T2R2 東京科学大学 リサーチリポジトリ
Science Tokyo Research Repository

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Data Sharing and Consent Management of Electronic Health Record based on the Blockchain Technology
著者(和文)	テイント ダラ
Author(English)	Dara Tith
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第11606号, 授与年月日:2020年9月25日, 学位の種別:課程博士, 審査員:小尾 高史,奥村 学,熊澤 逸夫,小池 康晴,長谷川 晶一
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第11606号, Conferred date:2020/9/25, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Powered by T2R2 (Science Tokyo Research Repository)

Data Sharing and Consent Management of Electronic Health Record based on the Blockchain Technology

Thesis Advisor

Associate Professor Takashi Obi

A dissertation presented

by

TITH Dara

to

Department of Information and Communication

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Engineering

Laboratory for Future Interdisciplinary Research of Science and Technology,
Institute of Innovative Research,
Tokyo Institute of Technology, Japan

To my family …

Acknowledgments

This thesis would have been impossible without the support and mentoring of my supervisor, associate prof. Takashi Obi and co-supervisor, associate prof. Joong-Sun Lee. Even after several years of working with them, I am constantly surprised by their amazing intelligence, infinite energy, boundless optimism, and genuine friendliness. I have also received a lot of input and support from assistant prof. Hiroyuki Suzuki who always give encouragement and useful critiques of this research work.

I would like to thank assistant prof. Chroeng Sopheap, who is the head of medical laboratory of Cambodia China Friendship Preah Kossomak hospital, for the actively providing documents and advising me about the medical system in Cambodia.

In addition, I would like to thank co-authors of my published research articles for their insight comments, hard questions, collaboration and contributions to this thesis. I learned a lot from them: prof. Nagaaki Ohyama, researcher Naoko Taira and Mr.W.M.A.B Wijesundara. I also thank my friends and colleagues who belong to the Ohyama-Obi laboratory.

Finally, I would like to thank to the Japan International Cooperation Agency (JICA) for the doctoral scholarship support.

Abstract

Electronic Health Record (EHR) has been increasingly used as an effective method to share patients' records based on the patient's consent among different hospitals. However, it is still a challenge to access scattered patient data through multiple EHRs; and patients have difficulties to modify consent after providing it to hospitals. In this thesis, I present my contributions to build a system to access patient records among EHRs without relying on the centralized supervisory system. In addition, patients can manage their consents flexibly, and healthcare organizations can get patient consents efficiently for variety of purposes. To build that system, I apply consortium blockchain to compose a distributed system using Hyperledger Fabric incorporating existent EHRs. Then, I introduce a new e-consent model using purpose-based access control scheme to this proposed blockchain system. I designed this system by adapting the security and privacy technology into the above blockchain platform and formalized a purpose-based consent model for allowing requestors for accessing to data flexibly. In my system, peer nodes hold the same blockchain in which address and metadata of patient record in EHR is written. To protect patient's privacy, I use proxy re-encryption scheme when their data are transferred. I designed and implemented various Chaincodes to handle business logic agreed by member organizations of the network. My system supports the scalability and availability in adapting to existing EHRs for enhancing security and privacy-preserving in managing patient records. It also provides the fine-grained way of medical staff's access request with diverse intended purposes of accessing to data.

Keywords: Electronic Health Record, E-Consent, Interoperability, Medical Information Sharing, Blockchain.

Contents

Abbreviation

ABAC	Attribute-Based Access Control
AES	Advanced Encryption Standard
AFGH	Ateniese, Fu, Green & Hohenberger
BBS	Blaze, Bleumer & Strauss
CA	Certificate Authority
CIA	Confidentiality, Integrity & Accountability
CRL	Certificate Revocation List
CT	Computed Tomography
DAC	Discretionary Access Control
DApp	Decentralized Application
EC-Elgamal	Elliptic Curve Elgamal
E-Consent	Electronic Consent
ECert	Enrolment Certification
EHR	Electronic Health Record
EMR	Electronic Medical Record
GDPR	General Data Protection Regulation
HIPAA	Health Insurance Portability and Accountability Act

HIT	Healthcare Information Technology
HLF	Hyperledger Fabric
MAC	Mandatory Access Control
PBFT	Practical Byzantine Fault Tolerance
PHR	Personal Health Record
PKI	Public Key Infrastructure
PoW	Proof-of-Work
PoS	Proof-of-Stake
PRE	Proxy Re-Encryption
RAdAc	Risk-Adaptable Access Control
RBAC	Role-Based Access Control
UTXO	Unspent Transaction Output

List of Figures

Chapter 1 Introduction

The Electronic Health Record (EHR) [1] has been increasingly used as an effective method to share patients' records among different hospitals. EHRs is a vital part of health IT. It can:

- Contain a patient's medical history, diagnoses, medications, treatment plans, immunization dates, allergies, radiology images, and laboratory and test results
- Allow access to evidence-based tools that providers can use to make decisions about a patient's care
- Automate and streamline provider workflow

In this chapter, I present the essential aspects of the EHR such as the interoperability of the medical systems and the consent of information sharing in the section 1.1 and section 1.2 respectively. In section 1.3, I talk about the problems that current EHR system are facing to deal with management of patient's records while preserving the patient's privacy for exchange medical records across different medical institutions. I then explain about my motivations, in section 1.4, for enhancing the medical records exchange and consent management across different EHRs. In section 1.5 and 1.6, I present about objectives and contributions of my research for overcoming challenges in the EHR for medical information sharing. In section 1.7, I describe my research methods for collecting, filtering and grouping resources, which I can learn from them for achieving my research's goals. Finally, in section 1.8 and 1.9, I show the brief descriptions of remained chapters, and my research publications and activities.

1.1 Interoperability of Medical Information System

Since the 1990s, advances in Information and Communication Technologies (ICTs) in healthcare have created new ways of managing patients' information through the digitization of health-related information. This improvement of digitization in healthcare has led to the generation of massive electronic records about patients. Such growth poses unprecedented demands healthcare system to have the interoperability with another parties and data protection while in use and exchange. In term of healthcare, interoperability is the ability of systems and devices to exchange and interpret data, which is showed in Figure 1.

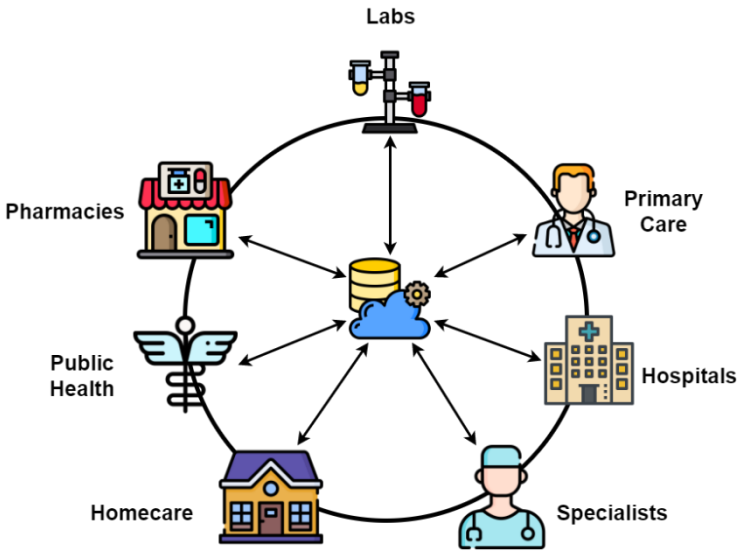


Figure 1: Interoperability of Healthcare

The interoperability or called integrate of healthcare is relied on the organizations where are having the healthcare system that is available for the medical information sharing. Otherwise, it is difficult for requesting and sharing the patient's records in the interoperability of healthcare network. The interoperability of healthcare allows healthcare providers for the exchange health information with other members in the same network about their patients for receiving the most up-to-date knowledge [2]–[4] about patient health conditions. For instance, healthcare providers need to exchange information, such as clinical notes, observations, laboratory tests, diagnostic

imaging reports, treatments, therapies, drugs administered, allergies and letters, x-rays, and bills. Then, they will make better treatment decisions [5], [6], reduce the cost of providing ambulatory care [7], enable the collection of data from many resources for research purposes, and simplify the data-management process for patients. In addition, the interoperability provides well-communicating systems, which can improve operational efficiency, reducing time spent on administrative tasks like manually entering data received from faxes [8]. It can also reduce duplicate clinical interventions like imaging studies or lab orders, decreasing overall health system cost, decreasing waste, and improving patient safety by reducing the exposure to radiation or invasive procedures [9], [10]. Finally, interoperability may also improve clinical care, by facilitating improved access to relevant, longitudinal clinical data at the point-of-care.

Achieving interoperability will certainly advance the adoption of eHealth [11]. However, when managing highly sensitive information such as health-related data, interoperability could lead to abuse of the patients' privacy [12]. The consequences could be irreversible and have a long-term impact on the patient, as well as on his social environment.

1.2 Consent for Information Sharing

In the context of healthcare information, a consent [13] refers to patient's granting the requester for the access to specific details of their records stored in the health record system. Simultaneously, in the context of healthcare professional and patient interaction, informed consent [14] represents the voluntary agreement given by the patient before having a medical treatment. In any case, the patient's consent is essential to the healthcare information access, and traditionally it has been given using paper-based forms with patient signature at most of healthcare and research institutions [15], [16]. Thus, once submitted, it is difficult for patients to change their decision,

which makes them cautious in signing the paper and tend to be reluctant to share their data with others. To sum up with above arguments, the patient consent is very essential to use as the rule of sharing patient medical records. Moreover, it is mandatory to get patient consent beforehand for sharing and utilizing patient data. However, the requirement for consent is underpinned by ethical principles of respect for persons and individual autonomy. In most countries, consent [17] is also the basis for data protection and privacy law although there may be exceptions for medical research.

In medical system for information sharing, it has many types of consent for allowing someone to access to medical data. They are a broad consent, blanked consent and specific consent. The broad consent [18] is an unspecified range of future research subject to a few content and/or process restrictions. In general, the broad consent can be described as a tool that enables research participants to consent to a variety of research projects. On the other hand, the blanket consent is an unspecified range of future research subject to a few contents of medical records and open-ended permission without any limitations. Nevertheless, the specific consent is to consent to a wide range of options.

So far, e-Consent system is created for the digital of consent management. There are essentially three broad requirements [19] of e-Consent:

1. *Patient consent has to be captured as a legal record.* The burden of maintaining of paper of patient's consent is a hard task for healthcare service providers. In addition, the consent form as the paper does not provide the availability to patients because they may require to come to hospitals for providing their consent. For these scenarios, an e-Consent system may be quite actively used in capturing the consent record.

Using structured data entry, clinicians could be guided to capture the specific form of consent granted, including inclusion and exclusion criteria from predefined lists where possible. Clinicians using an e-Consent record system cannot proceed with actions on the computer system until the consent process has been documented.

2. *An e-Consent system requires patients and clinicians to signify*. After patients agree to provide their consents, they need to sign on their consents. Then, the clinicians who receive these consents from patients also have to sign for accepting that patients provided the correct consent based on term and condition of the consent. The consent captures patient's purpose for accessing to data from data requestors. However, the decision to access information remains with the clinician. This type of e-consent system can generate an audit trail of accesses to information and can be used to both retrospectively check that patient consent is being observed or, in cases of dispute, can act as an authoritative record.
3. *The e-Consent system could act as a gatekeeper and permit only consented individuals to access information.* In a distributed information environment, many individuals may choose to access patient information. If I wish to ensure only those with proper authorization view patient data, then the e-Consent system will check to see that the individual who wishes to access the information is able to satisfy the conditions of consent before the access is granted. There would thus need to be a ''consent machine'' implemented that can read the consent record and match the conditions in the record with the individual seeking to access information.

However, the process of migrating the consent management from paper to digital is ongoing and characterized by several open challenges.

1.3 Challenges in Medical Information Sharing

I mentioned about the benefit of healthcare interoperability, the definition of consent in the medical system, and how patients use consents to allow medical staffs to claim medical records for the specific usage purpose. Nevertheless, it still has some open challenges in the current medical system in which demanding healthcare providers to improve their system for overcoming them.

Currently, patients may visit many different care providers' offices during their lifetime, and they met different hospital's clerks, nurses and doctors while they have the medical treatment. Eventually, their medical data are scattered to different healthcare providers' database. In consequence, it is difficult for patients to collect their medical data and doctors to access these scattered patient data through multiple EHRs. It is because of existing EHRs [20] are regionally limited one or belong to affiliated hospitals. In addition, based on the report published by the Office of the National Coordinator for Health Information Technology (ONC) [21], the main barrier to access patient records lies in the difficulty to find provider's addresses; and healthcare information was not always available when needed. The problems can be solved, if medical institutions interoperate with each other. Nevertheless, increasing costs of healthcare infrastructures and software in the industry have caused tremendous pressure on world economies [22].

Another problem is some medical institutions adapt different medical data standard. In consequence, it is hard for them to interconnect with each other under the supervision of the centralized system because of lacking of interoperable data standards enforcement. Without the enforcement of existing interoperable standards data, health data can vary in formats and structures that are hard to interpret and integrate into other systems [23].

Hospitals retain a right to allow the medical access based on the hospital's policy and patient's consent. Simultaneously, consents are still paper based and not sufficient, which meant that it is difficult for the patient to express his specific purpose for accessing to their data. Moreover, patients have difficulties [24] to give the new consent to medical treatment activities that were not foreseen. Thus, the exchange medical record is still having challenges, it insists many security and privacy preserving platforms to propose variety solutions to overcome above problems.

1.4 Motivation

"All EHR system should be ready or well-prepared for the revolution", it is the quote from the research paper of R. S. Evans [12]. This researcher said EHR system will be increased of use, storage and access of social, economic, behavioural and environmental data in the next 25 years from 2016. It is meant that EHR system will be integrated with other systems to make the big eco-system for enhancing social impact.

There have been several projects to overcome these problems, however, they are very hard tasks which involve redesign or upgrade of existing EHR systems requiring substantial amount of

Figure 2: Centralized system characteristic

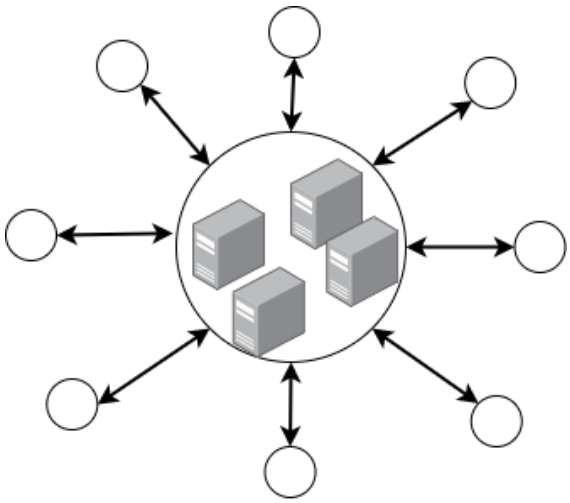
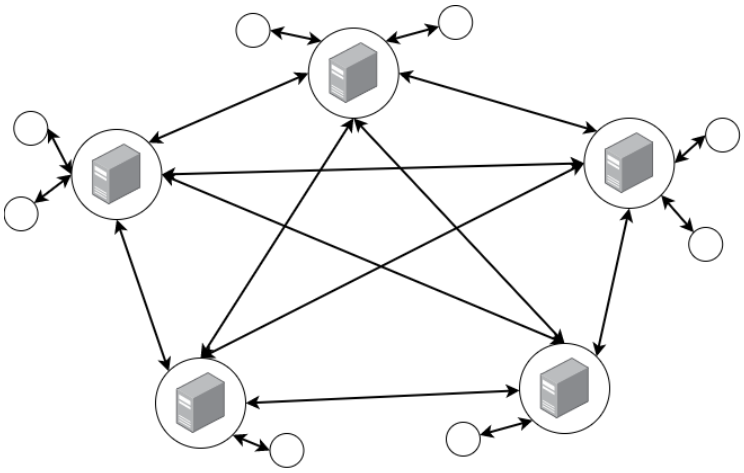


Figure 3: Decentralized system characteristic



expenses. Among them, one of the most actively ongoing programs is run by CommonWell Health Alliance [25] in USA, a non-profit association. They support EHRs, care providers and Healthcare Information Technology (HIT) vendors to connect to their nationwide interoperability network via certified integration platforms and intermediaries. As shown in Figure 2, its structures rely on one individual to make decisions, provide direction of healthcare interoperability, and it allows patients and doctors to search and receive the patient's scattered medical records [26] in different EHRs. This kind of system is convenience for the governance. In addition, scalability of the centralized system is possible when the improvements have the same objectives with the business idea of the centralized system.

I learnt the importance of sharing the medical data of the CommonWell Health Alliance so that I want to provide a new choice of the information system to core hospitals, small clinics, health insurance companies etc. for the interoperability of healthcare without relying on the top-tier system. I acknowledge that some of medical institutions have already the system infrastructure for managing the patient records, patients and medical staff's identity; exchange the medical records across the departments in the hospitals. Based on these existing system infrastructures, hospitals can use or improve the current capacity of the system to handle the additional tasks for data exchange. In Figure 3, it is the decentralized system characteristic, which all nodes communicate with each other without relying on the centralized system.

Another motivation of my research is when patient records are accessed for some reason, the history of all such events needs to be recorded in a log file of each hospital database for later audit on access histories. The log file is used for reconstructing of past state of medical records and it can be represented as a legal document [27]–[29]. Thus, I should firmly protect the log file from illegal access by making it transparent and immutable if possible.

Last but not least, various types of e-consent [19], [30] have been introduced which allow patients to give the consent electronically with their digital signatures and to withdraw it later if necessary. However, most of e-consent models are based on the centralized architecture, and some are built with trusted third-party delegation to evaluate patient consent and guarantee it [31]. There are also decentralized models that employ blockchain technology [32], [33]. Among them, Dwarna project [34] provide a well-designed web portal for dynamic consent that harnesses the blockchain ledger, which acts as a hub connecting participants in the biobank project. As long as I learnt deeply about this Dwarna project, I investigated that this project has some points that need to be improved.

So far, my system has not solved all of problems in the medical system, and yet fully had the functions of medical information sharing. However, I contribute a system that can allow all of medical institutions, governments and other organizations such as health insurances to communicate and exchange medical data with each other without relying on the supervision of centralized system. Likewise, all of the log activities are stored in the immutable data storage. Moreover, I illustrate the new decentralized e-consent scheme of medical record exchange.

1.5 Research Objectives

I start talking about the Alice's journey of her visiting hospitals, which is shown in the Figure 4. Since 2016, she started visiting the hospital_A and the following years she went to different hospitals, such as hospital B, hospital C and hospital D, for having the medical treatment. In 2020, she visits hospital_C again and the doctor, called him Bob, tends to review her health histories. Alice remembers the previous hospitals that she went, but she does not remember the type of medical records and the time she had visited in the past. In this case, she can claim her past histories

from hospital_D because both hospitals, hospital_C and hospital_D, have the healthcare interoperability with the same EHR_3 system. Unfortunately, she has to go to her past visiting hospitals, which has not connected with EHR_3 system, for requesting her past records. In addition, Alice wants to allow Bob to access to her past records in the future. However, she has no idea on how to access to her past consents for updating them in order to expand the availability of sharing her records.

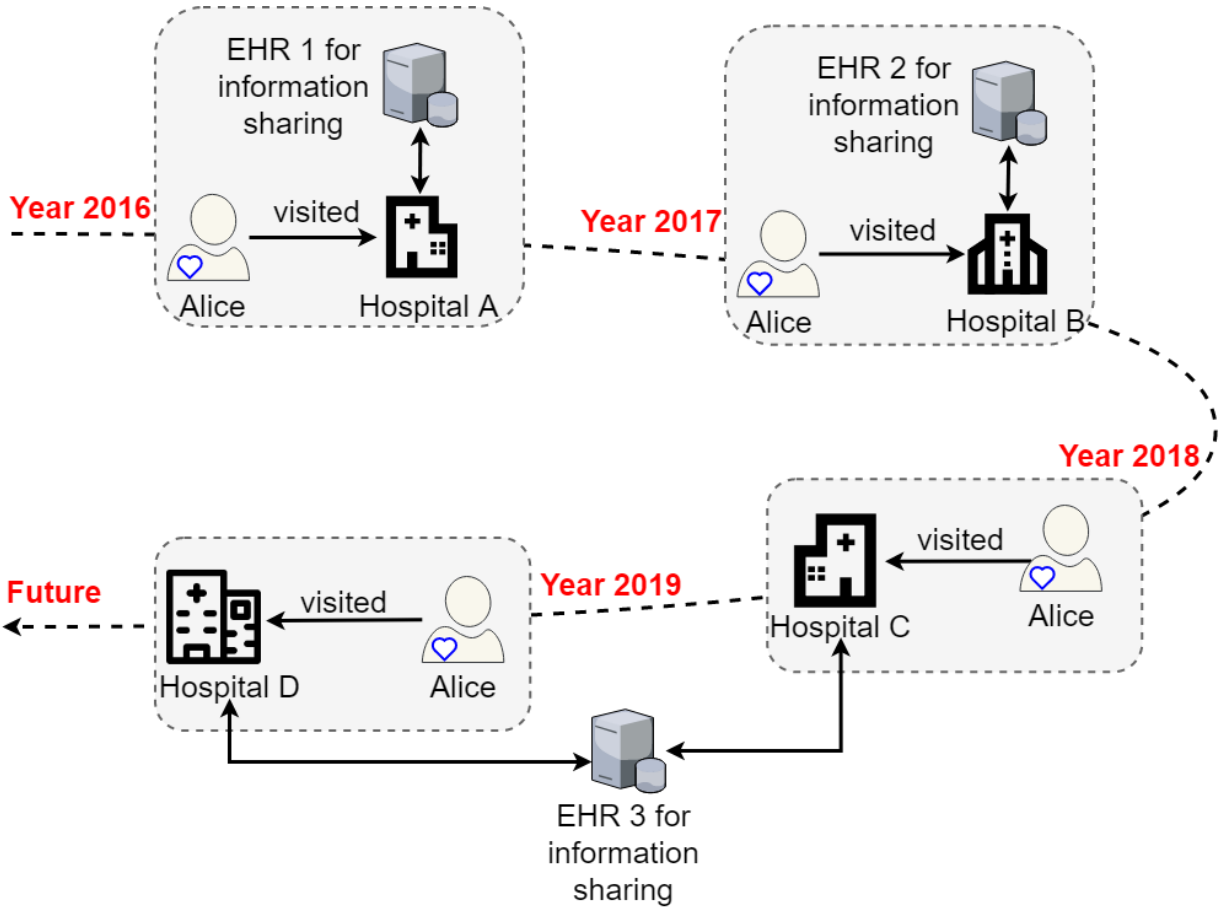


Figure 4: Histories of visiting hospitals of Alice

As given scenarios, these hospitals do not have a wide healthcare interoperability with different medical institutions. These hospitals also do not provide availability to doctors or other stakeholders for accessing and collecting medical records and patients have the difficulty to update or check their consents.

In this thesis, I propose a decentralized system to:

- Supporting the interoperability of EHRs for medical data sharing
- Enabling medical data sharing based on the patient consent stored in the secured database

To succeed goals of above proposed system, I present five major features of contributions to achieve my research objectives:

1. Providing a trusted directory of patient data in EHRs which guarantees an access as well as integrity of the data itself.
2. Allowing patients to manage their consent elaborately by assigning their intended purpose of use to each of the data. Patients can alter their consents flexibly or withdraw it afterwards. They can also monitor history of all their consents on the data.
3. Allowing healthcare organizations and research institutions to obtain patient records for future needs based on patient's consents. Since stored patient's consents on the shared and immutable ledger of blockchain, patient's consents have high transparency and traceability with data integrity.
4. Strengthening security in dealing with patient data by utilizing peculiar encryption scheme and providing transparent and undeniable audit trail based on immutable access log.
5. Providing scalability to cover multiple existing EHRs of regional or core hospitals with least modification and availability of the system without relying on centralized supervisory system.

In order to accomplish my goals, I adopt blockchain technology, especially consortium type [35], using Hyperledger Fabric (HLF) platform. Multiple hospitals gather to form a consortium having private peer-to-peer network and permissions to join it are determined based on consensus among the members.

Why Blockchain in my research?

Recent studies propose that blockchain is able to disrupt trusted business models mainly used in healthcare systems for information exchange purposes [36]. Considering the number of transactions (eg, information sharing) among health care entities and the expenses that hospitals experience in maintaining the Health Information Exchange (HIE) systems, the underlying blockchain technology of democratically sustained public ledgers of the records opens new and challenging opportunities for the health care industry. Blockchain can create an electronic context in which business transactions (such as information-sharing initiatives) between parties are conducted via a distributed community rather than a central authority or a single entity. This might essentially affect the transparency of the system and the role each entity plays [37].

Blockchains are widely categorized into three types, public, private and consortium according to purpose of use and features of members of the systems [38]. Blockchain provides trust by totally decentralized way, without relying on single actor [39]–[42]. All nodes or multiple independently designated nodes participate in operation based on a consensus made by them. With these characteristics, blockchain is gradually gaining popularity in healthcare industry [43], [44].

In addition, it maintains data integrity [43], [44] because all members have the copied of Blockchain and once the data are recorded in the Blockchain, it can no longer be modified neither deleted. If a transaction in the blockchain is modified, an updated data such as consents will be

recorded in the Blockchain. Moreover, the Blockchain can be a trustful log file [16], which is used to support the audition procedure if a malicious activity occurred in the medical information access system. While patient's data are stored in the decentralized system such as blockchain, it provides the wide accessibility to patients for creating, revoking and updating their consents of accessing data in the hospitals, which are members of blockchain network.

1.6 Research Questions and Contributions

In this work, I strive to develop the best approaches to achieve interoperability, support dynamic in healthcare system, and to ensure patient's privacy. This research is broken down into four main research questions (***RQ#***) and describe my contributions below.

- ***RQ1*:** Hospitals connect to different EHRs for storing and downloading patients records. How do I build a decentralized system for privacy-preserving, availability, scalability and interoperability in healthcare?

Contribution: To address this question, first I studied the existing Blockchain research concepts and implementations in term of medical information exchange and consent management. I see their limitations of the solving above conditions then, I proposed the Blockchain research. Second, I learnt the theory of existing Blockchain platform, and I choose a Blockchain platform, which is suitable for me to use it to achieve my research goals. The Blockchain platform provides the interoperability among members in the network, secure data storage and enforcement of the agreement between stakeholders. However, Blockchain cannot solve the vital problems of ensuring the patient's privacy of distributing medical records. In consequence, I proposed utilizing peculiar encryption scheme and purpose-based consent for preserving the patient's privacy. I have

done the prototype system and compare it with the existing blockchain systems for ensuring the privacy-preserving, availability, scalability and interoperability in healthcare.

- ***RQ2*:** Medical records are very sensitive, but the data in Blockchain is shared among members in the network. How do I preserve the patient's privacy and security?

Contribution: Data is stored in the Blockchain as the transaction belongs to the unique key, which is used for identification the data owner. To prevent the disclosure of the identification information of data owner to other users in the network, my system pseudonymizes the data owner's id by hashing it with the random number called Salt [45], [46]. For finding the correct patient records, data requestors have to know the patient's EID. Otherwise, they cannot find the patient records in the blockchain. My Blockchain does not store the big file such as medical image, video or prescription. But it conducts to preserve data integrity and data repository so that blockchain has patient's consents, medical records' address and records metadata such as hash values of the medical records. Based on these data, the medical staffs can search for location of records and verify the record's integrity by comparing hash value of receiving records with the hash value of record metadata stored in Blockchain.

- ***RQ3*:** How do I secure the processing of sending and receiving the data, which belong to different EHRs?

Contribution: Users in my system have the ECert (explain in section 3.4.3) and it adapts the concept of the Public Key Infrastructure, which supports the user identification and the public key (asymmetric) cryptography. In my system, the medical records in the EHRs have to be encrypted by the new symmetric key and this key is encrypted by the patient's public key. This method is called Hybrid Encryption. It conducts to secure the data stored in the EHRs, but the data owner

may take longer time for sending the medical records to requestor because the key to decrypt the medical records is encrypted by data owner's public key and he needs to decrypt and encrypt encrypted key repeatedly. To overcome the above problem, I use proxy to run re-encryption function, AFGH [47], for preserving patient's privacy and reducing the time of distributing encrypted keys. Moreover, I design proxy of each hospital, in Figure 5, to handles the communication from the host machine of the hospital to other hospitals or EHRs for requesting and downloading the data on the behalf of medical staffs.

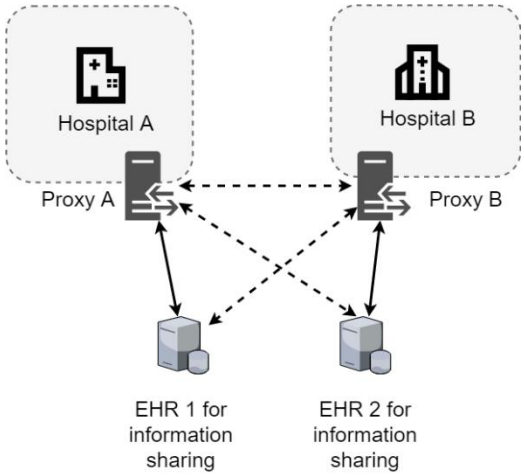


Figure 5: The proxy connects to different EHRs and another proxy

***RQ4*:** Sharing patient's data is based on the patient's consents stored with their records, which belong to different hospitals. What is the new approach of the consent model, which can be used in this scenario?

Contribution: I researched about existing access controls in term of: the allowance for accessing to the database based on (i) the occupation of user, (ii) the user's attributes and context, (iii) the purpose of data access. Simultaneously, I acknowledge that allowing accessing to data is based on the patient's consent, and the patient is the person who can decide the type of condition for allowing the data access. Thus, I cannot aim the authorization of the accessing data based on only user's role, context or purpose of data access. Each of them needs to be combined for

providing the availability to patients in order for them to allow various data requesters for accessing to their medical records. For instance, patients may share his Computed Tomography (CT) image of brain to pulmonologists. As a result, if I adapt the access based on role only, pulmonologists cannot access to that CT image because his role is not the neurology. Another scenario is patients provided his consent to the hospital that his medical records of stomach can be shared to the research students who want these data for doing a research. In consequence, it may be difficult for research students to receive that data because they are not the medical staffs of the hospital. In this research, I proposed the purpose-based consent, which is the modification version of existing purpose-based access control [48], [49], for using in the Blockchain for allowing the requester for accessing to data. This model provides the patient-centric method, which patients can allow the accessing to data based on the user's role, action of using the data and their purposes of record usages. Moreover, patients can retroactively withdraw consents at any time based on the function of activation and deactivation the consent in the Blockchain system.

1.7 Literature Survey Method

The research method, I have followed to collect and review papers, thesis, technical documents and relevant website, is inspired by the guideline of Kitchenham et al. [50], [51]. However, my method is not fully following to the Kitchenham method for the development of a review protocol. This literature survey method is based on my research field and system scenarios. This section, I begin with the presentation of the research survey protocol. I then show the survey materials, which is used to gather data systematically.

1.7.1 Survey Protocol

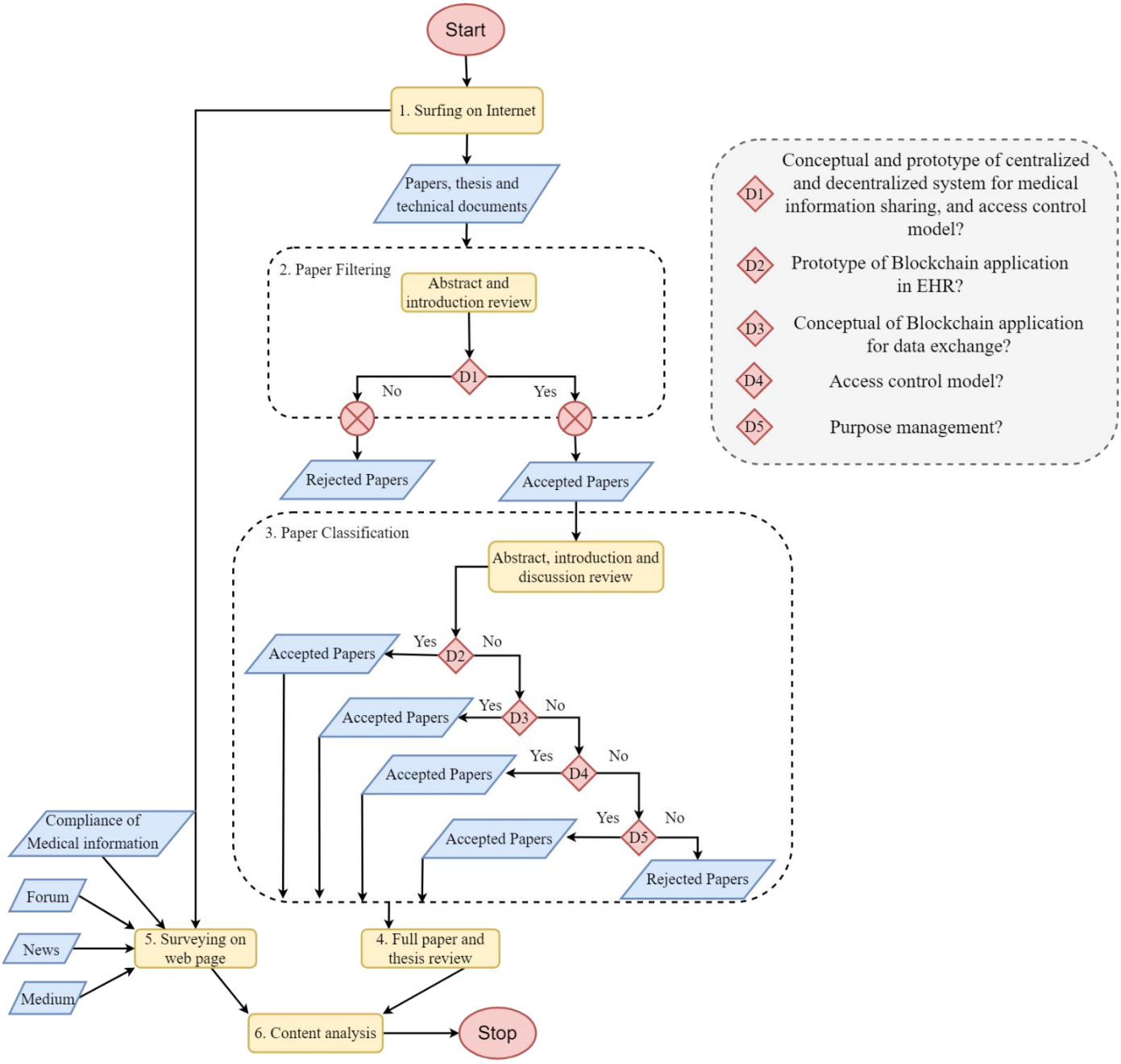


Figure 6: Survey protocol

The processing of the survey protocol is divided into six stages as shown in the flow diagram in the Figure 6. It starts from the top of the diagram, which I first surf on internet by providing some keywords for collecting research documents. I then download research papers, thesis and

technical documents, which are related to my research objectives. Among references of above research documents, I also find and select interesting research papers so that the amount of the collecting research documents is increasing.

Second step is the filtering. I start reviewing the abstract and introduction of above collecting research documents. I then select research documents, which are related to the conceptual and prototype of centralized and decentralized system for management and collection of medical records. At the same time, I measure their research objectives and results with the compliance of medical system for enhancing security and preserving-privacy of patient's records. In addition, I choose the research articles, which are related to the access control model.

Third step is the document classification. I expand my surveying method by reading discussion section, and re-reading the abstract and introduction of above filtered documents. I begin with the classification of the researches, which are about the concept of the Blockchain application for EHRs in general. I then group research documents which aimed to solve problems of medical exchange using Blockchain technology. After that, I classify filtered documents for the access control model and technique. I think that access control is one of the techniques to strengthen the security and preserving the privacy in the organizations. Finally, I group another research document, which are related to the consent and purpose model for accessing to the data.

Fourth step is to read deeply on the paper that I classified. At the same time, I proceed the fifth step to survey another source, which are related to my research objectives. They are website of the compliance of medical information, forum, news and medium, which explain about the latest information about current medical system and Blockchain application for the medical information

exchange. The last step, I analyze documents from each classification and information from the website, which I mentioned above, to answer my research questions.

1.7.2 Survey Materials

My research covers (i) research articles published in conferences and journals, (ii) thesis, books and manuscripts, such as technical reports published by the commercial and public institutions, (iii) sample source code and library for experiments, which are from Github.

I consider three input sources for my survey. Firstly, I focus on the research papers proposed by my advisor and numbers of thesis that are relevant to my addressing issue. Then, scanning the references section in each paper, a list of relevant papers (papers' titles) is extracted from the references section of those papers. Secondly, Elsevier, PubMed, IEEE, ACM and Springer databases are also the sources of my survey. I search through their databases by using my keywords and select the most relevant papers to my problem. Third source is Internet, with the help of Internet search engine, I am able to find number of interesting research papers.

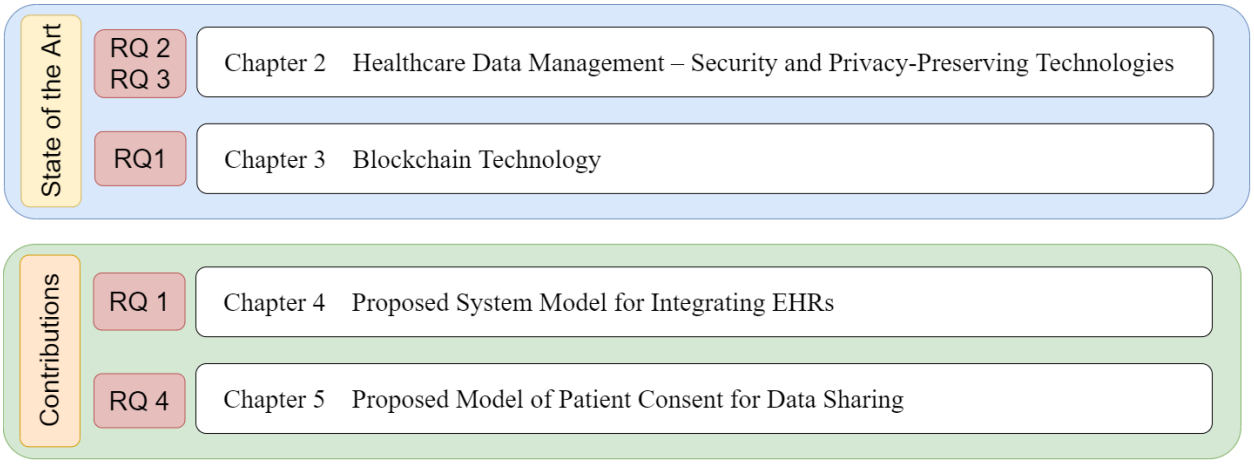


Figure 7: Thesis map

1.8 Reader's Guide

Chapters of this thesis are organized in layer, as shown in Figure 7. I group the chapters into two main parts. The following two chapters are about the state of the art and the works related to healthcare data management system, technologies which are used to enhancing security and preserving-privacy, and the blockchain technology. The second part that cover the rest of the chapters is the contribution. **Chapter 2** is divided into two parts. First part is about the basic regulations of building a medical system for the information sharing. Second part is about the technologies, which are invented by other researchers, such as encryption and decryption data methods, identity management in the information system and the access control techniques, which are used to manage the access in the organizations based on the role of user, context and access purpose.

In **Chapter 3**, I explain deeply about the blockchain technology, which is mainly used in my system. I describe the overview of this technology and other platforms, which are adapting the blockchain concept. Then, I do some survey about the importance of using blockchain technology in the medical system by reviewing the existing researches. After that I present about the existing research of prototype blockchain system for exchange, aggregation, traceability and patient consent management of medical data.

I finish explaining the state of the art so, I start presenting my contributions. In **Chapter 4**, it is about the proposed system model for integrating EHRs by showing the married between blockchain technology with existing security and privacy-preserving technologies, and applying them together. I also proposed some modification of the existing blockchain platform that I use for adapting to my use cases of medical system. I did the proof-of-concept of my proposed system to

make sure that the data is securely stored in the blockchain while maintaining the data owner's privacy and maintain the security when transferring data to other parties.

Finally, **Chapter 5** is about improving the patient e-consent management for medical records sharing. I learnt about the purpose-based access control, and then, I proposed the modification system for improving the access based on the accessing purpose that is not relied on the role hierarchy of the user. I then apply that method to the blockchain technology for providing the availability to patients for the consent management, and to doctors for retrieving medical data.

For the above two systems, I implement them in Java and Go programming language. Finally, **Chapter 6 & 7** discuss and concludes my research. I also addressed in **Chapter 7** the future work if I apply this proposed for GDPR and the hospitals.

1.9 Bibliographical Notes

The content of this thesis is based on the following publications and my researchers' activities so far.

Journal

1. Tith D, Lee J-S, Suzuki H, Wijesundara WMAB, Taira N, Obi T, Ohyama N. Application of Blockchain to Maintaining Patient Records in Electronic Health Record for Enhanced Privacy, Scalability, and Availability. Healthcare Informatic Research. 2020 Jan; 26(1):3–12.
2. Tith D, Lee J-S, Suzuki H, Wijesundara WMAB, Taira N, Obi T, Ohyama N. Patient Patient Consent Management by a Purpose-based Consent Model for Electronic Health

Record based on the Blockchain Technology. Healthcare Informatic Research Journal. **(To be published)**

Forum

3. Tith Dara, Takashi Obi, Joong-Sun Lee, Narin Piseth, Rin Darith, Lina Septiana. Blockchain-Based Blood Bank Ecosystem for Improving Public Health and Encouraging Voluntary Blood Donors. ASEAN IVO 2019. 20th November, 2019.

Chapter 2 Healthcare Data Management – Security and Privacy-Preserving Technologies

The improvement of the existing healthcare data system is to enhance the provision of healthcare service to patients, medical staffs, medical institutions and the government. Simultaneously, it has to be made to reach even bare minimum compliance with the privacy and security rules. First, I need to understand about the purpose and requirements of the system that I am going to build. At the same time, I need to take into account the regulation of medical system for preserving-privacy and security of patients and medical institutions. Next, I learn and choose the technologies, which can be used to satisfy the security and privacy protection of the specified requirements of the improvement of medical system. In this chapter 2, I present about the compliance of healthcare data management and the mandatory requirements for the building or improvement the medical system. I then illustrate the selection of existing technologies to strengthen the security and privacy protection in my proposed system.

2.1 Healthcare Data Management – Required Privacy and Security Properties

Securing healthcare data while meeting data privacy compliance demands has become a major pressure in the healthcare industry. Medical institutions are using information systems to reduce costs and to improve the quality of care. Easy-to-use technology enables healthcare staffs

to be more mobile and efficient, but also increases potential security risks. It is necessary to ensure all these new technological developments are reliable and secure because not only security, but also patient safety can be threatened. At least, those medical systems have to follow the fundamental of security properties, called CIA [52], [53], which are confidentiality, integrity, and availability of data. These properties can be defined as follows:

- ***Confidentiality*** [54] safeguards information that is gathered in the context of an intimate relationship. It addresses the issue of how to keep information exchanged in that relationship from being disclosed to third parties. Confidentiality, for example, prevents physicians from disclosing information shared with them by a patient in the course of a physician–patient relationship. Unauthorized or inadvertent disclosures of data gained as part of an intimate relationship are breaches of confidentiality.
- ***Integrity*** refers to the trustworthiness of data or resource, and it is usually phrased in terms of preserving improper or unauthorized change. Integrity includes the trustworthy data, which is the content of information, and origin integrity, which is the source of the data.
- ***Availability*** refers to the ability to use the information or desired resource whenever users want. Availability is an important aspect of reliability, as well as of system design.

Currently, when data are shared for research purposes, data anonymization can be an alternative to preserving the data owner's privacy. However, medical institutions cannot fully anonymize all of identity of the data owner when sharing the medical records. If they delete all of the sensitive information, these records will not be useful for learning or analyzing. Simultaneously, it is difficult for traceability of the records. In practice, traceability is required in

addition to interoperability and compliance with legislation and policies that regulate management of the personal data and, in particular, protected health information. Traceability of the data can be defined as the ability to retain the identities of the origin of the data, the entities who accessed the data, and the operations performed on the data. The data traceability will be particularly useful in legal cases, as well as for the patient, in defining and enforcing his access-control policy, allowing meaningful data aggregation for research purposes, and enabling reproducibility of research. Elgar et al. [55] discuss the main laws and guidelines that describe how to prepare data for use in medical research (including de-identification and pseudonymization). The authors propose the following definitions based on how the concepts of personal, identified, and identifiable data are formulated in the various legal documents:

- Personal data refers to the data that are about an individual who can reasonably be identified or identifiable.
- De-identification is the process of removing (or modifying) identifiers from the personal data so that identification is not reasonably possible.
- Pseudonymization is the step where a pseudonym or code is added to the de-identified data.
- Proportional or reasonable anonymity applies to de-identified or pseudonymized data that cannot reasonably be used to identify specific individuals.

To sum up with above arguments and requirements, before building my system, I have to learn and choose the technology, which are suitable with my use case for fulfilling the fundamental of security properties, CIA, and preserving the medical records when managing and sharing medical records across organizations.

2.2 Security and Privacy-Preserving Technologies

Currently, medical systems are having the system that are following to the security fundamental properties for enhancing the security in management and exchange of patient's health records. For instance, some system [56] encrypt EHR data, except data attributes (metadata), in order to increase security but other encrypt EHR data, identifiers (pseudonyms), keys and data attributes. Simultaneously, they have different technique to control and management the records and access of requestors based on the use case, environments and policies of the organizations. In this section, I present the technologies, which are used to enhance the security and privacy protection in current medical system. In addition, I also mention the existing access control models, which are used to grant the access to requestors who requested to the network.

2.2.1 Cryptography Primitives

The Oxford dictionary (2020) [57] defines cryptography as *the art of writing or solving codes*. Cryptography primitives are a set of mathematical techniques [58] related to the aspects of information security such as confidentiality, data integrity, entity authentication, and data origin authentication. For instance, it is used in the study [59] of methods for sending secret messages. In this section, I present the fundamental technologies, which I use for fulfilling the requirements of enhancing security and preserving-privacy in the medical system.

- **Hash Function**

Hash functions [60] function to map an input of arbitrary length to a string of fixed length, which is called the hash code. The basic idea of cryptographic hash functions *H* is that a hash-value serves as a compact representative image (message digest) of an input string, and can be used as if it were uniquely identifiable with that string. A hash function is a function which has at

least two properties, which are ease of computation and compression. In addition, a secure cryptographic hash function has the following properties:

i. One-way (pre-image resistance): $y\epsilon\{0,1\}^d$, it is hard to find an x such that $H(x) = y$

ii. String collision-resistance: It is hard to find any pair of input x, x' such that $H(x) = H(x')$

These mappings of hash function satisfy some additional cryptographic conditions, they can be used to protect the *integrity of information*. Other cryptographic applications, where hash functions are useful, are the optimization of digital signature schemes, the protection of passphrases and the commitment to a string without revealing it. Hash functions appeared in cryptographic literature when it was realized that encryption of information is not sufficient to protect its authenticity. The simplest example is the encryption with a block cipher, where every block is encrypted independently. It is clear that an active attacker can easily modify the order of the ciphertext blocks and hence of the corresponding plaintext blocks. It will be shown that cryptographic hash functions allow for efficient constructions to protect authenticity with or without secrecy.

- **Salt**

Salting technique [61] is a hedge against pre-computed dictionary attacks, the bedrock of which involves concatenating a random string of letters and numbers, namely a **Salt**, to the beginning or end of a password before hashing it. In other words, the system hashes H (password, salt), rather than solely hashing a password H (password). Hash function with no salting process unnecessarily provide defensive security guarantees, the ones armed prove indicate otherwise. In

regard of thwarting a myriad of varying online and offline attacks, salting is undoubtedly complementary, also of critical importance in conjunction with hash functions to take solace from.

- **Symmetric Cryptography**

Symmetric cryptography [62] is a type of encryption where only one key (a secret key) is used to both encrypt and decrypt electronic information. The entities communicating via symmetric encryption must exchange the key so that it can be used in the decryption process. By using symmetric encryption algorithms, data is converted to a form that cannot be understood, called *ciphertext*, by anyone who does not possess the secret key to decrypt it. Once the intended recipient who possesses the key has the message, the algorithm reverses its action so that the message is returned to its original, called *plaintext*, and understandable form.

There are two types of symmetric encryption algorithms:

Block algorithms: Set lengths of bits are encrypted in blocks of electronic data with the use of a specific secret key. As the data is being encrypted, the system holds the data in its memory as it waits for complete blocks.

Stream algorithms: it consists of a state machine that outputs at each state transition one bit of information. This stream of output bits is commonly called the running key. The state machine is nothing more than a pseudorandom number generator.

Following are the description of the Data Encryption Standard (DES) and Advance Encryption Standard (AES).

i. Data Encryption Standard (DES)

DES [63] is a symmetric algorithm developed by IBM in the early of 1970. It was the main standard for encrypting data. However, this has now been replaced by a new standard known as the Advance Encryption Standard (AES). I will explain AES in the next section. DES is insecure due to the relatively short 56-bit key size. In Janurary 1999, distributed.net and the Electronic Frontier Foundation collaborated to break a DES key successfully in 22 hours and 15 minutes. After that, it has been withdrawn as a standard by the National Institute of Standards and Technology (NIST).

DES is therefore a symmetric, 64 bit block cipher as it uses the same key for both encryption and decryption and only operates on 64 bit blocks of data at a time. The key size used is 56 bits, however a 64 bit (or eight-byte) key is actually input. The least significant bit of each byte is either used for parity (odd for DES) or set arbitrarily and does not increase the security in any way. All blocks are numbered from left to right which makes the eight bit of each byte the parity bit. Once a plain-text message is received to be encrypted, it is arranged into 64 bit blocks required for input. If the number of bits in the message is not evenly divisible by 64, then the last block will be padded. Multiple permutations and substitutions are incorporated throughout in order to increase the difficulty of performing a cryptanalysis on the cipher. However, it is generally accepted that the initial and final permutations offer little or no contribution to the security of DES and in fact some software implementations omit them.

ii. Advance Encryption Standard (AES)

AES [64] is a specification for the encryption of electronic data established by the U.S. NIST in 2001. It is a symmetric block cipher. The block and key can be chosen independently from 128, 160, 192, 224, 256 bits and need not be the same. However, the AES standard states that the

algorithm can only accept a block size of 128 bits and a choice of three keys - 128, 192, 256 bits. Depending on which version is used, the name of the standard is modified to AES-128, AES-192 or AES256 respectively. As well as these differences AES differs from DES in that it is not a feistel structure [65]. Recall that in a feistel structure, half of the data block is used to modify the other half of the data block and then the halves are swapped. In this case the entire data block is processed in parallel during each round using substitutions and permutations.

A number of AES parameters depend on the key length. For example, if the key size used is 128 then the number of rounds is 10 whereas it is 12 and 14 for 192 and 256 bits respectively. At present the most common key size likely to be used is the 128 bits key.

Asymmetric Cryptography

Asymmetric cryptography [62] is a form of encryption where keys come in pairs. One key is used for encryption, only the other can be used for decryption. There have been various of asymmetric algorithm. In this section, I mention only two algorithms that are popular to use in many applications. They are shown as the following.

i. RSA

RSA [66] was designed by Ron Rivest, Adi Shamir, and Leonard Adleman in 1978. It is one of the best known public key cryptosystems for key exchange or digital signatures or encryption of blocks of data. RSA uses a variable size encryption block and a variable size key. It is an asymmetric (public key) cryptosystem based on number theory, which is a block cipher system. It uses two prime numbers to generate the public and private keys. These two different keys are used for encryption and decryption purpose. Sender encrypts the message using Receiver public key and when the message gets transmit to receiver, then receiver can decrypt it using his own private

key [67]. RSA operations can be decomposed in three broad steps; key generation, encryption and decryption. Following are the procedures of key generation algorithm of RSA.

1. Generate two large random primes, p and q, of approximately equal size such that their product $n = p \cdot q$ is of the required bit length, e.g. 1024 bits. p and q are kept secret.
2. Compute $n = p \cdot q$ and $\varphi(n) = lcm(p-1, q-1)$. $\varphi(n)$ is kept secret.
3. Choose an integer $e, 1 < e < \varphi(n)$, such that $\gcd\big(e, \varphi(n)\big) = 1$; that is, e and $\varphi(n)$ are coprime.
4. Compute the secret exponent d, $1 < d < \varphi(n)$, such that $d \equiv e^{-1}\big(mod\ \varphi(n)\big)$
5. The public key consists of (n, e) and the private key consists of $(p, q, \varphi(n))$.

Encryption	=>	**Decryption**
Plaintext: $\boldsymbol{P}$		Ciphertext: $\boldsymbol{C}$
Ciphertext: $\boldsymbol{C = P^e\ (mod\ n)}$		Plaintext: $\boldsymbol{P = C^d\ (mod\ n)}$

RSA has many flaws in its design therefore not preferred for the commercial use. When the small values of p & q are selected for the designing of key then the encryption process becomes too weak and one can be able to decrypt the data by using random probability theory and side channel attacks. On the other hand, if large p & q lengths are selected then it consumes more time and the performance gets degraded. Further, the algorithm also requires of similar lengths for p & q, practically this is very tough conditions to satisfy. Padding techniques are required in such cases increases the system's overheads by taking more processing time.

ii. Elliptic Curve Cryptography

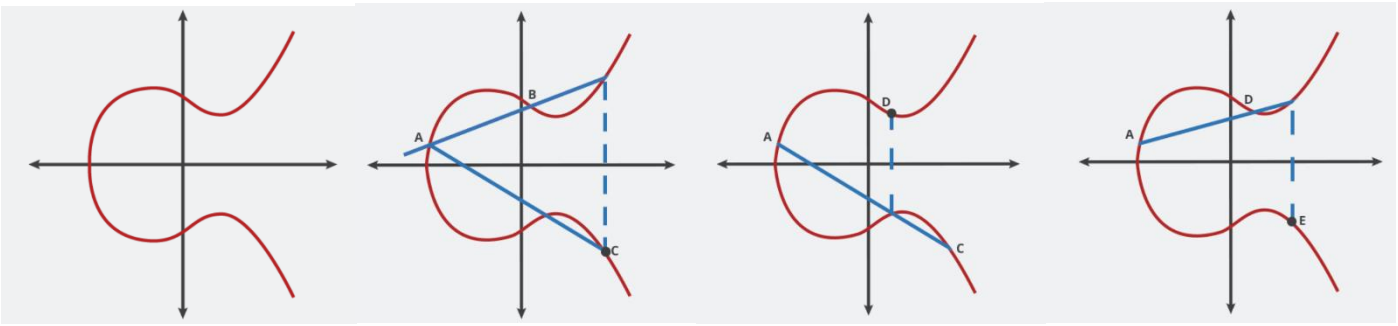


Figure 8: Elliptic Curve and technique to find public key and private key

Elliptic Curve Cryptography (ECC) [68], [69] is public-key cryptosystem like RSA. But it differs from RSA in its quicker evolving capacity and by providing attractive and alternative way to researchers of cryptographic algorithm. The security level which is given by RSA, can be provided even by smaller keys of ECC. For example, the 1024 bit security strength of a RSA could be offered by 163 bit security strength of ECC. Other than this, ECC is particularly well suited for wireless communications, like mobile phones and smart cards. Elliptic curves point of multiplication operation is found to be computationally more efficient than RSA exponentiation.

Elliptic curves defined over a finite-field provide a group structure that is used to implement the cryptographic schemes. The elements of the group are the rational points on the elliptic curve, together with a special point O (called the "point at infinity"). one of these is horizontal symmetry. Any point on the curve can be reflected over the x-axis and remain the same curve. A more interesting property is that any non-vertical line will intersect the curve in at most three places.

In Figure 8, an elliptic curve E defined over a field K can be made into an abelian group by defining an additive operation on its points. Thus, start with two points

$$A = (x_1, y_1), B = (x_2, y_2)$$

on an elliptic curve given by $y^2 = x^3 + \mathrm{a}x + \mathrm{b}$, where $4a^3 + 27b^2 \neq 0$, I can find a new point $C = (x_3, y_3)$ as follows. Draw the line L through the points A and B. The line L intersects the

curve in a uniquely determined point, which I denote as C'. Now reflect the point C' across the x-axis to obtain the point C. I define $A + B = C$. If I continue drawing the line L from C to A, it will intersect the curve on the new point D'. Then, reflect the point D' across the x-axis to obtain the point D. I define $A + C = D$. I can keep on drawing the line L to intersect the curve, I will get many new points. Hence, I have a group $G = \langle g, + \rangle$ that is the cyclic group of points that intersect of line L with the curve. Simultaneously, Elgamal cryptosystem uses the cyclic group for generating the key pairs. I will explain it later.

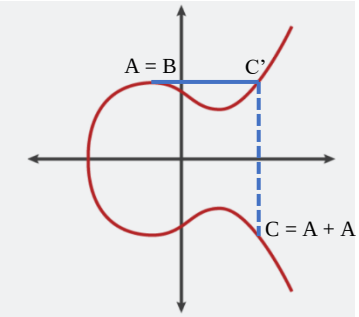


Figure 9: Elliptic Curve when A=B

In general, I avoid not to select two points of the curve E for generating the public key because it will make the public key's size become large. As shown in Figure 9, I randomly selected a point Q in the group $\langle G \rangle$ generated by $A = B$. So that, I can get the public key $Q = dA$, which d is the secret key that is the number of times that the tangent line L drawing from point A intersects with the curve E.

- **Hybrid encryption**

Hybrid encryption [70] incorporates a combination of symmetric and asymmetric encryption to benefit from the strengths of each form of encryption. The strength of the symmetric encryption is the speed of the encryption and decryption are faster than asymmetric encryption in case I apply both encryption method on the same file size. On the other hand, asymmetric encryption was

introduced to complement the inherent problem of the need to share the key in symmetrical encryption model, eliminating the need to share the key by using a pair of public-private keys.

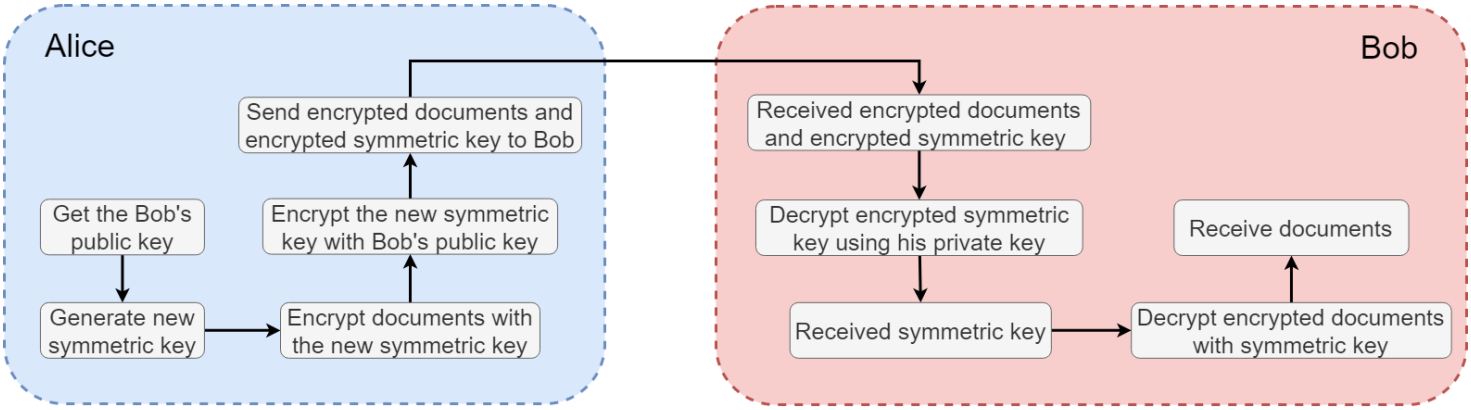


Figure 10: Procedure of hybrid encryption of sending data

The hybrid cryptosystem is itself a public-key system, who's public and private keys are the same as in the key encapsulation scheme. In place of public key system, I can use digital signature like message digesting function with symmetric key system to make hybrid crypto system. Note that for very long messages the bulk of the work in encryption/ decryption is done by the more efficient symmetric-key scheme, while the inefficient public-key scheme is used only to encrypt/decrypt a short key value. For example, to encrypt a message addressed to Bob in a hybrid technique [71]–[73] Alice does in Figure 10.

Public Key Infrastructure

PKI [74] is an abbreviation of the Public Key Infrastructure, it was developed to support the public key (asymmetric) cryptography. In this type of cryptography, the message is going to be encrypted by the sender using the public key of the receiver and then this receiver, presumably, is the only one who can decrypt this message using the corresponding private key. This direction in cryptography was introduced since 1976 [75], to solve the key management problem, using a directory called Public File in which entries are name, number and public key. The sender looks

the recipient up in the Public File by his name to find his public key. By this scenario, the sender does not have the complete confidence that the key truly belongs to desired recipient. Kohnfelder [76] proposed a solution by certificate or digitally sign each entry in the "Public File", so the certificates could be distributed through a network securely.

In the 1980's, International Telecommunication Union (ITU) decided to build a larger directory to cover all people and devices all over the world, so the result was a standard called X.500 [77], defining all characteristics of that directory. Another standard called X.509 was proposed for authentication purposes, nobody could change any entry in a directory except if he has permission. A X.509 standard defines the certificate format; it binds the identity of the key holder to the holder's public key. All these evolutions in public key cryptography have led to build a public key infrastructure (PKI) in which the digital certificates present the heart of it. For more trust authority, Certification Authority (CA) was introduced [78], which is a trusted party responsible for verify and sign the certificates. Therefore, PKI has helped the sender to receive the public key of desired recipient with the confidence that this key is really recipient's public key.

i. Components and Operations

The basic common operations in all PKIs are certification and validation. Certification is the binds of the value of public key to an entity in an authentic way. The other operation is validation, it is the process of verifying the validity of the certifications (still valid or not). A complete public key infrastructure is composed of several components which are: registration authority (RA), certificate authority (CA), security policy, PKI-enabled applications, distribution system, and certificate repository [79], [80]. The components of the PKI are shown in Figure 11 are described as following.

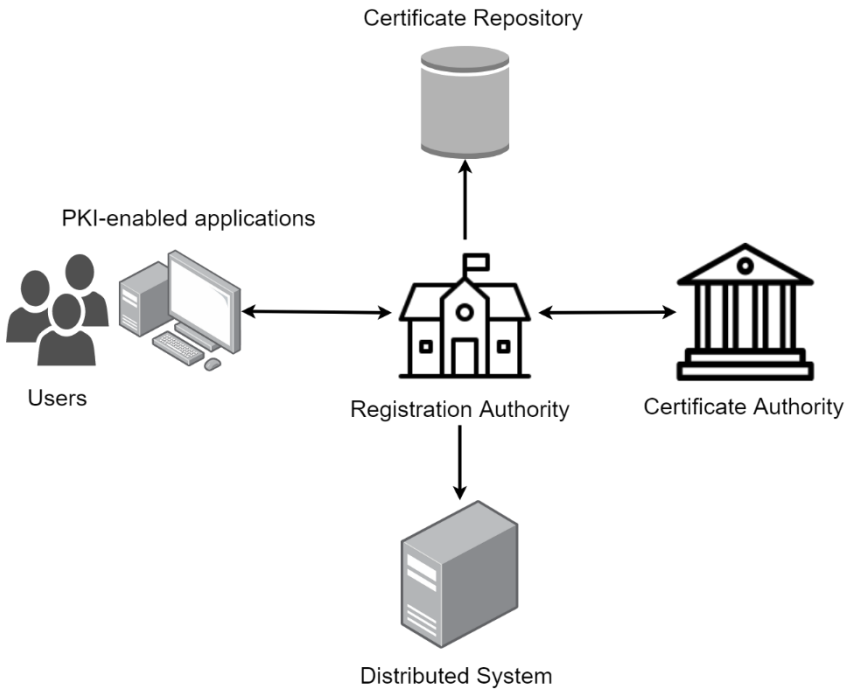


Figure 11: PKI's components

- Certificate Authority (CA): It's also called Certificate Issuer, which is used to issues the certificates and the revocation lists. A certificate is a data structure composed of both the public key value and the identified information that belong to the holder of the corresponding private key [78]. Each public key certificate is issued to an individual and each certificate has a digital signature of the issuing CA. The certificates have a lifetime of one or two years. The certificate might be revoked for several reasons such as loss of private key, compromise key or the lifetime of the certificate is terminated, etc. If any one of these situations happened; the entity who issued the certificate should be requested to invalidated (revoke) the certificate of public key. There are multiple revocation mechanisms to revoke the certificate and to allow the user to be able to check the validity of the certificate (the certificate still valid or has been revoked). All revocation mechanisms need to be timely and efficient. One of the revocation mechanisms is CRL (Certificate Revocation List)

which is a list contains certificate that have been revoked and signed digitally by the entity who had issued those certificates previously [79], [81], [82].

- Registration Authority (RA): Authenticates all the user's identities and registers the end user's information before certification, and it is used to submit all the requests to the CA. The services given by RA can be accessible through two ways: 1) Logging the administrator through the browser to the system. 2) Calling the web services interface through the application system. RA has only one super administrator which can access all the functions provided by RA where this super administrator can add more administrators if needed. Every administrator who wants managing the system must use its own smart card to prevent unauthenticated people from making any operations to the registration authority (RA) [80], [81].
- Distribution System and Certificate Repository: are used to provide storing mechanism for the certifications and CRL information [80]. The complexity of PKI can be hidden from client system by adding one more component such as Validation

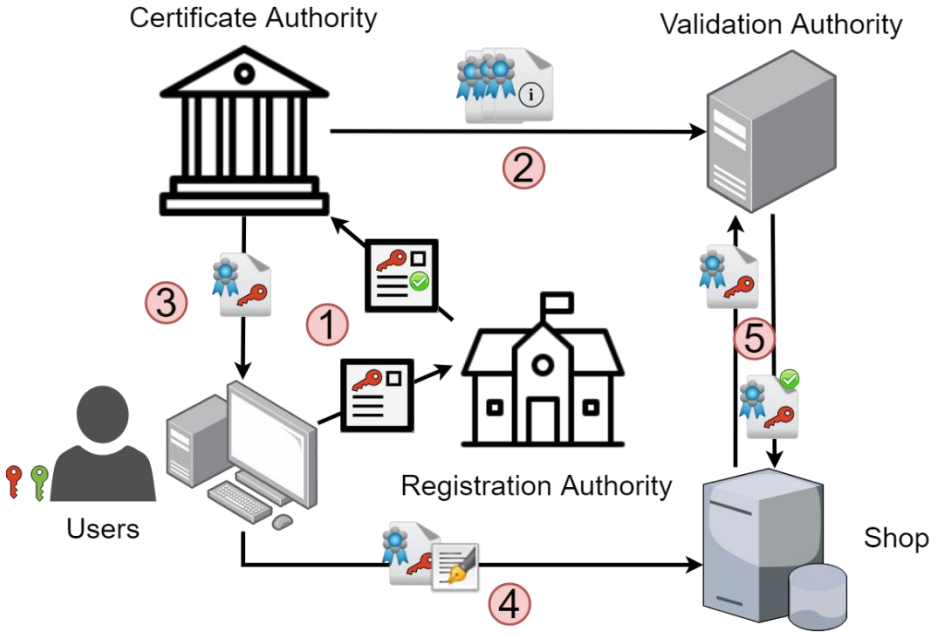


Figure 12: Protocol operations and actors of the PKI

Authority (VA) which responds to the client requests for certificates revocation status and doing the accessibility valuation of certificates on the behalf of the client system [81], [83].

The protocol operations of the PKI and actors are shown in Figure 12 where the protocol operations indicated by numbers. The protocol operations of the PKI in the figure above are:

- Certificate generation and Key: Registration Authority establishes the key holder identity and passes his/her information with the public key for certification to the CA. Then the CA or the owner of key can generate the key pair, where the most important point is the safe transportation of the private key to the holder of key. A typical lifetime is one year for a certificate after which a new certificate is issued.
- Revocation list generation: Making a list of all non-expired or revoked certificates where the CA signed these lists then sent them either to the VA or sent them immediately to the relying parties where there is no need for online status provider to doing the validation. The revocation lists generation interval is four hours.
- Signature generation: The holder of key making signs for a message including his/her public key certificate with the signature data. The frequency was chosen of the signed messages is 300 messages per day for every active user where this number has been chosen for the purpose of the forthcoming calculations.
- Certificate validation: The status of revocation of the signer's certificate must be checked by the relying party where it can be done either by checking the revocation list that most recently available or by querying the VA. The frequency of the operations that belongs to the validation is the aggregate frequency of the received messages—among all users—by the relying party.

ElGamal Cryptosystem

The ElGamal [84] cryptosystem was first described by Taher Elgamal in 1985. It is an asymmetric key encryption, which is used for key exchange. It consists of three components, which are the key generator, the encryption algorithm, and the decryption algorithm.

- **Key generation**: a sender generates a cyclic group G of order q with the generator g. Then, the sender selects a random integer x_a from $\{1, \dots, q-1\}$. Finally, the sender computes $y_a = g^{x_a}$. The sender makes y_a public while keeping x_a private. Similarly, the receiver selects a random integer x_b, from $\{1, \dots, q-1\}$, and computes $y_b = g^{x_b}$. The receiver makes y_b public while keeping x_b private.
- **Encryption**: To encrypt a message M of the sender for sending it to the receiver under receiver's public key is y_b:
    1. Map the message M to an element m of G using a reversible mapping function
    2. Choose an integer k randomly from $\{1, \dots, q-1\}$
    3. Compute $s := y_b^k$
    4. Compute $c_1 := g^k$
    5. Compute $c_2 := m \cdot s$
    6. Sender sends the ciphertext (c_1, c_2)
- **Decryption**: In order to decrypt the given ciphertext (c_1, c_2) with the receiver private key x_b
    1. Compute $s := c_1^{x_b}$. Since $c_1 := g^k, s := g^{x_b k} = y_b^k$. So that the sender and receiver shared the secret s that used by the sender in the encryption.

2. Since receiver received c_2 and $c_2 := m \cdot s$, $m := \frac{c_2}{s}$, $m := c_2 \cdot s^{-1}$. s^{-1} is the inverse of s in the group G. If G is a subgroup of a multiplicative group of inters modulo n, the modular multiplicative inverse can be computed using the Extended Euclidean Algorithm. An alternative is to compute s^{-1} as c_1^{q-x}.
3. Compute $m := c_2 \cdot s^{-1} = (m \cdot s) \cdot s^{-1} = m \cdot s \cdot c_1^{q-x} = m \cdot g^{x_b k} \cdot g^{(q-x_b)k} = m \cdot (g^q)^k = m \cdot e^k = m$

i. Blaze, Bleumer & Strauss, 1998

The Blaze, Bleumer & Strauss (BBS) is based on the ElGamal cryptosystem and introduces the notion of a "re-encryption key". Following is the process of re-encryption using BBS.

- **Key Generation**: a sender generates a cyclic group G of order q with the generator g. Then, the sender selects a random integer x_a, which is the private key, from $\{1, \dots, q-1\}$. After that, the sender computes $y_a = g^{x_a}$, which y_a is the public key. On the other hand, the receiver has $y_b = g^{x_b}$.
- **Encryption**: the sender compute $c_1 := g^{x_a k}$ and to find $c_2 := m \cdot g^k$.
- **Re-encryption**: the mail server or a proxy has the $RK_{A \to B} = \frac{x_b}{x_a} \ (mod\ q)$. Then, it re-encrypts the $C_a = (g^{x_a k}, m \cdot g^k)$ from sender to the receiver $C_b = ((g^{x_a k})^{RK_{A \to B}}, m \cdot g^k) = (g^{x_b k}, m \cdot g^k)$, which the receiver can use his private key x_b to decrypt that ciphertext.

ii. Ateniese, Fu, Green & Hohenberger, 2005

In 2005, Ateniese, Fu, Green & Hohenberger constructed the first unidirectional, collusion resistant re-encryption without any required pre-sharing between parties, based on the bilinear map. Now, I will explain the bilinear map before describing the procedure of AFGH.

Bilinear Maps

The Bilinear maps was created in 2001 by Boneh and Franklin for using them to construct a special type of encryption scheme.

Let $\mathbb{G}_1$, $\mathbb{G}_2$, $\mathbb{G}_3$ by cyclic groups of prime order q.

Function $e: \mathbb{G}_1 \times \mathbb{G}_2 \to \mathbb{G}_3$ is a bilinear map iff for all $g_1 \in \mathbb{G}_1, g_2 \in \mathbb{G}_2$, $a, b \in \mathbb{Z}_q$, that $e(g_1^a, g_2^b) = e(g_1, g_2)^{ab}$.

This re-encryption algorithm uses bilinear maps of the term $e: \mathbb{G}_1 \times \mathbb{G}_1 \to \mathbb{G}_2$, where $\mathbb{G}_1 = \langle g \rangle$. e must be efficiently computable. Also, e must be non-degenerate; that is, $\langle e(g, g) \rangle = \mathbb{G}_2$.

The AFGH Algorithm

- **Key Generation**: a sender generates a cyclic group G of order q with the generator g. Then, the sender selects a random integer x_a, which is the private key, from $\{1, \dots, q-1\}$. The sender computes the $Z = e(g, g)$ and the re-encryption key $RK_{A \to B} = (g^{x_b})^{1/x_a} = g^{x_b/x_a}$. The sender computes $y_a = g^{x_a}$, which y_a is the public key. On the other hand, the receiver has $y_b = g^{x_b}$.
- **Encryption**: the sender maps the message M to an element m of $\mathbb{G}_2$ using a reversible mapping function, and then compute $c_1 := g^{x_a k}$ and $c_2 := m \cdot Z^k$.
- **Re-encryption**: the mail server or a proxy re-encrypts the $C_a = (g^{x_a k}, m \cdot Z^k)$ from sender to the receiver $C_b = (e(g^{x_a k}, RK_{A \to B}), m \cdot Z^k) = (Z^{x_b k}, m \cdot Z^k)$, which the receiver can use his private key x_b to decrypt that ciphertext.

2.2.2 Access Control Model

Access control in computing is motivated by the need to divulge access to information and available computing resources and services to authorized entities only. An entity is a generic term that refers to an active agent capable of initiating or performing a computation of some sort (for example, an end user invoking a command or a program, a programming agent acting on behalf of a user, a running daemon process, a thread of execution, a hosting system, or a networking device). It consists of two main components: authentication and authorization. Authentication is the process of verifying the identity of users, while authorization is the process of verifying what they can access to. Normally, authentication is done before authorization. In this section, I present several concepts and models of access control.

2.2.2.1 Role-based Access Control

In RBAC [85], access decisions are based on the roles that individual users have as part of an organisation. Users take on assigned roles (e.g. doctor, nurse, etc.). Access rights are grouped by role name, and the use of resources is restricted to individuals authorised to assume the associated role. For example, within a hospital system, the role of doctor can include operations to perform a diagnosis, prescribe medication and order laboratory tests; the role of researcher can be limited to gathering anonymous clinical information for studies or for sharing those data. The use of roles to control access can be an effective means for developing and enforcing enterprise-specific security policies and for streamlining the security management process.

Under RBAC, users are granted membership into roles based on their competencies and responsibilities in the organisation. The operations that a user is permitted to perform are based on the user's role. User membership into roles can be revoked easily and new memberships

established as job assignments dictate. Role associations can be established when new operations are instituted, and old operations can be deleted as organisational functions change and evolve. This simplifies the administration and management of privileges; roles can be updated without updating the privileges for every user on an individual basis. When a user is associated with a role, the user can be given no more privilege than is necessary to perform the job; since many of the responsibilities overlap between job categories, maximum privilege for each job category could cause unauthorised access.

RBAC assumes that all permissions needed to perform a job function can be neatly encapsulated. In fact, role engineering has turned out to be a difficult task. The challenge of RBAC is the contention between strong security and easier administration. For stronger security, it is better for each role to be more granular, thus having multiple roles per user. For easier administration, it is better to have fewer roles to manage. Organisations need to comply with privacy and other regulatory mandates and to improve enforcement of security policies while lowering overall risk and administrative costs. RBAC products have sometimes proved challenging to implement.

Hierarchical RBAC [86]: under RBAC, roles can have overlapping responsibilities and privileges; that is, users belonging to different roles may need to perform common operations. Some general operations may be performed by all employees. In this situation, it would be inefficient and administratively cumbersome to specify repeatedly these general operations for each role that have a right of “Created”. Role hierarchies can be established to provide for the natural structure of an enterprise. A role hierarchy defines roles that have unique attributes and that may contain other roles; that is, one role may implicitly include the permissions that are associated with another role. Role hierarchies are a natural way of organising roles to reflect

authority, responsibility, and competency: the role in which the user is gaining membership is not mutually exclusive with another role for which the user already possesses membership.

2.2.2.2 Attribute-based Access Control

ABAC [87] is a logical access control model that is distinguishable because it controls access to objects by evaluating rules against the attributes of the subject, attributes of the object, environment conditions, and a formal relationship or access control rule defining the allowable operations for subject-object attribute and environment condition combinations. Attributes may be considered characteristics of anything that may be defined and to which a value may be assigned. All ABAC solutions contain these basic core capabilities to evaluate attributes and environment conditions, and enforce rules or relationships between those attributes and environment conditions. ABAC systems are capable of enforcing both Discretionary Access Control (DAC) and Mandatory Access Control (MAC) models. Moreover, ABAC systems can enable Risk-Adaptable Access Control (RAdAC) solutions, with risk values expressed as variable attributes.

2.3 Summary

When the medical system is built or updated, it needs to adapt the compliance and fundamental of security technique, and select the right security technologies for preserving-privacy and enhancing security to deal with the medical data management. Among the basic security properties, hash function is a technology that I use it for protecting data integrity of medical records and their metadata. Simultaneously, symmetric and asymmetric encryption are an efficient technology in which I use them for enhancing the data confidentiality of medical records in the system. At the same time, the proxy re-encryption function is used to reduce the processing time of decrypting and encrypting the encrypted data before sending it to another peers. So far, I

presented several access control models. They are used to rely on the difference of organization policy and management. In the Chapter 5, I will present about the novel type of using the purpose-based consent, which is the modification of existing purpose-based access control, for the blockchain application.

Chapter 3 Blockchain Technology

In Chapter 2, I presented the general concept of the medical system for the management of the medical records. I also mentioned the fundamental technology for preserving the privacy and enhancing the security in general term of the system, which are not only applied in medical system. In this chapter 3, I explain the overview of the blockchain and its types. From section 3.1 to section 3.4, I present detail of blockchain platforms and its components, which are currently being popular of use. After that, I show the use of blockchain in the medical system, in section 3.5, by reviewing the published research articles of many researchers. I then illustrate prototype designs for exchange, aggregation and traceability of medical data using blockchain. Finally, I show the existing blockchain applications of the consent management for data exchange in medical system.

3.1 Introduction

Blockchain [38] is a technology that enables immutability, and integrity of data. It is a bunch of blocks, which a block links [88] with the next block by the total hash value of a previous block stored in the next block as a sub-value. It has security to protect its transactions by using cryptographic primitives such as hash function, digital signature, and data encryption [89]. Blockchain technology has the consensus protocol employed in the implementation for adding a new block to the existing ledger. Due to the properties of the hash function, nodes can easily verify the integrity of the content of the other node because the Blockchain is replicated and maintained by every participant. With this decentralized approach there is no need for setting up a single trusted centralized entity for managing the registry.

3.2 Blockchain Types

Blockchain [38] is divided into three types such as permissionless or public, consortium and private Blockchain for using in different system's use case. The distinction among these types of Blockchains is the scheme of ledger sharing and the policy of allowing participates to access in a system [90].

Public Blockchains [91] like Bitcoin, Ethereum, etc. allow anyone to access and maintain the distributed ledger with permissions to validate the integrity of the ledger by running consensus mechanism. A public Blockchain network is completely opened and distributed; anyone can freely join, participate, and leave the system. Therefore, this system operates under unknown and untrusted nodes.

In private Blockchain, ledgers are shared in and validated by a predefined group of nodes. The system requires initiation or validation to nodes that want to be part of the system. Authorized nodes are responsible for maintaining consensus. Private Blockchain is suitable for closed systems, where all nodes are fully trusted. The owner has highest authority to control access to authorized nodes.

On the other hand, consortium Blockchains are hybrid between private and public Blockchains by incorporating many parties and the main nodes are initially and strictly selected. Consortium Blockchain is suitable for semi-closed systems consisting of a few enterprises, often organized in the form of a consortium. The degree of data openness varies, usually involving access controls, defined by the consortium, to control access in both participants and information inside Blockchains. Even though the system is not fully opened, the benefits of decentralization can be partially gained. For example, the system has some degree of false tolerance in the event of

some nodes acting maliciously. Hyperledger Fabric [92], Ripple [93] and Stellar [94] are examples of consortium Blockchain implementations.

Although they are different in setting, all types of Blockchain share the following similarities regarding the benefits that Blockchain technology provides. (1) They operate on Peer-to-Peer (P2P) network that provides some degree of decentralization, (2) multiple nodes maintain the integrity of the ledger through consensus mechanisms, and (3) data are stored in Blockchain which provides immutability, even when some nodes are faulty or malicious [95].

3.3 Distributed Consensus Algorithm

Distributed consensus mechanism is

"critical for blockchain since it determines which block can be accepted and inserted to the chain."

It is similar to the agreement on distributed power allocation because the node authoring the accepted block (hereafter referring to as the official validator) is able to change the state of the database shared by every other peer. Many blockchain platforms use different consensus mechanism for solving the various alternative agreements in the network. Following are the consensus algorithms of the blockchain technology.

3.3.1 Proof-of-Work (PoW)

Proof-of-work (PoW) is the original concept of Bitcoin, in which nodes have to compete by calculating a cryptographically sophisticated puzzle. It requires the power allocation associating with some cost and resources to prevent abuse. The characteristics of this puzzle ensures three properties: a node has to invest corresponding amount of computing power to complete it; the next

node to successfully solve the puzzle is random; and a node's claim on finding the answer of the puzzle can be easily verified by any other peer nodes. One additional issue is, however, malicious nodes controlled by an attacker could also be randomly selected as official validator as long as they follow the same process. Once chosen, a malicious node could still try to inject blocks of false transaction records into the blockchain. Therefore, there is a follow-up implicit consensus step after a peer node receives the block proposed by the official validator. In this step, the peer nodes can verify the transactions in the received new block, and if any anomaly is detected in it (such as inconsistency of the linked hash values, or mismatched transaction signature and identity), they can keep the prior state of the blockchain without accepting the new block. Otherwise if everything goes well, the node confirms the new block and accepts the updated blockchain. The likelihood of a block being rejected diminishes exponentially with the number of acceptance confirmations it receives from different nodes. After a certain number (e.g., 6 in the case of Bitcoin) of confirmations, the block is considered permanently committed to the blockchain.

3.3.2 Proof-of-Stake (PoS)

Proof of stake (PoS) [96] is a type of consensus algorithm by which a cryptocurrency blockchain network aims to achieve distributed consensus. In PoS-based cryptocurrencies, the creator of the next block is chosen via various combinations of random selection and wealth or age. The PoS algorithm uses a pseudo-random election process to select a node to be the validator of the next block, based on a combination of factors that could include the staking age, randomization, and the node's wealth. In PoS, blocks are said to be *forged* rather than mined. Users who want to participate in the forging process, are required to lock a certain amount of coins into the network as their stake. The size of the stake determines the chances for a node to be selected as the next validator to forge the next block - the bigger the stake, the bigger the chances. In order

for the process not to favour only the wealthiest nodes in the network, more unique methods are added into the selection process. The two most commonly used methods are *Randomized Block Selection* and *Coin Age Selection*.

When a node gets chosen to forge the next block, it will check if the transactions in the block are valid, signs the block and adds it to the blockchain. As a reward, the node receives the transaction fees that are associated with the transactions in the block. If a node wants to stop being a forger, its stake along with the earned rewards will be released after a certain period of time, giving the network time to verify that there are no fraudulent blocks added to the blockchain by the node.

3.3.3 Practical Byzantine Fault Tolerance (PBFT)

Practical Byzantine Fault Tolerance (PBFT) [97] is the feature of a distributed network to reach consensus (agreement on the same value) even when some of the nodes in the network fail to respond or respond with incorrect information. The objective of a PBFT mechanism is to safeguard against the system failures by employing collective decision making (both – correct and faulty nodes) which aims to reduce to influence of the faulty nodes.

PBFT tries to provide a practical Byzantine state machine replication that can work even when malicious nodes are operating in the system. Nodes in a PBFT enabled distributed system are sequentially ordered with one node being the primary (or the leader node) and others referred to as secondary (or the backup nodes). Note here that any eligible node in the system can become the primary by transitioning from secondary to primary (typically, in the case of a primary node failure). The goal is that all honest nodes help in reaching a consensus regarding the state of the system using the majority rule. A Practical Byzantine Fault Tolerant system can function on the

condition that the maximum number of malicious nodes must not be greater than or equal to one-third of all the nodes in the system. As the number of nodes increase, the system becomes more secure.

3.4 Blockchain Platforms

I presented about the type of the blockchain and the consensus algorithm in the distributed system using blockchain technology. In this section, I continue explaining about blockchain platforms, which are widespread of use in society.

3.4.1 Bitcoin

The world's first cryptocurrency, Bitcoin, was established in 2009 following the public release of a paper by Nakamoto [89], an author whose real identity remains unknown. Bitcoin is a decentralized crypto currency and remains the most important blockchain application today. It is believed that the inventor created Bitcoin to offer an alternative to the central-bank controlled monetary system. According to a technical report released by the European Commission [98],

"the major contribution of Bitcoin is the solution of how to establish trust between two mutually unknown and unrelated parties to such extent that sensitive and secure transactions can be performed with full confidence over an open environment, such as the Internet"

Most important, Bitcoin has opened a series of possibilities for blockchain-based innovative applications not only for financial transactions but also for transfer and trading of digital assets, aiming to guarantee safety, security and legitimacy.

A typical blockchain consists of a peer-to-peer network of computer nodes that maintain a decentralized shared database of records. In the original Bitcoin blockchain, the records contain transfer transactions of Bitcoin crypto currency between participating parties. Each party in the transaction has a Public Key Infrastructure (PKI) private key and public key pair. The hash value of the public key is used as the party's identity or transaction address. Transaction parties sign the transactions using their private key, which could later be verified by other parties using the signer's public key. The transactions are broadcast to all peer nodes in the network. Using a distributed consensus mechanism, the peer nodes agree on what transactions are valid and the sequence of those transactions that take place. Each block in the blockchain has its own timestamp and a cryptographic hash that connects it to the prior block. Blocks can only be appended, not deleted. The outcome is a shared database with an ever-growing list of records that are immutable and irreversible; tampering of any block information can be detected by peer nodes on the blockchain.

3.4.2 Ethereum

Ethereum is a distributed virtual machine that can execute smart contracts defined by a developer as a Bitcoin extension protocol [99]. When Ethereum was first developed in 2013, Vitalik Buterin proposed the extension of the Turing-incomplete Bitcoin script to a Turing-complete language capable of handling various types of smart contracts. Turing-completeness [100] facilitates the development of cryptocurrency, smart contracts and other decentralized applications (Dapps) by implementing various blockchain platforms. In Bitcoin, the unspent transaction output (UTXO) is simply expressed as spent/unspent; whereas, Ethereum represents everything as the state transformation of the blockchain or an account. Ethereum's smart contracts consider an account as a normal node. Therefore, a node is able to send coins to a smart contract account and receive coins from that account.

In Ethereum, the currency used is called Ether or Wei. Moreover, Ethereum uses a special programming language called Solidity. This language is a Turing-complete bytecode language called the Ethereum virtual machine bytecode and can use arrays, variables, constructors, etc., in the same manner as any typical Turing-complete programming language. The functions define the specific operation of these elements, and several types of Dapps, such as web sites or smartphone apps, can be developed using Ethereum smart contracts. Compared to Bitcoin, Ethereum has the advantageous features of scalability and flexibility.

$$Transaction\ fee = gas\ price \times gas\ limit \tag{1}$$

Gas is a characteristic element of Ethereum platform. Specifically, gas is to be considered while dealing with the cost of storage space or the computational power required to run smart contract functions, run the code, and so on. All operations are done on the network, such as deploying smart contracts or sending coins, creating transactions. The transaction fee is calculated by (1), and the amount of gas required to execute a code is defined in the Ethereum yellow paper [99]. If the gas price is high, the transaction has a high reward and will probably be mined relatively early. The creator of the transaction can set a gas limit to prevent too much gas from being consumed during the execution of a smart contract code. In Ethereum, the transaction creator node always pays the gas cost.

Smart Contract

A smart contract is a computer protocol intended to digitally facilitate, verify, or enforce the negotiation or performance of a contract. Unlike traditional contracts [101] that rely on the reputation of the counterparties, smart contracts can be made between untrusted, anonymous people. Also, the execution of contractual terms is automatic and does not rely on any third party.

The concept of smart contracts was introduced in 1990 by Wei Dai [102]. However, smart contracts were not possible in many traditional systems since the participating parties use to maintain separate databases and they do not rely on a proper trust model. However, the possibility to develop a trusted and shared database based on a blockchain has eliminated this limitation. A blockchain-based smart contract is a self-executing code on a blockchain that automatically implements the terms of an agreement between parties [103]. Blockchain-based smart contracts could offer a number of benefits, such as fast, dynamic and real-time updates, low cost of operation, high accuracy and fewer intermediaries [101], [103]. These benefits also fuel the adaptation of a smart-contract in different applications. Thus, blockchain-based smart contracts are getting a significant interest across a wide range of industries.

3.4.3 Hyperledger Fabric

Hyperledger Fabric [92] is an open-source framework for private, or consortium, blockchains developed initially by the Linux Foundation, and later supported by companies such as IBM. Hyperledger Fabric consists of a blockchain and State DB [92] (a.k.a World State). The former is a transaction log, while the latter holds current values of ledger states. Due to the State DB, program get readily values without traversing the entire transaction log. Transactions [104]–[106] are collected to form a block that is appended sequentially to the last block of the blockchain, which is immutable once made.

3.4.3.1 Membership Service Provider and Enrolment Certificate

In order to facilitate of the identification nodes in the blockchain system, Hyperledger Fabric allows consortium to group all nodes in the same channel for communication. In addition, it adapts the Membership Service Provider (MSP) for aiming to abstract all cryptographic mechanisms

and protocols behind issuing and validating certificates, and user authentication. There are two types of MSP; channel and local one. Channel MSP provides a method to validate Enrolment Certificate (ECert) among different organizations in the channel, while local MSP offers method

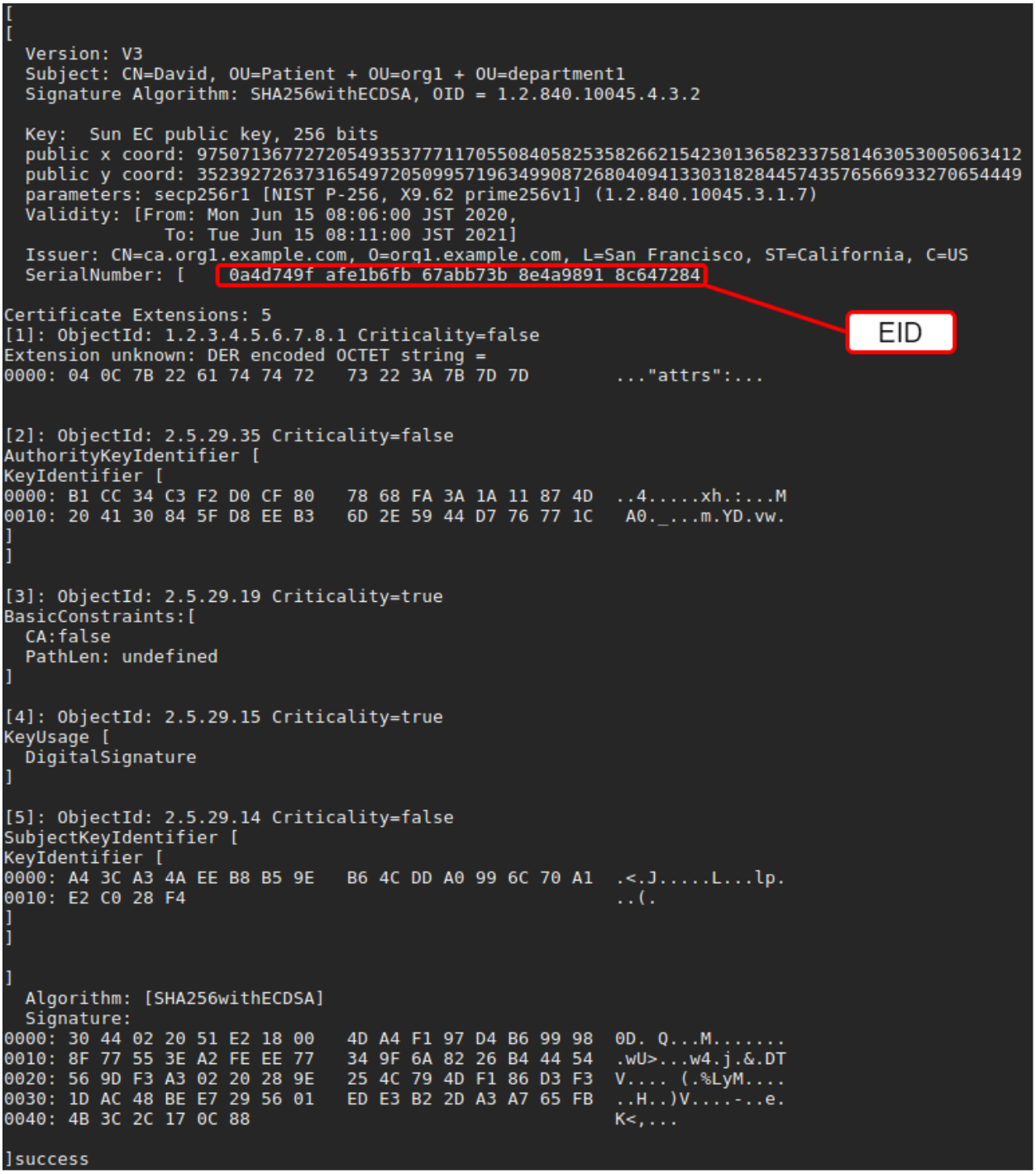


Figure 13: Enrolment Certificate

to verify user's identity in one organization. Thus, each organization has their local MSP having unique MSP ID and they issue ECert, an X.509 certificate as showed in Figure 13, to all the local participants with Enrolment IDs (EID) through their Certificate Authority (CA).

3.4.3.2 User role in Hyperledger Fabric

In Hyperledger Fabric, there are three main types of user roles; client, peer (endorsing and committing one), and Orderer. Peer is a network node, and endorsing peers, simply called endorsers, proceed endorsement with simulating client's Transaction Proposal. The Proposal is a tentative transaction before being accepted into a new block of the ledger. Orderer runs ordering service for creating new blocks with transactions, then broadcasts the block to all peers. Committing peer, also called as Committer, updates the ledger by appending the new block to it and revising the Sate DB with the write-sets of valid transactions. As shown in the Figure 14, it is the flow of HLF for appending a block to the blockchain by starting from the proposal creation

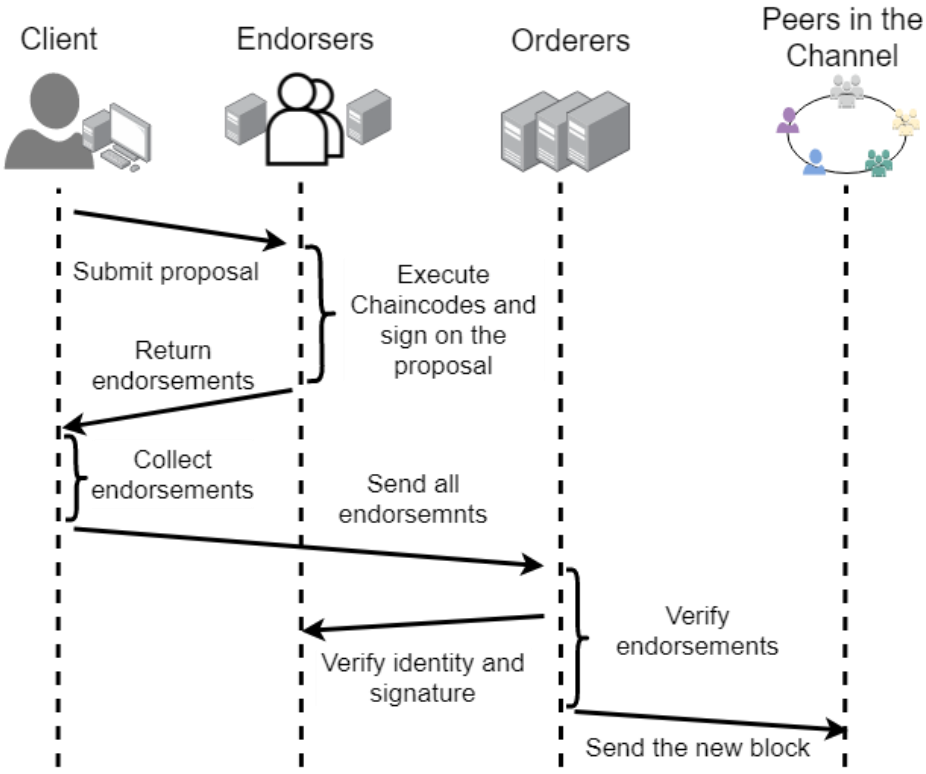


Figure 14: Flow of adding transactions to the blockchain

from clients and send it to endorsers for endorsing that proposal until Orderer validate the endorsements and broadcast a block that consists of the valid transaction to all peer in the channel network.

3.4.3.3 Chaincode

Like Ethereum, Hyperledger Fabric adopts the concept of smart contracts, or chaincode. It is an application program run by peers to facilitate, verify or enforce negotiation and agreement between users. Chaincode is otherwise known as smart contract in other blockchain platforms like Ethereum. Chaincode [107] has many programming functions in it and usually reads and updates the ledger state with all the business logic contained inside functions.

3.4.3.4 Endorsement Policies

Every chaincode has an endorsement policy which specifies the set of peers on a channel that must execute chaincode and endorse the execution results in order for the transaction to be considered valid. These endorsement policies define the organizations (through their peers) who must "endorse" (i.e., approve of) the execution of a proposal. As part of the transaction validation step performed by the peers, each validating peer checks to make sure that the transaction contains the appropriate number of endorsements and that they are from the expected sources (both of these are specified in the endorsement policy). The endorsements are also checked to make sure they're valid (i.e., that they are valid signatures from valid certificates). For instance, in Figure 15, it has two endorsers from each organization in the channel, so in total is four endorsers. They work to endorse the proposal sent by client. In the rule of this endorsement policy, it required at least one endorser from either organization one or organization two to execute the chaincode and endorse the client's proposal.

```
identities:  # list roles to be used in the policy
    user1: {"role": {"name": "member", "mspId": "Org1MSP"}} # role member in org with mspid Org1MSP
    user2: {"role": {"name": "member", "mspId": "Org2MSP"}}
    admin1: {"role": {"name": "admin", "mspId": "Org1MSP"}} # admin role.
    admin2: {"role": {"name": "admin", "mspId": "Org2MSP"}}

policy: # the policy  .. could have been flat but show grouping.
    1-of: # signed by one of these groups  can be <n>-of  where <n> is any digit 2-of, 3-of etc..
      - 1-of:
        - signed-by: "user1" # a reference to one of the identities defined above.
        - signed-by: "admin1"
      - 1-of:
        - signed-by: "user2"
        - signed-by: "admin2"
```

Figure 15: Endorsement policy

3.5 Literature Reviews of Blockchain application

I presented the blockchain's concept and its platform. In this section, I show existing research use cases in which its concepts are to apply the blockchain technology to the healthcare system. In addition, I present the state of the art of their researches, which I can use them as the fundamental of concept for designing my blockchain system.

3.5.1 Blockchain Use Case: Medical System

Fran et al. [108] presented the survey research paper related to the Blockchain application. They showed a comprehensive and in-depth classification of Blockchain-based applications. Among those applications, Blockchain technology could play a pivotal role in the healthcare industry with several applications in areas like public healthcare management, longitudinal healthcare records, automated health claims adjudication, online patient access, sharing patients' medical data, user-oriented medical research, drug counterfeiting, clinical trial, and precision medicine. They also stressed that the benefits of a Blockchain-based system for EHRs are manifold: records are stored in a distributed way (they are public and easily verifiable across non-

affiliated provider organizations), there is no centralized owner or hub for a hacker to corruptor breach, data is updated and always available whereas data from disparate sources is brought together in a single and unified data repository.

The review paper by Kuo et al. [106] presented several Blockchain applications in healthcare, such as improved medical record management and advanced healthcare data ledger, and their benefits for each described application. They then analyzed key challenges associated with using Blockchain technology for healthcare, including issues like confidentiality, scalability, and treat of a 51% attack [109] on the Blockchain network. According to the authors, some example implementation techniques that may mitigate the challenges are (i) encryption of sensitive data or dissemination of only meta data and storing sensitive data off-chain to protect confidentiality, (ii) keeping only partial, ongoing verified transactions on-chain rather than the entire transaction history to increase scalability of the Blockchain network, and (iii) the adoption of a virtual private network or HIPAA-compliant components.

To sum up with above arguments, one of the most prominent applications of blockchain technology is healthcare. The potential of blockchain in healthcare is to overcome the challenges related to data security, privacy, sharing and storage [110]. one of the requirements for the healthcare industry is interoperability. It is the ability of two parties, either human or machine, to exchange data or information precisely, efficiently, and consistently [111]–[114]. The goal of interoperability in healthcare is to facilitate the exchange of health-related information, such as electronic health record (EHR), among healthcare providers and patients so that the data can be shared throughout the environment and distributed by different hospital systems [104], [115]–[117]. Moreover, interoperability enables providers to securely share patient medical records (given patient permissions to do so), regardless of the provider's location and trust relationships

between them [118]. This is specifically important considering that the source of healthcare data is diverse. This aspect of interoperability is resolved by using blockchain technology which showed potential to store, manage, and share EHRs safely amongst healthcare communities [108]. Additionally, increasing costs of healthcare infrastructures and software in the industry have caused tremendous pressure on world economies [22]. In the healthcare sector, blockchain technology is positively affecting healthcare outcomes of companies and stakeholders to optimize business processes, improve patient outcomes, patient data management, enhance compliance, lower costs, and enable better use of healthcare-related data [119].

3.5.2 System Designs for Exchange, Aggregation and Traceability of Medical Data Using Blockchain

Ekblaw et al. [120] created a decentralized record management platform that enables patients to access their medical history in EMR. This platform used a so-called “consortium” blockchain (which is only accessible by authorized users, rather than one that is open to the public) to manage authentication, data sharing, and other security properties in the medical domain. Their blockchain design integrated with existing provider data storage to enable interoperability by curating a representation of patient medical records. Medical researchers were incentivized to contribute to mining of the blockchain by collecting aggregated metadata as mining rewards.

Similarly to the above research, Gaby et al. [104] created the consortium blockchain framework, called Ancile, for EHR management that could give ownership and final control of EHRs to the patient that they can access documents and track how records are used, allow for secure transfer of records. For the testing of the prototype system, they proposed different smart contracts. They suggested the interesting idea for sharing patient records is to apply the distributed

proxy re-encryption on the encrypted medical records. In addition, they use the Quorum consensus mechanism, which is different from the build-in consensus of Ethereum is PoW.

Peterson et al. [36] presented a healthcare blockchain also considers the integration with FHIR standards. They proposed a Merkle-tree based blockchain system that introduces “Proof of Interoperability” as the consensus mechanism during block mining. Proof of interoperability is based on conformance to the FHIR protocol, meaning that miners must verify the clinical messages sent to their blockchain to ensure they are interoperable with known structural and semantic standards.

Dubovitskaya et al. [121] also proposed a consortium blockchain framework on managing and sharing medical records for cancer patient care. Their design employed a membership service of Hyperledger Fabric to authenticate registered users using a username/password scheme. Patient identity was pseudonymized by combining of personally identifying information (including social security number, date of birth, names, and zip code) and encrypted for security. Medical data files were uploaded to a secure cloud server, with their access managed by the blockchain logic.

Unlike other blockchain designs, Gropper's “HIE of One” system Gropper [122] focused on the creation and use of blockchain-based identities to credential physicians and address the patient matching challenge facing health IT systems. Patients are expected to install a digital wallet on their personal devices to create their blockchain-based IDs, which can then be used to communicate with the rest of the network. Instead of storing patient information, Gropper's system would consume only the blockchain-based ID and use it to secure and manage access to patient data located in EHR systems.

To sum up, consortium blockchain is played for the important role of the interoperability of EHR. Although consortium systems may be prone to collusion due to the 51% attack problem [123], the consortium system used for healthcare would be maintained and managed by relatively large-scale entities/stakeholders within the healthcare industry. Unless majority of them (major hospitals, insurance companies, etc.) collude, therefore, the chance of experiencing this type of attack is quite low. Moreover, legal actions would most likely occur immediately upon the attack.

3.5.3 Blockchain-Based Patient Consent for Medical Data Exchange

Recently, there have been a research project is Dwarna [124], which is a web portal for 'dynamic consent' that acts as a hub connecting the different stakeholders of the Malta Biobank: biobank managers, researchers, research partners, and the general public. This portal stores research partners' consent in a blockchain to create an immutable audit trail of research partners' consent changes. Their blockchain's structure also presents a solution to the European Union's General Data Protection Regulation's right to erasure—a right that is seemingly incompatible with the blockchain model. Its transparent structure increases trustworthiness in the biobanking process by giving research partners more control over which research studies they participate in, by facilitating the withdrawal of consent and by making it possible to request that the biospecimen and associated data are destroyed.

3.6 Summary

I learnt the concept of blockchain and its applications. Blockchain technology is redefining data modelling and governance deployed in many healthcare applications. This is mainly due to its adaptability and abilities to segment, secure and share medical data and services in an unprecedented way. Since the medical records are very sensitive, the consortium blockchain is

playing in the important role for controlling the identity of members in the network. In addition, I can design the blockchain system for preserving the patient's privacy of the accessing by doctors or other stakeholders to medical records based on the patient's consent. In consequence, HLF is very suitable for building a good information system for medical information sharing because its policy divides different role of users and the processing time of adding the new block is faster than Bitcoin and Ethereum.

Chapter 4 Proposed System Model for Integrating EHRs

In Chapter 3, I presented the overview of blockchain technology and its applications. Blockchain technology is not only applied to solve the problems in Fintech but also in many domains such as medical system. I also show the selected blockchain platform, which is suitable for building the system for integrating EHRs. In this Chapter 4, I explain my contribution of the proposed system model adapting blockchain technology with the existing security concept for medical information sharing.

4.1 System Model

I build a private subnet of HLF network where the same ledger is shared among the hospital members (Figure 17), which is called a channel. Organizations or departments within them can constitute independent channels with relevant ledgers according to their needs. In practice, medical data is usually too big to handle directly in ledger thus kept in EHR, and only the address is recorded in the ledger. Such storage type is called as on-chain or off-chain according to whether the data is on the ledger or not [106]. Ledger also contains hash value of the data. This guarantees data integrity because once written in the ledger it becomes immutable and allows the user to check if the data is altered or not. As showed in Figure 16, an Ophthalmologist uploaded the encrypted medical data to the EHR database and its address is stored in the blockchain with the records metadata. The purpose of separating of data storing technique is because of the blockchain is

shared among members in the network so that if I store the large records in the blockchain, it is difficult for every member to keep the blockchain for long time.

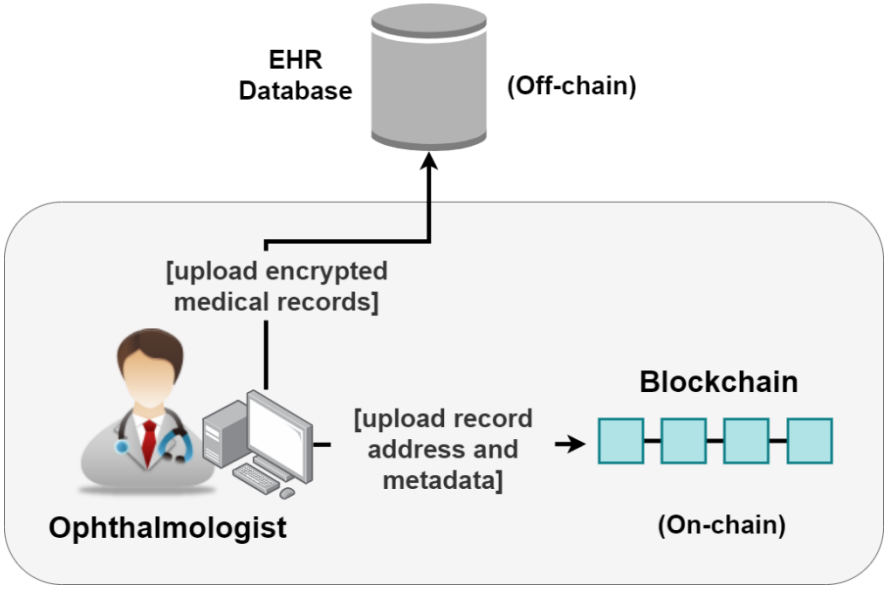


Figure 16: Off-chain and on-chain data storing

In my system, I assume client of HLF as doctor, nurse or clerk who helps patients to upload or share their medical records. Clients from medical institutions issue varieties of transactions and store them in the ledger. The Ledger consists patient metadata including demographics, and these data is used for retrieval request to find out transactions related with the patient during a specified period of timestamps of blocks in the ledger. Thus, the ledger plays a role as a registry of patient data 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 etc. In consequence, these data will be used for auditing purpose.

For individual patient, Enrolment ID (EID) issued by Membership Service Provider (MSP) is used as Channel Patient ID in the system. Each transaction in the ledger contains EID, which is

hashed after being concatenated with a random data so called salt [46] in the format as shown below.

$**n **$salt **$**hash (salt + EID)

This format is nearly the same with how Linux system stores its user's hashed passwords with salts. '$' is used as a delimiter between neighboring fields. 'n' represents hash algorithm type; 1, 5 and 6 correspond to MD5, SHA-256 and SHA-512 respectively. Salt is a string of random alphanumeric characters up to 16 letters.

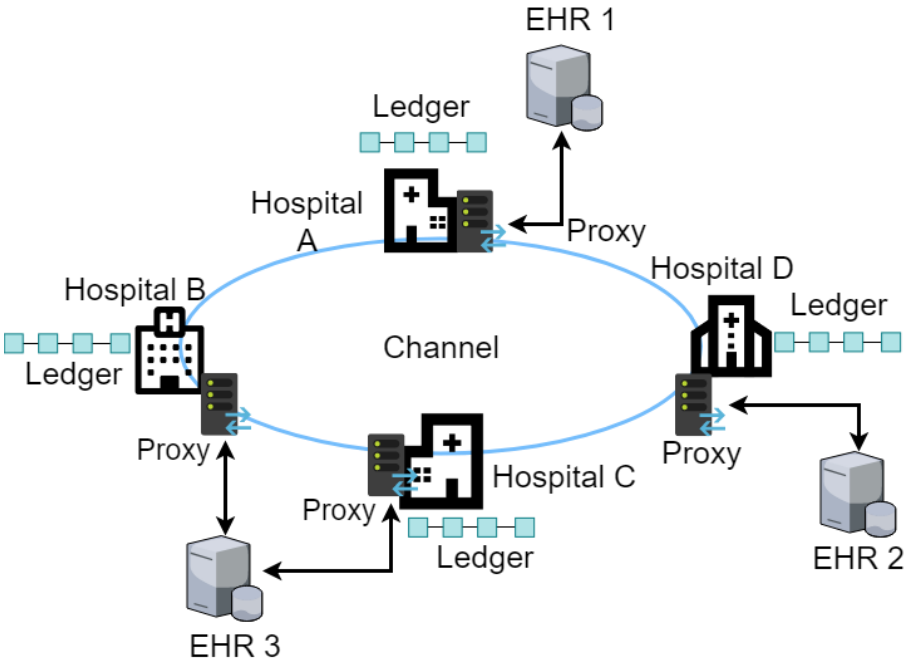


Figure 17: Design system

- **Proxy re-encryption**

i. What is proxy?

In computer networking, a proxy server [125] is a server application or appliance that acts as an intermediary for requests from clients seeking resources from servers that provide those resources. In my system, I suppose that each organization has their own proxy, which is used to handle the communication between organizations because each organization connects to different

EHRs and to allow staffs of the organizations, which are not the member of EHRs for accessing to data, it is the heavy task for them. Thus, the proxy of the organization is registered to the EHRs as a member, and if staffs of the organization request for retrieving the data based on the proxy installing in their organization. Moreover, the proxy executes re-encrypting function on the cipher text before sending it to requestors. In the next section, it presents the purpose of using the re-encryption function in the proxy.

ii. Proxy re-encryption

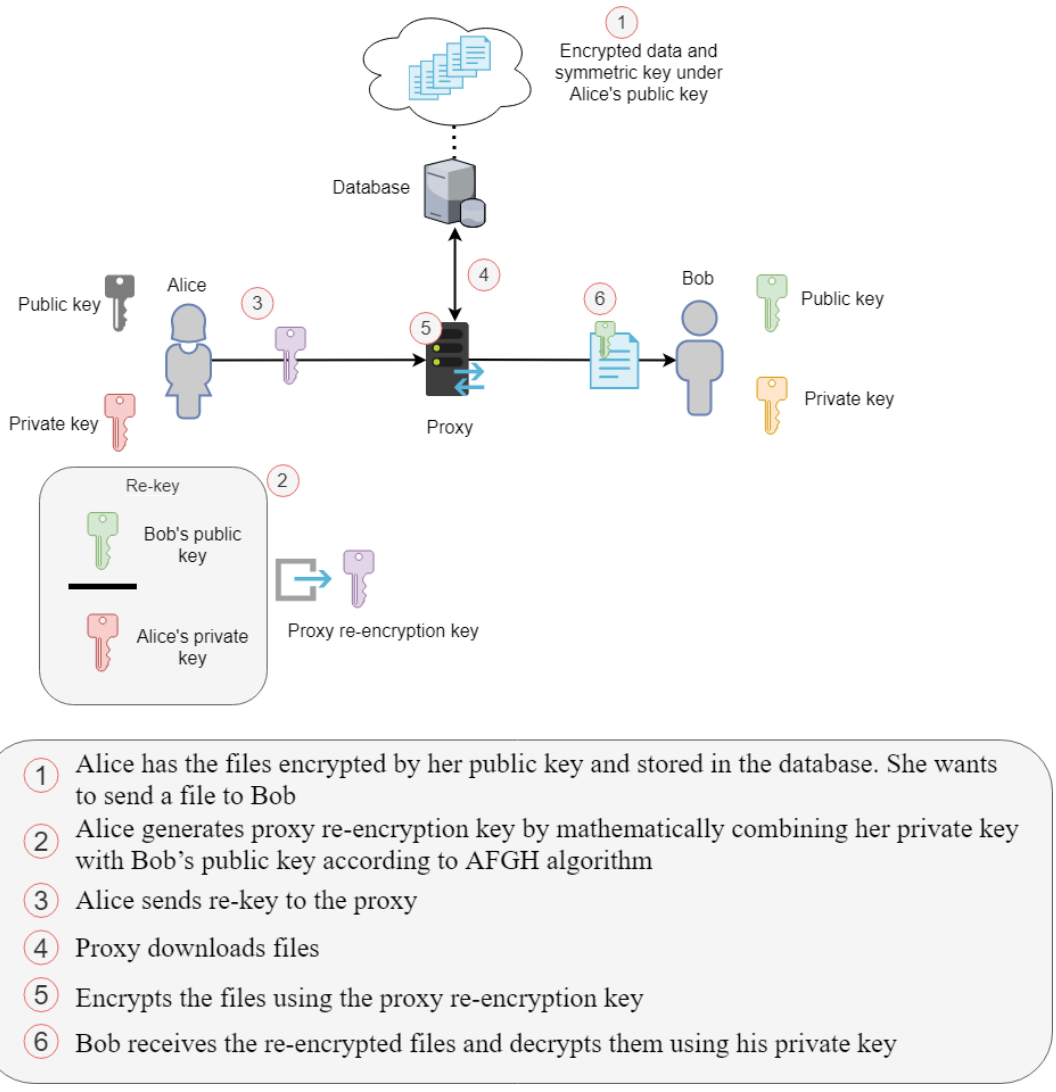


Figure 18: The proxy re-encryption scheme

Nowadays, Proxy Re-Encryption (PRE) [126] has been widely recognized as a promising solution to achieve secure cloud storage. Technically, PRE allows users to store their encrypted data in cloud and securely share their data with others even if cloud is unauthorized for accessing. In the PRE-based cloud, each user usually encrypts his data by his own public key and stores the resulted ciphertexts in this cloud. Suppose user Alice (Figure 18) wants to share her cloud data with user Bob. Alice delegates a re-encryption key to cloud; then cloud re-encrypts Alice's ciphertexts, and sends the re-encrypted ciphertexts to Bob, and finally, Bob decrypts these received ciphertexts by his own private key. Moreover, a secure PRE scheme guarantees that cloud cannot learn anything about Alice's data and share Alice's data to any unauthorized user. In Figure 19 below, it illustrates the result after applying AFGH algorithm of proxy re-encryption run by proxy. As the result, it does not change the plaintext.

```
Public key of User A is : 893169269892986527081175915135986601881400309288654874980369235125821600585546592
Secret key of User A is : 396942055552419626531130886309841956376679789733979843351250661370338690696609070
Public key of User B is : 121702200226706664342587750762365930425430833438632218661170587106535649834355839
Secret key of User B is : 663101575653509749491164360499150140812986593532136016040815050182978245835345306
Re-encryption key is : 317386336971429655980118391661317090678171496430158309414859578656345842729748789447
Message is: bxV3Jgb5Fh215Whbx1aReSQhSDNpavqIQ==
Message is encrypted by Public key A and published to the EHR Server : <552405453456400292051882106915512370
Proxy re-encryption the A's ciphertext <{x=290517948035457141902914983108022772246583603930470531651312262
User B decrypts by his secret key and gets message : bxV3Jgb5Fh215Whbx1aReSQhSDNpavqIQ==
```

Figure 19: Proxy re-encryption

4.2 Implementation

I explained the proposed system model and the technique to use for solving the integrating of EHRs for data sharing. In this section, I present the proposed chaincodes, explain the use case scenarios and analysis of this prototype system. I develop the prototype system using Eclipse Integrated Development Environment (IDE) for running the user interface and blockchain system (Backend side).

4.2.1 Chaincode for Exchanging Medical data

In my prototype system, I installed five chaincodes with which business logics are performed. Each chaincode has many programming functions in it and they usually read and update the ledger state with all the business logic contained inside functions. In actual system, each chaincode needs to get agreement among all the member hospitals before being deployed in the system. Following are detail of proposed chaincodes.

- **Record Manager Chaincode:** is the core Chaincode of the system, which get involved in other Chaincodes' execution, to simulate Transaction Proposal for validation and endorse the Proposal. This Chaincode helps client in preparing, uploading and sharing the patient's records.
- **Patient Identity Chaincode:** is called by Clients to register and query patient's identity from the ledger. Patients can find list of identity transactions containing their previous hospital visits. In addition, if patients lose their Ecert, they can provide identifiable attributes to Client for searching and recovering it. Hash value of EID and demographics can be stored into the ledger for identifying patients. Since patients would be given different patient ids from the hospitals they visited, this Chaincode also store and query for the patient ids based on the EID.
- **Permission Manager Chaincode:** works to authorize a third party's access to patient records based on patient consent. Patient consent contains a list of EIDs who are permitted to access, or conditions of comprehensive prior consent, which patient put in the transaction as metadata when their data are recorded in the ledger. For

instance, a patient can share a specific part of their records with an insurance company who is also a participant of the network by putting its EID in the transaction.

- **Personal Folder Chaincode:** helps doctors to collect all of the patient's transactions. It provides special query functions for searching for the transactions based on the multiple keywords such as the hash value of EID with Salt, hospital ID or timestamp.
- **Audition Chaincode:** for designated peers to audit on access histories of patient records by analysing the access log in the ledger. Thus, patient can realize how their data traversed among medical institutions and monitor whether each data transfer was made adequately in compliance with their consent. This Chaincode can also produce statistics based on the doctor's activities, timestamps of transactions, and patient meta data with demographics.

4.2.2 Use Case Scenario

I simulated use cases using the prototype system. In the following figures, which describe a practical situation, I assume that a patient, let's call him Alice, visits to Hospital_A for the first time. Alice is diagnosed with a cancer and his doctor, Dr. Bob, recommends him to go to the central hospital to see a cancer specialist. So, Dr. Bob uploads Alice's records with his consent to the hospital's EHR. Then, Alice moves to the central hospital and the cancer doctor access to Alice's data in the EHR which belongs to Hospital_A.

i. First visit to a hospital

Alice makes a first visit to Hospital_A (Figure 21). To enrol in the hospital, he provides his demographic information or the national insurance number to a clerk. This information will be used for registering him into the patient identity source of the hospital and issuing an ECert for

him. The ECert and private key need to be stored in a secure storage device, for instance, IC card or USB memory. After issuing the ECert by local CA, the clerk needs to store the hash value of Alice's EID and individual patient id in the ledger. In Figure 20, it is the endorsement of the proposal of adding Alice's EID to the blockchain from endorsers from different organizations.

```
Successful transaction proposal response Txid: 0eaa9982c4a5fd982b82d111ba9ed7799086a1f66fe405a3dbde3b611ad490bb from peer peer0.
Successful transaction proposal response Txid: 0eaa9982c4a5fd982b82d111ba9ed7799086a1f66fe405a3dbde3b611ad490bb from peer peer1.
Successful transaction proposal response Txid: 0eaa9982c4a5fd982b82d111ba9ed7799086a1f66fe405a3dbde3b611ad490bb from peer peer1.
Successful transaction proposal response Txid: 0eaa9982c4a5fd982b82d111ba9ed7799086a1f66fe405a3dbde3b611ad490bb from peer peer0.
Received 4 transaction proposal responses. Successful+verified: 4 . Failed: 0
```

Figure 20: Endorsement from endorsers

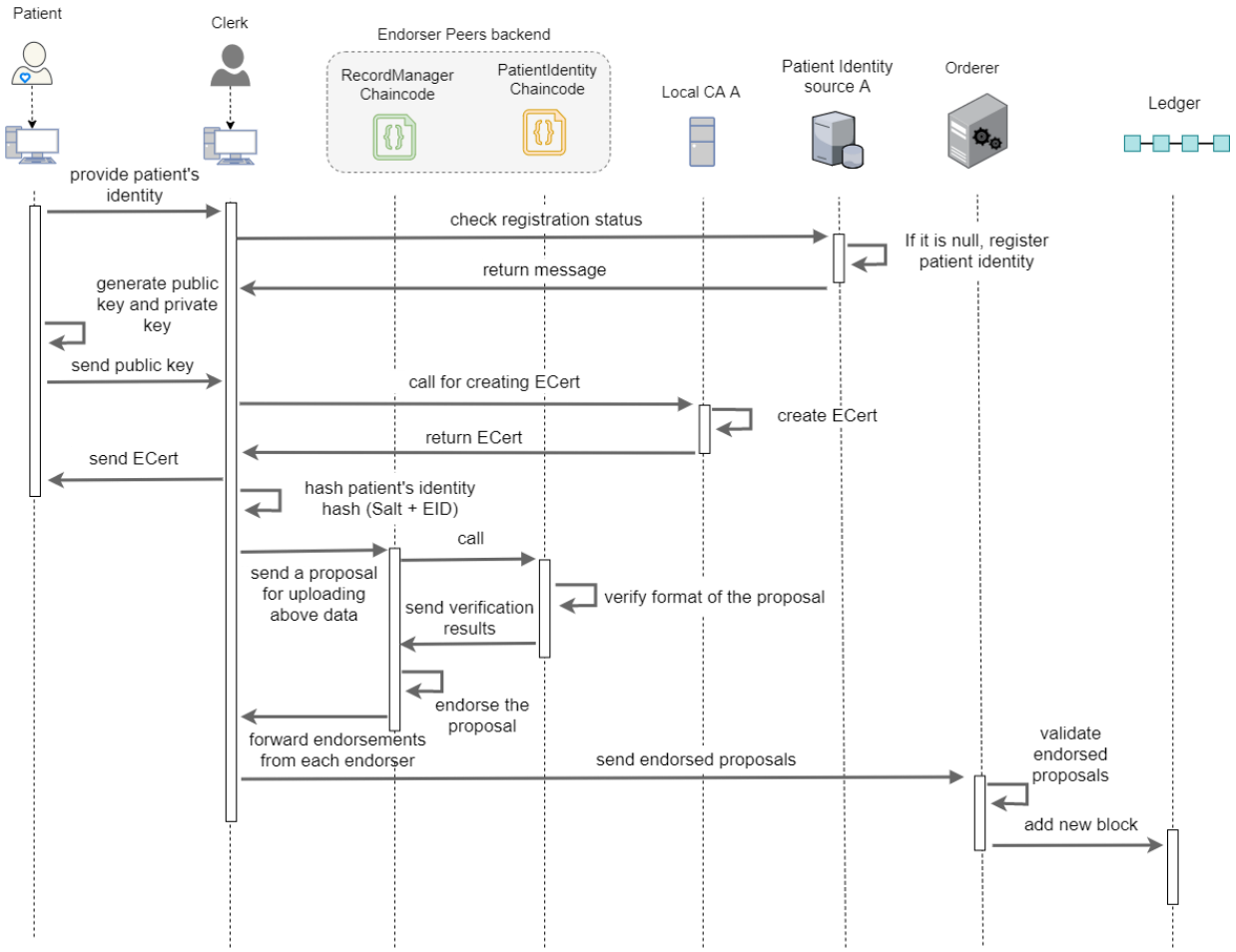


Figure 21: First visit to a hospital

ii. Uploading patient's record with metadata and consent

When a patient's records are uploaded to EHR (Figure 22), Alice provides the doctor his consent with conditions for sharing his records with other third party or his relatives. Then, the doctor

encrypts Alice's record using an adequate symmetric key, and encrypts the key this time using Alice's public key to attach it with the record. Finally, the doctor uploads Alice's record to Hospital_A's EHR and writes the record's consent and address of data location to the ledger.

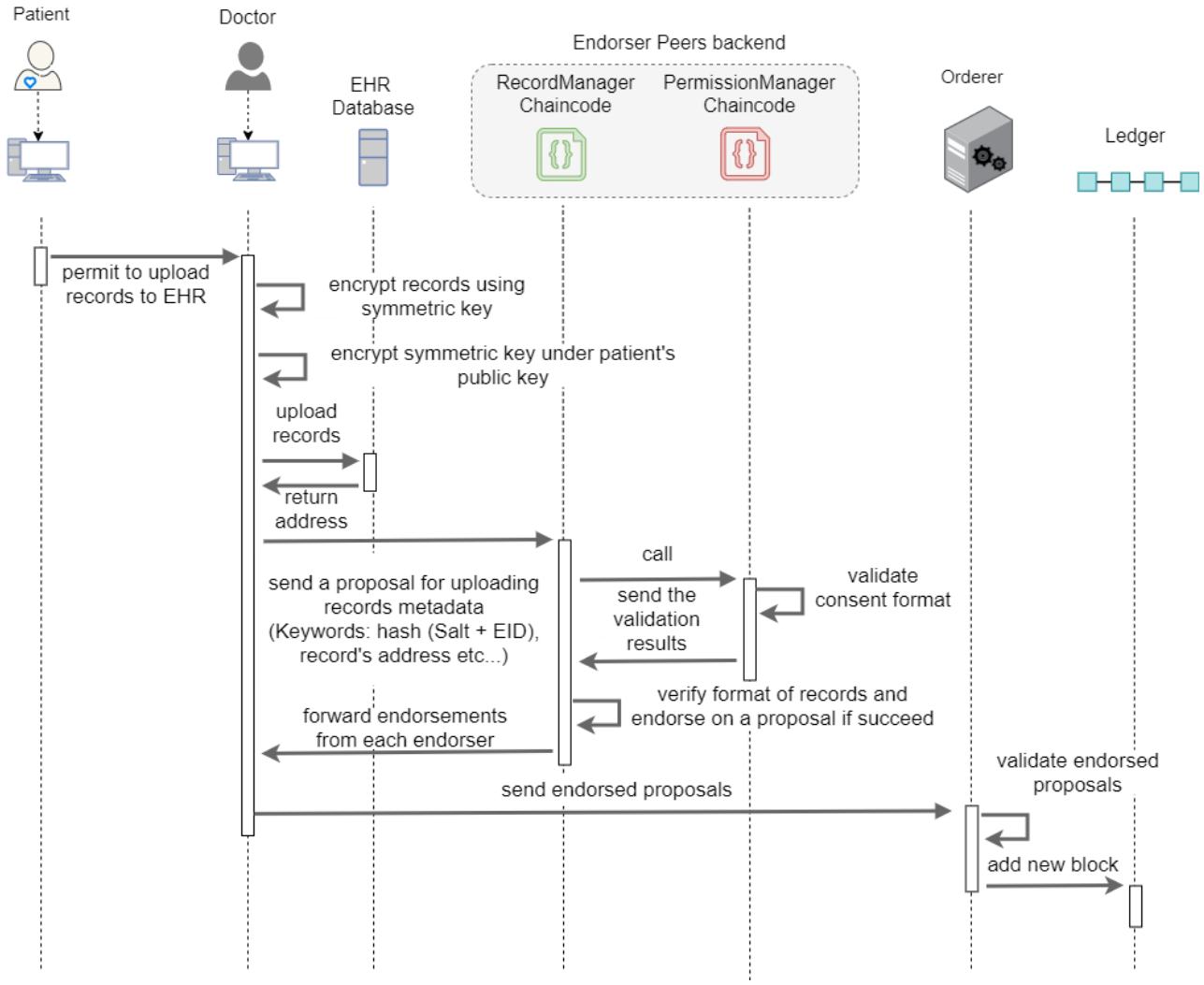


Figure 22: Uploading record with metadata and consent

iii. Requesting patient's record

Alice goes to see a specialist doctor in the central hospital (Figure 26), where she registers as a new patient, if needed, and suggests her ECert previously issued in Hospital_A. When treating Alice, the doctor wants to get Alice's previous records, so he sends a transaction proposal of request for getting Alice's records metadata during a certain period and previous hospital's id.

Then, each endorsing peer simulates the transaction proposal executing Chaincodes, and returns each result of the Chaincode to the proxy of the hospital where the client application run by the doctor. The application compares the query results and, if they are all matched, and lets the doctor select necessary records from them to make a list of patient's records, which he wants to get. For

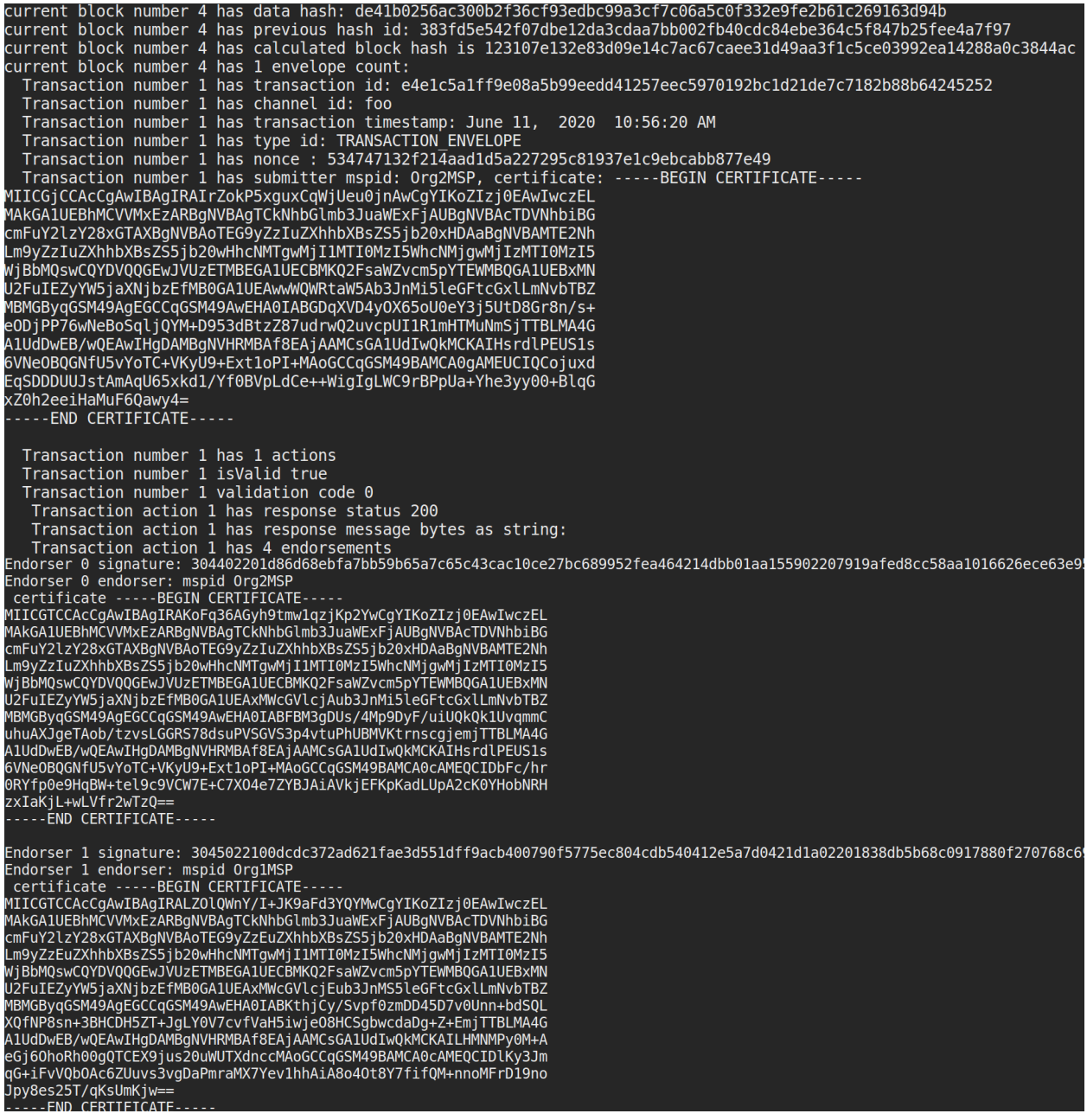


Figure 23: Query data from blockchain

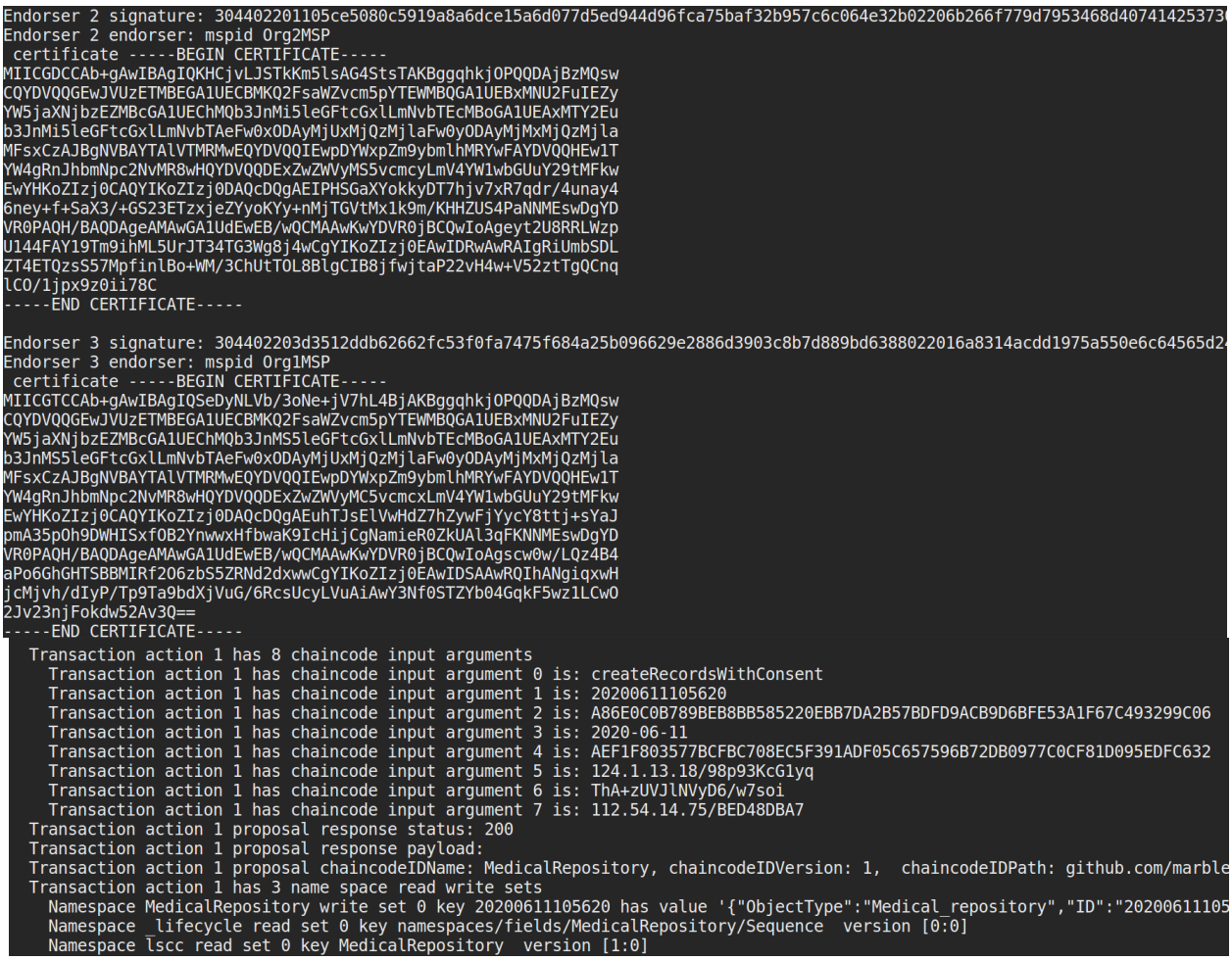


Figure 24: Query data from blockchain (Cont.)

instance, in Figure 25, they are the query results from endorsers such as records address and metadata.

Moreover, doctors can check the integrity of data based on the endorser's identity and their signature, and hash value of medical data, as showed in Figure 23 and Figure 24 . After receiving the list, the proxy asks Alice to generate the proxy re-encryption key. Then, the proxy downloads Alice's records in the list from relevant EHRs and re-encrypts every encrypted symmetric key at each record using the re-encryption. After that, the proxy sends Alice's records to the doctor.

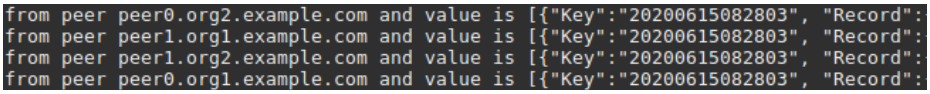


Figure 25: Query results from Endorsers

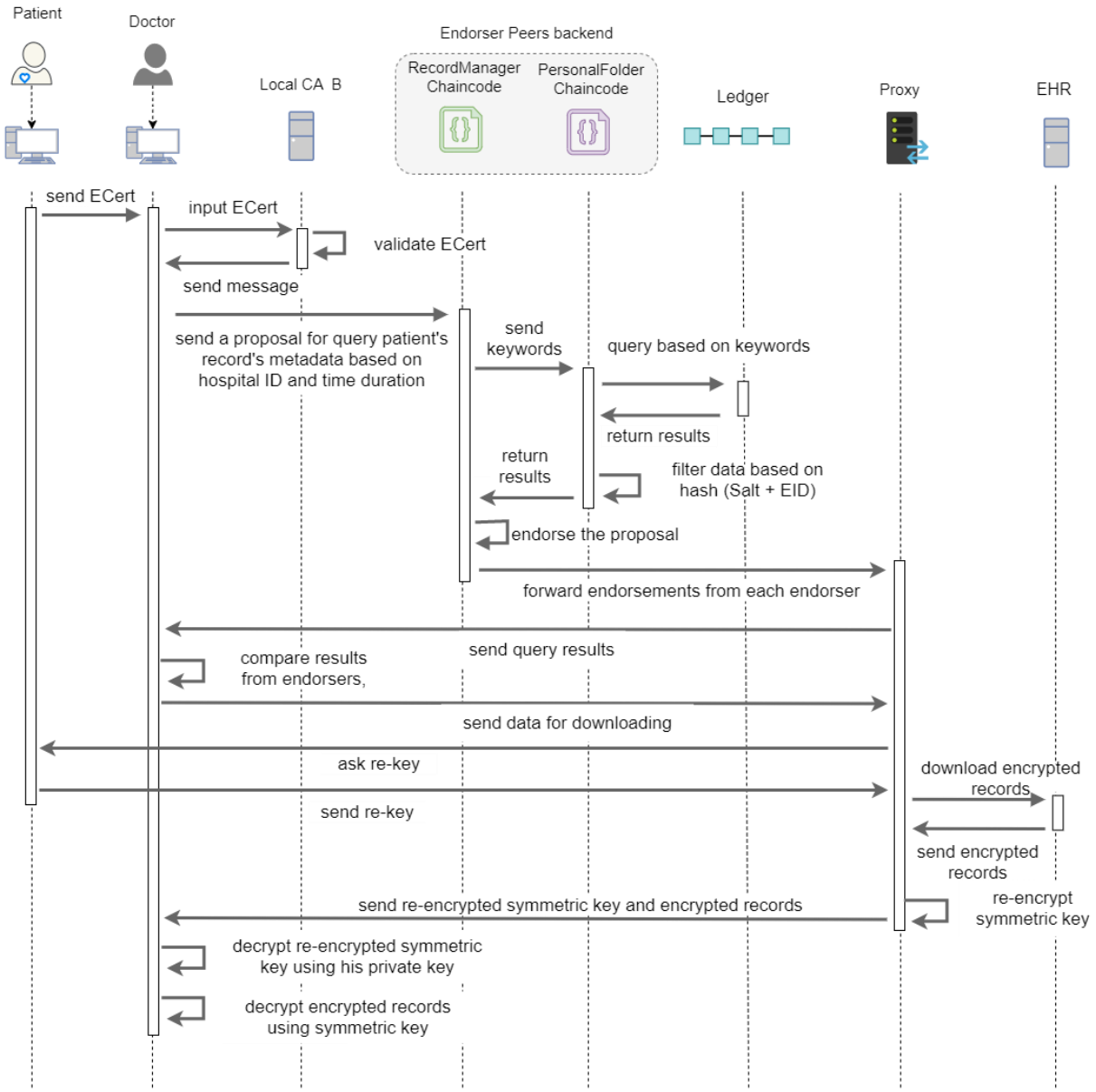


Figure 26: Requesting patients records

4.2.3 Prototype Analysis

A prototype system was built in a small scale for testing on a local network with four Window PCs for patients to use patient web application, four Linux PCs for doctors to use doctor web application and two proxies for two hospitals as shown in Figure 27. In addition, I used two Window PCs as EHRs. Hyperledger Fabric platform runs on Docker [127] for executing chaincodes. For EHR records, I deal with the standardized data such as HL7/CDA and DICOM

image data. I changed the system configuration with different number of PCs to conform the performance including chaincode logic. As a result, it took several second to succeed the request of uploading and querying data in the blockchain. In addition, it took time a little bit more with increasing number of PCs when querying data in blockchain, encrypting and decrypting the records and transferring files. It is because of the file size of the medical data and the number for transferring to other party. In addition, it is also depended on the length of the blockchain and complexity of keyword searching in the doctor's request. It does not matter if two or more client submit the requests at the same time. But all of proposals need to be in the queue until the other proposal is succeeded. For improving the performance, consortium of the network has to increase

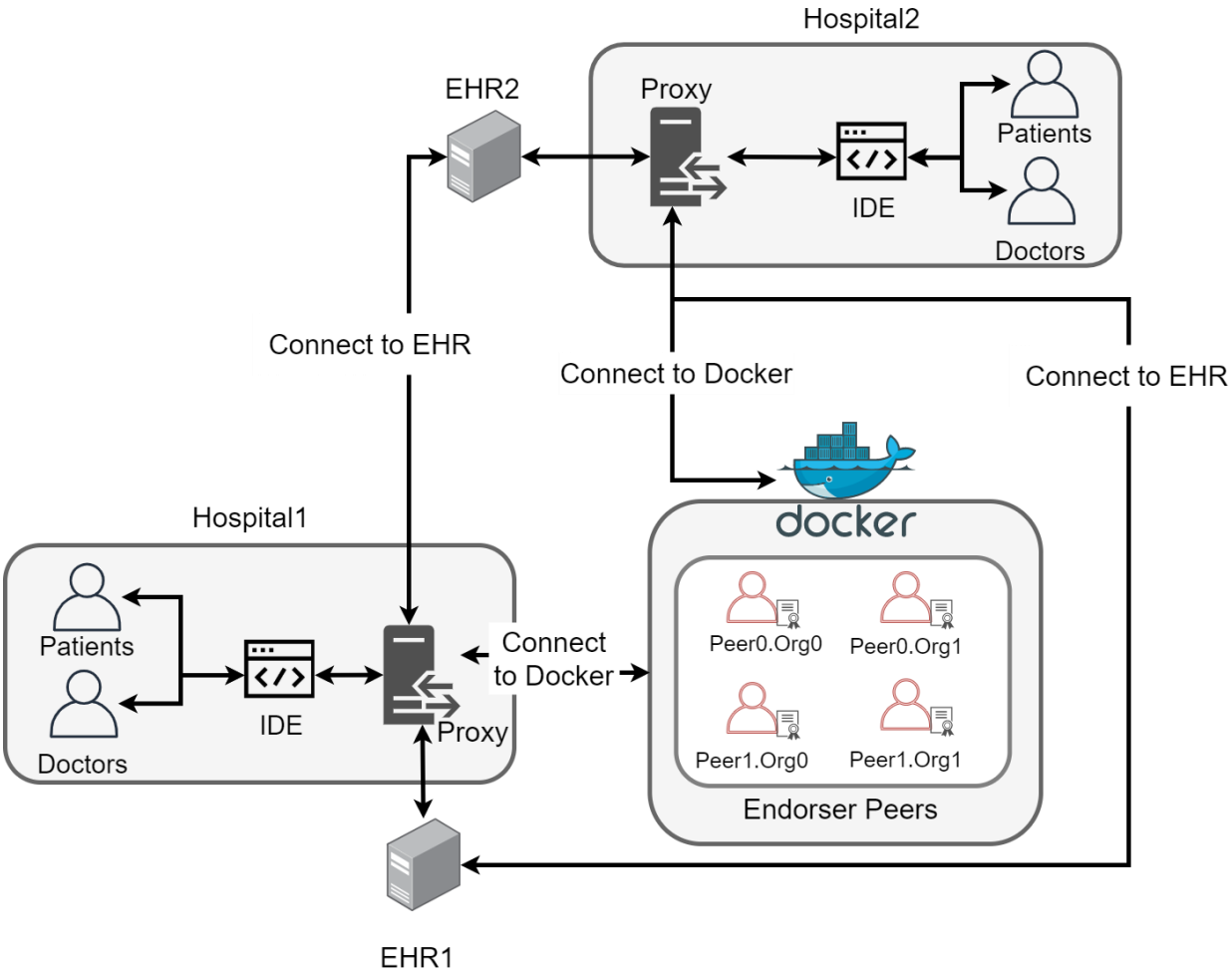


Figure 27: Prototype experiments

more endorser peers and avoid putting lots of endorsers to execute the chaincode in the endorsement policy.

The above prototype is not the same with the 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.

In the implementation of this system, all the verification steps are essential for the proof-of-concept of my blockchain system. For preserving patient privacy, I adopt AES algorithm for symmetric-key encryption of patient data and EC-ElGamal algorithm for asymmetric-key encryption of the symmetric key in the proxy re-encryption scheme. The asymmetric-key pair are also used for signature on the transaction proposal. However, for the purpose of further strengthening security, patient can have another key pair for the signature different from the one of the encryptions. The former is generated by using the Hyperledger Fabric function, the latter by importing a function of EC-ElGamal encryption using elliptic curve cryptography. When a patient chooses to have two pairs of keys, they bear the burden more to keep them secretly. In the case that a patient loses their private key, a key escrow system is assumed to be used for retrieving the lost key or symmetric keys from the ECert issuer or the hospital only for decryption of the patient data. After all, the received key must be used temporarily before new key and new ECert are issued for the patient.

I hash EID with salt to avoid for transactions of the records related with a patient having the same hash value of EID with which the patient records might be traced undesirably along the ledger. Meanwhile, this technique causes longer processing time to find out a patient in query of the data. To make it faster, doctors can input many relevant query keywords for obtaining the data.

These keywords include not only EID but timestamp and hospital id. The proxy's roles are to connect different EHRs through a secured communication network, download the medical records and re-encrypt patient's data. This scheme makes the processing time shorter in transferring a patient's data securely, otherwise the data needs to be sent to the patient to decrypt using the patient's private key and encrypt again using the receiver's public key before sending back to the proxy and then to the receiver. For proxy re-encryption, I adopt AFGH algorithm because it uses receiver's public key rather than the private key as in BBS algorithm [128] where the receiver's private key is created and used transiently only for receiving the data. To strengthen the privacy in access to records, patients can put the consents with conditions in the transaction of records for sharing them to third party. Furthermore, the ledger retains events of sharing data and relevant person's information, which facilitate the auditing procedure.

4.3 Related Works

There are several projects to establish a medical information sharing system based on the blockchain. Among them, Medrec [120] is an early research applying private Ethereum platform to Electronic Medical Record (EMR). In Ethereum, executable program run in the network is called as smart contract instead of Chaincode. Ethereum requires mining mechanisms to sustain the distributed ledger, which is time-delayed process with miners competing in Proof of Work even though it is not difficult to make private platform have short block time less ten seconds. Medical stakeholders, who are researchers, public health authorities, etc., need to be incentivized to participate actively as miners. To deal with these issues, MedRec 2.0 is currently under development [129].

Ancile [104] is another blockchain-based system using private Ethereum platform, which applies a similar technique with ours in medical records management adopting On-chain and Off-chain concept. Ancile uses the distributed proxies to re-encryption called a blinding re-encryption by splitting the ciphertext for the re-encryption between multiple nodes.

On the other hand, Alevtina et al. [121] uses Hyperledger Fabric in the cloud system. In their system, the data structure consists of key and value pair. The key is hash of a combination of symmetric key and a Uniquely Identifiable Information (UII) of the patient and the value is the records metadata. To reduce vulnerability of the system, patients encrypt each of their data using different symmetric keys. However, this incurs heavy burden of key management such that patients need to choose the corresponding symmetric key for generating a key number for every time they query for the data.

4.4 Summary

In conclusion, as business processes become more distributed, centralized workflow management is facing major challenges in meeting the conflicting requirements of scalability, security and openness. Interoperation and integration of cross-organizational business processes rely on distributed, autonomous, and also heterogeneous services for task executions. My system can be used to constitute a large-scale EHR system. It is flexibly configurable to be a layer belongs to existing EHR systems to strengthen security in the management and exchange of medical records. My system takes on the roles of a patient identifier, a trustee access log and registry of patient records. Even though my system does not offer explicit incentives to participants as other blockchain-based systems do by issuing a cryptocurrency, it will benefit users and stakeholders too including healthcare service providers and the government.

By using this system model, patients have to be in the hospital for providing the re-encryption key to doctors when doctors need patient's records. I see a problem in this system model so that I propose a new model of patient consent, which patients can store their consents of accessing to data in the blockchain, for allowing the accessing to patient's records if their access request is matched with the patient's consent. In the Chapter 5, I present about the proposed model of patient consent for data sharing and its components.

Chapter 5 Proposed Model of Patient Consent for Data Sharing

I presented my proposed system model for integrating different EHRs in the Chapter 4. That proposed system is not used for dealing deeply on the accessing to medical records based the patient's consent. Moreover, patients should deserve the application that they can use to design and manage their consents flexibly. Simultaneously, the preserving-privacy of patient when allowing someone for the accessing to resources is very important so that in this Chapter 5, I introduce the new proposed model of patient consent, which is adaptable to use in the blockchain system. I first explain the definition of word "Purpose" and its relationship with the consent, which is used for the formal policy of sharing the medical records in this research. I then describe my contribution of the purpose-based consent model and its components.

5.1 Definition of Purpose and Consent

This section I aim to study about the meaning of purpose. From Oxford dictionary,

"Purpose is the intention, aim, or function of something; the thing that something is supposed to achieve [130]*".*

One may ask or talk, for example, about the purpose of playing the video game "Why do you want to play the video game?", reading the newspaper "Why do you read the newspaper?", and sharing history of the medical records "Why do you share your histories of the medical

record?". I can receive various of answers from different people, for example, such as the purpose of playing the video game is for entertaining, the purpose of reading the newspaper is for gaining knowledge, and the purpose of sharing the history of the medical records is for donating to research laboratories or doctors who are working on the cancer research. At the same time, I see the role of purpose in the definition of the sense of the word *"wish"*. From Oxford dictionary,

> *"Wish is a thing that you want to have or to happen* [131]*".*

By observing the definition of *"purpose"* and *"wish"*, I see that they have similarity. They are all about reaching the goal or achieving something.

In term of the GDPR's regulation of the general data protection, I see that the word *"wish"* is mentioned in the definition of *"consent"*

> *"Any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her* [132]*".*

For instance, one may give a consent to doctors or researchers who want to collect the medical records for researching to improve the treatment of cancer. In other word, I can say that the data owner supposes to give medical to data requestors for the purpose of "cancer research activities". For retrieving the medical data, it requires data requestors to show their purpose or wish of using the data, which comply with the goal of the consent of data owners. With the above arguments, I acknowledge that *"purpose"* and *"consent"* have a relationship with each other. Thus, I can say *consent* is used to express the wish or purpose of the agreement of doing something.

In the next section, I present about the purpose model and its usage in the context of applying consent for data sharing.

5.2 Purpose Model

A purpose is defined as the reason for data collection and use [48]. It is the main keyword of patient's consent as I mentioned above, because patient decides to confine collection and use of their data within a certain range of specified purpose. In this section, I explain the purpose-based access control scheme and how to adapt it for my patient consent model, including the way how to make a data request in the model.

5.2.1 Purpose Tree

Purpose has its scope of coverage, narrow and wide, and can be organized in a hierarchical tree structure, so-called purpose-tree as in Figure 28.

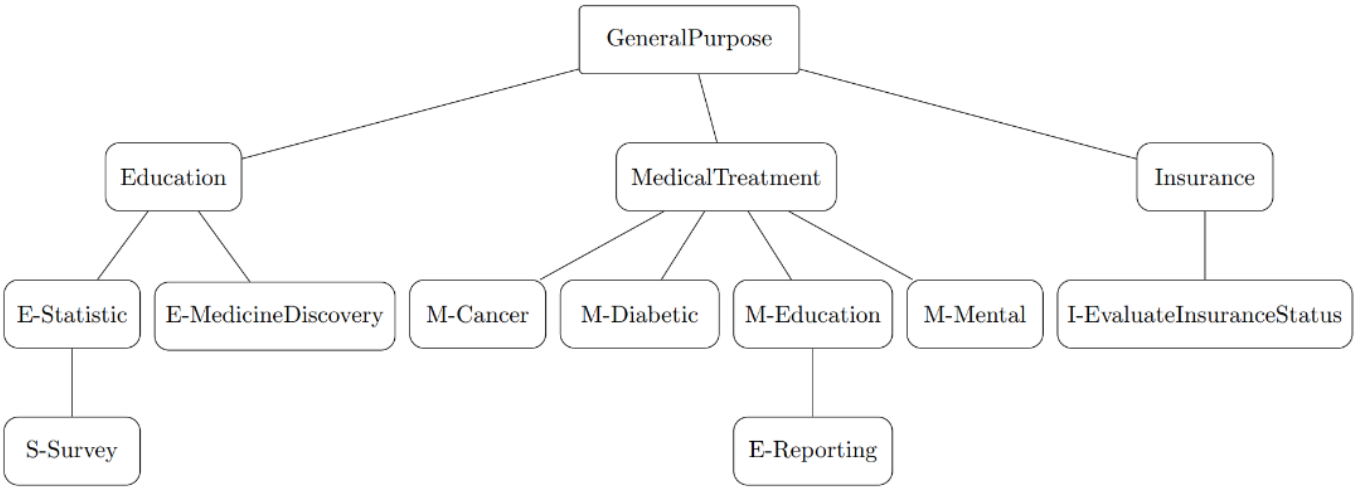


Figure 28: A purpose-tree

Definition 1: Purpose-tree illustrates the hierarchical and range of purposes.

The top node of the tree is the general purpose, which has the widest scope containing its descendants of purpose nodes. Each connection line from a node to another represents relationship

in purpose-tree. Each of purpose in the tree has different policies, which based on the organizations, for preserving patient's privacy when distributing data. In real practise, system developer can edit the name of each purpose of the tree to avoid the duplication of names, which is difficult for searching for purpose in the purpose-tree.

Definition 2: Purpose-tree represents the common goal of the organizations for exchanging data with each other.

It has various types of organization, which join in the system for medical information exchange. They have different intention of retrieving the medical records. For instance, core hospitals create the system for sharing medical records for the purpose of *"MedicalTreatment"* so that they will discuss with each other for making the purpose structure that have the *"MedicalTreatment"* as the new parent node. Next time, a research laboratory from a university wants to join in the above network for exchanging patient's records for education. Thus, all of members have to discuss to improve the previous purpose-tree by adding a new parent node of the accessing to medical data for *"Education"* purpose. In addition, Purpose-tree is so closely related with the privacy policy that organization members need to agree with the structure and attributes of it.

The purpose associated with data, regulating access to the data, is referred to as Intended-Purpose, and purpose for accessing data as Access Purpose [48], [49]. Usually, intended-purpose is described in patient consent in advance which states that for what purpose the data can be accessed. Thus, when a requestor asks an access to data, they should clarify the access purpose, that is checked against the intended-purpose of the data written in the patient consent. If both of the two purposes are matched, the system allows the requestor to access the data. The matching

rule is provided by the consent model that stipulates how the patient consent describes intended-purpose and how requestor prescribe an access request with their access purpose. In a general consent model, for instance, when a patient chose *"Insurance"* as intended-purpose for a data access, the system allows only the requestors to access the data when their access-purpose complies with intended-purpose, *"Insurance"* or its descendants in the purpose-tree.

While the blockchain support the (Key, Value) searching, I converted the purpose tree to the JSON array type and stored them in the blockchain, as shown in the Figure 29. In Figure 30, it is the purpose-tree stored in the world state database, which is used as a set of the purpose for accessing to data in the channel network.

```
"GeneralPurpose": {
    "Education": {
        "E-Statistic": {
              "S-Survey":{}
        },
        "E-MedicineDiscovery":{}
      },
    "MedicalTreatment": {
        "M-Cancer":{},
        "M-Diabetic":{},
        "M-Education": {
              "E-Reporting":{}
        },
         "M-Mental":{}
      },
    "Insurance": {
        "I-EvaluateInsuranceStatus":{}
      }
  }
```

Figure 29: JSON array type of the purpose-tree of Figure 16


```
"_id": "v1.3",
"_rev": "1-701d990993a8109b0a1a28932dbd4b89",
"ConsentTreeFull": "[{\"Education\":{\"E-Statistic\":{\"S-Survey\":\"null\"},\"E-MedicineDiscovery\":\"null\"},\"MedicalTreatment\":{\"M-Cancer\": \"null\",\"M-Diabetic\"
"ConsentType": "Consent_Tree",
"ConsentVersion": "v1.3",
"~version": "\u0000CgMBCQA="
```


Figure 30: Purpose-tree stored in world state database

5.2.2 Purpose-based Consent Model

As already mentioned, when patient gives their consent to usage of their data, they confine it within specified intended-purpose. In addition, they usually give healthcare professionals different levels of permissions according to the persons' role. In RBAC (Role Based Access Control) model [85], [133], the role represents job function or job title in the organization, and it is defined in role hierarchy. The access privilege is given based on the job title. It eases data owners to allow requestors data access based on the requestor's role rather than pointing to the user in the organization.

In my consent model, I adapt purpose-based access control scheme and RBAC's concept. The patient consent in my model contains intended-purpose of data access and specific user's role. However, I do not use role hierarchy to make my model as simple and clear as possible having generally acceptable structure. I modified the intended-purpose-based model of Byun et al. [48], [49]. Sharing patient data among many organizations is a complicated task, and each participant organization has their own privacy policy, which may be cause of discrepancy. Moreover, patient need to know how their consent is given and how it can affect their privacy, and doctor or healthcare personnel should explain it clearly until patients understand well enough.

The data access control in my consent model adopts basically whitelisting with an exception that patient can make some designated blacklists within a whitelist.

Each consent consists of four main tuples expressed as follows:

< Role; AdmitteeEIds; Action; Intended-Purpose>

- **Role**: Job title or job function of requestor who has their specific EID (it is id of Enrolment certificate). Examples are Cardiologist, Physician etc.

- **AdmitteeEIds**: Patient can add some designated doctors or healthcare professionals such as family doctor or medical specialist etc., who are allowed to access to the patient data. Their EIDs are listed here, or this element may remain blank.
- **Action**: The activity on the data. Examples are Copy, Read etc. Actions can have access privilege levels, so that the privilege of Copy includes that of Read, and the opposite is not allowed, i.e. Copy $\supset$ Read.
- **Intended-Purpose**: This element consists of two tuples as follows: < **AIP; PDP** > where AIP is Allowed Intended Purpose and PDP is Prohibited Descendant Purpose. AIP contains PDP, as the former is the ancestor of the latter.

There have been some rules of applying the patient's consent:

- Role and DoctorID are basic qualifiers necessary to specify requestor's legitimacy. One of these two and the other two tuples should be simultaneously complied by the requestor, i.e. $(Role \vee DoctorID) \wedge Action \wedge IntendedPurpose$
- A data access is allowed only for the allowed intended-purposes (AIP) that are explicitly written in a patient consent for the data, making all the other purposes implicitly prohibited one.
- Multiple AIPs constitutes a whitelist, for which data access is allowed.
- If an AIP has descendant purposes in the purpose-tree, then all of the descendants are also allowed purposes, belonging to the whitelist except some specific ones.
- Some of descendants of an AIP can be as prohibited descendant purpose (PDP), for which data access is not allowed, such that, $\exists\ PDP \in AIP$.
- Multiple PDPs under an AIP constitutes a blacklist (BlackList), consisting of a subset of the ancestor AIP, such that, $\forall\ PDP \in BlackList \subset AIP$.

To respond to various demands for data access from diverse requestors, patient can make a list of their consents for the data. The patient consent list is made by binding together multiple consents in combination with a variety of roles and intended-purposes. If an access request matches with one of the consents in the list, the data access is allowed. Figure 31 shows a simple example of it. In my system, the consent is stored on the blockchain with its hash value and metadata of the relevant patient records. However, it also can be stored off-chain along with the patient records.

In case, patient decides to change their consent, the consent list needs to be modified. However, blockchain is append-only storage, so new version of the consent list is newly appended into the ledger along with the transaction number of the old list in the blockchain. Thus, all the versions of the consent lists are connected by the transaction numbers with additional information. This provides traceability of the patient consents with the data integrity. The newest version is kept

```
"Consent": [
"consentId" : "#123",
  {
      "role": "nurse, physician",
      "action": "read",
      "intendedPurpose": "generalPurpose; M-Education, M-Mental"
  },
  {
      "role": "cardiologist, pharmacist",
      "action": "copy",
      "intendedPurpose": "generalPurpose; Education"
  },
  {
      "role": "health insurance staff",
      "admitteeIds": "#2",
      "action": "read",
      "intendedPurpose": "I-EvaluateInsuranceStatus; null"
  }
]
```

Figure 31: A simple example of patient's consent for a specific data in the state database

in state database, so-called World Sate in Hyperledger Fabric, and used when access requests are checked, while old versions used for management of history.

5.3 Access Request

When the requestor asks for data access, they have to possess proper qualification and appropriate purpose for the access. After requestor's role and the access purpose is validated successfully, the system allows the access within the designated activity on the data.

Actually, when requestors try to get some patient data with any usage purpose, they send the system a query having data attributes for data search together with the access request and additional helpful keywords. The data attributes and the keywords include patient-related information, hospital, department, doctor, disease, time and date, demography, etc. They have a wide variety of types and range, and get gradually narrowed down during search until finding out the relevant list of candidates of target data. After obtaining the target list, the system starts to validate the access request with the patient consent pertaining to each of the candidates. This order of searching procedure makes the system perform effectively and save time of data search and validation of access request. That's why my model separates the data attributes from the access request.

In my system, the access request has simply two tuples

<Access-Purpose; Action>

- **Access-Purpose**: The data requestor's purpose of using the data
- **Action**: The activity on the data. Examples are Copy, Read etc., having access privilege levels the same as in the patient consent.

Whether a data access is allowed or not depends on the relationship between requestor's Access Purpose (AP) and Intended-Purpose in the patient consent. The following is basic compliance rule to which access request is subject.

- If AP is included in Prohibited Descendant Purposes (PDP), the access request is rejected at all, i.e. $AP \notin PDP$
- Any of consent, which has Allowed Intended-Purposes (AIP) that is ancestor of AP, allows the access requests excluding PDP in the AIP, i.e. $AP \in AIP \ and \ AP \notin PDP \subset AIP$

Among the four main tuples in patient consent of my model, Role and requestor's EID are usually invariable, and registered in the system or the organization where they belong to. So, the access request contains only variable elements, access purpose and action. Thus, the requestor is authenticated by the system using their EID, and their role is also identified by the system which can consult participant organizations if necessary.

5.3.1 Validation the Access Request

Figure 32 shows a case when a nurse wants to read her patient's data, and submits an access request with data attributes and keywords for query. The system finds out a resultant list of data,

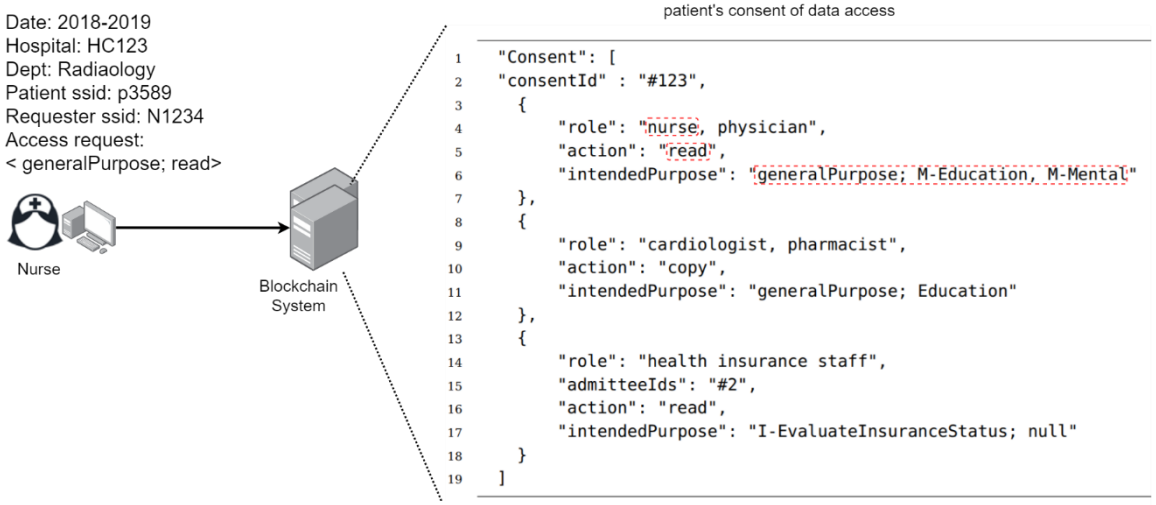


Figure 32: Validation of request's access request with patient's consent list

then checks the patient consent pertaining to each data in the list. If her access request complies with patient consent, then she can get access to the data and do action on the data. For this case, she can read a data for general purpose, except education and mental illness-related purposes. If she wanted to copy the data, her request should have been rejected, then she might have lowered access privilege to READ. She might have also lowered Access Purpose to get more data confining the purpose for medical treatment.

5.4 Implementation

I follow the same information network that I did experiments in the implementation of previous chapter. This implementation is to test the applying purpose-based consent for accessing to data.

5.4.1 Chaincode for Consent Management

Chaincode is a program that performs business logic agreed by members of the network, working the same as smart contract in Ethereum. It is run by endorsers to access blockchain through transactions [134]. I created a chaincode to deal with patient consent. The following describes the chaincode's operations.

i. Consent Management

This function manages patient consent allowing the patient to query for their past records and review the consent and update it in the blockchain as shown in Figure 33. At first, it checks the validity of the proposal's format and intended-purpose with the purpose-tree stored in the blockchain. Then, it queries for the consent policies and medical record transactions in the

blockchain based on the patient EID. Patient can update their consents by appending new ones to the relevant medical record transactions.

```
   Input: PatientProposal /* It has "eID" of patient and the "NewConsent"                    */
   Output: Message /* of successful or unsuccessful of creating or updating the "Newconsent"  */
1  Function ConsentActivities(PatientProposal):
2      if PatientProposal format is correct then
3          Query in the blockchain for patient's record transactions based on eID, then store
           these query results to the array of RecordTransaction
4          if Patient selects a transaction from the array of RecordTransaction then
5              if Patient wants to upload a Newconsent of the selected transaction then
6                  Append the Newconsent to the selected transaction
7                  Return Message
8              else if Patient wants to update a consent in the selected transaction then
9                  Deactivate the old consent
10                 Append the Newconsent to the selected transaction
11                 Return Message
```

Figure 33: Pseudocode of a part of the Chaincode for patient consent management

ii. Consent Check

```
   Input: PatientRecordTx, AccessRequest
   /* PatientRecordTx is the transaction of patient's records, which the doctor received after he
      queries in blockchain using keywords                                                    */
   /* AccessRequest is a request of the doctor, which contains "Action" and "AccessPurpose"   */
   Output: RecordURL, RecordMetaData
1  Function ConsentChecked(PatientRecordTx, AccessRequest):
2      Query in the blockchain for doctor's role
3      Compare attributes of patient's consent of each PatientRecordTx with the doctor's role
       and the AccessRequest; if a consent of PatientRecordTx is matched with AccessRequest
       and the doctor's role then
4          Return the RecordURL and RecordMetaData of that PatientRecordTx
```

Figure 34: Pseudocode of a part of the Chaincode for patient consent check

This function works to check requestor's access request with the patient's consent stored in blockchain as shown in Figure 34. After extracting the proposal, it checks the validity of proposal's format and requestor's role in the organization. Then, it queries for transactions in blockchain based on searching keywords in the proposal. After querying successfully, the chaincode compares entities of access request in the proposal with attributes of the patient's consent in the transactions. Finally, the chaincode only sends transactions that patient's consents are matched with the access

request. As mentioned early, transactions in blockchain contain records' URL in EHR, which is used to provide the location of patient medical data.

5.4.2 Use Case Scenario

I continue describing the Alice's journey of allowing doctors for the accessing to her medical data. In the scenarios of previous chapter, Alice needs to be in the hospital and generate the re-key for doing the re-encryption of her records if she wants to share her records to the doctors. In spite of that, the scenario of this model would allow Alice to provide the consent of accessing to her records to the requestors if their requests are matched with the attributes in Alice's consents. I explain as following about the procedure of Alice to manage her consents and how the system can validate the access request of doctor with Alice's consent.

i. Consents management

In Figure 35, it demonstrates the process of the consent management of Alice. She has transactions of medical record stored in the blockchain. She might want to create new consents or update her consents, in order to allow Bob to access to her records based on her specific purpose, that stored in the transactions of medical record. To do that, she needs to make a proposal that has the query keywords, and she sends it to the endorser peers. They execute the consent chaincode to query in the ledger based on the query keywords. After querying successfully, endorser peers send all query results to Alice. Then, Alice selects transactions that she wants to do the creation or updating of the consents. After that, she creates a new proposal and inserts the selected transactions, input action (Creating or updating) and the new consent that she wants to input in the blockchain. She then sends that proposal to endorser peers. If Alice wants to create a new consent, endorser peers will execute the consent chaincode for insert that new consent to the blockchain.

For appending a new consent to the transaction, it is unless that the transactions do not have the consent before. Otherwise, chaincode will not endorse that proposal.

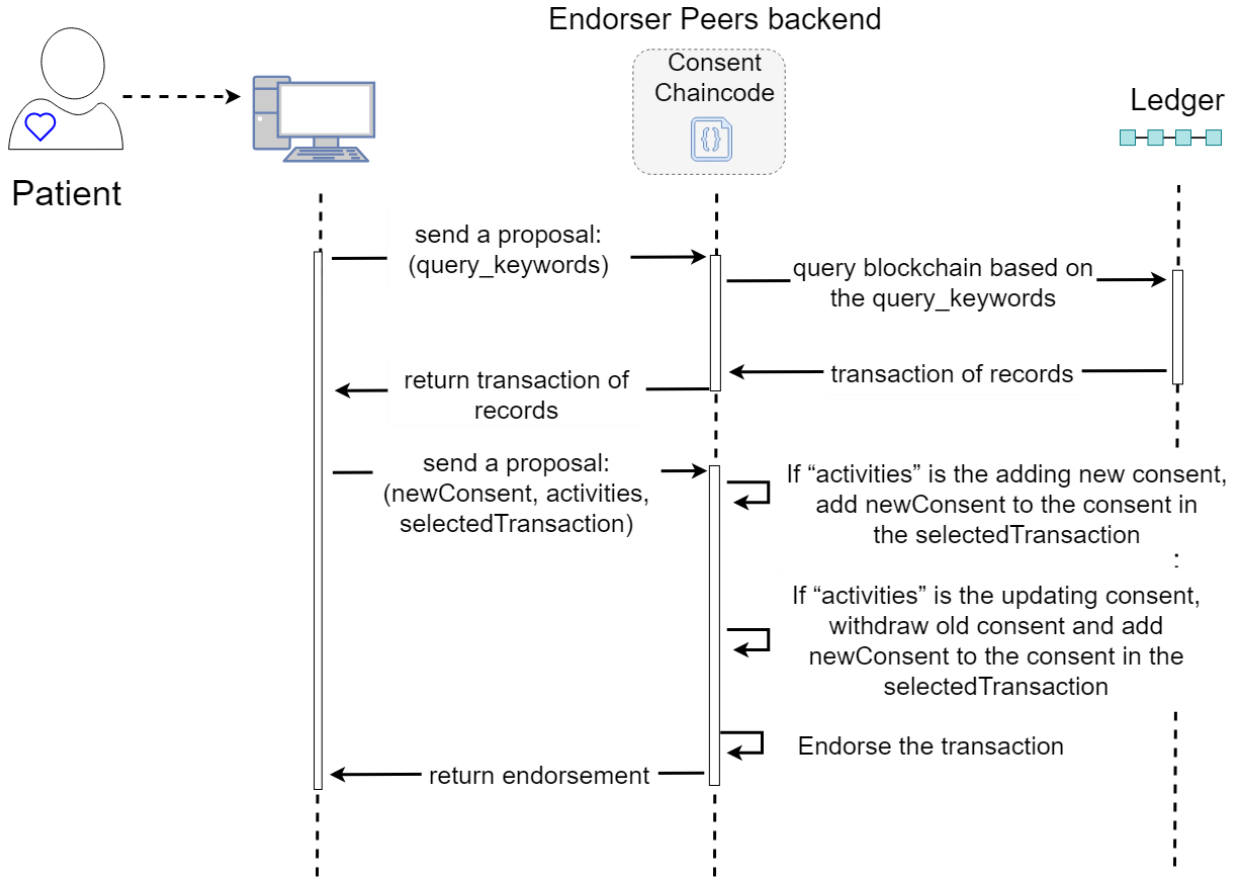


Figure 35: Patient manages the consents

If Alice wants to updating the new consent, the chaincode will withdraw old consent and add the new consent to the consent in the selected transactions, and chaincode of each endorser peers will endorse that proposal. Each endorser peer will send the endorsed proposal to Alice. Finally, Alice practices the activity of HLF for inserting her proposal to the ledger. As showed in Figure 37, it the patient's records stored in the world state database before updating the consent list. It has three conditions in the consent list, then it increases 2 conditions after updating the consent list as showed in Figure 36. Alice also can request to peers for querying her histories of

updating the records transaction in the blockchain. As showed in the Figure 38 and Figure 39, peers send Alice for the histories of updating her data of the past transactions.


```
"Consent": [
  {
    "action": "Read",
    "name": "PL1",
    "purpose": "Education; E-MedicineDiscovery",
    "role": "Nurse"
  },
  {
    "action": "Copy",
    "name": "PL2",
    "purpose": "MedicalTreatment, Education; M-Education",
    "role": "Cardiologist"
  },
  {
    "action": "Read",
    "name": "PL3",
    "purpose": "GeneralPurpose; Insurance",
    "role": "Nurse"
  }
],
"Date": "2020-06-15",
"EncryptedSymmetricKey": "tha+zuvjlnvyd6/w7soi",
"ID": "20200615082803",
"ObjectType": "Medical_repository",
"PatientECertHash": "$2a$10$zAEkYujJMGbbAamUUqjdsumqNEhTDnwKhqEaYudANcm6BoFecRTkC",
"RecordAddress": "192.168.145.2/oits8a+pkjj9",
"RecordHashValue": "d7330d6469a8aae057bce4a19a5a26ed81fb8d080182e85d3dba8321b65aadcf",
"RecordsOwnerAddress": "112.54.14.75/bed48dba7",
```


Figure 37: Patient records in world state database before updating the consent


```
"Consent": [
  {
    "action": "Read",
    "name": "PL1",
    "purpose": "Education; E-MedicineDiscovery",
    "role": "Nurse"
  },
  {
    "action": "Copy",
    "name": "PL2",
    "purpose": "MedicalTreatment, Education; M-Education",
    "role": "Cardiologist"
  },
  {
    "action": "Read",
    "name": "PL3",
    "purpose": "GeneralPurpose; Insurance",
    "role": "Nurse"
  },
  {
    "action": "Copy",
    "name": "PL4",
    "purpose": "Education; E-MedicineDiscovery",
    "role": "Research Student"
  },
  {
    "action": "Read",
    "name": "PL5",
    "purpose": "Insurance",
    "role": "Insurance"
  }
],
"Date": "2020-06-15",
"EncryptedSymmetricKey": "tha+zuvjlnvyd6/w7soi",
```


Figure 36: Patient records in world state database after updating the consent


```
[{"TxId":"e460b6764cb00020db20321f96eee524e4476d4b7fae8f034e5537f5b20d7051", "Value":{"ObjectType":"Medical_repository","ID":"20200615082803",
[{"TxId":"e460b6764cb00020db20321f96eee524e4476d4b7fae8f034e5537f5b20d7051", "Value":{"ObjectType":"Medical_repository","ID":"20200615082803",
[{"TxId":"e460b6764cb00020db20321f96eee524e4476d4b7fae8f034e5537f5b20d7051", "Value":{"ObjectType":"Medical_repository","ID":"20200615082803",
[{"TxId":"e460b6764cb00020db20321f96eee524e4476d4b7fae8f034e5537f5b20d7051", "Value":{"ObjectType":"Medical_repository","ID":"20200615082803",
"PatientECertHash":"$2a$10$zAEkYujJMGbbAamUUqjdsumqNEhTDnwKhqEaYudANcm6BoFecRTkC","Date":"2020-06-15","Consent":[{"name":"PL1","role":"Nurse","action
"PatientECertHash":"$2a$10$zAEkYujJMGbbAamUUqjdsumqNEhTDnwKhqEaYudANcm6BoFecRTkC","Date":"2020-06-15","Consent":[{"name":"PL1","role":"Nurse","action
"PatientECertHash":"$2a$10$zAEkYujJMGbbAamUUqjdsumqNEhTDnwKhqEaYudANcm6BoFecRTkC","Date":"2020-06-15","Consent":[{"name":"PL1","role":"Nurse","action
"PatientECertHash":"$2a$10$zAEkYujJMGbbAamUUqjdsumqNEhTDnwKhqEaYudANcm6BoFecRTkC","Date":"2020-06-15","Consent":[{"name":"PL1","role":"Nurse","action
{"name":"PL3","role":"Nurse","action":"Read","purpose":"GeneralPurpose; Insurance"}]
{"name":"PL3","role":"Nurse","action":"Read","purpose":"GeneralPurpose; Insurance"}]
{"name":"PL3","role":"Nurse","action":"Read","purpose":"GeneralPurpose; Insurance"}]
{"name":"PL3","role":"Nurse","action":"Read","purpose":"GeneralPurpose; Insurance"}]
```


Figure 38: A Tx that holds the past data

{"TxId":"a03ab6d61508966d0da32629fcf17cb8f11aa6ffe6ee2c21a86b1df5f83100cd", "Value":{"ObjectType":"Medical_repository","ID":"20200615082803",
{"TxId":"a03ab6d61508966d0da32629fcf17cb8f11aa6ffe6ee2c21a86b1df5f83100cd", "Value":{"ObjectType":"Medical_repository","ID":"20200615082803",
{"TxId":"a03ab6d61508966d0da32629fcf17cb8f11aa6ffe6ee2c21a86b1df5f83100cd", "Value":{"ObjectType":"Medical_repository","ID":"20200615082803",
{"TxId":"a03ab6d61508966d0da32629fcf17cb8f11aa6ffe6ee2c21a86b1df5f83100cd", "Value":{"ObjectType":"Medical_repository","ID":"20200615082803",
"PatientECertHash":"$2a$10$zAEkYujJMGbbAamUUqjdsumqNEhTDnwKhqEaYudANcm6BoFecRTkC","Date":"2020-06-15","Consent":[{"name":"PL1","role":"Nurse","action
"PatientECertHash":"$2a$10$zAEkYujJMGbbAamUUqjdsumqNEhTDnwKhqEaYudANcm6BoFecRTkC","Date":"2020-06-15","Consent":[{"name":"PL1","role":"Nurse","action
"PatientECertHash":"$2a$10$zAEkYujJMGbbAamUUqjdsumqNEhTDnwKhqEaYudANcm6BoFecRTkC","Date":"2020-06-15","Consent":[{"name":"PL1","role":"Nurse","action
"PatientECertHash":"$2a$10$zAEkYujJMGbbAamUUqjdsumqNEhTDnwKhqEaYudANcm6BoFecRTkC","Date":"2020-06-15","Consent":[{"name":"PL1","role":"Nurse","action
"Read","purpose":"Education; E-MedicineDiscovery"},{"name":"PL2","role":"Cardiologist","action":"Copy","purpose":"MedicalTreatment, Education;
"Read","purpose":"Education; E-MedicineDiscovery"},{"name":"PL2","role":"Cardiologist","action":"Copy","purpose":"MedicalTreatment, Education;
"Read","purpose":"Education; E-MedicineDiscovery"},{"name":"PL2","role":"Cardiologist","action":"Copy","purpose":"MedicalTreatment, Education;
"Read","purpose":"Education; E-MedicineDiscovery"},{"name":"PL2","role":"Cardiologist","action":"Copy","purpose":"MedicalTreatment, Education;
,{"name":"PL3","role":"Nurse","action":"Read","purpose":"GeneralPurpose; Insurance"},{"name":"PL4","role":"Research Student","action":"Copy","purpose
,{"name":"PL3","role":"Nurse","action":"Read","purpose":"GeneralPurpose; Insurance"},{"name":"PL4","role":"Research Student","action":"Copy","purpose
,{"name":"PL3","role":"Nurse","action":"Read","purpose":"GeneralPurpose; Insurance"},{"name":"PL4","role":"Research Student","action":"Copy","purpose
,{"name":"PL3","role":"Nurse","action":"Read","purpose":"GeneralPurpose; Insurance"},{"name":"PL4","role":"Research Student","action":"Copy","purpose
E-MedicineDiscovery"},{"name":"PL5","role":"Insurance","action":"Read","purpose":"Insurance"}],"RecordHashValue":"d7330d6469a8aae057bce4a19a5a26ed8
E-MedicineDiscovery"},{"name":"PL5","role":"Insurance","action":"Read","purpose":"Insurance"}],"RecordHashValue":"d7330d6469a8aae057bce4a19a5a26ed8
E-MedicineDiscovery"},{"name":"PL5","role":"Insurance","action":"Read","purpose":"Insurance"}],"RecordHashValue":"d7330d6469a8aae057bce4a19a5a26ed8
E-MedicineDiscovery"},{"name":"PL5","role":"Insurance","action":"Read","purpose":"Insurance"}],"RecordHashValue":"d7330d6469a8aae057bce4a19a5a26ed8

Figure 39: A valid Tx that holds the current data

ii. Doctor requests for patient's records

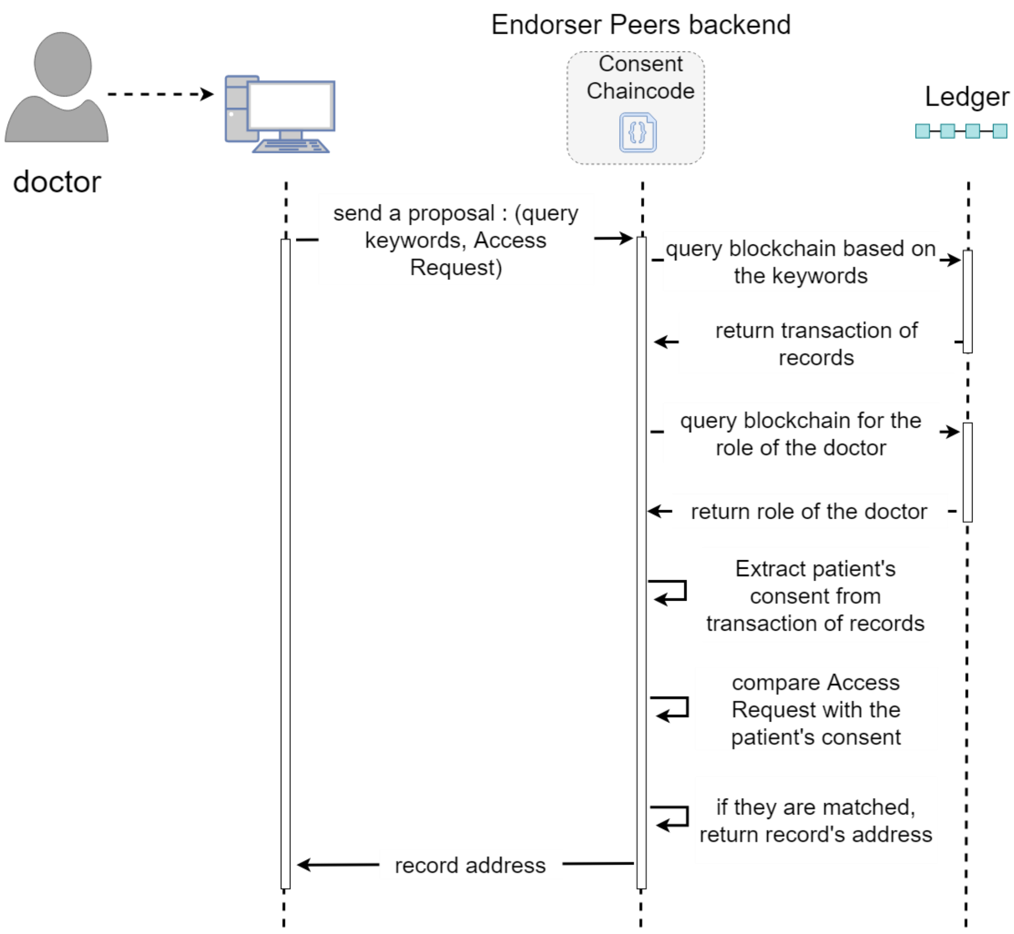


Figure 40: Doctor requests for record address

After Alice modified her consent for allowing Bob to access to her records. Then, Bob wants to get Alice records so that he creates a proposal, which has query keywords and access request, and send it to the endorser peers. Endorser peers start executing the consent chaincode for query transactions of records of Alice and Bob's role in the blockchain. After retrieving the Alice's transactions, chaincode compares the consent in each transaction with Bob's access request. If the transactions that have the consent and Bob's access request are matched, chaincode will send that transactions to Bob. Finally, Bob extracts Alice's records address from that transactions for using to access to Alice records.

5.4.3 Prototype Analysis

I evaluated the prototype system with following targets: (i) to check the integrity of data of history which includes creating, withdrawing and updating the patient consent; (ii) to check the chaincode's function of validation of the provided intended-purpose and access-purpose with the purpose-tree; (iii) to investigate whether the chaincode correctly worked in validating the access request of doctor with the patient's consent for data access.

I could find out a peer which had wrong blockchain because it was intentionally provided with wrong query results. During the process of building block, the consensus protocol of Hyperledger Fabric can check it by comparing results from all endorser peers. The processing time of validating the access request is depended on the complexity and length of the purpose-tree and consent list in the individual transaction. Simultaneously, speed of validating the proposal is based on the privacy policies and the legal requirements set by the consortium.

5.5 Related Works

In this research, I proposed a new type of e-consent system for patient to manage the consent elaborately in dealing with their data. I adopted the purpose-based access control scheme from RBAC model of relational database [48], [49], [133]. My system also has a hierarchical structure of user's intended-purpose, but does not have it in user's role, because combining both hierarchical structures together makes the consent so complicate for patient to understand when they set it to their data. RBAC, from its name, is role-based Hospital-Centric, meanwhile my concept is Patient-Centric.

My system is different from most purpose-based centralized ones, i.e. ours is fully decentralized blockchain. Dwarna project [34] also provides a blockchain solution for dynamic consent in biobanking, however, they use boolean-based consent to allow requestors to access data, because the purpose is very simple in their research.

5.6 Summary

The purpose-based consent shows the goal of patients for allowing and disallowing someone to access to the medical data. By storing it in the blockchain, it provides availability to patients to create and update their consent in the hospitals or organizations, which are the members of the blockchain network. The list of purpose is designed as a tree structure, which is represented the common goal or intension of sharing the medical system across members in the network.

The patient's consent plays a very crucial role for preserving patient's privacy in healthcare data sharing. If restrictively given, it would cause inconvenience in dealing with the data, while indulgently given, patient would be exposed to a risk of privacy disclosure. I also understand that

purpose and consent is corelated with each other. In result, based on the proposed purpose model and allowing for the accessing to the data based on the records, patients can monitor and design their consent efficiently and requestors who have different role from the medical staffs. They can access to patient's records if their access requests comply with the patient's consent. Moreover, storing consents in the blockchain is a very effective method for preserving the patient's privacy because it is transparent to all nodes that patient's consent is not altered, and other stakeholders can join to execute the chaincode for validating access request of requestors.

Chapter 6 Discussion

Interoperability of healthcare provides many advantages to medical institution for enhancing the public health. It sometime also has the communication with other organizations which are not related to the healthcare sector such as insurance companies, research institutes and government for sharing medical records in order to evaluate insurance cost, collect for research purposes etc. To be a part of social development, I first proposed the blockchain system model, which is used to support the integrating of healthcare among different EHRs and other organizations. After that, I proposed a new model that is suitable to use in the above blockchain system for the consent management and validation. There have been some reasons that both system models work well together. Moreover, they are still important for overcoming problems of other use cases.

- **Sharing medical data based on patient consent**

When doctors or medical staffs upload the medical records, they also upload the patient's consent at the same time. But some patients might refuse to give the consent for data sharing. However, most of the medical data stored in the EHRs always have the consents. Patient's consent is used in order to allow the access to data requestors, especially, when the interoperability of healthcare is created, the exchange of medical data will be actively happened in the network. This system will be expanded if there have been many organizations join so that the traffic of requesting and responding the medical data must be very busy. In chapter 3, I mentioned blockchain's essential components and advantages of applying it for building a decentralized system for giving benefits in dealing with health data in the medical system. After learning from these existing

blockchain applications, I came up with the idea to design the blockchain application for the contribution to the interoperability of healthcare. Regarding to my proposed system, it allows medical staffs to search for patient's records histories. In addition, when a doctor requests for medical records, patients have to decide and provide the consent for that doctor to access to their medical records. Simultaneously, patients have to be in the hospital in order to give the consent. In consequence, it is difficult for patients to give the consent. To overcome this problem, I design a new consent model, which is the purpose-based consent. With that consent model, patients can manage their consent flexibly.

Thus, my first model is the system architecture for supporting the decentralization of the interoperability of healthcare, and the second model is used for managing and validating consent for allowing the accessing to data among nodes in the system. So far, if my system is deployed to current medical system, dealing with the digital signature of patients are a very severe challenge. Patients sometime lose their smart cards or forgot the password, which is used to unlock their smart cards before doing the digital signature.

In reality, the purpose-tree might be updated by the agreement of participant organizations for some reason, such as, when their privacy policy is updated or new member organization is added. In that case, it is difficult to interpret patient's consent based on the new purpose-tree, and basically, all the organizations need to get new patient consents again. However, the contract between patients can prepare in advance a description about how to deal with the situation.

- **Expansion of models for information sharing**

Decentralized system does not rely on the single entity. It is meant that all entity can handle to receive and response requests of requestors by themselves. It is the advantage of this system to

provides the availability of accessing to resources in network to avoid the bottleneck of the data process management. Since my system is the decentralized system using the blockchain technology, it can be used to solve problems of many use cases of information sharing. For instance, in the homecare service, a doctor from a hospital wants to get patient's homecare data for the specific purpose so that he might requests patients or homecare service provider for the data. In this scenario, if homecare service providers want to join in the network, they need to make a request to register as a member of this network. Otherwise, non-member institution can communicate through the member institution. After becoming the member of the blockchain network, homecare service providers and hospitals can exchange patient's records, which all access histories are stored securely in the blockchain. Simultaneously, purpose-based consents are checked before allowing the accessing to data. Even my system can be expanded for solving the information sharing of PHRs or other systems, this proposed system needs to be redesigned to comply with different use cases and organization policies.

Chapter 7 Conclusion

So far, I presented the overview of medical system and its challenges in Chapter 1. Then, I mentioned my research motivations and objectives for the contribution of solving the challenges in current EHRs. After that, I presented the fundamental compliances for building the medical system for improving the security and preserving-privacy, in Chapter 2. To comply these regulations, I have to use the technology or security techniques for increasing the security level of storing and sharing medical records, authentication and authorization the access based on the user identity, context and consent. Since my research is mainly to design the decentralized system for integrating EHRs, I explained the concept of blockchain technology and its applications in medical system in Chapter 3. After I see the potential of using blockchain technology for building the decentralized system especially in medical field, I proposed the system model for improving the interoperability of EHRs in Chapter 4. I then contribute the new model of allowing access to requestors based on the consent in Chapter 5. In Chapter 6, I discussed how these two proposed models are works well together, which are not only because of the relationship between medical data and consent, but also advantages of applying these two models in many use cases for information sharing. In this Chapter 7, I summary of the thesis and future works of my research.

7.1 Summary of the Thesis

In this thesis, I address the challenges in the medical system, which are the interoperability of healthcare system and the availability of allowing the access to medical data based on the patient consent. I provide a novel decentralized system for supporting the communication between

medical institutions with another stakeholders. Actually, they can build a centralized system for supporting this technique of sharing medical data, but medical institution mays need the reliable central system such as government; to initiate this project. If comparing it with the decentralized system that organizations can agree among their alliance for building such system. They also have the choice to add resources (hardware or software) or not to their existing system for increasing the performance of information sharing. Simultaneously, all medical organizations make their own decision. Then, the final behaviour of the system is the aggregate of the decisions of the individual nodes. Moreover, there is no single entity that receives and responds to the request.

Narrow to the challenges of the interoperability of healthcare, I acknowledge that medical institutions are having difficulty to integrate with each other because they are using different standard of EHRs. To overcome this problem, I suppose that sharing the records address is very effective, practical and less complicated for sharing the data. Moreover, these data are very small for storing and transmitting across different organizations. Simultaneously, the patient wants to monitor how their data is used based on their consent. Without the appropriate measures, they tend to be inactive in data sharing and even reluctant in data donation for research purpose. By seeing above opportunities, I design the blockchain system, which is the decentralized system, for supporting the integrating of EHRs and providing the availability to patients for accessing and monitoring their consent based on their purpose flexibly.

My proposed blockchain system and consent model complete each other to overcome the problem of sharing medical data based on patient consent. Simultaneously, they need the modification if I want to expand my models for solving the information sharing of other use case. My proposed system, I use random number, called Salt, to hash with the EID for pseudonymizing real data owner's EID stored in blockchain. In consequence, if malicious people get the blockchain

and know one transaction belong to one specific person, they cannot search for another patient data. They need to know the real EID of patients to claim all records of the specific person. Normally, EID is stored in the ECert of patient secretly so that it is difficult for malicious people to get this EID. In addition, my system helps to reduce the time of sending the medical data to data requestors because the data stored in the EHRs were encrypted by the patient's public key. In fact, when patients want to send encrypted data, patients need to decrypt and encrypt their encrypted data repeatedly for data requestors. By using the proxy re-encryption in my system, this problem can be solved. My system adapts the consent purpose-based for data access. All purposes are initiated by consortium members and stored in the tree data structure, which make it easy for managing and finding the ascendant and descendant of the purpose; and assigning the new purpose. I implement above two models in Java and Go programming language, execute them in the local network environment. Moreover, I compare my proposed system models with the existing blockchain system. As the result, my proposed system models gain more advantages than another.

My designed system will not be used to replace the CommonWell Health Alliance system because my system is not designed to cover all of functions of CommonWell Health Alliance project has. It is built to contribute as a new choice not only for the CommonWell Health Alliance, but also for another medical institutions for the decentralization system of the searching for health data address across different EHRs and allowing the access to the data based on the purpose-based as the consent. I expect that my research can help patients to find their medical histories easier when they visit other hospitals. It also can be used as a solution not only in patient data sharing between hospitals, but also in data donation for research purposes such as in the biobanking etc.

7.2 Future Works

As future work, I am going to test my system in the real hospital environment. I will prepare to deal with non-standardized data when it comes in the real-world field test.

Applying to hospitals

My system is a consortium network. If other medical institutions want to access this network, they need to make a request to register as a member of this network. Otherwise, non-member institution can communicate through the member institution. Peers are the trustful elements from each medical institution. They need to strengthen their own security to protect the peers from illegal access. At the same time, every medical institution needs to agree on the Chaincode logic before deploying them in the system. Thus, my blockchain system also can be run effectively in cloud system even though its fundamental standpoint is opposite in terms of decentralization. Cloud computing can provide a solution to the blockchain size problem that ledger size gets gradually bigger with time and peers will have difficulty to keep and process.

To obtain patient consent, a full and detailed explanation should be given to the patient, which includes both good and bad points of data sharing and potential risks. The system needs to have adequate user interface to support those basic explanation as well as to teach how to fill out the consent electronically and how to manage the consent. I do not deal with the user interface in this study.

To share data among multiple organizations, participants need to reach an agreement with common privacy policy that might compromise each member's peculiar feature in their own policy. It may be desirable for the purpose-tree to cover all participant usable purposes with plenty of branches in each node, however, it leads to complexity and reduction of data usability. To absorb

these discrepancies, I tried to make rules of my model simple and generally acceptable. In addition, my system makes patient simply give their consent to the doctor on-site without checking the written consent after verifying the patient being at present. Using user's private key which is kept in IC card is one of usual ways to do that.

Blockchain technology has revealed its great potential to innovate business processes. The properties of persistency, validity, auditability, and disintermediary that Blockchain offers can greatly improve modern business processes to achieve digitalization, automation and transparency. However, efforts spent for integrating Blockchain into business processes is still at infancy.

Adapting to GDPR's for the right of erasure

European General Data Protection Regulation (GDPR) [41], [135] requires data subject, i.e. patient, has the right to request erasure of personal data related to them, so-called right of erasure. Apparently, this is considered to be incompatible with blockchain's immutability, so it is a big challenge for all blockchain-based system to comply with the request. To address this difficult problem, my system stores patient records off-chain in EHR. In addition, it makes each transaction on-chain have a unique hash number of the patient's EID with the random number so called Salt [45], [136] to thoroughly pseudonymize the data owner, even though this sacrifices data searching performance. The link that connects the randomized patient EID to off-chain records resides in the off-chain database [34], and in case patient asks to erase their data, then the system remove the link and off-chain record. The URL written in the transaction, which is the address of data site in EHR, might be a clue to specify the patient, however, it is too difficult to do that because the URL is shared with many other patients. The patient consent may be stored off-chain along with patient record the hash remaining on-chain to maintain the integrity.

References

[1] T. Greenhalgh, S. Hinder, K. Stramer, T. Bratan, and J. Russell, "Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace," *BMJ*, vol. 341, p. c5814, Nov. 2010, doi: 10.1136/bmj.c5814.

[2] D. B. Taichman *et al.*, "Sharing Clinical Trial Data: A Proposal from the International Committee of Medical Journal Editors," *PLOS Medicine*, vol. 13, no. 1, p. e1001950, Jan. 2016, doi: 10.1371/journal.pmed.1001950.

[3] E. Warren, "Strengthening Research through Data Sharing," *N Engl J Med*, vol. 375, no. 5, pp. 401–403, Aug. 2016, doi: 10.1056/NEJMp1607282.

[4] N. Geifman, J. Bollyky, S. Bhattacharya, and A. J. Butte, "Opening clinical trial data: are the voluntary data-sharing portals enough?," *BMC Medicine*, vol. 13, no. 1, p. 280, Nov. 2015, doi: 10.1186/s12916-015-0525-y.

[5] C. Castaneda *et al.*, "Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine," *J Clin Bioinforma*, vol. 5, pp. 4–4, Mar. 2015, doi: 10.1186/s13336-015-0019-3.

[6] G. D. Schiff *et al.*, "Diagnostic Error in Medicine: Analysis of 583 Physician-Reported Errors," *Archives of Internal Medicine*, vol. 169, no. 20, pp. 1881–1887, Nov. 2009, doi: 10.1001/archinternmed.2009.333.

[7] R. Kaushal, K. G. Shojania, and D. W. Bates, "Effects of Computerized Physician Order Entry and Clinical Decision Support Systems on Medication Safety: A Systematic Review,"

Archives of Internal Medicine, vol. 163, no. 12, pp. 1409–1416, Jun. 2003, doi: 10.1001/archinte.163.12.1409.

[8] Y. Zhou *et al.*, "The impact of interoperability of electronic health records on ambulatory physician practices: a discrete-event simulation study," *Inform Prim Care*, vol. 21, no. 1, pp. 21–29, 2013, doi: 10.14236/jhi.v21i1.36.

[9] B. A. Stewart, S. Fernandes, E. Rodriguez-Huertas, and M. Landzberg, "A preliminary look at duplicate testing associated with lack of electronic health record interoperability for transferred patients," *Journal of the American Medical Informatics Association*, vol. 17, no. 3, pp. 341–344, May 2010, doi: 10.1136/jamia.2009.001750.

[10] J. Walker, E. Pan, D. Johnston, J. Adler-Milstein, D. W. Bates, and B. Middleton, "The Value Of Health Care Information Exchange And Interoperability," *Health Affairs*, vol. 24, no. Suppl1, pp. W5-10, Jan. 2005, doi: 10.1377/hlthaff.W5.10.

[11] G. Eysenbach, "What is e-health?," *J Med Internet Res*, vol. 3, no. 2, pp. E20–E20, 2001, doi: 10.2196/jmir.3.2.e20.

[12] R. S. Evans, "Electronic Health Records: Then, Now, and in the Future," *Yearb Med Inform*, vol. Suppl 1, no. Suppl 1, pp. S48–S61, May 2016, doi: 10.15265/IYS-2016-s006.

[13] World Health Organization and Council for International Organizations of Medical Sciences, *International ethical guidelines for health-related research involving humans*. Geneva: CIOMS, 2017.

[14] World Medical Association, "World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects," *JAMA*, vol. 310, no. 20, pp. 2191–2194, Nov. 2013, doi: 10.1001/jama.2013.281053.

[15] J. Kaye, E. A. Whitley, D. Lund, M. Morrison, H. Teare, and K. Melham, "Dynamic consent: a patient interface for twenty-first century research networks," *Eur J Hum Genet*, vol. 23, no. 2, pp. 141–146, Feb. 2015, doi: 10.1038/ejhg.2014.71.

[16] G. Albanese, J.-P. Calbimonte, M. Schumacher, and D. Calvaresi, "Dynamic consent management for clinical trials via private blockchain technology," *Journal of Ambient Intelligence and Humanized Computing*, Feb. 2020, doi: 10.1007/s12652-020-01761-1.

[17] "OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data - OECD." http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm (accessed Apr. 16, 2020).

[18] C. Grady *et al.*, "Broad Consent for Research With Biological Samples: Workshop Conclusions," *Am J Bioeth*, vol. 15, no. 9, pp. 34–42, 2015, doi: 10.1080/15265161.2015.1062162.

[19] E. Coiera and R. Clarke, "e-Consent: The design and implementation of consumer consent mechanisms in an electronic environment," *J Am Med Inform Assoc*, vol. 11, no. 2, pp. 129–140, 2004, doi: 10.1197/jamia.M1480.

[20] A. Boonstra, A. Versluis, and J. F. J. Vos, "Implementing electronic health records in hospitals: a systematic literature review," *BMC Health Services Research*, vol. 14, no. 1, p. 370, Sep. 2014, doi: 10.1186/1472-6963-14-370.

[21] Y. Pylypchuk, C. Johnson, J. Henry, and D. Ciricean, "Variation in Interoperability among U.S. Non-federal Acute Care Hospitals in 2017," no. 42, p. 15, 2018.

[22] R. Sharma, “BLOCKCHAIN: the magic pill to alleviate the pain points of the healthcare industry?,” p. 25, 2018.

[23] R. L. Richesson and J. Krischer, “Data standards in clinical research: gaps, overlaps, challenges and future directions.,” *J Am Med Inform Assoc*, vol. 14, no. 6, pp. 687–696, Dec. 2007, doi: 10.1197/jamia.M2470.

[24] I. Budin-Ljøsne *et al.*, “Dynamic Consent: a potential solution to some of the challenges of modern biomedical research,” *BMC Medical Ethics*, vol. 18, no. 1, p. 4, Jan. 2017, doi: 10.1186/s12910-016-0162-9.

[25] K. Snow, “About CommonWell,” *CommonWell Health Alliance*. https://www.commonwellalliance.org/about/ (accessed Jul. 02, 2019).

[26] K. Snow, “Use Cases and Specifications,” *CommonWell Health Alliance*. https://www.commonwellalliance.org/connect-to-the-network/use-cases-and-specifications/ (accessed Jul. 02, 2019).

[27] H. van der Linden, D. Kalra, A. Hasman, and J. Talmon, “Inter-organizational future proof EHR systems: A review of the security and privacy related issues,” *International Journal of Medical Informatics*, vol. 78, no. 3, pp. 141–160, Mar. 2009, doi: 10.1016/j.ijmedinf.2008.06.013.

[28] “A Pandora’s Box: The EMR’s Audit Trail,” *Healthcare Litigation Consultants | Denver | EMRDiscovery*. https://www.emrdiscoveryintel.com/single-post/A-Pandoras-Box (accessed Jun. 25, 2019).

[29] T. Walsh and W. M. Miaoulis, “Privacy and Security Audits of Electronic Health Information (2014 update),” *Journal of AHIMA*, vol. 85, no. 3, pp. 54–59, Mar. 2014.

[30] K. Wuyts, R. Scandariato, G. Verhenneman, and W. Joosen, “Integrating Patient Consent in e-Health Access Control,” *IJSSE*, vol. 2, pp. 1–24, Jan. 2011, doi: 10.4018/jsse.2011040101.

[31] M. R. Asghar and G. Russello, “Flexible and Dynamic Consent-Capturing,” in *Open Problems in Network Security*, Berlin, Heidelberg, 2012, pp. 119–131.

[32] M. Benchoufi and P. Ravaud, “Blockchain technology for improving clinical research quality,” *Trials*, vol. 18, no. 1, p. 335, Jul. 2017, doi: 10.1186/s13063-017-2035-z.

[33] K. Rantos, G. Drosatos, A. Kritsas, C. Ilioudis, A. Papanikolaou, and A. P. Filippidis, “A Blockchain-Based Platform for Consent Management of Personal Data Processing in the IoT Ecosystem,” *Security and Communication Networks*, vol. 2019, p. 1431578, Oct. 2019, doi: 10.1155/2019/1431578.

[34] N. Mamo, G. M. Martin, M. Desira, B. Ellul, and J.-P. Ebejer, “Dwarna: a blockchain solution for dynamic consent in biobanking,” *European Journal of Human Genetics*, Dec. 2019, doi: 10.1038/s41431-019-0560-9.

[35] X. Xu *et al.*, “A Taxonomy of Blockchain-Based Systems for Architecture Design,” in *2017 IEEE International Conference on Software Architecture (ICSA)*, Apr. 2017, pp. 243–252, doi: 10.1109/ICSA.2017.33.

[36] K. Peterson, R. Deeduvanu, P. Kanjamala, and K. Boles, “A Blockchain-Based Approach to Health Information Exchange Networks,” p. 10.

[37] Y. Chen, “Blockchain Tokens and the Potential Democratization of Entrepreneurship and Innovation,” *Business Horizons*, vol. 61, pp. 567–575, Jul. 2018, doi: 10.1016/j.bushor.2018.03.006.

[38] W. Viriyasitavat and D. Hoonsopon, “Blockchain characteristics and consensus in modern business processes,” *Journal of Industrial Information Integration*, vol. 13, pp. 32–39, Mar. 2019, doi: 10.1016/j.jii.2018.07.004.

[39] S. Jiang, J. Cao, H. Wu, Y. Yang, M. Ma, and J. He, “BlocHIE: A BLOCkchain-Based Platform for Healthcare Information Exchange,” in *2018 IEEE International Conference on Smart Computing (SMARTCOMP)*, Jun. 2018, pp. 49–56, doi: 10.1109/SMARTCOMP.2018.00073.

[40] Y. Chen, S. Ding, Z. Xu, H. Zheng, and S. Yang, “Blockchain-Based Medical Records Secure Storage and Medical Service Framework,” *J Med Syst*, vol. 43, no. 1, p. 5, Nov. 2018, doi: 10.1007/s10916-018-1121-4.

[41] N. B. Truong, K. Sun, G. M. Lee, and Y. Guo, “GDPR-Compliant Personal Data Management: A Blockchain-Based Solution,” *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1746–1761, 2020, doi: 10.1109/TIFS.2019.2948287.

[42] A. A. Vazirani, O. O’Donoghue, D. Brindley, and E. Meinert, “Implementing Blockchains for Efficient Health Care: Systematic Review,” *Journal of Medical Internet Research*, vol. 21, no. 2, p. e12439, 2019, doi: 10.2196/12439.

[43] S. Tanwar, K. Parekh, and R. Evans, “Blockchain-based electronic healthcare record system for healthcare 4.0 applications,” *Journal of Information Security and Applications*, vol. 50, p. 102407, Feb. 2020, doi: 10.1016/j.jisa.2019.102407.

[44] F. Curbera, D. M. Dias, V. Simonyan, W. A. Yoon, and A. Casella, “Blockchain: An enabler for healthcare and life sciences transformation,” *IBM Journal of Research and Development*, vol. 63, no. 2/3, p. 8:1-8:9, Mar. 2019, doi: 10.1147/JRD.2019.2913622.

[45] D. Tith *et al.*, "Application of Blockchain to Maintaining Patient Records in Electronic Health Record for Enhanced Privacy, Scalability, and Availability," *Healthc Inform Res*, vol. 26, no. 1, pp. 3–12, Jan. 2020.

[46] B. Preneel, "Cryptographic Hash Functions: Theory and Practice," in *Progress in Cryptology - INDOCRYPT 2010*, 2010, pp. 115–117.

[47] S. Hohenberger and Z. Scott, "Lecture 17: Re-encryption," p. 6.

[48] J.-W. Byun and N. Li, "Purpose Based Access Control for Privacy Protection in Relational Database Systems," *The VLDB Journal*, vol. 17, no. 4, pp. 603–619, Jul. 2008, doi: 10.1007/s00778-006-0023-0.

[49] J.-W. Byun, E. Bertino, and N. Li, "Purpose Based Access Control of Complex Data for Privacy Protection," in *Proceedings of the Tenth ACM Symposium on Access Control Models and Technologies*, New York, NY, USA, 2005, pp. 102–110, doi: 10.1145/1063979.1063998.

[50] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering – A systematic literature review," *Information and Software Technology*, vol. 51, no. 1, pp. 7–15, Jan. 2009, doi: 10.1016/j.infsof.2008.09.009.

[51] B. Kitchenham, "Procedures for Performing Systematic Reviews," *Keele, UK, Keele Univ.*, vol. 33, Aug. 2004.

[52] K. T. Win, "A Review of Security of Electronic Health Records," *Health Information Management*, vol. 34, no. 1, pp. 13–18, Mar. 2005, doi: 10.1177/183335830503400105.

[53] J. Fruhlinger, “What is information security? Definition, principles, and jobs,” *CSO Online*, Jan. 17, 2020. https://www.csoonline.com/article/3513899/what-is-information-security-definition-principles-and-jobs.html (accessed Apr. 25, 2020).

[54] S. J. Nass, L. A. Levit, L. O. Gostin, and I. of M. (US) C. on H. R. and the P. of H. I. T. H. P. Rule, *The Value and Importance of Health Information Privacy*. National Academies Press (US), 2009.

[55] B. S. Elger *et al.*, “Strategies for health data exchange for secondary, cross-institutional clinical research,” *Computer Methods and Programs in Biomedicine*, vol. 99, no. 3, pp. 230–251, Sep. 2010, doi: 10.1016/j.cmpb.2009.12.001.

[56] J. L. Fernández-Alemán, I. C. Señor, P. Á. O. Lozoya, and A. Toval, “Security and privacy in electronic health records: A systematic literature review,” *Journal of Biomedical Informatics*, vol. 46, no. 3, pp. 541–562, Jun. 2013, doi: 10.1016/j.jbi.2012.12.003.

[57] “cryptography noun - Definition, pictures, pronunciation and usage notes | Oxford Advanced Learner’s Dictionary at OxfordLearnersDictionaries.com.” https://www.oxfordlearnersdictionaries.com/definition/english/cryptography (accessed Apr. 16, 2020).

[58] A. J. Menezes, S. A. Vanstone, and P. C. V. Oorschot, *Handbook of Applied Cryptography*, 1st ed. USA: CRC Press, Inc., 1996.

[59] S. S. Epp, *Discrete Mathematics with Applications*, 4th ed. USA: Brooks/Cole Publishing Co., 2010.

[60] B. Preneel, “1 CRYPTOGRAPHIC HASH FUNCTIONS : AN OVERVIEW.” https://www.semanticscholar.org/paper/1-CRYPTOGRAPHIC-HASH-FUNCTIONS-%3A-

AN-OVERVIEW-Preneel/f6770d8d53ee44781a6d292cba3595feac8fe41e#citing-papers (accessed Apr. 15, 2020).

[61] J. Zhang and S. Boonkrong, "Dynamic Salt Generating Scheme Using Seeds Warehouse Table Coordinates," in *2015 2nd International Conference on Information Science and Security (ICISS)*, Dec. 2015, pp. 1–6, doi: 10.1109/ICISSEC.2015.7370997.

[62] G. J. Simmons, "Symmetric and Asymmetric Encryption," *ACM Comput. Surv.*, vol. 11, no. 4, pp. 305–330, Dec. 1979, doi: 10.1145/356789.356793.

[63] "Data Encryption Standard," *Wikipedia*. Jun. 05, 2020, Accessed: Jun. 22, 2020. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Data_Encryption_Standard&oldid=960862279.

[64] "Advanced Encryption Standard," *Wikipedia*. May 29, 2020, Accessed: Jun. 22, 2020. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Advanced_Encryption_Standard&oldid=959595933.

[65] "What is Feistel Block Cipher? Definition, Encryption and Decryption Structure - Binary Terms." https://binaryterms.com/feistel-block-cipher.html (accessed Jun. 22, 2020).

[66] "RSA (cryptosystem)," *Wikipedia*. May 29, 2020, Accessed: Jun. 02, 2020. [Online]. Available: https://en.wikipedia.org/w/index.php?title=RSA_(cryptosystem)&oldid=959573567.

[67] Xin Zhou and Xiaofei Tang, "Research and implementation of RSA algorithm for encryption and decryption," in *Proceedings of 2011 6th International Forum on Strategic Technology*, Aug. 2011, vol. 2, pp. 1118–1121, doi: 10.1109/IFOST.2011.6021216.

[68] “Elliptic-curve cryptography,” *Wikipedia*. May 30, 2020, Accessed: Jun. 03, 2020. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Elliptic-curve_cryptography&oldid=959862207.

[69] M. Amara and A. Siad, “Elliptic Curve Cryptography and its applications,” in *International Workshop on Systems, Signal Processing and their Applications, WOSSPA*, May 2011, pp. 247–250, doi: 10.1109/WOSSPA.2011.5931464.

[70] P. Kuppuswamy and S. Q. Y. Al-Khalidi, “Hybrid Encryption/Decryption Technique Using New Public Key and Symmetric Key Algorithm,” *Int. J. Inf. Comput. Secur.*, vol. 6, no. 4, pp. 372–382, Mar. 2014, doi: 10.1504/IJICS.2014.068103.

[71] R. K. Gupta and P. Singh, *A New Way to Design and Implementation of Hybrid Crypto System for Security of the Information in Public Network*. .

[72] V. Kapoor and R. Yadav, “A Hybrid Cryptography Technique for Improving Network Security,” *International Journal of Computer Applications*, vol. 141, pp. 25–30, May 2016, doi: 10.5120/ijca2016909863.

[73] D.-D. Salama, H. Abd elkader, and M. M. Hadhoud, “Performance Evaluation of Symmetric Encryption Algorithms,” *Communications of the IBIMA*, vol. 10, Jan. 2009.

[74] A. Albarqi, E. Alzaid, F. Ghamdi, S. Asiri, and J. Kar, “Public Key Infrastructure: A Survey,” *Journal of Information Security*, vol. 06, pp. 31–37, Jan. 2015, doi: 10.4236/jis.2015.61004.

[75] W. Diffie and M. Hellman, “New Directions in Cryptography,” *IEEE Trans. Inf. Theor.*, vol. 22, no. 6, pp. 644–654, Sep. 2006, doi: 10.1109/TIT.1976.1055638.

[76] L. M. Kohnfelder, “Towards a Practical Public-key Cryptosystem,” p. 54.

[77] “X.500 : Information technology - Open Systems Interconnection - The Directory: Overview of concepts, models and services.” https://www.itu.int/rec/T-REC-X.500 (accessed Apr. 27, 2020).

[78] W. Ford, “Public-key infrastructure interoperation,” in *1998 IEEE Aerospace Conference Proceedings (Cat. No.98TH8339)*, Mar. 1998, vol. 4, pp. 329–333 vol.4, doi: 10.1109/AERO.1998.682203.

[79] S. Chokhani, “Toward a national public key infrastructure,” *IEEE Communications Magazine*, vol. 32, no. 9, pp. 70–74, Sep. 1994, doi: 10.1109/35.312846.

[80] Ke-feng Wang and Zhi-hong Zhang, “Design and implementation of a safe Public Key Infrastructure,” in *2010 International Conference on Future Information Technology and Management Engineering*, Oct. 2010, vol. 2, pp. 298–301, doi: 10.1109/FITME.2010.5656308.

[81] A. Fongen, “Optimization of a Public Key Infrastructure,” in *2011 - MILCOM 2011 Military Communications Conference*, Nov. 2011, pp. 1440–1447, doi: 10.1109/MILCOM.2011.6127509.

[82] P. Wing and B. O’Higgins, “Using public-key infrastructures for security and risk management,” *IEEE Communications Magazine*, vol. 37, no. 9, pp. 71–73, Sep. 1999, doi: 10.1109/35.790867.

[83] W. T. Polk, N. E. Hastings, and A. Malpani, “Public key infrastructures that satisfy security goals,” *IEEE Internet Computing*, vol. 7, no. 4, pp. 60–67, Aug. 2003, doi: 10.1109/MIC.2003.1215661.

[84] "ElGamal encryption," *Wikipedia*. Apr. 09, 2020, Accessed: Jun. 16, 2020. [Online]. Available: https://en.wikipedia.org/w/index.php?title=ElGamal_encryption&oldid=949973016.

[85] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman, "Role-Based Access Control Models," *Computer*, vol. 29, no. 2, pp. 38–47, Feb. 1996, doi: 10.1109/2.485845.

[86] D. F. Ferraiolo, R. Sandhu, S. Gavrila, D. R. Kuhn, and R. Chandramouli, "Proposed NIST Standard for Role-Based Access Control," *ACM Trans. Inf. Syst. Secur.*, vol. 4, no. 3, pp. 224–274, Aug. 2001, doi: 10.1145/501978.501980.

[87] C. T. Hu *et al.*, "Guide to Attribute Based Access Control (ABAC) Definition and Considerations," Feb. 2019, Accessed: Apr. 22, 2020. [Online]. Available: https://www.nist.gov/publications/guide-attribute-based-access-control-abac-definition-and-considerations-1.

[88] R. Zhang, A. George, J. Kim, V. Johnson, and B. Ramesh, "Benefits of Blockchain Initiatives for Value-Based Care: Proposed Framework," *Journal of Medical Internet Research*, vol. 21, no. 9, p. e13595, 2019, doi: 10.2196/13595.

[89] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," p. 9.

[90] V. Buterin, "A NEXT GENERATION SMART CONTRACT & DECENTRALIZED APPLICATION PLATFORM," p. 36.

[91] M. Andoni *et al.*, "Blockchain technology in the energy sector: A systematic review of challenges and opportunities," *Renewable and Sustainable Energy Reviews*, vol. 100, pp. 143–174, Feb. 2019, doi: 10.1016/j.rser.2018.10.014.

[92] “Introduction — hyperledger-fabricdocs master documentation.” https://hyperledger-fabric.readthedocs.io/en/latest/blockchain.html (accessed Jun. 25, 2019).

[93] “Instantly Move Money to All Corners of the World | Ripple.” https://ripple.com/ (accessed Apr. 20, 2020).

[94] “Stellar - an open network for money.” https://www.stellar.org/?locale=en (accessed Apr. 20, 2020).

[95] “The difference between public and private blockchain - Blockchain Pulse: IBM Blockchain Blog.” https://www.ibm.com/blogs/blockchain/2017/05/the-difference-between-public-and-private-blockchain/ (accessed Apr. 20, 2020).

[96] “Proof of Stake Explained,” *Binance Academy*. https://academy.binance.com/blockchain/proof-of-stake-explained (accessed Jul. 26, 2020).

[97] “practical Byzantine Fault Tolerance(pBFT) - GeeksforGeeks.” https://www.geeksforgeeks.org/practical-byzantine-fault-tolerancepbft/ (accessed Jul. 26, 2020).

[98] S. Muftic, I. Sanchez, E. D. Jrc, L. Beslay, and E. D. Jrc, “Overview and Analysis of the Concept and Applications of Virtual Currencies,” p. 52.

[99] D. G. Wood, “ETHEREUM: A SECURE DECENTRALISED GENERALISED TRANSACTION LEDGER,” p. 32.

[100] N. Atzei, M. Bartoletti, and T. Cimoli, *A Survey of Attacks on Ethereum Smart Contracts (SoK)*. 2017, p. 186.

[101] H. Watanabe, S. Fujimura, A. Nakadaira, Y. Miyazaki, A. Akutsu, and J. Kishigami, "Blockchain contract: Securing a blockchain applied to smart contracts," in *2016 IEEE International Conference on Consumer Electronics (ICCE)*, Jan. 2016, pp. 467–468, doi: 10.1109/ICCE.2016.7430693.

[102] S. Ølnes, "Beyond Bitcoin Enabling Smart Government Using Blockchain Technology," in *Electronic Government*, Cham, 2016, pp. 253–264.

[103] J. Cheng, N. Lee, C. Chi, and Y. Chen, "Blockchain and smart contract for digital certificate," in *2018 IEEE International Conference on Applied System Invention (ICASI)*, Apr. 2018, pp. 1046–1051, doi: 10.1109/ICASI.2018.8394455.

[104] G. G. Dagher, J. Mohler, M. Milojkovic, and P. B. Marella, "Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology," *Sustainable Cities and Society*, vol. 39, pp. 283–297, May 2018, doi: 10.1016/j.scs.2018.02.014.

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

[106] T.-T. Kuo, H.-E. Kim, and L. Ohno-Machado, "Blockchain distributed ledger technologies for biomedical and health care applications," *Journal of the American Medical Informatics Association*, vol. 24, no. 6, pp. 1211–1220, Sep. 2017, doi: 10.1093/jamia/ocx068.

[107] A. Manzoor, M. Liyanage, A. Braeken, S. S. Kanhere, and M. Ylianttila, *Blockchain based Proxy Re-Encryption Scheme for Secure IoT Data Sharing*. 2018.

[108] F. Casino, T. K. Dasaklis, and C. Patsakis, “A systematic literature review of blockchain-based applications: Current status, classification and open issues,” *Telematics and Informatics*, vol. 36, pp. 55–81, Mar. 2019, doi: 10.1016/j.tele.2018.11.006.

[109] H. S. Chen, J. T. Jarrell, K. A. Carpenter, D. S. Cohen, and X. Huang, “Blockchain in Healthcare: A Patient-Centered Model,” *Biomed J Sci Tech Res*, vol. 20, no. 3, pp. 15017–15022, 2019.

[110] M. A. Engelhardt, “Hitching Healthcare to the Chain: An Introduction to Blockchain Technology in the Healthcare Sector,” *TIM Review*, vol. 7, no. 10, pp. 22–34, Oct. 2017, doi: 10.22215/timreview/1111.

[111] C. N. Mead, “Data interchange standards in healthcare IT--computable semantic interoperability: now possible but still difficult, do I really need a better mousetrap?,” *J Healthc Inf Manag*, vol. 20, no. 1, pp. 71–78, Winter 2006.

[112] I. Olaronke, A. Soriyan, I. Gambo, and J. Olaleke, “Interoperability in Healthcare: Benefits, Challenges and Resolutions,” *International Journal of Innovation and Applied Studies*, vol. 3, pp. 2028–9324, Apr. 2013.

[113] I. Al Ridhawi, M. Aloqaily, Y. Kotb, Y. Al Ridhawi, and Y. Jararweh, “A collaborative mobile edge computing and user solution for service composition in 5G systems,” *Transactions on Emerging Telecommunications Technologies*, vol. 29, no. 11, p. e3446, Nov. 2018, doi: 10.1002/ett.3446.

[114] I. Al Ridhawi, M. Aloqaily, B. Kantarci, Y. Jararweh, and H. T. Mouftah, “A Continuous Diversified Vehicular Cloud Service Availability Framework for Smart Cities,” *Computer Networks*, vol. 145, Sep. 2018, doi: 10.1016/j.comnet.2018.08.023.

[115] L. Cardoso, F. Marins, F. Portela, M. Santos, A. Abelha, and J. Machado, "The next generation of interoperability agents in healthcare," *Int J Environ Res Public Health*, vol. 11, no. 5, pp. 5349–5371, May 2014, doi: 10.3390/ijerph110505349.

[116] W. J. Gordon and C. Catalini, "Blockchain Technology for Healthcare: Facilitating the Transition to Patient-Driven Interoperability.," *Comput Struct Biotechnol J*, vol. 16, pp. 224–230, 2018, doi: 10.1016/j.csbj.2018.06.003.

[117] P. Zhang, J. White, D. Schmidt, and G. Lenz, "Applying Software Patterns to Address Interoperability in Blockchain-based Healthcare Apps," Jun. 2017.

[118] P. Zhang, J. White, D. C. Schmidt, G. Lenz, and S. T. Rosenbloom, "FHIRChain: Applying Blockchain to Securely and Scalably Share Clinical Data," *Comput Struct Biotechnol J*, vol. 16, pp. 267–278, Jul. 2018, doi: 10.1016/j.csbj.2018.07.004.

[119] T. K. Mackey *et al.*, "'Fit-for-purpose?' – challenges and opportunities for applications of blockchain technology in the future of healthcare," *BMC Medicine*, vol. 17, no. 1, p. 68, Mar. 2019, doi: 10.1186/s12916-019-1296-7.

[120] A. Azaria, A. Ekblaw, T. Vieira, and A. Lippman, "MedRec: Using Blockchain for Medical Data Access and Permission Management," in *2016 2nd International Conference on Open and Big Data (OBD)*, Aug. 2016, pp. 25–30, doi: 10.1109/OBD.2016.11.

[121] A. Dubovitskaya, Z. Xu, S. Ryu, M. Schumacher, and F. Wang, "Secure and Trustable Electronic Medical Records Sharing using Blockchain," *AMIA Annu Symp Proc*, vol. 2017, pp. 650–659, Apr. 2018.

[122] A. Gropper, "Powering the Physician-Patient Relationship with HIE of One Blockchain Health IT," 2016.

[123] S. Sayeed and H. Marco-Gisbert, “Assessing Blockchain Consensus and Security Mechanisms against the 51% Attack,” *Applied Sciences*, vol. 9, p. 1788, Apr. 2019, doi: 10.3390/app9091788.

[124] N. Mamo, G. M. Martin, M. Desira, B. Ellul, and J.-P. Ebejer, “Dwarna: a blockchain solution for dynamic consent in biobanking,” *Eur J Hum Genet*, Dec. 2019, doi: 10.1038/s41431-019-0560-9.

[125] “Proxy server,” *Wikipedia*. Jun. 09, 2020, Accessed: Jun. 10, 2020. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Proxy_server&oldid=961702506.

[126] M. Blaze, G. Bleumer, and M. Strauss, *Divertible Protocols and Atomic Proxy Cryptography*, vol. 1403. 1998, p. 144.

[127] “Empowering App Development for Developers | Docker.” https://www.docker.com/ (accessed May 06, 2020).

[128] G. Ateniese, K. Fu, M. Green, and S. Hohenberger, “Improved Proxy Re-encryption Schemes with Applications to Secure Distributed Storage,” *ACM Trans. Inf. Syst. Secur.*, vol. 9, no. 1, pp. 1–30, Feb. 2006, doi: 10.1145/1127345.1127346.

[129] “MedRec.” https://medrec.media.mit.edu/technical/ (accessed Sep. 04, 2019).

[130] “purpose noun - Definition, pictures, pronunciation and usage notes | Oxford Advanced American Dictionary at OxfordLearnersDictionaries.com.” https://www.oxfordlearnersdictionaries.com/definition/american_english/purpose (accessed Apr. 23, 2020).

[131] "wish_2 noun - Definition, pictures, pronunciation and usage notes | Oxford Advanced American Dictionary at OxfordLearnersDictionaries.com." https://www.oxfordlearnersdictionaries.com/definition/american_english/wish_2 (accessed Apr. 24, 2020).

[132] "What is valid consent?," Feb. 24, 2020. https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/consent/what-is-valid-consent/ (accessed Apr. 24, 2020).

[133] Md. Enamul Kabir, H. Wang, and E. Bertino, "A conditional purpose-based access control model with dynamic roles," *Expert Systems with Applications*, vol. 38, no. 3, pp. 1482–1489, Mar. 2011, doi: 10.1016/j.eswa.2010.07.057.

[134] "Introduction — hyperledger-fabricdocs master documentation." https://hyperledger-fabric.readthedocs.io/en/latest/whatis.html (accessed Jan. 27, 2020).

[135] "Art. 17 GDPR - Right to erasure ('right to be forgotten')," *GDPR.eu*, Nov. 14, 2018. https://gdpr.eu/article-17-right-to-be-forgotten/ (accessed Apr. 08, 2020).

[136] P. Gauravaram, "Security Analysis of salt||password Hashes," in *2012 International Conference on Advanced Computer Science Applications and Technologies (ACSAT)*, Nov. 2012, pp. 25–30, doi: 10.1109/ACSAT.2012.49.