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 سائدًا إلى حد كبير بسبب التحسن في الصناعة والمستويات العالية من تنقل المرضى. نظرًا 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 آمنًا طوال الوقت. نظرًا 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 مخططًا آمنًا ومحسّنًا 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 منحنى 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 نظرًا 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 أيضًا كيف يمكّن 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.

أخيرًا، 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 قويًا بشكل لا يصدق لدعم العيادات وفي تحويل رعاية المرضى بشكل عام ويتم تطبيقه بشكل متزايد للكشف عن الميزات المهمة سريريًا في الصور بما يتجاوز ما يمكن أن تراه العين المجردة. تعد صور الصدر بالأشعة السينية واحدة من أكثر الطرق السريرية شيوعًا لتشخيص العديد من الأمراض. تعد ميزة التعلم العميق المدمجة والمعززة المقترحة حلاً خفيف الوزن مطورًا 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 مرضًا لتوفير نظرة ثاقبة لأبحاث التصوير الشعاعي للصدر في المستقبل.

أخيرًا، تم اختبار النظام المقترح في بيئة @ Microsoft Windows عن طريق تجميع نموذج أولي ذكي للعقد باستخدام Truffleونشره على Ethereum باستخدام .Web3 تم تقييم النظام المقترح من حيث تكاليف تخزين البيانات الطبية المتوقعة لـ IPFS علىblockchain ، 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.

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

………… ………… ………2023-03-13………….
Signature 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

LIST OF TABLES

LIST OF FIGURES

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.

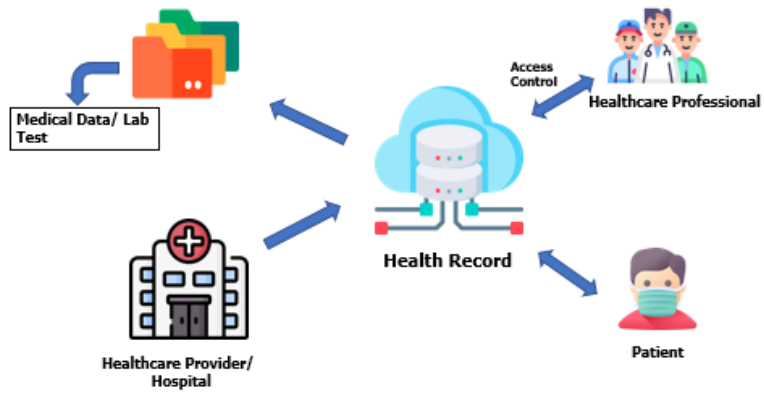

Figure 1.1. Overview of the current system.

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

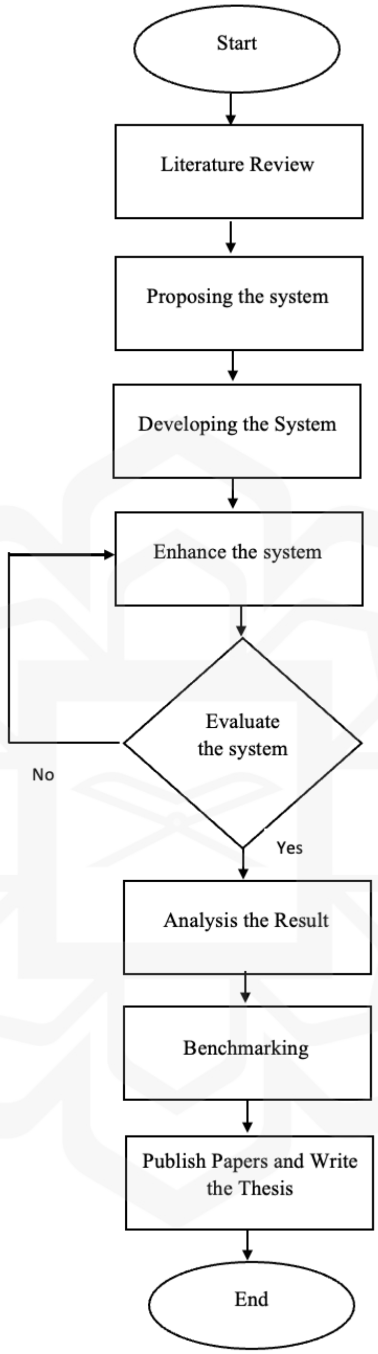

Figure 1.2 System Flowchart.

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