APPLYING BLOCKCHAIN AND DISTRIBUTED LEDGER TO IMPROVE THE ACCESSIBILITY, SECURITY AND PRIVACY OF ELECTRONIC HEALTH DATA

A Thesis Presented

by

Anania Mesfin Sileshi

to

The Faculty of Informatics

of

St. Mary's University

In Partial Fulfillment of the Requirements
for the Degree of Master of Science
in
Computer Science

July, 2020

ACCEPTANCE

APPLYING BLOCKCHAIN AND DISTRIBUTED LEDGER TO IMPROVE THE ACCESSIBILITY, SECURITY AND PRIVACY OF ELECTRONIC HEALTH DATA

By

Anania Mesfin Sileshi

Accepted by the Faculty of Informatics, St. Mary's University, in partial fulfillment of the requirements for the degree of Master of Science in Computer Science

Thesis Examination Committee:

_____

Internal Examiner

_____

External Examiner

_____

Dean, Faculty of Informatics

July 2020

DECLARATION

I, the undersigned, declare that this thesis work is my original work, has not been presented for a degree in this or any other universities, and all sources of materials used for the thesis work have been duly acknowledged.

_____
Full Name of Student

_____
Signature

Addis Ababa
Ethiopia

This thesis has been submitted for examination with my approval as advisor.

_____
Full Name of Advisor

_____
Signature

Addis Ababa
Ethiopia

July 2020

ACKNOWLEDGEMENTS

First and foremost, I would like to thank God and Virgin Mary for their valuable support to accomplish this thesis and being with me in all directions of my life. I give heartfelt thanks to my dearest parents for their love, support and patience during my thesis work and for being the source of motivation in my life.

I would like to give special thanks to my advisor, Asrat Mulatu (PhD), for the time and energy he has invested on me throughout the work with the thesis. This work would not have been possible without his support, guidance, invaluable feedback and advice on my thesis. In fact, I'm very grateful to have had his trust in my ability and for inspiring me to take part in the exciting field of Block chain and distributed ledger technology which is a very recent technology.

I would like to take this opportunity to express my gratitude to Freadam Moges Eshete (MSc degree on Digital Currency) from University of Cyprus for commenting on this study work and giving me a tremendous advice the whole time.

And finally It is my pleasure to express my deepest gratitude to organization, doctors and nurses at Tikur Ambessa hospital, Bethel Hospital, Hallelujah Hospital, Kadisco Hospital, Torhayloch Hospital, Amen Hospital, Yehuleshet Higher Clinic, Dr. Khalid and his family Higher clinic and all my friends and my classmates who have helped and supported me in various aspects of conducting the required materials.

ABSTRACT

As the concept Transparency is key to information exchange amongst systems running on different platforms. And the healthcare is from one of the key industries that seek transparency, confidentiality and integrity of stored data. Beyond the security concerns of the E-health industry interoperability is also major issue and in current trend systems operate within the hospital domain cannot communicate with other health facilities in order to share information. This makes it hard for health practitioners to share patient data and access medical history which facilitate evidence based decision-making at all levels of the system especially at the point of origin. This research sought to investigate why hospitals, that are using he manual system, and developers, who try to automate, faced with challenges to standardize existing systems, to integrate legacy with modern platforms and to achieve interoperability. To this end, interviews were conducted to get general information on sharing of patient data. Findings indicate that current systems do not allow sharing of health data. This thesis work, therefore, developed a platform that uses Blockchain technology and distributed file systems to integrate the existing health information systems so as to facilitate a fast and secure data exchange, facilitating interoperability at the end. Based on the performance evaluation made using primary end users revealed that the framework prototype allowed patients to port data and share it with different doctors on demand. Moreover, it ensured that a permanent reference of the data is stored in a distributed ledger that is sharable and interoperable using different application frontends.

Keywords: Blockchain technology, distributed ledger, distributed framework, eHealth, Health Data.

Table of Contents

List of Acronyms

EHR ----------Electronic Health Recording

HDR ----------Health Data Recording

DLT ----------Distributed Ledger Technology

RAID----------Redundant Array of Independent Disks

DL ----------Distributed Ledger

SME ----------Small and Medium Enterprise

P2P ----------Peer-to-Peer

HLA ----------High Level Architecture

EVM----------Ethereum Virtual Machine

DApp ----------Decentralized Application

EHRs ----------Electronic Health Records

AES----------Advanced Encryption Standards

CINA----------Confidentiality, Integrity, Non-repudiation, Authentication

BGP----------Byzantine Generals Problem

BFT----------Byzantine Fault Tolerance

FOD----------Function Oriented Design

List of Tables

List of Figures

CHAPTER ONE

INTRODUCTION

1.1 Background

Electronic Health Data Recording (EHR) were never designed to manage multi-institutional, lifetime medical records. Patients leave data scattered across various organizations as life events take them away from one provider's data to another. In doing so they lose easy access to past data, as the provider, not the patient, generally retains primary stewardship (either through explicit legal means in over 21 states, or through default arrangements in the process of providing care) [1]. Through the rules that are internationally set by the HIPPA patients are eligible to get their own medical record on hand within 24hours of request. Beyond the time delay, record maintenance can prove quite challenging to initiate as patients are rarely encouraged and seldom enabled to review their full record [1,2]. Patients thus interact with records in a fractured manner that reflects the nature of how these records are managed.

Due to the growing complexity of health care, patient data are more and more in demand for purposes such as research, education, post marketing surveillance, quality assessment and outcome analysis. Many of these records need patient data to be available in a structured electronic format. The rapid advances in computer technology, which allow patient data to be organized analyzed and shared, majority of health care facilities use the paper based data recording system which is old fashioned.

Apparently most physicians still perceive the paper records are still more suitable for their task than present day computerized versions. Both the short comings and strength of paper based HDR's have been identified and it proves difficult to design a computerized HDR that exploits the strengths of computers without losing the advantages of the paper chart. Basically the structure of HDR is an area of high interest since structure determines how physicians or other health workers take care of the input.

1.2 Motivation

Current practice in Medical Data recording in Ethiopia is very poor, in saying poor, most of the health care institutions in Ethiopia use the paper based data handling mechanism and this

mechanism has lots of drawbacks, some of them are.

- Time: costs more time searching for medical records if they are in hard copies.
- Interoperability: patients can't get continuous medical treatments across different medical treatment organizations.
- Consistency: data consistency is also another problem when it comes to medical record assessment, different medical institutions handle records differently.
- Security: medical records in an organizations may face problems both human made or natural, records might get stolen or medical histories might burn due to some reason of fire, or in a worst case scenario an earth quake might collapse the medical institution with all the records.

This paper tried to show a model that could potentially resolve this problem and get us to a better data handling environment, some of the solutions that could be addressed are:

(i) Strengthen Patient Safety

Check & monitor current **medical prescription** & medication lists.

Reduces risk of forgetting/missing important details.

(ii) Continuity & Legibility

Avoid mistakes caused by illegible writing

Access Patient information such as vacations, absences

(iii) Flexibility

Streamline interoffice communications

Access data from anywhere at anytime

(iv) Improves Patient Health Outcomes

Deliver chronic disease management

Access Lab results

Generate reminders & decision support for physicians

Efficient Administrative task

1.3 Statement of the Problem

As technology is getting further and further and current the electronic world takes over everything information plays a big role and with it our universe is facing issues like security, privacy and confidentiality, where any data that is useful is considered as information Patients medical

information is also an area that needs focus and need lots of work. As Ministry of Health is the biggest health data facilitation institution in Ethiopia it has a mandate of controlling as well as protecting any and every medical information in medical facilities. Talking about medical records management institutions often are use manual medical records management or web based data management which has lots and lots of vulnerabilities. even if privacy concerns could be addressed with the system based approach, there is no broad consensus around the specific technical infrastructure needed to support such a task. Finally Medical data should be possessed, operated, and allowed to be utilized by data subjects other than hospitals. This is a key concept of patient-centered interoperability could not be solved up until now.

1.4 Research Question

The researcher put four major research questions that could help in organizing the rest of the study.

- To what extent Ethiopian health care facilities are taking care of patients medical history?
- How to achieve interoperability of medical data between medical institutions?
- How Blockchain can support EHR process within the medical sector?
- How to let a patient access his/her medical information anywhere anytime.

1.5 Objective of the study

1.5.1 General Objectives

The main objective of this study is to use the Blockchain and ensure security of medical information. Ensuring the Transparency, integrity and confidentiality of medical records are critical aspects in security to maintain privacy and immutability of data, the last but not the list using Blockchain technology interoperability in electronic health data recording.

1.5.2 Specific Objectives

- To review literature on related research works in order to have an understanding on concepts, principles and technologies of Distributed Ledgers in case of Blockchain.
- To extract the domain knowledge, which will be, used in identification of electronic health data recording.
- To analyze basic functionality of Blockchain systems in E-Health system domain structured and gain new knowledge.

1.6 Methodology

The study will follow design science research approach to discuss What, How, and Why at each step of the research process. Since it allows the development of model based on the existing problem domain and helps to add knowledge regarding digitally recording of medical records. The fundamental principle of design science research is that knowledge and understanding of a design problem and its solution are acquired in the building and application of an artifact. (Alan Hevner, 2010).

The general model used in this study is the model provided by Omar Valdez-de-Leon. It is used as the general guideline to further identify specific criterion questionnaires that are aligned with the given seven dimensions.(Omar Valdez de Leon 2016).

1.7 Scope of the Study

The scope of this research is limited to the retrieval of patient data and access granting between two parties (Nodes) on the system for instance patient and doctor, patient and pharmacist. Cyber Threats facing medical organizations include; Distributed Denial of Service (DDoS) attacks, targeted attacks; and also natural disaster are the main causes of data loss.

The study is limited to develop a prototype of EHR management platform for the purpose of securing medical records the best way while making it available for everybody because every actor in the system is treated as a node.

1.8 Significance of the study

The study helps in managing of the interoperability of medical data since the structure of the data has to be the same across all medical organizations, reducing the time spent in managing medical data in different medical organizations. The most recent and secure data storage place is the cloud and organizations are satisfied not only the security of cloud storage but also the storage capability of cloud based storages but the Blockchain technology can solve the risk management and storage demand of customers the very easy way. The developed prototype of Blockchain based EHR managing platform could save institutions from serious damages to their data and make them feel safe when it comes to their data because the data is spread across all the organizations and patients so that patients will have better satisfaction because they can access their medical history anywhere

anytime on the go, increase their reliability on privacy issues and inconveniences, and will have better reliance on the company services provided by the organizations. This research can serve as a base for future researchers interested on this area. This study also initiates other organizations to turn their face to intelligence Blockchain based data monitoring environment.

1.9 Organization of the Thesis

The rest of this thesis report is organized as follows. Chapter 2 presents the review of related literatures for the understanding of basic concepts related to Blockchain and Distributed ledger implementations, and comparison of the most common trends while working on Blockchain based applications, data handling capability of the Blockchain systems. Moreover, it discusses related research works for assessing the state of the art Blockchain applications. Chapter 3 presents the analysis made for the design of a multi-script text editing. In this chapter, design alternatives, objectives and decisions will be presented before modeling the text editing that integrates the Latin and Ethiopic scripts and finally the proposed model is presented.

The next chapter, Chapter 4, assesses the different implementation alternatives, and forwards a recommendation on how the developer can implement the designed model in each alternative. It also presents the characteristics of different handheld platforms. A discussion on how a font is designed also presented in this chapter. Chapter 5 presents experimental methods and results obtained. Finally, the thesis report will be concluded by forwarding conclusion and recommendations in chapter 6 of the document.

CHAPTER TWO

Literature Review

2.1 Electronic Health Data Recording

As discussed on chapter one the data management of health care records is pretty backward and the interoperability challenges between different provider and hospital systems pose additional barriers to effective data sharing. This lack of coordinated data management and exchange means health records are fragmented, rather than cohesive [3]. Patients and providers may face significant hurdles in initiating data retrieval and sharing due to economic incentives that encourage "health information blocking." A recent ONC report details several examples on this topic, namely health IT developers interfering with the flow of data by charging exorbitant prices for data exchange interfaces [4].

When designing new systems to overcome these barriers, we must prioritize patient agency. Patients benefit from a holistic, transparent picture of their medical history [3]. This proves crucial in establishing trust and continued participation in the medical system, as patients that doubt the confidentiality of their records may abstain from full, honest disclosures or even avoid treatment. In the age of online banking and social media, patients are increasingly willing, able and desirous of managing their data on the web and on the go [3]. However, proposed systems must also recognize that not all provider records can or should be made available to patients (i.e. provider psychotherapy notes, or physician intellectual property), and should remain flexible regarding such record-onboarding exceptions [5].

Medical records prove critical for research. The ONC's report emphasizes that biomedical and public health researchers "require the ability to analyze information from many sources in order to identify public health risks, develop new treatments and cures, and enable precision medicine" [4]. Though some data trickles through to researchers from clinical studies, surveys and teaching hospitals, we note a growing interest among patients, care providers and regulatory bodies to responsibly share more data, and thus enable better care for others [4].

In this work, Blockchain technology is applied to EHRs. Prototype is built on distributed ledger protocol originally associated with Bitcoin [8]. The Blockchain uses public key cryptography to create an append-only, immutable, timestamped chain of content. Copies of the Blockchain are

distributed on each participating node in the network. The Proof of Work algorithm used to secure the content from tampering depends on a "trustless" model, where individual nodes must compete to solve computationally intensive "puzzles" (hashing exercises) before the next block of content can be appended to the chain. These worker nodes are known as "miners," and the work required of miners to append blocks ensures that it is difficult to rewrite history on the Blockchain.

2.2 The potential of Blockchain in managing healthcare data

The most important thing that makes the use of blockchain revolutionary in healthcare is the lack of a central administrator. Why? Because a database is still a tangible thing- consisting of bits and bytes. If the contents of the database are stored in the physical memory of a particular system, anyone who has access to that system could corrupt the data within.

With blockchain, all the users are in control of all their information and transactions. Plus, there will be no need for a central administrator- eliminated by clever cryptography. Since healthcare deals with confidential patient information and requires quick access to information, blockchain can streamline these medical records and enable their sharing in a secure way. Blockchain, in a single go, offers access security, scalability, and data privacy.

In case of Ethiopia in most health care facilities, patient history is managed by data clerks meaning the manual way and anyone who has access to those physical health data recording ledgers has the power to manipulate the data or the worst case scenario destroy it. The same thing works for current E health systems that are applicable in hospitals which have backends and databases for data storage and there will be a person to manage and migrate and backup the system who is the system admin. Who also has access to all the records in the system. So one way or the other the security or the Privacy of patients medical history is in danger.

2.3 Peer-to-Peer Networking

A peer-to-peer (P2P) network is a type of network connection where two or more PCs are connected and share resources without going through a separate server computer. In preparing a P2P network there is one thing to keep in mind which is none of the computers has to be a client or a slave instead a P2P network can be an ad hoc connection a couple of computers connected via

a Universal Serial Bus or an Ethernet connection or a local ad hoc network to transfer files. A P2P network also can be a permanent infrastructure that links a half-dozen computers in a small office over copper wires.

A P2P network can be a network on a much grander scale in which special protocols and applications set up direct relationships among users over the Internet.

The initial use of P2P networks in business followed the deployment in the early 1980s of free-standing PCs. In contrast to the mini mainframes of the day, such as the VS system from Wang Laboratories Inc., which served up word processing and other applications to dumb terminals from a central computer and stored files on a central hard drive, the then-new PCs had self-contained hard drives and built-in CPUs. The smart boxes also had onboard applications, which meant they could be deployed to desktops and be useful without an umbilical cord linking them to a mainframe.

2.4 Interoperability

2.4.1 What is Interoperability

In broad terms, interoperability is the ability of different information and communications technology systems and software applications to communicate, to exchange data accurately, effectively, and consistently, and to use the information that has been exchanged [8]. Data interoperability is the ability to correctly interpret data across systems or organizational boundaries [9].

The key points are illustrated below in Figure 2.1. In the scenario below, it is assumed that the people on the left have information needed by the people on the right, and that data in one system is accessible to the other. Hence, interoperability will only be achieved if the receiving system and users properly understand the meaning of information they receive and they are able to use this information [10].

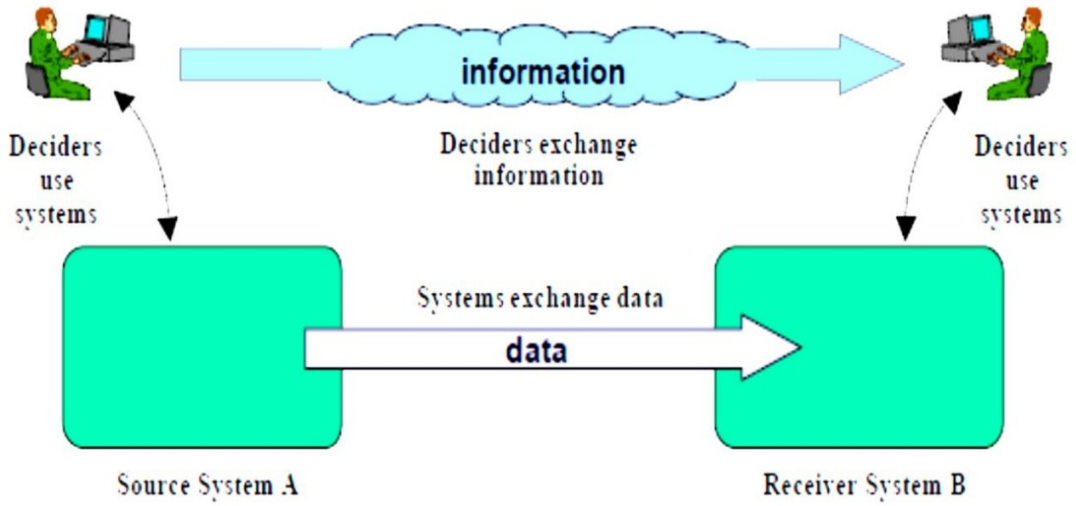


Fig 2.1 Concept of Interoperability

In general, there are seven basic levels of different levels of interoperability [11]. These levels include:

- Level 0 or No Interoperability: This is usually characterized by stand-alone systems which have no interoperability.
- Level 1 or Technical Interoperability: This level of interoperability involves the use of a communication protocol for the exchange of data between systems. Technical interoperability establishes harmonization at the plug and play, signal and protocol level.
- Level2 or Syntactic interoperability: This is the ability of two or more systems to exchange data and services using a common interoperability protocol such as the High Level Architecture (HLA).
- Level3 or Semantic Interoperability: Semantic interoperability refers to the ability of two or more systems to automatically interpret the information exchanged meaningfully and accurately in order to produce useful results as defined by the end users of the systems [12]. Semantic interoperability is also used in a more general sense to refer to the ability of two or more systems to exchange information with an unambiguous and shared meaning [13]. Semantic interoperability implies that the precise meaning of the exchanged information is understood by the communicating systems. Hence, the systems are able to recognize and process semantically equivalent information homogeneously, even if their instances are heterogeneously represented, that is, if they are differently structured, and/or using different

terminology or different natural language [47]. Semantic interoperability can thus be said to be distinct from the other levels of interoperability because it ensures that the receiving system understands the meaning of the exchange information, even when the algorithms used by the receiving system are unknown to the sending system.

- Level4 or Pragmatic Interoperability: This level of interoperability is achieved when the interoperating systems are aware of the methods and procedures that each other are employing. This implies that the use of the data or the context of its application is understood by the participating systems.
- Level5 or Dynamic Interoperability: Two or more systems are said to have attained dynamic interoperability when they are able to comprehend the state changes that occur in the assumptions and constraints that they are making over time, and they are able to take advantage of those changes.
- Level6 or Conceptual Interoperability: Conceptual interoperability is reached if the assumptions and constraints of the meaningful abstraction of reality are aligned.

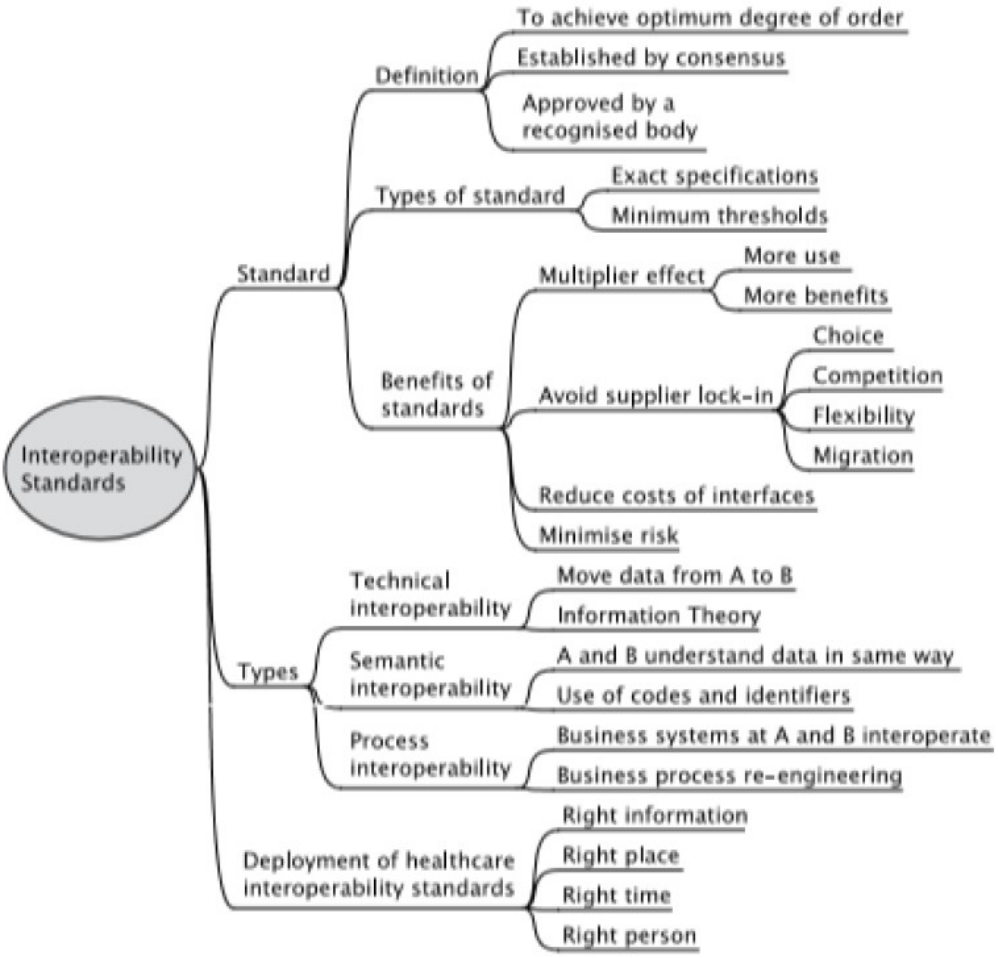


Figure 2.2. Interoperability standards (Adopted from Benson 2010)

2.3.2 What is Interoperability in Healthcare

Das and mhapatra regards e-government as a complex context because it has to deal with policy, legal, politics and sociocultural issues. They identified legal and political factors to be among other things influencing interoperability of systems.

According to ISO(2004)a standard is document established by consensus and approved by recognized body that provides for common and repeated use, rules, guidelines or characteristics for activities or their results, aimed at the achievement of the optimum degree of order in given context. Several institutions in the world have developed standards around medical information exchange, medical terminologies and electronic health record systems processes. Some of the standards have been tailored to cover the security of health information systems (Ministry of Health, 2010).

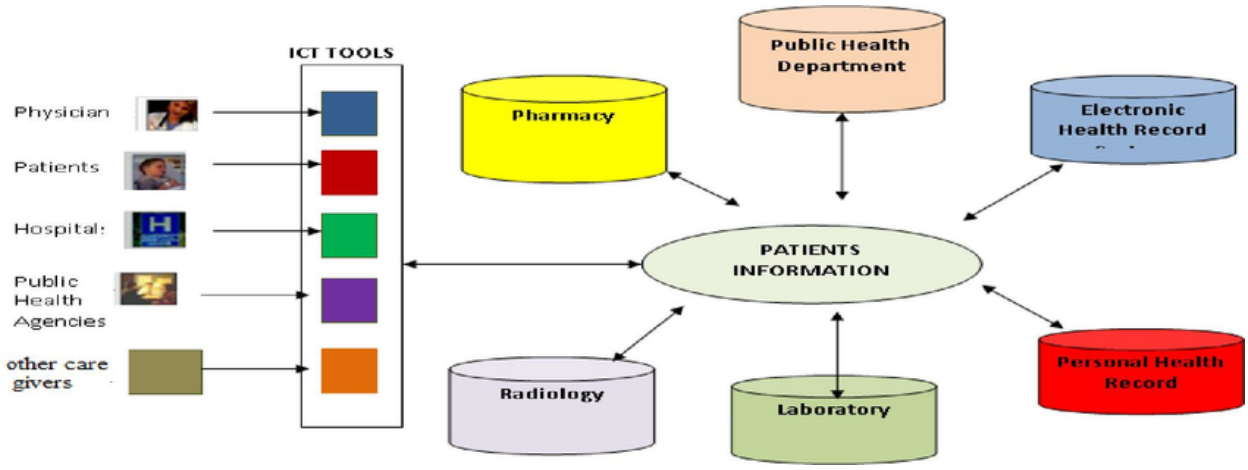


Fig 2.3 Healthcare Interoperability

2.3.3 Barriers that affect the interoperability of Healthcare

According to a report prepared by the ONC there are six overarching barriers limiting the electronic exchange of health information between hospitals today.

There was a 22-page report by the ONC that show nationwide trends in health information exchange in 2018, 2019, including the adoption of EHRs and other technologies that support electronic access to patient information.

"HHS is committed to the use of health IT to support the free flow of health information for patients, healthcare providers and payers as well as to promote competition in healthcare markets," the report reads.

Six challenges inhibiting electronic data exchange in healthcare, as described by the ONC:

1. Technical barriers. "This limit interoperability through

> Ex. a lack of standards development, data quality, and patient and healthcare provider data matching."

2. Financial barriers. "These relate to the costs of developing, implementing and optimizing health IT to meet frequently changing requirements of healthcare programs," including lack of incentives for sharing information and need for business models for secondary uses of data.

3. Trust barriers. "Legal and business incentives to keep data from moving present challenges. Health information networks and their participants often treat individuals' electronic health information as an asset that can be restricted to obtain or maintain competitive advantage."

4. Administrative requirements. "Federal documentation and administrative requirements (including billing requirements) contribute to health IT burden due to outdated guidelines for evaluation and management codes that unnecessarily link payment to documentation."

5. Reporting requirements. "Federal reporting requirements in some cases add burden to healthcare providers by requiring them to report on quality measures that are not relevant or meaningful."

6. IT usability. "Health IT system design and usability barriers identified by stakeholders include ... variations in the design [of user-interfaces] that make day-to-day use complicated when a healthcare provider uses multiple systems and the lack of developer engagement with end users of health IT regarding design needs[37].

2.5 Technologies used to achieve Interoperability

2.5.1 Distributed Databases

In recent years trends show that organizations have seen a continuous growth in the amount of data they handle and the level of sophistication it has achieved. Therefore the use of a centralized database can no longer serve the organizations effectively. This therefore led to the development of distributed databases. A distributed database is a single logical database that is spread physically across computers in multiple locations that are connected by a data communications network [49]. Here by we will show the structure of distributed database [55].

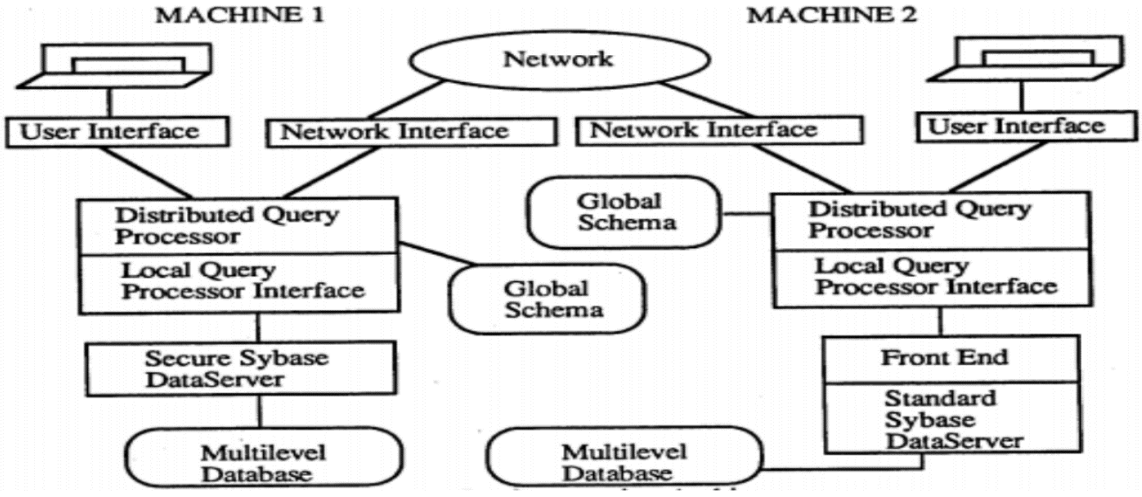


Fig 2.4 Distributed database implementation structure (Rubinovitz and Thuraisingham, 2010)

Distributed databases have improved the performance of systems tremendously. This is because of the extensive data fragmentation, the fragmentation enables data to be stored in close proximity to the users which benefits the system by reducing the amount of power needed to manage and query the database as compared to centralized databases. This also insures that there is reduction in the delays caused by delays in the networked environment. K. Tripathi and M. Tripathi 2012 [49], notes that centralized databases run on very complex and expensive mainframes which pose a great challenge when it comes to scaling. it also does not have capabilities to allow users to add processing and storage capacity. This means that organizations needs scalable databases that can allow system administrators to handle change in demand with less friction. Accordingly it's noted that replicated database systems provide a level of fault tolerance which cannot be achieved by centralized databases or traditional means such as redundant array of independent disks (RAID). To achieve this level of tolerance K. Tripathi and M. Tripathi 2012 [49] suggests replicating the database so that it's on two separate machines in different physical locations on the network so that the probability of losing the data is reduced significantly. This two options for failure recovery are implemented for database replication.

a) Warm standby user's synchronous replication to maintain the standby server in a state nearly consistent with that of the primary server. Due to the lag between transactions being committed on the primary server and replication on the standby server, a small number of transactions are normally lost during a primary server failure and switch over to the standby server
b) Hot standby users synchronous replication to maintain the standby server in a state always consistent with the primary server. From an availability perspective this is the preferred solution but the higher costs and potential lower performance of the synchronous replication databases cause many organizations to select a warm standby solution.

2.5.2 Blockchain

If this technology is so complex, why call it "Blockchain?" At its most basic level, Blockchain is literally just a chain of blocks, but not in the traditional sense of those words. When we say the words "block" and "chain" in this context, we are actually talking about digital information (the "block") stored in a public database (the "chain").

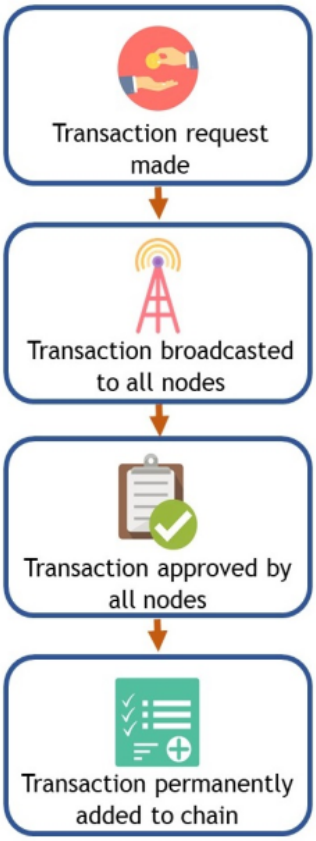


Fig 2.5 Generalized workflow of the blockchain process

In principle, a Blockchain should be considered as a distributed append-only timestamped data structure. Blockchain allow us to have a distributed peer-to-peer network where non-trusting members can verifiably interact with each without the need for a trusted authority [10]. To achieve this one can consider Blockchain as a set of interconnected mechanisms which provide specific features to the infrastructure, as illustrated in Fig. 2.6. At the lowest level of this infrastructure, we have the signed transactions between peers. These transactions denote an agreement between two participants, which may involve the transfer of physical or digital assets, the completion of a task, etc. At least one participant signs this transaction, and it is disseminated to its neighbors. Typically, any entity which connects to the Blockchain is called a node. However, nodes that verify all the Blockchain rules are called full nodes. These nodes group the transactions into blocks and they are responsible to determine whether the transactions are valid, and should be kept in the Blockchain, and which are not.

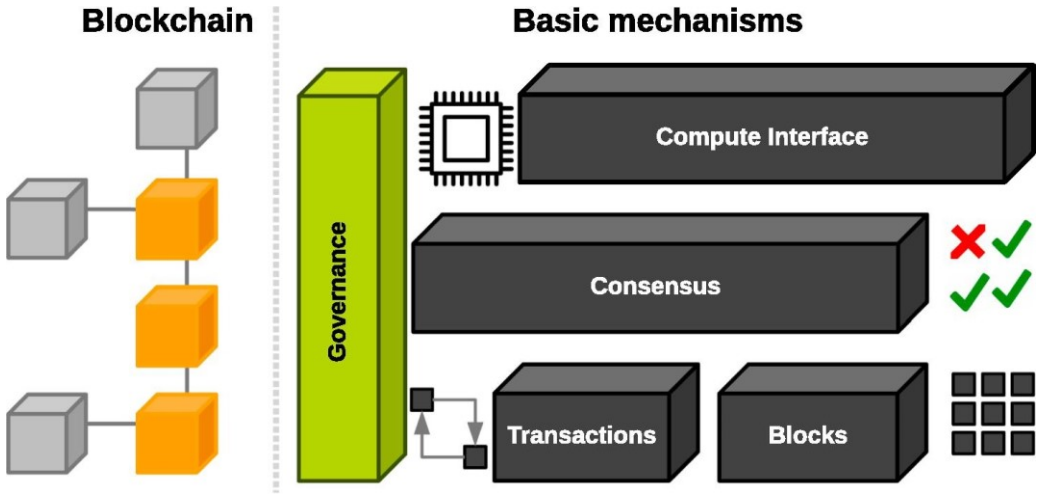


Fig 2.6 An illustration of basic Blockchain architecture.

“Blocks” on the Blockchain are made up of digital pieces of information. Specifically, they have three parts:

1. Blocks store information about transactions like the date, time, and dollar amount of your most recent purchase from Amazon. (NOTE: This Amazon example is for illustrative purchases; Amazon retail does not work on a Blockchain principle)
2. Blocks store information about who is participating in transactions. A block for your splurge purchase from Amazon would record your name along with Amazon.com, Inc. Instead of using your actual name, your purchase is recorded without any identifying information using a unique “digital signature,” sort of like a username.
3. Blocks store information that distinguishes them from other blocks. Much like you and I have names to distinguish us from one another, each block stores a unique code called a “hash” that allows us to tell it apart from every other block. Let’s say you made your splurge purchase on Amazon, but while it’s in transit, you decide you just can’t resist and need a second one. Even though the details of your new transaction would look nearly identical to your earlier purchase, we can still tell the blocks apart because of their unique codes.

While the block in the example above is being used to store a single purchase from Amazon, the reality is a little different. A single block on the Blockchain can actually store up to 1 MB of data.

Depending on the size of the transactions, that means a single block can house a few thousand transactions under one roof.

2.5.2.1 Classification of Blockchain

Blockchain is broadly classified in to three based on the type of transaction it handles and the nodes.

A. Private (permissioned) Blockchain

A permissioned Blockchain is one in which the interaction within the business is restricted to users who have the access rights provided by the network owner/s. In such a network, non-anonymous validation of blocks or interaction with the Blockchain is not permitted. Usually, a Certificate Authority (CA) is used to manage access to such a network. A Blockchain platform running its network as a permissioned network, will determine who can be validators and what privileges are given to what users. Hyperledger Fabric [6] is one of the most prominent example of a permissioned Blockchain framework.

B. Public Blockchain (Permission less)

A public Blockchain is a Blockchain that can be accessed by anyone (often, anonymously). There are no restrictions on who can join and what transaction they can post if the transactions are mathematically valid. Although members can join the network anonymously (revealing only their public key), every transaction that they undertake is visible to everyone (public), which can be carefully studied to identify the users. Bitcoin is the most famous example of public Blockchain's [3]. In such a network, there is typically an incentive given to the participants for executing a computing resource intensive consensus protocol (e.g., validate a block using Proof-of-work).

C. Consortium Blockchain

It is possible that a single (originating) organization will maintain the Blockchain (centralized) and provide predefined access rights to interacting parties. Such a network typically suits government or regulatory bodies who have legal purview over other participants. However,

It is debatable whether to call such a network a Blockchain network, as the ledgers are stored centrally and are not distributed among participating nodes. A consortium Blockchain provides some of the benefits affiliated with permissioned Blockchain — efficiency and transaction privacy, for example without consolidating power with only one company.

Since the third type of Blockchain is not often accepted as a Blockchain that provides openness and have that censorship resistant feature we are about to compare the private and public Blockchain on the table below.

Table 2.1 Comparison of private and public Blockchain

Criteria	Private permissioned	public permission less
Access	Read and write upon invitation only	Read and write public to anyone
Network Actors	know each other	Don't know each other
Native Token	Not necessary	Yes
Security	Legal contracts proof of authority	Economic incentives proof of work proof of stake proof of space proof of burn etc.
Speed	Fast	Slow
Examples	R3(Banks) EWF(Energy) B3i(Insurance) corda.	Bitcoin Etherium Monero Zcash Steemit Dash LiteCoin Stellar etc.
Effects	Reduces transaction costs and data redundancy and replaces legacy systems, simplifying document handling and getting rid of semi manual compliance mechanisms.in that sense it can be seen as equivalent to SAP in 1990's reduces costs but not disruptive.	potential to disrupt current business models through disintermediation. Lower infrastructure cost: no need to maintain servers or system admins radically reduces the cost of creating and running decentralized application(dApps).

2.5.2.2 How Blockchain Works?

When a block stores new data it is added to the Blockchain. Blockchain, as its name suggests, consists of multiple blocks strung together. In order for a block to be added to the Blockchain, however, four things must happen:

1. A transaction must occur. Let's continue with the example of your impulsive Amazon purchase. After hastily clicking through multiple checkout prompt, you go against your better judgment and make a purchase.
2. After making that purchase, the transaction must be verified, your transaction must be verified. With other public records of information, like the Securities Exchange Commission, Wikipedia, or your local library, there's someone in charge of vetting new data entries. With Blockchain, however, that job is left up to a network of computers. These networks often consist of thousands (or in the case of Bitcoin, about 5 million) computers spread across the globe. When you make your purchase from Amazon, that network of computers rushes to check that your transaction happened in the way you said it did. That is, they confirm the details of the purchase, including the transaction's time, dollar amount, and participants.
3. That transaction must be stored in a block. After your transaction has been verified as accurate, it gets the green light. The transaction's dollar amount, your digital signature, and Amazon's digital signature are all stored in a block. There, the transaction will likely join hundreds, or thousands, of others like it.
4. That block must be given a hash. Not unlike an angel earning its wings, once all of a block's transactions have been verified, it must be given a unique, identifying code called a hash. The block is also given the hash of the most recent block added to the Blockchain. Once hashed, the block can be added to the Blockchain.

When that new block is added to the Blockchain, it becomes publicly available for anyone to view even you. If you take a look at Bitcoin's Blockchain, you will see that you have access to transaction data, along with information about when ("Time"), where ("Height"), and by who ("Relayed By") the block was added to the Blockchain.

A Blockchain is a distributed transaction ledger [13]. The Blockchain itself is composed of blocks, with each block representing a set of transactions. As a data structure, a Blockchain has several interesting properties. First, blocks are provably immutable. This is possible because each block contains a hash, or numeric digest of its content, that can be used to verify the integrity of the containing transactions. Next, the hash of a block is dependent on the hash of the block before it.

This effectively makes the entire Blockchain history immutable, as changing the hash of any block n - 1 would also change the hash of block n.

The Blockchain itself does not depend on a central, trusted authority. Rather, it is distributed to all nodes participating in the network. Because no centralized authority may verify the validity of the Blockchain, a mechanism for reaching network consensus must be employed. In Bitcoin, a Proof of Work function is used to ensure network consensus [13]. This strategy requires that any node wishing to add a block to the Blockchain must complete a computationally expensive (but easily verifiable) puzzle first. At a high level, this ensures consensus of the network because there is an opportunity cost (the computation time) to building a block. There are several other techniques used, such Proof of Stake [14] and Proof of Activity [15], but all are designed to drive the network to consensus on Blockchain validity.

Miners are nodes that assemble the blocks and add them to the Blockchain. It is through the miners that the consensus strategy is enacted, usually via some incentivisation protocol. In Bitcoin, for example, miners are incentivized by collecting transaction fees and also by a reward for adding the block to the Blockchain. In general, however, there should exist an incentive for them to only build on top of valid blocks, which in turn drives the entire network to consensus.

A Blockchain is essentially a distributed database of records, or public ledger of all transactions or digital events that have been executed and shared among participating par-ties. Each transaction in the public ledger is verified by consensus of a majority of the participants in the system. Once entered, information can never be erased. The Blockchain contains a certain and verifiable re-cord of every single transaction ever made. To use a basic analogy, it is easier to steal a cookie from a cookie jar, kept in a secluded place, than stealing the cookie from a cookie jar kept in a market place, being observed by thousands of people. Bitcoin is the most popular example that is intrinsically tied to Blockchain technology. It is also the most controversial one since it helps to enable a multibillion-dollar global market of anonymous transactions without any governmental control. Hence it has to deal with a number of regulatory issues involving national governments and financial institutions. However, Blockchain technology itself is non- controversial, has worked flawlessly over the years, and is being successfully applied to both financial and non-financial world applications.

The Blockchain poses some unique characteristics as compared to any other ledger technology.

Below are some of its characteristics.

I. Decentralized

The major pillar of the Blockchain is not only its hashed chain of blocks but it's decentralized according to [39], this means that no single entity is in charge of the records. For the Blockchain there is no abuse concentrated power which is normally controlled by a central authority instead the power is distributed among a network of independent nodes and constantly validated by the network participants.

II. Network Effect

On most Blockchain technologies the network effect takes the biggest part. The greater the adoption of the technology means increased security, robustness, value and attractiveness.

III. Distributed Consensus

Distributed systems are highly expected to be reliable and to achieve this, protocols have to be put in place to ensure if failure occurs in one end the system should still be able to run. These protocols and other components need to cooperate in order to achieve this, but the problem comes in where systems have to decide which data will be used in computation [25]. For example, a vehicle control system should be able to be in sync with the sensor outputs. So in distributed consensus the concern is not what the process exchange but the fact in that all of them should have the same conclusion. And the Blockchain is built up putting these thing in consideration. It's able to get distributed system to come to an agreement regarding the state of the data without requiring a central authority. This is achieved by a P2P consensus based exchange and record keeping [39].

IV. Immutability

Going to the definition of Immutability on the oxford dictionary "immutable" means unable to change or unchangeable. Immutability in Blockchain presumably means it's impossible to change or alter any previous data in the chain of blocks. Miners are the once who are capable to add data to the Blockchain which is agreed and validated by all the nodes on the network. This makes the process and applications running on Blockchain to operate with a high degree of confidence that they have a complete and original.

This is what makes the block Blockchain the most amazing technology since the invention of the internet and the most significant applications of the technology are yet to be developed. In this section we will try to see the most famous applications of Blockchain starting from the most famous one:

A. Bitcoin:

According to Nakamoto bitcoin is a chain of digital signatures [45]. Where each owner transfers the coin to the next bi digitally signing hash of previous transaction and the public key of the next owner and adding these at the end of the coin. A payee can verify the signature to check the chain of ownership. And each individual node on the network can back trace the reliability of the chain of blocks up-to the genesis block. In Blockchain all blocks are mined (added) based on consensus except the genesis block the first block, why? it's because of the genesis block has no previous block that its chained to and to mine a block the miner need to verify the previous blocks hash [45].

In bit coin each block with multiple sets of transactions is hashed with SHA256 and a Nonce is added to it and hashed by the same hashing algorithm SHA256 again. A Nonce is a one-time use set of characters. And SHA256 provides an output of 64 digits of hexadecimal which is 256 digits of binary, which makes it very secure and immutable which is briefly described on security of Blockchain [7].

According to Goldman Sachs Global Investment Research [58]. The bitcoin transaction as illustrated in figure 2.7

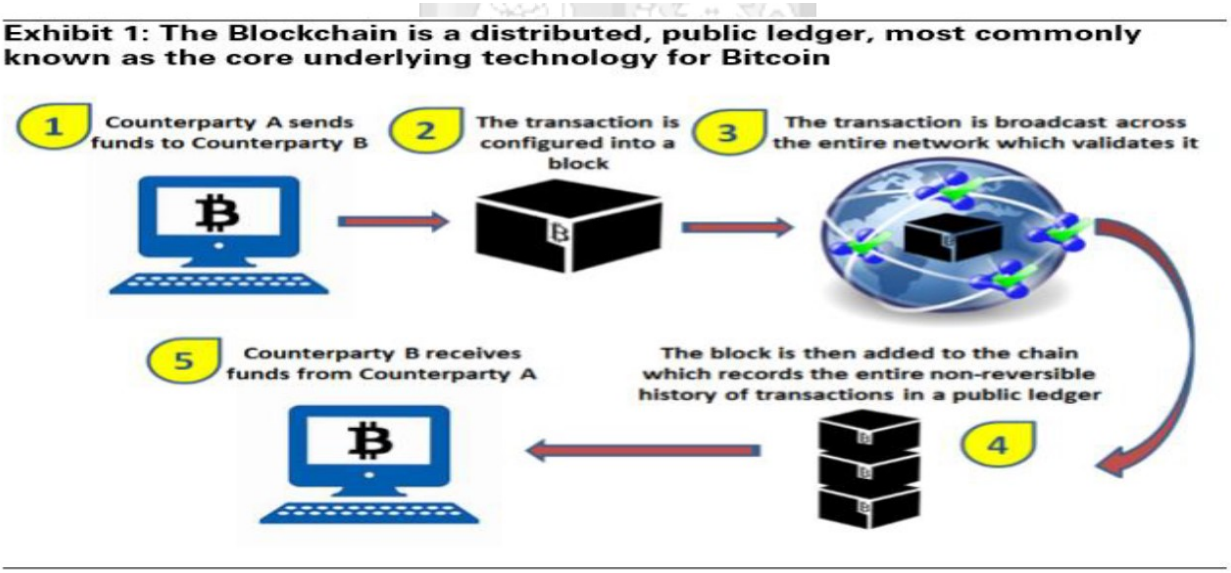


Fig 2.7 Overview of the Bitcoin Transaction[58]

Validity of a Bitcoin Transaction

A valid transaction means, for instance, that Bob received one bitcoin from Alice. However, Alice

may have tried to transfer the same bitcoin, as it is a digital asset, to Carol. Therefore, nodes must reach to an agreement on which transactions must be kept in the Blockchain to guarantee that there will be no corrupt branches and divergences[46]. This is actually the goal of the second Consensus layer. Depending on the Blockchain type, different Consensus mechanisms exist[46]. The most well-known is the Proof-of-work (PoW). PoW requires solving a complicated computational process, like finding hashes with specific patterns, e.g. a leading number of zeroes (Antonopoulos, 2014), to ensure authentication and verifiability. Instead of splitting blocks across proportionally to the relative hash rates of miners (i.e., their mining power), Proof-of-Stake (PoS) protocols split stake blocks proportionally to the current wealth of miners (Pilkington, 2016). This way, the selection is fairer and prevents the wealthiest participant from dominating the network. Many Blockchains, such as Ethereum[60], are gradually shifting to PoS due to the significant decrease in power consumption and improved scalability. Other consensus approaches include Byzantine Fault Tolerance (BFT) [14] and its variants.

An additional layer, the Compute Interface, allows Blockchain to offer more functionality. Practically, a Blockchain stores a state which consists e.g. of all the transactions that have been made by the users, thereby allowing the calculation of each user's balance. However, for more advanced applications we need to store complex states which are updated dynamically using distributed computing, e.g. states that shift from one to another once specific criteria are met. This requirement has given rise to SCs which use nodes of the Blockchain to execute the terms of a contract.

Finally, the Governance layer extends the Blockchain architecture to cover the human interactions taking place in the physical world. Indeed, although Blockchain protocols are well defined, they are also affected by inputs from diverse groups of people who integrate new methods, improve the Blockchain protocols and patch the system. While these parts are necessary for the growth of each Blockchain, they constitute off-chain social processes. Therefore, Blockchain governance deals with how these diverse actors come together to produce, maintain, or change the inputs that make up a Blockchain.

B. Ethereum

Capgemini explains Ethereum as a platform that takes the Blockchain concept a step further. It creates an open secure model for decentralized and generalized-transaction ledger. The creators of

the Ethereum envision it as a world computer that is immune from censorship and anyone can program it and pay exclusively for what they use. The Ethereum concept comes three years later the bitcoin which is 2012 GC by Vitalik Buterin and his friends and got into business on 2015. The Ethereum platform was created for those who have an idea the can be put on the Blockchain. Which would allow developers to create consensus based applications which are scalable, standardized, and easy to develop and be able to exchange information with other systems seamlessly [60]. The platform works by building an abstract foundation layer which lies on Blockchain with a built-in Turning complete programming language called solidity that allows anybody to write smart contracts and decentralized applications where they can create their own rules, transaction formats and state transaction functions. Figure 2.8 illustrates the state of transaction of Ethereum.

Smart contracts are data-sharing agreements between patients and care providers that are automatically enforced. Blockchain can put the patient at the center of the health care data ecosystem, enabling them to hold their own record and control providers' access to it. This may include having clinical photos and flexible control over who accesses them and in what context (e.g., diagnostic, research, before and after, academic, operative planning).

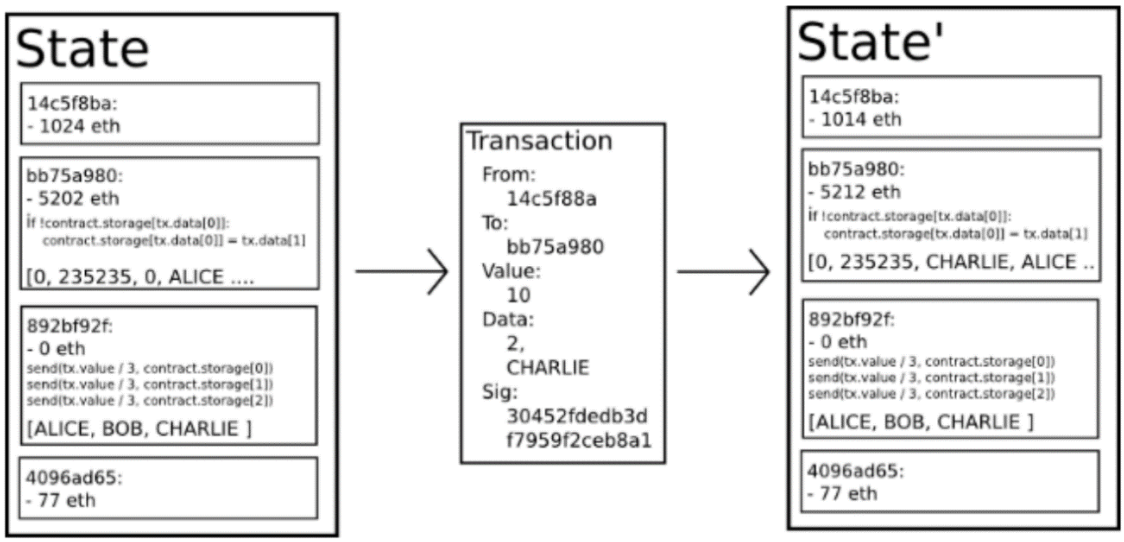


Fig 2.8 A state of an Ethereum transaction

C. Hyperledger Fabric

Hyperledger is an open source collaborative effort created to advance cross-industry Blockchain technologies. It is a global collaboration, hosted by The Linux Foundation, including leaders in

finance, banking, Internet of Things, supply chains, manufacturing, and Technology. Hyperledger does not support Bitcoin or any other cryptocurrency. But the platform is thrilled by Blockchain technology.

D. Storj

According to Wilkinson, Bosheveski, Brandoff and Vitalik Buterin story is described as peer-to-peer cloud storage network that implements end-to-end encryption to allow users to transfer and share data without reliance on a third party data provider [60]. By removing the central authority, the traditional data failures and outages are eliminated and security, privacy and data control is improved. To achieve this a peer-to-peer network and basic encryption are used to solve the problem, also users are given incentive for participating in the network.

2.5.2.3 How secure is Blockchain?

The whole point of security is letting people interact with each other with no fear, people who don't trust one another. Share valuable data in a secure, tamperproof way. That's because Blockchain stores data using sophisticated math and innovative software rules that are extremely difficult for attackers to manipulate. To understand why, we shall start with what makes Blockchain "secure" in principle. Bitcoin is a good example, which is the first application of the Blockchain technology back in 2009. In Bitcoin's Blockchain, the shared data is the history of every Bitcoin transaction ever made: for instance, let's take an accounting ledger or going specific a bank ledger (Bank Book). In a bankbook each and every one of our transaction is written in a separate line showing us all the transactions performed, but the bank book was never mandatory to get the history of our transaction of get the current balance on the account b/c all the transaction is stored on the bank's system (which makes is centralized).but the whole point of Blockchain was decentralization as well us security so In bitcoin Satoshi Nakamoto came up with a brilliant idea in which all transactions of everybody stays on everybody's ledger and no central authority has the right or mandate to follow-up those transactions, instead transactions bundled in chunk and sent out to the rest of the world if they meet certain rules of consensus, these set of rules are kept on every other copy of application that runs the bitcoin software [7]. Transactions chunked together are called blocks of the Blockchain. Chunked transactions are hashed with SHA256 which yields a 64-character hexadecimal. Hashes of blocks are saved on headers of their successor blocks and secret non repeatable character called Nonce is added on the footer of each block and hashed again

by SHA256 to give the block hash shown on the image below.

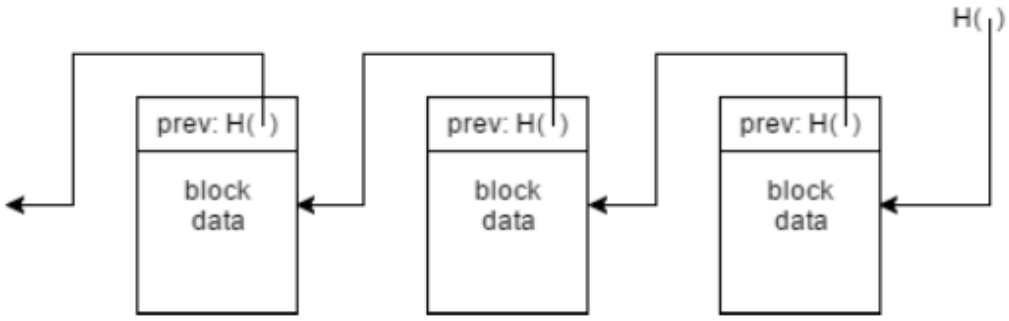


Fig 2.9 A Continuously Growing List of Ordered and Validated Transactions

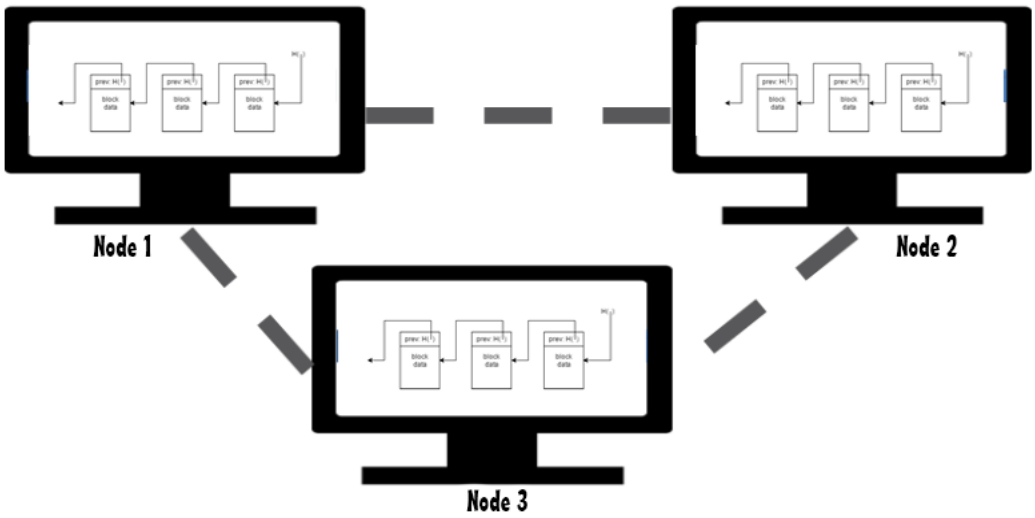


Fig 2.10 Nodes having their own identical ledger.

All the transactions on bitcoin take the following 7 steps before they get added to the Blockchain which are

I. Transactions are requested.
II. A block that represents the transaction is created.
III. A block is sent to every node on the network.
IV. Nodes validate the transaction. (Special nodes called miners)
V. Nodes receive reward for their proof of work.
VI. Block is added to the existing Blockchain.
VII. The transaction is complete and node get ready to proceed to the next transaction.

In Blockchain the ledger is stored in multiple copies on a network of computers, called “nodes.”

Each time someone submits a transaction to the ledger, the nodes check to make sure the transaction is valid—that whoever spent a bitcoin had a bitcoin to spend. A subset of them compete to package valid transactions into "blocks" and add them to a chain of previous ones. The owners of these nodes are called miners. Miners who successfully add new blocks to the chain earn bitcoins as a reward.

So what makes Blockchain tamperproof, the fact that every block on the Blockchain is interlocked with previous blocks hashes it's very difficult to hack, even if some smart head has the capability to hack the Blockchain and change record on a certain block he can't, because as soon as the change is made on one node it broadcasts the change in which the other nodes on the network disagree on the change and its discarded.

2.5.2.4 Consensus in Blockchain

Consensus is a mechanism or a method by which all nodes in the Blockchain network agree upon which block (transaction) gets added to the chain. As discussed earlier on P2P Networks distributed computing existed well before Blockchain, but it is these consensus mechanisms which makes Blockchain new technology in which all the nodes in the network agree which is so robust. This technology provides nodes the ability to create honest self-correcting systems without the need of third-party to enforce the rules is what makes Blockchain so powerful. To enforce the rules, several variations of consensus algorithms/protocols are used, each with their pros and cons.

A. Proof of Work (PoW)

The most famous consensus algorithm is Bitcoin's Proof of Work (PoW), which make sure that the subsequent blocks in the chains are the only true versions. The correctness of the transactions can be verified by any participant using network consensus methods and cryptographic technologies in the Blockchain network. So, effectively the trust is established continuously within the network and not by any external central authority or auditor [10,47].

The Proof of Work protocol involves the following:

- The miners solve cryptographic and complex puzzles to "mine" a block.
- These puzzles are designed in such a way which makes it hard and taxing on the system as the process requires immense amount of energy and computations usage.

- After solving a puzzle, the miner must present the block to the network for verification. Then, it is concluded whether are not this block belongs to the chain, which is not a simple process.

B. Proof of Stake (PoS)

Proof of Stake (PoS) is ideologically different from PoW, wherein the complete mining process is done virtually and the miners are replaced with validators. The validators must lock up some of their coins as a stake before the validation process is begun. During the validation process, if a block is discovered which they think can be added to the Blockchain, it would be validated by placing a bet on it. If the block's validation is successful, the block gets appended and the validators get a reward which is proportionate to the bets they placed [10].

C. Proof of Activity (PoA)

Proof of Activity (PoA) is a hybrid approach that combines the previous two consensus algorithms namely PoW and PoS Here, the mining is commenced in the traditional Proof-of-Work way, where the miners compete to solve a cryptographic puzzle. Importantly, here depending on specific implementations, the 'mined blocks' does not contain any transactions but are more like templates. The successfully mined and validated block contains only a header and the miner's reward address.

D. Proof of Burn (PoB)

As can be derived easily, the first three consensus algorithms are quite resource-intensive, both computationally, financially, and energy-wise (massive electricity used for upkeep of expensive computer hardware and Application Specific Integrated Circuit or ASIC cards). To circumvent this drawback, the Proof-of- Burn (PoB) algorithms lets you 'burn' the coins by dispatching them to irretrievable addresses. The miners are selected randomly to mine on the system. Depending on the implementation, the miners may burn the native currency or the currency of alternative chain such as Bitcoin. The miners have better chance of being selected to mine the next block depending on how many coins they have burnt. The following table compares well known consensus algorithms in Blockchain and explains the pros and cons of each of them.

Table 2.2 Comparison of Blockchain consensus algorithms

Consensus Algorithm	Brief Description	Pros/Cons

Proof of Work (PoW)	Nodes must solve complex cryptographic puzzles to get the right to append new blocks to the chain and get the rewards.	Pros: Being the first algorithm, it's currently the most popular. It's also highly scalable, which makes it attractive.	Cons: Resource-intensive (Computational, financial, energy). Vulnerable to "51% attack"
Proof of Stake (PoS)	Validators lock up some of their coins as stake after successful validation block is added to the chain.	Pros: No need to solve complex cryptographic puzzles. Fast, efficient and uses less hardware	Cons: Vulnerable: Someone with enough money to invest exclusively into the destruction of this system can do so by investing only
Proof of Activity	Hybrid approach combining PoW and PoS The successfully mined and validated block contains only a header and the miner's reward address.	Pros: Combines best features of both PoW and PoS	Cons: Less resource-intensive (Computational, financial, energy)
Proof of Burn	Nodes must send their coins to an irretrievable address to mine a new block. The miner sending the largest number of coins get the chance to mine a new block.	Pros: No need to solve complex cryptographic puzzles.	Cons: Burning coins is expensive as there is loss of coins. Less resource intensive.
Proof of Capacity	Large number of plots generated in hard disk on stake, to get the right to mine the next block.	Pros: Miners does not need specialized hardware to mine. It decentralizes the mining process.	Cons: Need lots of hard-disk space
Proof of Elapsed Time	All nodes receive different waiting time duration, and the node with shortest duration will mine a new block.	Pros: Highly energy efficient as no cryptographic puzzle to be solved.	Cons: Reliance on third-party (Intel). Relies on specialized hardware

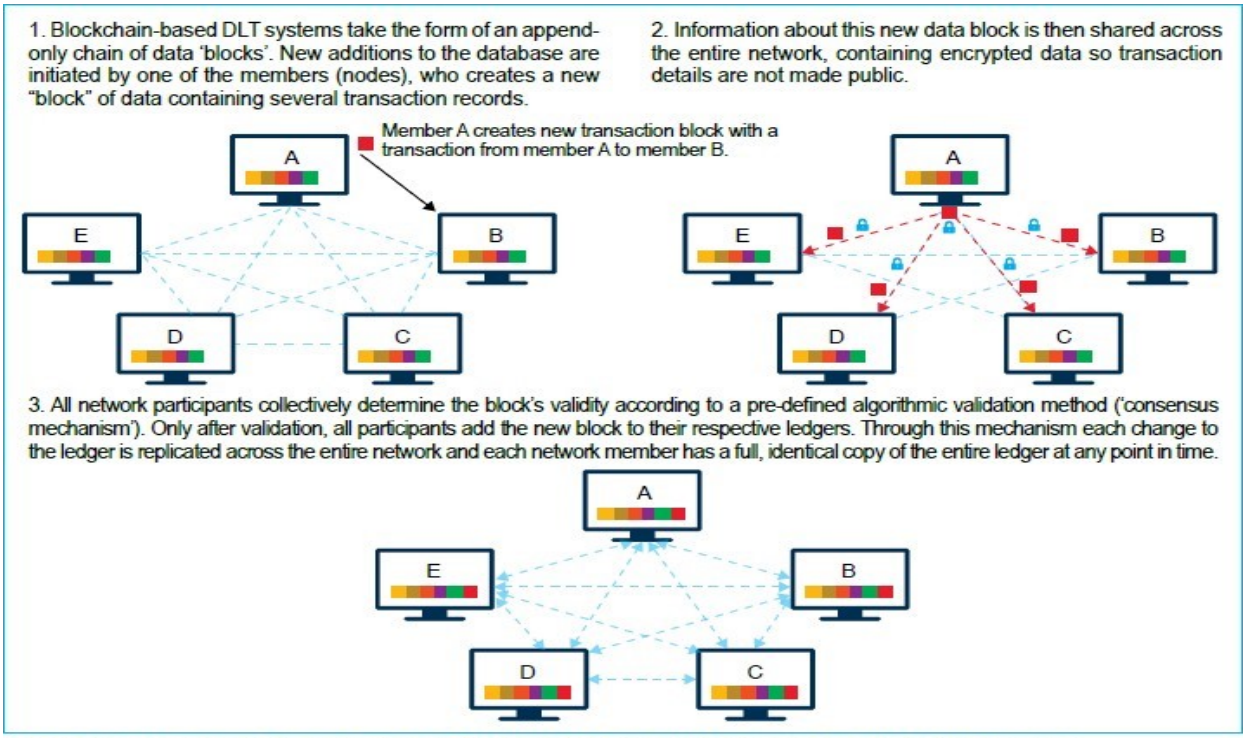
2.6 Distributed Ledger Technology (DLT)

DLT comes on the heels of several peer-to-peer (P2P) technologies enabled by the internet, such as email, sharing music or other media files, and internet telephony. However, internet-based transfers of asset ownership have long been elusive, as this requires ensuring that its true owner only transfers an asset and ensuring that the asset cannot be transferred more than once, i.e. no

double-spend. The asset in question could be anything of value. In 2008, a landmark paper written by an as yet unidentified person using the pseudonym Satoshi Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System”, proposed a novel approach of transferring “funds” in the form of “Bitcoin” in a P2P manner [45].

The underlying technology for Bitcoin outlined in Nakamoto’s paper was termed Blockchain, which refers to a particular way of organizing and storing information and transactions[45]. Subsequently, other ways of organizing information and transactions for asset transfers in a P2P manner were devised – leading to the term “Distributed Ledger Technology” (DLT) to refer to the broader category of technologies. DLT refers to a novel and fast-evolving approach to recording and sharing data across multiple data stores (ledgers), which each have the exact same data records and are collectively maintained and controlled by a distributed network of computer servers, which are called nodes. One way to think about DLT is that it is simply a distributed database with certain specific properties (see section 3). Blockchain, a particular type of DLT, uses cryptographic and algorithmic methods to create and verify a continuously growing, append-only data structure that takes the form of a chain of so called ‘transaction blocks’ – the Blockchain – which serves the function of a ledger. One of the members (nodes), who creates a new “block” of data, for example containing several transaction records, initiates new additions to the database.

Information About this new data block is then shared across the entire network, containing encrypted data so transaction details are not made public, and all network participants collectively determine the block’s validity according to a pre-defined algorithmic validation method (‘consensus mechanism’). Only after validation, all participants add the new block to their respective ledgers. Through this mechanism each change to the ledger is replicated across the entire network and each network member has a full, identical copy of the entire ledger at any point in time. This approach can be used to record transactions on any asset, which can be represented in a digital form.



Source: Adapted from: "Dubai Aims to Be a City Built on Blockchain", By Nikhil Lohade, 24 April 2017, Wall Street Journal https://www.wsj.com/articles/dubai-aims-to-be-a-city-built-on-blockchain-1493086080

Fig 2.11 Distributed Ledger Technology

Table 2.3 shows the strengths and weaknesses of the Distributed technologies used in achieving interoperability.

Table 2.3 Comparison of Distributed Technologies used to achieve Interoperability

Technology	**Strengths**	**Weaknesses**	**Author**
Blockchain	Immutable Decentralized Distributed consensus Scalability Fault tolerance Transparency	Latency	(Needham & Company LLC, 2015)
Distributed databases	Decentralized Scalability	Concurrency Privacy Costly	(Depardon, Mahec, & S´eguin, 2013)

	Fault tolerance Transparency		
Distributed File systems	Decentralized. Fault tolerant. Self-sustained scalable	Privacy Latency Data integrity made on P2P host Distributed Denial of Service attacks can be made of P2P host	(Ragib, Zahid, William, Larry, & Roy, 2008)

2.6.1 Distributed peer to peer file system

There has been a growth in research on peer-to-peer systems. The use of large scale distributed network of nodes has becoming an important component of distributed computing due to the increased use by peer-to-peer platforms such as Napster. Peer-to-peer file systems are becoming popular in the area of research, this is because they offer a decentralized, self-sustained, scalable, fault tolerant and symmetric network of nodes offering an effective balance in storage and bandwidth resources [44].

Peer-to-peer architecture has been regarded as a major component for implementing distributed file system. In such a network users share resources via a direct exchange with other nodes, this means the information is distributed among the member nodes instead of being in a single central server [44].

The proposed solution is going to provide a fast and secure means of storing and retrieving medical records for all systems connected to the network. As discussed in the literature there are a number of implementation developed to enhance interoperability, but most of them have dwelled on semantic interoperability. Therefore, the combination of Blockchain technology, distributed file system, E-Health standards and web services will provide a system that takes care of the semantic and technical interoperability.

2.7 Byzantine General's Problem(BGP)

In a few words, the Byzantine Generals' Problem was conceived in 1982 as a logical dilemma that illustrates how a group of Byzantine generals may have communication problems when trying to agree on their next move [36].

The dilemma assumes that each general has its own army and that each group is situated in different locations around the city they intend to attack. The generals need to agree on either attacking or retreating. It does not matter whether they attack or retreat, as long as all generals reach consensus, i.e., agree on a common decision in order to execute it in coordination [36].

Therefore, we may consider the following requirements:

- Each general has to decide: attack or retreat (yes or no);
- After the decision is made, it cannot be changed;
- All generals have to agree on the same decision and execute it in a synchronized manner.

The aforementioned communication problems are related to the fact that one general is only able to communicate with another through messages, which are forwarded by a courier. Consequently, the central challenge of the Byzantine Generals' Problem is that the messages can get somehow delayed, destroyed or lost.

In addition, even if a message is successfully delivered, one or more generals may choose (for whatever reason) to act maliciously and send a fraudulent message to confuse the other generals, leading to a total failure.

If we apply the dilemma to the context of Blockchain's, each general represents a network node, and the nodes need to reach consensus on the current state of the system. Putting in another way, the majority of participants within a distributed network have to agree and execute the same action in order to avoid complete failure [14].

Therefore, the only way to achieve consensus in these types of distributed system is by having at least ⅔ or more reliable and honest network nodes. This means that if the majority of the network decides to act maliciously, the system is susceptible to failures and attacks (such as the 51% attack).

2.7.1 Byzantine Fault Tolerance

In a few words, Byzantine fault tolerance (BFT) is the property of a system that is able to resist the class of failures derived from the Byzantine Generals' Problem. This means that a BFT system is

able to continue operating even if some of the nodes fail or act maliciously.

There is more than one possible solution to the Byzantine Generals' Problem and, therefore, multiple ways of building a BFT system. Likewise, there are different approaches for a Blockchain to achieve Byzantine fault tolerance and this leads us to the so-called consensus algorithms.[14]

2.8 Related works

On the paper "Blockchain for cities: by *CHARLES SHEN AND FENIOSKY PENA-MORA* who are famous scientists from the university of NYU having lots of research papers and publications around Blockchain and other cryptology related articles have stated that many cities around the world have reported Blockchain-related initiatives, such as those in Australia, China, Denmark, United Arab Emirates, Estonia, Georgia, Ghana, Honduras, Malta, Russia, Sweden, Singapore, Spain, Switzerland, United Kingdom, Ukraine, United States(US). Cities and states set up various goals and employ many approaches in the race to lead the Blockchain wave. For example, Dubai is building a single software plat-form through which city public sector can launch Blockchain projects, as part of the ambition to become paperless by 2020; in contrast, Illinois takes a more experimental approach, launching multiple separate Blockchain pilots across different industrial sectors including governance, education, health-care and energy, each selecting their own Blockchain platform as appropriate. Zug is developing itself to be a ''crypto valley'' through establishing a crypto-friendly business ecosystem. New York City announced plans to launch the Blockchain Resource Center as a hub for the city's Blockchain industry and to convene both government and citizen stakeholders in developing a regulatory environment that stimulates the overall Blockchain industry.

Despite all the ongoing Blockchain efforts, many also believe that our current understanding of Blockchain is pre-mature and there is a lack of knowledge on where Blockchain technology can provide mentionable societal effects. Sometimes the field is even described as ''an innovative technology searching for use cases'' because it is largely unknown how Blockchain could be incorporated to existing digital services, processes and infrastructures. In a testimony to the US congress, US Department of Homeland Security's Science and Technology Division Director *Douglas Maughan* also pointed out specific concerns in the Blockchain space for the asymmetries between knowledge and action. Biased use of the buzzword in fragmented or superficial ways will lead to more confusion than clarity. Falling into the tendency to believe that innovative technologies like Blockchain can automatically transform the ecosystem around us will actually

hinder the achievement of the technology's real potential. Under this mixed backdrop, this paper attempts to advance the understanding towards how Blockchain can fit in the next level of urban development initiatives, by combining foundational frameworks on sustainable and smart cities with Blockchain domain knowledge accumulated by the research community. Through helping city policy makers, industrial practitioners and all stakeholders better understand Blockchain use cases in cities, we hope to facilitate decision makers in planning of Blockchain strategy and drive actions in the most pertinent industrial domains that contribute to meet the urban growth challenges.

And also other papers are also reviewed which are listed below.

Table 2.4 Related Papers and their reviews.

Titles	Author	Technology used	Achievements
Hashcash-A denial of service countermeasure,2002	Adam Back	Hashcash for POW and SHA256	Discourages Spammers by giving them exhaustive mathematical calculations
Bitcoin: A Peer-to-Peer Electronic Cash System	Satoshi Nakamoto, 2009	Blockchain, Hash Algorithm SHA-256	Revolutionize the banking system where no banker is needed and peers trust each other with consensus mechanisms
How to time-stamp a digital document	S. Haber, W.S. Stornetta,1991	Hash Algorithms SHA-256,MD-5	Digital documents get their timestamps regardless of their time zone
The Ethereum Blockchain	Vitalik Buterin, 2012	Blockchain, distributed ledger, Pos for consensus	Makes the blockchain programmable by introducing smart contracts.
BASIC: Towards a Blockchained Agent-Based Simulator for Cities	Luana Marrocco et al., 2016	Ethereum Blockchain, Smart contracts	blockchain in simulated urban scenarios by considering the communication between agents through smart contracts.
Blockchain Technologies for the Internet of Things: Research Issues and Challenges	Mohamed Amine Ferrag,2008	Ethereum, IOT	Recommends how to resolve major security issues that the IOT might face.

CHAPTER THREE
METHODOLOGY

3.1 Introduction

The study follows design science research approach for the overall work of this study. The researcher, on the next sections has discussed What, How, and Why is each step of the research process done in detail. The fundamental principle of design science research is that knowledge and understanding of a design problem and its solution are acquired in the building and application of an artifact [48].

This study aims at finding out the challenges that health care organizations may experience in keeping and exchanging patient's medical history amongst themselves and how to develop a platform that can request data from the Blockchain system and get records on hand. This chapter explains the research methodology that was used, location of the research, study population and sampling , the stages in the research, and purpose of the research, data collection techniques used. The general model used in this study is the model provided by Omar Valdez-de-Leon. It is used as the general guideline to further identify specific criterion questionnaires that are aligned with the given seven dimensions.

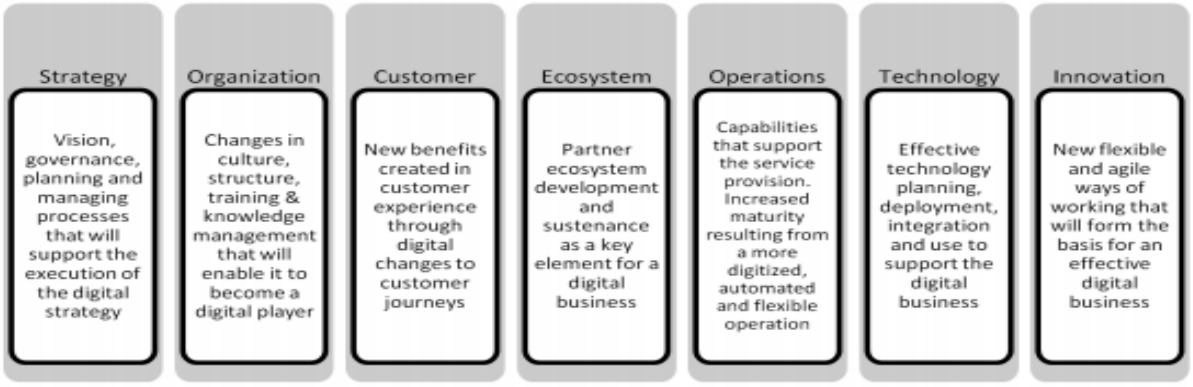


Fig 3.1 multi stage research model by Omar Valdez-de-Leon

3.2 System development methodology

The development methodology of the system is an Agile methodology that provides an opportunities to assess the direction of the project on each and every step of the lifecycle. This is

achieved through iterations or sprints, in which it helps us get a more stable outcome at the end.

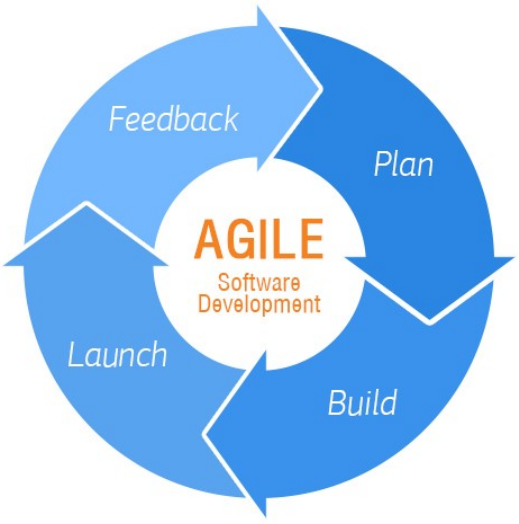


Fig 3.2 Agile Development cycle

Agile software development is based on several values that are meant to increase communication among developers and between developers and clients. They are also intended to allow clients to have a better sense of project progress so that revisions can be addressed earlier in the process [46].

Clear Communication: Programming team members need to be in regular communication so that questions can be answered quickly and instructions delivered directly[17]. This makes communication more efficient, keeping everyone on the same page. It also streamlines the development process.

Short-term Goals over Long-term Plans: Agile software development encourages breaking projects down into smaller pieces. Partners and teams are assigned small chunks of programming with short deadlines. Short-term goals allow for greater flexibility. As changes arise, goals can be more easily modified. [46] *Active Collaboration between Client and Developers*: One of the frustrations for clients was not being able to see the progress on a project. At the same time, developers would be frustrated when a finished project was sent back for multiple revisions. By bringing the client in as a partner, revisions can be brought in as part of the goal-setting process during the building phase [46].

Shorter Feedback Loops: In the agile world, testing becomes a regular part of the process. Small pieces of the project are tested and presented on a regular basis. This gives everyone a better sense of the project's timeline. In addition, this frequent testing allows developers to catch bugs before they become deeply entrenched in the code. [13]

3.3 Research Design

The research design used in this thesis is experimental research design, this design allows one to analyze the prior achievement in order to establish an equivalent solution for the study (Ross & Morison). Coming up with this design was such a difficult task because the requirements for the proposed system has very few previous implementations since the technology is very new? The analysis formed a basis for the system requirements.

3.3.1 System analysis

Since the intended users of the system are mostly patients the study used is object oriented analysis approach, this approach combines data and process into single entities called object.

This approach was most commonly used in developing applications. The study used use-case modeling and sequence diagrams. It helped in gaining a clear understanding of functional requirements of the system.

The system requirement of the system were obtained through analysis of documents, interviews with domain experts and also questionnaires conducted. Domain experts were interviewed from different organizations and different fields i.e. Medical Doctors, Health officers, Pharmacists, Data clerks, patients and the last but not the least Software Developers to achieve the overall skeleton of the system. This was done in order to create a balance in the kind of solution to be developed. The documentation analyzed contained specifications and standard guidelines on previous implementation of previously developed systems. This analysis gave us an abroad perspective of the solution to be developed including the features that should be added and that should be deprecated.

3.3.2 System Design

We have used a Function oriented design (FOD) techniques to refine the functional requirements identified during system analysis and to decompose the design into sets of interacting units where each unit has clearly defined functions. Dataflow diagrams were used to show how the system will

handle the different data flows between the process and the entities.

The system sequence diagram shows information passing b/n the main entities was used to model the system flow. It showed how objects interact with each other, which helped us to explain the different components of the system work together. Use case diagrams were used to model the system functionalities. This enabled the researcher to separate the system actors from use cases.

3.3.3 System Implementation

For this project we have used different components which help in achieving certain functionalities. NodeJs provide server side scripting using JavaScript. JavaScript is chosen for because it's the most widely used programming language nowadays especially when it comes to lightweight applications running on a server, proceeding to the pillar of the Blockchain which is the smart contact, its written with ethereum's special programming language "solidity" which is only used to write smart contracts that run on the Blockchain. The Blockchain was used as a public ledger because of its tamperproof nature and it also cost money to put data on the Blockchain, this means the number of people spamming the network will not be many. In our case anyone who needs to talk to the Blockchain uses an ordinary browser like Firefox or Chrome with the "Metamask" add-on which is used to make the ordinary browser a Blockchain browser so it will talk to the Blockchain.

3.3.4 Testing

The system combines two types of testing which are System testing and usability testing. System testing refers to checking and evaluating each transaction in the Blockchain if they can pass the rules on the smart contracts. Usability testing refers to evaluating a product or service by testing it with representative users. Typically, during a test, participants completed typical tasks while observers watch, listen and took notes. The goal was to identify any usability problems, collect qualitative and quantitative data and determine the participant's satisfaction with the product (Improving the User Experience). Usability testing was done to ascertain that the developed application was user friendly and easy to navigate through without any challenge.

3.4 Location of study

The study was carried Addis Ababa capital city of Ethiopia and some other sub urban areas of Ethiopia which are Adama and Hawassa. Those cities were chosen because of its massive number of Governmental and non-governmental health care organizations exist. And also the system we

are working on needs a good internet connection and these cities have the best internet coverage in Ethiopia, this means it will facilitate the deployment of the platform without much resistance.

3.5 Target population

In order to estimate the target population multistage sampling was used. The first stage was to identify the number of health facilities in the cities selected. According to Ethiopian Ministry of health publication on 2019, the total number of public and private health facilities in those cities were 171 and from those 30 were picked.

The second stage was to identify the personnel to conduct the study. The study picked 1 people per facility based on criteria and availability.

3.6 Data collection

The data collected was the test results and interview feedback. This usability questionnaires results provided insights on how the users felt about the application and whether it solved their problem. There were also interviews conducted. Because of the Novel Corona Virus pandemic is spreading very fast Google forms and Google docs were used in the questionnaire data collection. The process involved the use of prepared guidelines in Appendix I, II, III. The guidelines assisted the researcher to be able to derive the requirements and limitations of the existing system.

3.7 Research Instruments

The study was carried out using questionnaires, Observation, reading of documents and interviews. The data collected was used to analyze the current systems, come up with requirements of the proposed system and finally determine the usability of the system developed.

3.7.1 Interviews

Interview guide in Appendix A and B were used to provide a roadmap on the kind of questions to be asked on the questionnaire so doctors and developers interact with the current systems and whether they experience any difficulties when sharing data. The interview was used to identify how the doctors and developers feel about the current system and determine what improvements should be made to make data sharing a seamless process.

3.7.2 Questionnaires

Questionnaires were administered to a select sample population composed of patients, developers and doctors. The main medium of administering the questionnaires was through representatives.

Filled questionnaires are properly compiled and analyzed. The questionnaires were used to analyze how medical record flows in an organization and how they are capable of managing it.

3.8 Data analysis

Exploratory data analysis was performed in order to make connections between the technologies which have been used in the past and how the current technologies can be incorporated to make a difference. And how current technologies could enhance to current trend.

3.9 Research Quality Aspects

3.9.1 Research Validity

Validity determines whether the research truly measures that which it was intended to measure or how truthful the research results are [26]. To validate the study, a simulation to depict the nature of transaction that occurs between two individual actors in the system. The validity of the system is also dependent on the hash of the transaction mined on the system. And it's based up on feedbacks from domain individuals (Doctors, Nurses, Midwives…), Patients, Developers, and the overall validation of the thesis focuses on three major stages and they are listed below.

1. Validate against those individuals who filled the questionnaire in the first place.
2. Validate against those individuals who didn't fill the questionnaire.
3. Validate against review of related works where you are going to do comparative features analysis.

Details of the research validity is discussed on chapter 5 of this paper.

3.9.2 Reliability

Reliability is the extent to which results are consistent over time and an accurate representation of the total population under research is referred to as reliability and if the results of a research can be reproduced under a similar methodology, then the research instrument is considered to be reliable [26].

In this research reliability was attained by giving respondents questionnaires to fill. After a few system features were added, a second survey was gathered through interviews so that it helps to check whether the system requirements are met. This gave the researcher a go ahead with the study.

3.10 Tools and Technologies Used

3.10.1 Draw.io

Draw.io is an online modelling software which integrates with many platforms. This tool was used come up with application designs and mock. It's a web-based proprietary platform that is used to allow users to collaborate on drawing, revising and sharing charts and diagrams. (Google, 2016).

3.10.2 NodeJs

NodeJs is a JavaScript server side scripting language. NodeJs uses an event-driven, non-blocking I/O model that makes it lightweight and efficient (Nodejs Foundation, 2016). The technology was used to interface the middleware application with the Bitcoin Blockchain.

3.10.3 Truffle Suit + Ganache

Ganache is a truffle framework that works locally building virtual nodes and act as a Blockchain working locally. Truffle Ganache is used for quickly firing up a personal Ethereum Blockchain which you can use to run tests, execute commands, and inspect state while controlling how the chain operates.

3.10.4 ReactJs

React (also known as React.js) is a JavaScript library for building user interfaces. It is maintained by Facebook and a community of individual developers and companies.

React can be used as a base in the development of single-page or mobile applications. However React is only concerned with rendering data to the DOM and so creating React applications usually requires the use of additional libraries for state management, routing, and interaction with an API. Redux, React Router and Axios are respective examples of such libraries.

3.10.5 IPFS

The Interplanetary File System (IPFS) is a new hypermedia distribution protocol, addressed by content and identities. IPFS enables the creation of completely distributed applications. It aims to make the web faster, safer, and more open (Protocol Labs, n.d.). IPFS is a distributed file system was used to store the electronic health records for individual patient records. The distributed nature of the platform allows the records to be replicated across several geographical regions.

3.10.6 INFURA

Infura is a scalable back-end infrastructure for building Dapps on the Ethereum Blockchain. It is a method for connecting to the Ethereum network without having to run a full node, and in this project Infura is used to get connected to the IPFS network.

CHAPTER FOUR

SYSTEM DESIGN AND ARCHITECTURE

4.1 Introduction

In this chapter we are defining the architecture, modules, interfaces, and data for a system to satisfy specified requirements. Systems design could be seen as the application of systems theory to product development.

Respondents from the data collection and analysis overall replied that the current trend in the health care data storage is not serving them very well and it should be enhanced in whatever the way, most of the respondents are having interoperability issues and others have concerns in security and privacy of medical history storage. The overall response of respondents is presented separately in chapter 6.

4.2 Requirement Analysis

4.2.1 Functional Requirements

The functional requirements define how the system is going to behave. The following table shows the functional requirements used in development of the system and used to guide the system work flow.

Table 4.1 Functional requirements

№	Functional requirements
1	All users shall be able to access their account through meta mask.
2	All users should login to authenticate their health system.

3	Patient should be able to access their past medical record anywhere any time as long as they are connected to the Blockchain.
4	Patient should be able to allow doctor to view there medical record.
6	All recodes shall be accessed independent of the organization. Medical data's shall be accessed across any health facility.
8	All users should be able to logout.

4.2.2 Non-Functional Requirements

Any aspect of the system that is not related with the behavior of the system is called non-functional requirement. the main target of the functional requirement is showing how the system is supposed to be. The main objective of the system is not only fulfilling the functional requirements but to make an immutable platform that can handle the basic workflow of medical facilities, and also making a suitable and convenient platform that gives emphasis on the components that need to make it usable and valuable to the user. The main source of these requirements is as result of analyzing the responses from the interview with the selected sample population. The table 4.2 shows the non-functional requirements.

Table 4.2 Non-Functional requirements

№	Non-Functional requirements
1	The application must run on any browser that's capable of accessing the Blockchain.
2	The application should be easy and intuitive to use.
3	The application should ensure that the all requests have their own transaction hash for confirmation.
4	The application should ensure the integrity and verification of the patient's data.
5	The application should provide an intuitive error handling and notification interface.
6	The system should allow the patient to secure their records while sharing.
7	The system should always be available for the user to access their data.
8	The system should perform validation for all the data entered.

4.3 Proposed System

The proposed system gives provides the Electronic Health data recording system basically two things. Firstly secure storage for patients medical record by using the raw Blockchain technology and secondly the file storage system in which the blockchain technology is incapable of providing so to support the blockchain system in storing big files like images IPFS is used which provides big file storage.

4.4 Prototype System Design

4.4.1 Prototype Framework Structure

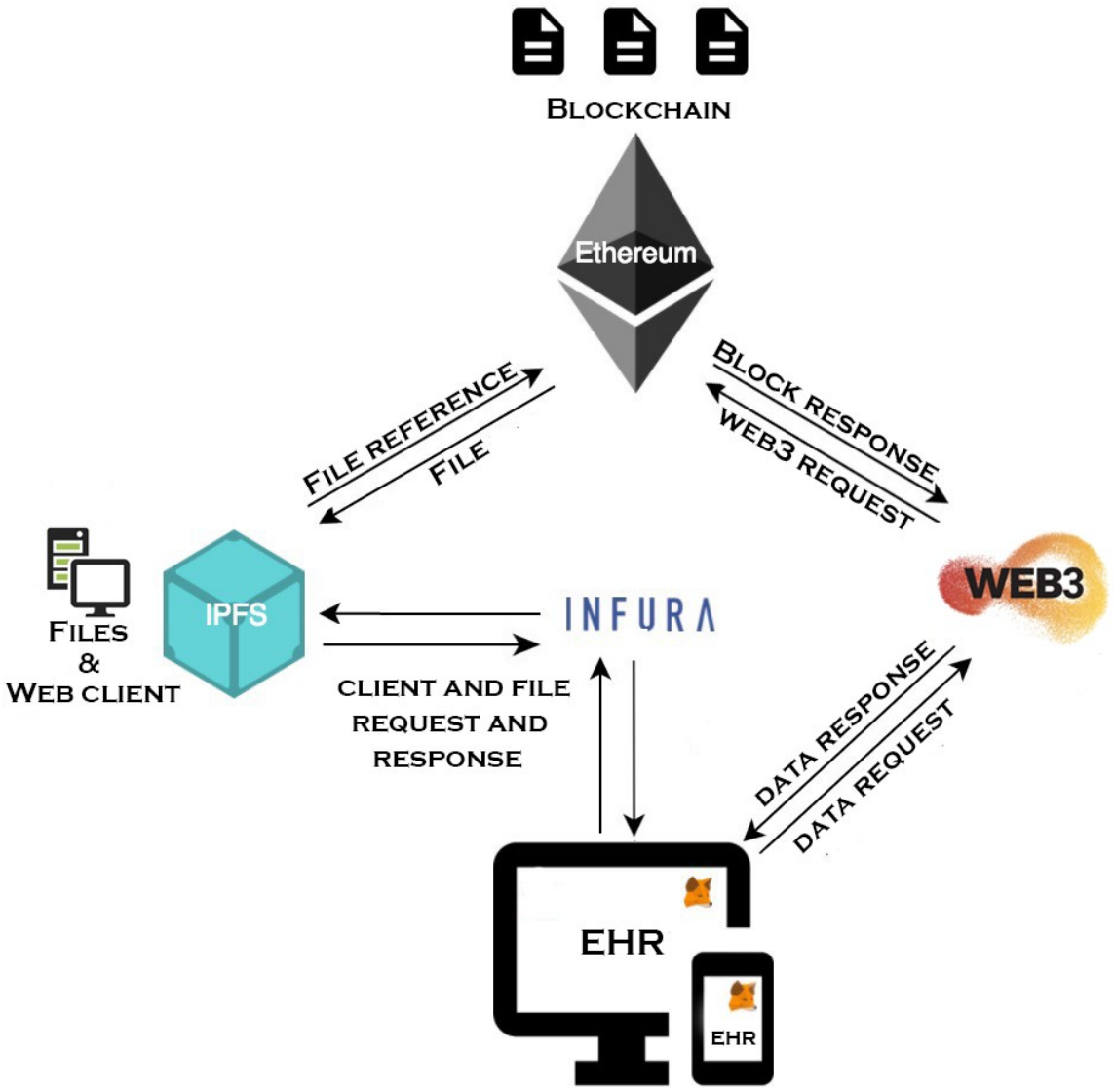


Fig 4.1 Prototype Structural Framework

As its shown on the figure the client machine requests for the webpage to the network in then Infura sends the request to IPFS then gets all the distributed client applications built in ReactJs to the client machine, by this instance client gets the webpage loaded on the browser if the client is using a Blockchain browser meaning a browser with meta mask extension. Then the client becomes able to communicate to the Ethereum Blockchain via web3. Web3 requests the Blockchain for any responses or if it's a block write operation web3 requests and awaits for the transaction to get accepted if everything succeeds a block response or a transaction hash is created. IPFS comes in place only when files are uploaded to the system.

4.4.2 System Architecture

This section explains the key components that help heterogeneous E-Health systems to interoperate. Figure 4.1 shows how the different components of system that are going to communicate. The system is made up of 4 key components plus INFURA being a middleware:

I. EHR System Client

The first layer consist of the electronic medical software. These software run on different platforms and databases. The main components of this layer are clients that facilities have been using to manage their electronic data. And any browser with Ethereum wallet is able to access the application right away.

II. Web3 API

The Web3API provides an interface between the EHR system running distributed on the IPFS and the Ethereum Blockchain. This is the main point of data exchange between the systems and the two components which store the reference (Blockchain) and the real data (IPFS).

III. Ethereum Blockchain

By using Ethereum Blockchain and its private ledger the platform used to store permanent references to the health data. This ensured data integrity is in place and a public tamper proof record for every patient file was created.

IV. IPFS

We use IPFS to store patient records that are very big and cant be stored on the blockchain. IPFS provides a distributed file system that ensured the data stored is accessible in different geographical locations. This provides the files stored on different nodes on the network come together to form the complete file. IPFS provides the EHR client and files saved on the system while reference to the files are put on the Ethereum Blockchain.

4.5 System Analysis

4.5.1 Use Case Diagram

In the following system patients are needed to be dominant because of the file to be stored on the Blockchain belongs to them. Since the Blockchain provides an immutable block of records all parties are safe from data being lost, and also secrecy and privacy is a primary concern in the system health records are only visible for the patient unless the patient grants permission to a doctor to visit his/her medical record history, and this is achieved by signing the transaction.

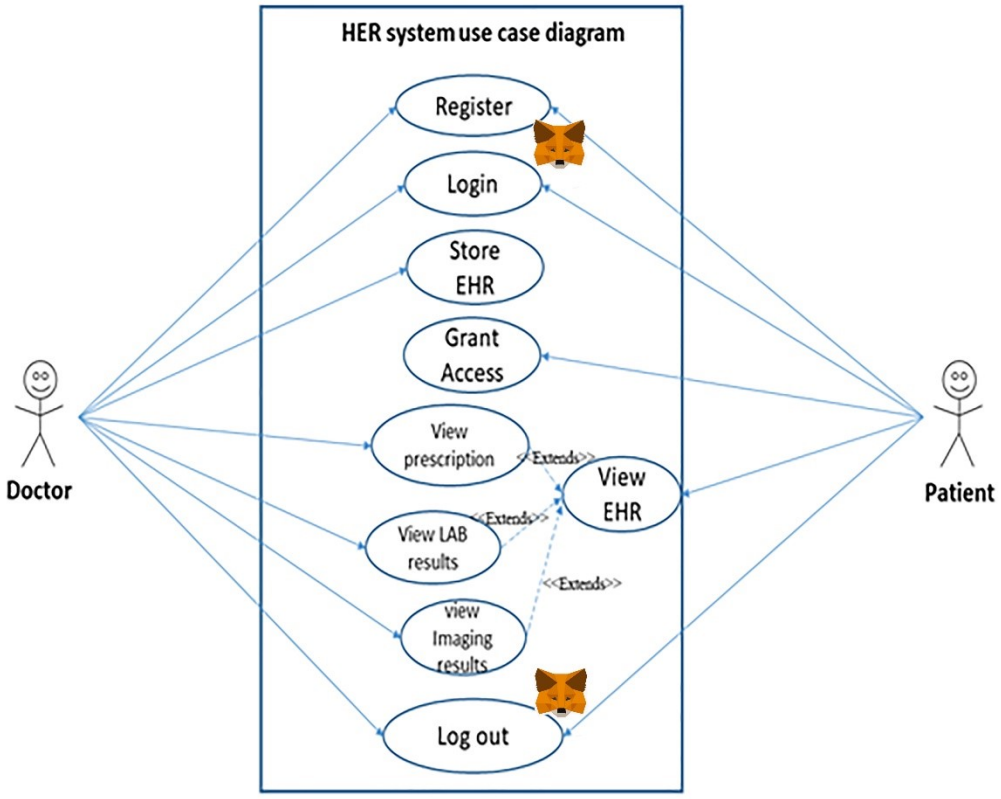


Fig 4.2 Use case diagram for EHR system

Use Case description

The following tables show all conditions that shall be mate for a certain gets access to the system.

Table 4.3 Registration Use Case

Use case:	Registration
Primary actor:	Doctor/Patient
Stakeholders:	a) Patient want to create their own secure wallet. b) Doctor wants to register to be able to access patient's data.
Preconditions:	a) Patient/doctor should have an active account. b) Patient/doctor has not registered with same device.
Post condition:	The patient and the doctor should be able to pass registration

Table 4.4 Login Use Case

Use case:	Login
Primary actor:	Doctor/Patient
Stakeholders:	a.) Patient wants to access their health records. b.) Doctor wants to review patient's data.
Preconditions:	a.) Patient/doctor have already registered b.) Patient/doctor put valid login credentials
Post condition:	Patient/doctor should login successfully and access protected features.

Table 4.5 Grant permission Use Case

Primary actor:	Doctor/Patient
Stakeholders:	a.) Patient wants to grant permission for the past records. b.) Doctor is willing to review patient's data.
Preconditions:	a.) Patient signs smart contract and passes authority for the doctor to his EHR. b.) Doctor gets proper authority on the EHR
Post condition:	Patient/doctor should follow certain steps and conditions in granting permissions.

Table 4.6 Store block record to EHR Use Case

Use case:	Login
Primary actor:	Doctor
Stakeholders:	b.) Doctor wants to diagnose the patient
Preconditions:	b.) Doctor gets the diagnostics report and adds the block record under patient's address that grants him.
Post condition:	Doctor should leave and returns back the grant.

Table 4.7 Patient visits his EHR Use Case

Use case:	Login
Primary actor:	Patient

Stakeholders:	b.) Patient wants to view medical records.
Preconditions:	b.) Patient gets all the medical records including recent once.
Post condition:	Patient leaves the system and comeback whenever he wants.

4.5.2 Sequence Diagram

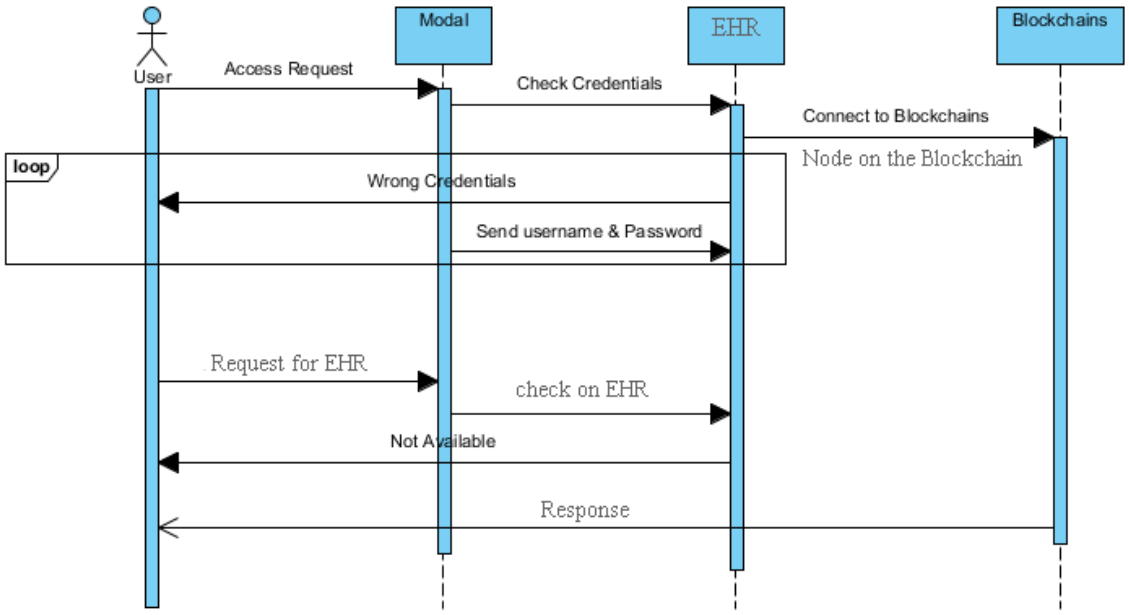


Fig 4.3 Sequence diagram for EHR system

As we its shown on the above diagram the user first requests to the client if there is any record available on the system then the client requests the EHR system which is which is on the EVM the EHR system first checks if the accessing client has is using a blockchain browser or not just by

checking if there is any wallet application connected to the browser, if so the request is forwarder to any Ethereum node on the network requesting for the data. Then the EHR system authenticates the user if he provides his account for the Ethereum wallet which is discussed on 4.7.2 Authentication of these section. If authentication succeeds. Then the respective data will be provided.

4.6 Stored data format and structure.

All health files are stored on the Ethereum blockchain inside solidity structs and on the client side they are converted into plain JavaScript objects and sent to the client as simple JSON objects. The figure below shows how the EHR data sent to the client looks like stored in json format.


```
"Patient Hisory": {
        "Diagnosis": {
            "Timestamp":"05-27-2020 05:33:21",
            "priscription": [{
                "name": "paracetamol",
                "dosage": "50mg"
                       },{
               "name": "mebendazol",
                "dosage": "100mg"
                       },{
               "name": "Amoxi Clavulanic acid",
                "dosage": "100mg"
                       }
               ],
            "labtest": [{
                "name": "stool test",
                "result": "nothing found"
            },{
                "name": "blood test",
                "result": "H pillory bacteria"
            },
               }],
            "Imaging": [{
                "name": "MRI",
                "result": "QmaQn9NqGAvzg1gamjRtVYL4hWzRmqBoxJtTj9Ubkns4wa"
            },{
                "name": "Xray",
                "result": "VYL4hWzRmqBoxJtTQmans4waQn9NqGAvzg1gamjRtj9Ubk"
            },
               }],


        },
```


Fig 4.4 Json data for single EHR Transaction

4.7 Security Design

The data being transmitted on the web based application which and each and every post from each input field is encrypted before it gets stored on the Blockchain since the EHR data is very sensitive and by all means it should be secured from eavesdropping during transmission and in storage. Several components have been incorporated so as to ensure total security of the data. AES encryption algorithm is used to encrypt the health file. The owner of the health data obtains all the transaction and block hashes of his own record so he will be able to see only his own medical data. While dealing with security of any application we need to focus on the CINA rule where Confidentiality, Integrity, Non-repudiation and Authentication play the major role. To say an application is temper proof these parameters need to get satisfied.

4.7.1 Data Integrity

To ensure data integrity the patient wallet stores hashes which act as health data file references. These references are generated when the data is saved from an external system. The integrity of the data is as strong as the hash function used which is SHA256. A reference also stored on the Blockchain to ensure that the data files cannot be tampered.

CHAPTER FIVE
PROTOTYPE IMPLEMENTATION, TESTING AND EVALUATION

5.1 Introduction

Since all the necessary data is gathered up until now we are ready to go to the development process and we will iterate through each process until we met the requirements, as discussed on chapter four the main purpose of using Agile software development methodology is that we can go back to the first stage whenever its needed.

languages and a set of web technologies. The programming language used is javascript while the system is built for web browsers and since the UI is written with REACT reusable components it can be shipped to any other environment and works perfect.

The programming language used for development of this project is Solidity which helps us develop blockchain applications for the Ethereum ecosystem. Vyper is also another programming language that is used to develop applications that are capable of running on the Ethereum Virtual Machine(EVM).

Other web technologies are also used for the development which are , JAVASCRIPT ES6, JSX JSON, IPFS and Nodejs with other bunch of libraries

This is the point of interaction between the distributed application and the platform which is composed of a public ledger and a distributed file system. The data of a patient visit which forms the health data file is stored distributed in nodes across the entire Network.

First thing First the components that are mandatory for every electronic health recording program.

1. Symptom Collection
2. Diagnosis and lab test
3. Imaging(if needed)
4. Prescription

In this system we need to implement and deploy these four modules inside a blockchain. And combining all the technologies below the Electronic Health Data recording for Ministry of health

Ethiopia is developed.

1. Ethereum (Solidity)
2. IPFS
3. Truffle Suite (Truffle + Ganache)

Overall Construction of the EHR system is simply described on figure 5.1

Fig 5.1 Overall Construction of the EHR system.

Deployment of the projects could be handled in three types of networks. Which are

a. Remix web based Test Network
b. Local Blockchain Network
c. Remote Test Network
d. Main Net

5.2 Application Mockups

There are certain things that shall be included while dealing with overall Health Data. When a patient go to medical center to he is asked and investigated for certain things, and all the information gathered from the patient is documented. The process in which patients medical record is as follows.

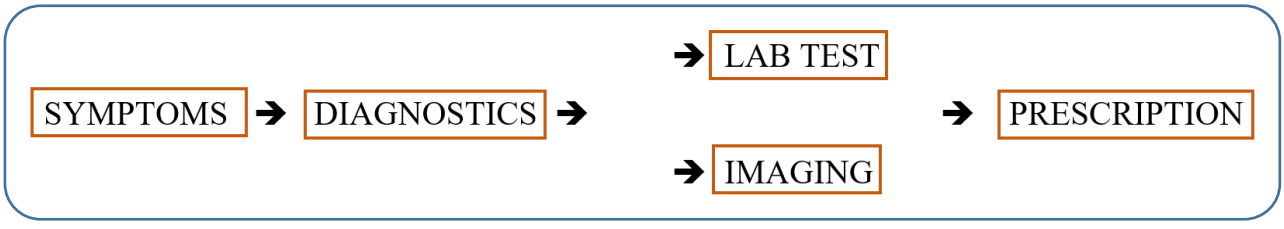


Fig 5.2 Process inside medical institutions

The process for the EHR system follows the same pattern as the real world experience as its shown on figure 5.2

The DApp is developed considering these 5 steps in mind so it has 5 modules for data management.

5.3 Setting up the project

First of all we need to setup the backend of the application before diving into the development. Since it's a new project it might have bugs. So it needs to be tested on an already configured Ethereum test network which is Remix.

a. Working with Remix

Remix is a powerful, open source tool that helps you write Solidity contracts straight from the browser. Written in JavaScript, Remix supports both usage in the browser and locally. Remix also supports testing, debugging and deploying of smart contracts and much more. And any body who wants to test his or her Solidity project before deploying it to the Ethereum blockchain can test it on the remix web based test project. Remix project with all its features is available at remix.ethereum.org.

There are certain steps that we need to follow while testing our projects on Remix.

i. Create a project
ii. Compiling the project.
iii. Deploying the project
iv. Testing the project

We might ask, what if we don't test our project on test network and just deploy it on the real Ethereum network?

While doing other projects there are trends where we don't test our projects and just deploy them and they work out of the box. But in the Ethereum blockchain each transaction is worth of money, so if we put a bad code it might work as intended but charges us a lot for every transaction which is bad trend. The official Remix Ethereum website shown below in fig 5.2

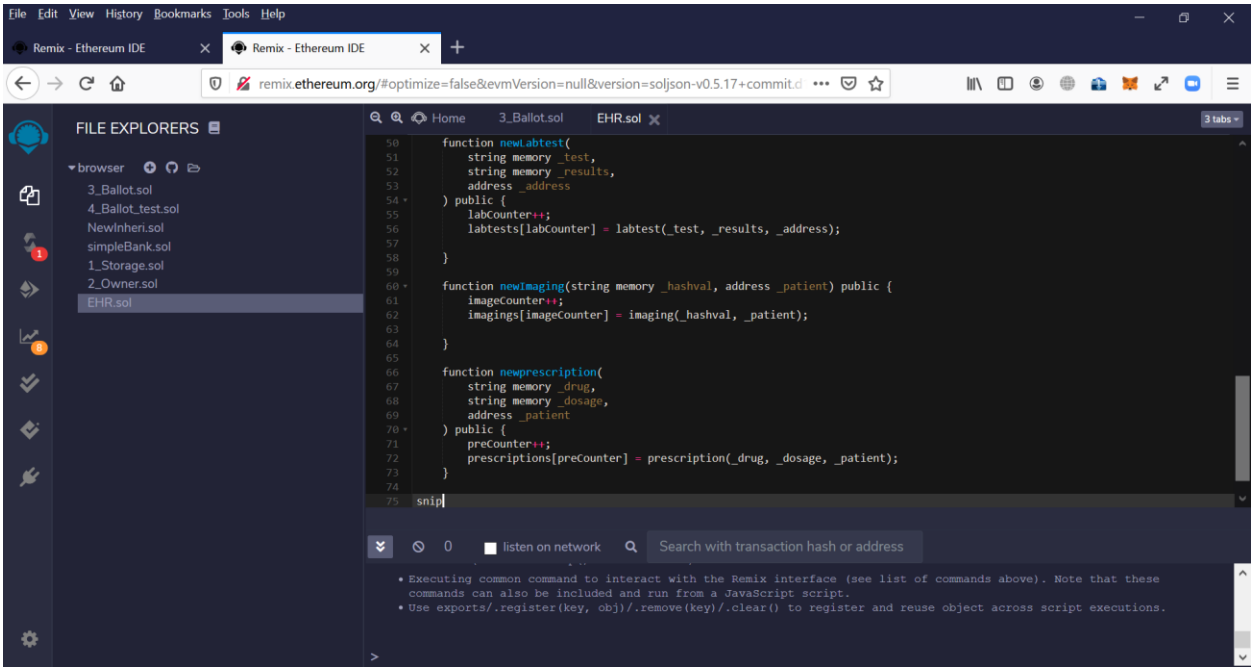


Fig 5.3 remix.etereum.org page

A fully functional blockchain testing web application is provided as soon as we submit the URL.

b. Local Blockchain Network

Working with the local blockchain network is even better trend to develop an application to be deployed on the Ethereum Blockchain or the Ethereum Virtual Machine(EVM).

This portion of the paper covers the most valuable and crucial part of the study where we can see how the bloc chain system works from end to end and it will be very interesting. There are certain steps to get started and they are described as follows.

Step 1. Install Node + NPM on our machine

Step 2 install GIT

Step 2. Install the Truffle suite

Step 3. Install Truffle Ganache

Step 4. Install Metamask browser extension

After installing the following applications and setting up everything we are ready to go ahead. The application we are developing contains both front end and backend, so the description is divided into two.

Front-end: Developed using React

Back-end: Developed using Solidity and JavaScript.

First we will download and setup a Truffle box with React embedded in it. This Box has nothing

but the structure for a basic Blockchain application. After setting up the project we will copy all the solidity code in the contracts directory so that the application is ready for deployment.

5.4 DApp Deployment (Migration)

Deploying the DApp is quite easy if its tested first. Keep in mind that deploying DApp is migrating it to the EVM. And everything write operation on the EVM coasts Gas fee so we need to compile our project before migrating it to the network(fig 5.3).

```
ETNTLGHLP036539+Local Admin@ETNTLGHLP036539 MINGW64 /d/projs/EHR (master)
$ truffle compile

Compiling your contracts...
===========================
> Compiling .\src\contracts\ElecHCare.sol
> Compiling .\src\contracts\Migrations.sol
> Artifacts written to D:\projs\EHR\src\abis
> Compiled successfully using:
   - solc: 0.5.16+commit.9c3226ce.Emscripten.clang
```

Fig 5.4 Compiling The DApp

By running the following command we can migrate the DApp(fig 5.6). as discussed on chapter four Ganache gives us 10 accounts with 100 ETH for free and while migration of a DApp some Ether will be discounted form the first Account in ganache for Gas fee(fig 5.5).

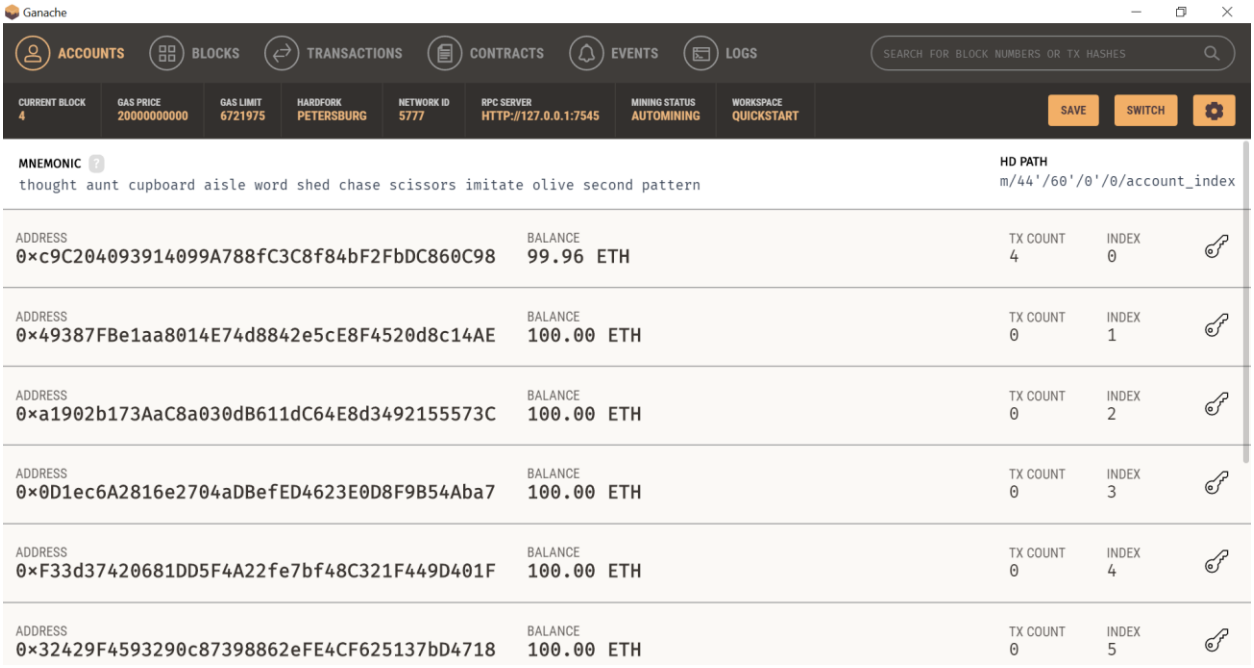


Fig 5.5 Ganache discounts Deployment Gas fee

```
ETNTLGHLP036539+Local Admin@ETNTLGHLP036539 MINGW64 /d/projs/EHR (master)
$ truffle migrate --reset

Compiling your contracts...
===========================
> Compiling .\src\contracts\ElecHCare.sol
> Compiling .\src\contracts\Migrations.sol
> Artifacts written to D:\projs\EHR\src\abis
> Compiled successfully using:
   - solc: 0.5.16+commit.9c3226ce.Emscripten.clang



Starting migrations...
======================
> Network name:    'development'
> Network id:      5777
> Block gas limit: 6721975 (0x6691b7)

1_initial_migration.js
======================

   Replacing 'Migrations'
   ----------------------
   > transaction hash:    0x579d117b8ff256ada709ee24d2f15f66c9bc22cc81bdc8945413018629410ea9
   > Blocks: 0            Seconds: 0
   > contract address:    0xdCeD2aad3F5Dac1627C72a9572127554aaDcFF23
   > block number:        46
   > block timestamp:     1594580267
   > account:             0xc9C204093914099A788fC3C8f84bF2FbDC860C98
   > balance:             99.57134974
   > gas used:            263741 (0x4063d)
   > gas price:           20 gwei
   > value sent:          0 ETH
   > total cost:          0.00527482 ETH


   > Saving migration to chain.
   > Saving artifacts
   -------------------------------------
   > Total cost:          0.00527482 ETH
```

Fig 5.6 Initial Migration of DApp from Terminal

Every block transaction inside the EVM has its own: Transaction hash, Blocks, Contract Address, Block Number, Block timestamp, Account, Balance, Gas used, Gas price, Value sent. Running the truffle console command from the terminal will give us access to the web3 object that gets access to all the methods from the deployed smart contract. Figure 5.6 shows deploying the smart contract to the blockchain and Figure 5.7 shows accessing all accounts from terminal which are provided by Ganache.

```
2_dapp_deploy.js
================

   Replacing 'ElecHCare'
   ---------------------
   > transaction hash:    0xc2840dc0114fd4fe5a6788ae954e8e351105033cae74a21b8c06d3158b8c1c50
   > Blocks: 0            Seconds: 0
   > contract address:    0xF7e1683393D0Fbb3354BbA833BE4718e36186C64
   > block number:        48
   > block timestamp:     1594580268
   > account:             0xc9C204093914099A788fC3C8f84bF2FbDC860C98
   > balance:             99.5366598
   > gas used:            1692474 (0x19d33a)
   > gas price:           20 gwei
   > value sent:          0 ETH
   > total cost:          0.03384948 ETH


   > Saving migration to chain.
   > Saving artifacts
   -------------------------------------
   > Total cost:          0.03384948 ETH


Summary
=======
> Total deployments:   2
> Final cost:          0.0391243 ETH
```

Fig 5.7 Deploying the smart contract to the blockchain

```
ETNTLGHLP036539+Local Admin@ETNTLGHLP036539 MINGW64 /d/projs/EHR (master)
$ truffle console
truffle(development)> const Dapp = await ElecHCare.deployed()
undefined
truffle(development)> Dapp
```

Fig 5.8 Calling Deployed Smart Contract from Terminal

```
truffle(development)> const Accts = await web3.eth.getAccounts()
undefined
truffle(development)> Accts
[ '0xc9C204093914099A788fC3C8f84bF2FbDC860C98',
  '0x49387FBe1aa8014E74d8842e5cE8F4520d8c14AE',
  '0xa1902b173AaC8a030dB611dC64E8d3492155573C',
  '0x0D1ec6A2816e2704aDBefED4623E0D8F9B54Aba7',
  '0xF33d37420681DD5F4A22fe7bf48C321F449D401F',
  '0x32429F4593290c87398862eFE4CF625137bD4718',
  '0xAc2705FdB899D8e0A0098C1527FB6a45D016BD1C',
  '0x78FCe8768b6e9b36D12a6EE3255e634c0cfa22Be',
  '0x6d30650c280A04317eaFEdEdeF93e6537C350214',
  '0x0a4a0Dfb0FA3B1D730aB7Ee3A8DecE5b9844714B' ]
```

Fig 5.9 return all accounts from blockchain.

5.5 User Authentication

While developing the suitable environment on the application, Authentication of users comes to mind but we don't have to worry about it because an Ethereum wallet is used for the authentication of users

Authentication is the most important thing on blockchain applications and its done with public key cryptography algorithms.

Public-key cryptography, or asymmetric cryptography, is an encryption scheme that uses two mathematically related, but not identical, keys - a public key and a private key. Unlike symmetric key algorithms that rely on one key to both encrypt and decrypt, each key performs a unique function. The public key is used to encrypt and the private key is used to decrypt.

It is computationally infeasible to compute the private key based on the public key. Because of this, public keys can be freely shared, allowing users an easy and convenient method for encrypting content and verifying digital signatures, and private keys can be kept secret, ensuring only the owners of the private keys can decrypt content and create digital signatures. On any Etherium based wallets the public key is used as the identity of the user and the private key is used for authentication of a certain user.

On the Ethereum network public keys are 32 characters long and private keys are 42 characters long sample public and private keys are shown on Fig 5.6

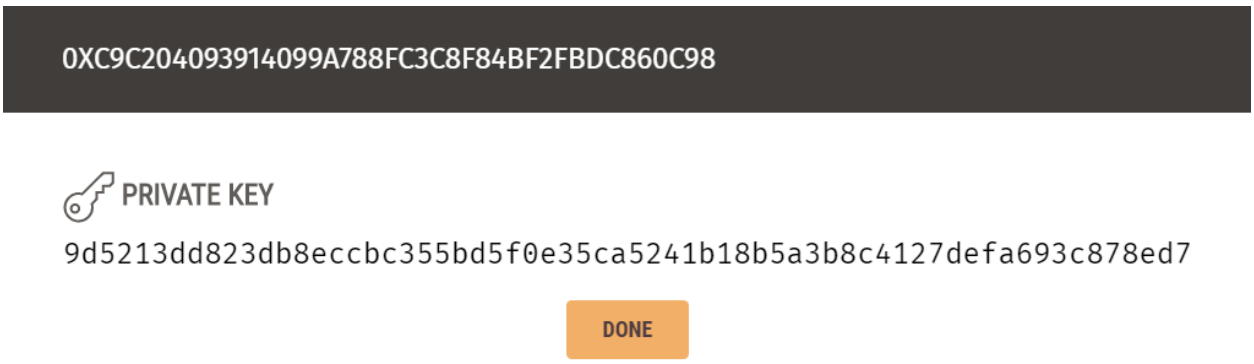


Fig 5.10 Public and Private key for DApp Deployer

5.6 System Testing

In blockchain once a system is deployed on the EVM Ethereum virtual machine it holds an address so updating the smart contract is unachievable even deleting the smart contract can't be done so the system shall be tested recursively using different tools so that it is Bug free.

For testing purposes different module is created. The type of testing used in this module was unit testing and JavaScript has special plugin libraries.

Installing the mocha unit testing library and chai assertion library each individual method is unit tested and the test result turns out positive.

5.7 DApp Deployment and Usage

Running the react development server by running npm run start command web app starts running at port 3000 and if the browser has an Ethereum wallet extension already installed(Metamask) we it will ask for password for the wallet to get connected(fig 5.11). After inserting the password we will get another popup from Metamask and we will be prompted if we want to connect to the DApp that's deployed.(fig 5.12)

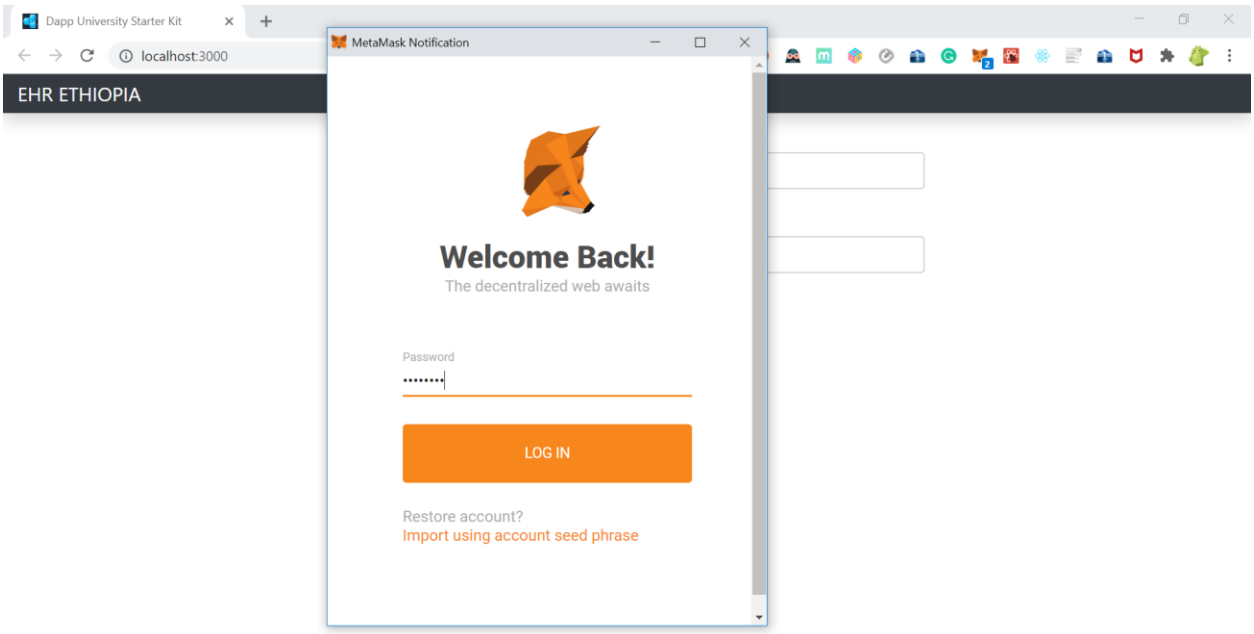


Fig 5.11 Metamask asking password for the wallet

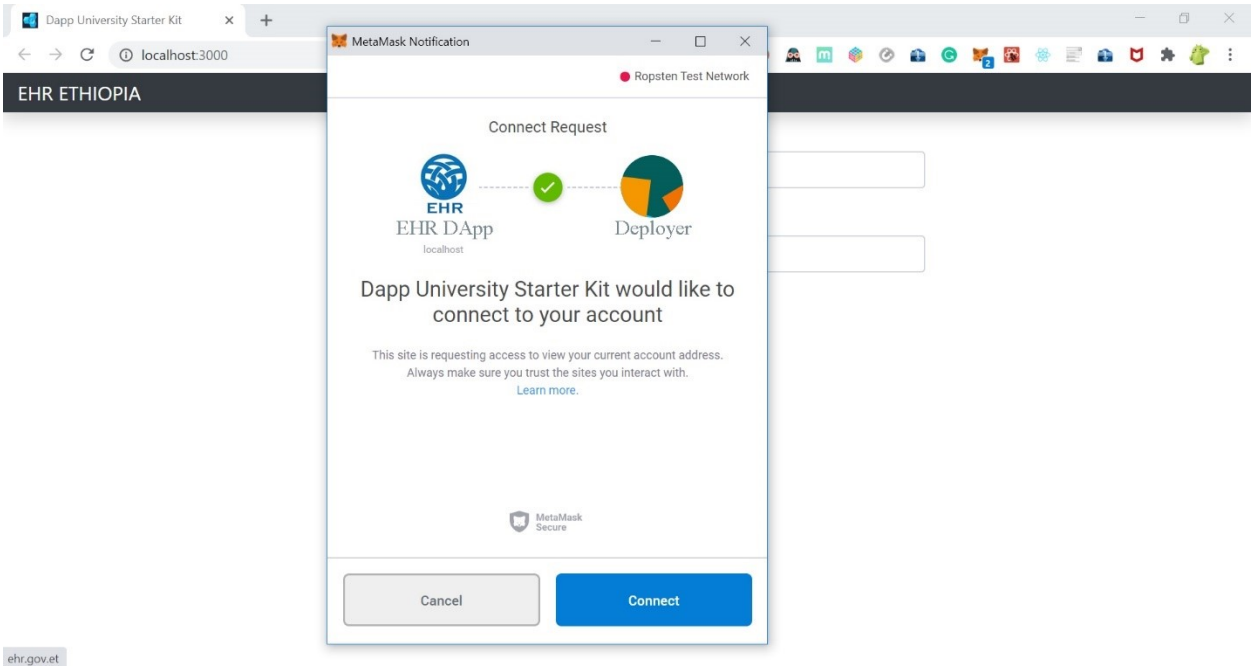


Fig 5.12 Prompt to connect to the DApp

After pressing the connect button we will be redirected to the DApp's home page. Then the doctor choses from the top links adds information to the blockchain about the patient.

Inserting symptoms looks as follows.

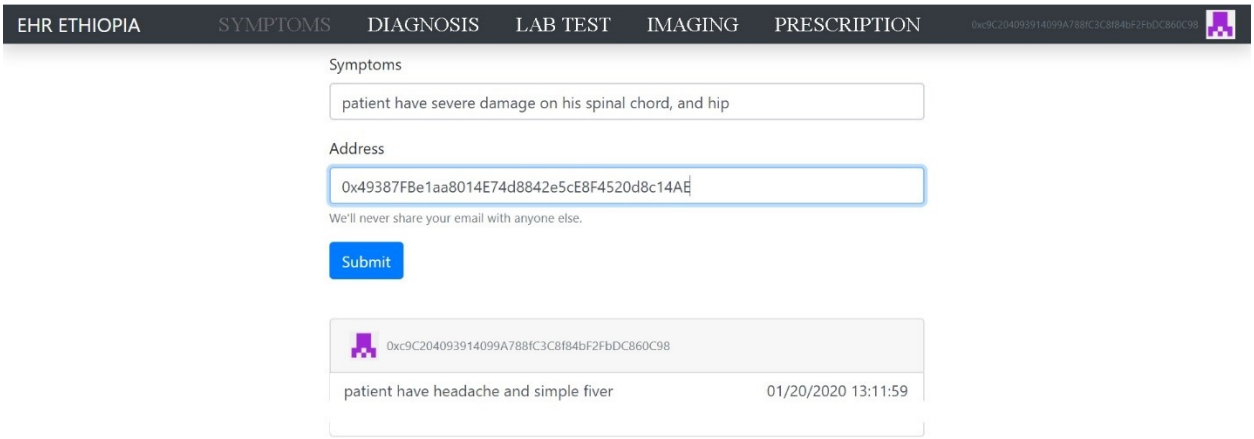


Fig 5.13 submitting symptoms to the blockchain

The same manner the doctor inserts other data to the blockchain. The only data that the blockchain is incapable of handling is the medical image.

One might ask why we can't put images to the blockchain. The reason why we cant put large files on the blockchain is each and every block on the EVM copies itself on every node across the

network so anybody who is part of the network has his or her own copies of all the transactions in the network this helps the network to be secure and immutable and no one is unable to erase or modify any data inside the blockchain. So to put every block starting from block one or the genesis block to the latest block on the network is possible cause texts wont take large data in the storages of our system. But if we put images on the blockchain the blockchain will run out of storage immediately. Even syncing to the blockchain will be very hard task to do. So the system comes up with some way wise, which is using the IPFS.

IPFS-inter planetary file system is so similar with the blockchain but its intended to put files in it and returns hush so whenever we want our file back we give the IPFS the hash and it returns our file.

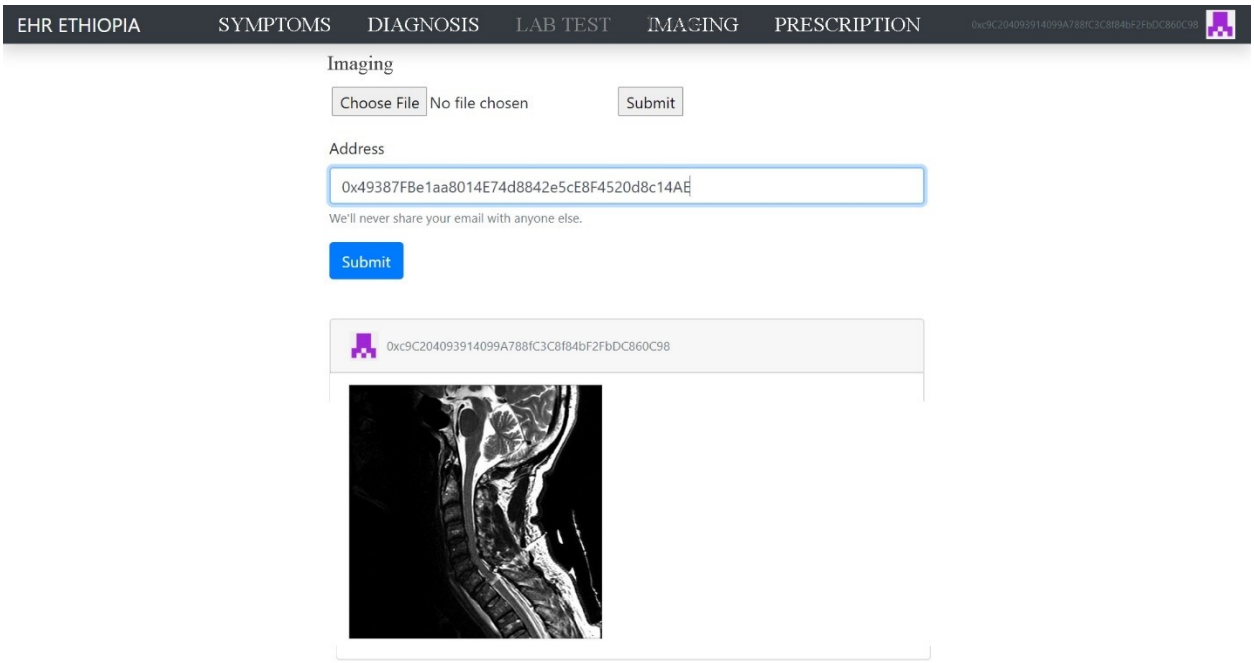


Fig 5.14 medical Imaging Prototype.

5.8 Validation

Validation of this thesis work as well as the prototype designed have gone through three steps as discussed on validity section on chapter three. And for further clarification its based on feedbacks, feedbacks from domain individuals (Doctors, Nurses, Midwives…), Patients, feedbacks from Developers.

1. Validate against those individuals who filled the questionnaire in the first place.
2. Validate against those individuals who didn't fill the questionnaire.
3. Validate against review of related works where you are going to do comparative features analysis.

1. Validation against the intended individuals meaning Medical professionals(domain experts). And other colleagues who filled the first questionnaire which was designed to spot out what major system and functional requirements needed in order to solve the problem observed on medical information archives and interoperability of them that was stated on the problem statement. As opinion of the domain individuals and other colleagues differ in different ways the researcher is forced to treat their responses differently. and responses are summarized as follows.

   Count of all the respondents is summarized on table 5.1 bellow.

Table 5.1 summary of domain expert responses

N°	Count of Responders	strongly Disagree	disagree	Neither agree nor Disagree	Agree	strongly Agree
1	Domain Experts(docs, Nurses …)	0	1	5	61	40
2	Others	0	0	16	81	33

2. Validating the prototype from other individuals that didn't even know that this research is being done and validating it against these individuals gives the researcher another view.

   The reason why individuals are questioned against the prototype though they didn't have any idea what it is or what it's about is, to

   individuals that ware questioned first have lots of questions in their mind and expect lots of things from the prototype but the individuals that didn't know about the problem and the prototype gets another and good kind of look and expected to give good feedback when they see the project at first glance.

   based up on the behavioral psychology of human beings expectations have great weight when it comes to feedbacks.

   According to a study in the university of Massachusetts in 1997 GC scientists put two group of

individuals and asked the first group for their advices on a certain matter and didn't ask the individuals on the second group. After sometime the scientists came up with a certain project , showed both groups and asked for their feedbacks. Then the result of the study came up with the first group of people gave a 47.8% positive feedbacks and average feedback from individuals from the second group become 89.8% positive so the scientists experimentally proved expectations play the biggest role in the human beings psychology.

So the researcher expects positive feedbacks from most of these individuals observing the prototype for the first time compared to the validation condition above. According to the questionnaire the researcher get results from Domain experts, Developers, Others the positive feedback 88%,78%,91% consecutively.

3. Validate against review of related works where for comparative features analysis.

And finally going through other research's

Table 5.2 summary of related works

N°	Titles	Author	Technology used	Achievements
1	Hashcash-A denial of service countermeasure,2002	Adam Back	Hashcash for POW and SHA256	Discourages Spammers by giving them exhaustive mathematical calculations
2	Security confidentiality and privacy of healthcare data	Jomin George et al., 2019	Distributed Filesystem	Checked the major concerns in security of data and applied it to health data.
3	Bitcoin: A Peer-to-Peer Electronic Cash System	Satoshi Nakamoto, 2009	Blockchain, Hash Algorithm SHA-256	Revolutionize the banking system where no banker is needed and peers trust each other with consensus mechanisms

4	How to time-stamp a digital document	S. Haber, W.S. Stornetta,1991	Hash Algorithms SHA-256,MD-5	Digital documents get their timestamps regardless of their timezone
5	The Ethereum Blockchain	Vitalik Buterin, 2012	Blockchain, distributed ledger, Pos for consensus	Makes the blockchain programmable by introducing smart contracts.
6	BASIC: Towards a Blockchained Agent-Based SImulator for Cities	Luana Marrocco et al., 2016	Ethereum Blockchain, Smart contracts	blockchain in simulated urban scenarios by considering the communication between agents through smart contracts.
7	Blockchain Technologies for the Internet of Things: Research Issues and Challenges	Mohamed Amine Ferrag,2008	Ethereum, IOT	Recommends how to resolve major security issues that the IOT might face.

Based on this validation criteria of validating against related works the research passes its validation test where it shows a significant outcome and usage.

CHAPTER SIX
CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

Lots of effort has been put in this study to ensure EHR systems can solve major security issues like transparency and confidentiality of patients medical record. The use of distributed ledger made the system possible even reliable. Several consortiums have come up with standards and technologies to ensure data is exchanged seamlessly. Doctors insisted that this data is very important in making key decisions about the patient health status. Most doctors and developers haven't been able to implement this standard in their systems. From the interviews conducted most developers expressed a difficulty in integration because of the different systems deployed in the health facilities and also different infrastructure. On the other side doctors were fearful of losing data hence the deadlock.

Since the blockchain is scalable by default and the distribution of data among nodes makes it, immutable and easy to do integration with. The solution plugs in to the existing compatible E-Health Recording systems. During the development of the system several factors had to be considered in order to answer the research questions. To create an interoperable ecosystem, the study consider the use of distributed file systems. But then again another problem emerged concerning the integrity of the data, this is because someone can put false information into the platform or modify the existing records. To curb this problem a distributed file system that uses distributed hash tables was incorporated. Apart from the integrity of records, the issue of locating health records brought in the use of an immutable distributed public ledger (Blockchain) which allows stakeholders to get location of the health data by searching the using supplied identifiers.

By employing the platform the health facilities and other stakeholders creates an ecosystem of interoperable systems. If adopted this platform enables health facilities who either own E-Health systems or are developing new ones to have a common data exchange point.

6.2 Recommendations

The findings of the study carried out were a success in an effort of evaluating the use of web/mobile technologies to enhance E-Health interoperability for health facilities. The system was able to simulate a successful creation of sample health records which were replicated to different servers across the globe. However the research felt the need to have more features and made the following recommendations:

people could easily share the records with others. Since there is no central store for the health data references, the users suggested that a secure vault should be included to ensure that patients can be able to retrieve their health wallets even if they lose data.

iii. Include the use of soap messaging. This will make integration with other web platforms easy and more Robust. This will also give developers an easy time to understand the way messages are exchanges and the kind of response they will receive.

6.3 Future works

This study had its own limitations. The main focus was developing the platform which can be used to enhance E-Health interoperability and a proof-of-concept mobile application. For further research the following aspects should be put into consideration:

I. Development of analytic tools that can present the summary of health data files in an intuitive way by tapping into the EHR blockchain.

ii. To develop systems that can help in performing health data forensics.

iii. To develop a secure vault which the patient can use to keep their health data references.

iv. Develop a messaging system based on the health data exchange standards put in place e.g.

References

[1] "Who Owns Medical Records: 50 State Comparison?" Health Information and the Law. George Washington University Hirsh Health Law and Policy Program. Aug. 20, 2015. [Online] Available: http://www.healthinfolaw.org/comparative-analysis/who-owns-medical-records-50-state-comparison

[2] U.S. Department of Health and Human Services, Office of Civil Rights. (2013). 45 CFR Parts 160, 162, and 164. "HIPAA Administrative Simplification." [Online] Available: http://www.hhs.gov/sites/default/files/hipaa-simplification-201303.pdf

[3] Mandl, Kenneth D., David Markwell, Rhona MacDonald, Peter Szolovits, and Isaac S. Kohane. "Public Standards and Patients' Control: how to keep electronic medical records accessible but private." Bmj 322, no. 7281 (2001): 283-287.

[4] Office of the National Coordinator for Health Information Technology. (2015). Report to Congress. "Report on Health Information Blocking." [Online] Available: https://www.healthit.gov/sites/default/files/reports/info_blocking_040915.pdf

[5] "Individuals' Right Under HIPAA to Access their Health Information 45 CFR § 164.524." U.S. Department of Health and Human Services. [Online] Available: http://www.hhs.gov/hipaa/for-professionals/privacy/guidance/access/ . Accessed: Aug. 8, 2016.

[6] Grossmann, Claudia, W. Alexander Goolsby, LeighAnn Olsen, and J. Michael McGinnis. Institute of Medicine of the National Academies. "Clinical Data as the Basic Staple of Health Learning: Creating and Protecting a Public Good." Workshop Summary (Learning Health System Series). National Academies Press, (2010).

[7] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system.

[8] Marco Iansiti and Karim R Lakhani. The truth about Blockchain. Harvard Business Review, 95(1):118–127, 2017.

[9] Steven Norton. Cio explainer: What is Blockchain? The Wall Street Journal, 2, 2016.

[10] Christian Cachin and Marko Vukoli´c. Blockchains consensus protocols in the wild. arXiv preprint arXiv:1707.01873, 2017.

[11] Martin A Makary and Michael Daniel. Medical error-the third leading cause of death in the us. BMJ: British Medical Journal (Online), 353, 2016.

[12] IBM publications on healthcare

[13] Agile development important things to know, https://jumpgrowth.com/agile-development-important-things-you-need-to-know/

[14] Things to know about BFT, https://www.binance.vision/Blockchain/byzantine-fault-tolerance-explained

[15] Capgemini. (2015, November 13). Blockchain: A Fundamental Shift for Financial Services Institutions. Retrieved

[16] December 15, 2015, from Capgemini: https://www.capgemini.com/resources/blockchain-a-fundamental-shift-forfinancial-services-institutions

[17] CPrime. (2014). Hybrid Projects: How Can Waterfall and Agile work Together. Retrieved from CPrime: www.cprime.com

[18] Das, K. T., & Mahapatra, K. D. (2012). Addressing the Influencing Factors for Interoperability in a Global Domain .

[19] IOSR Journal of Computer Engineering (IOSRJCE) , 41-48.

[20] David, C., David, F., Ed, M., & Pat, P. (2005, August). Some Current Approaches to Interoperability. Retrieved November 2015, from http://www.sei.cmu.edu/reports/05tn033.pdf

[21] DeNardis, L. (2011). Standards and eHealth. ITU-T Technology Watch Report.

[22] DeNardis, L. (2012). E-health Standards and Interoperability. ITU.

[23] Depardon, B., Mahec, G. L., & S´eguin, C. (2013, February 15). Analysis of Six Distributed File Systems. Retrieved December 2015 , from Hal Archives:

[24] https://hal.archivesouvertes.fr/file/index/docid/789086/filename/a_survey_of_dfs.pdf

[25] Fischer, M. J. (1983, June). Yale University Computer Science. Retrieved November 2015, from Yale University:

http://cpsc.yale.edu/sites/default/files/files/tr273.pdf

[26] Golafshani. (2003). Understanding Reliability and Validity in Qualitative Research .

[27] HL7. (n.d.). Introduction to HL7 Standards. Retrieved December 15, 2015, from Health Level Seven International: http://www.hl7.org/implement/standards/

[28] Hoffer. (2001). Modern Systems Analysis and Design.

[29] IBM. (2015, February 9). RESTful Web services. Retrieved May 20, 2016, from https://www.ibm.com/developerworks/library/ws-restful/ws-restful-pdf.pdf

[30] Improving the User Experience. (n.d.). Retrieved January 12, 2016, from Usability.gov: http://www.usability.gov/how-to-and-tools/methods/usability-testing.html

[31] International Health Standards Development Organization. (2014, 10 14). Guidance on use of SNOMED CT and

[32] LOINC together. Retrieved November 2015, from International Health Standards Development Organization:

[33] http://ihtsdo.org/fileadmin/user_upload/doc/download/xdoc_SnomedCtAndLoincGuide_Current-enUS_INT_20141014.pdf?ok

[34] Kilwake , J., Matoke , N., Waliaro , A., Wanyembi , G., & Ogao , P. (January 2012). Current Status of E-Health in Kenya and Emerging Global Research Trends. International Journal of Information and Communication Technology Research.

[35] Kothari. (2004). Research Methodology: Methods and Techniques, . New Delhi: New Age International limited. .

[36] The Byzantine Generals Problem, LESLIE LAMPORT, ROBERT SHOSTAK, and MARSHALL PEASE SRI International

[37] Lee, J. A., Long, A., Steiner, J., Handler, S. G., & Wood, Z. (2015). Blockchain Technology and Legal Implications of ‘Crypto 2.0’. Retrieved December 2015, from

[38] http://www.gibsondunn.com/publications/Documents/Lee-LongBlockchain-Technology-BNA-Banking-03.31.2015.pdf

[39] Ministry of Health. (2010). EMR SYSTEM AND INFORMATION EXCHANGE - INTEROPERABILITY. 10-12

[40] Needham & Company LLC. (2015, October). The Blockchain Report: Welcome to the Internet of Value. RetrievedDecember 2015, from Stotj: http://storj.io/TheBlockchainReport.pdf

[41] Nodejs Foundation. (2016). Node.JS. Retrieved May 17, 2016, from https://nodejs.org/en/

[42] Özsu, T. M., & Valduriez, P. (1991, August). DISTRIBUTED DATABASE SYSTEMS: WHERE ARE WE NOW? Retrieved January 2016, from Pennsylvania State University: http://wps.pearsoned.co.uk/wps/media/objects/10977/11240737/Web%20chapters/Chapter%2012_WEB.pdf

[43] Protocol Labs. (n.d.). IPFS. Retrieved May 17, 2016, from https://ipfs.io/

[44] Ragib, H., Zahid, A., William, Y., Larry, B., & Roy, C. (2008). A Survey of Peer-to-Peer Storage Techniques forDistributed File Systems. Retrieved October 2015, from VirginiaTech department of computer science http://courses.cs.vt.edu/cs5204/fall08-kafura/Papers/FileSystems/Survey.pdf

[45] Ross, S. M., & Morrison, G. R. (n.d.). EXPERIMENTAL RESEARCH METHODS. Retrieved February 2016, from The University of Memphis: http://www.aect.org/edtech/ed1/38.pdf

[46] Satoshi, N. (2007). Bitcoin: A Peer-to-Peer Electronic Cash System. Retrieved February 2016, from Bitcoin https://bitcoin.org/bitcoin.pdf

[47] Scott, W. A. (2008). The Agile System Development Life Cycle (SDLC). Retrieved January 2016, from Ambysoft http://www.ambysoft.com/essays/agileLifecycle.html

[48] Scholer, K. (2016, February). An Introduction to Bitcoin and Blockchain Technology. Retrieved March 23, 2016, from http://www.kayescholer.com/docs/IntrotoBitcoinandBlockchainTechnology.pdf

[49] Sparks. (2001, 2001). Methods and Tools.

[50] Tripathi, P. K., & Tripathi, M. (2012, April). A Framework of Distributed Database Management Systems. Retrieved December 2015, from ijarcsse: http://www.ijarcsse.com/docs/papers/April2012/Volume_2_issue_4/anand_tripathi.pdf

[51] Turnkey Linux. (2016, April 14). LAMP Stack - Web Stack (MySQL). Retrieved May 2016, from

https://www.turnkeylinux.org/lampstack

[52] UN. (2016, January 22). United Nations Statistics Division. Retrieved May 17, 2016, from http://data.un.org/Data.aspx?d=POP&f=tableCode%3A240

[53] University of Missouri. (2001). Information Systems Analysis . Retrieved from University of Missouri http://www.umsl.edu/

[54] Wang, J. (2001). University of Missouri-St. Louis. Retrieved January 2016, from Object-Oriented Analysis Methodology: http://www.umsl.edu/~sauterv/analysis/488_f01_papers/wang.htm

[55] Were, E., & Jamah, A. (2013, May 11th). The shocking truth on 'killer doctors'. Retrieved December 2015, from Standard Media: http://www.standardmedia.co.ke/article/2000083356/the-shocking-truth-on-killer-doctors

[56] Wilkinson, S., Boshevski , T., Brandoff, J., & Buterin, V. (2014, December 15). Storj A Peer-to-Peer Cloud Storage Network. Retrieved November 2016, from Storj: http://storj.io/storj.pdf

[57] Belatrix. (2015). Functional Testing Best Practices. Retrieved from Belatrix Software http://www.belatrixsf.com/

[58] Benson, T. (2010). Principles of Health Interoperability HL7 and SNOMED. London: Springer. Bodenreider, O. (2008). Issues in Mapping LOINC Laboratory Tests to SNOMED CT . AMIA 2008 Symposium Proceedings (pp. 51-55). U.S. National Library of Medicine.

[59] Boroujerdi, R., & Wolf, C. (2015, December). What if I Told You…. Retrieved December 16, 2015, from Goldman Sachs Global Investment Research: http://www.theblockchain.com/docs/Goldman%20Sachs%20Blockchain%20Report.pdf

[60] Buterin, V. (2015, November 21). A NEXT GENERATION SMART CONTRACT & DECENTRALIZED APPLICATION PLATFORM. Retrieved January 13, 2016, from Github https://github.com/ethereum/wiki/wiki/White-Paper

[61] CPrime. (2014). Hybrid Projects: How Can Waterfall and Agile work Together. Retrieved from CPrime www.cprime.com

[62] A.Back, "Hashcash- a denial of service counter measure

"http://www.hashcash.org/papers/hashcash.pdf, 2002

[63] Prasad Patil, Exploratory data analysis https://towardsdatascience.com/exploratory-data-analysis-8fc1cb20fd15

[64] Blockchain technology and healthcare https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6517629/

[65] Ins and outs of blockchain https://innovatemedtec.com/digital-health/blockchain

[66] Application of Blockchain to Maintaining Patient Records in Electronic Health Record for Enhanced Privacy, Scalability, and Availability https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7010942/

[67] Blockchain Technology in Healthcare in 2019 https://theblockbox.io/blockchain-technology-in-healthcare-in-2019/

[68] A medical records managing and securing blockchain based system supported by a Genetic Algorithm and Discrete Wavelet Transform https://www.sciencedirect.com/science/article/abs/pii/S1389041718301177

Appendix I: Questionnaire I

Date 03/11/2019

Dear Participant

My name is Anania Mesfin Sileshi and I am a post graduate student at St. Mary University. For my final thesis, I am working on a Blockchain technology and its impact in the healthcare sector and what major roles it could play in terms of data management, security, interoperability, building an immutable platform and how it could help the society , I am inviting you to participate in the research study by completing the attached surveys.

There is no compensation for responding nor is there any known risk. In order to ensure all information will remain confidential, please do not include your name. copies of this thesis will be provided to my St. Mary University Advisor and St. Mary University School of Computer Science. If you choose to participate in this thesis, please answer all questions as honestly as possible and return the completed questionnaires promptly in person. Participation is strictly voluntary and you may refuse to participate at any time.

Thank you for taking time to assist me in my educational endeavors. The data collected will provide full information regarding the work on how medical facilities like Hospitals, Clinics, Pharmacies take care of patients data and bringing the system and medical record(Data) to safe environment. completion and return of the questionnaire will indicate your willingness to participate in this study. If you require additional information or have questions, please do not hesitate to contact me at my contact listed below.

Sincerely

Anania Mesfin

+251911066609

Coofrozen@gmail.com

Asrat Beyene(PhD)

ambnn2001@gmail.com

Questionnaire

1. Type of the organization you are working in

   ☐ Governmental ☐ Non-Governmental ☐ Private

2. Work type of the organization

   ☐ Clinic ☐ Hospital ☐ Pharmacy ☐ Other H.C.O

3. Does the organization have any data management trend(Digital/Traditional)

   ☐ Yes ☐ No

4. Comment on the data Management of the organization/organization

_____

_____

_____

_____

_____

5. How do you exchange patients health data with other medical institutions

_____

_____

_____

_____

6. What patients comment on the data management of the organization (or complain)?

_____

_____

_____

_____

_____

7. To what extent the organization keeps the secrecy of patient's medical record?

   ☐ Very much ☐ Somehow ☐ No it doesn't

8. Has the organization ever faced a patient's medical history loss?

   ☐ Yes ☐ No

9. If your answer for question no 7 was YES then what kind of major the organization took to resolve the problem?

_____

_____

_____

_____

10. What kind of action would be taken if the organization faces a total patients data (Medical History) destroyed by some kind of natural reasons like earthquake, fire or theft. Do you think the data could be recovered? If yes then how?

_____

_____

_____

_____

11. What do you think is the best option to handle patients medical record the safest way?

_____

_____

_____

_____

11.please provide any other comment if you have one (not mandatory)

_____

_____

_____

_____

Thank you for your willingness for spending your precious time in filling the questionnaire and be part of the survey

Appendix II: Interview Question

Since patients in hospitals are in a rush collecting questionnaires from patients is quite impossible so in this thesis work interview questions are presented for patients so that they can provide answers on the go.

1. How often do you visit Medical Facilities like Hospitals for medical examinations?
2. Do you remember when the last time was you went to a hospital for checkups and do you keep record of your visits to hospitals?
3. Do you keep copies of your medical results every time you visit a hospital?
4. Do you believe keeping all copies of your medical record in hand helps to let you know where your medical status is at and directs you to check up on a regular basis or else do you think it just makes you panic every time you see your record?
5. How do you describe the interoperability of medical records in medical institutions you have ever visited?
6. Please provide any other comment regarding data archive in medical facilities?

Thank you for your willingness for spending your precious time in answering the interview questions and be part of the survey

Appendix III: Questionnaire II

Part I: General information

- ❖ You can put tick (√) "□" mark sign for your response from the given choices.

1. please specify your position? Patient ☐ Developer☐ Other ☐

For each of the questions below, circle the response that best characterizes how you feel about the statement, where : 1 = strongly Disagree, 2 = disagree, 3 = Neither agree nor Disagree, 4 = Agree, and 5 = strongly Agree

Part II: Questions related to strategy

	strongly Disagree	disagree	Neither agree nor Disagree	Agree	strongly Agree
	1	2	3	3	5
1. would you say the application is usable					
2. did you Face any challenges on while logging in the system					
3. The prototype can handle the secrecy of once personal health info.					
4. The prototype has a strong authentication system					
5. The prototype addresses major Medical data handling transactions.					
6. The application can handle the immutability of health data					
7. The application provides more security than the current systems?					

8. Basic security majors are considered in the prototype					
9. The prototype solves the issues of interoperability?					
10. Evaluate the overall system from your perspective(observation).					

11.please provide any other comment if you have one (not mandatory)

_____

_____

_____

Appendix III: Code snippet EHR Smart Contract

```
pragma solidity >=0.4.21 <0.6.0;
contract ElecHCare {
    uint256 public sympCounter;
    uint256 public labCounter;
    uint256 public imageCounter;
    uint256 public preCounter;
    mapping(uint256 => symptom) public symptoms;
    mapping(uint256 => labtest) public labtests;
    mapping(uint256 => imaging) public imagings;
    mapping(uint256 => prescription) public prescriptions;
    constructor() public {
```

```
}
struct symptom {
    string symp;
    address patient;
}
struct labtest {
    string testname;
    string results;
    address patient;
}
struct imaging {
    string hashval;
    address patient;
}
struct prescription {
    string drugname;
    string dosage;
    address patient;
}
function newSymptom(string memory _symp) public {
    require(bytes(_symp).length > 0);
    sympCounter++;
    symptoms[sympCounter] = symptom(_symp, msg.sender);
}
function newLabtest(
    string memory _test,
    string memory _results,
    address _address
```

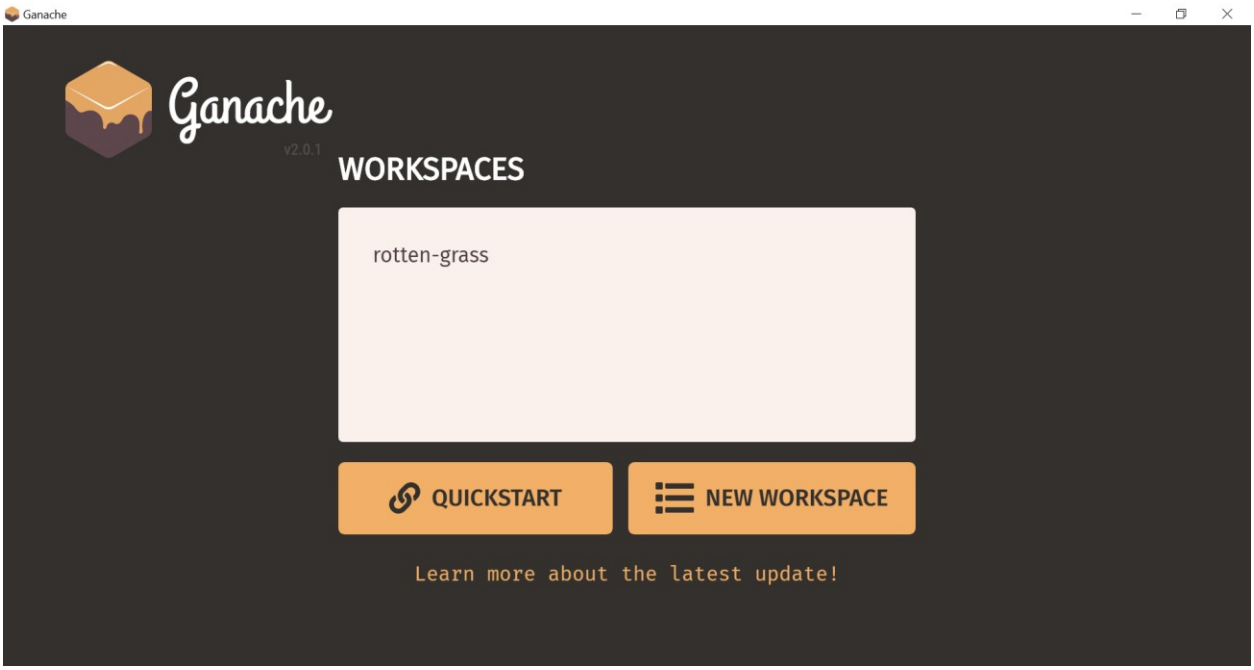
```
) public {
    labCounter++;
    labtests[labCounter] = labtest(_test, _results, _address);
}
function newImaging(string memory _hashval, address _patient) public {
    imageCounter++;
    imagings[imageCounter] = imaging(_hashval, _patient);
}
function newprescription(
    string memory _drug,
    string memory _dosage,
    address _patient
) public {
    preCounter++;
    prescriptions[preCounter] = prescription(_drug, _dosage, _patient); }
}
```

Get pending Transactions from the Blockchain

```
web3.eth.getPendingTransactions().then(console.log);
> [
    {
        hash: '0x9fc76417374aa880d4449a1f7f31ec597f00b1f6f3dd2d66f4c9c6c445836d8b',
        nonce: 2,
        blockHash: '0xef95f2f1ed3ca60b048b4bf67cde2195961e0bba6f70bcbea9a2c4e133e34b46',
        blockNumber: 3,
        transactionIndex: 0,
        from: '0xa94f5374fce5edbc8e2a8697c15331677e6ebf0b',
        to: '0x6295ee1b4f6dd65047762f924ecd367c17eabf8f',
        value: '123450000000000000',
        gas: 314159,
        gasPrice: '2000000000000',
        input: '0x57cb2fc4'
        v: '0x3d',
        r: '0xaabc9ddafffb2ae0bac4107697547d22d9383667d9e97f5409dd6881ce08f13f',
        s: '0x69e43116be8f842dcd4a0b2f760043737a59534430b762317db21d9ac8c5034'
    },...,{
        hash: '0x9fc76417374aa880d4449a1f7f31ec597f00b1f6f3dd2d66f4c9c6c445836d8b',
        nonce: 3,
        blockHash: '0xef95f2f1ed3ca60b048b4bf67cde2195961e0bba6f70bcbea9a2c4e133e34b46',
        blockNumber: 4,
        transactionIndex: 0,
        from: '0xa94f5374fce5edbc8e2a8697c15331677e6ebf0b',
        to: '0x6295ee1b4f6dd65047762f924ecd367c17eabf8f',
        value: '123450000000000000',
        gas: 314159,
        gasPrice: '2000000000000',
        input: '0x57cb2fc4'
        v: '0x3d',
        r: '0xaabc9ddafffb2ae0bac4107697547d22d9383667d9e97f5409dd6881ce08f13f',
        s: '0x69e43116be8f842dcd4a0b2f760043737a59534430b762317db21d9ac8c5034'
    }
]
```

Appendix III: Truffle Ganache

The home page of the Ganache local blockchain.



New workspace setup with 10 accounts having 100 ETH each one of them for simulation of blockchain transactions.