**PROJECT REPORT (PHASE-1)**

*on*

# BLOCKCHAIN AND MACHINE EARNING RECOMMENDING SYSTEM TO FIND A DOCTOR

*Submitted in partial fulfillment for the award of the degree*

*of*

**BACHELOR OF TECHNOLOGY**

*in*

**COMPUTER SCIENCE AND ENGINEERING**

*by*

**Jaskirat Singh Bhatia (Reg. No: RA1511003010011)**

*Under the supervision of*

**Dr. M. Prakash**
**(Associate Professor, Department of Computer Science and Engineering)**


**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**
**FACULTY OF ENGINEERING AND TECHNOLOGY**
**SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**
(Under section 3 of UGC Act, 1956)
SRM Nagar, Kattankulathur- 603203
Kancheepuram District
**NOVEMBER 2018**

# BONAFIDE CERTIFICATE

Certified that this project report titled **"BLOCKCHAIN AND MACHINE LEARNING RECOMMENDING SYSTEM TO FIND A DOCTOR"** is the bonafide work of Jaskirat Singh Bhatia (Reg. No: RA1511003010011) who carried out the project under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form part of any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion or any other candidate.


**Signature of the Guide**
**Dr. M. Prakash**
Associate Professor
Department of Computer Science and Engineering
SRM Institute of Science and Technology
Kattankulathur- 603203

**Signature of the HOD**
**Dr. B. AMUTHA**
Professor& Head
Department of Computer Science and Engineering
SRM Institute of Science and Technology
Kattankulathur- 603203


**DATE: 02/11/2018**

# TABLE OF CONTENTS

# ABSTRACT

Electronic Health Records (EHRs) are entirely controlled by hospitals instead of patients, which complicates seeking medical advices from different hospitals. Patients face a critical need to focus on the details of their own healthcare and restore management of their own medical data. The rapid development of blockchain technology promotes population healthcare, including medical records as well as patient-related data. This technology provides patients with comprehensive, immutable records, and access to EHRs free from service providers and treatment websites. In this paper, to guarantee the validity of HERs encapsulated in blockchain, we present an attribute-based signature scheme with multiple authorities, in which a patient endorses a message according to the attribute while disclosing no information other than the evidence that he has attested to it. Furthermore, there are multiple authorities without a trusted single or central one to generate and distribute public / private keys of the patient, which avoids the escrow problem and conforms to the mode of distributed data storage in the blockchain. By sharing the secret pseudorandom function seeds among authorities, this protocol resists collusion attack out of N from N−1 corrupted authorities. Under the assumption of the computational bilinear Diffie-Hellman, wealso formally demonstrate that, in terms of the unforgeability and perfect privacy of the attribute-signer, this attribute-based signature scheme is secure in the random oracle model. The comparison shows the efficiency and properties between the proposed method and methods proposed in other studies.

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF SYMBOLS AND ABBREVIATIONS

| ABBREVIATION | | EXPANSION |
|---|---|---|
| ICT | : | Information and communication technology |
| PMS | : | Pharmacy Management System |
| WHO | : | World Health Organization |
| DFD | : | Data Flow Diagram |
| HIS | : | Hospital Information System |
| EHR | : | Electronic Health Record |

# INTRODUCTION

## 1.1 Blockchain and Machine Learning Recommending System

Now a day's Blockchain and ML, plays a great role in different fields or areas among the Health Care System. This leads to various studies and researches being conducted to selected health care facilities. It is necessary to ensure a technologically appropriate, equitable, affordable, efficient, and environmentally adaptable and consumer friendly system, designed to fully utilize the Blockchain and Machine Learning technology for the maximum benefit in the health care society.

Here computers have great relevant on storing data's securely and ease access on them in short period of time.

In order to store the consumer information securely, a blockchain is being build. The blockchain is a robust, integrated, secure and trustworthy technology. The blockchain system encrypts the data stored in each layer with the data stored in the next layer, making it almost impossible to crack it without the key which brings us to the next technology in Blockchain that is the public/private key system. The public/private key of Blockchain works like any other system, that is all the hospitals have the public key of the patient and can use it to enter the details of the patient, and the patient will have his private key which he and only he can use to access his data, hence making it very secure.

Machine Learning using the algorithms of Deep Learning will be used to a high level of extent in our project. Deep Learning algorithms on clustering to figure out the type of disease the patient has and informing the patient about it. Also using Machine Learning Recommending algorithms to recommend the most suitable doctor for that particular disease nearest to the patient.

### 1.2 Problem Statement

Data is one thing in this world everyone is after. The ley man is afraid to give up his personal details into the world for privacy reasons. The lay man in this world is also very busy and doesn't has enough time to spend on his health and take care of it properly. Thus, our aim is to develop a system which is secure and also one which saves the time of the user by predicting his disease from his symptoms and recommend the most suitable doctor accordingly.

### 1.3 Feasibility Study

A feasibility analysis involves a detailed assessment of the need, value and practicality of a proposed enterprise, such as systems development. The process of designing and implementing record keeping systems has significant accountability and resource implications for an organization. Feasibility analysis will help you make informed and transparent decisions at crucial points during the developmental process to determine whether it is operationally, economically and technically realistic to proceed with a particular course of action.

Most feasibility studies are distinguished for both users and analysts. First, the study often presupposes that when the feasibility document is being prepared, the analyst is in a position to evaluate solutions. Second, most studies tend to overlook the confusion inherent in system development – the constraints and the assumed attitudes.

### 1.4 Operational feasibility

People are inherently resistant to change, and computers have been known to facilitate change. An estimate should be made of how strong a reaction the user staff is likely to have toward the development of a computerized system. It is common knowledge that computer installations have something to do with turnover, transfers, retraining, and changes in employee job status. Therefore, it is understood that the introduction of a candidate system requires special effort to educate, sell and train the staff on new ways of conducting business.

# LITERATURE SURVEY

Healthcare may be a huge application situation of blockchain, and blockchains utilized in aid area unit known as health blockchain. In general, blockchain blocks area unit open and therefore the transactions in them area unit public. If some privacy data are involved in these transactions, they will be leaked. Owing to aid system involving an excellent deal of privacy information, certain security mechanisms should be designed to shield this privacy information in health blockchain. Furthermore, because the core of security mechanisms is that the key management schemes, the suitable key management schemes ought to be designed before blockchains may be utilized in aid system. Here, according to the features of health blockchain, the authors use a body detector network to style a light-weight backup and economical recovery theme for keys of health blockchain. The authors' analyses show that the scheme has high security and performance, and it can be used to protect privacy messages on health blockchain effectively and to market the appliance of health blockchain.

Based on the traditional dictionary learning methods, that neglects the link between the sample and therefore the wordbook atom, we have a tendency to propose the weighted mechanism to attach the sample with the wordbook atom during this paper. Meanwhile, the traditional dictionary learning method is prone to cause over-tting for patient classication of the limited training data set. Therefore, this paper adopts l2-norm regularization constraint, which realizes the limitation of the model space, and enhances the generalization ability of the model and avoids over-tting to some extent. Compared with the previous shallow dictionary learning, this paper proposed the greedy deep dictionary learning.

Based on the traditional dictionary learning methods, that neglects the link between the sample and therefore the wordbook atom, we have a tendency to propose the weighted mechanism to attach the sample with the wordbook atom during this paper. Meanwhile, the traditional dictionary learning method is prone to cause over-tting for patient classication of the limited training data set. Therefore, this paper adopts l2-norm regularization constraint, which realizes the limitation of the model space, and enhances the generalization ability of the model and avoids over-tting to some extent. Compared with the previous shallow dictionary learning, this paper proposed the greedy deep dictionary learning.

Patient similarity learning aims to find appropriate distance metrics to measure patient pairs for a specific task. To capture the historical information of patient' record, a proper way to represent longitudinal EHR is necessary. Moreover, we need a way to learn the similarity degree or distance between each pair of patients. In this paper, we propose two patient similarity learning frameworks on EHR dataset. The raw EHRs are feed into a CNN model which captures the consecutive sequential information to learn a vector representation. Then soft-max based supervised classification method and triplet loss based distance metric learning method are used to learn the similarity of patient pairs. Experimental results on disease prediction and patient clustering show that CNN can better represent the longitudinal EHR sequences, and our endtoend similarity frameworks outperform state-of-the-art distance metric learning methods.

The immense potential of this technology shows up wherever, until now, a trusted third party was necessary for the settlement of market services. With Blockchain, direct transactions suddenly become possible, whereby a central actor, who controlled the data, earned commission or even intervened in a censoring fashion, can be eliminated. This disruptive character, which underlies Blockchain technology, will strongly affect the balance of power between existing market players in healthcare. It will also promote new digital business models and digital health initiatives. Due to the fact that, in the future, (data) intermediaries can be avoided, this technology opens new doors with respect to how market interactions in healthcare can be conducted. Blockchain thus has an immense potential for the future and will show disruptive changes in the healthcare industry.

# 3. PROPOSED SYSTEM

## 3.1 Existing System

The modern is still very naïve and the general approach if any person is sick is that first the person realises that he is out of his comfort zone and is not feeling well then her gives it some time sometimes a few days to confirm the fact that he is not feeling well. Then he looks for a general doctor nearby and goes to him for consultation and then either the doctor gives him medicine or tells him to go to a specialist as the knowledge of a human is limited to what all he has learnt in his college and from his experience. After this, the patient goes to the specialist and takes various testes and gets his diagnosis and the doctor recommends him some medicine and a dosage for the medicine.

The customer then goes to the shop and purchases the medicine required. So a lot of time is wasted and the person gets tired. If he wants to exchange the product, once again he goes to the shop and replaces them. The complete process depends on the physical interactions.

The existing system is paper-based involving high amount of paper work and manpower requirement. Even though computerized systems are used in some places, they are not web-based and are very insecure and improperly managed. So, the current patient management procedure is very uneconomical, inflexible and insecure to meet user demands.

### 3.1.1 Drawbacks

- Expensive Medical Details
- Improper diagnosis by inexperienced doctors
- Lack of proper technology to predict the system
- Inaccurate data of medicine stock and requirement
- Potential of information decay
- Data is not immediately visible
- No standardization

## 3.2 Proposed System

### 3.2.1 Technical feasibility

Technical feasibility centers around the existing computer system (hardware, software, etc.) and to what extend it can support the proposed addition. For example, if the current computer is operating at 80 percent capacity – an arbitrary ceiling – then running another application could overload the system or require additional hardware. This involves financial considerations to accommodate technical enhancements. If the budget is a serious constraint, then the project is judged not feasible.

### 3.2.2 System Analysis

It is the most creative and challenging phase of the system life cycle. The analysis phase is used to design the logical model of the system whereas the design phase is used to design the physical model.

Many things are to be done in this phase, we began the designing process by identifying forms, reports and the other outputs the system will produce. Then the specify data on each were pinpointed. we sketched the forms or say, the displays, as expected to appear, on paper, so it serves as model for the project to begin finally we design the form on computer display, using one of the automated system design tool, that is PyCharm.

After the forms were designed, the next step was to specify the data to be inputted, calculated and stored individual data items and calculation procedure were written in detail. File structure such as paper files were selected the procedures were written so as how to process the data and procedures the output during the programming phase. The documents were design ion the form of charts.

Output design means what should be the format for presenting the results. It should be in most convenient and attractive format for the user. The input design deals with what should be the input to the system and thus prepare the input format. File design

deals with how the data has to be stored on physical devices. Process design includes the description of the procedure for carrying out operations on the given data.

### 3.2.3 System Requirements

The system services and goals are established by consultation with system user. They are then defined in details and serve as a system specification. System requirement are those on which the system runs.

| | | |
|---|---|---|
| Python version | 3 | 5.0 or greater |
| Memory | 512 MB | 1 GB or more |
| Free disk space | 300 MB | 1 GB or more |
| Processor speed | 800 Mhz | 1.5 Ghz or faster |

### 3.2.4 Cost/ Benefit analysis

Economic analysis is the most frequently used method for evaluating the effectiveness of a candidate system. More commonly known as cost benefit analysis, the procedure is to determine the benefits and savings that are expected from a candidate system and compare them with costs. If benefits overweigh costs, then the decision is made to design and implement the system. Otherwise, further justification or alterations in the proposed system will have to be made if it is to have a chance of being approved. This is an ongoing effort that improves in accuracy at each phase in the system life cycle.

**Costs:**

1. Cost of new computer approximately Rs. 22,000/-
2. Cost of operating system approximately Rs. 5000/-

**Benefits:**

- Standardized data fields
- Immediate document retrieval
- Saving storage space
- Keeps data secure
- Easy to use, update and maintain

### 3.2.5 Proposed Architecture

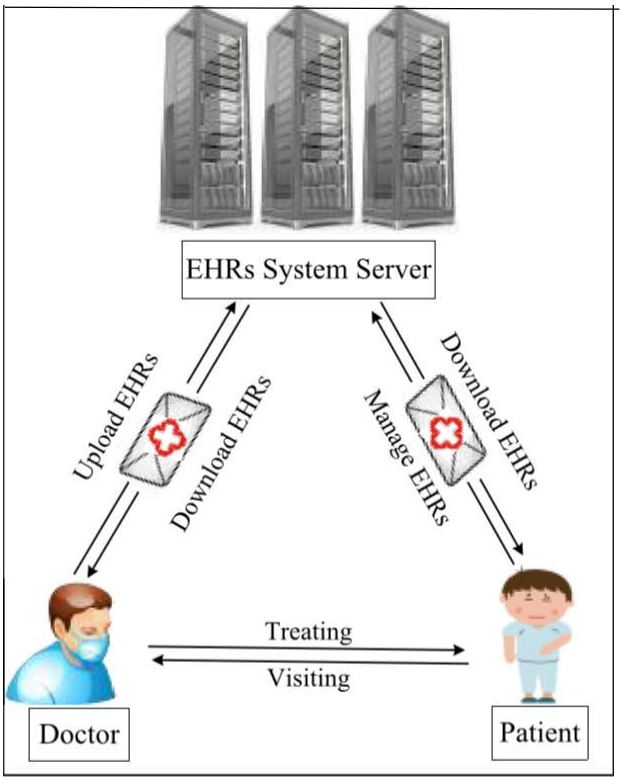


**Fig 1. 0th Level DFD**

In the $0^{th}$ level of the DFD the patient goes to the doctor and tells him all his symptoms and then the doctor diagnoses his symptoms and treats him. The doctor then uploads the EHRs to the EHR System Server and uses the patients public key for his records. Then if the patient wants to view or modify his EHRs, he can do so with his private key.

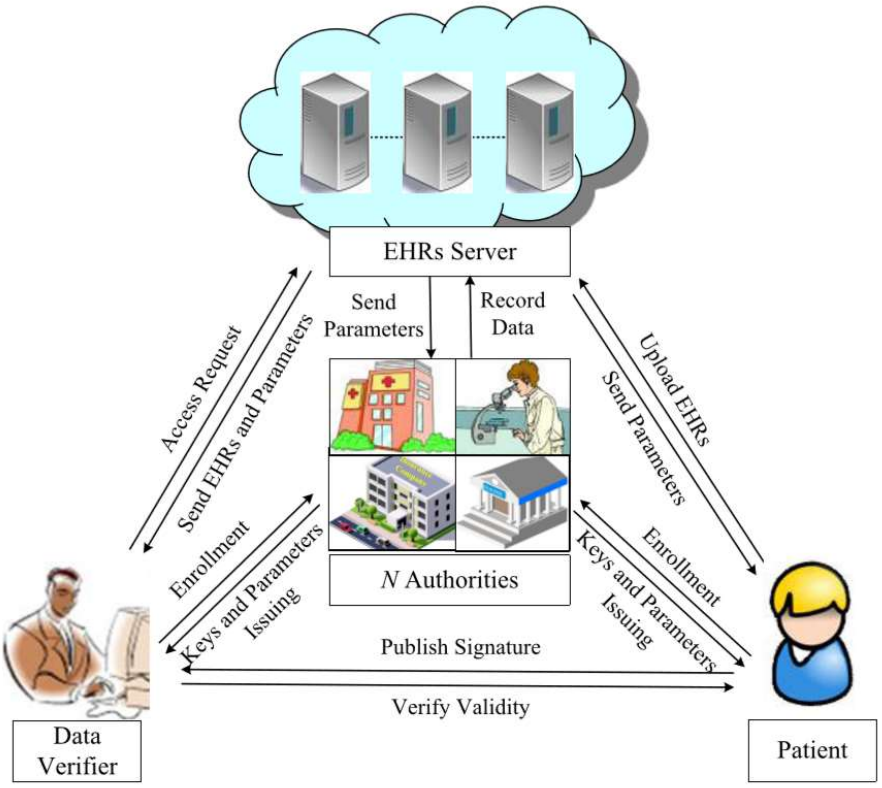


**Fig 2. $1^{st}$ Level DFD**

In level one of the DFD the client select the mode of the action i.e. how the keys and the parameters will be passed in the EHRs and the EHRs system. The flow of data in the above diagram will be as follows:

First the doctor will treat the patient and give all the details regarding the patients disease and his symptoms to a data verifier. The data verifier will then check the data and remove any flaws in it and make it syntactical to be stored in the blockchain.

The data verifier will then access the EHRs server. If the verifier is not enrolled or is new, he will enrol to the EHRs and then get his own unique key and other parameters. Then the verifier sends a request to the EHRs Server to enter the details, if the request is completed, the verifier then sends his key and other parameters along with the public key of the patient.

The EHRs server is just like a cloud storage server, which is responsible for storing and transmitting the EHRs. N authorities are various organizations, such as hospitals, medical insurance organizations, medical research institutes, etc., which are responsible for accepting the enrolment and exchange of patient information. Patients may create, manage, control and sign their own EHRs and define the predicate while the data verifier is allowed to access this signature and verify the correctness.

After the data has been entered on the EHRs Server, the patient can then access the data through his private key. If the patient is not enrolled in the EHRs System, he sends his information to the N Authorites and they enrol him as required. Then to access the data, he sends his private key to the EHRs Server and if the key is confirmed, the Server asks for various other parameters and then finally, the patient can access and modify his own records.

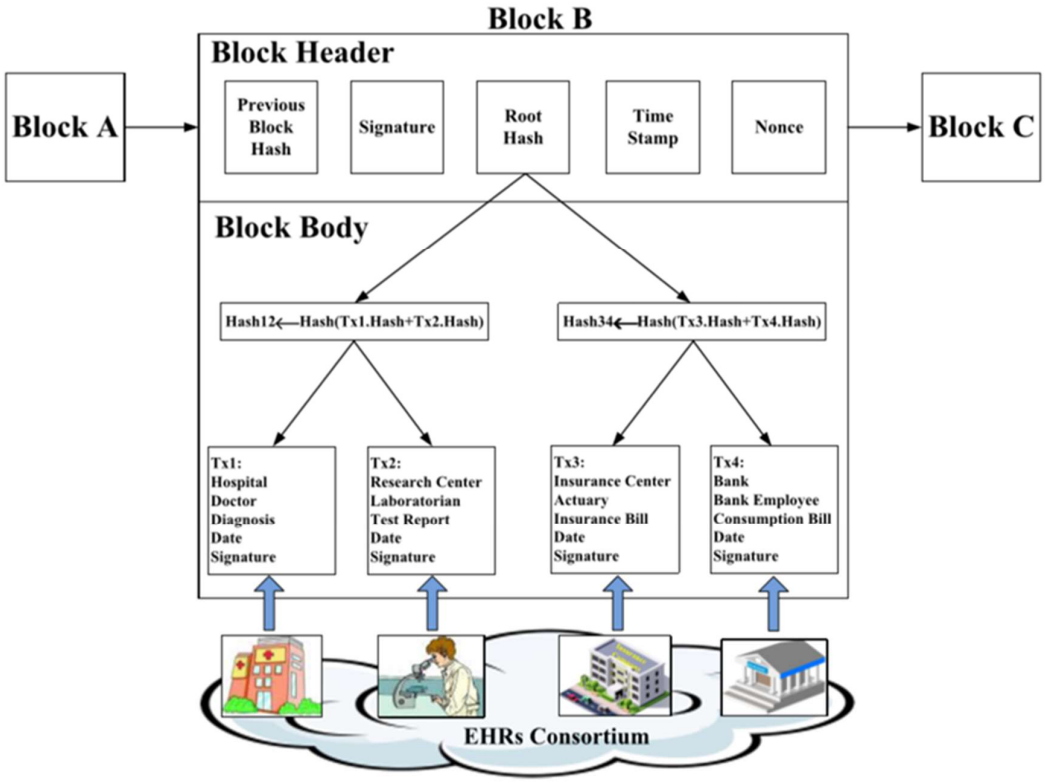


**Fig 3. 2nd Level DFD**

The EHRs Server is explained in the second level of DFD

Assuming that there is an EHRs system in a cloud storage platform, which consists of some departments, such as hospitals, pharmaceutical departments, insurance departments, disease research departments and so on, EHRs systems can be jointly managed. All departments can offer services for patients together and restrict the rights of each department to prevent EHRs abuse. Thus, an EHRs system with a blockchain structure is designed as shown in Fig. 3. Suppose that every patient owns one blockchain of healthcare alone. After being treated in a hospital, all the information including EHRs, consumption records, insurance records, etc. is encapsulated in one block. Patient treatments at different times will be generated in different blocks. Then, a series of blocks are EHRs system in blockchain. Every patient owns this chain by himself, after being treated in one hospital, all the information related to patient is encapsulated in one block.
generated according to the time sequence and a healthcare blockchain of this patient is constructed. Authorized entity might look over the health records of this patient by means of his blockchain, and has powerless to tamper the data in established block (such as drug allergy and dosage). When the patient goes to be treated in other clinical departments or hospitals next time, the new entity needs to identify this patient and authenticate his available blockchain, which could save the medical resources and avoid the repeated detection. To meet the requirement of distributed structure in EHRs system, we employ attributes based signature with multiple authorities to address the above application.

### 3.2.6 Algorithm

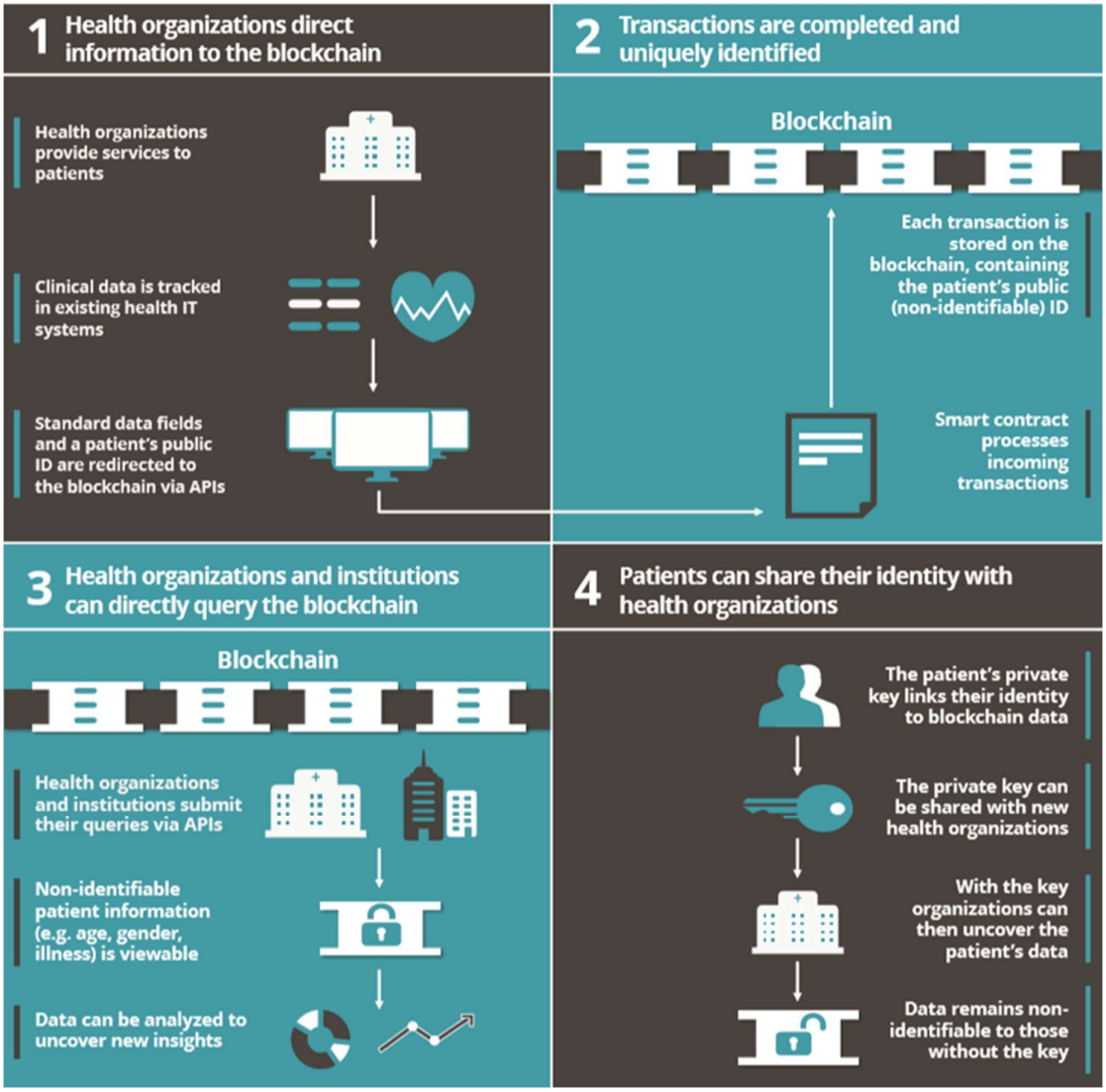


**Fig 4. Pharmacy System Analysis**

- Health organisations direct information to the blockchain
  - Health organisations provide services to patients
  - Clinical data is tracked in existing health IT systems
  - Standard data fields and a patients public ID are redirected to the blockchain via APIs
- Transactions are completed and uniquely identified
  - Each transaction is stored on the blockchain, constraining the patients public ID
  - Smart contract processes incoming transactions
- Health organisations and institutions can directly query the blockchain
  - Health organisations and institutions submit their queries via APIs
  - Non-identifiable patient information is viewable
  - Data can be analysed to uncover new insights

- Patients can share their identity with health organisations
    - The patients private key links their identity to blockchain data
    - The private key can be shared with new health organisations
    - With the key organisations can uncover the patients data
    - Data remains non-identifiable to those without the key

# REFERENCES

1. Abdo-Rabbo A, Al-Ansari M, Gunn BC, Suleiman BJ. The use of medicines in Oman: Public knowledge, attitudes and practices. Sultan Qaboos Univ Med J. 2009;9:124–31.

2. Balamurugan E, Ganesh K. Prevalence and pattern of self-medication use in coastal regions of South India. Br J Med Pract. 2011;4:a428.

3. Harrison R, Walton M, Manias E et al. The missing evidence: a systematic review of patients' experiences of adverse events in health care. Int J Qual Health Care 2015;27:424–42.

4. K. Khile, A. Abedaullah A study on the performance of the pharmacy information system within the Moroccan hospital sector. IEEE Explore/ 7731721

5. World Health Statistics 2015 - World Health Orgnization

6. Organization of Pharmaceutical Producers of India (OPPI) 48th Annual Report 2013-2014

7. Kotwani A, Ewen M, Dey D, Iyer S, Lakshmi P, Patel A, et al. Medicine prices and availability at six sites in India: Using the WHO-HAI methodology. Indian J Med Res. 2007;125:645–54.