability
(Nguyen DC et al. 2021)	using Federated Learning (FL) for smart healthcare	coordinating various customers, such as hospitals,.. etc, employing a distributed collaborative AI paradigm is very	Implementation

		appealing for smart healthcare.	
--	--	---------------------------------	--

2.5.4 Inclusion and Exclusion Criterion

The authors selected a clear finding centered on the new technological implications of technology and applications for the development by incorporating AI and blockchain into existing health data management systems. Only those studies meeting the first requirements, which must be updated and published in English, should be selected. The findings received from all electronic databases are evaluated based on the developed parameters, and the papers for this systematic literature review are selected from the aforementioned databases.

The criteria of inclusion and exclusion studies have been defined in the following:

Table 2.6 Inclusion And Exclusion Criterion.

Inclusion	Exclusion
<p>IC 1: Original research study.</p> <p>IC 2: Publication related to the topics of AI-blockchain in healthcare data management system.</p> <p>IC3: The study provides ample and strongly correlating research findings in the domain of healthcare data management.</p> <p>IC4: The publication year for the study should be between 2016 and 2021.</p>	<p>EC 1: Review papers that are based on secondary data or are irrelevant to the targeted domain.</p> <p>EC 2: Studies published in the magazine, discussion, and interviews.</p> <p>EC 3: Studies not published in English</p>

2.5.5 Privacy and Security Issues

After reviewing the literature for information extraction, we should answer the following pressing questions.

Q1: How does blockchain ensure privacy by utilizing anonymity?

We see varying degrees of privacy and anonymity (Tzanou, M., 2017) depending on the implementation type of the blockchain: public, private, or licensed. According to (Ahmed and A., 2019), Corda (Hoepman and J.H., 2014) protects the transaction's privacy by requiring validation to be performed only by the persons participating in the transaction. In the field of Industry 4.0, we discover the blockchain-based Secure Mutual Authentication System (BSeIn) (Lieshout et al. 2011), which aims to provide privacy and security assurances such as anonymous authentication, audit capabilities, and secrecy. It demonstrates the scalability enabled by Smart Contracts. They enable privacy via the various consensus methods employed in blockchain (Appari et al. 2010). In other instances, anonymity is used in (O'Keefe et al. 2010). While the work in (Cavoukian and A., 2020) emphasizes conditional privacy, it considers traceability of operations important in the event of a public audit by all entities participating in the blockchain.

The first references we found to anonymization were through pseudonymization (Tzanou and M., 2017), which is the process of obliterating some of the information required to identify an entity. Although they assert in (Sweeney and L., 2005, Vicotia 2018, Skinner et al. 2004) that blockchain does not guarantee completely anonymous transactions and that transactions can be traced using a pseudonym. In The studies of (2019) state that distributed consensus and anonymity are two critical characteristics of blockchain. Cryptography is critical for ensuring the anonymity of participants on the blockchain, with various levels of anonymity achievable depending on the cryptographic methods utilized. Pseudonymization is one method of implementing blockchain technology (Tzanou and M. 2017, Skinner. 2004, Victoria. 2018, Spiekermann et al. 2009). A mechanism in which the identity of the sender is frequently concealed behind a public key, but other transaction characteristics are made public. This presents a difficulty for health data. One option to limit public exposure is to utilize approved blockchain technology. One way to safeguard sensitive data is to implement an out-of-chain solution (Tzanou and M. 2017, Gkoulalas-Divanis et al. 2015). The approach entails locating sensitive data in a system other than the blockchain and anchoring it to the blockchain's link. This technique is advantageous for systems that manage enormous amounts of data, and it would be impractical to incorporate these data into the blockchain

structure. Additionally, it is recommended for systems that handle highly sensitive data and require greater access control, such as health data.

In order to ensure trust and privacy in vehicular communication networks, there is a need for a mechanism that can protect cars from forgeries while also ensuring privacy from surveillance threats. To address this issue, (Dang et al. 2019) proposes a Blockchain-Based Anonymous Reputation System (BARS) that establishes a trust model while preserving the privacy of Vehicular Ad Hoc Networks (VANETs) by using a public key as a pseudonym for anonymous communication. The system aims to prevent the propagation of false messages by using a reputation evaluation algorithm to assess the quality of messages. On the other hand, it exploits the properties of a lexicographical Merkle and eliminates the chance of the public key being linked to the real identity. Such system can be replicated for EHR privacy handling too by taking advantage of the features used.

To accomplish anonymization, the approach presented in (Lieshout et al. 2011) (BSeIn) uses broadcast encryption and multi-receiver encryption to ensure safe communication between an entity and a collection of previously designated receivers. Additionally, it ensures the confidentiality and anonymity of messages between recipients. It produces one public/private key pair at a time for each transaction, allowing it to withstand replay assaults efficiently. Thus, the system can also be replicated for HER to guarantee the user's privacy without jeopardizing it.

Privacy by design (PbD) Privacy by design (PbD) refers to a set of procedures designed to ensure the highest level of privacy and data protection throughout the development of various products, services, and processes. PbD integrates privacy and data security considerations into the development process, from start to finish, for sensitive information like healthcare data. The concept of PbD was introduced by Ann Cavoukian in the mid-1990s, and since then, it has been accredited by data protection specialists and regulatory authorities (Dang et al. 2019, Iachello et al. 2007, Victorian Information3 2019, Cavoukian et al. 2012, Cavoukian et al. 1996).

In 2010, at the International Conference of Data Protection and Privacy Commissioners held in Jerusalem, Privacy by Design (PbD) was adopted as an international standard for privacy (Spitzer et al. 2019). PbD has also been recognized by the Commercial Privacy Bill

of Rights Act in the United States and included in the General Data Protection Regulation (GDPR) in the European Union. Furthermore, PbD has been acknowledged by data protection commissioners globally as a crucial concept to ensure sufficient privacy protection in a world where information technology systems can gather and process vast amounts of data (Donnelly and C., 2019). EHR can definitely benefit from PbD implementation in order to ensure inherent data protection and privacy features throughout the designed system levels.

The strategies for privacy by design are classified into two categories:

A. Data-Driven Approaches

1- Keep it simple: Minimize is the simplest privacy design technique, suggesting that just the barest minimum of personal data should be processed. In (Gürses et al. 2011) describe this method in detail. As a result, it is critical to avoid collecting unneeded data; hence, the probable influence on a system's privacy is minimal.

-Design patterns: "choose before you collect" (Jacobs and B. 2005) and the usage of pseudonyms and anonymization (Pfitzmann et al. 2010) are examples of design patterns that put this technique into effect.

2. Hide: This method emphasizes the need of keeping personal information and its interrelationships hidden from plain view. The idea for this method is based on the fact that hiding personal data prevents a variety of abuses (ISO/IEC 29100, 2011).

-Design patterns: Within the confines of the "hide strategy", design patterns take on a variety of forms. One such pattern is data encryption (in transit or at rest, anonymization or pseudonymization), which refers to strategies that disentangle certain related events. Data encryption is a type of security that encrypts data so that it may be accessed only with the correct encryption key. It converts data to another format and hence requires a decryption key to retrieve the data (Pfitzmann et al. 2010).

3. Separate: This technique stipulates that personal data should be stored in distinct partitions and, if possible, spread out. By segregating the storage and processing of personal information from a variety of sources associated with the same person, an individual's complete profile cannot be derived (Warren et al. 1990). This technique necessitates a

distributed processing solution rather than a centralized one. Data from multiple sources should be stored independently and separately.

- Design patterns: No specific design pattern for this strategy has been identified to date (Hoepman et al. 2014).

4. Aggregate: According to this technique, personal data should be managed with the fewest feasible details and at the highest level of aggregate possible. As a result, this data becomes less sensitive. When the data is sufficiently uneven, the group over which it is aggregated is large, and only a small quantity of data can be ascribed to a single individual, resulting in privacy protection (Hoepman et al. 2014).

- Design patterns: There are two common strategies used. “Granularity of location” design pattern that changes dynamically enables the collection and delivery of data to be as efficient as possible (J.H ,2014). “K-anonymity” design pattern, on the other hand, is a critical model for privacy protection since it protects against joint attacks. It is a dataset characteristic that is used to describe the dataset's degree of anonymity (Sweeney and L., 2002).

B. Process-Oriented Approaches

1. inform: This technique embodies the critical concept of transparency. If personal data is processed, data subjects' information should be kept current. When a user interacts with a system, they should be appropriately informed about the data that is processed and why it is processed. This includes information on the mechanism used to protect the data in question and transparency regarding the system's security (Hoepman et al. 2014).

-Design patterns: Both platforms for privacy preferences and notifications of data violation are examples of this type of design pattern. The work in (Graf et al. 2010) presented an unusual array of privacy design patterns intended to educate the user from the perspective of human-computer interaction.

2. Control: This technique is a necessary complement to the “inform strategy”. It serves little purpose to tell the user that personal data is being gathered unless the user has a realistic means of limiting the use of his data (Deng et al. 2010). Users frequently have the right to access, amend, and request deletion of personal data gathered under data protection

regulations. This technique accentuates this point and enables users to exercise their data protection rights (Hoepman et al. 2014).

-Design patterns: There are no specific design patterns that fit the strategy (Hoepman et al. 2014).

3. Implement: This technique ensures that the system operates in a manner that respects user privacy. More significantly, the policy must be implemented. To ensure that the privacy policy is not violated, adequate technical protection measures are developed. Additionally, the policy must be formed through an effective governance system (J.H ,2014).

-Design patterns: This method is carried out using design patterns such as access control and privacy rights management, and license to personal data, which includes the form for managing digital rights (Hoepman et al. 2014).

4. Exhibit: This approach establishes the relationship between a data controller and the monitoring of compliance with privacy policies and applicable regulations. In the event of issues, the user should promptly be able to determine the amount of any potential privacy infringement.

-Design patterns: Examples of design patterns that support this strategy include the usage of logging and auditing, as well as a privacy management system (Graf et al. 2010).

Q2: What are the drawbacks of using blockchain to comply with the GDPR?

There are several practical challenges and obstacles associated with the implementation of blockchain technology in the healthcare industry, which include ensuring compliance with GDPR regulations. Compliance with an individual's right to be forgotten is one such challenge. When a transaction is authenticated on the blockchain, it becomes permanent, and information about a patient cannot be deleted if the patient exercises his right to be forgotten. This limitation may compromise the patient's privacy and right to control their data (Tzanou and M., 2017, OECD. 2019).

Although the identity associated with the transaction introduced into the blockchain is anonymized, the remainder of the transaction's information is accessible. This feature enables auditing of the entire blockchain when necessary, which may result in the exposure of sensitive information such as EHR, and ultimately the determination of the transaction's

identity (Gkoulalas-Divanis et al. 2015, Fernández-Alemán et al. 2013, Cavoukian et al. 2020).

Q3. How were the arisen issues resolved?

Depending on how blockchain is implemented, various privacy concerns may occur, making it easy to track an entity's transactions. A notable example is given in (Tzanou and M. 2017), where an entity's public key corresponds to its identity in the blockchain system, allowing for the discovery of all transactions linked with that public key. This scenario would be catastrophic in a public blockchain and might also present an issue in a private blockchain, as not all members may require access to transaction data. The case in (Tzanou and M. 2017) refers to specific blockchain implementations that enable selective publication of private information and rely on zero-knowledge cryptography for verification. How to apply the GDPR-mandated right to be forgotten for a patient's data is one of the disadvantages demonstrated when implementing blockchain in the health area. Among the downsides of blockchain technology are the costs involved with authenticating connected data, auditing different entities and transactions, and the cost of interoperability provided to the network of participants. The pseudonym does not ensure transaction privacy, and it is even feasible to de-anonymize a user's identity through analysis of incoming and outgoing transactions.

Privacy by Blockchain Design (PbBD): Privacy by Blockchain Design develops on data privacy solutions for the disruptive and rapidly growing new tech ecosystem. Blockchains can not only be GDPR-compliant, but they can also help raise data protection levels and truly give back data ownership to individual patients or their legal guardians (e.g., family members or the state), by establishing general principles and methods for handling personal data in blockchain ecosystems. PbBD specifies technical and organizational measures for data protection while taking into account the principle of "privacy by design" as well as specifications that are inspired by legal frameworks, such as GDPR. As such, the Blockchain as a great tool for privacy and want to encourage the industry to take the lead in this area.

2.6 TAXONOMY OF AI-BLOCKCHAIN

2.6.1 Decentralized Applications

AI applications are self-contained and execute intelligent decisions by making use of a range of strategizing, discovery, improvement, training, pattern recognition, and information management methodologies. Decentralizing AI activities, on the other hand, is a tough and time-consuming task.

2.6.1.1 AUTONOMOUS COMPUTING

One of the primary aims of AI applications is to facilitate the complete or partial autonomous process. This is achieved when a large number of intelligent agents in the form of small size computer programs identify their component ecosystems, sustain their internal environments, and conduct set actions to produce a response (Ye D and Zhang M, 2016). Modern computer systems must be able to handle tremendous heterogeneity across all verticals to operate autonomously, which often includes datasets, instruments, data processing, storage services, and application linkages, to name a few. Not only the usage of a multiagent approach across all layers makes it more convenient to deal with heterogeneity, but it also enables the easier establishment of the inter-and intra-layer functionality across the whole systems (Rizk Y et al. 2018). By ensuring operational decentralization and retaining perpetual records of interactions between the data, users, devices, apps, and systems, the blockchain architecture is significant in developing wholly decentralized autonomous systems.

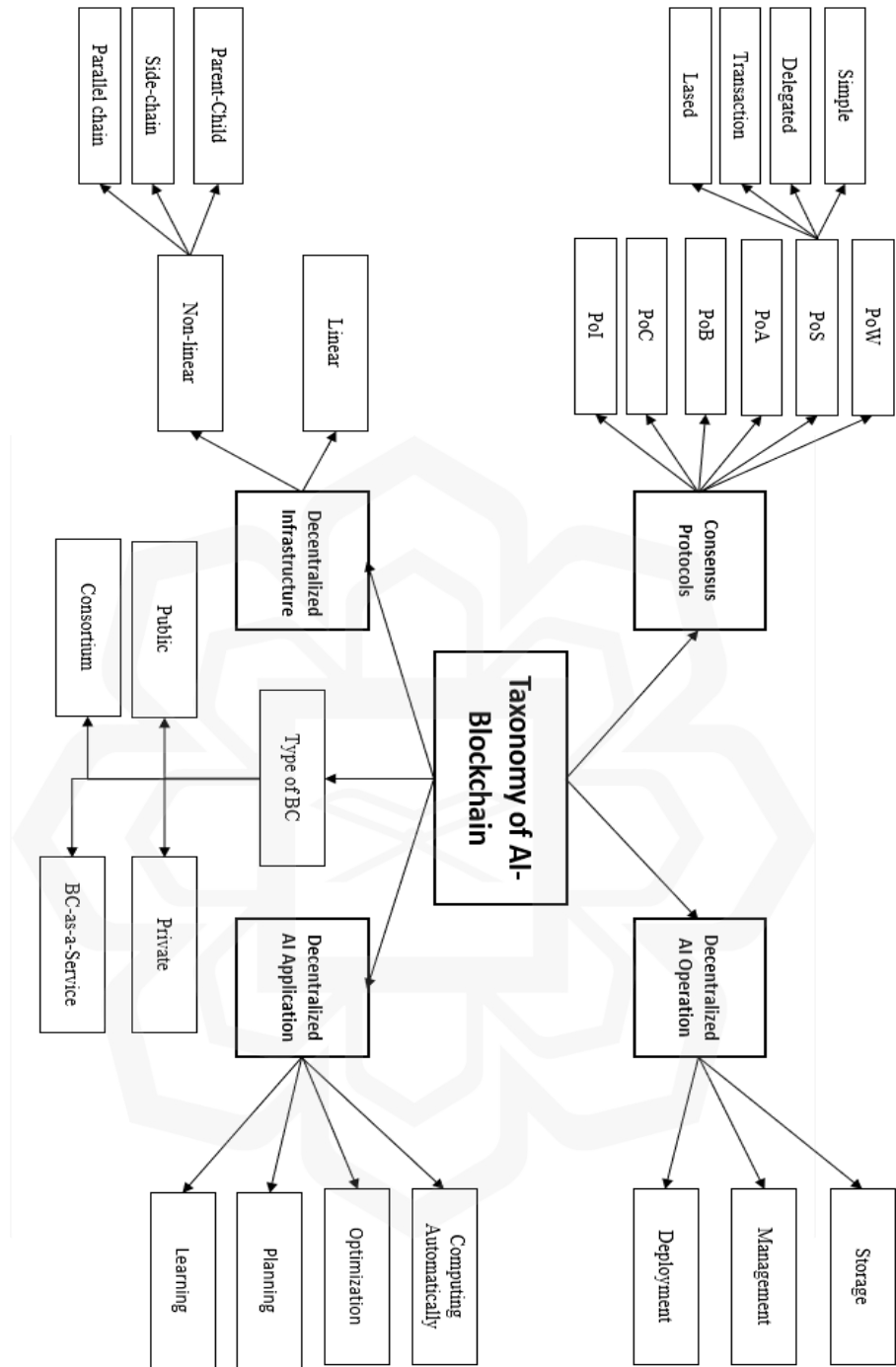


Figure 2.7 Taxonomy AI- Blockchain

2.6.1.2 Optimization

Among the primary characteristics of AI-enabled apps and schemes is the discovery of a collection of optimum solutions from all available alternatives (Fioretto F and Pontelli E, 2018). Modern AI applications and systems can be found in ubiquitous computing such as edge computing systems, infrastructure-restrained environments in mobile devices, spatially confined ecosystems such as wireless local area networks and personal area networks, and centralized enormous parallel computing systems distribution as applied in cloud computing (Rehman M ur and Liew C et al. 2017, Rehman M ur and Batool A et al. 2017). The optimization algorithms operate in confined or unconstrained environments based on application- and system-level objectives (Rehman M ur and Batool A et al. 2017). These strategies facilitate the discovery of the most suitable solutions in identifying the pertinent data sources in pervasive environments, the best cloud or edge servers for processing the data and application, as well as in allowing resource-efficient information management in extensive distribution of computing settings.

The optimization process at present is implemented by centralizing control and taking into account system-wide and application-wide enhancement objectives, causing unnecessary and irrelevant management of data and poor performance of the system or the application itself (Bottou L et al. 2018). The application of blockchain enables decentralized optimization methodologies to bring up new research and development possibilities. By analyzing highly applicable data, the decentralized optimization techniques are advantageous in terms of improving system performance, particularly when numerous techniques are executed concurrently across the systems and applications.

2.6.1.3 Planning

AI apps and systems use planning approaches to collaborate with other systems and applications, as well as to solve complex problems in new situations. Planning strategies improve the operational efficiency and resilience of AI systems by gathering current input conditions and performing different logic and rule-based algorithms to achieve preset goals (Contreras-Cruz M et al. 2017). Currently, centralized planning is a laborious and time-

consuming activity. Consequently, decentralized AI planning techniques based on blockchain are required to provide a higher degree of robustness with provenance history and continuous monitoring. It is worth noting that the blockchain ecosystem can also be used to create immutable and critical blueprints for task-essential systems and relevant applications.

2.6.1.4 Learning

Learning algorithms, with models such as unsupervised, semi-supervised, supervised, reinforcement, transfer, ensemble, and deep learning, remain to be the heart of AI systems in facilitating knowledge discovery and autonomous processes. These learning models tackle a wide range of machine learning issues, from classification to clustering, besides regression analysis to frequent pattern mining. Traditional learning models are taught and released by utilizing centralized infrastructure to achieve global intelligence.

Dispersed learning models can help in the construction of highly propagated and automated learning systems in contemporary AI systems, allowing for the complete co-ordination of local intelligence across all verticals (Kurtulmus AB et al. 2018, Kim H et al. 2019). Furthermore, by maintaining data provenance and history, the blockchain enables irreversible and highly secure configuration of learning models. Because smart contracts are irrevocable, learning models must be extensively trained and evaluated before they can be implemented on the blockchain.

2.6.2 Decentralized Operation

Large volumes of data are typically handled by AI applications to make superior and more versatile decisions. However, when it comes to designing highly secure and privacy-preserving AI systems, centralized data retention via clouds, data centers and clusters presents a significant challenge. In other cases, learning model development and deployment might also be arduous.

2.6.2.1 Storage

A centralized data server raises the issue of vulnerability in terms of privacy and security concerning the users' personal and sensitive data, such as financial information, health records, whereabouts, and activities. Furthermore, as AI applications attempt to analyze, transform, and store massive information, wide-scale data collection would reveal the centralized infrastructure's scalability and capacity constraints. Blockchain-based decentralized storage architecture enables reliable cryptographic data storage across collaborating networks (Mcconaghy T et al. 2016, Shafagh H et al. 2017, Cui S et al. 2018). To maintain data availability for desired clients consisting of an application, user, or a node on the blockchain, every node in the system maintains a client-centric openly secured version of the whole library, which the clients can harvest and utilize their data as needed.

The key technologies for decentralized storage are sharding and swarming (Cui S et al. 2021, Zamani M and Movahedi M. 2018, Rıfat"ozyılmaz K et al. 2018). Sharding is a technique of dividing a database into logical parts and assigning each one a unique key to be accessed. The shards are then grouped, with the accumulated storage is supported by a swarm of network nodes. In AI applications, the swarms reduce latency by allowing numerous nodes in the network to access data simultaneously. In addition, geographically dispersed multiparty decentralized storage systems would improve storage scalability and dependability.

2.6.2.2 Data Management

AI applications must manage data in such a way that is highly applicable and precise, with full datasets obtained from credible data sources, along with effective decentralized storage. In the underlying network, AI applications traditionally have used centralized data management techniques operated across all nodes (Vo H et al. 2018). These strategies include but are not limited to, data segmentation, filtration, context-aware storage systems and transmission in underlying architecture, as well as temporal and intelligent management of data systems. When considering decentralized storage networks and blockchain immutability requirements, inefficient centralized data management may arise,

resulting not only in data redundancy in terms of small modifications but also in the transfer of comparable information several times. In the event of large datasets is being utilized, the massive size of data transfer would cause bandwidth to overload quickly and raise the issue of backhaul network traffic, thus, necessitating decentralized data processing for AI systems based on blockchain structure. By taking into account the data's temporal and spatial features, decentralized data infra-structure strategies are intended for application at the network node level. Furthermore, decentralized data management systems may place metadata on the blockchain network to assure data security and provenance while the conventional large-capacity storage solutions, including cloud clusters and data centers, might be utilized to store actual data. For client-centric small datasets, the metadata and real data are maintained on the blockchain, with the management of data being done through the network via token-based incentives for nodes carrying various shards or participants in swarms.

2.6.2.3 Deployment

A trained model's true performance is evaluated after the distribution in production settings. Model deployment, on the other hand, is a regular and repetitive process as the developers must constantly improve the models and rectify bias by generating a certain set of findings while disregarding the rest of the options to provide particularly useful and educated judgments. Model deployment is considered a simple iterative process in centralized systems. In decentralized systems, however, poses quite a challenge (Lai L and Suda N. 2018). Intelligent contract-based model deployment overcomes these difficulties by constantly logging changes and preserving unchangeable model versioning. Furthermore, a model collaboration between various AI applications would be safer and more reliable since developers can monitor the origin and traces of a specific model version.

2.6.3 Blockchain Types for AI Application

The two types of blockchain technologies consist of Permission and Permissionless structures. For the Permission type, only authorized users would be able to handle the blockchain applications in a consortium, cloud-based, or private environment, while it is openly usable for all users over the Internet for the Permissionless type.

2.6.3.1 Public

Users may retrieve the blockchain codes and save them to their terminal for editing and utilizing based on their needs using the public blockchains (NN-A at S. 2017, GW-E. 2014). To add to this effect, public blockchains can be easily accessed and available to all network participants, particularly for read and write operations. Because of this feature, blockchains employ complicated security and consensus methods, as well as anonymity and bogus data on the network to handle user credentials and private transactions. For any public blockchain, innate tokens such as valuable pointers and cryptocurrencies are used to move assets and data. Due to its huge decentralization and transparency, public blockchains are extensively used, even though the users and validators are constantly being anonymous. It was worth noting that due to the obscurity, hostile security assaults such as significant data and value theft on these blockchains are always a possibility.

To reach a consensus, public blockchains would require 51 percent validators at the very least and would perform complicated mathematic works in the background to attempt cracking the security codes, which often results in high energy expenditure and the issue of vulnerability if the attackers obtain control on the 51 percent shares on the network. This might also be the reason for the higher transaction approval times on public blockchains as compared to the consortium and private blockchains.

On a public blockchain, a transaction is often approved in 10 minutes or above, depending on the number of network users and the mathematical complexity of the consensus algorithms used.

2.6.3.2 Private

A single organization manages a private blockchain, which is structured as a Permission system so that the acknowledged users and participants would be pre-authorized for read and write activities within the network (Dinh TTA et al. 2017). Since the credentials of pre-approved network participants and validators are known, private blockchains are comparably faster than public blockchain as it requires fewer mathematical operations for transaction validation purposes on the network. In addition, within the network, private

blockchains can broadcast any type of indigenous assets, data, and values. Voting or multiparty consensus algorithms are used to approve transactions and asset transfers, which require minuscule energy consumption, allowing for a quick transaction process. For example, on private blockchains, transaction approval times typically take less than one second.

2.6.3.3 Blockchain-As-A-Service

Due to widespread usage and approval by governments and large corporations, blockchain technologies are drawing the attention of cloud service vendors. Customers of major cloud suppliers such as Microsoft, Amazon, and IBM can now create and experiment with blockchain services in their environments (Lai L and Suda N. 2018).

The emergence of BaaS is projected to benefit both consortium and private blockchain firms by allowing them to concentrate on creating value through apps development, validation, and implementation rather than worrying about the infrastructures associated with the storage, underlying network, and computation. Besides the fact that the installation of BaaS facilitates the formation of new cross-industry private-public partnerships, it also helps in the development of new opportunities and company-customer interaction models. To construct smart contracts, developers have access to a single-click setup of BaaS services. On that note, the incorporation of BaaS with AI services opens up a new world of possibilities for apps developers, considering that the main cloud providers currently are offering a plethora of cloud services for AI applications.

2.6.4 Decentralized Infrastructure

Traditional blockchain systems built a linear infrastructure using a mixture of a connected list of data frameworks and hashing algorithms. Nonlinear infrastructure, built upon graph theory and buffering data modeling, on the other hand, is growing to meet the needs of instantaneous applications and to manage massive volumes of data.

- Linear: Blockchain system based on a single chain that expands linearly, with new blocks inserted at the chain's end. The early adoption phase of a decentralized system usually uses single chains despite several flaws associated with it. For example, single chains would

scale sluggishly, affecting the real-time performance of decentralized applications (NN-A at S, 2017, GW-E. 2014). Furthermore, because each business situation has its single chain, information, asset, and value exchange in different chains would be a challenging task. Single-chain blockchains instead, may be used for single-task AI systems that conduct search, refinement, and training, as well as autonomous AI applications that function in homogenous environments. Rather than the AI programs themselves being executed via smart contracts, single-chain blockchains may be more advantageous when just the performance records of AI apps need to be preserved in perpetuity. For instance, in radiology applications, a model for deep learning can be used to deliver accurate results for diagnosing liver cancer. The successful search footprints of distant industrial robots could be another example of its use. Since AI applications typically function in unrestrained contexts, placing the entire components on blockchain structure is not a viable option.

- Non-Linear: Multichain architectures are utilized to construct nonlinear blockchain architectures, using topologies and different types of chains such as parallel, parent-child, and main-side (King S, 2012). Multi-chain architectures not only offer a broad range of business cases and inter-chain value transfer, but they are also scalable for live performance. One or more chains would serve as the primary chain in a multi-chain structure, holding the data concerning other chains while the remaining ones would be employed as the parallel, side, or child chains. Side and child chains are typically similar in operation, with the principal difference being that the business scenarios in child chains are firmly related to parent chains while the side chains can operate completely independent from the main chains. As for the parallel chains, they could function separately from one another. To transfer the value between several chains, the "pegging" approach is implemented by integrating a two-way peg procedure that allows bidirectional value transfer at a fixed exchange rate. It should also be noted that in the blockchain, native currencies or tokens would represent the exchange value. For interested readers, the following studies provide a full discussion of nonlinear blockchains.

In decentralized apps, nonlinear blockchains for the AI apps domain grant the operation of several related and independent AI tasks.

Furthermore, the scalability property allows AI applications to be developed and deployed in parallel such that AI parts are installed on the main or parent chain in a production context, while the testing and training apps are loaded on the test nets or side chains. Emerging apps, such as those in adapting and reinforcing learning algorithms, benefit from nonlinear architectures since the principal applications must continually improve their productivity by reconfiguring the learning models. In this case, learning models are built on the side chains and subsequently deployed on the main chains.

2.6.5 The Role of Consensus Protocol

2.6.5.1 Proof-of-Work (Pow)

The PoW consensus mechanism is used by popular public blockchain systems, namely Ethereum and Bitcoin to verify transactions after the participation of at least 51 percent of nodes on the underlying network (NN-A at S, 2017, GW-E. 2014).

Because the validating nodes run anonymously and in vast numbers, they must produce the blocks by deciphering complicated and arbitrary mathematical problems, as well as cracking the hash code to access the transactions on the blockchain network. To receive the prizes, the successful nodes send the answer through a peer-to-peer network. Additional transactions and data are irrevocably joined to the blockchain when 51 percent of the nodes successfully solve the mathematical problem. Although PoW has shown to be a standard consensus protocol, it consumes a lot of energy in large networks and causes delays in transaction approvals time. As astute algorithms regularly streamline decision structures to make an educated judgment, AI applications would have a higher prevalence of write operations. As a result, in real-time AI applications, PoW protocols would become a performance barrier, besides the fact that an attack on 51% of the nodes in the underlying network could jeopardize the reliability of AI applications.

2.6.5.2 Proof-of-Stake (Pos)

Consensus PoS-based techniques attempt to address the problem of PoW's excessive energy consumption (King S, 2012). The PoS protocols function by identifying key players on the

blockchain network to allow them to generate new blocks. These methods select validators based on a variety of factors, such as delegated, high frequency transacting, random, or those that maintain coins for a longer period.

PoS has shown to be more energy-efficient than PoW, and it also solves the vulnerability issue by eliminating pseudonymous validators and allowing only those who possess the blockchain's native currency to participate. Validators, on the other hand, have little to risk if they do not authenticate the transactions on the blockchain, which may delay the development of new blocks. Although PoS is useful for the lag-tolerant AI apps, it is not ideal for AI systems, especially in the management of flowing data, changing the identification, and making intelligent decisions on a real-time basis.

2.6.5.3 Proof-of-Activity (Poa)

PoAc is a mixture of PoW and PoS protocols. Such protocol aims to address the 51 percent attack problem by implementing the PoW algorithm on blank blockchains (Bentov I, 2014). This is done by PoAc protocol solves complicated mathematical problems first and validators begin to receive incentives, increasing their holding on the ledger. This allows for the validators with a sufficient stake in the blockchain to use the PoS algorithm. Additionally, PoAc is effective in terms of security, memory, and network connectivity.

As a result, it may be advantageous for AI programs that require less data accessibility and higher security.

2.6.5.4 Proof-of-Burn

According to the PoB protocol, validators can only spend their coins if they send them to a public, valid, unusable, and faulty address. After burning their money, users are instantly authorized to develop new blocks and collect incentives (NN-A at S, 2017). Users could benefit from PoB since it allows them to contribute in advance and earn interest on the chain while also becoming approved validators. The protocol also gives an advantage by fixing the PoW's energy use problem. Furthermore, coin burning lowers the number of coins on the ledger, resulting in a gradual increase in coin value, amount balancing of currencies on the blockchain, the spending of unsold coins, and payment of the transaction cost. PoB protocols

can be used by AI systems to urge participants to keep the value of the underlying judgments. Applications needing a specific degree of precision, a set amount of clusters or items recognized, for example, can consume learning models and search trees to keep value over the ledger.

2.6.5.5 Proof-of-Capacity (Poc)

Traditional PoW algorithms become computationally expensive since they must obtain randomized nonce values to decrypt the blocks. The Proof-of-Concept protocol, commonly called proof of space, is a substitute mechanism for determining the space amount of hard drive on the blockchain network's nodes (Tschorsch F, 2022).

Rather than utilizing random numbers, it stores the potential nonce values on the hard drive and looks for matching nonce-hash combinations to decrypt the blocks. Nodes that are having a large amount of disc space would obtain a lot more stake and a high chance of winning with PoC.

2.6.5.6 Proof-of-Authority (Poa)

PoA could be used to address the problem of PoW's high energy usage, as well as the issue of the validators should possess a portion of capital invested in the blockchain network. A PoA protocol delegates authoritative power to specified nodes, forming a consensus based on the absolute majority to create additional blocks on the ledger (Angelis S De, 2018).

PoA has been shown as being a resource-effective and low-latency consensus system, albeit it is better suited for networks in private since it allows authorized stakeholders to delegate validation authority. Consequently, blockchain implementers must consider the validators' legal identities, well-defined eligibility requirements, and a common qualification condition for each shareholder to operate as validators. PoA security risks are always present, owing to security attacks on validators, which could be a source of assault on the network, notwithstanding their energy economy and fiscal efficacy. Alternatively, PoA might be used as a substitute consensus approach for AI systems that run on private or consortium networks because all validators are recognized over the system.

2.7 E2E ENCRYPTION

End-to-end encryption (E2EE) is a secure communication method that ensures data is inaccessible to third parties during transfer between two devices or systems. E2EE involves encrypting data on the sender's device, which can only be decrypted by the intended recipient. This ensures that messages are protected from hackers, ISPs, application service providers and other entities.

Many messaging service providers such as Facebook, WhatsApp and Zoom employ E2EE, but its adoption has been controversial as it makes it difficult for providers to share user information with authorities and could facilitate private messaging among individuals involved in illegal activities.

To encrypt and decrypt messages in end-to-end encryption, the cryptographic keys are stored on the endpoints and public key encryption is used. This encryption method utilizes a public key and a private key.

The public key can be shared with others, who can use it to encrypt a message and send it to the owner of the public key. The message can only be decrypted by the corresponding private key, which is also known as the decryption key.

When exchanging messages online, there is typically an intermediary server that facilitates communication between the parties. These servers are usually operated by internet service providers, telecommunications companies or other organizations. However, with end-to-end encryption that uses public key infrastructure, these intermediaries are prevented from intercepting the messages being exchanged.

To ensure that the public key belongs to the intended recipient, it is embedded in a certificate that is digitally signed by a trusted certificate authority (CA). The CA's public key is widely known and its authenticity can be trusted.

Therefore, a certificate signed by that public key can be considered valid. The certificate links the recipient's name and public key, so the CA would not sign a certificate that associates a different public key with the same name.

2.7.1 How Does E2EE Differ From Other Types Of Encryption?

What makes end-to-end encryption unique compared to other encryption systems is that only the endpoints -- the sender and the receiver -- are capable of decrypting and reading the

message. Symmetric key encryption, which is also known as single-key or secret key encryption, also provides an unbroken layer of encryption from sender to recipient, but it uses only one key to encrypt messages.

The key used in single-key encryption can be a password, code or string of randomly generated numbers and is sent to the message recipient, enabling them to unencrypt the message. It may be complex and make the message look like gibberish to intermediaries passing it from sender to receiver. However, the message can be intercepted, decrypted and read, no matter how drastically the one key changes it if an intermediary gets ahold of the key. E2EE, with its two keys, keeps intermediaries from accessing the key and decrypting the message.

Another standard encryption strategy is encryption in transit. In this strategy, messages are encrypted by the sender, decrypted intentionally at an intermediary point -- a third-party server owned by the messaging service provider -- and then re-encrypted and sent to the recipient. The message is unreadable in transit and may use two-key encryption, but it is not using end-to-end encryption because the message has been decrypted before reaching its final recipient.

Encryption in transit, like E2EE, keeps messages from being intercepted on their journey, but it does create potential vulnerabilities at that midpoint where they are decrypted. The Transport Layer Security encryption protocol is an example of encryption in transit.

2.7.2 Benefits Of Using E2EE

The main advantage of end-to-end encryption is a high level of data privacy, provided by the following features:

- Security in transit. End-to-end encryption uses public key cryptography, which stores private keys on the endpoint devices. Messages can only be decrypted using these keys, so only people with access to the endpoint devices are able to read the message.
- Tamper-proof. With E2EE, the decryption key does not have to be transmitted; the recipient will already have it. If a message encrypted with a public key gets altered or

tampered with in transit, the recipient will not be able to decrypt it, so the tampered contents will not be viewable.

- Compliance. Many industries are bound by regulatory compliance laws that require encryption-level data security. End-to-end encryption can help organizations protect that data by making it unreadable.

2.8 DISCUSSION

1) To what extent has the blockchain been developed for the management of EHRs, and how has it evolved over time?

Authors in the literature attempted to propose solutions for managing EHRs from various perspectives. For data encryption, many people used symmetric encryption schemes, while others used asymmetric encryption schemes. A few authors provided solutions for the blockchain's scalability when managing EHRs. Some people brought smart contracts, while others used chain-code for EHR preservation mechanisms. When it comes to EHR storage, there are two options: on-chain storage and off-chain storage. An on-chain storage scheme is focused on storing data on the blockchain, whereas an off-chain storage scheme stores data in the cloud or a local database and links the data's address to the blockchain.

From 2016, when blockchain-based solutions for managing EHRs first became available, to 2021, there has been tremendous progress. The idea of using blockchain as a platform to manage health data was first mentioned in the article (Ekblaw A, 2016). Later that year, an article (Xia Q, 2017) discussed the use of private blockchain for EHRs. Following that, researchers attempted to demonstrate the utility of AI-blockchain for handling EHRs.

2) What standards are used to store EHRs in the blockchain?

The issue of data format and interoperability standards continues to be a challenge for sharing and storing EHRs. While most authors have considered FHIR and HL7 as potential standards for EHRs data format, only a few have followed the HL7 standard, and a small number have considered FHIR. The use of a standardized EHR data model can help support interfacing with clinical decision support systems. Some authors have

described the standard of ISO 27789, HL7, and HIPAA, but have not implemented those principles.

Despite these efforts, achieving a standard for EHRs exchanging, uploading, storing, authenticity checking, and formatting remains a critical challenge for blockchain-enabled EHRs solutions. This could be due to the evolving nature of blockchain technology and the lack of standardized development platforms. While blockchain shows promise for EHR management, it still has a long way to go before it can be considered stable enough to support a standardized framework.

3) How large amounts of EHR data are handled?

Dealing with massive amounts of data is a significant challenge, and it becomes more difficult when it comes to handling data via blockchain due to its high storage costs. While blockchain was initially designed for small-sized financial transactions, researchers have devised solutions to overcome the limitations of data storage capacity. However, slightly less than half of the papers reviewed for this topic did not address the major data storage issue. Some authors have addressed the issue but did not discuss data storage services. Others have chosen the Interplanetary File System (IPFS) as a medium of data storage and then linked the address with the blockchain, while others have proposed using private blockchain or off-chain storage to handle scalability issues.

Overall, while solutions have been proposed to address the challenge of data storage on the blockchain, it remains a significant issue that needs to be addressed to fully realize the potential of blockchain in managing large amounts of data.

4) What blockchain platforms/mechanisms are used to manage EHRs?

Because EHRs include sensitive personal information, a private blockchain is at the top of the popularity ranking. Furthermore, a private blockchain can enable access control rules, allowing only particular persons to join the network while adhering to good security policies. A public blockchain, on the other hand, does not have strong access control rules,

so anyone can join the web and read the data. A consortium blockchain also provides a private network and restricts access to network data.

The literature review included several potential models or architectures. The majority of the authors concentrated on EHR integrity, availability, transparency, privacy, and security. Almost all of the models offered to support for the storage of EHRs from medical institutions as well as wearable devices. A significant number of papers used the Ethereum platform for the proposed solutions. The number for Hyperledger Fabric was only one paper (Sukhwani H, 2017). The rest of the offered solutions include Bitcoin (NN-A at S, 2017), (GW-E project, 2014), (Bentov I, 2014) ,(Tschorsch F, 2016) ,(S. Nakamoto, 2008) consortium blockchain (Lai L, 2018), (Li Z, 2017), private & consortium blockchain (Lai L and Suda N, 2018), (GW-E, 2014), Multichain (King S, 2012), private blockchain (Cyran MA, 2019), (Lima C. 2018), (Dinh TTA, 2017), (Tzanou, M. 2017), and Permissioned Blockchain (K. Gai, 2019).

2.9 OPEN CHALLENGES AND FUTURE RESEARCH OPPORTUNITIES

One may define numerous issues of healthcare Blockchain-based applications based on the proposed prototypes and developed applications discussed above.

With the introduction of wearables and a slew of new IoT devices with data flows harnessed, improved security is required to be readily available to healthcare providers (Ferrag, 2018). These issues might be addressed with blockchain technology, which offers interoperability, integrity, and security, as well as portable user-owned data.

Interoperability refers to a system's capacity to seamlessly integrate with another system to share critical data. The ease of transformation of the medical records and the healthcare data information from one provider to another is referred to the interoperability in the EHR system. While health care organizations can connect in a variety of ways, the EHR is generally regarded as one of the simplest and most secure methods that do not result in information blocking (Shafagh H. 2017, Cui S, Asghar M. 2018]. To begin with, the EHR must have core interoperability. This enables the entire system to send data to another system while also receiving data. While the data received will not need to be analyzed as part of this

level of interoperability, it will be available within the system immediately. This is the lowest degree of functional interoperability, allowing only the most basic data exchange.

Second, the EHR must have structural interoperability, which means that data must flow appropriately through the system so that providers may see unmodified patient data. To establish a new EHR database utilizing structured messages, this intermediate region of health care data exchange ensures that patient information is provided and received in a relevant and shareable fashion. Furthermore, even if the data changes hands, the facts, and meanings will not be altered.

Third, the EHR must have semantic interoperability, which allows data to be accurately reorganized and codified so that any system can receive and interpret the new information. This means that the language used by one EHR system must be readable by the next system. This is the highest level of interoperability possible with significant implications for patients, clinicians in a health system, and scientists and researchers who collect data to study patient populations. Due to the adoption of standardized coding, information is transferrable and usable at this level. In contrast to studies (Liang X. 2018, Guo R. 2018), which lacked the possibility of interoperability and is not discussed in EMR systems as a result, medical and health data experts must perform manual inspection and mapping of predefined ontologies. At the same time, clinical malpractice is uncontrollable. Furthermore, scalability and interoperability concerns are at the forefront of current and future research in this area. The lack of standards for designing healthcare applications based on blockchain technology is revealed by the interoperability challenge.

Guo R. (2018) has pointed out that privacy and security pose a significant challenge for blockchain technology. The decentralized nature of the blockchain, where data is distributed to all nodes, can lead to non-compliance with privacy regulations and create security vulnerabilities. As a result, to protect data privacy and security, data must be stored off-chain. New privacy technologies, such as homomorphic and attribute-based encryption, secure multiparty computation, zero-knowledge proof, obfuscation, and format-preserving encryption, and may be able to accomplish data privacy (Jiang S. 2018).

Designing using hybrid privacy approaches and leveraging security-enhancing technology, such as a homomorphic signature, which works better than public-key

certificates, could speed up the different security levels in a system. More significantly, any malicious attacker can manipulate health data acquired from hospitals, clinical labs, and patients, rendering AI learning useless. As a result, utilizing federated learning mixed with blockchain technology, it is necessary to collect health data from many sources without any privacy leaks. Each healthcare organization's central entity is responsible for any legal difficulties as well as the overall seamless operation of the centralized healthcare systems. A decentralized, patient-centric system, on the other hand, makes it difficult to resolve any legal disputes or inconsistencies in the public blockchain architecture. When personal data is run on converging AI and blockchain platforms, for example, copyright infringement and defamation issues occur. On the other hand, scalability is the main issue in blockchain-based healthcare systems (Kim H. 2019, GW-E. 2014, Dinh TTA et al. 2017, Li Z et al. 2017, Hwang GH et al. 2018), especially when dealing with large amounts of medical data. Due to the high volume of healthcare data, it is not feasible to store it on-chain, as this would result in significant performance degradation. To achieve consensus and ensure ledger replication across all network participants, blockchain networks have always faced limitations in terms of scalability, as stated by Houtan et al. (2020). These limitations have been particularly challenging for blockchain-based networks in the healthcare sector, which require fast adoption of new technologies. Besides the performance bottleneck, the capacity issue with blockchain should be seriously considered. As the size of a blockchain expands, the amount of storage required by all blocks expands as well. As a result, complete nodes, which keep all the network's block data, demand a lot of storage space (Sun Y et al., 2017). Similarly, as the blockchain history grows, the Bootstrap time will climb linearly, slowing the process of new nodes joining the system. All these constraints reduce a blockchain's availability and decentralization and should be carefully considered when creating a large-scale blockchain. Not every entity in such a network needs a comprehensive blockchain ledger. As a result, the strategies should concentrate on interactions between just those in the network who need to know, i.e. on a need-to-know basis. Innovative technologies typically face the challenge of scalability, and blockchain networks are no exception. Scalability is typically measured by factors like throughput, latency, storage, and block size. In order to address this challenge in blockchain networks, various performance metrics such as throughput, consensus latency,

and the number of transactions per unit time should be analyzed. A higher number of verifiers during the block verification phase ensures greater security, but it also results in increased latency. Healthcare demands a high level of security with minimal verification time.



CHAPTER THREE

RESEARCH METHODOLOGY

3.1 INTRODUCTION

in this chapter will explain the general approach of the proposed system. Which combines x-ray image analysis using Artificial Intelligence methods for prediction of group of common diseases together with the E2E Encryption of data exchange throughout blockchain which based on patient-centric control. The following sections will explain the proposed system in detail.

3.2 OVERVIEW OF THE PROPOSED SYSTEM

The proposed system has multiple stockholders that can request access to EHR. Patient who owns his records only can determine who will grant or revoke permission access in own medical records. This is patient centric control which who have full control on his own records.

In addition, for enhanced security and a decentralized approach, medical data will be stored on the Interplanetary File System (IPFS), which is a protocol for distributing files. This protocol enables all computers worldwide to store and share files as part of a vast peer-to-peer network. Any computer, anywhere in the world, can install the IPFS program and begin hosting and sharing files.

If anyone runs IPFS program on computer and downloaded the file to the IPFS network, then can anyone in the world view or download the file who also running IPFS For that, to keep EHR in more secure and to provide high degree of confidentiality, The medical data exchange will made by hyper encryption End-two-End methods using AES-ECC. The following sections will explain the proposed system steps in detail.

3.2.1 Overview Blockchain and EHR management based on patient centered control.

The proposed system is designed to provide patients with full control over their records, while ensuring confidentiality, robustness, and security through a permissioned

infrastructure within the Ethereum Blockchain framework. Health information is stored mainly as hashes on the blockchain, while the original data is kept off-chain in IPFS to ensure efficiency and scalability.

To ensure authorized access, will create a smart contract protocol called the Patient-Centric Healthcare Data Management Access Control-Smart Contract, which uses role-based access control chain code for authorized stakeholders.

The protocol does not involve any incentive mining, and it ensures equal access for all parties. During registration, unique role-based IDs are generated for stakeholders, and each is provided with a public and private key pair for secure health information storage and transfer.

In the proposed healthcare data management system, doctors are responsible for creating a patient's health record, which is then securely stored off-chain in IPFS and its hash value is stored on the Ethereum blockchain for permanent record keeping.

If doctors need to make changes to a patient's record, they can be granted or revoked access to do so. The patient-centric view of the health record is created from IPFS, allowing doctors to update the documentation before the patient commits to the update using their key pairs to store the updated files in IPFS.

Patients can selectively grant access to relevant stakeholders to view the records in a patient-centric view from the IPFS system.

To ensure data privacy, a doctor's session would expire before the hash value is committed to the ledger, preventing unauthorized access by uninvolved personnel. Smart contracts are also created for various healthcare procedures in the system's backend.

The use of role-based access control ensures the protection of patient privacy, while the system's scalability and interoperability features make it superior to existing systems.

3.2.2 A Background of the Proposed System

The proposed system in Figure 3.1 illustrates the general architecture and access control for sharing EHRs using blockchain technology. To ensure privacy, the EHR data is

encrypted with the EHR owner's secret key and stored in IPFS. The secret key is encrypted using a public key and stored on the blockchain for authentication purposes. The EHR owner and healthcare professionals can access the essential data. This blockchain-based security framework satisfies the requirements for shared EHR systems by protecting patient privacy and ensuring data integrity.

Figure 3.1 compares the proposed system with the conventional EHR system.

In the conventional EHR system:

- (1) The patient sees the physician (Doctor).
- (2) The patient is then given medical attention by the doctor.
- (3) The doctor uploads the EHR to the server following treatment.
- (4) The doctor can download the EHR for later use.

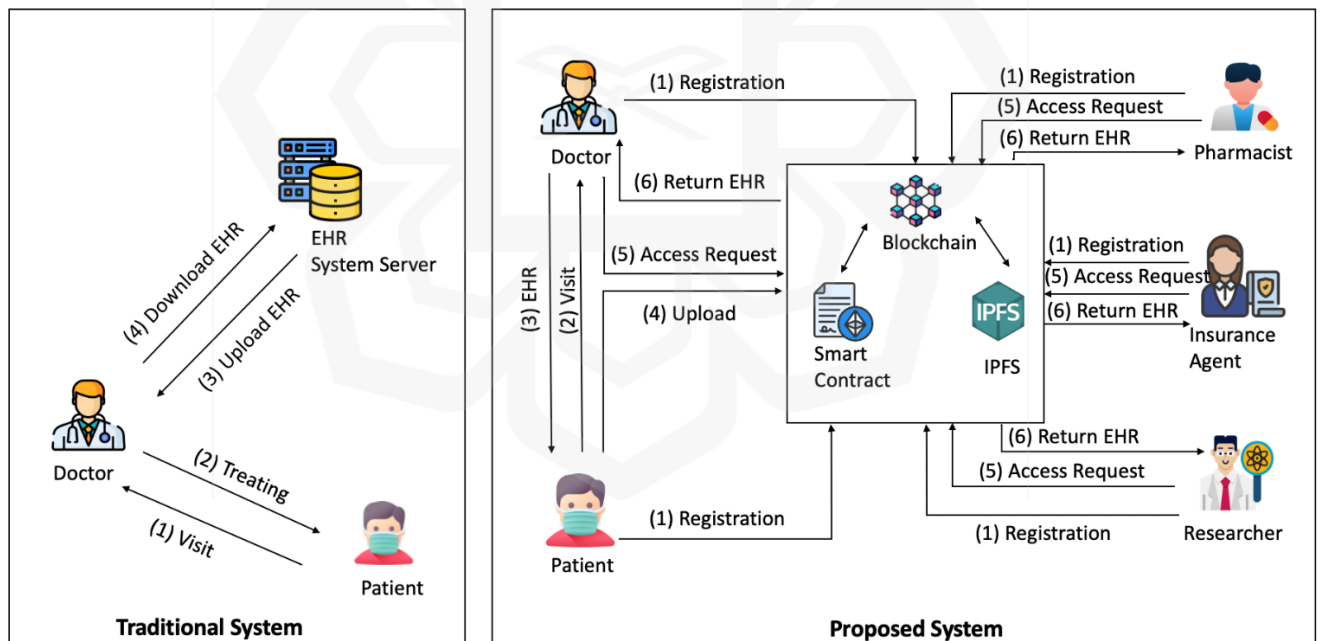


Figure 3.1 Overview of the traditional and the proposed solution

Our strategy might be described as revolutionary secure electronic health records for patients that they can privately share. where the patient can independently manage, download, and trade his or her EHRs.

In our proposed study:

- (1) To produce and save keys, all users must be registered with the system.
- (2) The patient seeks medical attention from a doctor at a hospital or other healthcare facility.
- (3) He or she acquires their electronic health records, which will contain their private health information and be created following the assessment.
- (4) The patient submits the healthcare data to the blockchain and uploads the encrypted files straight to IPFS.
- (5) The user asks access to the file (which might be a researcher, pharmacist, or doctor).
- (6) The user obtains the original file from the IPFS after receiving the encrypted data.

As a result, only the patient and others involved in the access control procedure can view the patient's EHR documents. Moreover, a doctor is only permitted to look through the EHR of patients that they have already seen. Last but not least, a doctor or other users can only access an EHR if the owner has given them permission to do so.

The encrypted text in the electronic health record is being attempted to be decoded. Health records may be taken, changed, or falsified by a malevolent enemy. IPFS and the data requestors will agree to derive the EHR's plain text. The security targets are as follows according to the threat model:

- Data privacy: The original EHR of the owner cannot be revealed to unauthorised people.
- Data authenticity: The patient's EHR can be verified by those who have access to the data.
- Integrity: Patient electronic health records (EHRs) can be stored securely to prevent tampering.
- Data confidentiality: Patient electronic health records are kept secret from outsiders and stored securely.

- Customizable access control: Patients can choose how to access their EHRs, and only individuals with permission can do so.
- Authentication: Users must first authenticate themselves in order to access EHR.

Algorithm 1 is used to enable patients to grant access to their health record to doctors through the use of PCEHRM-SC. This algorithm ensures that only specific fields of the health record are viewed and updated, rather than granting unrestricted access to the entire record. The patient-centric view is generated, and a session key (Sk) is created for use by both the patient and the doctor during the session. The session key is encrypted using the public keys of the patient and the doctor. Algorithm (1) calls the create_Update_HR() function from Algorithm (2) to initiate the update of the health record. The doctor's and patient's session keys are decrypted, and the modifications are uploaded into the updated patient-centric view (UP_Pacenvn). Once the update is completed, the health record HRn is saved in IPFS after the patient approves the changes. The Sk and Pacvn are then terminated, and the IPFS generates the health record hash value HRn_hash, which is saved in blocks within the Ethereum blockchain.

Algorithm 1: System Function()

Input: Doctor D_n , with their Public key D_{pubk_n} , with their Private key D_{prk_n} , with session key S_k of HR_n Health Record. Patient Pa_n with their Public key Pa_{pubk_n} , and Private key Pa_{prk_n} .

Output: Boolean (True or False)

1. Function for storing and updating health records.
2. **For** user U have Access permission to HR
3. Check PCEHRM-SC
4. **If** (permission=="Grant" && role=="Doctor") then
5. Create $Pacenv_n$ for HR_n in IPFS
6. $Pacenv_n \rightarrow$ Decryption (Encryption (HR_n))
7. Create S_k
8. send Encrypted ($Pa_{pubk_n}(S_k)$, $D_{pubk_n}(S_k)$, $Pacenv_n(S_k)$) to Pa_n ,
9. D_n and $Pacenv_n$.
10. create Update HR()
11. $HR_n \rightarrow$ [(Decryption Pa_{pubk_n} (Encrypted Pa_{pubk_n} (HR_n)) + Encryption
(UP $Pacenv_n$))]
12. $Pa_n \rightarrow$ Commit (IPFS (HR_n))
13. IPFS \rightarrow HR_n_hsh
14. $HR_n_hsh \rightarrow$ Ethereum Blocks
15. **Return** True
16. **Else**
17. Permission=Deny
18. **Return** False
19. **End if**
20. **End For**
21. **End Function**

3.2.2.1 Actors

Our architecture contains the following components: stakeholders which they are (patient, doctor, pharmacist, analyst, physician, or researcher) as shown in Figure 3.2, IPFS, blockchain, an encryption mechanism, smart contract, electronic health record, and a web portal.

Actors. In the proposed system, the actors concerned are:

- **Patient:** a person seeking medical attention from a doctor in a hospital or other healthcare facility. He or she has access to their electronic health records, which will contain their private health information generated following consultation and treatment. Therefore, the patient's profile, address, and location, diagnoses, physician recommendations, notes for the next review, names of the doctors, medicine, scan, and test results are all contained in the medical records.

Table 3.1 Patient roles.

Patient	Grant-Revoke-Commit, Read Record
	Revoke permission from Doctor/Service Providers.
	Permission to Doctor to Read/write of their her.
	Able to search for Doctor/Labs.

- **A doctor** can add an observation, if necessary, issue prescriptions with the patient's permission, and examine the patient's medical record via a web application.

- **Other users** who wish to examine the information in the medical record have gained permission from the relevant data owner, such as pharmacies, labs, insurance, or researchers. For instance, the pharmacist gets patient prescriptions to provide patients their medications.

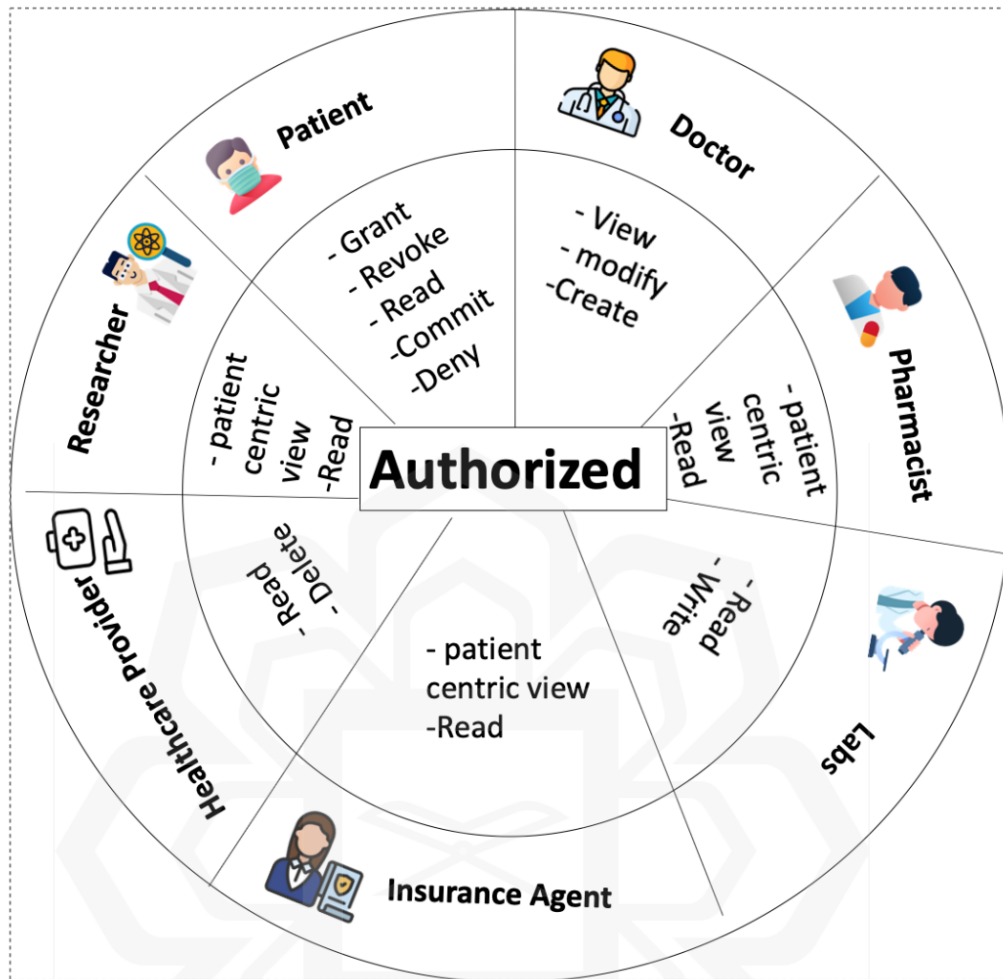


Figure 3.2 Actors and rule-based access in the proposed system.

3.2.2.2 Electronic Health Record (EHR).

Sharing EHRs across different healthcare organizations is crucial for effective healthcare coordination and management. EHRs serve as a comprehensive data source containing information from multiple clinicians, making them transportable and accessible to patients. Protecting the privacy and security of EHRs is paramount, and distributed EHR sharing can empower patients to take control of their healthcare by giving them easy online access to their medical information. By sharing their EHRs with other clinicians, patients can participate actively in the coordination of their treatment and management of their health information.

Algorithm 2: create_Update_HR ()**Input:** $D_n, D_{pubk_n}, D_{prk_n}, S_k$ **Output:** Storage of HR

1. Function Doctor D_{pubk_n}
2. **For** Doctor with D_{pubk_n}, S_k
3. $D_n \rightarrow \text{Decrypt}(D_{pubk_n}(S_k))$
4. $D_n \rightarrow \text{Decrypt}(Pacenv_n(S_k))$
5. $Pacenv_n \rightarrow UP_Pacenv_n$
6. IPFS Encrypt($UP_Pacenv_n(S_k)$)
7. **End For**
8. **End Function**

3.2.2.3 Blockchain

The healthcare industry can benefit from blockchain technology in various ways. Our approach to using blockchain technology ensured that patient data was kept transparent, decentralized, and immutable. Nevertheless, blockchain also provides privacy and confidentiality by concealing individuals' identities through complex and secure mechanisms to protect the confidentiality of medical data. The decentralized nature of blockchain allows patients, doctors, and other healthcare professionals to share information efficiently and securely. The Ethereum blockchain was chosen because it supports smart contracts and may be used as a foundation for decentralised apps. The idea of smart contracts was originally introduced on the Ethereum blockchain platform, which explains why decentralised healthcare applications based on smart contracts are so popular.

3.2.2.4 Website Portal

This portal serves as the first level of security. By linking a username and password, it provides access to specific features and EHR data. Patients will have access to some of the information that healthcare practitioners send them as well as information about their own health data. Depending on their position within the EHR process, other actors have access to the applications created just for them. In our proposed system the meaning of web portal

is using truffle and ganache and metamask and web3 applications, these applications I will explain them it in the following section.

3.2.2.5 Smart Contract

A smart contract is a self-executing code that is installed into the blockchain to execute a specific task when certain conditions are met. It allows for transactions and agreements to be made without the need for a central authority or external enforcement, providing a decentralized system. The Smart Contract has been selected to link with the blockchain and healthcare providers for managing the patient's healthcare information as per their requirements. It also verifies the user's access rights and authentication. The Smart Contract plays a crucial role in executing the agreement between different parties in the system, making it essential for implementation. A smart contract may be created by creating the codes, and these codes describe the contract that the patient has signed. Once it has been approved, the contract transmits a transaction to be added to the blockchain. With our system, smart contracts are created using the Solidity programming language, and then they are deployed on the Ethereum test network. At the conclusion of the section, the code smart contract will display.

3.2.2.6 Access control

guarantees the privacy and accuracy of EHR. Only authorised healthcare professionals and patients should be able to access medical information on our suggested system. Patients should control how their data is collected and who has access to it. These access decisions are made by the patient listed in the smart contract. The smart contract will deny any access requests from unauthorised parties, and the system will be shut down as a result.

The patient's consent and approval access is required before the user can retrieve health data. The only individual who can add another person to view the patient's record is the patient. To ensure that the doctor has access to the patient's data, the smart contract checks the access only before receiving the health data. If the doctor doesn't have access control, the system delivers a bogus message and ends the session.

Scenario

To communicate with the system, all actors use the web portal. To produce and store keys, each user needs to be registered with the system. In Figure 3.3, several interactions are shown:

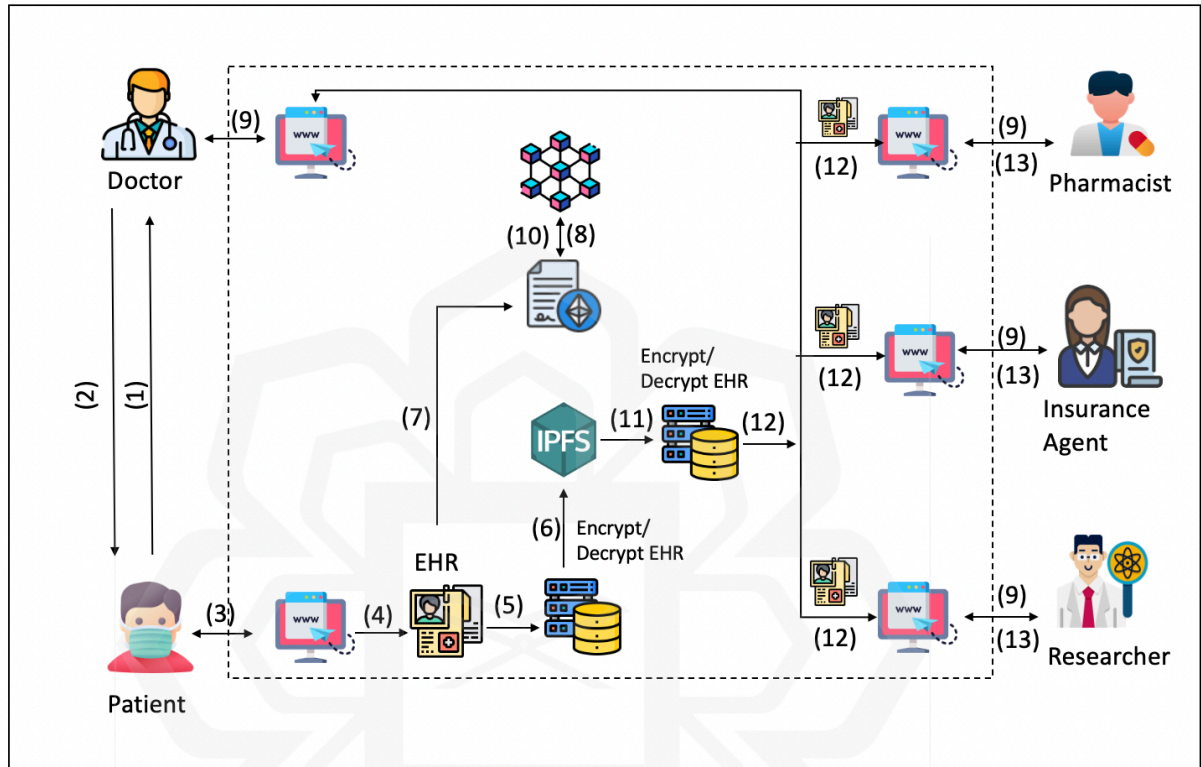


Figure 3.3 Interactions in the system

- 1) Patients see doctors in hospitals or other healthcare facilities.
- 2) The patient receives access to his or her electronic health records, which include any private health information created following encounters with the doctor.
- 3) The patient enters his user ID and password to access the portal.
- 4) The patient adds the system with his EHR.
- 5) The patient adds a user's access rights in accordance with the user's role to his or her EHR.
- 6) The EHR will be encrypted and posted to the IPFS in encrypted form.
- 7) The patient uploads the key that has been encrypted along with other data to the blockchain.

- 8) All requests for transactions are logged to the blockchain.
- 9) In order to retrieve EHR, users submit an access request to the system with necessary data.
- 10) The smart contract authenticates the user with access and decrypts the encrypted key.
- 11) Pulls encrypted EHR from IPFS and decrypts it.
- 12) The user, a doctor, downloads and consults the EHR.
- 13) Depending on his or her access privileges, the user can modify the patient's electronic health record.

3.2.2.7 Hybrid Encryption.

Data is encrypted and decrypted using the same key in symmetric encryption. Although very quick, algorithms using this method are not as secure as those using asymmetric encryption. Asymmetric encryption is thought to be more safe because it does not require key sharing, but it is slower and takes longer to complete. We decided to combine symmetric and asymmetric encryption as a result. Because symmetric encryption is needed to convert plaintext to ciphertext, hybrid encryption is necessary. This utilises the speed of symmetric encryption. The symmetric key is encrypted using the asymmetric key to take advantage of the security of asymmetric encryption, making sure that only the intended recipient can decrypt the symmetric key.

3.2.2.7.1 Definition of ECC and AES

ECC (Elliptic Curve Cryptography)

The use of elliptic curve cryptography (ECC) is a well-known method of encrypting data to prevent unauthorized access. ECC employs pairs of public and private keys to ensure data security. This technique uses two-dimensional fields such as binary and prime fields to provide security, and its enhanced operations and relation between binary and primary fields make it difficult to hack. ECC is characterized by its small key size, and the appropriate field for cryptographic implementation is determined by the maximum number of points. ECC reduces the complexity of operations and is mainly used for key generation. Compared to other cryptographic techniques, ECC has a higher level of enhancement due to its small key size.

- **AES**

The symmetric key cryptography was developed by Joan Daemen and Vincent Rijmen, two Belgian cryptographers (S. N. Mendonca, 2018). usually used to encrypt or decrypt large amounts of data more quickly. This is due to the fact that it uses the same key for both encrypting and decrypting operations rather than generating a new one.

The AES encrypts 128-bit blocks of data using 128-bit keys with 10 encryption rounds, 192-bit keys with 12 encryption rounds, or 256-bit keys with 14 encryption rounds (E. Conrad et al. 2017). It has been demonstrated to have a higher level of security than DES or 3-DES, come with a larger key size, and encrypt communications more quickly (T. B. Azad, 2008). The following methodical processes are used in the encryption and decryption operations: To decode it, perform byte substitution, shift rows, mix columns, add a round key, and finally the opposite of these operations.

1) Encryption Method

Each round of the encryption process is composed of the 4 sub-processes as the following:

- **Byte Substitution**

Data substitution using a substitution table is the first adjustment made to the AES encryption algorithm. The fixed table, called the Substitution Box (S-box), which contains every likely combination of eight-bit order, replaces the 16 input bytes. The new 16 bytes that result are set up in a matrix with four columns and rows (S. N. Mendonca, 2018).

- **Shift Rows**

The second transformation involves shifting data rows. In a matrix created from the byte, each row is shifted or altered to the left. Any entries that go off to the right are reinserted there. The first row is left in tact, and the second row is moved to the left by a byte. The third row then moves leftward by 2 locations, while the fourth row moves leftward by 3 (byte) positions. The same 16 bytes are then used in the resulting matrix, although in different locations (S. N. Mendonca, 2018).

- **Mix Column**

The third transformation now combines the columns. Then, using a special Galois field mathematical function, each of the four-byte columns is altered (GF). The function inputs 4 bytes to represent one column and outputs 4 new bytes to represent the new column. The action is not carried out in the last round (S. N. Mendonca, 2018).

- Insert Round Key

The last, most straightforward modification involves using a separate portion of the encryption key to perform an exclusive OR operation on each column. Following the mix column stage, the resulting matrix's 16 bytes are regarded as 128 bits. At this point, a 128-bit round key is bitwise EX-ORed with the 128-bit state. If it's the last line, the output is the encryption text. If not, the resulting 128 bits will be read as 16 bytes and a new process of substitution would likely begin. It is a column-wise action between the state column's four bytes and the round key's one word (S. N. Mendonca, 2018).

2) Decryption Method

Until plaintext or the original data is obtained, this includes performing the opposite of the encryption procedure.

- Insert a round key

The round keys are chosen in reverse order for this function's own inverse. –

- Inverse Shift Row (b)

In reverse order, the inverse shift row functions in the same manner. The first row is left in place, and the second, third, and fourth rows are then moved to the right by one byte, two bytes, and three bytes, respectively.

- Inverse byte replacement

Utilizing the inverse s-box substitution table, this is accomplished.

- The Inverse Mix Column

Utilizing polynomials of less than 4 degrees over the Galois field (GF) 28, whose coefficients are the components of the state column, the reverse mix column transition is accomplished.

3.2.2.7.2 A Proposed Hybrid Encryption Approach ECC-AES

The two separate algorithms are combined to increase security and reduce the possibility of data loss due to hackers. The ECC method replaces the AES encryption by encrypting the AES key in the cloud, as indicated in the block diagram in Figure 3.4, Figures 3.5. The AES algorithm will first be implemented.

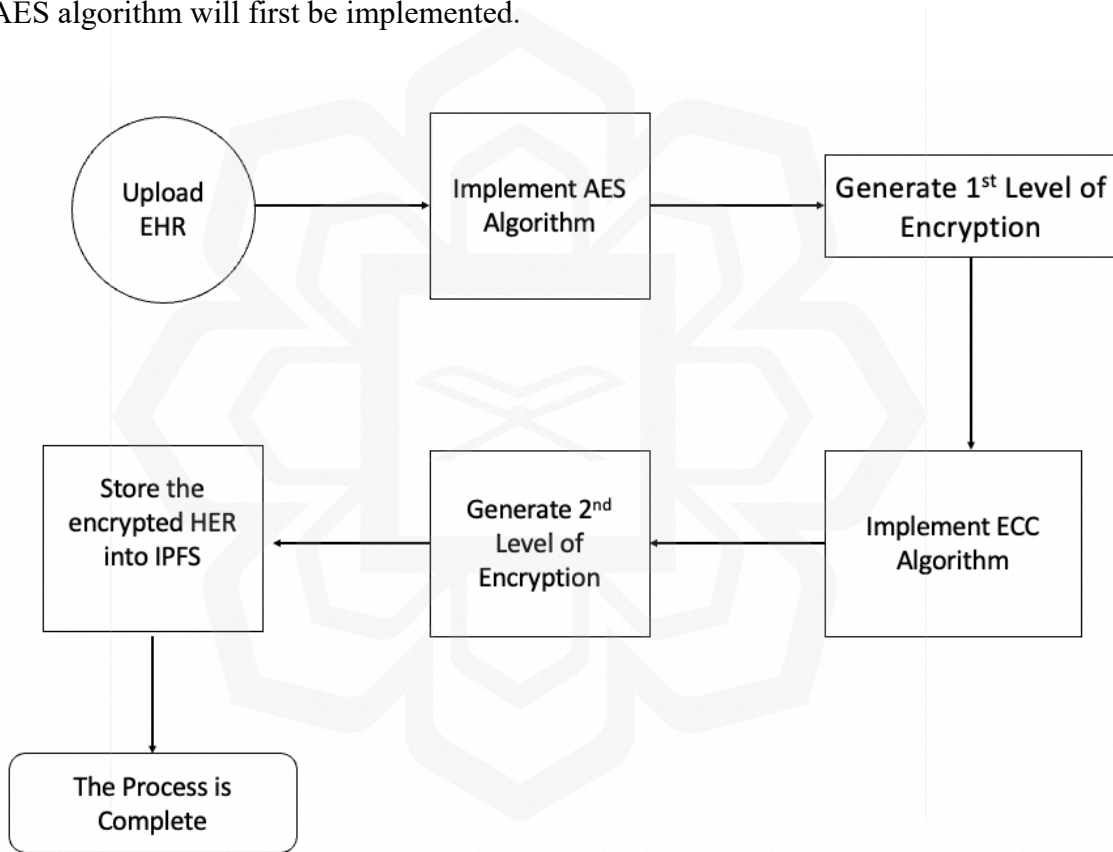


Figure 3.4. Hybrid Encryption steps

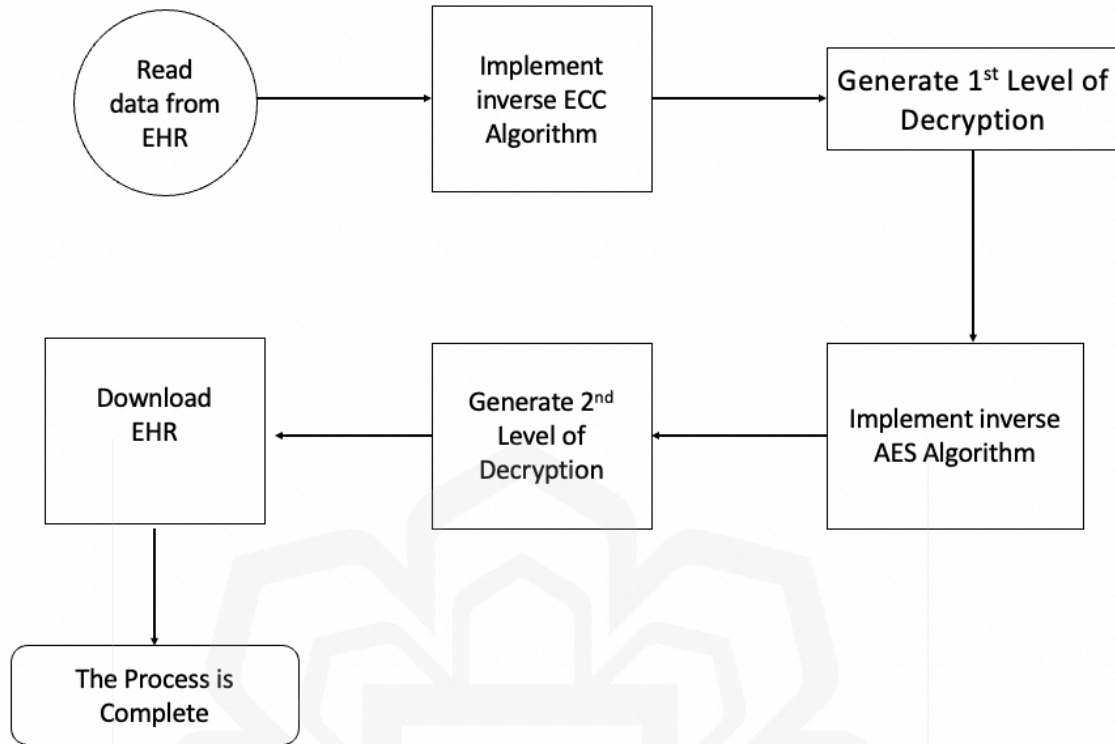


Figure 3.5 Hybrid Decryption steps

ECC and AES are used in combination to create an efficient cryptographic technique for secure cloud storage (IPFS). The hybrid (ECC-AES) method is faster and has a smaller key size than using a single AES due to ECC's small key size feature. ECC employs encryption and decryption key standards to establish a secure key system and reduce key size, making it ideal for use with AES in protecting data from unauthorized access (Mendonca & S.N., 2018). Once the key size is determined, ciphertext is generated using AES for data encryption and decryption with the key generated by ECC.

The proposed technique utilizes the combined effect of ECC and AES to create a secure system for cloud storage, which helps in reducing the size of secure data storage. Figure 3.6 shows the block diagram of the proposed algorithm.

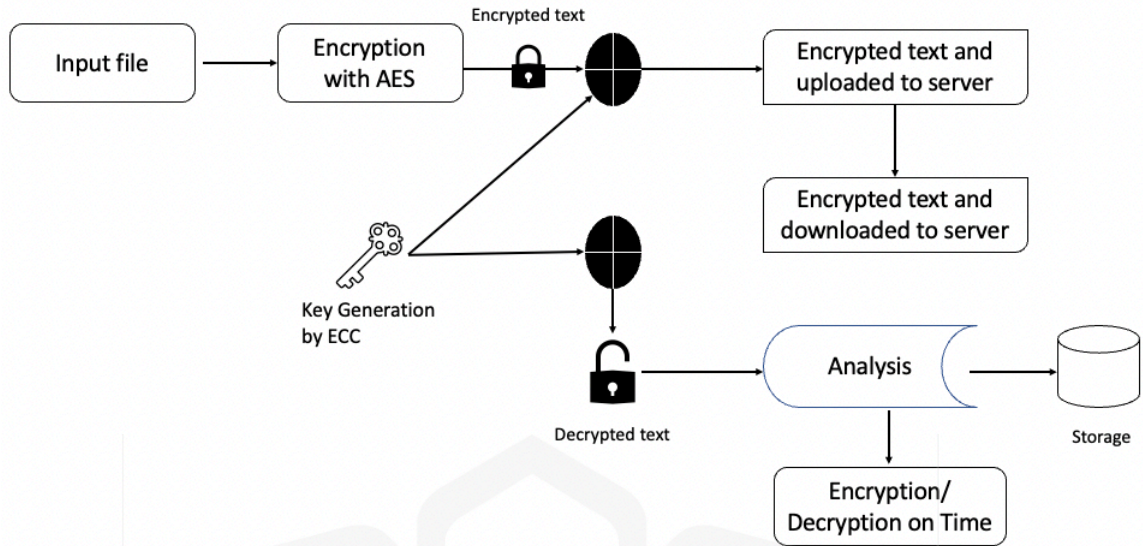


Figure 3.6 ECC and AES algorithm.

The above diagram clearly shows how AES, in conjunction with ECC, effectively secures data stored in the cloud.

The proposed method is a novel approach for secure data transmission and storage. The system diagram provides an innovative solution for protecting user data during transmission to the server and secure storage of the encrypted data.

The novelty of the method can also be evaluated based on its computational cost and time. To prevent attacks, the user's personal information is first encrypted using AES encryption when they upload the input file, making the text fully encrypted.

This ensures that the information is protected in case an attacker tries to access it. Additionally, even if an attacker does manage to obtain the encrypted file, they can't decrypt it, thereby safeguarding the data from attacks.

3.2.2.7.3 Implement the Hybrid encryption algorithm.

Firstly to implement the hybrid encryption mechanism in our proposed system, users should register unique accounts and create their keys. First, a symmetric key (SK) 128 bits in

length is generated for each patient' EHR (EHR_P). The pseudo-random number generation (PRNG) algorithm known as SHA1PRNG used by the SUN provider produces the SK as its result. To create a continuous stream of random numbers, the hash function is utilised. The EHR is encrypted using the SK of the AES advanced encryption standard. To complete data sharing transactions, each user on the blockchain obtains key pairs Public key (PuK), Private key (PrK), by hashing a random number (RN) using the 256-bit SHA-1 hash method. SK was encrypted and the original EHR was signed using the key pair of elliptic curve cryptography (ECC), an asymmetric key authentication method.

As demonstrated in Equations (1) - (2), the EHR P generates all the keys, encrypts the EHR using the SK to produce the ciphertext CEHR, and then encrypts his or her symmetric encryption key using the public key PuK to produce the ciphertext key CK.

$$CEHR = \text{Enc}_{EHR} (IPFS, SK) \quad (1)$$

$$CK = \text{Enc}_{key} (SK, PuK) \quad (2)$$

The encrypted EHR is then hashed using Equation (3), where MD stands for message digest, and then signed. The MD is then signed using the private key, and the digital signature is the encrypted hash (SIG). Equation describes how the patient EHR sends the encrypted EHR (CEHR) to the IPFS once the signature method is finished (4).

$$MD = H(CEHR) \quad (3)$$

$$SIG = (MD, PrK) \quad (4)$$

Afterwards, he or she transmits to the blockchain both SIG and encrypted keys (CK). Also, as mentioned in 3.2.2.5 above, he or she transmits the smart contract the access permissions. For instance, the system database contains the public keys for each user. The SK will be re-encrypted using the public key of C if B (the owner) wishes to share data with C (add C to the list of authorised users). C can use its private key to decrypt the data when it needs to access it. No one else is able to decrypt the data since only C has access to C's private key. In Algorithm 1, the storage procedure is displayed.

Algorithm 1. Data storing.

1. Input EHR_i , Access control, PuK , PrK , SK , SHA-2
2. For each EHR data do
3. Use SK to encrypt EHR $C_{EHR} = Enc_{EHR}(EHR_i (i \in [1;8]), SK)$.
4. Use PuK to encrypt SK , $CK = Enc_{Key}(SK, PuK)$.
5. Use SHA-2 to create MD on encrypted EHR, $MD = H(Enc_{EHR})$.
6. Use PrK to sign MD, $SIG = (MD, PrK)$.
7. Store user's PuK in the system's database.
8. Upload C_{EHR} to the IPFS.
9. Upload CK and SIG to the BC.
10. End for;
11. Output $CEHR$, CK , SIG .

To enable the safe sharing of EHR, the EHR owner predefines access rights in smart contracts, including access privileges, access actions, and access rights (like read and write).

The smart contract is activated as soon as the access need is satisfied, ensuring the accuracy and justice of the data sharing to carry out the associated procedure. The following two steps make up the EHR sharing process:

A. Blockchain access:

- *EHR access request:* The process of requesting access to EHR involves initiating a blockchain network transaction called EHR exchange request (Req) by the user (U), as per the equation. This request must contain access target (ID), access EHR_i , and PrK (5). The blockchain network will then receive and authenticate the transaction request to verify the identity of the EHR user (U). The transaction data will be recorded on the blockchain, ensuring transparency and accountability, and only the authorized user (U) will have access to it.

$$\text{Req} = (\text{ID} \parallel \text{EHR}_i \parallel \text{PrK}); i \in [1 ; 8]) \quad (5)$$

- *Smart contract execution:* Equation describes how the SK will be given to the user after being encrypted using the U's private key if the Req is valid (6).

$$\text{SK} = \text{DecCK} (\text{CK}, \text{PrK}) \quad (6)$$

B. IPFS storage EHR sharing:

The U will recover the EHR_i from the IPFS, as shown in Algorithm 2. The U then generates an MD2 hash of the encrypted EHR to ensure its validity and integrity, as illustrated in Equation (7). The SIG is then decrypted using the EHR P's public key, and the outcome is displayed in Equation (8).

$$\text{MD2} = \text{H}(\text{CEHR}) \quad (7)$$

$$\text{DecSIG} = (\text{SIG}, \text{PK}) \quad (8)$$

The EU will decrypt the EHR and carry out its access action, as specified in Equation, if this decrypted MD matches MD2, indicating that the signature is valid (9). If not, the user can alert the system to the possibility that the data has been altered.

$$\text{EHR}_i = \text{DecCEHR} (\text{CEHR}, \text{SK}) \quad (9)$$

Algorithm 2. Data sharing (Decryption).

1. Input SK, PrK.
2. If Req is not valid then
3. 'return failure'.
4. Else
5. Decrypt EncKey, SK = DecCK (C_k, PrK).

6. Retrieve CEHR from the CS.
7. Create MD2= H(CEHR).
8. Decrypt SIG, DecSIG = (SIG, PK) to get the MD.
9. IF the two MD do not match then
10. 'return failure'.
11. Else
12. Decrypt CEHR, EHR_i= Dec_{CEHR} (CEHR, SK).
13. End If
14. End If
15. Output EHR_i.

3.2.2.8 IPFS for storing the health data

The Interplanetary File System (IPFS) is a distributed file storage protocol that allows computers all over the globe to store and serve files as part of a giant peer-to-peer network. In our proposed system IPFS used to store encrypted EHRs uploaded by the owner.

The code in implement Hybrid AES-ECC to secure EHR and achieve the goal of the proposed system:

Key Generators

1. Symmetric Key

// generate Symmetric Key

```
public String generateSK() throws Exception {
```

```

/*
The SK is the output of the hash function (SHA-1) employed by the SUN provider's
pseudo-random number generation (PRNG) algorithm termed SHA1PRNG. The hash
function is used to generate a stream of random numbers. The SK of an advanced
encryption standard (AES) is used to encrypt the EHR
*/
KeyGenerator keygen = KeyGenerator.getInstance("AES");
SecureRandom secureRandom = SecureRandom.getInstance("SHA1PRNG",
"SUN");
secureRandom.setSeed("EHR 3.0".getBytes());

//generate 128bit random key
    keygen.init(128, secureRandom);
    SecretKey sk = keygen.generateKey();
    return Base64.getEncoder().encodeToString(sk.getEncoded());
}

```

2. ASymmetric Key

```

function getKeys() {
    // A new random 32-byte (256 bits) private key
    const pr = eccrypto.generatePrivate();
    // Corresponding uncompressed 65-byte (520 bits) public key
    const pb = eccrypto.getPublic(pr);
    return { pr: pr.toString("hex"), pk: pb.toString("hex") };
}

```

Encryption

1. Create Cipher C_{EHR} (AES)

$CEHR = EncEHR (IPFS,SK)$

```

public String encryptEHR(String _SK, String EHR) throws Exception {
    byte[] decodedKey = Base64.getDecoder().decode(_SK);
    SecretKey SK = new SecretKeySpec(decodedKey, 0, decodedKey.length, "AES");
    Cipher cipher = Cipher.getInstance("AES");

    // Encrypt SK

    cipher.init(Cipher.ENCRYPT_MODE, SK);
    byte[] cipherEHR = cipher.doFinal(EHR.getBytes());

    //return Cipher SK
    return Base64.getEncoder().encodeToString(cipherEHR);
}

```

2. Encrypt Symmetric Key CK (ECC)

ECIES

```

// Encrypy SK using Public Key of Doctor

function encryptSK(_sk, _pb) {
    let pub = Buffer.from(_pb, "hex");
    let str = eccrypto.encrypt(pub, Buffer.from(_sk)).then((e) => {
        let enc = {
            iv: e.iv.toString("hex"),
            ephemPublicKey: e.ephemPublicKey.toString("hex"),
            ciphertext: e.ciphertext.toString("hex"),
            mac: e.mac.toString("hex"),
        };

        return JSON.stringify(enc).toString("hex");
    });

    // return ck in string format

```



```
return str;
```

```
}
```

3. Message Digest of CEHR

ECDSA

```
// Returns Message Digest
```

```
crypto.createHash("sha256").update(_CEHR).digest();
```

4. Sign Message Digest (MD) using Private Key of Patient /

```
// Sign Message digest using ECDSA
```

```
    eccrypto.sign(Buffer.from(_PRK, "hex"), H).then((sign) => {  
      _sign = Buffer.from(sign).toString("hex");  
      return _sign;  
    });
```

Decrypting

1. Decrypt Ck

ECIES

```
function decryptSK(_ck, _prk) {  
  let encJ = _ck;  
  let encB = {  
    iv: Buffer.from(encJ.iv, "hex"),
```

```

ephemPublicKey: Buffer.from(encJ.ephemPublicKey, "hex"),
ciphertext: Buffer.from(encJ.ciphertext, "hex"),
mac: Buffer.from(encJ.mac, "hex"),
};
// B decrypting the message.
let sk = eccrypto.decrypt(Buffer.from(_prk, "hex"), encB).then((s) => {
  return s.toString();
});

return sk; }

```

2. Verify signature

ECDSA

```

function verifySign(_sign, _cehr, _pub) { var pb = Buffer.from(_pub, "hex");

var msg = crypto.createHash("sha256").update(_cehr).digest(); var sig =
Buffer.from(_sign, "hex");
let v = eccrypto

.verify(pb, msg, sig) .then(function () {

console.log("Verified");

return true; })

.catch(function () { console.log("invalid"); return false;

}); return v;

}

```

3. Decrypt CEHR

```
public String decryptEHR(String _SK,String _CEHR) throws Exception{  
  
    // decode the base64 encoded string//  
  
    byte[] decodedKey = Base64.getDecoder().decode(_SK);  
    // rebuild key using SecretKeySpec//  
  
    SecretKey SK = new SecretKeySpec(decodedKey, 0, decodedKey.length, "AES");  
    Cipher cipher = Cipher.getInstance("AES");  
    cipher.init(Cipher.DECRYPT_MODE,SK);  
    byte[] cipherEHR = cipher.doFinal(Base64.getDecoder().decode(_CEHR));  
  
    // return decrypted cipher EHR  
    return new String(cipherEHR);  
}
```

Implement smart contract code:

```
contract Contract {  
    // uses Roles library form openzeppelin for role based access control  
    using Roles for Roles.Role;  
    // mainly three roles admin,doctor and patient  
    Roles.Role private admin;  
    Roles.Role private doctor;  
    Roles.Role private patient;  
  
    struct Doctor {  
        // doctor details  
  
        address id;  
        string prk;  
        string sk;  
    }  
}
```

```

struct Patient {
address id;
string patHash;
string prk;
string sk;
}
// Struct to store Medical doctors
struct MedRec {
// owner :patient
address owner;
string docID;

string cEHR;

address[] doctors;
}
// struct for sharing EHR
struct ShareDoc {
address owner;
address docID;
string sign;
string ck;
uint8 access;
}

// mappings

mapping(address => Doctor) Doctors;
mapping(address => Patient) Patients;
mapping(address => MedRec) Records;
mapping(address => mapping(address => ShareDoc)) Shares;
address[] public Dr_ids;
address[] public Patient_ids;
string[] public RecordHashes;
address accountId;
address admin_id;
address get_patient_id;
address get_dr_id;

```

```

    constructor() {
// set the account which deploy the contract as admin
// In this case first account on ganache
        admin_id = msg.sender;
        admin.add(admin_id);
    }

//get Admin

    function getAdmin() public view returns (address) {
        return admin_id;
    }

//Add Doctor

    function addDoctor(
        address _newdr,
        string memory _prK,
        string memory _sk
    ) public {
        require(admin.has(msg.sender), "Only For Admin");
        doctor.add(_newdr);
        Doctor storage doc = Doctors[_newdr];
        doc.id = _newdr;
        doc.prk = _prK;

doc.sk = _sk;

        Dr_ids.push(_newdr);
    }

// get Doctor

    function getDoctor(address _docID) public view returns (Doctor
memory) {
        return Doctors[_docID];
    }
    function delDoctor(address docID) public {

```

```

require(admin.has(msg.sender), "Only For Admin");
doctor.remove(docID);

}

// check is Doctor
function isDr(address id) public view returns (string memory) {
    require(doctor.has(id), "Only for Doctors");
    return "1";
}

// Check is Patient
function isPat(address id) public view returns (string memory) {
    require(patient.has(id), "Only for Doctors");
return "1"; }

// Add patient
function addPatient(address _newpatient) external onlyAdmin {
    patient.add(_newpatient);
}

// Get Patient Information => return IPFS hash of patient details
function getPatInfo(address id) public view returns (Patient memory)
{
    return (Patients[id]);
}

// Add patient Information to BlockChain
function addPatInfo(
    address pat_id,
    string memory _patInfoHash,
    string memory _prK,
    string memory _sK
) public {
    Patient storage patInfo = Patients[pat_id];
    patInfo.id = pat_id;
    patInfo.patHash = _patInfoHash;
    patInfo.sk = _sK;
}

```

```

    patInfo.prk = _prK;
    Patient_ids.push(pat_id);
    patient.add(pat_id);
}

// Add Medical record to block chain
function addMedRecord(
    address _pat_id,
    string memory _doc_id,
    string memory _cEhr
) public {

    require(doctor.has(msg.sender) == true, "Only Doctor Can Do
That");
    MedRec storage record = Records[_pat_id];
    record.owner = _pat_id;
    record.docID = _doc_id;
    record.cEHR = _cEhr;
}

// View Medical record returns EHR for given patient id
function viewMedRec(address id) public view returns (MedRec memory)
{
return (Records[id]);
}

//Share EHR with given docID with encrypted SK and signature
function shareMedRec(
    address _owner,
    address _docID,
    string memory _sign,
    string memory _ck,
    uint8 _access
) external onlyPatient {
    ShareDoc storage share = Shares[_owner][_docID];
    share.owner = _owner;
}

```

```

share.docID = _docID;
share.sign = _sign;
share.ck = _ck;
share.access = _access;
Records[_owner].doctors.push(_docID);
}
// get Shared Document for Doctor
function getPatRecord(address _owner, address _docID)
    public

view

    returns (ShareDoc memory)
    {
        return Shares[_owner][_docID];
    }
// get shared Doctors for given patient
function getSharedDoctors(address _id)
    public

view

    returns (ShareDoc[] memory)
    {
        MedRec storage rec = Records[_id];
        ShareDoc[] memory _shareDocs = new
ShareDoc[](rec.doctors.length);
        for (uint256 i = 0; i < rec.doctors.length; i++) {
            if (Shares[_id][rec.doctors[i]].owner != address(0)) {
                _shareDocs[i] = Shares[_id][rec.doctors[i]];
            }
        }

    }

    return _shareDocs;
}
//update shared access for shared EHR
function updateAccess(
    address _owner,
    address _docID,

```



```

    uint8 _access
) public {
    Shares[_owner][_docID].access = _access;
}
//delete shared doctor access
function deleteAccess(address _owner, address _docID) public {
    delete Shares[_owner][_docID];
    MedRec storage rec = Records[_owner];
    for (uint256 i = 0; i < rec.doctors.length; i++) {
        if (rec.doctors[i] == _docID) {
            delete rec.doctors[i];
        }
    }
}

//update Shared Signature
function updateSharedSign(address _owner, string memory _sign)
public {
    MedRec storage rec = Records[_owner];
    for (uint256 i = 0; i < rec.doctors.length; i++) {
        if (rec.doctors[i] != address(0)) {
            Shares[_owner][rec.doctors[i]].sign = _sign;
        }
    }
}
}
}

```

3.2.2 AI Model

The main goal of our proposed system based on Blockchain and AI and how can integrating these two techniques to build immutable and secure system of patient's medical information. Furthermore, we improve the system with integrating both technologies for making a significant difference in healthcare. Deep learning (DL) and artificial intelligence (AI) give potential to develop solution that address highly particular business requirements. Deep learning in healthcare has significantly improved clinic support and changed patient care overall. Deep learning is being used more frequently to identify clinically significant elements in photos that go beyond what the human eye can see.

Chest X-ray images are commonly utilized in clinical practice to diagnose various illnesses, including pneumonia, lung cancer, and abnormalities such as lesions and fractures. The aim of our proposed system is to create a method to identify 14 various chest diseases from an X-ray image as indicated in the following Figure 3.7 Our classifier generates a label vector from an input X-ray image that indicates which of 14 illness groups the image belongs to. The stages that follow will describe how to analyse x-ray images and verify using CNN to predict the disease.

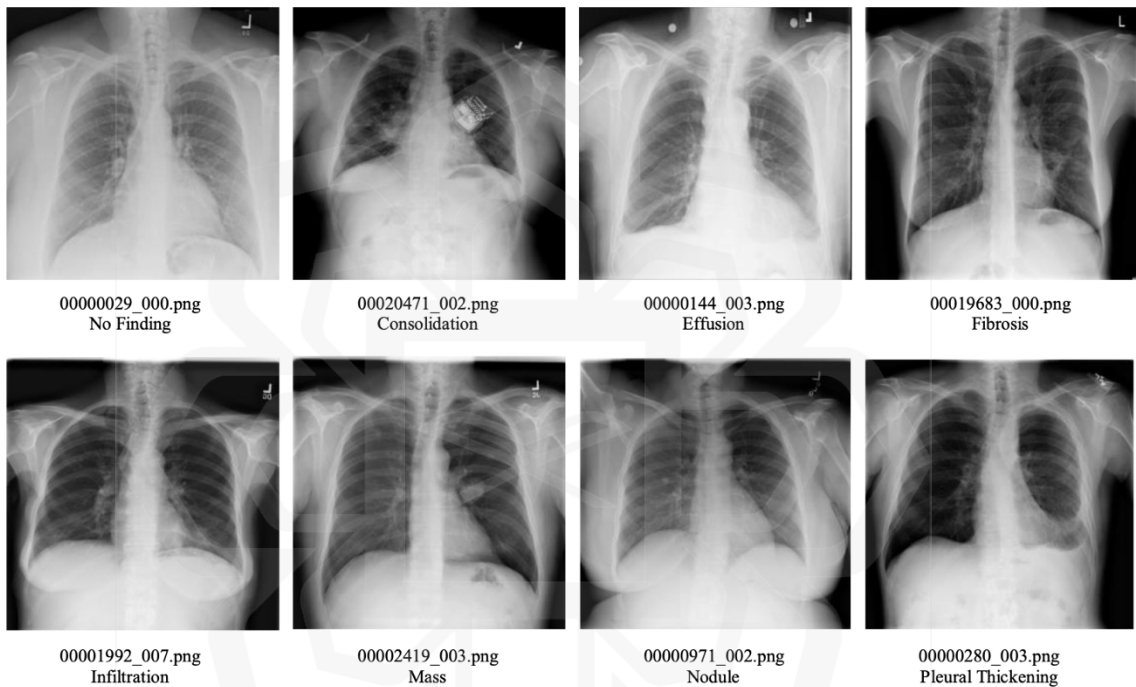


Figure 3.7 Sample Images in NIH ChestX-Ray8 Dataset.

3.2.2.1 Dataset

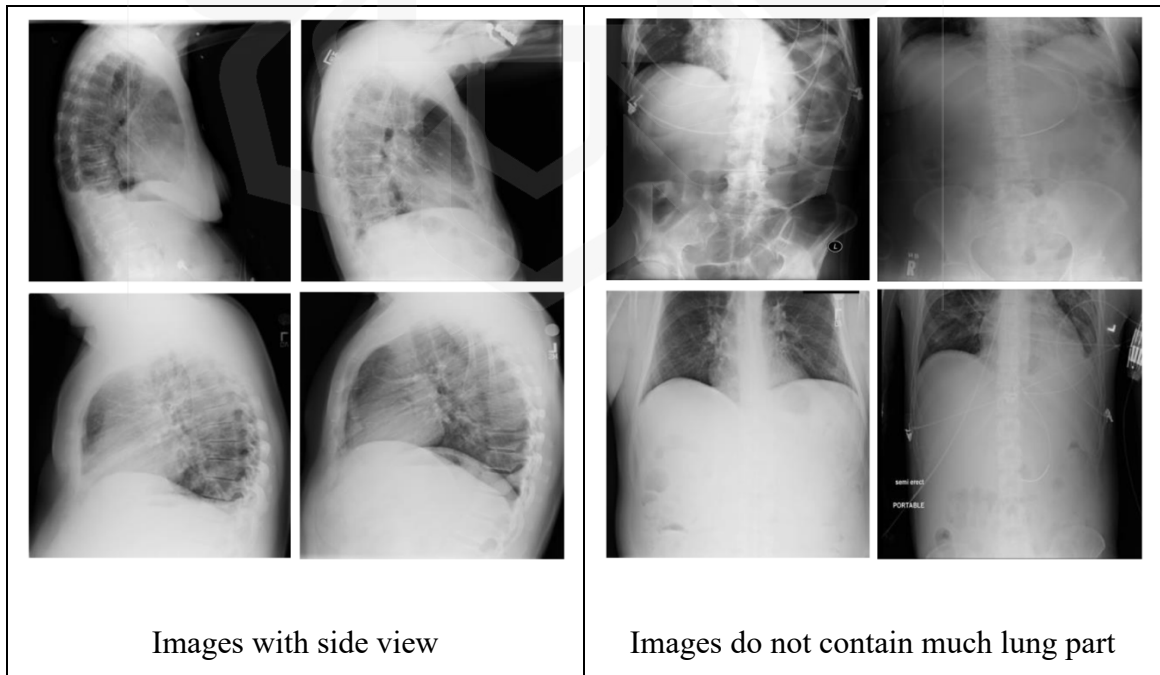
One of the largest public chest x-ray databases for thorax disease detection research purposes to date is the NIH ChestX-Ray8 dataset (Wang et al. 2017). This dataset includes 112,120 frontal view chest x-ray pictures of 30,805 unique people with 14 thoracic diseases that were taken from the clinical PACS database at the National Institutes of Health Clinical Center (atelectasis, consolidation, infiltration, pneumothorax, edema, emphysema, fibrosis,

effusion, pneumonia, pleural thickening, cardiomegaly, nodule, mass and hernia). For convenience, all of the photos in this collection have already been preprocessed to the same size of 1024 1024. This dataset, which contains both typical instances and those connected to TB, includes some sample raw photos in the previous Figure .

Clinical readings sorted by image names and saved in a single Comma Separated Values (CSV) file called "Data Entry 2017" contain patient ID, follow-up number, age, gender, view position, and anomaly information.

The NIH ChestX-Ray8 dataset, which has a big amount of data, extensive annotations, and a wide range of thorax disorders it covered, has been widely used by deep learning researchers for medical applications, although there are still some issues.

The chest x-ray image quality variance is the first and biggest issue, and it significantly adds to the workload associated with data cleaning. The first step is to delete any side views, photographs with little meaningful information in the lung region, rotated images, and images with poor pixel quality. If not, these "poor data" will influence how the deep learning models train, which will affect how well they diagnose diseases in general. Figure 3.8 contains examples of photographs that feature the aforementioned issues.



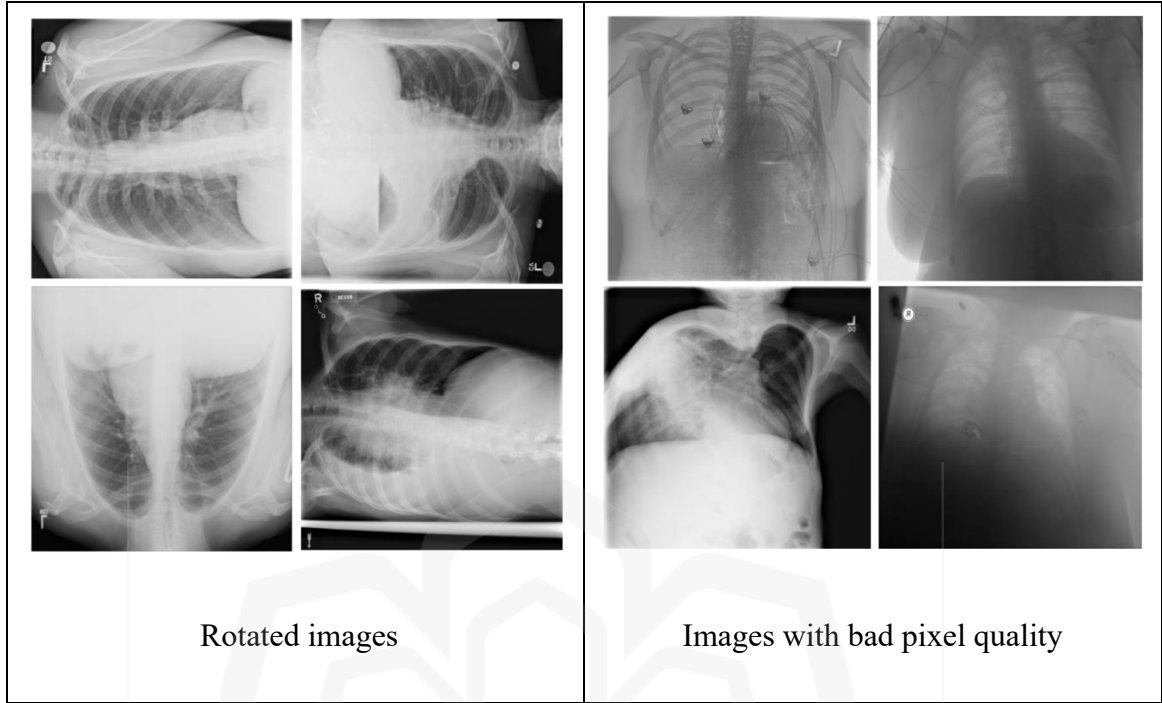


Figure 3.8 sample of images with bad quality in NIH ChestX-Ray8 dataset.

3.2.2.2 Pre-Processing

In order to prepare the image for future processing, it must be resized and normalised at the image pre-processing stage. Based on the need for model creation, several image pre-processing techniques may be discovered in earlier literature. Image scaling, image normalisation, and covert level to category algorithms are among those that are frequently utilised. Using the ImageDataGenerator() Python function, images were scaled in this study to ensure the same size and identical pixel. In this study, photographs with 320 by 320 pixel sizes are taken into account. Also, by adjusting the pixel intensity, we may make the image appear more realistic. To do this, we often use the image's average and standard deviation. This method lessens the computational complexity involved in modelling training. However, images were normalised using Eq. (10).

$$X_{\text{norm}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}, \quad (10)$$

Where X_{\min} and X_{\max} refer to the minimum and maximum pixel values.

After normalisation, we will set "batch_size" which divides the total number of images in train set in every epoch, then set the image size to be 320px by 320px.

By using ImageDataGenerator() function, will build a separate generator for validation and testing data. This is because we want to normalize one image at a time rather than in batches for test and validation data. In the interest of time, we'll take a sample of the dataset to calculate the mean and standard deviation rather than all of it.

```
def get_train_generator(train_df, img_dir, x_col, y_cols, shuffle=True, batch_size=8,
seed=1, t_width = 320, t_height = 320):
    // Returns generator for training set.
    # normalize images
    image_generator = ImageDataGenerator(
        samplewise_center=True,
        samplewise_std_normalization=True)
    # taking the dataframe and the path to a directory + generating batches
    generator = image_generator.flow_from_dataframe(
        dataframe=train_df,
        directory=img_dir,
        x_col=x_col,
        y_col=y_cols,
        class_mode="raw",
        batch_size=batch_size,
        shuffle=shuffle,
        seed=seed,
        target_size=(t_width,t_height))
    return generator
```

3.2.2.3 Classification

Convolutional Neural Network (CNN) is a type of deep learning algorithm which can recognize patterns and features in an input image by assigning weights to different parts of the image. One of the main benefits of CNN is its ability to capture spatial and temporal dependencies in an image using filters that are applied to the image (Simonyan et al., 2014). The CNN architecture is composed of three layers: convolutional, pooling, and fully

connected. The convolutional layer is responsible for extracting features from the input image. It can contain multiple convolution kernels, and its calculations are carried out as follows:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l\right), \quad (11)$$

where x_i^{l-1} is the characteristic map of the output of previous layer, x_j^l is the output of the i th channel of the j th convolutional layer and $f(\cdot)$ is called the activation function. M_j is the subset of input feature maps, k_{ij}^l is a convolutional kernel and b_j^l is its corresponding weight.

3.2.2.3.1 DenseNet

Convolutional networks can be significantly deeper, more precise, and easier to train if they have shorter connections between layers that are near to the input and those that are close to the output, according to recent research.

The Dense Convolutional Network (DenseNet), which connects each layer to every other layer in a feed-forward manner, is a component of our system that takes advantage of this observation.

Our network features $L(L+1)/2$ direct connections as opposed to standard convolutional networks with L layers having L connections, one between each layer and its succeeding layer. The feature-maps of all layers before it are utilised as inputs for each layer, and its own feature-maps are used as inputs into all levels after it.

DenseNets have several advantages, such as solving the vanishing-gradient problem, enhancing the propagation of features, promoting the reuse of features, and reducing the number of parameters. DenseNets have been shown to achieve significant improvements over state-of-the-art models with less memory and processing power required to achieve high performance.

DenseNet is a modern CNN architecture that achieves state-of-the-art performance with fewer parameters for visual object recognition. It shares many similarities with ResNet, but the key difference lies in the way they merge the output of previous layers with subsequent

ones. While ResNet uses additive attribute (+), DenseNet employs concatenated (.) attributes to tightly link all layers, aiming to eliminate the problem of vanishing gradients.

Among the different DenseNet (DenseNet-121, DenseNet-160, DenseNet-201), this study employed DenseNet-121 [5 + (6 + 12 + 24 + 16) × 2] = 121] architecture. Details of the DenseNet-121 is following: 5—convolution and pooling layers, 3—transition layers (6,12,24), 1—Classification layer (16) and 2—denseblock (1 × 1 and 3 × 3 conv).

Generally, traditional CNNs calculate the output layers (*l*th) using a non-linear transformation $H_l(.)$ to the output of the previous layer X_{l-1}

$$X_l = H_l(X_{l-1}). \quad (13)$$

DenseNets concatenate rather than sum up the layer output functionality maps with the inputs. An easy communication model for enhancing information flow across layers is provided by DenseNet: The features of all preceding layers provide input to the *l*th layer: The equation is then changed once more to:

$$X_l = H_l[(X_0, X_1, X_2, \dots, X_{l-1})], \quad (14)$$

where $[X_0, X_1, X_2, \dots, X_{l-1}]$ is a single tensor formed by the concatenation of the output maps of previous layers. Out of the functions, $H_l(.)$ represents a non-linear transformation function. This function consists of three major operations, batch normalization (BN), activation (ReLU) and pooling and convolution (CONV). DenseNet architecture is presented in Figure 3.9. However, the growth rate *k* helps to generalize the *l*th layer in following manner: $k[l] = (k[0] + k(l - 1))$. Where $k[0]$ is known as the number of channels.

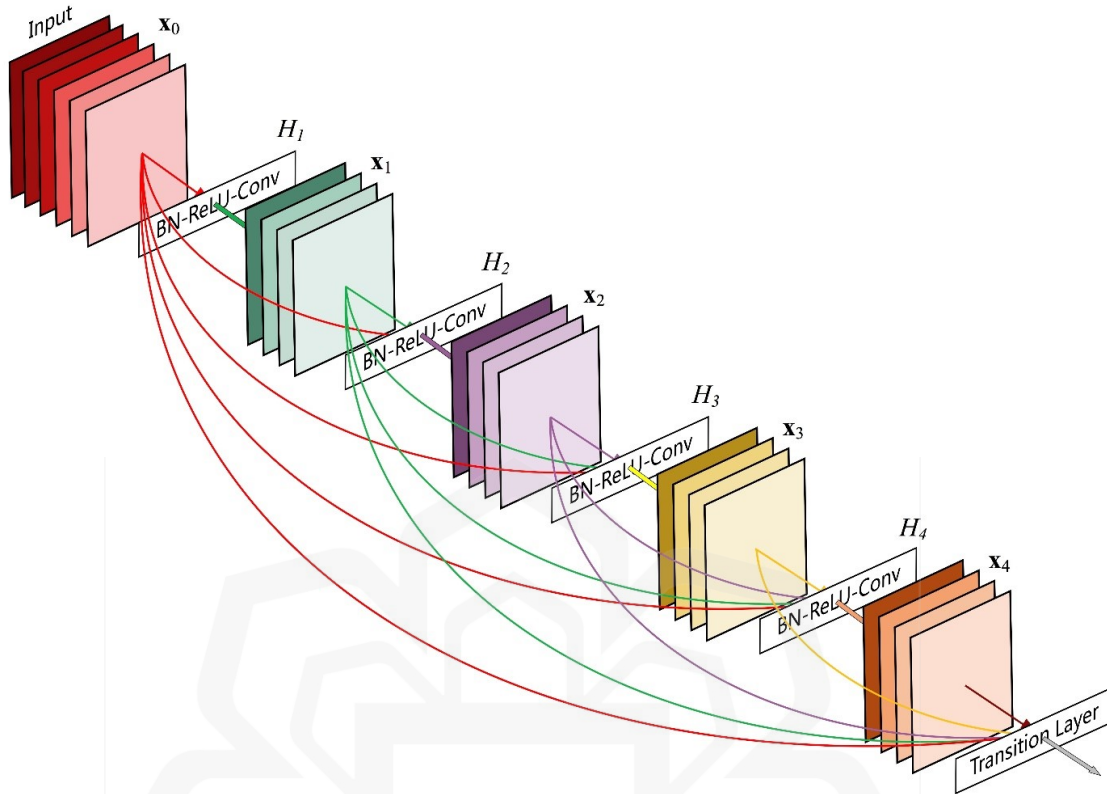


Figure 3.9 DenseNet architecture (Huang G, et al.2017)

Our system will utilize a DenseNet121 model obtained from Keras, which we will enhance by adding two layers on top of it. The first layer, a GlobalAveragePooling2D layer, will enable us to compute the average of the last convolution layers. The second layer will be a Dense layer, with a sigmoid activation function, to make predictions for each of the classes in our dataset.

The implementation code:

```
# create the base pre-trained model
base_model = DenseNet121(weights='./gib/densenet.hdf5', include_top=False)
x = base_model.output
x = GlobalAveragePooling2D()(x)
predictions = Dense(len(labels), activation="sigmoid")(x)
model = Model(inputs=base_model.input, outputs=predictions)
model.compile(optimizer='adam', loss=get_weighted_loss(pos_weights, neg_weights))
```


3.2.3 Tools Used to Implement The Proposed System

In general, our proposed system will be based on Ethereum Platform, Then, when the system is running, and after user login Authentication, the user can upload the medical file, then The uploaded file then will pass through AngularJS as front-end software that contains the user interfaces. After that, the back-end system will be with Python, IPFS will fetch the file from AngularJs. At the same time, AngularJS will connect to the blockchain through Web3 and Ganache blockchain will also integrate to the MetaMask through web3.

Following that, both AngularJs and Ganache will receive the IPFS hash that was previously sent. The user will receive the transaction permission on the Ganache blockchain from Ganache via MetaMask. Ganache will then deploy the smart contract when the patient has given his or her consent. At this point, each transaction will result in the crediting of some gas.

The Encryption process we can execute it by implement the library eccrypto from NodeJS. In addition, AI section we will use Libraries support CNN algorithm to predict and achieved the requirements.

In the following section, we will describe the the tools in details:

3.2.3.1 Blockchain Platform



Figure 3.10 Ethereum Logo.

Ethereum is open source software platform based on blockchain, and it serves the most second largest cryptocurrency on the world after Bitcoin (IHS S., 2015). Ethereum shares data and decentralised public ledger like Bitcoin, however it has expanded its functionality where the user can utilise the platform to build, publish, monetise, and use the applications.

Furthermore, when we talk about the application, Ethereum is not cryptocurrency only, it is a platform for developers to develop decentralized application (DApp) and publish the smart contract.

DApp is like YouTube and Twitter but stores the data in decentralized methods and allows for the user to get access on his own data without any centralized entity. Each of DApp and smart contracts are free from downtime risks, fraud, and being tampered off.

The user of DApps will be charged in terms of Ether while using DApps like file uploading because Ethereum also has its own coin, called Ether (ETH). There will be a fee for each transaction made on the blockchain, including those made on Ethereum. The cost varies according to the amount of data that is uploaded to the programme. The uploaded data will be secured with the gas fee. The term "gas" refers to this computation of execution. Additionally, the maximum amount of gases that can be utilised for each transaction can be specified by the user or account owner.

However, if the gas is insufficient, the user will not be able to use the apps properly even though the gas will still be utilised. On the other hand, if the fee is less than the cost of the gas, the gas will be given back to the owner.

There are five basic parts of Ethereum. The parts are as follows:

- **Programming Language:** Solidity
- **Smart Contract** (explained in Chapter 2): Since Ethereum uses the programming language Solidity, smart contracts are written in Solidity. Cryptocurrency includes smart contracts, and the system will operate in accordance with the contracts that the developer has programmed. Additionally, there shouldn't be any adjustments possible once this smart contract has been implemented on the blockchain.
- **Ethereum Virtual Machine (EVM):** Typically, contracts are written in high-level languages like Solidity, then they are compiled into EVM bytecode. As an EVM instance is run by each node in the Ethereum blockchain, nodes can agree on the

same set of instructions to be executed. In other words, the smart contract will actually be carried out by EVM.

- **Ether:** On the Ethereum platform, transactions and smart contract execution both require the Ethereum token.
- **Consensus Algorithm** (explained in Chapter 2).

3.2.3.2 AngularJs

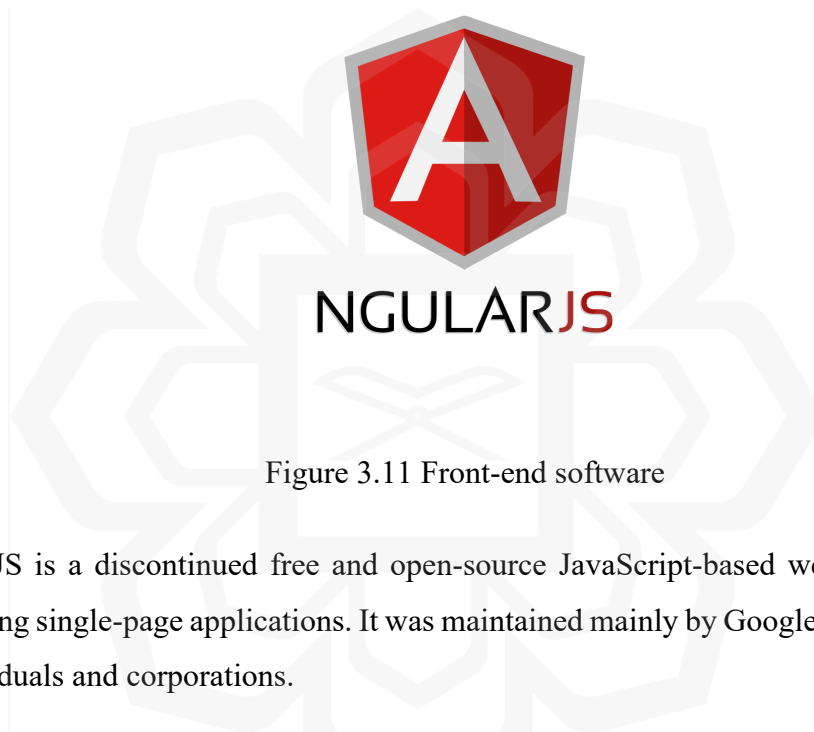


Figure 3.11 Front-end software

AngularJS is a discontinued free and open-source JavaScript-based web framework for developing single-page applications. It was maintained mainly by Google and a community of individuals and corporations.

AngularJS is a toolset for building the framework most suited for our application development. It is fully extensible and works well with other libraries. Every feature can be modified or replaced to suit your unique development workflow and feature needs.

3.2.3.3 Truffle

The developer's ability to create the blockchain application was made easier using truffle. For the Ethereum blockchain using EVM, Truffle is utilised as a development environment, a framework for testing, and an asset pipeline. Additionally, Truffle has built-in smart

contract compilation, linking, deployment, and binary management features. Run the following command on the terminal to instal the Truffle on a global scale.



TRUFFLE

Figure 3.12 Truffle Logo

3.2.3.4 Ganache



Figure 3.13 Ganache Logo

To emulate the behaviour of a public blockchain, ganache is a privalbe Ethereum blockchain. Decentralized apps (DApps), smart contracts, software testing, and state inspection are all possible with Ganache while retaining chain management. The user in Ganache is given ten accounts, each with 100 ETH. Each transaction on the blockchain requires the payment of gas.

3.2.3.5 *MetaMask*



Figure 3.14 MetaMask Logo

The MetaMask extension is obtained for this project by browser the Google Chrome webstore. The Ethereum wallet and blockchain application gateway MetaMask operates as a web browser add-on. With the help of this extension, the individual account might control the price of gas or ether by being linked to the blockchain. In fact, the web browser has evolved into a blockchain browser with the aid of MetaMask.

3.2.3.6 *IPFS*

What follows is a general explanation of how IPFS functions. First, the data will be broken up into smaller pieces after being uploaded to IPFS. In order to prevent network duplication, the system will encrypt (cryptographic hash) each and every data shard. The system will also keep track of each data shard's version history. Later, the network's nodes will individually store the encrypted shards. Additionally, the content IDs from which the Merkle Directed Acyclic Graph (Merkle DAG) will be created are included in the encrypted hash. As a result, this Merkle tree makes it simple to identify archived material using the root hash. Finally, file locations and node connection information are obtained using the distributed hash table (DHT).

There are two types of IPFS data structures that are used for retrieving files: Data and Link. It will handle the massive amounts of unstructured binary data for the data and examine

each file to make sure it isn't more than 256kB. However, Link includes a Link structure array that consists of Name, Hash, and Size. The first data field is called Name and contains the name of the Link. The second data field is called Hash and contains the hash of the IPFS entity that is connected. The last data field is called Size and contains the total size of the IPFS object.

The IPFS can be initialised using either the IPFS Daemon or the IPFS Infura methods. However, in this project, the latter approach was chosen for the system since it is much better in integrating process.



Figure 3.15 IPFS Logo.

3.2.3.7 Web3

The decentralised application (DApp) that powers the blockchain is created using Web3. The non-blockchain website often used Web2, and this website uses an intermediary entity that may not have operated as intended. The following command is entered into the console to install the web3.

```
> npm ls web3
```

CHAPTER FOUR

RESULT AND ANALYSIS

4.1 INTRODUCTION

In order to properly conclude this study, the results are described in Chapter 3 must be analyzed to test the proposed hypothesis for answering the research questions defined in the problem statement and to determine the extent to which we are or are not compromising other performance metrics in achieving our research goals.

4.2 RESULTS OF THE PROPOSED SYSTEM

Firstly, the proposed system will be executed via the command line, both on the server side and the client side. Run truffle migrate because it provides the compiler for smart contracts. We need it to convert the Solidity code into machine-readable code that can be deployed on Ganache blockchain and connect the tools together for our proposed system.

In order to engage with smart contracts on our own private blockchain, we can imitate the Ethereum blockchain using Ganache, a private Ethereum blockchain environment. The following are some characteristics offered by Ganache:

- Shows output from the blockchain log.
- Offers advanced mining control.
- Environment for the Ethereum blockchain; built-in block explorer
- There is a desktop application for Ganache.

Each block in Ganache has public and private key used for connecting purpose with blockchain.

4.2.1 Results of the proposed system

When execute our proposed system. First step as we mentioned in the previous, the patient visits the doctor. The user interface for this step will be shown in the following Figure 4.1

patient ID from Ganache, which is the private key, then can choose the specialty doctor and add doctor's public key, date of appointment and time.

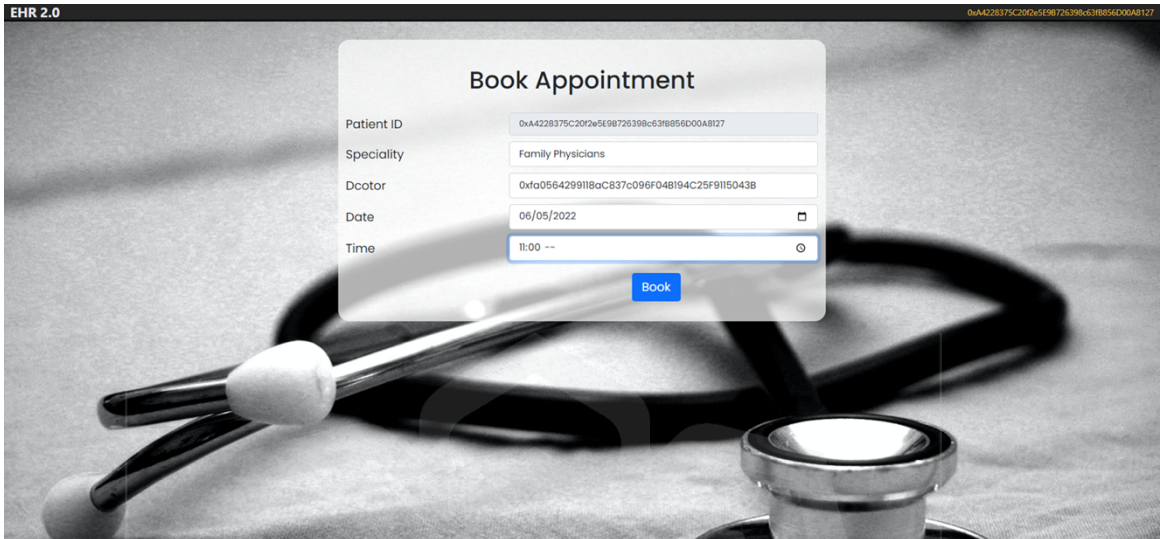


Figure 4.1 . Book appointment interface.

In turn, on the part of the doctor all patients' appointments will be displayed. Once the appointment is chosen, the system sends a request to the patient to grant access, as shown in the following figures.

- If the doctor lacks the authorization to view or modify the EHR, they will not be able to carry out the consultation process. When attempting to conduct a consultation through the EHR, the doctor will receive a warning message indicating their lack of editing access, as shown in the figure 4.2.

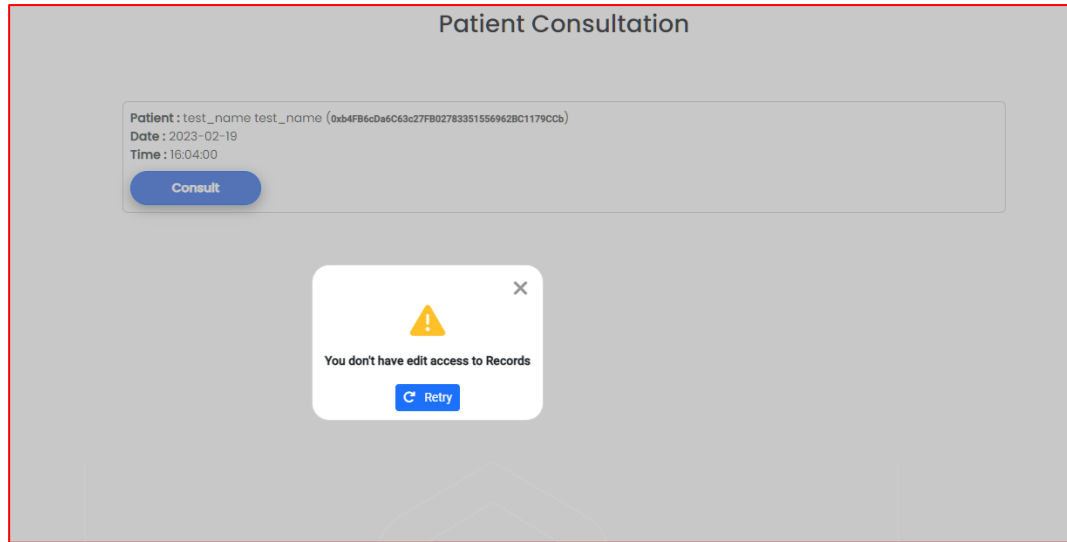


Figure 4.2 Doctor didn't have edit access.

- When the patient manages their own medical records and grants the doctor viewing access, the doctor is only able to view the records for the purpose of reading them. As shown in the Figure 4.3.

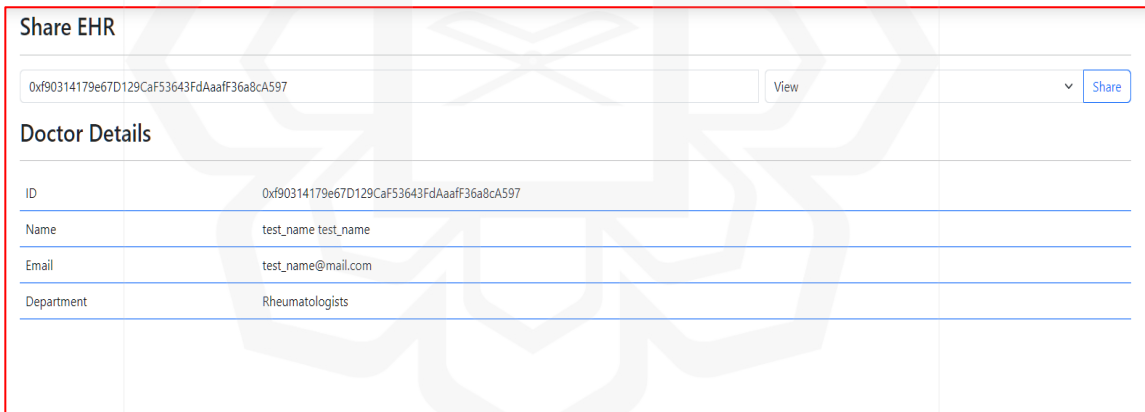


Figure 4.3 Sharing record with view access.

- When the patient is in charge of managing their own medical records and gives the doctor permission to make edits, the doctor can only make changes to the records for the purpose of updating them. This is illustrated in Figure 4.4.

Share EHR

0xf90314179e67D129CaF53643FdAaaff36a8cA597 Edit Share

Doctor Details

ID	0xf90314179e67D129CaF53643FdAaaff36a8cA597
Name	test_name test_name
Email	test_name@mail.com
Department	Rheumatologists

Figure 4.4 Sharing record with edit access.

- **Manage Shared EHR:** Managing a shared EHR involves displaying a list of doctors (identified by their ID) along with their respective access permissions. The patient should be able to remove or modify these permissions as needed. As illustrated in Figure 4.5.

Manage Shared EHR



#	Doctor	Access	Action
1	0xf90314179e67D129CaF53643FdAaaff36a8cA597	Edit	 

Figure 4.5 Manage shared EHR.

- **Update Shared EHR:** After granting the doctor permission to access the medical records, if the patient needs to update the doctor's details, they can do so using Figure 4.6. This allows the patient to modify not only the doctor's information but also their access permissions.

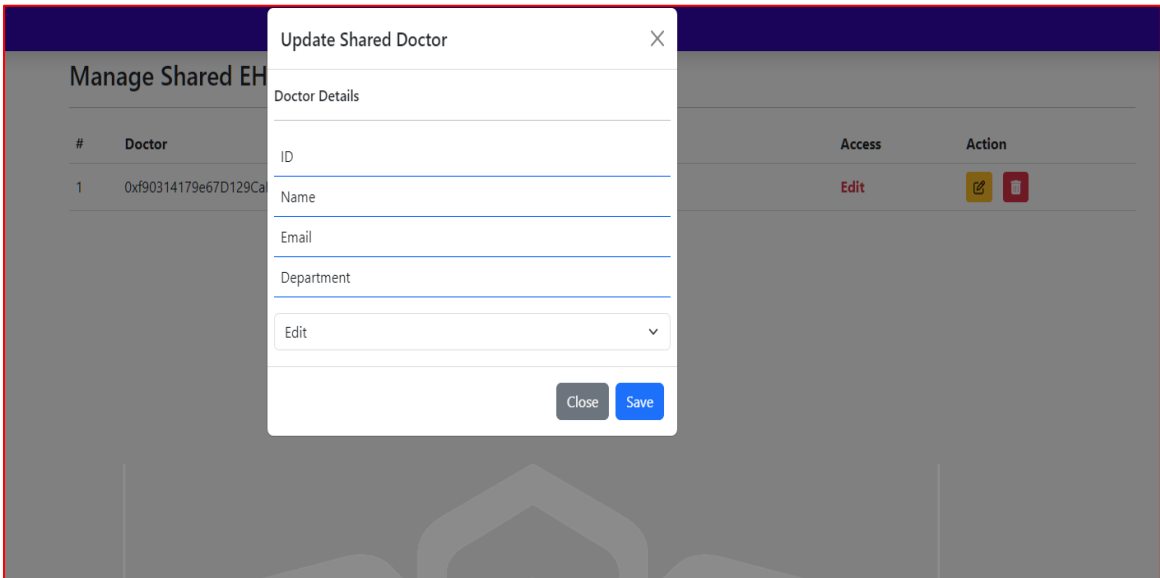


Figure 4.6 Update Shared record.

Once doctor have permission to edit EHR, now to make our proposed system unique from other works in the open current literature, we added a new feature beside the use of Blockchain and the process to manage the EHR which help doctor in diagnosis and save time and cost for decision making process. We use an AI diagnostic tool that we have developed based on convolutions network algorithm DenseNet121 to predict any present disease in the patient's x-ray image. More on this in subsection 4.3.3 below.

The following figures, shows how decryption process is done once access control is granted, in addition to the user interface to edit EHR.

- Doctor receives Signed EHR: Once the doctor has been granted permission to view or edit the encrypted EHR, they can request to obtain the signed version of the EHR and decrypt it. Figure X illustrates the process that the doctor can follow to view the signed EHR.

The screenshot shows a 'Patient Medical Record' interface. The patient profile includes a name, ID, and contact information. Below this is a 'Disease Prediction' section with an 'Upload Chest X-Ray' button and a 'Predict Disease' button. A table for 'Medicines' is partially visible. On the right, a developer console shows a JSON object representing a signed EHR, including fields for 'consultation.component.ts:58', 'doctor.service.ts:97', and 'EHR' data.

Figure 4.7 Access and received signed EHR.

- In the following figure show how signed EHR From Blockchain: To decrypt the encrypted EHR, the doctor's ID must be used to initiate the request. The first step is to verify the validity of the request ID, followed by the decryption of the Symmetric key (CK) using the Doctor ID and the patient's private key with the Symmetric key. The next step is to verify the signature of the EHR, which involves generating the hash of the encrypted EHR. This is a critical step that ensures the validity and integrity of the EHR. The decrypted signature of the EHR is then performed using the patient's public key. The final step involves decrypting the decrypted signature of the EHR and Symmetric key. This is illustrated in figure 4.8.

```

> EHR
< {access: '1', ck: '{"iv":"3a33a4c0994666200c086db47c50f982","ephemPub...2558d5437a740110ed
1ef552d2d887701065fa1d9780d68"}', docID: '0xf90314179e67D129CaF53643FdAaafF36a8cA597', o
wner: '0xb4FB6cDa6C63c27FB027833515569628C1179CCb', sign: '30450221009c379c0987c669fcc34
bc5848ad2a801523e9d27...9590aaed061a47e7a959476fd0a2b39d376db572d55d6f78c'}
  access: "1"
  ck: "{\"iv\": \"3a33a4c0994666200c086db47c50f982\", \"ephemPublicKey\": \"04fbf602c8b6690-
docID: \"0xf90314179e67D129CaF53643FdAaafF36a8cA597\"
owner: \"0xb4FB6cDa6C63c27FB027833515569628C1179CCb\"
sign: \"30450221009c379c0987c669fcc34bc5848ad2a801523e9d2724407ba93aa9443f754a35f002201-
  ▶ [[Prototype]]: Object

```

Figure 4.8 illustrated signed EHR.

- Uploaded Xray images and disease prediction, with accuracy of the result of prediction, shown in the following figures.

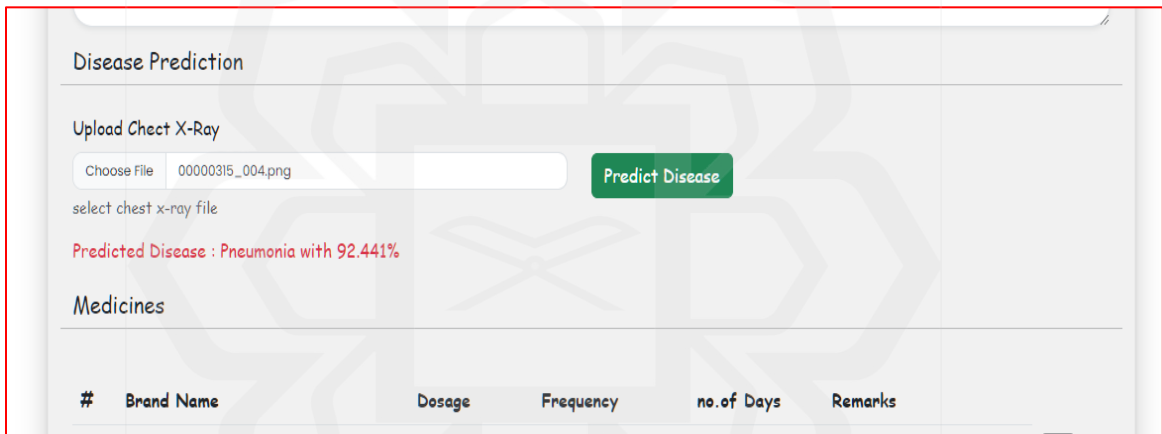


Figure 4.9 uploaded x ray image and predict disease.

Then once the disease prediction is carried out and the doctor updates the patient EHR, one should confirm the consumed estimated gas fees to save the changes on EHR, as shown in following figure from metamask.

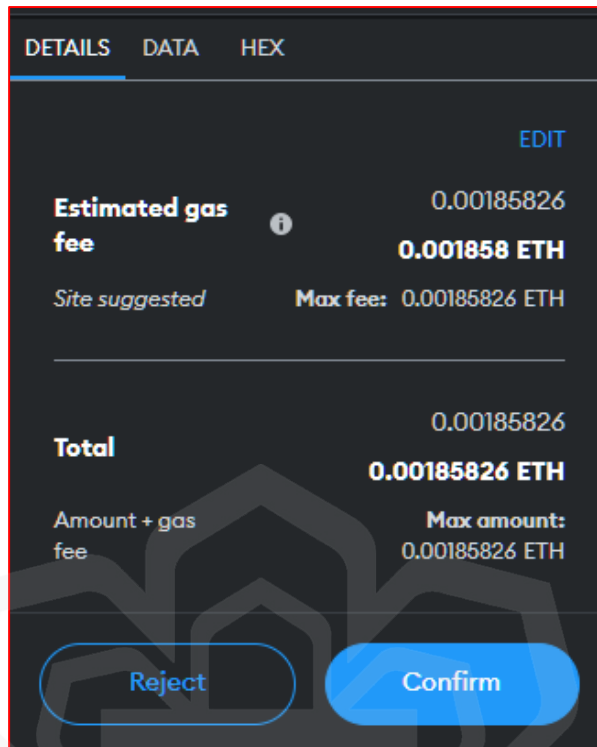


Figure 4.10 Confirms estimated gas consumed to update EHR.

- Stored EHR in IPFS: Figure 4.8 illustrates how storing the EHR in IPFS involves creating a list of doctors who have edited the EHR, including their Doctor ID, as well as a record of the consultation process. This includes a list of medications prescribed, medical reports added to the patient's EHR, and other relevant information.

```
⏪ ⏩ 🏠 🌐 bafybeifdzulghpxern5dl7ywjnznzjmdszw2wjyefpchbirgy4nytmey.ipfs.localhost:8080
{
  "MedRecord": [
    {
      "doctor": "0xf90314179e670129Caf53643FdAaafF36a8cA597",
      "data": {
        "diagnosis": "tetr",
        "medication": [],
        "tests": [],
        "files": [],
        "predictedDisease": ""
      },
      "date": 1676719436116
    },
    {
      "doctor": "0xf90314179e670129Caf53643FdAaafF36a8cA597",
      "data": {
        "diagnosis": "tes",
        "medication": [],
        "tests": [],
        "files": [],
        "predictedDisease": ""
      },
      "date": 1676720216420
    },
    {
      "doctor": "0xf90314179e670129Caf53643FdAaafF36a8cA597",
      "data": {
        "diagnosis": "tesers",
        "followup": "After 1 Weeks",
        "medication": [
          {
            "name": "zsdffsedfzse",
            "dose": 1,
            "frequency": "1-1-0",
            "nofDays": 5,
            "remarks": "cfgyx"
          },
          {
            "name": "dfgs",
            "dose": 1,
            "frequency": "0-1-0",
            "remarks": "47"
          }
        ],
        "tests": [
          {
            "name": "CT Scan - Thoracic"
          }
        ]
      }
    }
  ]
}
```

Figure 4.11 EHR stored in IPFS.

the detailed implementation of the proposed hybrid ECC_AES technique is discussed below, where we should first mention the Postman. Postman allows us to perform different tasks on API requests and test scripts where we might apply encryption and many other possibilities when encryption can be used. then Postman can be called as an API platform for building and using APIs.

we used Crypto.js that is one of the most libraries in encryption and decryption and it supports our hybrid encryption for the proposed system.

The following figures 4.12- 4.17 show how to execute API for the proposed hybrid encryption steps such as: encrypt EHR, Decrypt EHR, Sign EHR,

- 1- Encrypt SK: The symmetric key encryption process involves using the original symmetric key with the patient's public key to generate CK.

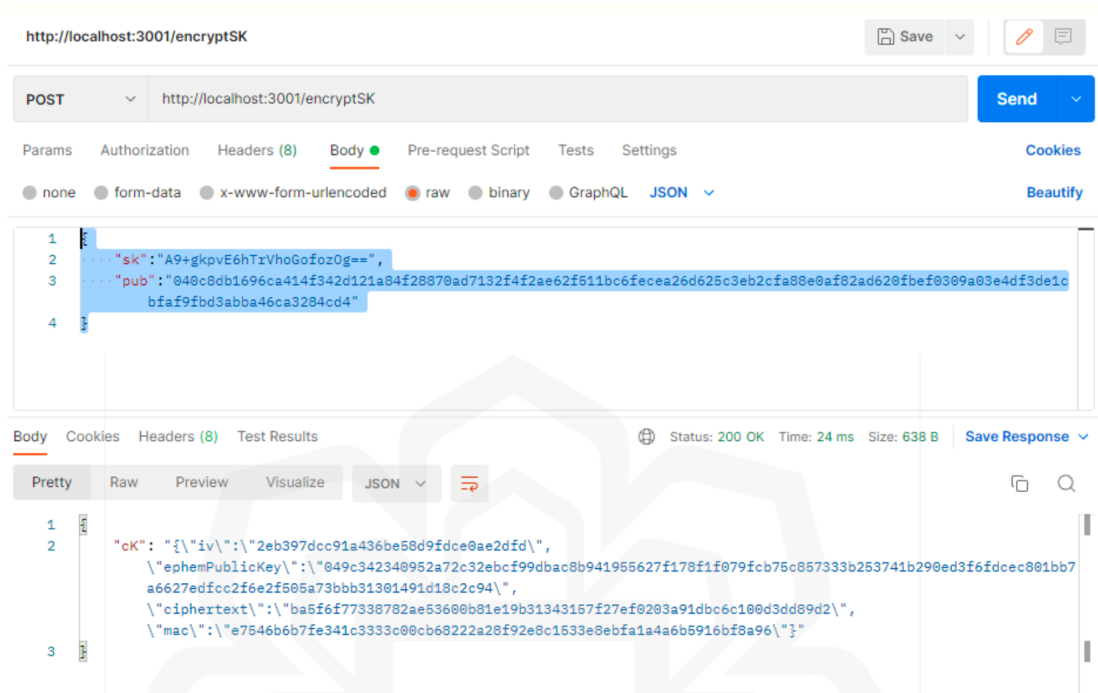


Figure 4.12 Show how encrypt SK.

- 2- Encrypt EHR: The encrypted EHR is generated by using the symmetric key to produce CEHR.

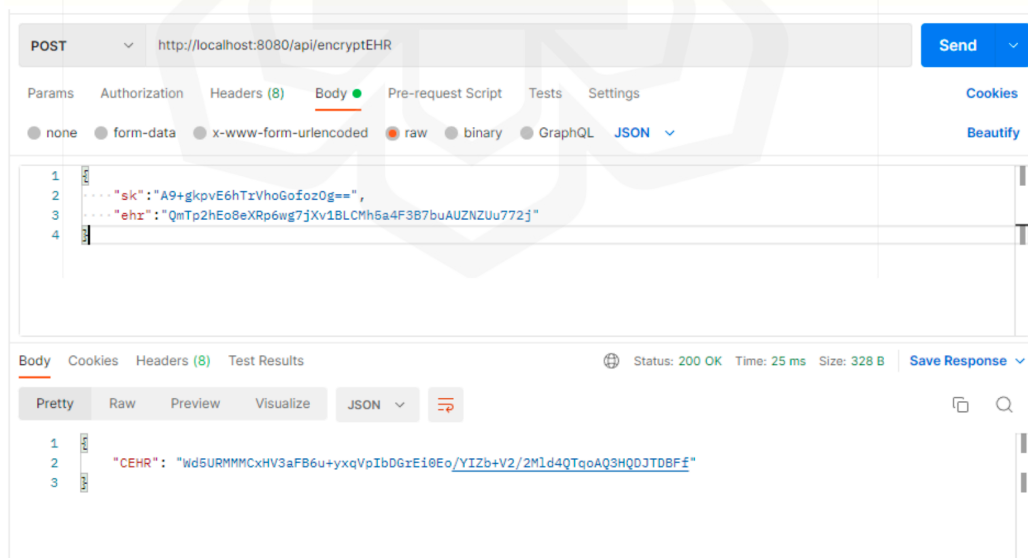


Figure 4.13 show how encrypt EHR.

- 3- Sign HER : To sign the encrypted EHR, it is first hashed to generate the message digest (MD). Using this message digest and the private key, the EHR can then be signed.

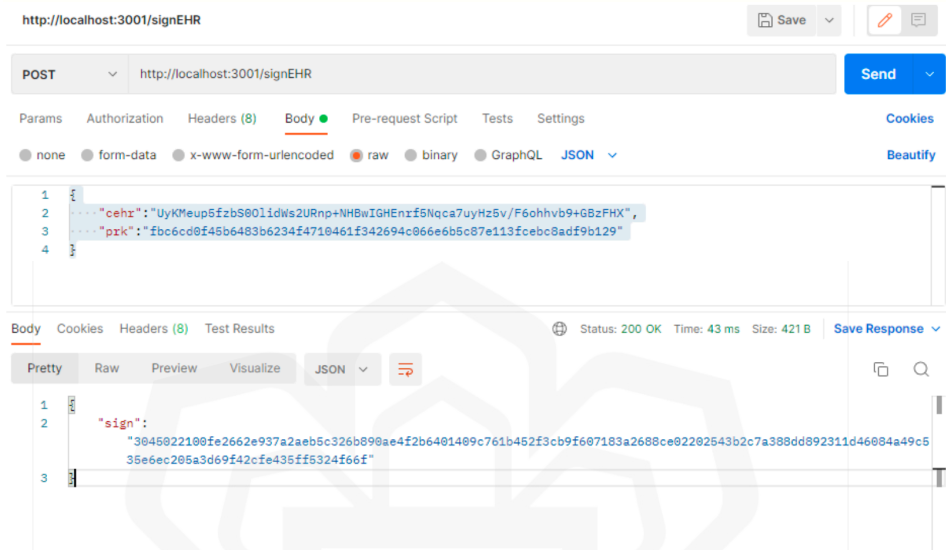


Figure 4.14 show how Sign EHR.

- 4- Decrypt SK : To decrypt the encrypted EHR, the doctor's ID must be used to initiate the request. The first step is to verify the validity of the request ID, followed by the decryption of the Symmetric key (CK) using the Doctor ID and the patient's private key with the Symmetric key.

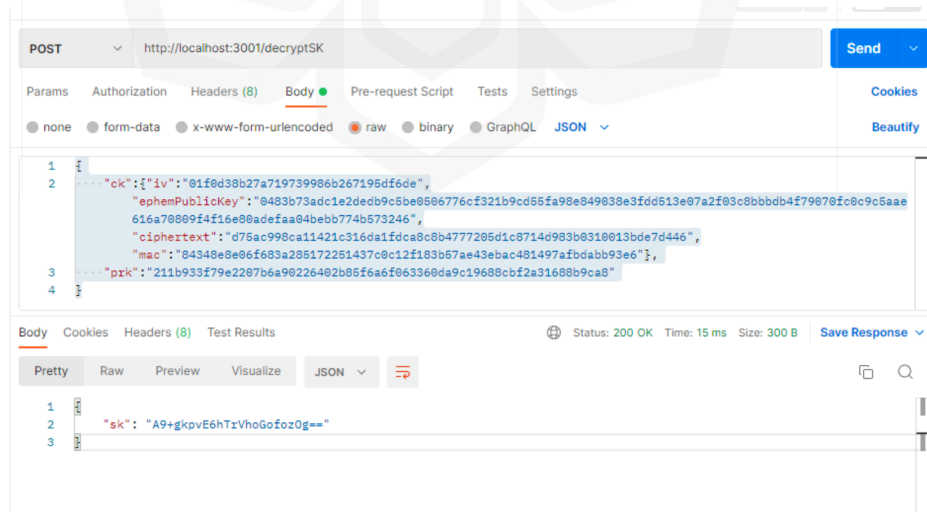


Figure 4.15 show how decrypt SK.

- 5- Verify the sign: The next step is to verify the signature of the EHR, which involves generating the hash of the encrypted EHR. This is a critical step that ensures the validity and integrity of the EHR. The decrypted signature of the EHR is then performed using the patient's public key.

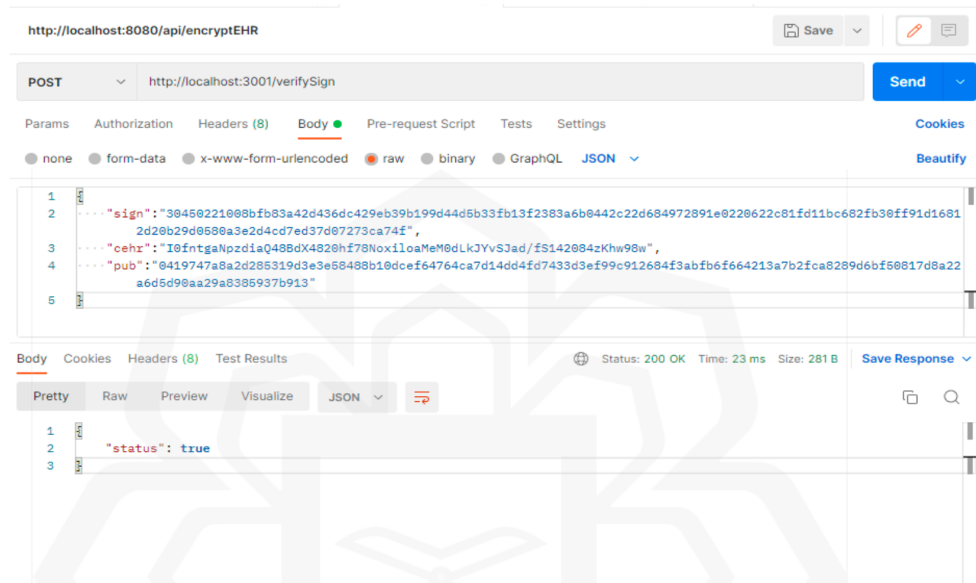


Figure 4.16 show how verify sign.

- 6- Decrypt HER : The final step involves decrypting the decrypted signature of the EHR and Symmetric key.

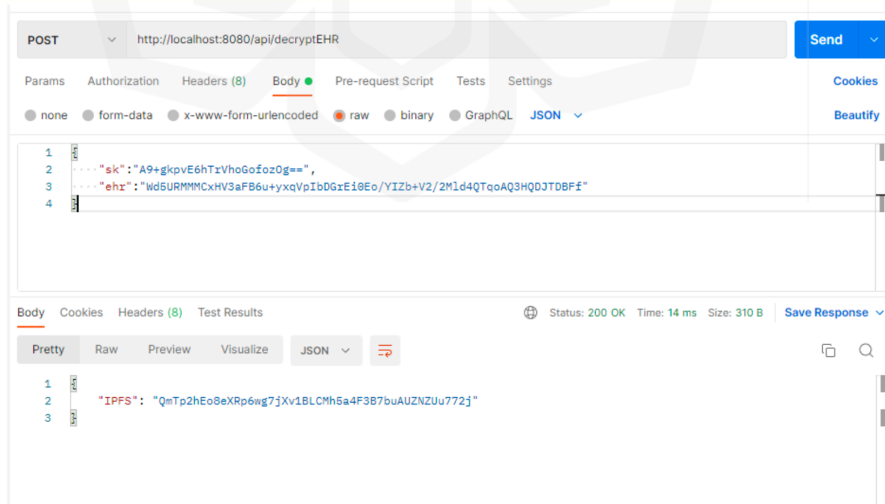


Figure 4.17 show how decrypt EHR.

4.3 ANALYSIS

Our results in table 4.1 show that processing time and cost after encryption and decryption process time and cost before applying the E2E hybrid encryption. Table 4.2, however, shows the same results after applying the E2E hybrid encryption.

It is noticed that when we save the medical patient record using hybrid encryption, it's required execution time is more than when didn't use hybrid encryption. The saving cost, however, is approximately 73.4172% in execution time. As expected, a delay penalty is paid for enjoying a high level of E2E security, which is the one of the main contributions of our proposed system.

Table 4.1 Execution time and cost without hybrid encryption ECC-AES.

id	fnName	timeTaken	userID	time	gasUsed	totalCost
1	addDoctor()	10.703 s	0x3787a1Cc61EF97d1ae26d1eBDe95Fa79da1294c1 : (Admin)	1676631203868	44721	0.00089442 ETH
2	addPatient()	10.24 s	0xb4FB6cDa6C63c27FB02783351556962BC1179CCb : (Patient)	1676631238879	148329	0.00296658 ETH
3	getpatientDetails()	0.198 s	0xb4FB6cDa6C63c27FB02783351556962BC1179CCb : (Doctor)	1676631280021	0	0 ETH
4	saveMedicalRecord()	11.082 s	0xf90314179e67D129CaF53643FdAaafF36a8cA597 : (Doctor)	1676631283744	168010	0.0033602 ETH
5	getPatientRecords()	1.393 s	0xb4FB6cDa6C63c27FB02783351556962BC1179CCb : (Doctor)	1676631307183	0	0 ETH

Table 4.2. Execution time and cost with hybrid encryption ECC-AES.

id	fnName	timeTaken	userID	time	gasUsed	totalCost
26	getpatientDetails()	0.267 s	0xf90314179e67D129CaF53643FdAaafF36a8cA597 : (Doctor)	1676560236120	0	0 ETH
27	getpatientDetails()	0.347 s	0xf90314179e67D129CaF53643FdAaafF36a8cA597 : (Doctor)	1676560236123	0	0 ETH
28	getpatientRecords()	1.268 s	0xf90314179e67D129CaF53643FdAaafF36a8cA597 : (Doctor)	1676560240942	0	0 ETH
29	updateSharedSign()	19.095 s	0xf90314179e67D129CaF53643FdAaafF36a8cA597 : (Doctor)	1676560262012	61942	0.00123884 ETH
30	getpatientRecords()	1.2 s	0xf90314179e67D129CaF53643FdAaafF36a8cA597 : (Doctor)	1676560315741	0	0 ETH
31	getpatientDetails()	0.142 s	0xf90314179e67D129CaF53643FdAaafF36a8cA597 : (Doctor)	1676560422903	0	0 ETH
32	getpatientDetails()	0.149 s	0xf90314179e67D129CaF53643FdAaafF36a8cA597 : (Doctor)	1676560422906	0	0 ETH
33	getpatientRecords()	1.077 s	0xf90314179e67D129CaF53643FdAaafF36a8cA597 : (Doctor)	1676560426401	0	0 ETH
34	savePatientMedicalRecords()	21.995 s	0xf90314179e67D129CaF53643FdAaafF36a8cA597 : (Doctor)	1676560426400	44667	0.00089334 ETH
35	updateSharedSign()	21.789 s	0xf90314179e67D129CaF53643FdAaafF36a8cA597 : (Doctor)	1676560448406	61918	0.00123836 ETH
36	getpatientDetails()	2.226 s	0xb4FB6cDa6C63c27FB02783351556962BC1179CCb : (Patient)	1676560516171	0	0 ETH
37	getpatientDetails()	2.031 s	0xb4FB6cDa6C63c27FB02783351556962BC1179CCb : (Patient)	1676560517579	0	0 ETH
38	shareMedicalRecords()	21.308 s	0xb4FB6cDa6C63c27FB02783351556962BC1179CCb : (Patient)	1676560521382	118097	0.00236194 ETH
39	getpatientDetails()	2.207 s	0xb4FB6cDa6C63c27FB02783351556962BC1179CCb : (Patient)	1676560905628	0	0 ETH
40	getpatientDetails()	2.05 s	0xb4FB6cDa6C63c27FB02783351556962BC1179CCb : (Patient)	1676560952178	0	0 ETH
41	getpatientDetails()	2.036 s	0xb4FB6cDa6C63c27FB02783351556962BC1179CCb : (Patient)	1676561147787	0	0 ETH
42	getpatientDetails()	2.036 s	0xb4FB6cDa6C63c27FB02783351556962BC1179CCb : (Patient)	1676561206061	0	0 ETH
43	getpatientDetails()	2.047 s	0xb4FB6cDa6C63c27FB02783351556962BC1179CCb : (Patient)	1676561278687	0	0 ETH
44	getpatientDetails()	2.048 s	0xb4FB6cDa6C63c27FB02783351556962BC1179CCb : (Patient)	1676561530590	0	0 ETH
45	getpatientDetails()	2.031 s	0xb4FB6cDa6C63c27FB02783351556962BC1179CCb : (Patient)	1676561581652	0	0 ETH
46	addDoctor()	20.051 s	0x3787a1Cc61EF97d1ae26d1eBD95Fa79da1294c1 : (Admin)	1676561680933	177704	0.00355408 ETH
47	addPatient()	20.073 s	0x3fa4Caf64d1748c1ED3ca2526424659CAE88c385 : (Patient)	1676561734783	218773	0.00437546 ETH

4.3.1 Encryption and decryption time

For confirmation, we also compared the encryption and decryption times for our proposed hybrid method using various key sizes and with the already-in-use techniques (AES, DES). Several keys, including 64 bits, 128 bits, 192 bits, and 256 bits, were used for the tests. Using various keys, the proposed and current encryption techniques were tested for text data. Firstly, the following Table 4.3 show key generation time for asymmetric and symmetric types of different sizes:

Table 4.3 shows key generation time for ECC-AES In different size.

Key Size	Time in nano seconds (Encryption)	Time in nano seconds (Decryption)
64	2178500	3876199
128	2386100	4083799

192	2384900	4082599
256	2301300	3998999

The encryption and decryption times of all the values are displayed in nanosecond in Tables 4.4 and Table 4.5 For a better understanding, we may also view the visual analysis of these values.

Table 4.4. Compare encryption time with other methods

	AES	DES	Proposed methods Hybrid ECC-AES
64	3.52	3.67	0.654
128	3.59	4.2	0.65
192	3.45	4.25	0.93
256	3.61	4.43	0.89

Table 4.5 Compare the decryption time with other methods

	AES	DES	Proposed methods Hybrid ECC-AES
64	2.69	3.11	0.6
128	2.82	3.30	0.844
192	2.93	3.38	0.843
256	3.08	3.58	0.835

Figure 4.18 compares the times required for the encryption process of the identical text data by the AES, DES, and hybrid methods. The hybrid ECC-AES model was

found to take less time on key sizes in bits (64, 128, 192, and 256), but the other algorithms in use took more time on key sizes in bits (64, 128, 192, and 256).

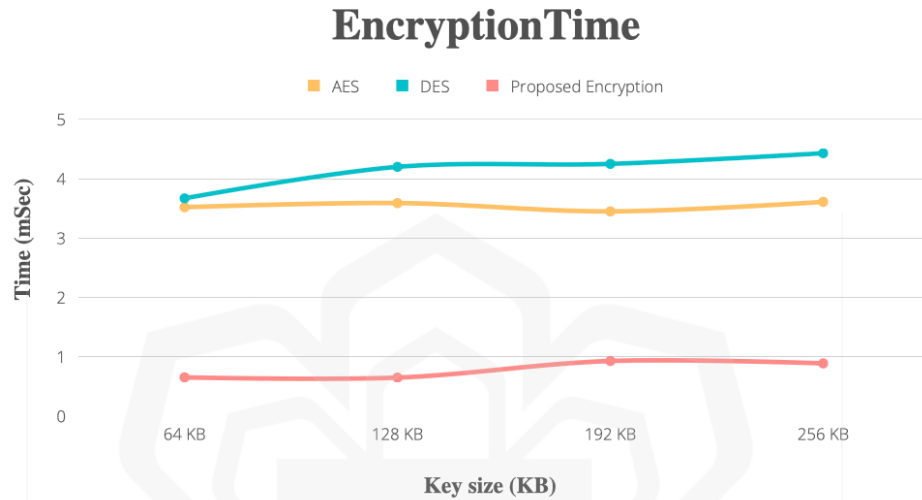


Figure 4.18 Compare time with different key size for encryption process with other methods.

Figure 4.19 illustrates how long the decryption of the same text data using the AES, DES, and hybrid methods takes in comparison. The hybrid model was found to take less time on keys of 64, 128, 192, and 256 bits, but the two existing algorithms (AES, DES) spent more time on keys of those same sizes (64, 128, 192, and 256).

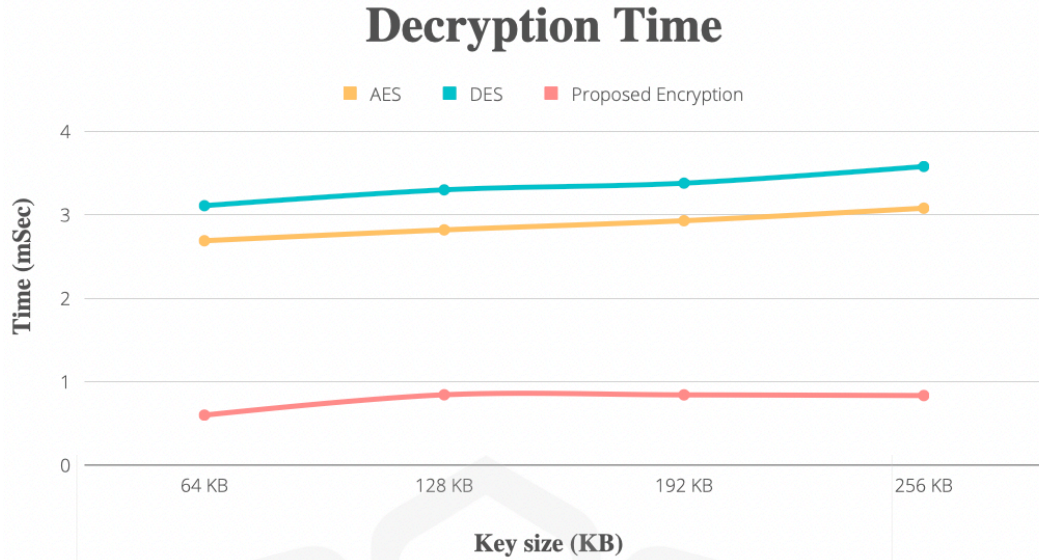


Figure 4.19 Compare time with different key size for decryption process with other methods.

Figures 4.18 and 4.19 show how, in comparison to previous cryptographic algorithms, our scheme will aid in memory space optimization and minimise computational complexity. In comparison to other cryptographic methods, the hybrid algorithm requires a medium amount of memory and takes less time to encrypt and decrypt data.

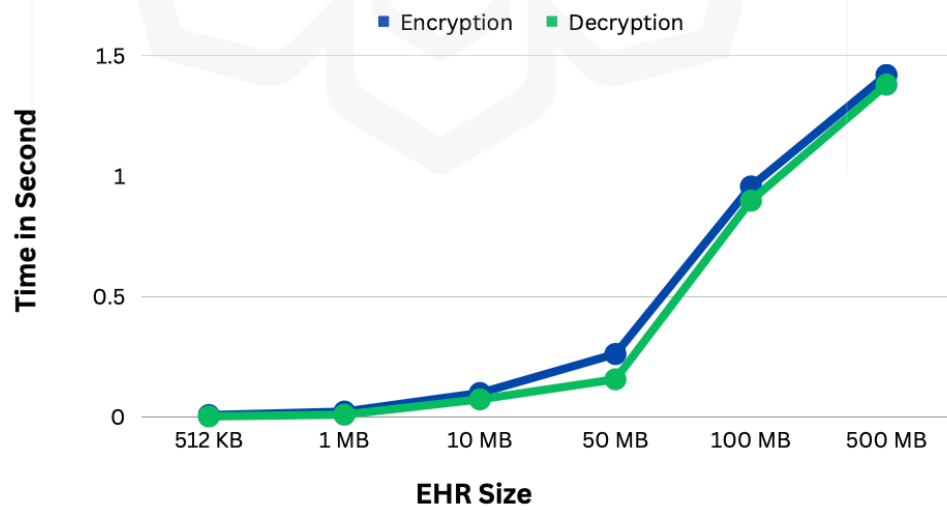


Figure 4.20 Encryption and decryption time consumption with different EHR size.

Record encryption time is the amount of time required to encrypt a health record based on the owner's request, whereas record decryption time is the amount of time required to decrypt a health record based on the user's request. In our proposed encryption and decryption methods save time around 74% in encryption and decryption process in execution time.

Figure 4.20 depiction of encryption and decryption times shows a consistent rise in time consumption with record size. This indicates that a file's size affects how long it takes to encrypt and decode it.

Table 4.6 benchmarks the length of time required for encryption and decryption of blockchain works of various EHR sizes utilizing with Thwin and Vasupongayya (2019) and Boumezbeur I,et al. (2022). Figure 4.20 illustrates how our proposed hybrid encryption and decryption times relate to the size of the HER. The encryption and decryption times somewhat rise along with the EHR size. Moreover, the encryption and decryption processes take less than 1 second to complete when the EHR is around 100 MB in size. Even if the EHR is more than 100 MB, the extra time required is only around 1 second.

We compared the test results in Figure 4.20 with the amount of time that Thwin and Vasupongayya (2019) Boumezbeur I,et al. (2022) spent encrypting and decrypting data. Figures 4.21, 4.22 present the findings of the comparison. The encryption and decryption procedures take less time than in the other tasks.

Our encryption and decryption efficiency are much better than those of Thwin and Vasupongayya (2019) and Boumezbeur I,et al. (2022) when the EHR is large. X-rays images are just examples of the numerous huge image files included in EHRs. Based on these comparisons, our technique is superior to earlier work on health record encryption and decryption.

Table 4.6 compare encryption and decryption time with different EHR size.

File Size	Thwin and Vasupongayya (2019)		Boumezbeur I, et al. (2022).		Proposed system	
	Encryption	Decryption	Encryption	Decryption	Encryption	Decryption
512 KB	0.094	0.0064	0.0158	0.0027	0.0053	0.0017
1 MB	0.101	0.01662	0.0452	0.0157	0.0126	0.0089
10 MB	0.152	0.069	0.075	0.058	0.0489	0.0227
50 MB	0.503	0.438	0.406	0.401	0.261	0.165
100 MB	1.428	1.414	1.149	1.228	0.957	0.898

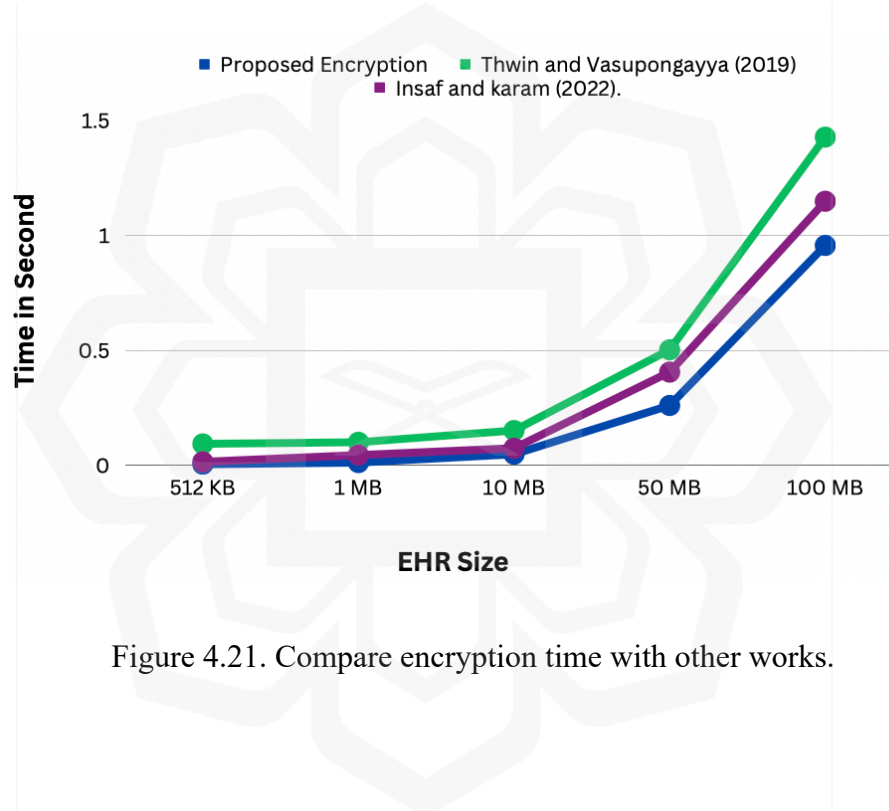


Figure 4.21. Compare encryption time with other works.

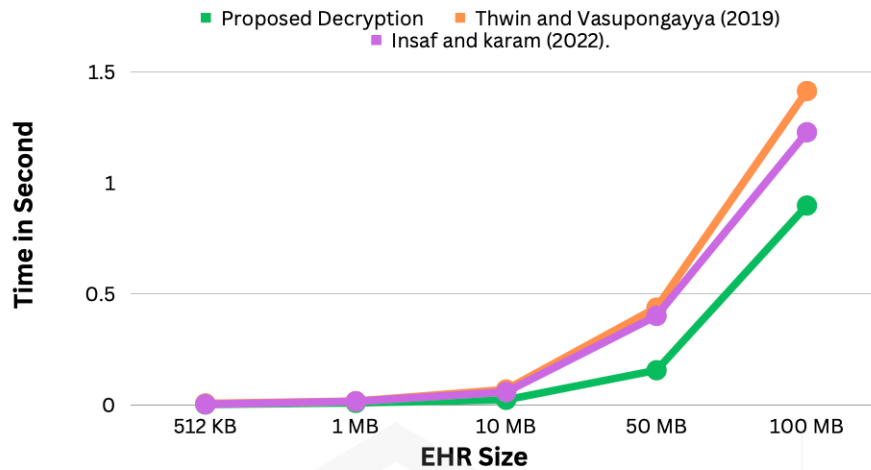


Figure 4.22. Compare decryption time with other works.

4.3.2 IPFS uploading and downloading time.

The length of time it takes to upload and download a health record to and from the IPFS is the basis for the system evaluation. The performance of uploading health records to the IPFS for various record sizes is shown in Figure 4.23. The graph below displays all of the constituent times. The size of a health record has a direct relationship with how long it takes to upload it to the IPFS. By stating that the key's calculation time is unaffected by the file's size, we can see that it has remained stable. However sometimes the record upload time only slightly increased, which can be related to the network's status at the time.

In other hand, process of original health record that is downloaded from IPFS is the exact opposite of what is uploaded. Also, Figure 4.23 shows the outcomes of the IPFS download procedure and the decryption process. The download time continuously increases from 512KB to 500 MB as the record size does. At the smallest record size of 512KB, it takes only 0.285 seconds, but the larger record computation of size 500 MB significantly lengthens the download time to 13.26 s. The outcomes trend is therefore the same as the uploading process. Moreover, both download and decryption times are evolving.

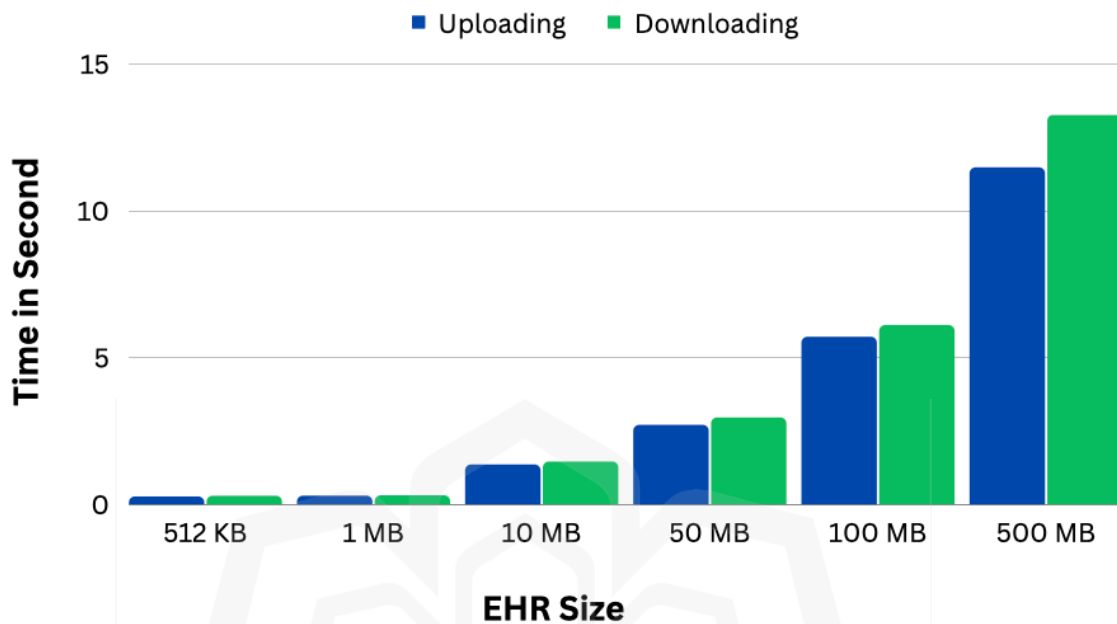


Figure 4.23 Time to upload and download different EHR size.

Finally, To demonstrate the usefulness of our study, we compare the key generation time, record encryption time, record decryption time, record upload time, and record download time for different file sizes ranging from 512 KB to 500 MB. A comparison of our proposed system to other nearby works is also included, based on a set of security criteria such as integrity, privacy, access control, encryption, and key encryption.

The suggested system is shown in Table 4.7 along with the upload and download process response times for various studies, where "UP" and "DN" stand for "upload" and "download," respectively. The proposed system performs better than the other current approaches as a result. It provides superior upload and download performance, albeit the numbers may differ significantly depending on the speed of internet connection and the size of records. Table 4.7's comparisons of response times for encryption and decryption show that the proposed system performs better than other approaches due to the lack of complex calculations. The suggested procedure is at least twice as quick as the others.

Table 4.7 Comparison of uploading and downloading time for encrypt/decrypt files.

File size	Hema and Kesavan (2019)		Boumezbeur I,et al. (2022)		Proposed system	
	UP	DN	UP	DN	UP	DN
512 KB	0.80	0.82	0.32	0.35	0.26	0.285
1 MB	1.20	1.24	0.35	0.39	0.287	0.30
10 MB	5.60	5.68	1.82	1.89	1.35	1.45
50 MB	8.25	8.78	3.2	3.25	2.7	2.95
100 MB	16.35	18.98	7.69	8.01	5.70	6.1
500 MB	32.10	38.22	14.36	18.03	11.47	13.26

4.3.3 Classification Results of The Diseases Prediction

The analysis result for DenseNet121 to predict disease from 14 different diseases using NIH ChestX-Ray8 dataset shown in ROC curve in the Figure 4.24. An ROC curve (receiver operating characteristic curve) is a graph showing the performance of our classification model at all classification thresholds. This curve plots two parameters: True Positive Rate. False Positive Rate.

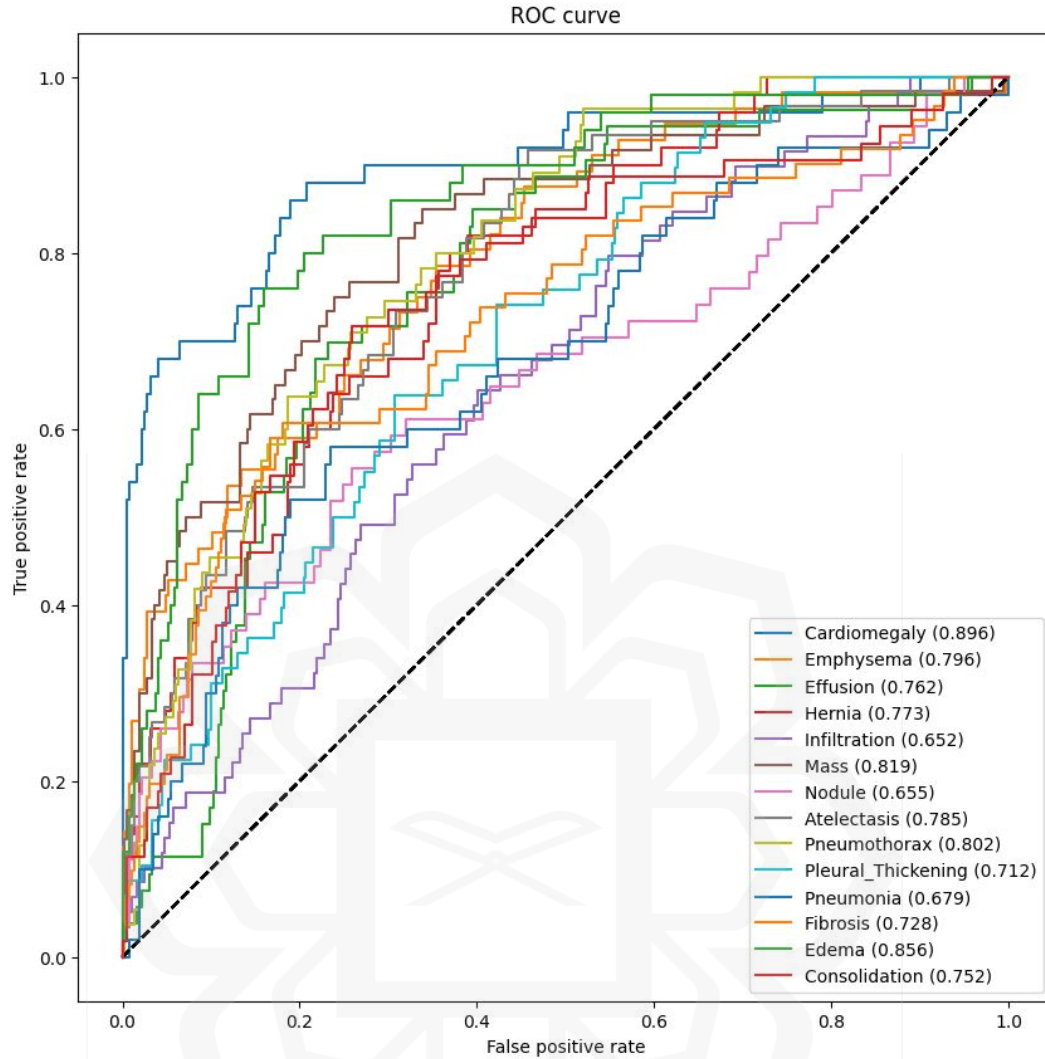


Figure 4.24 ROC for the performance of classification with 14 diseases.

4.4 DISCUSSION

There are several authors who discussed blockchain based healthcare systems to secure healthcare data sharing. Unlike Zhao et al. (2019) model, which encrypt all files and stored it on blockchain, in our proposed methods used IPFS to reduce the pressure on blockchain storage and meet real-world deployment requirements.

in another hand, there are developed systems that make the patient had a private key for doctor for access data in Chen et al. (2019). However, this transmission mechanism cannot guarantee the confidentiality of private key from hostile nodes which can access to key any time.

We employed a hybrid encryption strategy, as opposed to Qin et al. (2021) to assure better security and to benefit from each encryption. Our technology uses completely secure encryption keys. Before being transferred to IPFS, the original EHR is encrypted. As a result, the blockchain's storage capacity issue can be resolved, and there is a significantly lower chance that original electronic health data could leak sensitive information.

Although many of the proposed systems can provide the necessary levels of privacy and integrity, not all of them can offer access control, which is a crucial security objective in an EHR sharing system. Furthermore, only a small number of planned solutions use smart contracts. Our solution uses the smart contract to securely store the encryption key, maintain the EHR signature, confirm authentication, and manage the user's access control, in contrast to Qin et al. (2021), Khalaf et al. (2020).

The system we provide safeguards patient privacy by enabling them to establish precise smart contract-based access controls to their EHR data. It also has a single point of failure and a decentralised network topology. Our strategy, in contrast to Chen et al. (2019), relies on defined user responsibilities to guard against security risks such as unauthorised access to EHR data. Malicious users won't be able to access EHR data as a result. In accordance with the patient's preferences, it also enables quick and secure access to EHR data. It guarantees that EHR data elements are accessible without the requirement for external validation.

In comparison to these similar studies, our system performs better in terms of features like personal data protection, access control, and data integrity. It also offers a feasible alternative for upgrading current electronic healthcare systems, including smart contracts, hybrid encryption, and encryption keys.

The following Table 4.8 demonstrates how the proposed system ensures security objectives compared to others.

The protection of confidentiality ensures that information is kept private and secure. In our system, only those with permission can view a patient's electronic health records (EHRs). Details are in the EHR sharing paragraph, and only the entity that has authorisation is permitted to access data from the cloud IPFS.

By taking advantage of the blockchain and smart contract encryption capabilities, the blockchain account is anonymous and cannot be linked to a real identity. Hence, blockchain

privacy can stop public information from disclosing a person's real identify. Our access control system protects people's personal information. Smart contracts' user authentication and permission features prevent malicious access, protecting our cloud servers from further attacks.

The integrity ensures that patient data is transferred between authorised users without being altered. To avoid change, the electronic health records are still encrypted in our scheme.

Users of EHRs cannot alter or modify the quality of transactions signed or registered in smart contracts for EHR sharing, and neither can anyone else. Moreover, each node keeps a copy of the blockchain data, making it simple to identify changes to a particular block if a node interacts with other nodes. Most crucially, in our situation, EHR users don't have the authority to alter or update the access policies and smart contract. As a result, if the EHR is altered, Merkle's root value would also change, causing the block content to change. As a result, EHRs can be maintained utilising the blockchain safely and accurately without being tampered with.

The connected contracts must first be added to the blockchain when the patients authenticate their EHR with a signature. Based on the signatures and integrity offered by the blockchain, data authenticity can be provided.

The user sends the request to what he or she is interested in in order to be granted access to a patient's EHRs. If the user is able to access the information, it will be returned together with the encrypted EHRs that were requested, ensuring access control.

So, to ensure the best possible health care experience for the patient, the patient should be able to access or grant access to his or her medical information as needed. This is currently not possible when the data is stored in a hospital's proprietary Electronic Health Record (EHR). As a result, there is a requirement for patient operated where the details of the patient's treatment are with the patient. This paper proposes a secure, interoperable patient-centric data access management system based on blockchain. In our proposed system, the patient has complete control over his or her health record-related data, which is stored securely using IPFS. In addition to, stored the EHR in IPFS with high level of security which combining symmetric and asymmetric encryption technique to create hybride ECC-AES encryption. Furthermore, we integrate the system with AI techniques to support doctor or

health providers in decision making process .The following Table 4.6 explains the summary of the comparison between the existing and suggested system focusing on privacy, security, and integrity. The authors in (Rajput,A. et al. 2021) reviewed current strategies for healthcare management using blockchain technology and its effects. The existing architectures in (Shen, B. et al. 2019, Dwivedi, A.D. 2019, Egala, B.S. 2021) were examined. Each block includes a health record hash value that would transform whenever the record is updated. This ensures the records are immutable as it is computationally expensive to manipulate the ledger. On top of that, stakeholder access to patients’ medical records is prohibited by access control rights and level. We have benchmarked our proposed system to them. The proposed system architecture, as can be seen, shares all features proposed in other systems.

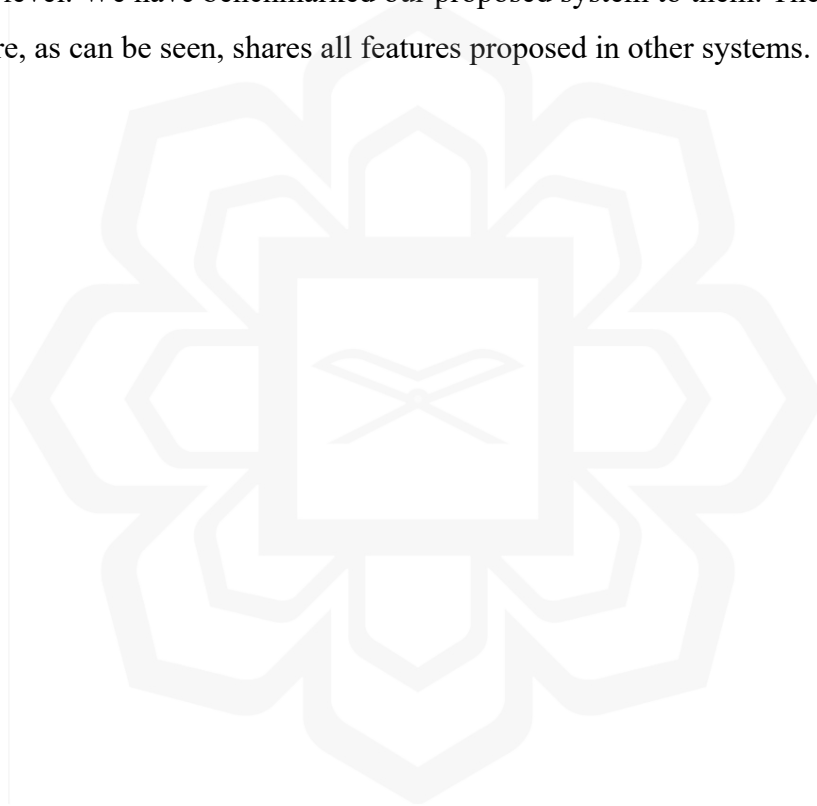


Table 4.8 Compare our proposed system with the related literature.

	Blockchain- Based	Confidential	Privacy	Integrity	Access control	Hybrid Encryption	Smart Contract	Patient Centric	Predict Xray Disease
(Chen et al. 2019)	yes	yes	yes	yes	no	no	yes	no	no
(Shen, B. et al. 2019)	yes	yes	yes	yes	no	no	no	no	no
(Dwivedi, A.D. 2019)	yes	no	yes	yes	no	no	yes	no	no
Khalaf et al. (2020)	yes	yes	yes	No	Yes	no	no	no	no
(Rajput,A. et al. 2021)	yes	yes	yes	No	Yes	no	yes	no	no
(Egala, B.S. 2021)	yes	yes	yes	yes	yes	no	yes	no	no
(Mohsan SA, et al. 2022)	yes	no	yes	no	yes	no	yes	yes	no
(Mahajan HB, et al. 2023)	yes	no	yes	no	no	no	no	no	yes
(Peng G, et al. 2023)	yes	yes	yes	yes	yes	no	yes	yes	no
(Khan AA, et al. 2023)	yes	yes	yes	no	yes	No	yes	yes	no
Proposed system 2023	yes	yes	yes	yes	yes	yes	yes	yes	yes

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

5.1 CONCLUSION

Medical records are an important asset in the healthcare industry, but their dispersion across different platforms can make it difficult to share information and create a cohesive system. Decentralized designs and system interoperability are becoming more important as the dispersed nature of healthcare services is becoming more apparent. In this study, a patient-centered EHR management system was developed, which uses a decentralized framework for sharing access and storing medical data. This system is built on the Ethereum blockchain and IPFS architecture, and includes an access control system to provide authorized parties access to the relevant blockchain data. First, a framework is proposed for sharing Electronic Health Records (EHR) among various entities, with a focus on utilizing blockchain and cloud storage via IPFS. In this study, cloud storage stores the encrypted EHR while EHR signatures are stored on the Ethereum EHR blockchain. The proposed system aims to enhance patients' rights by giving them complete control over their healthcare records through a smart contract protocol.

This unique approach enables patients to manage their medical reports and authorize or deny access to them for clinical trials or research. To ensure data confidentiality and privacy preservation, the system uses symmetric and asymmetric encryption.

Finally, the system is integrated with Artificial Intelligence (AI) to revolutionize the E-Healthcare Records (EHR) system and offer a possible solution to various technical challenges the healthcare industry is currently addressing. Chest X-ray images are one of the most common clinical methods for diagnosing a number of diseases. The goal of this integration is to develop a solution to detect 14 different chest conditions from an X-ray image. Given an X-ray image as input, our classifier outputs a label vector indicating which of 14 disease classes the image falls into. Original study focuses on predicting 14 diseases, a type of deep learning technique, was employed for automated feature extraction and disease

identification to assist in the decision-making process based on the patient's X ray image to potentially decipher the challenging nature of this task

The proposed framework was also subjected to analysis and evaluation, and based on the results, it was found to be efficient and compliant with various security requirements. The use of E2EE hybrid encryption ECC-AES provides a high level of privacy, security, confidentiality, and scalability. Additionally, the framework offers improved productivity, data integrity, efficient audit, and shared access to medical data, making it easier for patients to access a secure and unalterable medical database.

To establish a global patient-centric EHR management system and evaluate related regulations and standards for integrating blockchain technology into the healthcare system, our ultimate goal for this study is to implement the proposed system architecture using real-world scenarios. Additionally, we believe that the inclusion of artificial intelligence will help clinicians analyze diagnostic medical data more effectively and improve communication with patients.

5.2 NOVELTY OF THE WORK

The proposal is to develop a Patient-Centered Blockchain-Based EHR Management (PCBEHRM) system that allows patients to manage their healthcare records across multiple stakeholders, without the need for a centralized infrastructure. We achieve this using an Ethereum smart contract called Patient-Centric Access Control (PCAC) protocol, which ensures patient privacy and control. The system also employs the Inter Planetary File System (IPFS) to store medical records, which is distributed, secure, and immutable. To further enhance security and reduce computational power, we use an End to End Encryption (E2EE) functionality that combines Elliptic Curve Cryptography (ECC) and Advanced Encryption Standard (AES) methods. The proposed system aims to provide a secure and optimized scheme for sharing medical data while maintaining data security and integrity. In the long-term, we aim to implement this system architecture on a larger scale to establish a global patient-centered EHR management system, leveraging the benefits of artificial intelligence for diagnostic analysis and communication with patients.

Finally, our study enhances the proposed PCBEHRM system with deep learning artificial intelligence capabilities to revolutionize the management of the EHR and offer an add-on diagnostic tool based on the captured EHR metadata. The proposed enhancement integrated deep learning feature is a developed a solution that can detect 14 different chest conditions from classification of X-ray image. The results show that the proposed system achieves a reduced storage cost of 73.4172% and around 76% of time efficient in comparison to other proposed systems in the open literature.

5.3 THESIS CONTRIBUTION

- 1) a detailed literature review study that covered the benefits of applying Blockchain and benefits of applying AI to the problem of health record management each of them individually. Also provide an updated discussion on privacy and security issues related to EHR management. Finally, A taxonomy of proposed AI-Blockchain solutions to the problem of EHR.
- 2) Design of a patient-centered on EHR management (PCEHRM) system that comprises ensures that the patient has full control over his health record and who has access to their health record and make practical use of the data and adheres to new standards in data integrity and privacy.
- 3) Enhancing the proposed system with E2E Encryption using hybrid ECC-AES. Uses symmetric and asymmetric algorithms to encrypt EHR and secret keys for ensuring confidentiality and privacy.
- 4) integrate the proposed system and add AI model. By deep learning algorithm can help the doctor in diagnosis decision-making and predict the diseases from 14 lung different diseases by classification X ray images using DenseNet121.
- 5) Performance analysis of the proposed system with time and cost consumed before and after encryption EHR. in addition, analysis the time need to upload or download from IPFS cloud storage with different size of EHR.
- 6) Benchmarking of the proposed system against other recent proposals in the literature.

5.4 FUTURE WORK

In general, there are few future directions that stand clear for the EHR management research efforts, namely, applications of Big Data, AI, Edge Computing and IoMT. Below is our take on these directions.

5.4.1 Big Data

A significant problem for healthcare data systems seeking to improve the quality of healthcare services is acquiring, processing, and analyzing huge volumes of personal healthcare data, particularly from commonly used mobile and wearable devices, while minimizing privacy violations. Blockchain technology has the potential to address the security concerns associated with big data techniques by providing immutability, security, and traceability. Big data can make the best use of all healthcare data assets to assist necessary improvements in areas of prediction in healthcare diagnosis, analysis in magnetic resonance imaging, and other applications (Otero P, et al. 2014).

Two broad categories of big data analysis are data management and data analysis. For data management, blockchain technology can be utilized to securely maintain immutable healthcare records. For data analysis, the blockchain's transactions and records can be extracted and studied for potential trading behaviors.

5.4.2 Artificial Intelligence

When blockchain technology is combined with AI in a variety of real-world healthcare applications, the resulting systems become more efficient and stable (M. N. K. Boulos, 2018). Machine learning (ML) and deep learning (DL) are two major branches of AI that are assisting in the automation of real-world applications. In the near future, machine learning will be used in concert with blockchain to manage EHRs. Despite the difficulties associated with storing, distributing, and training vital EHR data to design practical applications, interest among researchers in developing machine learning and blockchain-based EHR applications has grown tremendously (X. Zheng, 2018), (S. H. Lee and C. S. Yang, 2018). IBM has

announced intentions to implement an intelligent blockchain, in which an AI agent performs various duties such as enforcing laws, improving records, detecting suspicious activity, and making recommendations for upgrading smart contracts over a broad network. In the MATRIX project (L. Tzu., 2017), AI is employed to construct a next generation blockchain that enables the automated development of intelligent contracts, enhances protection against malicious attacks, and enables highly scalable operations. Various machine learning techniques can be used to detect fake EHR data, ensuring that only authentic EHRs are maintained on the blockchain. Deep learning enables the recovery and storage of previously damaged scanned medical records in blockchain for the sake of knowledge enhancement (e.g., drug analysis and prediction) (D. E. O’Leary, 2013). Additionally, deep learning as-a-service (DaaS) is employed on stored EHRs to accurately forecast future diseases based on current patient diagnosis reports (P. Bhattacharya, 2021). Machine learning techniques can also be employed to protect blockchain networks from large-scale attacks (S. Dey, 2018). There are some established projects that mix AI with blockchain. For example, SingularityNET (Singularitynet, 2018) focuses on developing AI and blockchain-based networking for the robot brain, while DeepBrainChain focuses on developing a platform for developing AI algorithms. Additionally, several machine learning and deep learning-based health-related projects are underway, including the Gamalon project, TraneAI (2017), and Neureal (2021).

5.4.3 Edge Computing

Due to network congestion and data size, sharing huge volumes of EHRs among diverse health care companies is problematic. Recent options for EHR management are limited in terms of scalability, computing cost, and reaction time. Edge computing may provide a solution to these difficulties. It can process a vast amount of data from multiple locations, as edge computing is comprised of a set of servers/computers (A. Awad Abdellatif, 2021). Researchers in (K. Gai, 2019) propose using edge computing to extend cloud services to the network's edge, thereby increasing processing capacity and device QoS. Edge Processing offers the advantages of large data storage, extensive networking, and high computing power, while also enabling secure and regulated scaling for distributed EHR applications. While

edge computing has several drawbacks, including security, vulnerability to various attacks during message transmission, and integrity, blockchain-based solutions face several challenges, including storage, scalability, block size constraints, and block creation time, all of which can be addressed using edge computing. Similar approaches for decentralized technology can improve privacy, security, and resource management on an automatic basis (R. Yang, 2019). Combining the two can provide several advantages. For example, blockchain can first be used to implement distributed controls across multiple edge nodes. The blockchain mining process verifies the accuracy, consistency, and dependability of data. Then, user privacy can be enhanced further by allowing people to control data using cryptographic keys. Finally, edge computing entails resource sharing across nodes, which can be accomplished securely via blockchain-based smart contracts (P. De Filippi, 2016).

5.4.4 Internet of Medical Things (IoMT)

The IoMT is a collection of medical devices and software that connect to various healthcare providers via online computer networks. The Internet of Medical Things is built on the concept of Machine-to-Machine (M2M) communication between wireless medical devices. Medical care providers and authorities can obtain real-time health updates on patients from remote places using wearable devices via the IoMT. Apart from the benefits of IoMT, there are some disadvantages, as IoMT devices are susceptible to security attacks. Not only has demand for novel medical devices surged dramatically during the Covid-19 outbreak, but so have cyber risks associated with them (P. Dialani, 2021). Blockchain technology might be viewed as a savior against the hazards posed by IoMT devices. Blockchain's decentralized key management, inseparability, and integrity qualities enable the secure communication of intelligent medical equipment.

APPENDIX X: Bibliography

REFERENCES

- Kumari, A., Tanwar, S., Tyagi, S., & Kumar, N. (2018). Fog computing for Healthcare 4.0 environment: Opportunities and challenges. *Computers & Electrical Engineering*, 72, 1-13.
- Campanella, P., Lovato, E., Marone, C., Fallacara, L., Mancuso, A., Ricciardi, W., & Specchia, M. L. (2016). The impact of electronic health records on healthcare quality: a systematic review and meta-analysis. *The European Journal of Public Health*, 26(1), 60-64.
- Tanwar, S., Parekh, K., & Evans, R. (2020). Blockchain-based electronic healthcare record system for healthcare 4.0 applications. *Journal of Information Security and Applications*, 50, 102407.
- Mistry, I., Tanwar, S., Tyagi, S., & Kumar, N. (2020). Blockchain for 5G-enabled IoT for industrial automation: A systematic review, solutions, and challenges. *Mechanical systems and signal processing*, 135, 106382.
- Tanwar, S., Parekh, K., & Evans, R. (2020). Blockchain-based electronic healthcare record system for healthcare 4.0 applications. *Journal of Information Security and Applications*, 50, 102407.
- Tagde, P., Tagde, S., Bhattacharya, T., Tagde, P., Chopra, H., Akter, R., ... & Rahman, M. H. (2021). Blockchain and artificial intelligence technology in e-Health. *Environmental Science and Pollution Research*, 28, 52810-52831.
- Vora, J., Italiya, P., Tanwar, S., Tyagi, S., Kumar, N., Obaidat, M. S., & Hsiao, K. F. (2018, July). Ensuring privacy and security in e-health records. In *2018 International conference on computer, information and telecommunication systems (CITS)* (pp. 1-5). IEEE.
- Tanwar, S., Parekh, K., & Evans, R. (2020). Blockchain-based electronic healthcare record system for healthcare 4.0 applications. *Journal of Information Security and Applications*, 50, 102407.
- Kubo, T., Yanasan, A., Herbosa, T., Buddh, N., Fernando, F., & Kayano, R. (2019). Health data collection before, during and after emergencies and disasters—the result of the

- Kobe expert meeting. *International journal of environmental research and public health*, 16(5), 893.
- Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future healthcare journal*, 6(2), 94.
- Feng, Q., He, D., Zeadally, S., Khan, M. K., & Kumar, N. (2019). A survey on privacy protection in blockchain system. *Journal of network and computer applications*, 126, 45-58.
- Lin, C., He, D., Huang, X., Khan, M. K., & Choo, K. K. R. (2020). DCAP: A secure and efficient decentralized conditional anonymous payment system based on blockchain. *IEEE Transactions on Information Forensics and Security*, 15, 2440-2452.
- Ma, S., Deng, Y., He, D., Zhang, J., & Xie, X. (2020). An efficient NIZK scheme for privacy-preserving transactions over account-model blockchain. *IEEE Transactions on Dependable and Secure Computing*, 18(2), 641-651.
- Lamport, L., Shostak, R., & Pease, M. (1982). The Byzantine Generals Problem—ACM Transactions of Programming Languages and Systems. *The Byzantine Generals Problem ACM Transactions on Programming Languages and Systems*, 4(3).
- Wang, W., Hoang, D. T., Hu, P., Xiong, Z., Niyato, D., Wang, P., ... & Kim, D. I. (2019). A survey on consensus mechanisms and mining strategy management in blockchain networks. *Ieee Access*, 7, 22328-22370.
- Siyal, A. A., Junejo, A. Z., Zawish, M., Ahmed, K., Khalil, A., & Soursou, G. (2019). Applications of blockchain technology in medicine and healthcare: Challenges and future perspectives. *Cryptography*, 3(1), 3.
- Shahnaz, A., Qamar, U., & Khalid, A. (2019). Using blockchain for electronic health records. *IEEE access*, 7, 147782-147795.
- Nguyen, D. C., Ding, M., Pathirana, P. N., & Seneviratne, A. (2021). Blockchain and AI-based solutions to combat coronavirus (COVID-19)-like epidemics: A survey. *Ieee Access*, 9, 95730-95753.
- Cyran, M. A. (2018). Blockchain as a foundation for sharing healthcare data. *Blockchain in Healthcare Today*.
- Kaye, J., Whitley, E. A., Lund, D., Morrison, M., Teare, H., & Melham, K. (2015). Dynamic consent: a patient interface for twenty-first century research networks. *European journal of human genetics*, 23(2), 141-146.

- Lima, C. (2018). Blockchain GDPR privacy by design. *IEEE blockchain group*, 4(5).
- Villani, C., Bonnet, Y., & Rondepierre, B. (2018). *For a meaningful artificial intelligence: Towards a French and European strategy*. Conseil national du numérique.
- Dimitrov, D. V. (2019). Blockchain applications for healthcare data management. *Healthcare informatics research*, 25(1), 51-56.
- Zhang, P., White, J., Schmidt, D. C., Lenz, G., & Rosenbloom, S. T. (2018). FHIRChain: applying blockchain to securely and scalably share clinical data. *Computational and structural biotechnology journal*, 16, 267-278.
- Salah, K., Rehman, M. H. U., Nizamuddin, N., & Al-Fuqaha, A. (2019). Blockchain for AI: Review and open research challenges. *IEEE Access*, 7, 10127-10149.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112, p. 18). New York: springer.
- Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and vascular neurology*, 2(4).
- Ortín López, M. (2023). Study and prediction of time of recovery of consciousness after general anaesthesia.
- Resta, M., Sonnessa, M., Tànfani, E., & Testi, A. (2018). Unsupervised neural networks for clustering emergent patient flows. *Operations Research for Health Care*, 18, 41-51.
- Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health information science and systems*, 2, 1-10.
- Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. *Jama*, 319(13), 1317-1318.
- Gehrmann, S., Deroncourt, F., Li, Y., Carlson, E. T., Wu, J. T., Welt, J., ... & Celi, L. A. (2018). Comparing deep learning and concept extraction based methods for patient phenotyping from clinical narratives. *PloS one*, 13(2), e0192360.
- Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., ... & Ng, A. Y. (2017). Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*.

- Omar, I. A., Jayaraman, R., Salah, K., Yaqoob, I., & Ellahham, S. (2021). Applications of blockchain technology in clinical trials: review and open challenges. *Arabian Journal for Science and Engineering*, 46, 3001-3015.
- Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning deep features for discriminative localization. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2921-2929).
- Ford, E., Carroll, J. A., Smith, H. E., Scott, D., & Cassell, J. A. (2016). Extracting information from the text of electronic medical records to improve case detection: a systematic review. *Journal of the American Medical Informatics Association*, 23(5), 1007-1015.
- Shivade, C., Raghavan, P., Fosler-Lussier, E., Embi, P. J., Elhadad, N., Johnson, S. B., & Lai, A. M. (2014). A review of approaches to identifying patient phenotype cohorts using electronic health records. *Journal of the American Medical Informatics Association*, 21(2), 221-230
- Carrell, D. S., Schoen, R. E., Leffler, D. A., Morris, M., Rose, S., Baer, A., ... & Mehrotra, A. (2017). Challenges in adapting existing clinical natural language processing systems to multiple, diverse health care settings. *Journal of the American Medical Informatics Association*, 24(5), 986-991
- Kirby, J. C., Speltz, P., Rasmussen, L. V., Basford, M., Gottesman, O., Peissig, P. L., ... & Denny, J. C. (2016). PheKB: a catalog and workflow for creating electronic phenotype algorithms for transportability. *Journal of the American Medical Informatics Association*, 23(6), 1046-1052
- Foster, K. R., Koprowski, R., & Skufca, J. D. (2014). Machine learning, medical diagnosis, and biomedical engineering research-commentary. *Biomedical engineering online*, 13, 1-9.
- Asperti, A., & Mastronardo, C. (2017). The effectiveness of data augmentation for detection of gastrointestinal diseases from endoscopic images. *arXiv preprint arXiv:1712.03689*.
- Wong, S. C., Gatt, A., Stamatescu, V., & McDonnell, M. D. (2016, November). Understanding data augmentation for classification: when to warp?. In *2016 international conference on digital image computing: techniques and applications (DICTA)* (pp. 1-6). IEEE.
- Kemp, R., & Prasad, V. (2017). Surrogate endpoints in oncology: when are they acceptable for regulatory and clinical decisions, and are they currently overused?. *BMC medicine*, 15(1), 1-7.

- Lorbieski, R., & Nassar, S. M. (2018). Performance Assessment of Multi-Level Ensemble for Multi-Class Problems. *International Journal of Computer and Information Engineering*, 12(3), 149-155.
- Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. *Computational and structural biotechnology journal*, 15, 104-116.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5, 115-133.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088), 533-536.
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117.
- Liu, S., Liu, S., Cai, W., Pujol, S., Kikinis, R., & Feng, D. (2014, April). Early diagnosis of Alzheimer's disease with deep learning. In *2014 IEEE 11th international symposium on biomedical imaging (ISBI)* (pp. 1015-1018). IEEE.
- Cheng, Y., Wang, F., Zhang, P., & Hu, J. (2016, June). Risk prediction with electronic health records: A deep learning approach. In *Proceedings of the 2016 SIAM international conference on data mining* (pp. 432-440). Society for Industrial and Applied Mathematics.
- Zhang, J., Gong, J., & Barnes, L. (2017, July). HCNN: Heterogeneous convolutional neural networks for comorbid risk prediction with electronic health records. In *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)* (pp. 214-221). IEEE.
- Hossain, M. E., Khan, A., Moni, M. A., & Uddin, S. (2019). Use of electronic health data for disease prediction: A comprehensive literature review. *IEEE/ACM transactions on computational biology and bioinformatics*, 18(2), 745-758.
- Rallapalli, S., & Suryakanthi, T. (2016, November). Predicting the risk of diabetes in big data electronic health Records by using scalable random forest classification algorithm. In *2016 International Conference on Advances in Computing and Communication Engineering (ICACCE)* (pp. 281-284). IEEE.
- Uddin, S., Khan, A., & Baur, L. A. (2015). A framework to explore the knowledge structure of multidisciplinary research fields. *PloS one*, 10(4), e0123537.

- McCormick, T., Rudin, C., & Madigan, D. (2011). A hierarchical model for association rule mining of sequential events: An approach to automated medical symptom prediction.
- Uddin, S., Khan, A., & Piraveenan, M. (2015, January). Administrative claim data to learn about effective healthcare collaboration and coordination through social network. In *2015 48th Hawaii international conference on system sciences* (pp. 3105-3114). IEEE.
- Khan, A., Choudhury, N., Uddin, S., Hossain, L., & Baur, L. A. (2016). Longitudinal trends in global obesity research and collaboration: a review using bibliometric metadata. *Obesity Reviews*, *17*(4), 377-385.
- Khan, A., Uddin, S., & Srinivasan, U. (2018). Comorbidity network for chronic disease: A novel approach to understand type 2 diabetes progression. *International journal of medical informatics*, *115*, 1-9.
- Khan, A., Uddin, S., & Srinivasan, U. (2016, February). Adapting graph theory and social network measures on healthcare data: A new framework to understand chronic disease progression. In *Proceedings of the Australasian Computer Science Week Multiconference* (pp. 1-7).
- Khan, A., Uddin, S., & Srinivasan, U. (2019). Chronic disease prediction using administrative data and graph theory: The case of type 2 diabetes. *Expert Systems with Applications*, *136*, 230-241.
- Kang, E., Kim, S., Rhee, Y. E., Lee, J., & Yun, Y. H. (2021). Self-management strategies and comorbidities in chronic disease patients: associations with quality of life and depression. *Psychology, health & medicine*, *26*(8), 1031-1043.
- Gupta, S., Tran, T., Luo, W., Phung, D., Kennedy, R. L., Broad, A., ... & Venkatesh, S. (2014). Machine-learning prediction of cancer survival: a retrospective study using electronic administrative records and a cancer registry. *BMJ open*, *4*(3), e004007.
- Velez-Serrano, J. F., Velez-Serrano, D., Hernandez-Barrera, V., Jimenez-Garcia, R., Lopez de Andres, A., Garrido, P. C., & Alvaro-Meca, A. (2017). Prediction of in-hospital mortality after pancreatic resection in pancreatic cancer patients: A boosting approach via a population-based study using health administrative data. *PloS one*, *12*(6), e0178757.
- Huang, Z., Dong, W., Duan, H., & Liu, J. (2017). A regularized deep learning approach for clinical risk prediction of acute coronary syndrome using electronic health records. *IEEE Transactions on Biomedical Engineering*, *65*(5), 956-968.

- Jin, B., Che, C., Liu, Z., Zhang, S., Yin, X., & Wei, X. (2018). Predicting the risk of heart failure with EHR sequential data modeling. *Ieee Access*, 6, 9256-9261.
- Pham, Q. V., Nguyen, D. C., Huynh-The, T., Hwang, W. J., & Pathirana, P. N. (2020). Artificial intelligence (AI) and big data for coronavirus (COVID-19) pandemic: a survey on the state-of-the-arts. *IEEE access*, 8, 130820-130839.
- Xu, W., Zhang, J., Zhang, Q., & Wei, X. (2017, February). Risk prediction of type II diabetes based on random forest model. In *2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)* (pp. 382-386). IEEE.
- Forssen, H., Patel, R., Fitzpatrick, N., Hingorani, A., Timmis, A., Hemingway, H., & Denaxas, S. (2017). Evaluation of machine learning methods to predict coronary artery disease using metabolomic data. In *Stud Health Technol Inform* (Vol. 235, pp. 111-115). IOS Press.
- Liang, X., Zhao, J., Shetty, S., Liu, J., & Li, D. (2017, October). Integrating blockchain for data sharing and collaboration in mobile healthcare applications. In *2017 IEEE 28th annual international symposium on personal, indoor, and mobile radio communications (PIMRC)* (pp. 1-5). IEEE.
- Zheng, X., Mukkamala, R. R., Vatrupu, R., & Ordieres-Mere, J. (2018, September). Blockchain-based personal health data sharing system using cloud storage. In *2018 IEEE 20th international conference on e-health networking, applications and services (Healthcom)* (pp. 1-6). IEEE.
- Lee, S. H., & Yang, C. S. (2018). Fingernail analysis management system using microscopy sensor and blockchain technology. *International Journal of Distributed Sensor Networks*, 14(3), 1550147718767044.
- Yaji, S., Bangera, K., & Neelima, B. (2018, December). Privacy preserving in blockchain based on partial homomorphic encryption system for AI applications. In *2018 IEEE 25th International Conference on High Performance Computing Workshops (HiPCW)* (pp. 81-85). IEEE.
- Juneja, A., & Marefat, M. (2018, March). Leveraging blockchain for retraining deep learning architecture in patient-specific arrhythmia classification. In *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)* (pp. 393-397). IEEE.
- Ekblaw, A., Azaria, A., Halamka, J. D., & Lippman, A. (2016, August). A Case Study for Blockchain in Healthcare: "MedRec" prototype for electronic health records and medical research data. In *Proceedings of IEEE Open & Big Data Conference* (Vol. 13, p. 13).

- Yue, X., Wang, H., Jin, D., Li, M., & Jiang, W. (2016). Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. *Journal of medical systems*, 40, 1-8.
- Beninger, P., & Ibara, M. A. (2016). Pharmacovigilance and biomedical informatics: a model for future development. *Clinical therapeutics*, 38(12), 2514-2525.
- Dubovitskaya, A., Xu, Z., Ryu, S., Schumacher, M., & Wang, F. (2017). Secure and trustable electronic medical records sharing using blockchain. In *AMIA annual symposium proceedings* (Vol. 2017, p. 650). American Medical Informatics Association.
- Randall, D., Goel, P., & Abujamra, R. (2017). Blockchain applications and use cases in health information technology. *Journal of Health & Medical Informatics*, 8(3), 8-11.
- Rifi, N., Rachkidi, E., Agoulmine, N., & Taher, N. C. (2017, October). Towards using blockchain technology for eHealth data access management. In *2017 fourth international conference on advances in biomedical engineering (ICABME)*(pp. 1-4). IEEE.
- Wang, Z., & O'Boyle, M. (2018). Machine learning in compiler optimization. *Proceedings of the IEEE*, 106(11), 1879-1901.
- Ye, D., Zhang, M., & Vasilakos, A. V. (2016). A survey of self-organization mechanisms in multiagent systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 47(3), 441-461.
- Rizk, Y., Awad, M., & Tunstel, E. W. (2018). Decision making in multiagent systems: A survey. *IEEE Transactions on Cognitive and Developmental Systems*, 10(3), 514-529.
- Fioretto, F., Pontelli, E., & Yeoh, W. (2018). Distributed constraint optimization problems and applications: A survey. *Journal of Artificial Intelligence Research*, 61, 623-698.
- ur Rehman, M. H., Liew, C. S., Wah, T. Y., & Khan, M. K. (2017). Towards next-generation heterogeneous mobile data stream mining applications: Opportunities, challenges, and future research directions. *Journal of Network and Computer Applications*, 79, 1-24.
- ur Rehman, M. H., Batool, A., Liew, C. S., & Teh, Y. W. (2017). Execution models for mobile data analytics. *IT Professional*, 19(3), 24-30.

- Bottou, L., Curtis, F. E., & Nocedal, J. (2018). Optimization methods for large-scale machine learning. *SIAM review*, 60(2), 223-311.
- Contreras-Cruz, M. A., Lopez-Perez, J. J., & Ayala-Ramirez, V. (2017, June). Distributed path planning for multi-robot teams based on artificial bee colony. In *2017 IEEE congress on evolutionary computation (CEC)* (pp. 541-548). IEEE.
- Kurtulmus, A. B., & Daniel, K. (2018). Trustless machine learning contracts; evaluating and exchanging machine learning models on the ethereum blockchain. *arXiv preprint arXiv:1802.10185*.
- Kim, H., Park, J., Bennis, M., & Kim, S. L. (2019). Blockchained on-device federated learning. *IEEE Communications Letters*, 24(6), 1279-1283.
- McConaghy, T., Marques, R., Müller, A., De Jonghe, D., McConaghy, T., McMullen, G., ... & Granzotto, A. (2016). Bigchaindb: a scalable blockchain database. *white paper, BigChainDB*, 53-72.
- Shafagh, H., Burkhalter, L., Hithnawi, A., & Duquennoy, S. (2017, November). Towards blockchain-based auditable storage and sharing of IoT data. In *Proceedings of the 2017 on cloud computing security workshop* (pp. 45-50).
- Cui, S., Asghar, M. R., & Russello, G. (2018, July). Towards blockchain-based scalable and trustworthy file sharing. In *2018 27th International Conference on Computer Communication and Networks (ICCCN)* (pp. 1-2). IEEE.
- Luu, L., Narayanan, V., Zheng, C., Baweja, K., Gilbert, S., & Saxena, P. (2016, October). A secure sharding protocol for open blockchains. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security* (pp. 17-30).
- Zamani, M., Movahedi, M., & Raykova, M. (2018, October). Rapidchain: Scaling blockchain via full sharding. In *Proceedings of the 2018 ACM SIGSAC conference on computer and communications security* (pp. 931-948).
- Özyılmaz, K. R., & Yurdakul, A. (2018). Designing a blockchain-based IoT infrastructure with Ethereum, Swarm and LoRa. *arXiv preprint arXiv:1809.07655*.
- Vo, H. T., Kundu, A., & Mohania, M. K. (2018, March). Research Directions in Blockchain Data Management and Analytics. In *EDBT* (pp. 445-448).
- Lai, L., & Suda, N. (2018). Rethinking machine learning development and deployment for edge devices. *arXiv preprint arXiv:1806.07846*.
- Nakamoto, N. (2017). Centralised bitcoin: a secure and high performance electronic cash system. *Available at SSRN 3065723*.

- Wood, G. (2014). Ethereum: A secure decentralised generalised transaction ledger. *Ethereum project yellow paper*, 151(2014), 1-32.
- Dinh, T. T. A., Wang, J., Chen, G., Liu, R., Ooi, B. C., & Tan, K. L. (2017, May). Blockbench: A framework for analyzing private blockchains. In *Proceedings of the 2017 ACM international conference on management of data* (pp. 1085-1100).
- Li, Z., Kang, J., Yu, R., Ye, D., Deng, Q., & Zhang, Y. (2017). Consortium blockchain for secure energy trading in industrial internet of things. *IEEE transactions on industrial informatics*, 14(8), 3690-3700.
- Hwang, G. H., Chen, P. H., Lu, C. H., Chiu, C., Lin, H. C., & Jheng, A. J. (2018). InfiniteChain: A multi-chain architecture with distributed auditing of sidechains for public blockchains. In *Blockchain-ICBC 2018: First International Conference, Held as Part of the Services Conference Federation, SCF 2018, Seattle, WA, USA, June 25-30, 2018, Proceedings 1* (pp. 47-60). Springer International Publishing.
- King, S., & Nadal, S. (2012). Ppcoin: Peer-to-peer crypto-currency with proof-of-stake. *self-published paper*, August, 19(1).
- Bentov, I., Lee, C., Mizrahi, A., & Rosenfeld, M. (2014). Proof of activity: Extending bitcoin's proof of work via proof of stake [extended abstract] y. *ACM SIGMETRICS Performance Evaluation Review*, 42(3), 34-37.
- P4Titan. (2014). Slimcoin. A peer-to-peer crypto-currency with proof-of-burn 'Mining without powerful hardware'.
- Tschorsch, F., & Scheuermann, B. (2016). Bitcoin and beyond: A technical survey on decentralized digital currencies. *IEEE Communications Surveys & Tutorials*, 18(3), 2084-2123.
- De Angelis, S., Aniello, L., Baldoni, R., Lombardi, F., Margheri, A., & Sassone, V. (2018). PBFT vs proof-of-authority: Applying the CAP theorem to permissioned blockchain. In *CEUR workshop proceedings* (Vol. 2058). CEUR-WS.
- Zhang, J., Xue, N., & Huang, X. (2016). A secure system for pervasive social network-based healthcare. *Ieee Access*, 4, 9239-9250.
- Xia, Q. I., Sifah, E. B., Asamoah, K. O., Gao, J., Du, X., & Guizani, M. (2017). MeDShare: Trust-less medical data sharing among cloud service providers via blockchain. *IEEE access*, 5, 14757-14767.
- Jiang, S., Cao, J., Wu, H., Yang, Y., Ma, M., & He, J. (2018, June). Blochie: a blockchain-based platform for healthcare information exchange. In *2018 IEEE international conference on smart computing (smartcomp)* (pp. 49-56). IEEE.

- Li, H., Zhu, L., Shen, M., Gao, F., Tao, X., & Liu, S. (2018). Blockchain-based data preservation system for medical data. *Journal of medical systems*, 42, 1-13.
- Fan, K., Wang, S., Ren, Y., Li, H., & Yang, Y. (2018). Medblock: Efficient and secure medical data sharing via blockchain. *Journal of medical systems*, 42, 1-11.
- Wang, H., & Song, Y. (2018). Secure cloud-based EHR system using attribute-based cryptosystem and blockchain. *Journal of medical systems*, 42(8), 152.
- Guo, R., Shi, H., Zhao, Q., & Zheng, D. (2018). Secure attribute-based signature scheme with multiple authorities for blockchain in electronic health records systems. *IEEE access*, 6, 11676-11686.
- Uddin, M. A., Stranieri, A., Gondal, I., & Balasubramanian, V. (2018). Continuous patient monitoring with a patient centric agent: A block architecture. *IEEE Access*, 6, 32700-32726.
- Sun, Y., Zhang, R., Wang, X., Gao, K., & Liu, L. (2018, July). A decentralizing attribute-based signature for healthcare blockchain. In *2018 27th International conference on computer communication and networks (ICCCN)* (pp. 1-9). IEEE.
- Zhang, X., & Poslad, S. (2018, May). Blockchain support for flexible queries with granular access control to electronic medical records (EMR). In *2018 IEEE International conference on communications (ICC)* (pp. 1-6). IEEE.
- Yang, G., & Li, C. (2018, December). A design of blockchain-based architecture for the security of electronic health record (EHR) systems. In *2018 IEEE International conference on cloud computing technology and science (CloudCom)* (pp. 261-265). IEEE.
- Thakkar, P., Nathan, S., & Viswanathan, B. (2018, September). Performance benchmarking and optimizing hyperledger fabric blockchain platform. In *2018 IEEE 26th international symposium on modeling, analysis, and simulation of computer and telecommunication systems (MASCOTS)* (pp. 264-276). IEEE.
- Sukhwani, H., Martínez, J. M., Chang, X., Trivedi, K. S., & Rindos, A. (2017, September). Performance modeling of PBFT consensus process for permissioned blockchain network (hyperledger fabric). In *2017 IEEE 36th symposium on reliable distributed systems (SRDS)* (pp. 253-255). IEEE.
- Thakkar, P., & Natarajan, S. (2020). Scaling hyperledger fabric using pipelined execution and sparse peers. *arXiv preprint arXiv:2003.05113*.

- Chen, L., Lee, W. K., Chang, C. C., Choo, K. K. R., & Zhang, N. (2019). Blockchain based searchable encryption for electronic health record sharing. *Future generation computer systems*, 95, 420-429.
- Nguyen, D. C., Pham, Q. V., Pathirana, P. N., Ding, M., Seneviratne, A., Lin, Z., ... & Hwang, W. J. (2022). Federated learning for smart healthcare: A survey. *ACM Computing Surveys (CSUR)*, 55(3), 1-37.
- Keshta, I., & Odeh, A. (2021). Security and privacy of electronic health records: Concerns and challenges. *Egyptian Informatics Journal*, 22(2), 177-183.
- Cifuentes, M., Davis, M., Fernald, D., Gunn, R., Dickinson, P., & Cohen, D. J. (2015). Electronic health record challenges, workarounds, and solutions observed in practices integrating behavioral health and primary care. *The Journal of the American Board of Family Medicine*, 28(Supplement 1), S63-S72.
- Win, K. T. (2005). A review of security of electronic health records. *Health Information Management*, 34(1), 13-18.
- Ancker, J. S., Silver, M., Miller, M. C., & Kaushal, R. (2013). Consumer experience with and attitudes toward health information technology: a nationwide survey. *Journal of the American Medical Informatics Association*, 20(1), 152-156.
- Perera, G., Holbrook, A., Thabane, L., Foster, G., & Willison, D. J. (2011). Views on health information sharing and privacy from primary care practices using electronic medical records. *International journal of medical informatics*, 80(2), 94-101.
- Wikina, S. B. (2014). What caused the breach? An examination of use of information technology and health data breaches. *Perspectives in health information management*, 11(Fall).
- Kruse, C. S., Smith, B., Vanderlinden, H., & Nealand, A. (2017). Security techniques for the electronic health records. *Journal of medical systems*, 41, 1-9.
- Liu, V., Musen, M. A., & Chou, T. (2015). Data breaches of protected health information in the United States. *Jama*, 313(14), 1471-1473.
- Yüksel, B., Küpçü, A., & Özkasap, Ö. (2017). Research issues for privacy and security of electronic health services. *Future Generation Computer Systems*, 68, 1-13.
- Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Decentralized business review*.
- Keele, S. (2007). Guidelines for performing systematic literature reviews in software engineering.

- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group*. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, 151(4), 264-269.
- Houtan, B., Hafid, A. S., & Makrakis, D. (2020). A survey on blockchain-based self-sovereign patient identity in healthcare. *IEEE Access*, 8, 90478-90494.
- Ferrag, M. A., Derdour, M., Mukherjee, M., Derhab, A., Maglaras, L., & Janicke, H. (2018). Blockchain technologies for the internet of things: Research issues and challenges. *IEEE Internet of Things Journal*, 6(2), 2188-2204.
- De Aguiar, E. J., Faiçal, B. S., Krishnamachari, B., & Ueyama, J. (2020). A survey of blockchain-based strategies for healthcare. *ACM Computing Surveys (CSUR)*, 53(2), 1-27.
- Rundo, L., Pirrone, R., Vitabile, S., Sala, E., & Gambino, O. (2020). Recent advances of HCI in decision-making tasks for optimized clinical workflows and precision medicine. *Journal of biomedical informatics*, 108, 103479.
- Tzanou, M. (2017). *The fundamental right to data protection: normative value in the context of counter-terrorism surveillance* (Vol. 71). Bloomsbury Publishing.
- Ahmed, A., Parvez, M. R., Hasan, M. H., Nur, F. N., Moon, N. N., Karim, A., ... & Jonkman, M. (2019, January). An intelligent and secured tracking system for monitoring school bus. In *2019 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-5). IEEE.
- Hoepman, J. H. (2014, June). Privacy design strategies. In *IFIP International Information Security Conference* (pp. 446-459). Berlin, Heidelberg: Springer Berlin Heidelberg.
- van Lieshout, M., Kool, L., van Schoonhoven, B., & de Jonge, M. (2011). Privacy by Design: an alternative to existing practice in safeguarding privacy. *info*, 13(6), 55-68.
- Appari, A., & Johnson, M. E. (2010). Information security and privacy in healthcare: current state of research. *International journal of Internet and enterprise management*, 6(4), 279-314.
- O'Keefe, C. M., & Connolly, C. J. (2010). Privacy and the use of health data for research. *Medical Journal of Australia*, 193(9), 537-541.
- Cavoukian, A. (2020). Understanding how to implement privacy by design, one step at a time. *IEEE Consumer Electronics Magazine*, 9(2), 78-82.
- Sweeney, L. (2005). Privacy-preserving surveillance using selective revelation. *IEEE Intelligent Systems*, 1, 83-84.

- by Design, P. (2018). Effective Privacy Management in the Victorian Public Sector. *Commissioner for Privacy and Data Protection: Melbourne, Victoria, Australia*.
- Semantha, F. H., Azam, S., Yeo, K. C., & Shanmugam, B. (2020). A systematic literature review on privacy by design in the healthcare sector. *Electronics*, 9(3), 452.
- Skinner, G., Chang, E., McMahon, M., Aisbett, J., & Miller, M. (2004, November). Shield privacy Hippocratic security method for virtual community. In *30th Annual Conference of IEEE Industrial Electronics Society, 2004. IECON 2004* (Vol. 1, pp. 472-479). IEEE.
- Spiekermann, S., & Cranor, L. F. (2008). Engineering privacy. *IEEE Transactions on software engineering*, 35(1), 67-82.
- Semantha, F. H., Azam, S., Yeo, K. C., & Shanmugam, B. (2020). A systematic literature review on privacy by design in the healthcare sector. *Electronics*, 9(3), 452.
- Dang, L. M., Piran, M. J., Han, D., Min, K., & Moon, H. (2019). A survey on internet of things and cloud computing for healthcare. *Electronics*, 8(7), 768.
- van Lieshout, M., Kool, L., van Schoonhoven, B., & de Jonge, M. (2011). Privacy by Design: an alternative to existing practice in safeguarding privacy. *info*, 13(6), 55-68.
- Iachello, G., & Hong, J. (2007). End-user privacy in human-computer interaction. *Foundations and Trends® in Human-Computer Interaction*, 1(1), 1-137.
- Cavoukian, A. (2012). Privacy by design [leading edge]. *IEEE Technology and Society Magazine*, 31(4), 18-19.
- Cavoukian, A., & Tapscott, D. (1996). *Who knows: safeguarding your privacy in a networked world*. McGraw-Hill Professional.
- Spitzer, J. (2019). 6.1 M healthcare data breach victims in 2018: 5 of the biggest breaches so far. *Becker's Health IT & CIO Report: Chicago, IL, USA*.
- Donnelly, C. (2019). The GDPR: Why you Need to Adopt the Principles of Privacy by Design. *IT Governance*, 16.
- Gürses, S., Troncoso, C., & Diaz, C. (2011). Engineering privacy by design. *Computers, Privacy & Data Protection*, 14(3), 25.

- Pfitzmann, A., & Hansen, M. (2010). A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management.
- Semantha, F. H., Azam, S., Yeo, K. C., & Shanmugam, B. (2020). A systematic literature review on privacy by design in the healthcare sector. *Electronics*, 9(3), 452.
- Hoepman, J. H. (2014, June). Privacy design strategies. In *IFIP International Information Security Conference* (pp. 446-459). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Sweeney, L. (2002). k-anonymity: A model for protecting privacy. *International journal of uncertainty, fuzziness and knowledge-based systems*, 10(05), 557-570.
- Graf, C., Wolkerstorfer, P., Geven, A., & Tscheligi, M. (2010, November). A pattern collection for privacy enhancing technology. In *The 2nd Int. Conf. on Pervasive Patterns and Applications (PATTERNS 2010)* (pp. 21-26).
- Deng, M., Wuyts, K., Scandariato, R., Preneel, B., & Joosen, W. (2011). A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. *Requirements Engineering*, 16(1), 3-32.
- Organization for Economic Co-operation and Development,. (1980). OECD guidelines on the protection of privacy and transborder flows of personal data. *Paris, France*.
- Fernández-Alemán, J. L., Señor, I. C., Lozoya, P. Á. O., & Toval, A. (2013). Security and privacy in electronic health records: A systematic literature review. *Journal of biomedical informatics*, 46(3), 541-562.
- Cavoukian, A. (2020). Understanding how to implement privacy by design, one step at a time. *IEEE Consumer Electronics Magazine*, 9(2), 78-82.
- Otero, P., Hersh, W., & Ganesh, A. J. (2014). Big data: Are biomedical and health informatics training programs ready?. *Yearbook of medical informatics*, 23(01), 177-181.
- Boulos, M. N. K., Wilson, J. T., & Clauson, K. A. (2018). Geospatial blockchain: promises, challenges, and scenarios in health and healthcare. *International journal of health geographics*, 17.
- Zheng, X., Geng, X., Xie, L., Duan, D., Yang, L., & Cui, S. (2018, February). A SVM-based setting of protection relays in distribution systems. In *2018 IEEE Texas Power and Energy Conference (TPEC)* (pp. 1-6). IEEE.
- Lee, S. H., & Yang, C. S. (2018). Fingernail analysis management system using microscopy sensor and blockchain technology. *International Journal of Distributed Sensor Networks*, 14(3), 1550147718767044.

- Tzu, L. (2017). MATRIX Technical Whitepaper. *MATRIXTechnicalWhitePaper.pdf*.
- O'Leary, D. E. (2013). Artificial intelligence and big data. *IEEE intelligent systems*, 28(2), 96-99.
- Bhattacharya, P., Tanwar, S., Bodkhe, U., Tyagi, S., & Kumar, N. (2019). Bindaas: Blockchain-based deep-learning as-a-service in healthcare 4.0 applications. *IEEE transactions on network science and engineering*, 8(2), 1242-1255.
- Dey, S. (2018, September). Securing majority-attack in blockchain using machine learning and algorithmic game theory: A proof of work. In *2018 10th computer science and electronic engineering (CEEC)* (pp. 7-10). IEEE.
- Gupta, I. (2020). Decentralization of artificial intelligence: analyzing developments in decentralized learning and distributed AI networks. *arXiv preprint arXiv:1603.04467*.
- Corea, F. (2017). The convergence of AI and Blockchain: what's the deal. Retrieved October, 6, 2020.
- Wang, R., Luo, M., Wen, Y., Wang, L., Raymond Choo, K. K., & He, D. (2021). The applications of blockchain in artificial intelligence. *Security and Communication Networks*, 2021, 1-16.
- Abdellatif, A. A., Samara, L., Mohamed, A., Erbad, A., Chiasserini, C. F., Guizani, M., ... & Laughton, J. (2021). Medge-chain: Leveraging edge computing and blockchain for efficient medical data exchange. *IEEE Internet of Things Journal*, 8(21), 15762-15775.
- Gai, K., Wu, Y., Zhu, L., Xu, L., & Zhang, Y. (2019). Permissioned blockchain and edge computing empowered privacy-preserving smart grid networks. *IEEE Internet of Things Journal*, 6(5), 7992-8004.
- Yang, R., Yu, F. R., Si, P., Yang, Z., & Zhang, Y. (2019). Integrated blockchain and edge computing systems: A survey, some research issues and challenges. *IEEE Communications Surveys & Tutorials*, 21(2), 1508-1532.
- De Filippi, P. (2016). The interplay between decentralization and privacy: the case of blockchain technologies. *Journal of Peer Production*, Issue, (7).
- Nosowsky, R., & Giordano, T. J. (2006). The Health Insurance Portability and Accountability Act of 1996 (HIPAA) privacy rule: implications for clinical research. *Annu. Rev. Med.*, 57, 575-590.
- Li, X., Jiang, P., Chen, T., Luo, X., & Wen, Q. (2020). A survey on the security of blockchain systems. *Future generation computer systems*, 107, 841-853.

- Li, C., & Palanisamy, B. (2019, June). Incentivized blockchain-based social media platforms: A case study of steemit. In *Proceedings of the 10th ACM conference on web science* (pp. 145-154).
- Madine, M. M., Battah, A. A., Yaqoob, I., Salah, K., Jayaraman, R., Al-Hammadi, Y., ... & Ellahham, S. (2020). Blockchain for giving patients control over their medical records. *IEEE Access*, 8, 193102-193115.
- Saidi, H., Labraoui, N., Ari, A. A. A., Maglaras, L. A., & Emati, J. H. M. (2022). DSMAC: Privacy-aware Decentralized Self-Management of data Access Control based on blockchain for health data. *IEEE Access*, 10, 101011-101028.
- Makridakis, S., & Christodoulou, K. (2019). Blockchain: Current challenges and future prospects/applications. *Future Internet*, 11(12), 258.
- Mendonca, S. N. (2018). Data security in cloud using AES. *Int. J. Eng. Res. Technol*, 7.
- Kumari, A., Yahya Abbasi, M., Kumar, V., & Khan, A. A. (2019). A secure user authentication protocol using elliptic curve cryptography. *Journal of Discrete Mathematical Sciences and Cryptography*, 22(4), 521-530.
- Feldman, J., Misenar, S., & Conrad, E. (2016). *Eleventh Hour CISSP®: Study Guide*. Syngress.
- Azad, T. B. (2008). Understanding XenApp Security. *Securing Citrix Presentation Server in the Enterprise*, 259-316.
- Mendonca, S. N. (2018). Data security in cloud using AES. *Int. J. Eng. Res. Technol*, 7.
- Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE conference on computer vision and pattern recognition*(pp. 2097-2106).
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- Hendriks, S. (2016). *Internet of Things: how the world will be connected in 2025* (Master's thesis).

- Zhao, Y., Cui, M., Zheng, L., Zhang, R., Meng, L., Gao, D., & Zhang, Y. (2019). Research on electronic medical record access control based on blockchain. *International Journal of Distributed Sensor Networks*, 15(11), 1550147719889330.
- Chen, L., Lee, W. K., Chang, C. C., Choo, K. K. R., & Zhang, N. (2019). Blockchain based searchable encryption for electronic health record sharing. *Future generation computer systems*, 95, 420-429.
- Qin, Q., Jin, B., & Liu, Y. (2021). A secure storage and sharing scheme of stroke electronic medical records based on consortium blockchain. *BioMed Research International*, 2021.
- Khalaf, O. I., Abdulsahib, G. M., Kasmaei, H. D., & Ogudo, K. A. (2020). A new algorithm on application of blockchain technology in live stream video transmissions and telecommunications. *International Journal of e-Collaboration (IJeC)*, 16(1), 16-32.
- Thwin, T. T., & Vasupongayya, S. (2019). Blockchain-based access control model to preserve privacy for personal health record systems. *Security and Communication Networks*, 2019.
- Boumezbeur, I., & Zarour, K. (2022). Privacy-Preserving and Access Control for Sharing Electronic Health Record using Blockchain Technology. *Acta Informatica Pragensia*, 11(1), 105-122.
- Shen, B., Guo, J., & Yang, Y. (2019). MedChain: Efficient healthcare data sharing via blockchain. *Applied sciences*, 9(6), 1207.
- Dwivedi, A. D., Srivastava, G., Dhar, S., & Singh, R. (2019). A decentralized privacy-preserving healthcare blockchain for IoT. *Sensors*, 19(2), 326.
- Rajput, A. R., Li, Q., & Ahvanooy, M. T. (2021, February). A blockchain-based secret-data sharing framework for personal health records in emergency condition. In *Healthcare* (Vol. 9, No. 2, p. 206). MDPI.
- Egala, B. S., Pradhan, A. K., Badarla, V., & Mohanty, S. P. (2021). Fortified-chain: a blockchain-based framework for security and privacy-assured internet of medical things with effective access control. *IEEE Internet of Things Journal*, 8(14), 11717-11731.
- Sri Vigna Hema, V., & Kesavan, R. (2019). ECC based secure sharing of healthcare data in the health cloud environment. *Wireless Personal Communications*, 108, 1021-1035.
- Mohsan, S. A. H., Razzaq, A., Ghayyur, S. A. K., Alkahtani, H. K., Al-Kahtani, N., & Mostafa, S. M. (2022). Decentralized Patient-Centric Report and Medical Image

Management System Based on Blockchain Technology and the Inter-Planetary File System. *International Journal of Environmental Research and Public Health*, 19(22), 14641.



APPENDIX Y:
LIST OF PUBLISHED PAPERS

Haddad, A., Habaebi, M. H., Islam, M. R., & Zabidi, S. A. (2021, August). Blockchain for Healthcare Medical Records Management System with Sharing Control. In *2021 IEEE 7th International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)* (pp. 30-34). IEEE.

Haddad, A., Habaebi, M. H., Islam, M. R., Hasbullah, N. F., & Zabidi, S. A. (2022). Systematic review on ai-blockchain based e-healthcare records management systems. *IEEE Access*.

Haddad, A., Habaebi, M. H., Suliman, F. E. M., Elsheikh, E. A., Islam, M. R., & Zabidi, S. A. (2023). Generic Patient-Centered Blockchain-Based EHR Management System. *Applied Sciences*, *13*(3), 1761.

