2021 MS-TM

A proposal on blockchain-based electronic records sharing system for healthcare big data

Hwang, Byoung Yong

June 2022

2021 MS-TM

블록체인 기반 의료 빅데이터 전자기록 공유 시스템 제안

황 병 용

June 2022

Executive Summary (English)

In our society, innovative changes are rapidly taking place due to the convergence of cutting-edge information and communication technologies such as artificial intelligence (AI), Internet of Things (IoT), Big Data, and Cloud Computing.

The 4th Industrial Revolution will create a hyper-connected society between people and people, between things and people, and between things and things. By this disruptive change, if the society becomes a place where all information is fluid and easy to access, the security of information must be an essential factor. The development of internet technology has given our lives an excellent result of the convenience of using information. On the other hand, however, it brought the paradoxical result of the risk of information leakage due to easy access to information. Numerous costs and efforts are being spent to strengthen information security, from individuals, businesses, and government organizations. It seems to be an incompatible topic for everyone to use information easily and maintain information security at the same time.

However, in order to solve this dilemma, blockchain is attracting attention as an innovative technology. The centralized system for information management by using 3rd generation technology could be a way of controlling access rights through various security equipment and software for storing and sharing information. On the other hand, the core concept of blockchain as 4th generation technology is not for managing information centrally, but for managing information through decentralized approach. In other words, distributed ledger technology can enhance security by sharing encrypted data to all participants. In particular, the application of blockchain technology can benefit all participants in terms of speed, transparency, security, convenience, and cost.

Blockchain technology is currently most widely used in the financial sector, and efforts are underway to introduce it in the medical sector. In particular, research is being actively conducted to apply it to Hospital Information System (HIS) and Electronic Health Record (EHR). Medical information used under the current centralized management system has a risk of unauthorized leakage or misuse of it by external attacks or internal persons. In order to solve this single point of failure problem, blockchain technology is being applied. In addition, blockchain technology is needed to efficiently utilize medical information beyond simply protecting information. This technology is the most efficient and safe way to transfer ownership over medical information from medical institutions to patients. If blockchain technology is fully utilized, patients will be able to independently own and use their medical information.

However, most of the blockchain research applied to the medical field is focusing on the concept,

definition, characteristics, applicability, and prospect of the technology. At the present stage, empirical studies are insufficient. Relevant institutions and venture companies in each country are researching platforms to utilize medical information, but they remain at the entry level of technology that simply proves patient's medical records and shares it in a narrow range of feasibility. It could be inferred that the blockchain technology in medical fields is still in the early stage in terms of commercialization.

The '**blockchain-based electronic records system** for healthcare big data' is created based on Ethereum. Because the Ethereum blockchain platform is easy to apply various applications and has a great potential for continuous development. In order to utilize medical big data, it must be able to accommodate the development of surrounding technologies and various trials. And openness that all stakeholders can participate must be guaranteed. In this respect, Ethereum is an optimized blockchain technology. Ethereum blockchain technology is developing even today. 'Ethereum 2.0' is being prepared to solve the 'scalability' problem, which is the biggest limitation of the current blockchain technology. The development of scalability in blockchain technology will provide a practical basis for processing medical information which have large-volume characteristics in the future. But this proposal has limitations based on the current Ethereum 1.1 version. However, if the goal is to propose only a conceptual design rather than an in-depth technical approach, the current version of Ethereum platform can sufficiently give a good foundation to achieve the purpose.

In addition, it can be said that this proposal is useful in terms of not just a study to maintain security through encryption of medical information, but a practical study that can easily utilize medical information through mobile platform. The goal of this proposal is to create the concept of a blockchain-based system that provides easy access to EHRs by healthcare professionals or patients without relying on a centralized supervisory system. This study suggests a platform that can store and share medical big data on the blockchain.

The composition of this proposal is as follows. Chapter 1 describes the simple history of medical records and the motivation for this study is presented. At end of this chapter, the limitations in using medical big data are described, and the utility of the blockchain platform to overcome those limitations is summarized. In addition, a methodology is presented on what type of block chain platform should be designed in order to effectively utilize medical Big Data.

In Chapter 2, the background of the study was explained to expand on the research motivation. In particular, the reason of difficulties in using big data are described and it would be summarized that how blockchain technology can be used for the purpose of solving them, based on existing studies.

Chapter 3 summarizes the type of medical record system currently used under centralized management system and its problems. And to solve these problems, the Ethereum platform and Hyperledger Fabric platform applied to the EHR system is explained. As a core part of this chapter, two papers in which

research was conducted using each platform are reviewed. After reviewing two studies, the current research challenge and critical points for improvement was clearly evident for the further research. At the end of this chapter, it is given various trials for integration of the blockchain into medical big data, which is the central topic of this proposal. Those reviews set a good example for a novel blockchain-based EHR system of medical Big Data.

In conclusion, it was found that the direction of this proposal should be for improving scalability to utilize medical Big Data on the block chain and for being designed on mobile environment to increase usability and accessibility. However, since the level of professional technology is required for holistic design of blockchain-based mobile EHR sharing system, it must be decided to which level of technical approach is proper in this proposal. The in-depth technical approach for realizing the proposed system will be conducted with an expert group in future studies.

In Chapter 4, firstly the previous trial of a blockchain-based EHR sharing system to utilize medical big data is presented. To give a foundation for better understanding the blockchain-based EHR system design, relative concepts would be explained, and the system architecture and operation flow is explained. The structure of smart contract, which is the core part of this platform, and an actual simulation which proposed platform can be run in a mobile environment would be explained.

In Chapters 5, in the conclusion part, the core benefits of the proposed blockchain-based EHR are summarized. Some points which could be improved compared to the existing platforms, and the usefulness of this proposed platform are explained, and the limitations of the proposed system are discussed, and future research directions are suggested. General issues which would be faced in medical field to develop EHR system are also mentioned regarding ‘Scalability limitations’, ‘Adoption and enticements for stakeholders’, and ‘Regulatory consideration and compliance’.

Executive Summary (Korean)

현재 우리 사회는 인공지능(Artificial Intelligence), 사물인터넷(IOT: Internet of Things), 빅데이터(Big Data), 클라우드 컴퓨팅(Claud computing)등 첨단정보통신기술이 사회 경제 전반에 융합되어 혁신적인 변화가 빠르게 일어나고 있다.

4차 산업혁명은 사람과 사람간, 사물과 사람간, 사물과 사물간 초연결사회를 창조할 것이다. 이러한 환경에서, 모든 정보들이 유동적이고 접근이 쉬워지는 사회가 된다면 정보의 보안은 필수적인 요인이 된다. 정보 기술의 발전은 정보의 활용 편의성이라는 훌륭한 결과물을 우리의 삶에 부여했다. 그러나 다른 면에서는 쉬운 정보 접근성에 의한 정보 유출의 위험이라는 역설적인 결과를 가져왔다. 개인, 기업체, 정부단체에 이르기까지 정보의 보안을 강화하기 위해 수많은 비용과 노력이 소비되고 있다. 정보를 만인이 보편적으로 이용하면서 동시에 정보의 보안을 유지해야 하는 것은 양립할 수 없는 주제인 것으로 보인다.

그러나 이러한 딜레마(dilemma)를 해결하기 위해서 블록체인이 혁신적인 기술로써 주목을 받고 있다. 3차 혁명 기술을 이용하는 중앙 집중관리 방식은 정보의 저장 및 공유를 위해 다양한 보안장비와 소프트웨어를 통해 접근 권한을 통제하는 방식이라고 할 수 있다. 반면에 블록체인의 개념은 거래 정보를 중앙에서 처리하는 것이 아니라 탈중앙화를 통해 정보를 관리하는 것이다. 그리고 암호화된 데이터를 모든 참여자들에게 공유함으로써 보안을 강화한다. 다시 말해서, 모든 참여자들이 암호화된 문서를 공유하게 하는 '분산원장' 기술을 통해서 보안을 강화할 수 있다. 특히 블록체인 기술을 적용함으로써 속도, 투명성, 보안성, 편의성, 그리고 비용적인 측면에서 모든 참여들에게 이익을 가져다 줄 수 있다.

이러한 장점을 지닌 블록체인 기술은 현재 금융업 분야에 가장 널리 사용되고 있으며 또한 의료분야에서도 도입하기 위한 노력이 진행 중이다. 특히, 병원정보시스템(Hospital Information System; HIS)과 전자의무기록(Electronic Health Record; EHR) 등에 적용하기 위한 연구가 활발하게 진행되고 있다. 현재의 중앙집중식 관리시스템 아래에서 통용되는 의료정보는 외부의 공격이나 내부 관리자에 의해 의료정보가 무단으로 유출되거나 오남용 될 위험이 있다. 이러한 단일지점 실패 문제(a single point of failure problem)를 해결하기 위해서 블록체인 기술을 적용하려는 것이다. 또한, 단순하게 정보를 보호하는 것을 넘어서서 의료정보를 효율적으로 활용하기 위해서도 블록체인 기술이 필요하다. 이 기술은 의료정보에 대한 권한을 의료기관에서 환자에게 옮길 수 있는 가장 효율적이고 안전한 방법이다. 블록체인 기술이 완전하게 활용된다면 환자는 자신의 의료정보를 주체적으로 소유하고 활용할 수 있게 된다.

그러나 의료분야에 적용되는 대부분의 블록체인 연구는 블록체인의 개념, 정의, 특성,

그리고 적용 가능성, 전망 등에 관한 연구가 주류를 이루고 있다. 현 단계에서는 실증적인 연구가 미흡한 편이다. 각국의 관련 기관이나 벤처기업들이 의료정보를 활용하기 위한 플랫폼을 연구하고 있지만 단순히 환자의 의료기록을 증명하고 간단한 기록을 공유하는 정도의 기술 수준에 머물러 있다. 이것은 아직 블록체인 기술이 상용화 측면에서 초기 단계에 해당하기 때문이라고 유추할 수 있다.

본 제안서에서 제시되는 '블록체인 기반 의료 빅데이터 전자기록 공유 시스템 (blockchain-based electronic records system for healthcare big data)'은 이더리움 (Ethereum) 기반을 통해서 만들어졌다. 왜냐하면 이더리움 블록체인 플랫폼은 다양한 애플리케이션을 적용하기가 용이하고 지속적인 발전 가능성을 가지고 있기 때문이다. 의료 빅데이터를 활용하기 위해서는 주변 기술의 발전을 잘 수용할 수 있어야 하고 다양한 시도를 자유롭게 행할 수 있어야 한다. 그리고 모든 stakeholder들이 참여할 수 있는 개방성(openness)이 보장되어야 한다. 이러한 점에 있어서 이더리움은 최적화된 블록체인 기술이다. 현재도 이더리움의 블록체인 기술은 발전을 계속하고 있다. 특별히 현재의 블록체인 기술에 있어서 가장 큰 한계인 확장성(scalability) 문제를 해결하기 위해 2.0 버전을 준비중이다. 이 부분의 발전은 앞으로 대용량의 특성을 가진 의료정보를 처리할 수 있는 실제적인 기반을 마련해 줄 것이다. 따라서 본 제안서는 현재의 이더리움 1.1 버전의 기반으로 제안되는 한계를 가지고 있다. 그러나 깊이 있는 기술적 접근이 아닌 설계 아이디어만 제안하는 것을 목표로 한다면, 현재의 이더리움 기반으로도 충분하게 목적을 달성할 수 있다고 할 수 있다.

또한, 이 제안이 단순히 의료정보의 암호화를 통한 보안성 유지만을 위한 연구가 아니라 의료 정보를 모바일 플랫폼을 통해 손쉽게 활용할 수 있는 실증적인 연구라는 관점에서 유용성이 있다고 할 수 있다. 이 제안서의 목표는 중앙 집중식 감독 시스템에 의존하지 않고 EHR에 의료전문가나 환자가 쉽게 접근할 수 있는 블록체인 기반 시스템의 개념을 만드는 것이다. 또한, 빅데이터의 특성을 가지는 의료데이터를 블록체인 상에서 활용할 수 있는 플랫폼을 제안하는 것이다.

본 제안서의 구성은 다음과 같다. 1장에서는 의무기록(Medical records)의 간략한 역사와 본 연구를 제안하게 된 동기가 제시되었다. 이를 위해서 의료 빅데이터를 활용하는 점에 있어서의 한계를 기술하였고, 이를 극복하기 위한 블록체인 플랫폼의 활용성에 대해서 정리하였다. 또한, 의료 빅데이터를 유용하게 활용하기 위해서는 어떤 형태의 블록체인 플랫폼을 연디자인 해야 할 지에 대한 방법론을 제시하였다.

2장에서는 연구동기를 보다 확장하기 위해서 연구의 배경에 대해서 설명하였다. 특별히 빅데이터를 활용에 있어서의 난제와 이를 해결하기 위한 목적으로 블록체인 기술이 어떻게 활용될 수 있는 지를 기존 연구들을 토대로 분석하였다.

3장에서는 현재 사용되는 중앙 집중식 관리 하의 의무기록시스템의 형태와 그에 따른 문제점들에 대해서 정리하였다. 그리고 이러한 문제점을 해결하기 위해서 EHR 시스템에 적

용하고 있는 Ethereum 플랫폼과 Hyperledger Fabric 플랫폼의 특징에 설명하였다. 또한 각각의 플랫폼을 사용하여 실제로 연구가 진행된 논문 2편을 검토하였다. 이 검토를 통해서 현재의 연구 상황과 추후 개선해야 할 점들을 찾아 내려고 하였다. 이 장의 마지막 부분에서는 본 제안의 중심 논제인 의료분야의 빅데이터를 활용하기 위해 블록체인에 integration 한 사례들을 검토해 보았다. 이러한 검토는 블록체인 기반의 빅데이터를 위한 EHR을 제시하는데 좋은 본보기를 제시해 주었다.

결론적으로, 이 시스템의 방향은 빅데이터를 블록체인 상에서 활용하기 위한 확장성(scalability)을 향상시켜야 하는 것과 활용성을 높이기 위해 모바일 환경을 기반(mobile-based)으로 디자인되어야 것임을 알게 되었다. 다만, 이 블록체인 기반 mobile EHR sharing system을 완성하기 위해서는 전문적인 기술의 수준이 요구됨으로, 개념적인 설계의 수준까지 제시하는 것으로 하였다. 제안된 시스템을 실현하기 위한 기술적인 방면은 기회가 된다면 추후 연구에서 전문가 그룹과의 연구를 통해 진행할 것이다.

4장에서는 선행연구들의 검토를 통해서 이상적인 의료 빅데이터를 활용하기 위한 블록체인 기반 EHR sharing system의 디자인이 제시되어 있다. 설계를 이해하기 위한 개념의 설명과 System Architecture 및 구동 흐름에 대해서 설명했다. 이 플랫폼의 핵심 부분인 Smart contract 부분과 이 플랫폼이 mobile 환경에서 구동될 수 있는 실제 예시를 들어 설명하였다.

마지막으로 5장에서는 결론의 부분으로서, 제안된 블록체인 기반의 EHR 시스템이 기존 방법에 비해서 개선된 점들에 대해 설명한다. 또한, 제안된 시스템의 유용성에 대해 정리해서 언급하고 있다. 또한, 제안된 시스템의 한계점에 대해서 논의하고 앞으로의 연구방향에 대해서 제시했다. 의료 분야에서 EHR 시스템을 개발할 때 당면하게 되는 일반적인 문제들(Scalability limitations, Adoption and enticements for stakeholders, Regulatory consideration and compliance)에 관해서도 언급했다.

Table of Contents

List of Tables

List of Figures

I. Introduction

1. Motivation of the Proposal

Currently, our society is undergoing innovative changes due to the convergence of advanced information and communication technologies. In this era, digital technology is being applied to various industrial fields, leading to a different type of innovation (Kim, Park, & Yang, 2020). It could be said that the use of Big Data and blockchain technology is by far the center of this revolution.

In particular, the use of Big Data in the medical field is becoming more important. But there are various challenges and issues associated with Big Data techniques and applications in healthcare domain. Security and privacy have been considered as important issues since medical data often involves different types of sensitive personal information, e.g., age, addresses, disease, family information.

Recently, blockchain technology has emerged as an attractive solution for providing security and privacy in Big Data management. Blockchain has the great potential to transform current Big Data management systems by providing flawless security and network management for enabling newly emerging Big Data services and relative applications in healthcare field.

However, the blockchain technology in the healthcare field is still at the level of simply transferring personal medical information to a mobile platform. The technological approach for large-capacity data remains at a rudimentary level. This current state of development in this domain presents a practical difficulty in sharing data for multi-institutional research or being used for accurate statistical data in national institutions. These results cause a major hindrance to the democratization of information, which is the strongest point of blockchain technology, thereby fading the meaning of the blockchain-based EHR system.

In addition, most of the existing studies do not use an online platform, and most of them propose a mechanism that is used only on web pages and between researchers. This aspect might be centered on only one side of the EHR and PHR systems using blockchain technology. When patients easily use and participate in the creation of medical information, valuable medical information will be accumulated, and utilization will increase. It is necessary to develop a mobile-based EHR system for such patient participation.

A new platform is needed to compensate for the shortcomings in the above-mentioned major two point. It is a platform that allows for unlimited sharing of large-capacity data and allows patients to conveniently use Medi-Cal information. This proposal was motivated to solve these unmet needs.

2. Method of research

Based on the review of previous studies, this proposal consider how blockchain can be used for medical Big Data and present a model. The main purpose of this paper is to examine the latest research and the preceding studies of blockchain for Big Data applications and propose a novel blockchain-based EHR sharing system. The contribution and functions provided by this work can be described as follows.

- Presenting an overview of blockchain and Big Data as well as the motivations behind the use of blockchain for Big Data.
- Reviewing two main blockchain platform for medical Big Data, including example architectures on medical field.
- Proposing concepts of EHRs sharing system based on blockchain and interplanetary file system (IPFS) on a mobile environment.
- Discussing a number of key features and challenges of proposed blockchain-based EHR sharing system for the medical Big Data. This proposal also high-light open research opportunities that provide a roadmap for future research.

In order to achieve the above object, this study was conducted as follows. First, the overall relative blockchain technology that has been developed so far would be summarized up. Second, through the previous review, it would be examined what are the limitations for the utilization of medical Big Data and the usefulness of the blockchain to overcome those challenges. Third, the current EHR system would be reviewed to find the things to improve, and two latest and advanced blockchain-based platforms are studied as case studies in detail. After comparing the pros and cons of the two platforms, better and more advanced model to previous studies would be presented as a core part of this proposal.

II. Literature survey

1. The brief history of healthcare records

The earliest medical records began in Egypt around 3000, B.C. The Edwin Smith Papyrus, discovered in 1862 outside of Luxor, Egypt, is the oldest known surgical text in the history of civilization. The history of these records developed rapidly during the Roman period, and after that, scientific records could be produced along with the development of modern scientific revolution. The medical records began to be systematized in earnest in the 20th century. Standardization of medical records was established by the American College of Surgeons (ACOS).

Until the mid-20th century, patient charts were prepared in the form of handwritten files. Throughout the late 20th century, patient charting began to be moved into electronic systems. In the 21st century, patient data can be accessed and shared seamlessly among providers caring for the patient, through EHRs. As the digitalization of medical records advances, it gives stakeholders a dilemma. Information is concentrated in the central hub, and there are various cases in which information is easily leaked or altered by suspicious third parties. EHRs may contain sensitive patient information, and if the medical records are leaked by an attacker, it can cause serious economic and human damage. Therefore, a secure EHR sharing system is required, and a cloud based EHR sharing system is being actively studied (Kim, M., Yu, S., Lee, J., Park, Y., & Park, 2020).

Cloud-based systems can share information efficiently, but due to their centralized nature, they can become a major target for attackers (Park, Y., and Park, Y., 2016) and a single point of failure problem may occur. EHRs can be maliciously accessed by unauthorized entities without patient consent, which has detrimental effects on data integrity, privacy, and security of cloud e-health systems. Moreover, in recent years, the number of patients using medical information in a mobile environment is increasing. In such a mobile cloud environment, there is no choice but to become more vulnerable to such external attacks. Therefore, a decentralized EHR sharing system is required.

Blockchain-based sharing system provides various security features for EHRs with great advantages over conventional solutions. First, the blockchain constructs immutable ledgers of transactions for data sharing system (M. Hölbl, M. Kompara, A. Kamišalic, and L. N. Zlatolas, 2018). Second, sharing system using blockchain can achieve the transparency property with the ability of solving effectively the issue of data leakage which can be caused by curious servers. Third, the use of blockchain-based smart contract can achieve the authentication and user verification property. Final, blockchain coupled with the smart contract technology eliminates the reliance on central servers or third parties to ensure fairness among transaction parties.

Additionally, the advantage of the blockchain-based EHR sharing system can be suitable for the study of Big Data in which data is diversified and vast. The most important point in Big Data analysis is the process of data collection and processing. Acquiring reliable data could be the most important factor in making decisions, which is the purpose of Big Data. The access control by smart contract of blockchain is useful for analyzing and structuring such Big Data collected from various IoMT devices or mobile gadgets.

But in the sharing of EHRs, the limitations of current blockchain technology are revealed. The problem of acceptability (scalability) of Big Data is the apex that the latest blockchain technology is solving (T. McConaghy, R. Marques, A. Muller, D. De Jonghe, T. McConaghy, G. McMullen, R. Henderson, S. Bellemare, and A. Granzotto, 2016). In relation to acceptability, the use of Big Data on the blockchain causes the problem of delay. In many studies, it has been found that as the amount of information increases, the latency on the block chain increases (X. Fan and Y. Huo, 2020). The limitations of blockchain technology could be an important point to develop a model for sharing medical Big Data. Therefore, this proposal suggests a blockchain-based EHR sharing system focused on medical Big Data to improve such limitations (scalability and latency).

The research space where EHR and blockchain intersect is still in its infancy stage, with the first blockchain EHR introduced in 2016 (Azaria A, Ekblaw A, Vieira T, and Lippman A.). Research on the integration of medical Big Data and blockchain has not progressed much. Although there are many existing studies, it could not be found anything about the application model of blockchain for use of medical Big Data. The study of (E. Karafiloski and A. Mishev, 2017) reviews the blockchain for Big Data applications, but it has been a long time and no follow research has been carried out since then.

The research in (Alrebdi, N., Alabdulatif, A., Iwendi, C., and Lian, Z., 2022) mainly examines how blockchain is used to solve security problems. Other surveys in (D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne, 2020) also mentioned the interaction between blockchain and Big Data but provide only a brief introduction to the topic without presenting models or going into in-depth discussions. A review of these prior studies gave a motivation to research a blockchain-based data sharing system for Big Data in the medical field.

2. Conventional Methods for Healthcare Records

With the development of cloud computing and Information and Communications Technology (ICT) technology, numerous scientific works are being carried out to develop secure EHR system. The most basic tool is Attribute-Based Encryption (ABE)[1], which is used to store personal medical information in

[1] Attribute-based encryption is a type of public-key encryption in which the secret key of a user and the ciphertext are dependent upon attributes (e.g., the country in which they live, or the kind of subscription they have). In such a system, the decryption of a

quasi-trusted data centers.

In 2013, Li M, Yu S, Zheng Y, Ren K, and Lou W. (2013) have suggested a patient-centered system for managing healthcare data and for improving security features by using the ABE method to encode EHR information. Hua J., Zhu H., Wang F., Liu X., and Lu R. (2019) have suggested that *CINEMA*: *Efficient and Privacy-Preserving Online Medical Primary Diagnosis with Skyline Query*, a protected online healthcare evaluation system, allows users to initiate cloud server request activities without decoding personal information using fast, secure permutation and related technologies. Nevertheless, *CINEMA* needs immense computational and storage capacity from cloud data centers to allow users from worldwide to access the web services simultaneously.

Healthcare professionals today primarily use EHR to observe patient data using a client-server architecture. In this form of medical data management system, hospitals are the main data source. Although these strategies provide secure cloud storage and fine-grained access control, some issues still exist in the system, such as preventing internal malicious activity and cloud server malfunctions.

ciphertext is possible only if the set of attributes of the user key matches the attributes of the ciphertext.

II. Research background

This section discusses background information on blockchain and Big Data to get better understanding of this proposal. This section also reviews the characteristics and limitations of Big Data and summarizes the benefits of blockchain technology to overcome limitations of use of Big Data in medical field.

1. Big Data and the Challenges of Big Data Utilization

Big Data could be defined as large amounts of data, both structured and unstructured, usually stored in the cloud or in data centers, which are then utilized by companies, organizations, startups, and even the government for various purposes. To utilize data means cleaning it and then analyzing it, finding patterns, connection, trends and correlations, to produce insights and decision. This is what's called Big Data analytics. Big Data is also commonly described by its qualities, also known as the 4Vs (A. Jindal, N. Kumar, and M. Singh, 2020).

- **Volume**: Volume simply means the insurmountable amounts of data due to improvements to technology and data storage (cloud storages, better processes, etc.). Some questions benefit from huge amounts of data, with the sheer volume of data, it negates small messiness or inaccuracies.
- **Velocity**: Data is generated at astonishing rates, related to computer's speed and capability increasing (Moore's Law). Real-time information makes swift decisions based on updated and informed predictions.
- **Variety**: Variety is the characteristic of Big Data in wide range of data of different formats and types easily collected, in an era of social media, the internet and various IoT. Variety gives us ability to ask new questions and form new connections, questions that were previously inaccessible
- **Veracity**: Veracity is inconsistencies and uncertainty of data (unstructured data — images, social media, video, etc.). Messy and unstructured data give rise to the possibility of hidden correlations. Perhaps the most promising benefit of more data is to identify hidden correlations.

In addition to the advantages of Big Data described above, there are various challenges in analyzing Big Data due to its characteristics. Regarding big data processing, one usually faces several challenges, which may include the curse of modularity (i.e., not available to store/load the complete data in memory and hard disk), the curse of class imbalance (i.e., there may exist different data distributions), the curse of dimensionality (i.e., the dataset has many features and attributes) (A. Oussous, F.-Z. Benjelloun, A. A. Lahcen, and S. Belfkih, 2018). Moreover, data non-linearity, variance and bias, and computing availability are also considered as challenges associated with the **volume**. The major challenges caused by **variety** may include data locality, data heterogeneity, dirty and noisy data (H. V. Jagadish, J. Gehrke, A. Labrinidis, Y. Papakonstantinou, J. M. Patel, R. Ramakrishnan, and C. Shahabi, 2014). Here, data locality expresses that the complete data cannot be stored in a data center and is typically distributed over many

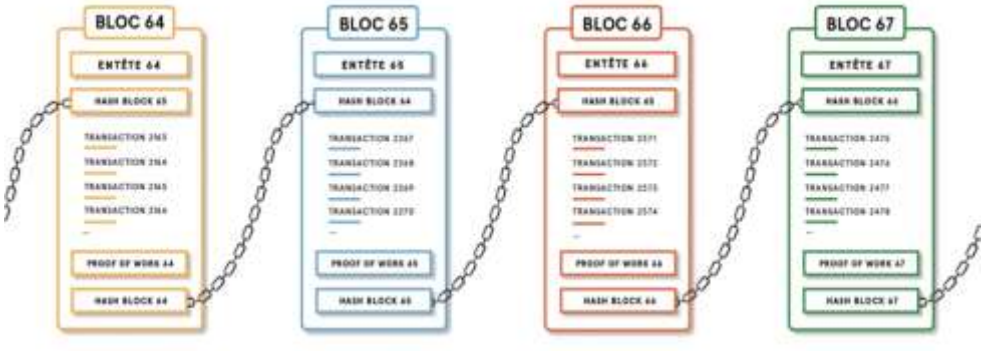
physical locations. Data heterogeneity is referred to as various heterogeneous sources of data, thus having different data types, formats, models, and semantics. Dirty and noise data means that the data can contain inaccurate or distorted information, which would be caused by data collection methods, data sources, and generation time. **Veracity** refers to the quality aspect since the data can be collected from multiple sources, which may include low-quality and noisy samples. To improve the quality and analytical accuracy of Big Data, the challenges of data provenance, uncertainty, dirty and noisy data should be effectively tackled.

Challenges encountered in Big Data analysis could be more severe in medical records. This is because most of medical records exist as large files, and information is gathered through various tools in real time. Along with these problems, this research focuses on how the blockchain technology helps in relation to those challenges.

2. Blockchain and The Benefits of Blockchain

A blockchain could be thought of as a shared (or distributed) database that is spread across multiple sites and participants. In order for new data to be added to a blockchain, they are first compiled into a "block" which is simply a collection of records to be added to the database. The block is then combined with some data (a "hash key") from the previous block through a cryptographic technique called "hashing" before it is added. Because it combines the previous block's hash key, each new block is tied to all its predecessors in the form of a chain – hence the term "blockchain".

<Figure 1. Sketch of blockchain>



Source: Quora Homepage (https://www.quora.com)

Just refer to <Figure 1>, blockchain is mainly a chain or linked list kind of data structure where each block contains a set of approved transactions and every block except genesis block is linked to its previous block using the 'Hash pointer' of previous block (The first block in a blockchain is called 'Genesis' block and it will not contain in hash pointer as there is no previous block to genesis block). The

data meaning the transactions in the blockchain are not hashed but the complete block is hashed, and the hash value is stored in next block to point to previous block. So, block no 67 contains Hash of Block 66, Block 66 contains hash of block 65 and so on.

The hash pointer mechanism helps to find out, if the data is tampered in any blocks in the blockchain and also if all the peers in the blockchain contain the same chain. This makes the blockchain immutable. Before data can be added to a blockchain, its users need to agree, or reach "consensus". This is achieved through a "consensus algorithm". A well-known consensus algorithm is the Proof-of-Work (PoW) algorithm[2]. PoW is used in the Bitcoin and Ethereum blockchain network protocols. In the PoW algorithm, users (also known as "miners") compete on computational tasks in order to reach consensus. The winning miner of each block's task would usually be given a reward.

Blockchains can be classified into three types, depending on which participants are allowed in the consensus algorithm (Zheng, Z, Xie, S, Dai, H, Chen, X, and Wang, H., 2017).

- **Public**: Anyone can participate in the consensus algorithm.
- **Consortium**: A select (or permissioned) group of entities can participate in the consensus algorithm.
- **Private**: Only a single entity operates the consensus algorithm and controls adding of new data.

Some of the most important features of blockchain are as follows:

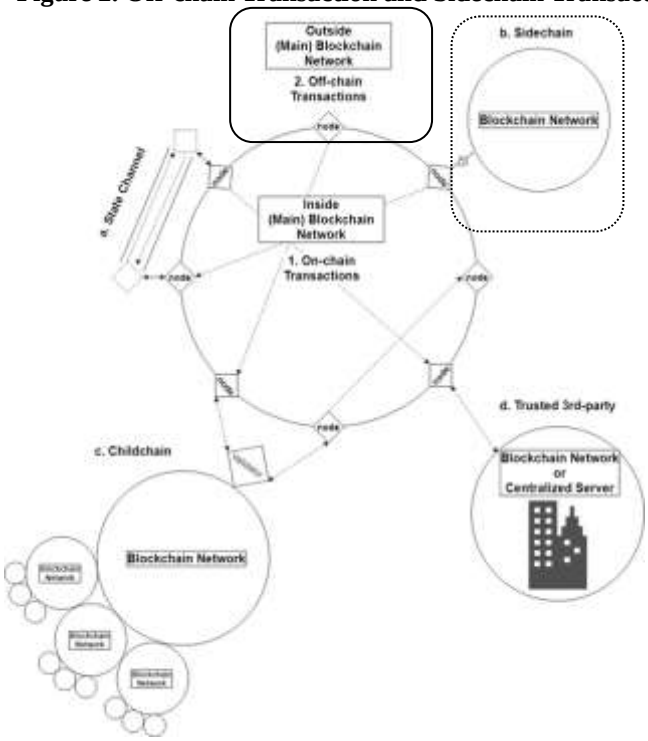
- **Immutability**: Blockchain is almost impossible to corrupt due to a permanent and unalterable network. When any transaction is initiated, nodes check its validity and authenticate to add to the ledger.
- **Decentralization**: The network is not governed by a single authority but a group of nodes that are responsible for maintaining the network. This decentralized approach allows participants to access the blockchain from the web and store their replicated information using private keys.
- **Security**: Blockchain with its decentralized and immutable natures can provide high degrees of security. Each information is hashed which hides its actual nature and also provides a unique identification for each data.
- **Consensus**: The operation of the blockchain frameworks relies on associated consensus algorithms, which is responsible for deciding the group of active nodes on the network. This makes the validation process for a transaction faster and like a voting system.

Finally, it would be explained that what are the concepts of "*scalability*" and "*smart contracts*" which will be relevant to subsequent parts of this paper. **Scalability** refers to the capacity of the blockchain to store and process transactions. It generally relates to the size and frequency of the transactions a

[2] Proof of work (PoW) is a decentralized consensus mechanism that requires members of a network to expend effort solving an arbitrary mathematical puzzle to prevent anybody from gaming the system. PoW is used widely in cryptocurrency mining, for validating transactions and mining new tokens. Due to PoW, Bitcoin and other cryptocurrency transactions can be processed peer-to-peer in a secure manner without the need for a trusted third party.

blockchain can handle. Various solutions have been proposed to improve scalability. One such solution is to store data "off-chain[3]" (instead of on-chain[4]; *indicated by a solid line*), and another is to use "side-chain[5]" (linked to the main chain; *marked with a dotted line*) to enable larger transaction volumes to be processed in parallel.

<Figure 2. Off-chain Transaction and Sidechain Transaction>



Source: Steemit Hompate (https://steemit.com/kr-dev/@modolee/onchain-offchain)

As mentioned earlier, since medical information data is often large, improving scalability in the block chain is a very important vertex. Given that healthcare data is estimated to reach as much as 2,314 exabytes generated yearly by 2020, it becomes crucial for almost all blockchain-based healthcare applications to achieve a certain level of scalability (Banks MA., 2020). **Smart contracts** are programmable computer rules. Blockchain being a digital database allows for the implementation of smart contracts, which can be automatically triggered to execute when predefined conditions are satisfied.

[3] It is simply a transaction that occurs outside of the main blockchain. Example: From the point of view of the Ethereum network, transactions sent and received on the Bitcoin network are off-chain transactions. Conversely, from the point of view of the Bitcoin network, a transaction on the Ethereum network will be an off-chain transaction. After a transaction occurs, it takes a long time for the transaction to be propagated to the blockchain network and confirmed. Therefore, services that require fast processing cannot be processed on-chain but must be processed off-chain.

[4] A transaction that takes place on the chain. By the way, the chain here means the main (single) blockchain network. Example: Transactions that occur within the block chain constituting its own network such as Bitcoin and Ethereum and are recorded in blocks.

[5] By constructing a blockchain network that adopts a high-TPS (Transaction per second) consensus method (eg DPos), off-chain transactions are performed quickly, and the final result value is reflected on the main chain. Example: Loom Networks.

3. Benefits of Using Blockchain Technology on Big Data

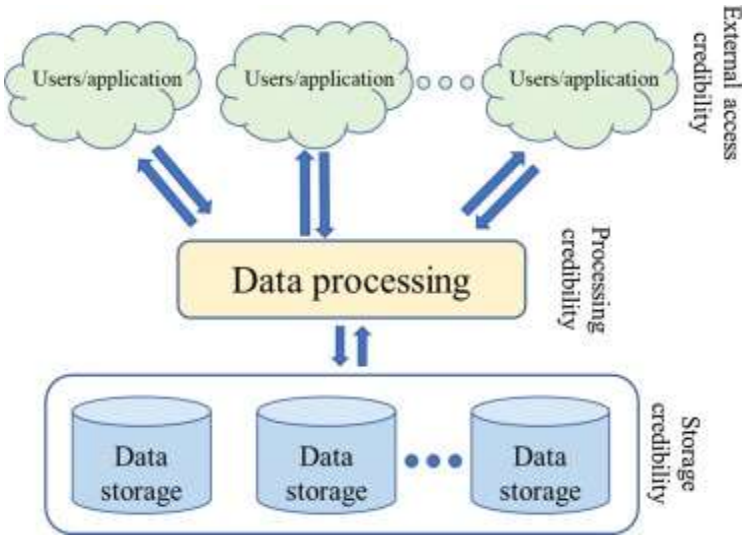
Using blockchain adds another data layer to the Big Data analytics process. Most importantly, this data layer complies with 2 main demands of the Big Data analysis. First, blockchain-generated Big Data is secure, as it cannot be forged due to the network architecture. Second, blockchain-based Big Data is valuable, meaning it is structured, abundant and complete, making it a perfect source for further analysis. Big Data is an incredibly profitable business, with revenues expected to grow to $203 billion by 2020. To say even more, by 2030 the data contained in the blockchain ledger might be worth up to 20% of the global Big Data market and generate up to $100 billion in annual income (Gil Press, 2017). Among the market creation effects of the use of Big Data, medical record is one of the most important factors due to the desire to extend the lifespan of people. The real question nowadays is who will be the first to provide the most suitable and best trained AI/machine learning model operating on top of distributed, transparent, and immutable blockchain-generated data layers. The business to do this will roll in investments and generate immense profits.

The motivations of integrating blockchain with big data are discussed as follows:

- **Improving Big Data Security and Privacy**: The traditional security solutions like firewalls cannot address this issue of Big Data since the organizations have no control over the data as it is not stored within the network perimeter of the organizations. The usage of blockchain to store the Big Data has the potential to address this issue. The encrypted and decentralized storage of the data in the blockchain network makes it very difficult for any unauthorized access to the data. Because medical information includes particularly sensitive information, this characteristic of blockchain is a very important strength.
- **Improving Data Integrity**: The immutability property of the blockchain ensures that it is next to impossible to tamper with the data stored in the blockchain network. If someone wants to modify the data in the blockchain network, they have to modify the data in at least 50% of the nodes in the blockchain network, which is nearly impossible in practice. Since medical information is relevant to the patient's life, there should be no fraudulent correction. In addition, data integrity is very important in order to develop new treatments or new drugs by processing medical Big Data.
- **Real-Time Data Analytics**: Since the blockchain stores every transaction, it makes the real-time analytics of Big Data achievable. The hospital and medical institutes can make the cross-border information sharing including large data like MRI or PET-CT in real-time as the blockchain integrated Big Data analytics enables the medical institutes to get the information quickly. Also, healthcare professionals can monitor the changes in patients' data in real time, thus enabling them to make decisions in real time.

- **Enhancement of the Quality of Big Data**: Data scientists spend most of their time on data integration as different sources follow different formats in data collection. By using blockchain for data storage, the quality of the data can be improved as it is structured and complete. Hence, data scientists can work on the quality data to come up more accurate predictions in real time.
- **Streamlining the Data Access**: The use of blockchain would simplify the life cycle of Big Data analytics by online streamlining the data access. Indeed, by involving multiple departments in an organization in a common blockchain, authorized users can get access to the secure, trusted data without having to go through several checks.

<Figure 3. The architecture of trusted database management system>



Source: Guo, L., Xie, H., Li, Y. (2020), Data Encryption based Blockchain and Privacy Preserving Mechanisms towards Big Data, Journal of Visual Communication and Image Representation, vol. 70

In data management, the essence of blockchain technology is a data block system which is built on peer-to-peer network and provides trusted data management function. A reliable database management system consists of three key components: the credibility of storage, processing, and external access, as shown in <Figure 3>.

III. Case study

This section presents the review works related to the conventional smart healthcare architecture and recent EHRs management based on blockchain. In recent blockchain-based EHR management, Ethereum platform and Hyperledger Fabric-based platform are the mainstays. According to a systemic review of Fang HSA, Tan TH, Tan YFC and Tan CJM (2020), Ethereum-based blockchain was the commonly used (n=26) with Hyperledger Fabric (HF) being the next most common (n=20). Based on this systemic review (total case was 58 cases), the research on two platforms would be summarized and compared to gain pros and cons within each platform. As a result, it would be a good background for the proposed framework, which is a key part of this proposal.

1. Case Study on Blockchain-based EHR

Blockchain is an advanced computer technology that allows users or nodes on a network to instantly analyze information and share transactions with other nodes. Several studies have used blockchain-based platforms to correct deficiencies in the current EHR. Various research papers suggest new data encoding or decoding techniques, unique digital signature method, protected scheme for information transmission and keys generator method for the blockchain-based healthcare record system.

Azaria et al. (2016) have suggested a scalable healthcare storage platform (*MedRec*) that utilizes a blockchain-based method for sharing healthcare information. Some researchers have suggested an Ethereum framework to design smart contracts that provide access policy for the healthcare system (Dagher GG, Mohler J, Milojkovic M, and Marella PB., 2018). Other researchers have presented a medical information sharing method called *MedBlock*, which utilizes blockchain technology to ensure the easy extraction of EHRs by using a Practical Byzantine Fault Tolerance (pBFT)[6] consensus method as opposed to Proof of Work (Fan K, Wang S, Ren Y, Li H, and Yang Y., 2018).

In the next section, the frequently used Ethereum-based blockchain technology and Hyperledger Fabric-based blockchain technology would be described and compared for better understanding current technical situation and getting insight on advanced model. In addition, the relevant blockchain models which is used for Big Data technology in the medical field would be summarized.

[6] Practical Byzantine Fault Tolerance (pBFT) is an algorithm that optimizes aspects of Byzantine Fault Tolerance (in other words, protection against Byzantine faults) and has been implemented in several modern distributed computer systems, including some blockchain platforms. These blockchains typically use a combination of pBFT and other consensus mechanisms.

(1) Ethereum-based EHR: *Ancile*

Ethereum offers an extremely flexible platform on which to build decentralized applications using the native Solidity[7] scripting language and Ethereum Virtual Machine. Ethereum's large user base encourages developers to deploy their applications on the network, which further reinforces Ethereum as the primary home for decentralized applications. In the future, the Ethereum 2.0 protocol, currently under development, will provide a more scalable network on which to build decentralized applications that require higher transaction throughput.

The most advanced example of building an EHR sharing system using Ethereum-based blockchain technology is the *Ancile* model proposed by Dagher GG, Mohler J, Milojkovic M, and Marella PB (2018). *Ancile* uses six unique types of smart contracts for operation: Consensus, Classification, Service History, Ownership, Permissions, and Re-encryption.

Using smart contracts, *Ancile* maintains cryptographic hashes and query links, confirming the integrity of EHR Databases. Patients can also view and control who has permissions for their private information by using a smart contract to manage access control. Moreover, patients may give transfer permissions to other nodes. *Ancile* consists of three main software components: Database Manager, Cipher Manager and Ethereum-Go Client.

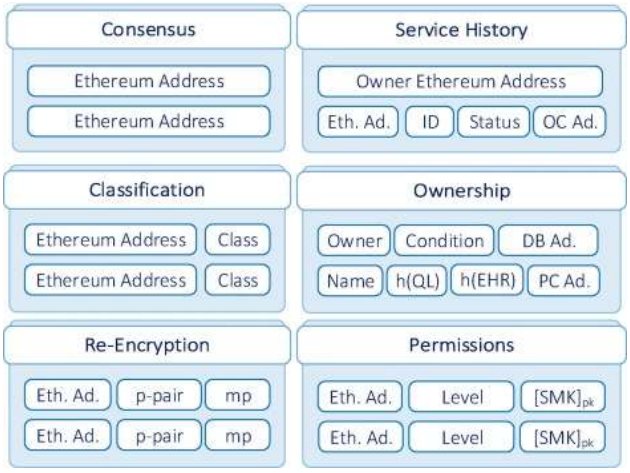
- **Database Manager** is used to navigate existing EHR Databases and for generating the link that maps to a record. Moreover, it will also create hashes of both the record and the query link to place on the blockchain.
- **Cipher Manager** is responsible for all cryptography in *Ancile*. Decryption of all files, three different forms of encryption: symmetric key, public key, and proxy re-encryption.
- **Ethereum-Go Client** (Go Ethereum, 2017), sometimes called Geth, is the main Ethereum command line interface in the Go (programming language)[8]. Geth is an access point to Ethereum networks. *Ancile* is designed to function on a permissioned Ethereum blockchain; thus, the Geth client would be used by permitted nodes to access the private blockchain.

Ancile consists of six smart contracts in <Figure 4>: Consensus Contract, Classification Contract, Service History Contract, Ownership Contract, Permissions Contract and Re-encryption Contract.

[7] Solidity is an object-oriented programming language for implementing smart contracts on various blockchain platforms, most notably, Ethereum.

[8] Go, also known as Golang, is an open source, compiled, and statically typed programming language designed by Google. It is built to be simple, high-performing, readable, and efficient.

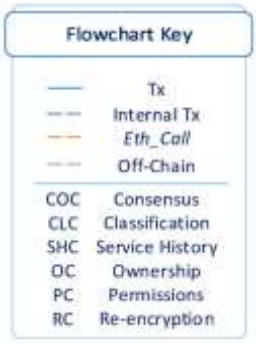
<Figure 4. Six Smart Contracts of Ancile>



Source: Dagher GG et al (2018), Ancile: privacy-preserving framework for access control and interoperability of electronic health records using Blockchain technology, Sustain Cities Soc., 39:283–97

The following diagrams demonstrate the architecture of *Ancile* by assessing how the framework would be used in various situations. The framework uses four distinct forms of actions, as seen in <Figure 5>.

<Figure 5. A Key for the different forms of action and abbreviation for *Ancile*>



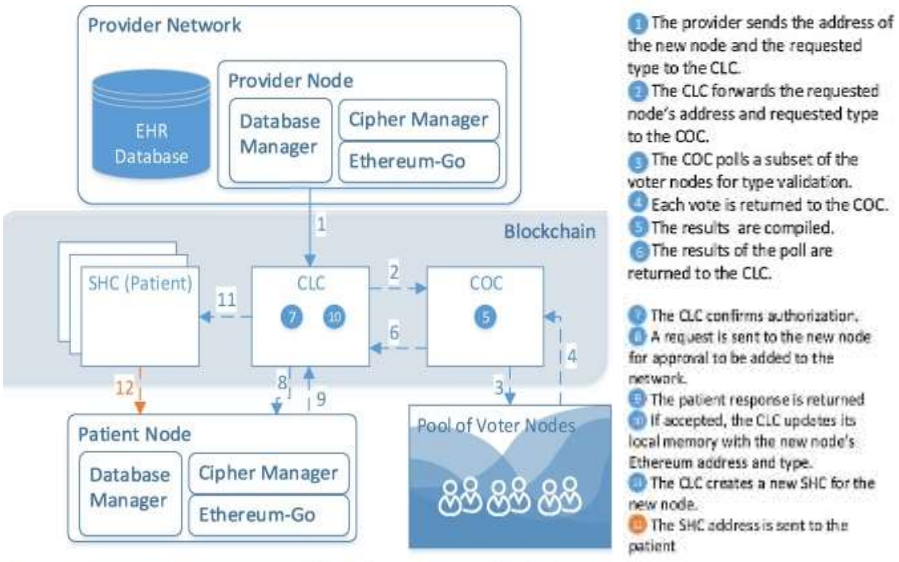
- Solid blue: a standard blockchain transaction (Tx)
- Dashed blue: internal transactions (Internal Tx)
- Dashed orange: an *eth_call*[9], which is used when data needs to be sent to a smart contract, but does not need to be written to the blockchain
- Dashed gray: a non-blockchain action (Off-Chain). This could represent data being transmitted over HTTPs or something happening internally to a node.

Source: Dagher GG et al (2018), Ancile: privacy-preserving framework for access control and interoperability of electronic health records using Blockchain technology, Sustain Cities Soc., 39:283–97

In order to explain how *Ancile* actually works, it would be summarized that how to create a node for a new patient on the blockchain. The process of adding a node begins by having voter nodes validate that the public ID suits the requested classification. The process for adding a new node can be seen in <Figure 6>. This process could be understood by referring to the types and colors of the lines shown in <Figure 5>.

[9] Executes a new message call immediately, without creating a transaction on the block chain. The *eth_call* method can be used to query internal contract state, to execute validations coded into a contract or even to test what the effect of a transaction would be without running it live.

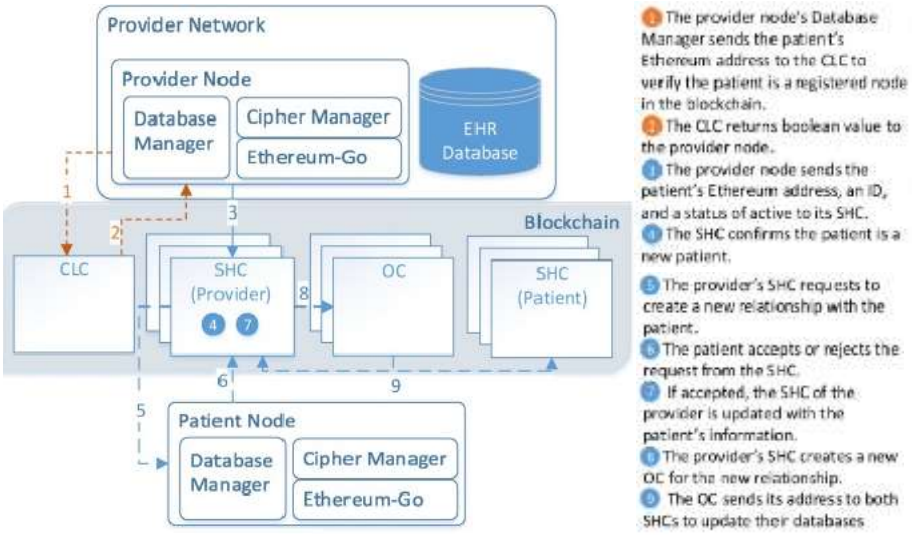
<Figure 6. The process for adding a new node>



Source: Dagher GG et al (2018), Ancile: privacy-preserving framework for access control and interoperability of electronic health records using Blockchain technology, Sustain Cities Soc., 39:283–97

Through the newly created patient node, the procedure for registering a patient on the blockchain would be completed. This process must be completed every time a new patient visits a provider. Using the SHC, *Ancile* documents each time a new relationship is formed. As seen in <Figure 7>, registration begins by confirming that the patient is a registered node in the system.

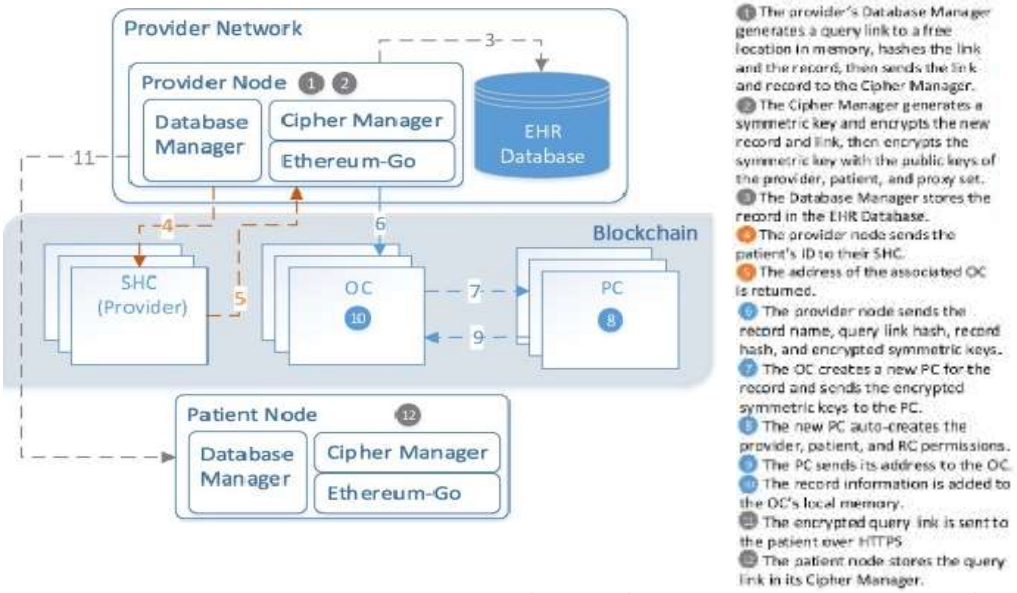
<Figure 7. The process for registering a patient>



Source: Dagher GG et al (2018), Ancile: privacy-preserving framework for access control and interoperability of electronic health records using Blockchain technology, Sustain Cities Soc., 39:283–97

From now on, it would be summarized that how to make the procedure for registering the patient's medical record presented by the authors on the blockchain. The process for adding a record begins with internal encryption in a provider node. <Figure 8> depicts the process of adding a record to an EHR Database and then using *Ancile* for data integrity and access control.

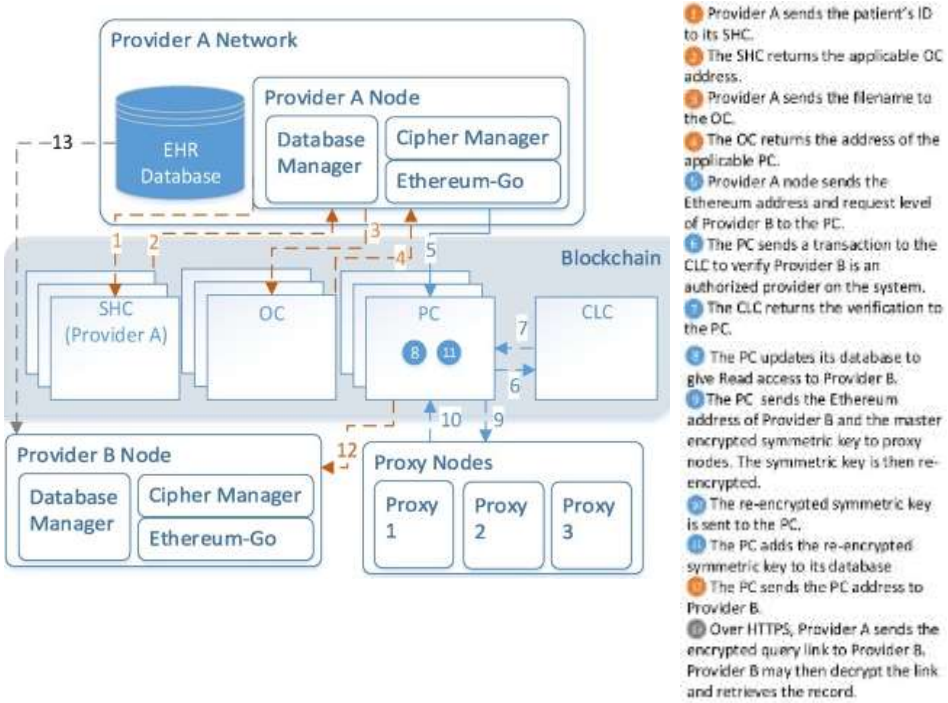
<Figure 8. The process for adding a new record>



Source: Dagher GG et al (2018), Ancile: privacy-preserving framework for access control and interoperability of electronic health records using Blockchain technology, Sustain Cities Soc., 39:283–97

Finally, *Ancile* can transmit patient's medical data to another provider (hospital or insurance company). *Ancile* uses proxy re-encryption to balance the need for accessibility while maintaining security. <Figure 9> depicts the process of one provider sending a record to another. It should be noted that the process for transferring a record could technically occur by retrieving a record, decrypting, and sending to another party. Thus, *Ancile* cannot ensure all data movement is tracked, but can verify who is permitted to share records and with whom records can be shared. In this way, *Ancile* may be used as an indisputable ledger should external actions need to be taken.

<Figure 9. The process sending a patient's record to another provider>



Source: Dagher GG et al (2018), Ancile: privacy-preserving framework for access control and interoperability of electronic health records using Blockchain technology, Sustain Cities Soc., 39:283–97

The *Ancile* (Ethereum-based EHR system) presented by the authors has a very high degree of completeness compared to previous research and presents detailed protocols and implementations. This is considered a realistically executable module. However, an accurate method for improving the scalability of the big data characteristics of medical records has not been presented. This is an area that needs further study. In addition, a specific front-end technology for implementing in a mobile environment was not applied. Realistically, in order for individual patients to actively use medical information, it must be developed on a mobile environment.

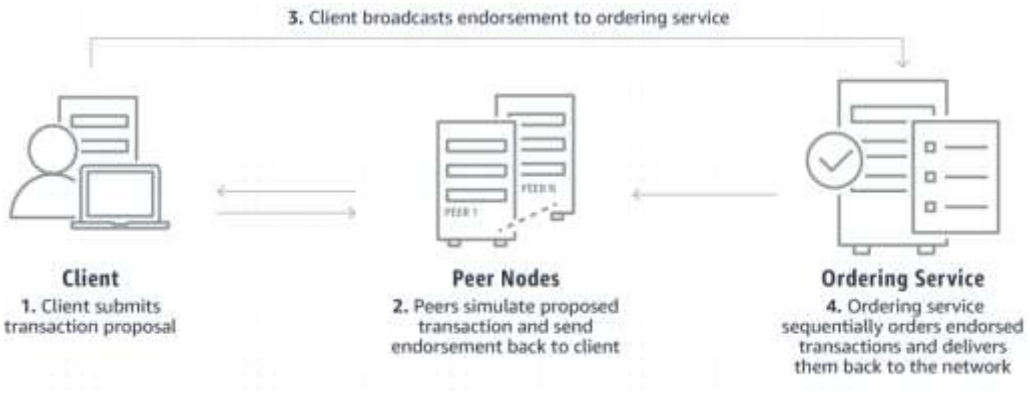
The authors say that *Ancile* can achieve better results if the background technology for this platform develops further in the future. In the context of the authors' opinion, if version 2.0 of Ethereum is released, smart contracts and permissioned in the Ethereum environment will be improved. This advancement will greatly lower the gas price and the latency problem will also be greatly overcome.

The Ethereum-based *Ancile* platform could be the good foundation of the proposed platform to be presented in Chapter IV, so it has been dealt with in detail. So, the key contribution of this proposal is to establish blockchain based EHR sharing model for addressing such scalability issue and setting a mobile environment platform. In the next section, it would be focused on the platform based on Hyperledger Fabric, which is the frequently used platform along with the platform based on Ethereum.

(2) Hyperledger Fabric (HF)-based EHR

Hyperledger Fabric is an open source, permissioned blockchain framework, started in 2015 by The Linux Foundation. It is a modular, general-purpose framework that offers unique identity management and access control features, which make it suitable for a variety of industry applications such as track-and-trace of supply chains, trade finance, loyalty, and rewards, as well as clearing and settlement of financial assets (AWS, 2021). A Hyperledger Fabric network is comprised of unique organizations (or members) that interact with each other on the network. From a Fabric component perspective, each organization has a Fabric certificate authority and one or more peer nodes. A Fabric network also has an ordering service shared by all organizations in the network, and this component helps process transactions for the network. Hyperledger Fabric Transaction flow is below:

<Figure 10. Hyperledger Fabric Transaction Flow>



Source: Amazon Web Service (AWS) Homepage (https://aws.amazon.com/ko/blockchain)

1. The transaction flow begins when a client application sends a transaction proposal to peers in each organization for endorsement.
2. The peers verify the submitting client's identity and authority to submit the transaction. Next, they simulate the outcome of the proposed transaction and if it matches what was expected, it sends an endorsement signature back to the client.
3. The client collects endorsements from peers, and once it receives the proper number of endorsements defined in the endorsement policy, it sends the transaction to the ordering service.
4. Lastly, the ordering service checks to see if the transaction has the proper number of endorsements to satisfy its policy. It then chronologically orders and packages the approved transactions into blocks and sends these blocks to peer nodes in each organization. Peer nodes receive new blocks of transactions from the ordering service, and then do a final validation for transactions in that block. Once this is complete, the new block is added to the ledger and the state of the ledger is updated. The new transactions are now committed.

Benefits of Hyperledger Fabric can be summarized into four main categories as below:

- *Open Source*: It has an active and growing community of developers.
- *Permissioned*: All participating member's identities are known and authenticated. This benefit is particularly useful in industries including healthcare and insurance where data cannot be exposed to unknown entities.
- *Governance and Access Control*: Members on the network can transact in a private and confidential way. Each transaction on the blockchain network is executed on a channel, where each party must be authenticated and authorized to transact on that channel. This provides an additional layer of access control and is especially useful when members want to limit exposure of the data.
- *Performance*: Hyperledger Fabric is a permissioned blockchain framework, it does not need to solve for Byzantine Fault Tolerance[10] which can cause slower performance when validating transactions on the network.

In addition to this basic HF blockchain knowledge, it is needed to add a description of the components of HF presented in their article to better understand *the case of Tith D. et al (2020).*

<Table 1. Components of the Hyperledger Fabric>

Ledger	- It consists of a blockchain and state database (DB). - DB holds the fact of a business object and can be created, update and delete. - DB can be generated at any time from the blockchain and maintained by Peers. - Each block contains a sequence of transactions representing a query. - Each block header includes a hash of blocks transactions, as well as a copy of a hash of the previous block's header.
User roles(nodes)	- There are three main types of user roles: client, peer, and orderer. - Clients communicate with both peers and the ordering service. - Peer commits transactions and maintains the state and a copy of the ledger. - Orderer implements a delivery guarantee, such as atomic or total order broadcast.
Chaincode	- Multiple smart contracts can be defined within a single chaincode. - Chaincode is a technical container of a group of related smart contracts for installation and instantiation. - Every chaincode has endorsement policy attached to it, which applies to every smart contract defined within it.

[10] Byzantine Fault Tolerance is a computer system's ability to continue operating even if some of its nodes fail or act maliciously. The term comes from a hypothetical called the Byzantine Generals Problem. The blockchain needs to be able to function even if it has nodes that aren't working correctly or are providing false information. However, it requires communication between nodes at every step of the process. This takes time, which can be a problem from a scalability standpoint.

Membership Service Provider (MSP)	- It is a component that defines the rules in which, identities are validated, authenticated, and allowed access to a network. - There are two types of MSPs: Local MSP and Channel MSP. - Local MSP defines users (Clients) and nodes (peers, orderers) and defines who has administrative or participatory rights at that level. - Channel MSP defines administrative and participatory rights at the channel level.

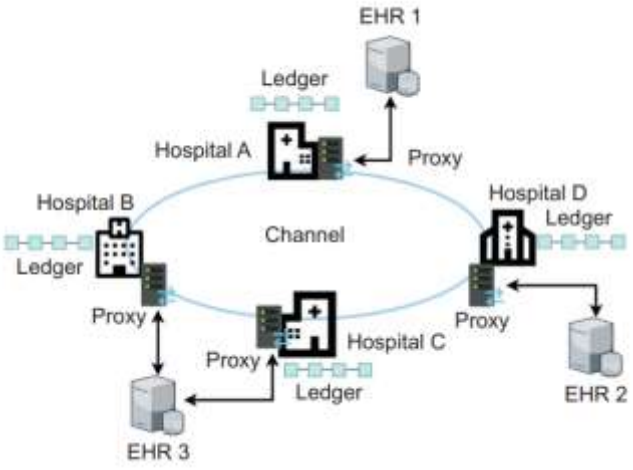
Tith D et al, Application of blockchain to maintaining patient records in electronic health record for enhanced privacy, scalability, and availability, Healthcare Informatics, 2020 Jan;26(1):3-12.

The authors suggest that their HLF based EHR system has three major features:

- A trusted directory of patient data in EHRs which guarantees access as well as the integrity of the data itself
- Strengthened security in dealing with patient data by utilizing a particular encryption scheme and providing a transparent and undeniable audit trail based on an immutable access log
- Providing scalability to cover multiple existing EHRs of regional or core hospitals with the least modification and availability of the system without relying on a centralized supervisory system

HLF provides a variety of special designated chaincodes called 'system chaincodes' to perform certain privileged tasks. Examples of 'system chaincodes' are Configuration, Life Cycle, Query, Endorser, and Validator 'system chaincodes'. The authors designed several prerequisite chaincodes and implemented them in their prototype system. They built a private subnet of an HLF network where the same ledger is shared among the hospital members shown as in <Figure 11>.

<Figure 11. Channel of network among medical institutions with the same ledger >



Source: Tith D. et al (2020), Application of blockchain to maintaining patient records in electronic health record for enhanced privacy, scalability, and availability, Healthcare Informatics, 2020 Jan;26(1):3-12

The authors who designed this model tried to implement the platform through the following conditions. In order to use the characteristics of the closed blockchain platform, government-run national public hospitals or large hospitals with franchise hospitals are targeted, and each hospital is premised on having the same ledger.

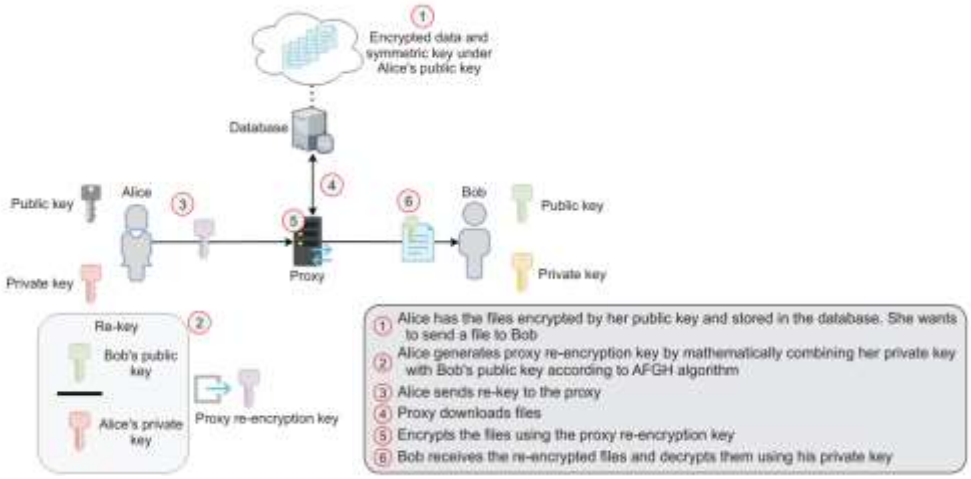
Organizations or departments within them can constitute independent channels with relevant ledgers according to their needs in the system. In practice, medical data is usually too big to handle directly in a ledger; therefore, they use on-chain or off-chain strategy to run the system efficiently. Data is kept in an EHR (EHR1, 2, 3 in <Figure 11>), and only the address is recorded in the ledger. A ledger also contains the hash values of data. This guarantees data integrity because once a piece of data is written in a ledger, it becomes immutable, and this allows the user to check whether the data has been altered or not.

The ledger consists of patient metadata, including demographics, and these data are used for retrieval requests to find transactions related to a specific patient during a specified period of timestamps of blocks in the ledger. Thus, the ledger functions as a registry of patient IDs for doctors to search for their patient's records stored in other EHRs.

In addition, each transaction contains the client's request metadata, chaincode execution results, and medical record metadata, such as hospital ID, hash of medical records stored in an EHR, and so forth. In consequence, these data will be used for auditing purpose. For an individual patient, the enrolment ID (eID) issued by a membership service provider (MSP) is used as the channel patient ID in the system.

In their suggested system, to read patient data, a proxy downloads it from the relevant EHR and sends it to the receiver. However, in case the receiver is different from the patient, the encrypted symmetric key at the data should be transformed, so that it can be decrypted by the receiver's private key. To do this, user use a proxy re-encryption scheme shown as in <Figure 12>.

<Figure 12. Proxy re-encryption scheme>



Source: Tith D. et al (2020), Application of blockchain to maintaining patient records in electronic health record for enhanced privacy, scalability, and availability, Healthcare Informatics, 2020 Jan;26(1):3-12

The authors defined their framework as a web-based application. In addition, it was also said that this web-based application improved the user's convenience. But, if it is not a mobile environment platform, it cannot be called a convenient EHR sharing system anymore. When this research was designed, mobile-based technologies had already become a core foundation in many fields and commercialized but limiting the system to web-based applications will reduce the utilization of it.

The authors tell their prototype: "The prototype is not the same as an actual working environment. The system and 'chaincode' functionality may require specific modification to suit consortium privacy policies and the legal requirements set by the governing authority". And in the conclusion, they tell the excellence of their research by comparing it with other medical information-sharing system projects. For comparison, there are three projects in their comparison: '*Medrec* (Azaria A. et al, 2016), *Ancile* (Dagher GG. Et al, 2018), and *Dubovitskaya A, Xu Z, Ryu S, Schumacher M, and Wang F (2018)*. The first two projects are Ethereum-based EHR systems, and the third project is the same HLF-based system.

The authors insist that their prototypes are superior in feasibility compared to each system, especially emphasizing that their system can be implemented in cloud systems. However, they did not provide a detailed description of how the system can be implemented in a cloud environment. This proposed study is significant in that it provided a complete cycle of implementation for HLF-based EHR.

However, detailed studies on the implementation of the system in the mobile environment for users' accessibility and on the cloud computing to improve scalability have not been considered and designed. This gap will be filled by the advanced EHR system proposed in the next chapter (IV).

<Table 2. Comparison of the Ethereum vs. Hyperledger Fabric>

	Ethereum	Hyperledger Fabric
Public vs. Private	Public	Private
Permissions	Permissionless	Permissioned
Governance	Decentralized	Federated
Private Transactions	No	Yes

Before entering the full-scale case study, <Table 2> can explain why the Ethereum-based platform is suitable for the EHR sharing system. This table summarizes the characteristics of the Ethereum platform and the HF platform. As seen above, an Ethereum-based EHR system is more suitable because of strong characteristics of permissionless and decentralized process. In contrast, the HF-based platform is more useful as a method for sharing information between companies or for communication between specific interest groups.

2. Integration of Blockchain and Big Data in Medical field

Most of the previous studies might be in the early stages in relation to mobile environment platform for user convenience and accessibility. In addition, the blockchain technology for processing medical Big Data, which is highlighted in this proposal, has not yet been discussed in earnest.

This chapter would be concluded with the review of the latest research using blockchain for medical Big Data. For this objective, three systemic reviews (Sen H, Fang A, Hwee T, Tan Y, and Tan M, 2020; E. Karafiloski and A. Mishev, 2017; Mazlan A. A, Daud S. M., Sam S. M, 2020) were selected and studied. All of these reviews are effective for obtaining the latest research flow. Summarization on those systemic reviews are as follows:

- Most of the research in this field is conducted by computer scientist, not by medical professionals and China and India are leading the way in this area (Sen H. et al, 2020)
- The development of Ethereum-based and HLF-based EHR system has become more gravitated and research to solve data storage has been also active (Sen H. et al, 2020).
- One of the major areas regarding blockchain EHR that is still undergoing much research is scalability (Sen H. et al, 2020; Mazlan A. A. et al, 2020; E. Karafiloski et al, 2017). The targets to improve scalability are block size, high volume of transactions, number of nodes and protocol, etc (Mazlan A. A. et al, 2020).

From the above findings, it was clear that there are not many mobile-based studies yet. However, in recent two years, due to the spread of the coronavirus, a lot of mobile-based research has been conducted (Santos J. A, Pedro R. M, Silva B., 2021; Kassab M and Destefanis G, 2021; Bittins S, Kober G, Margheri

A, Masi M, 2018; Pratima P, Jindal R, and Borah M, 2021). Based on this summary, in the next chapter (IV), an Ethereum-based mobile EHR sharing system would be proposed a for medical Big Data management.

Since the application of blockchain technology for such medical Big Data is still in the early stages of research, there are few commercialized cases as a business profit model. Even if such an attempt is being made through some venture companies, it is still in the research stage, and technical difficulties need to be resolved for commercialization.

In this initial stage, it is true that it is very difficult to find a commercialized model for case study for this proposal. Therefore, this proposal focused on supplementing the problems in the previously designed blockchain-based EHR sharing system and mentioning the elements necessary for commercialization, rather than investigating the commercialization model.

IV. Proposed framework

In this proposal, except for the technical part, it would be given a concept for the blockchain-based EHR sharing system on the background obtained through previous case studies. As mentioned before, this proposed EHR sharing system is based on Ethereum network and will utilize on-site and off-site to improve data scalability. Also, it will be designed based on IPFS for its utilization. Lastly, in order to promote user convenience in line with the current trend, it will be created on the mobile base. In the next section, terms and background knowledge used for this design are briefly summarized.

1. Preliminaries of the Proposed Platform

In this proposal, an Ethereum blockchain platform would be employed for building our EHR sharing system. A big advantage of Ethereum is its adaptable and flexible features, which allow to build any blockchain applications such as e-medical records.

The HLF-based EHR system reviewed in the previous section has a problem of data compatibility with all medical institutions or government or third party like as insurance company due to its closed system and may have security vulnerabilities because it is cloud-based. If these points are considered, the Ethereum-based EHR system is superior in information access to all stakeholders through its openness feature, and also through compatibility with IPFS (Inter-Planetary File System)[11]-based data storing system which would be a great feature of this proposal. The security vulnerabilities of cloud systems can be overcome through convergence with IPFS technique.

Main components of an Ethereum network which would be used in this study are as follow:

- **Ethereum account**: Ethereum has two different account, externally owned accounts (EOAs) and contract accounts. Every account is indexed by an address and defined by a pair of keys, a private key and a public key. To interact with Ethereum blockchain, each user needs to create an account to become an entity or node in the network.

- **Smart contract**: A smart contract is a kind of self-operating computer program, which can be executed automatically when specific conditions are met. Functions defined in smart contracts can be triggered by a new transaction sent from an account. This property allows entities to implement their job functionalities such as data transmission, request handling or access management.

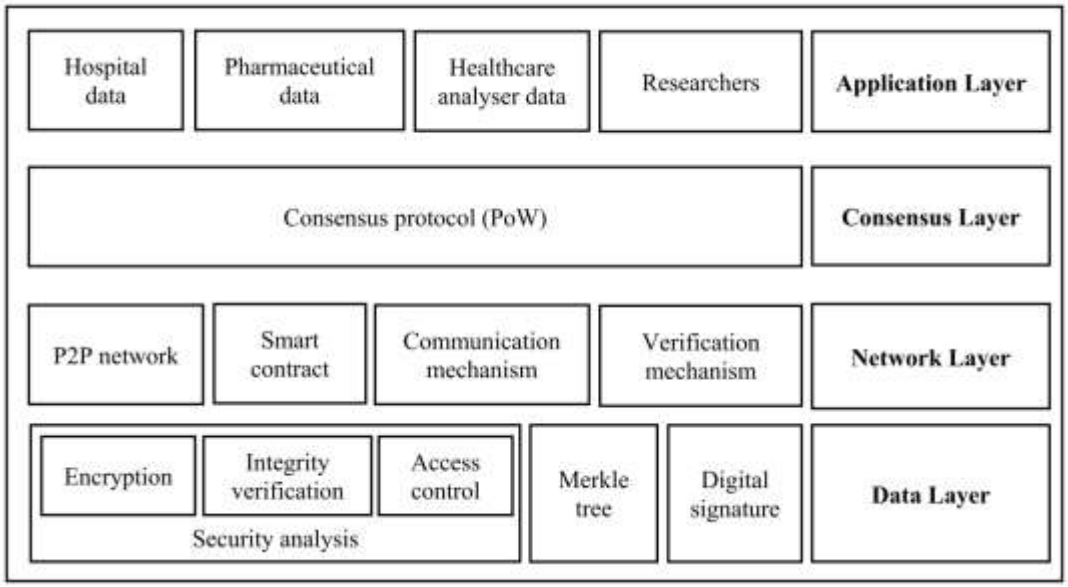
[11] The InterPlanetary File System (IPFS) is a protocol and peer-to-peer network for storing and sharing data in a distributed file system. IPFS uses content-addressing to uniquely identify each file in a global namespace connecting all computing devices.

- **Transaction**: An Ethereum transaction is a data packet to transfer *ether* (Ethereum native token) from an account to another.

2. System Architecture and Design Goals

In this section, the entire proposed scheme has been discussed in detail. The proposed model is divided into four layers: (1) data layer (2) network layer (3) consensus layer, and (4) application layer. The layered structure is shown in <Figure 13>.

<Figure 13. Layered structure of the proposed scheme>



- **Data Layer**: The data layer is the bottom layer, which manages the data before moving onto the network layer. The Ethereum network generates public-private keys for the users, and a digital signature scheme ensures the authenticity of the users. All the transactions and data flow are compiled in the form of hash chains or chains of transactions. Here, access control is managed using the credentials of the users, and only authorized users can upload healthcare Big Data. Users can request to check the integrity of a document, and eventually, security is ensured. The healthcare data is encoded using the Advanced Encryption Standard (AES) algorithm[12] before it is stored on IPFS.

[12] AES has been adopted by the U.S. government. It supersedes the Data Encryption Standard (DES), which was published in 1977. The algorithm described by AES is a symmetric-key algorithm, meaning the same key is used for both encrypting and decrypting the data.

- **Network Layer**: The network layer is responsible for establishing a connection between the data layer and the application layer. It contains a P2P network[13] for communication, including smart contracts, which work as the backend of the proposed scheme. There are various users in the proposed scheme, such as hospitals' admin and doctors, who are authorized to work on the network. The hospital's admin uploads the data, and only authorized members of the hospital can view it. Doctors and other staff of the hospital can also upload the data. Admin can share the medical Big Data from one hospital with other hospitals, so that other hospitals' staffs can view the data, if required. This is how data is shared over the network. The smart contract acts as a controller to provide services to various users and manages unauthorized users.

- **Consensus Layer**: As in a blockchain network, each node may receive necessary data at different times, a consensus mechanism[14] is required to determine which node should add the new block in the blockchain. The proposed scheme is built on top of the Ethereum Blockchain network, which uses the PoW (Proof of Work)[15] method to manage all the activity. Ethereum takes around 10–15 seconds to confirm a transaction, which may vary according to the user's spending, i.e., the gas value.

- **Application Layer**: The application layer is the topmost layer that provides services to the end users and doctors. This layer helps all the users to communicate with the blockchain network. Admins can register the users. The authorized users can upload medical data in excel format and can view the uploaded data accordingly. Other network's peers can demand a file or data from any node by using a Distributed Hash Table (DHT)[16].

3. Workflow

There are few main entities in the proposed scheme: (1) users (hospitals, doctors, staff and researchers) (2) admin (3) smart contract, and (3) IPFS storage. All these entities are connected to form a P2P blockchain network. The workflow of the proposed scheme is shown in <Figure 14>.

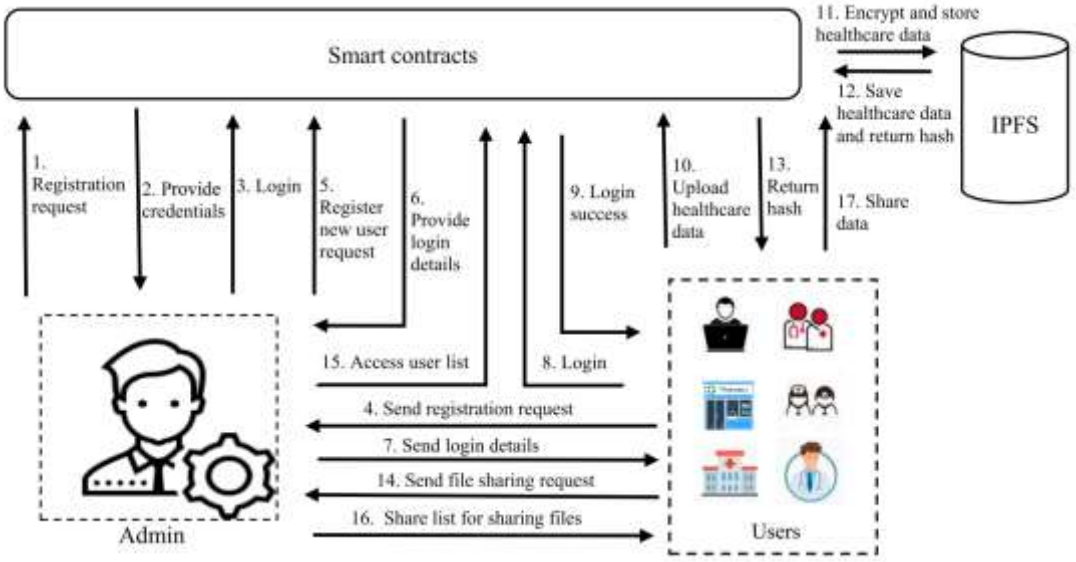
[13] Peer-to-peer (P2P) computing or networking is a distributed application architecture that partitions tasks or workloads between peers. Peers are equally privileged, equipotent participants in the application. They are said to form a peer-to-peer network of nodes. A peer-to-peer (P2P) network in which interconnected nodes ("peers") share resources amongst each other without the use of a centralized administrative system.

[14] Proof of stake (PoS) or Proof of work (PoW) protocols are a class of consensus mechanisms for blockchains that work by selecting validators in proportion to their quantity of holdings in the associated cryptocurrency.

[15] Proof of work (PoW) is a form of cryptographic proof in which one party (the prover) proves to others (the verifiers) that a certain amount of a specific computational effort has been expended.

[16] A distributed hash table (DHT) is a distributed system that provides a lookup service similar to a hash table. The main advantage of a DHT is that nodes can be added or removed with minimum work around re-distributing keys.

<Figure 14. Architecture of the proposed scheme>



Here, at first, (1) admins send key generation requests to generate global parameters and public and private keys to register in the proposed architecture. Then, (2) the blockchain network executes the function of the registration smart contract, generates credentials, and the credentials are shared with the admin. In the third step, (3) the admin utilizes the shared credentials to login into the proposed architecture and accesses the services. Next, (4) the user sends a registration request to the admin with some details like name, email ID, hospital, etc. In response, (5) the admin saves the user's details in the proposed architecture and executes the corresponding user's registration function to generate the credentials. After executing the function of the user's registration smart contract, (6) the admin receives the login details of the user. (7) The admin is responsible for sharing the login credentials with the user. In the next step, (8) the user can login into the proposed application using shared details and access the services.

After the successful login, (9) the user can access the services of the proposed application to upload, download and share the healthcare data. To upload healthcare data, (10) the user sends the healthcare data file to the blockchain network using suitable options available on the application portal. In the eleventh step, (11) the Blockchain network executes the corresponding smart contract function to encrypt the healthcare data using the AES algorithm and stores it in a distributed manner by using IPFS storage.

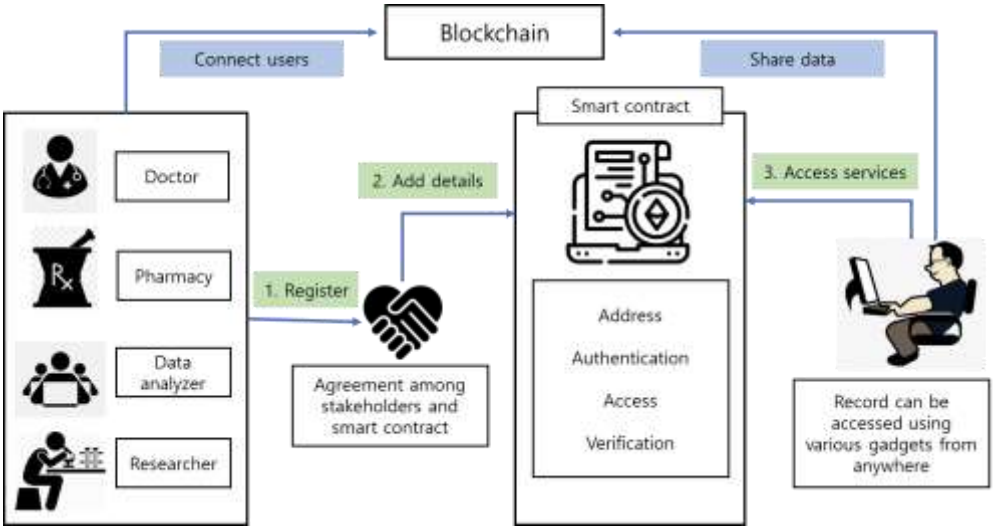
On successfully storing healthcare data on the IPFS, (12) the unique file hash is returned by the smart contract's function. Then, (13) the generated unique file hash is shared with the user. The user can utilize the file hash for the data accessing process. In the next step, (14) the user is allowed to share the uploaded healthcare files with the other users. The user can obtain the registered users' details by sending a request

to the admin. In response, (15) the admin obtains the registered user list from the Blockchain network. Then, (16) the admin sends the registered user list to the user for sharing the healthcare data. Finally, (17) the user can utilize the service of the portal to share data with the respective user.

4. Smart Contract of Proposed Model

Now, it would be explained in more detail about the smart contract modeling, which is the most important part of this workflow. Smart contracts are a package of codes that encode and replicate real world contractual agreements in the computer domain. The basic principle of contracts is to create a legal agreement between two or more parties. Each party's roles and activities are as shown in <Figure 15>.

<Figure 15. Smart contract of proposed model>



In the smart contract module, all the backend functionalities of the proposed scheme are implemented. The smart contract manages record, and transfers healthcare record among different users.

Smart contracts are a core part of the proposed application. These smart contracts are designed for registering users, authentication, data storage and integrity check functions. The blockchain publishes the smart contract's functions and provides efficient, reliable and security features. All communications are occurred using signed messages in the blockchain network. The users of the blockchain network receive the notification of deployment of the functions of the smart contracts. The functions of the designed smart contracts are programmable assets that are executed as decided in the code, and these functions are mainly dependent on the user's responsibility or role. For admin, the smart contract provides the function to register the user. For other users, smart contract manages upload, access, share and check integrity functions. The users of the application execute these functions to utilize its services.

The proposed Ethereum framework generates the public and private keys for the users and also specify the procedure for signing and verifying the users' signature. Here, the first step is to generate the hash of the message or file that needs to be signed. In the next step, Ethereum JSON RPC[17] is utilized to sign the message hash from the generated Ethereum address and the signature is generated in the "sign" variable. The last step, i.e., verification step returns the Ethereum address that is used to sign the message. The changes in the message hash or signature depict a different address than the original address (In this part, designing the algorithms using the actual programming language is a specialized area, so it would be not realized and would be reinforced through future research). The smart contract defines different functions for the registration process, sharing data, file uploading and integrity checking.

At first (1), in this smart contract, the registration process of hospitals and doctors are involved. It checks the hospital's details, i.e., if the entered details are already presented in the contract, it discards the request and returns false. Then, it checks the admin's validity, i.e., whether the caller of the function is a valid admin or not. It also takes the admin address, hospital's Identity (ID) and doctor's credentials to the registration process of doctors or staff or researchers of a particular hospital. After verification, it checks whether the entered doctor's details are already presented in the contract or not. If entered doctor's details are already added, it discards the request and returns false. Otherwise, the entered doctor's details are added in the agreement and mapped to the added doctor's unique address. After successful registration, doctors are allowed to access the portal of the proposed application and can perform various operations. Doctors are also entitled to access the data shared by other registered hospitals.

Secondly (2), the healthcare file uploading process are involved in this smart contract. To begin with, it checks whether the function is called by an authorized doctor/admin or not. It takes the doctor's address and healthcare excels file as the input parameters. It uses a predefined decrypt hash function to generate a unique hash for the record. After this hash calculation, the data is uploaded on the IPFS storage network, and the returned hash is stored on the smart contract.

Lastly (3), it would be involved in the steps for sharing stored healthcare data among different users of the proposed application. It takes the hospital's ID as an input parameter. This smart contract checks the user's identity, and then, allows the sharing process. In the similar way, multiple functions can be used for sharing the data among different users.

[17] JSON-RPC is a stateless, light-weight remote procedure call (RPC) protocol. Primarily this specification defines several data structures and the rules around their processing. It is transport agnostic in that the concepts can be used within the same process, over sockets, over HTTP, or in many various message passing environments.

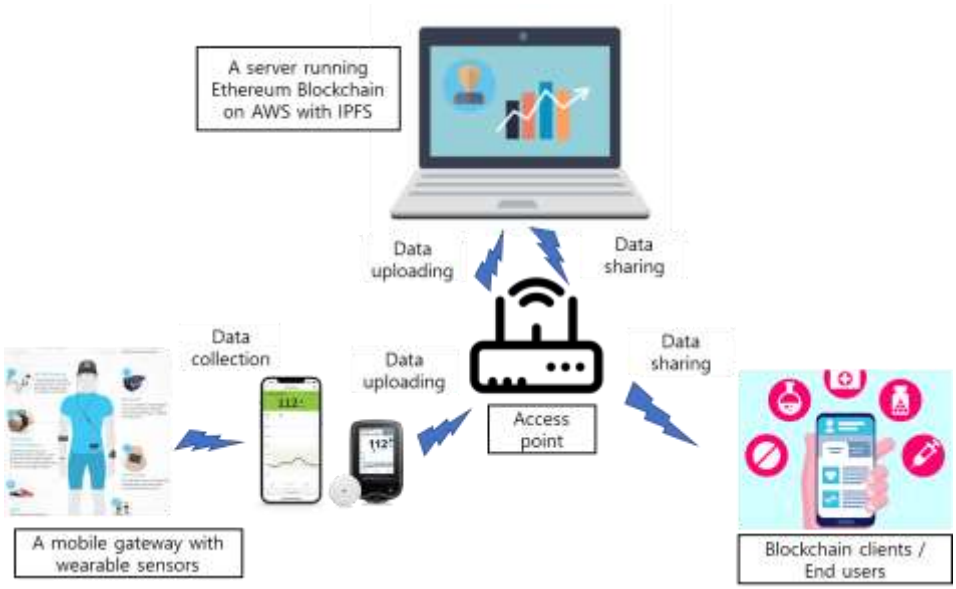
5. System Setting on the Mobile Platform

As a final step to design the proposed scheme, it would be discussed how to implement this platform on mobile. In this platform, users can interact with smart contracts through their Android phone where a Geth client (a command line interface implemented in the Go language) was installed to transform each smartphone into an Ethereum node.

By using the Geth client, a mobile user can create an Ethereum account to communicate with our blockchain network for accessing data. The web3.js library, a lightweight Java library for working with smart contracts and blockchain, would be also used for developing the mobile application to connect with the Ethereum blockchain network.

Next, IPFS would be set up on Amazon cloud to build a decentralized storage system. Currently, IPFS can be built on top of the cloud computing on AWS. A delicate configuration for the IPFS system on AWS will be suggested in the follow up study. In order to store and utilize patient's clinical data in a real mobile environment, proposed mobile healthcare platform would include a network of wearable sensors or mobile self-blood glucose monitoring system and a mobile Android application to collect and process patient data for cloud storage. it is assumed that medical records were collected by sensors and stored in the IPFS system. Note that all data files were encrypted by the public key of EHRs manager in data uploading and then decrypted by EHRs manager with its private key in the data sharing process in an asymmetric manner as explained before. Schematic EHRs sharing framework on mobile cloud as shown in <Figure 16>.

<Figure 16. Scheme of mobile EHRs sharing>



V. Discussion & conclusion

1. Beauty of Proposed Mobile EHR Sharing System

This proposal discusses the smart contract architecture on Ethereum platform for the healthcare information sharing system to provide security and privacy of broad-scale medical big data by mobile cloud computing and blockchain. This study identifies critical challenges of current EHRs sharing systems and propose efficient solutions to address these issues through a real prototype implementation.

In this paper, it was focused that how blockchain technology could be leveraged to improve and enrich the existing EHRs systems in healthcare. This proposal gave a perspicacity on a Ethereum-based architecture for the secure and efficient EHRs systems. In addition, the basis for a feasible platform for the characteristics of medical Big Data has been provided. It enables us to create a peer-to-peer blockchain network of various identified and registered healthcare stakeholders to achieve maximum interoperability, security, privacy, scalability. The other aspects of the paper lie in the illustration and discussion of functional transaction activities and events performed between various stakeholders such as patient, healthcare professional and hospital.

As a result, the proposed novel EHR sharing has some advanced point compared to the previous research in the following areas:

- A blockchain platform for researching medical big data is practically suggested. Previous studies related to the medical field have focused on simply using blockchain technology to exchange medical information, improve payment systems, and logistics systems. In addition, previous research on the EHR system using blockchain technology to utilize medical big data was not advanced and remained at the rudimentary steps. However, it is a step forward compared to previous studies to present a feasible EHR platform to utilize medical big data through detailed workflow.

- Another advanced point is to present a IPFS storing and sharing technique to improve the scalability and security, which is a blind spot of the block chain technology for the utilization of Big Data. As mentioned at the outset of this proposal, the simultaneous existence of security and scalability is the same as the relationship between shield and spear. The two factors compete each other in blockchain technology. However, for the utilization of medical big data, both characteristics of the blockchain are required reciprocally.

  By simply using cloud technology, scalability could be increased, but it can also provide a risk

factor to security. To solve this complicated problem, IPFS, a distributed storage technology, is employed and now it can be easily used through programming on AWS. In this study, it was possible to apply the coexistence of scalability and security of the blockchain technology, which had not been solved before, by reflecting the development of the latest technology.

- Finally, this study is advanced in realization within a mobile environment. In the present era, the flow of information is rapidly moving from monitors to the mobile. In order to develop a platform for real-time utilization of medical information, it must be implemented in a mobile environment. In relation to this, this proposal designed an architecture that can be practically implemented in a mobile environment.

Based on the merits of our model, it is believed that our blockchain enabled solution is a step towards efficient management of e-health records on mobile clouds, which is promising in many healthcare applications.

2. Limitation and Future work of Proposed System

Apart from the advantages mentioned in the previous section, this study has the following limitations. To begin with, it was not possible to set professional design using a specific programming language to lay out each step of smart contract. This limitation can be eliminated in future studies with the help of experts or additional studies.

In addition, since it was not possible to design an actual smart contract, there is a short point that could not be compared with other studies by operating it in reality. In future research, it will be necessary to check the efficiency and safety of the proposed platform by measuring gas consumption, information transfer speed, and level of security at each stage.

To ensure the better exploitation and implementation of blockchain architecture on the medical Big Data, it necessitates a good understanding of the technology as well as what it entails to achieve the desired objectives. The blockchain-based architecture poses several significant challenges and opportunities for the healthcare industry as it’s not a fully matured platform nor a remedy solution to be implemented proximately. It requires us to address different organizational, technical, and performance-related challenges before a blockchain-enabled EHRs solution can be implemented and adopted by various organizations worldwide.

Some of the problems current researchers and business organizations facing today are described below:

- **Scalability limitations**: Medical information is being accumulated exponentially in the amount and size through the advancement in storing capacity of each device. In order to accept various and huge data, it is necessary to implement a disruptive smart contract technology that can accept and process information in real time while maintaining security.

- **Adoption and enticements for stakeholders**: In order for the blockchain-based EHR system to become universal, improvements in interface are needed for the convenience of users. The development of a user-interface that can be easily used by elderly patients or people who are incompetent in information technology is also an important point. No matter how great a technology, if it does not provide convenience to users, it can only stay in the laboratory.

- **Regulatory consideration and compliance**: It is necessary for the government to take the lead, to incentivize individual users and medical organizations to accommodate these systems, and to readjust the relevant laws to create an environment for technology development.

References

김종식, 박민재, 양경란(2020), 디지털 트랜스포메이션 전략, 지식플랫폼

A. Jindal, N. Kumar, and M. Singh (2020), A unified framework for big data acquisition, storage, and analytics for demand response management in smart cities, Future Generation Computer Systems, vol. 108, pp. 921–934.

A. Oussous, F.-Z. Benjelloun, A. A. Lahcen, and S. Belfkih (2018), Big data technologies: A survey, Journal of King Saud University-Computer and Information Sciences, vol. 30, no. 4, pp. 431–448

Alrebdi, N., Alabdulatif, A., Iwendi, C., and Lian, Z.(2022), SVBE: Searchable and verifiable blockchain-based electronic medical records system. Scientific Report, 12, 266.

Amazon Web Servise (AWS, 2021), What is Ethereum?, Retrieved Jan 12 from https://aws.amazon.com/ko/blockchain/what-is-ethereum/

Azaria A, Ekblaw A, Vieira T, and Lippman A. (2016), MedRec: Using blockchain for medical aata access and permission management, 2016 2nd International Conference on Open and Big Data (OBD), 2016:25-30.

Banks MA.(2020), Sizing up big data, Nat Med. 2020;26(1):5-6.

C. H. Liu, Q. Lin, and S. Wen, “Blockchain-enabled data collection and sharing for industrial IoT with deep reinforcement learning,” IEEE Transactions on Industrial Informatics, vol. 15, no. 6, pp. 3516–3526, Jun. 2019.

Chen J, Lv Z, and Song H (2019), Design of personnel big data management system based on blockchain, Future Generation Compter System, 101(2019):1122-1129

D. C. Nguyen, P. N. Pathirana, M. Ding, and A. Seneviratne (2020), Integration of blockchain and cloud of things: Architecture, applications and challenges, IEEE Communications Surveys & Tutorials.

Dagher GG, Mohler J, Milojkovic M, and Marella PB (2018), Ancile: privacy-preserving framework for access control and interoperability of electronic health records using Blockchain technology, Sustain Cities Soc., 2018;39:283–97.

Dagher GG, Mohler J, Milojkovic M, and Marella PB. (2018), Ancile: privacy-preserving framework for access control and interoperability of electronic health records using Blockchain technology, Sustain Cities Soc 2018;39:283–97.

E. Karafiloski and A. Mishev (2017), Blockchain solutions for big data challenges: A literature review, in IEEE EUROCON 2017-17th International Conference on Smart Technologies, Ohrid, Macedonia, pp. 763–768.

Fan K, Wang S, Ren Y, Li H, and Yang Y. (2018), MedBlock: efficient and secure medical data sharing via Blockchain, J Med Syst 2018;42(8):136.

Fang HSA, Tan TH, Tan YFC and Tan CJM (2020), Blockchain personal health records: Systematic review, J Med Internet Res 2021;23(4):e25094

Gil Press (2017), 6 Predictions For The $203 Billion Big Data Analytics Market, Retrieved Jan 2, 2022 from https://www.forbes.com/sites/gilpress/2017/01/20/6-predictions-for-the-203-billion-big-data-analytics-market/?sh=613f2f692083

H. V. Jagadish, J. Gehrke, A. Labrinidis, Y. Papakonstantinou, J. M. Patel, R. Ramakrishnan, and C. Shahabi (2014), Big data and its technical challenges, Communications of the ACM, vol. 57, no. 7, pp. 86–94.

Healthcare Data Breach Statistics (2020), 2020 Healthcare Data Breach Report: 25% Increase in Breaches in 2020, from https://www.hipaajournal.com/2020-healthcare-data-breach-report-us/[Accessed on 03 Jan. 2022].

Hua J., Zhu H., Wang F., Liu X., and Lu R. (2019), CINEMA: efficient and privacy-preserving online medical primary diagnosis with skyline query. IEEE Internet Things J 2019;6(2):1450–61.

Kim, M., Yu, S., Lee, J., Park, Y., and Park, Y. (2020), Design of secure protocol for cloud-assisted electronic health record system using blockchain, Sensors, vol. 20, no. 10, p. 2913

Li M, Yu S, Zheng Y, Ren K, and Lou W. (2013), Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption, IEEE Trans Parallel Distrib Syst 2013;24(1):131–43.

M. Hölbl, M. Kompara, A. Kamišalic, and L. N. Zlatolas(2018), A systematic review of the use of blockchain in healthcare, Symmetry, vol. 10, no. 10, p. 470

Marc S, Arlene S, & Steven D. S. (2006), The Edwin Smith Papyrus: the birth of analytical thinking in medicine and otolaryngology, Laryngoscope, Feb. 2006;116(2), 182-188

MarketsandMarkets (2020), Big Data Market Worth $229.4 Billion by 2025, Retrieved January 2, 2022, from https://www.marketsandmarkets.com/PressReleases/big-data.asp

Park, Y., and Park, Y. (2016), Three-factor user authentication and key agreement using elliptic curve cryptosystem in wireless sensor networks, Sensors, vol. 16, no. 12, p. 2123

T. McConaghy, R. Marques, A. Muller, D. De Jonghe, T. McConaghy, G. McMullen, R. Henderson, S. Bellemare, and A. Granzotto (2016), “BigchainDB: a scalable blockchain database,” white paper, BigChainDB

Tith D, Lee J, Suzuki H, A. M, Taira N, Obi T, and Ohyama N (2020), Application of blockchain to maintaining patient records in electronic health record for enhanced privacy, scalability, and availability, Healthcare Informatics, 2020 Jan;26(1):3-12

X. Fan and Y. Huo (2020), Blockchain based dynamic spectrum access of non-real-time data in cyber-physical-social systems, IEEE Access, vol. 8, pp. 64 486–64 498.

Zheng, Z., Xie, S., Dai, H., Chen, X., and Wang, H.(2017), An overview of blockchain technology: Architecture, consensus, and future trends, IEEE International Congress on Big Data (BigData Congress), 2017:557-564..

Acknowledgements

I would like to express my deepest gratitude to Professor Jong-Sik Kim and Professor Young-Hee Koh for giving illuminating and novel advice and guides to elevate the rudimentary proposal to a sound level.