Research Project Report

Intelligent system for supporting patient and physician

Siting Chi 12870212

Jiayan Zhao 12868574

Xinchen Wang 12772149

Table of Content

1. Abstract

SmedicalS is an intelligent system that can be used to store patient's EMR/EHR, reduce patients waiting time for primary diagnosis, and build a relationship between different hospitals and physicians and patients. This system includes three data models to store different types of data. Ontology has been used in the system to stand at patients' side, to help patient get quickly diagnosis, and direct patients to the right department of the nearest hospital. Relational data is used to check medicine information. DLT is used to store patients' EMR/EHR. The goal of this research is to design a new system to improve the efficiency of primary diagnosis, to help patients to find the most suitable department and hospital, and to conduct a primary diagnosis on the platform, especially in populated cities as Shanghai in China.

Keywords: Urban–rural gap, Healthcare resources, Allocation, Inequity

2. Introduction

2.1 Research Problem and Objectives

Our project is to design a platform by cooperating ICT (Information Communication Technology) and DLT to help improve the efficiency of outpatient as well as organize the medical data for further supervision and research. This is an artificial design of how EMR is read, stored, updated and retrieved to support patient and physician activity such as primary diagnosis and making an appointment online. This project also shows how EMR data and other relational data is separately managed according to their different required level for security.

2.2 Background

The imbalance in the allocation in healthcare resources between urban and rural areas has become a main focus of the recent medical reforms adopted in China.

Financial, material, and human resources are the foundation of the healthcare services provided by the Ministry of Health of the People's Republic of China to ensure the accessibility of quality healthcare services.

The number of the healthcare clinics located in towns and townships declined in the decade from 2000 to 2010 from 49229 to 37836, respectively, and the number of healthcare clinics located in villages has rapidly decreased from 709,458 in 2000 to 514,920 in 2003, but had since grown to 648,424 by 2010. Between 1980 and 2006, the number of beds in both hospitals and clinics in China increased from 2.184 million to 3.271 million, which including a 249.83% increase in cities from 903,000 to 2.256 million, while decreasing by 79.08% in rural areas from 1.2817 million to 1.103 million.

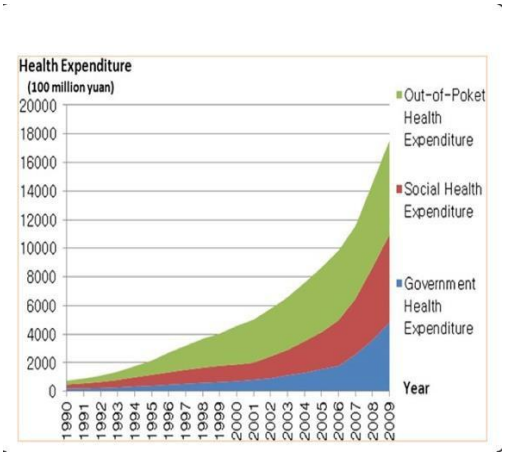

Figure 2.1

From 2000 to 2010, the average number of healthcare technicians available in rural healthcare centers has steadily increased from 23.0 in 2002 to 25.9 in 2007 and to 30.4 in 2010. Furthermore, the number of doctors per village and per 1000 rural population has gradually in- creased over this period, but is still far below the na- tional average.

Healthcare is considered a social welfare service, thus the government is required to improve and maintain equal access of healthcare services to both urban and rural residents. Generally speaking, the reasonable allo- cation of healthcare resources can be thought of as an “equilateral triangle,” which refers to allocation of healthcare resources that flow into rural areas with greater populations and greater healthcare needs. As a matter of fact, the allocation of healthcare resources in China is actually an “inverted triangular”(Yiyi Chen,2014).

Large population also emerges large volume of medical data. It cannot be effectively used if without proper management. The adoption of EMR has become a popular way to manage medical records in developed countries. In 2018 China Electronic Medical Record Industry Analysis Report, it is mentioned that over 30 hospitals in Beijing have shared EMR. Physicians in those hospital can check patient’s EMR at any time without additional operation. It is convenient when there is emergency but it also risky because it doesn’t need patient’s authorization through this process. We found there is a gap in data security.

Figure 2.2

The security problem for EMR not only shows in China. According to our former research, current healthcare service providers in developed countries such as AdvancedMD usually focus on the billing management and appointment management. There are few service providers have a integrated management solution for EMR. However, separated system means high security risk for such sensitive data as EMR.

From applying this intelligent system, we expect to improve the efficiency of hospital outpatient departments by reducing the workload. By research, we identify primary diagnosis which is following a set of static steps can be automatically done on the platform at home through website/application. Patient can finish the few steps of primary diagnosis online and make appointments with physician without waiting in the hall for hours. The adoption of ICT also can provide lower cost for both hospital and patient. Conbination of relative functions together into one single system and use DLT to enhance the security performance make it easier to manage.

3. Literature Review

Development of ICT has great impact on the way of healthcare. The common way of seeing a doctor is go to the hospital, and register, and wait until the doctor says you are the next one. In populated city such as Shanghai in China, number of physicians is much less than patients which means patients have to wait for a long time before they get proper treatment. In order to improve the efficiency of service, some hospital started adopting ICT, such as EHR to transform the healthcare service.

EHR (Electronic Health Record), also called EMR (Electronic Medical Record)has many definitions. Amy Merlino (p, 140) addresses a basic but important statement of EHR: “The EHR is where we all come together, even if we are located in separate clinical facilities, to document the work we do, to order the services we need, to receive the results that will inform our decisions, and to educate one another about the details of our patient’s situation and progress.” This indicates EHR has no limitations of locations. The difference between paper chart and EHR is on a paper chart, the provider may refer back to the problem list the next time the patient visits, but EHR is a shared resource between multiple doctors and the patients. In Louis Capponi (p, 151)’s article, it is believed that EHR is not only accesses between physicians, but also helps patients choose a physician to visit. It will become easy to access a provider virtually rather than give patients a physician’s office. In this model, patients can save time for travelling between different hospitals and seeing different doctors. Also, doctors will find themselves free to practice in any way and in any location.

The first step of improvement of patient medical care is the catalyst for the electronic health record. Instead of the hand-written health records, the premise of electronic health records is to improve the quality of health care. Hamilton (2011, p. 22) considers some core functions of an EHR. Firstly, EHR records health information and data which includes past medical information, narratives, and diagnoses. Some patients might allergic to something. Some patients might have adverse drug reaction of drug interaction. All these issues have to be clearly record. Secondly, EHR improves patient support. The original EHR improves care quality, and reduces medical costs in some way. However, different hospitals do not share the same EHR. This caused many issues like extra time and money cost caused by duplicated testing. This is contrary to EHR’s original intention. Medical decision support is the

third functionality which improves the care quality such as primary diagnosis. In some simple situation, it also helps drug prescribing. Assessed information such as prescription details, drug interaction, allergies, dosing information diagnoses, management of diseases and symptoms, and adverse reactions can be present through EHR. Lastly, EHR helps the communication between physicians and patients more smoothly. Also, physicians can communicate with other professionals smoothly. For example, information between specialists, primary care physicians, radiology, laboratories, and pharmacies are available.

However, the question is firstly focus on how we enter the data into the system. With the rapidly development of computer technology, EHR becomes more complete. Keyboard is not only the tool that can use to enter data, other methods such as voice recognition systems, electronic handwriting recognition, as well as a series of templates.

Secondly, the birth of a new concept often comes with barriers and benefits. At the beginning, a lack of standards for EHR systems persecute people. It is because a complete EHR includes various program such as patient's information, checking history, medical history, and so on. Various programs offered different features and the security of confidential information data is hard to set. Then, the birth of new technology is exiting but not acceptable to everyone at the same time. Physicians may think EHR takes more time to enter data rather than handwriting. Also, patients may worry about their data security, and they do not want other people to know what is wrong with their body.

On the other hand, President George W. Bush in his 2004 stated, "By computerizing health records, we can avoid dangerous medical mistakes, reduce costs, and improve care." Different people's writing styles are different. Instead of handing writing, EHR creates better clinic information. Physicians have to write down information about the patient immediately with hand writing. Otherwise, physicians might forget some information. Moreover, physicians must go back to the office to find the patient's documents within the trading hours. However, with EHR, it can access to the patient's medical information without limitation because EHR is based on database.

Another important benefit driving the use of EHR creates better patient care. Improved both the quality and quantity of patient care thorough clinical information. For patients, nothing is more important than get diagnose early and quick, and get the diseases fixed. In order to reduce waste of source, time, and money, EHR records all medical test information. This can avoid repetition of labs and tests.

Since EHR is full of information, how to protect it from unauthorized using? The answer we found is blockchain technology. Blockchain technology addresses the interoperability challenges of today's healthcare IT systems and becomes a technology standard that enables individuals, healthcare organizations, healthcare organizations and medical researchers to share the potential of electronic health data securely. Blockchain will run on widely used and reliable commercial hardware. Commodity hardware provides useful calculations at the lowest cost. The hardware is based on open standards and is manufactured by multiple vendors. It is the most cost-effective and efficient architecture for health and genomic research. Excessive blockchain hardware capabilities can be shared with healthy researchers and facilitate the discovery of new drugs and treatments faster.

Blockchain technology also addresses interoperability challenges in a healthy IT ecosystem. The medical IT system will use the open API for integration and data exchange with healthy blockchain. The open API is based on industry best practices. They are easy to use and eliminate the need for complex point-to-point data integration between different systems.

Blockchain will enable patients, healthcare professionals and researchers to access a shared data source to obtain timely, accurate and comprehensive patient health data. The Lake of Blockchain data structure combined with data can support a wide variety of health data sources, including mobile applications from patients, wearable sensors, electronic medical records, documents and image data. The data structure is flexible, scalable and will be able to accommodate unforeseen data, which will be available in the future.

4. User requirement

In this section, we have conducted face-to-face interviews and E-mail interviews as well as send out questionnaires on the Internet. Our questions will focus on the needs of patients and physicians during the primary medical diagnosis. In this phase, the interview would be well-structured, clear, and coherent. Also, the questionnaire would be the quantitative research method which would contain both closed-ended and open-ended questions. However, this phase provides basic questions that the physician might ask, and personal identity will not be collected from the user research process.

Through questionnaire and interviews, we expected to get information about

1) Whether patients will accept such intelligent system and what is their biggest concern about using ICT technology
2) What is the biggest problem in visiting a doctor for primary diagnosis
3) What is the normal flow of a primary diagnosis.
4) Whether physician will accept such intelligent system and what is their biggest concern about using ICT technology

Then we developed use case diagram to locate the user requirements.

4.1 Questionnaires

Topic : Did you ever meet some problems in hospital registration?

(1) Gender [Multiple choice question]
1. male
2. Female
Result:

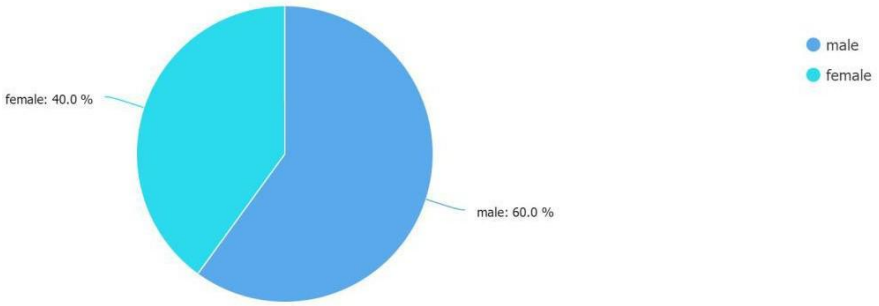

Figure 4.1

(2) What kind of illness do you usually go to the hospital to register? [Multiple choice questions]
1. Hand, foot and mouth disease
2. Cold and fever
3. Dizziness and nausea
4. Gastric discomfort
5. Trauma
Result:

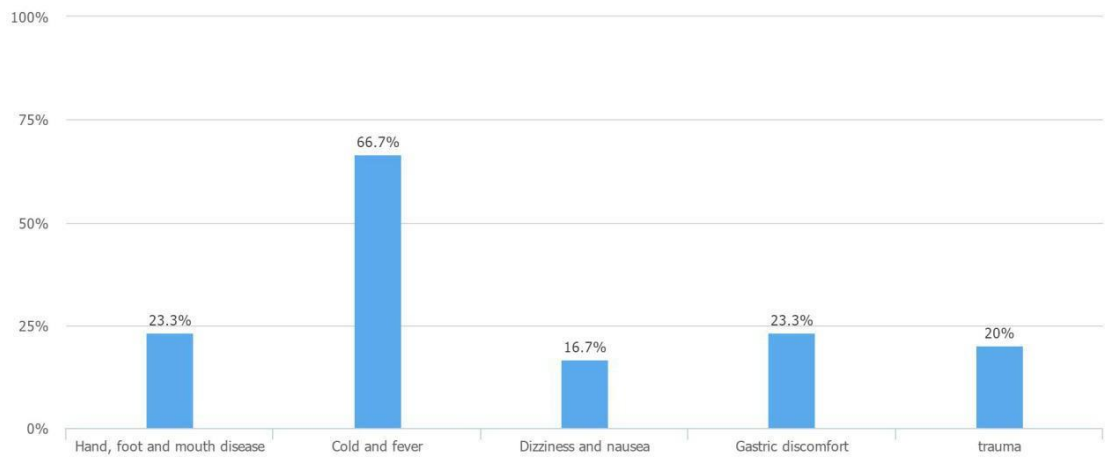

Figure 4.2

(3) How long does it usually take you to queue for the hospital registration? [Multiple choice question]
1. fast
2. normal
3. very long time
Result:

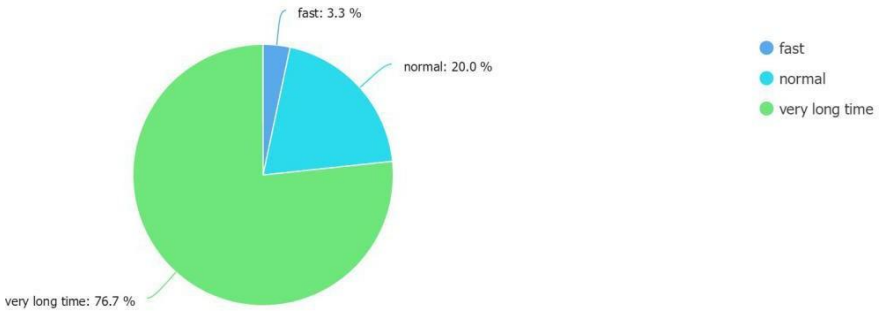

Figure 4.3

(4) Which sections can be improved in the hospital registration in your opinion? [Multiple choice questions]

1. Different hospitals have different electronic medical records
2. Doctor does not have enough diagnosis time
3. Queue time is too long
4. Checking procedures are too complicated

Result:

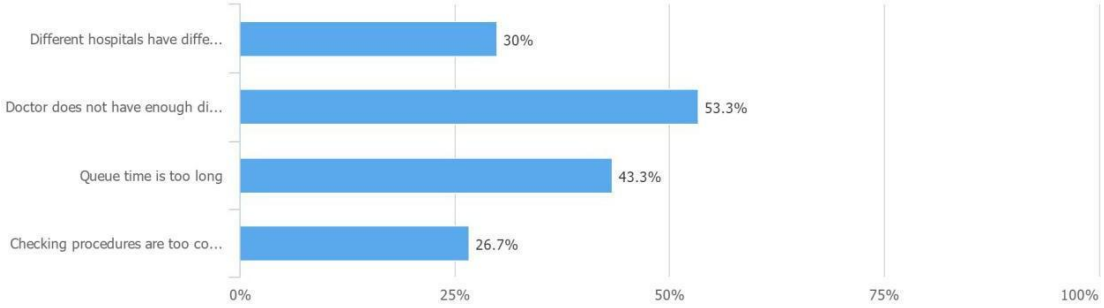

Figure 4.4

(5) Would you like to use the electronic medical record to improve medical efficiency before seeing a doctor? [Multiple choice question]

1. Yes
2. No

Result:

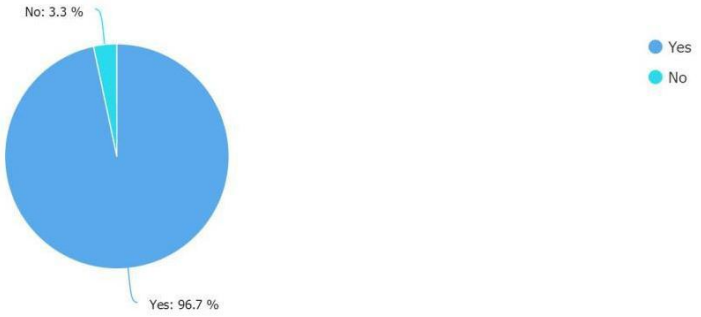

Figure 4.5

4.2 Interview

(1) Under what circumstances do you usually go to the hospital to register for medical treatment?
Interviewee (patients) answer:
No. 1 Patient : I usually go to the hospital if I feel unwell and have taken medicine for a day or two without obvious improvement.
No. 2 Patient : I generally have a fever or get cold then I would go to see a doctor.
No. 3 Patient : I usually go to the hospital to see a doctor when I feel sick.

(2) Do you meet some unreasonable things in the hospital registration process? E.g?
Interviewee (patients) answer:
No .1Patient: I usually have to queue for a long time for the hospital registration, especially when I am not feeling well, I can not stand this circumstance.
No .2Patient: If I want to see the doctor in another hospital, I always need to re-register the medical registration and pay the bill again. I think this process not only cost money but also waste my time.

(3) What is the most convenient and quick way for hospital registration in your opinion? why?
Interviewee (patients) answer:
No .1 patient: In my opinion, I think we should make full use of the online medical pre-registration system to reduce the long-term queue in real life.
Interviewee (physician) answer:
No. 1 physician: I personally think that if most patients have already pre-registered online or established a personal medical record, then my work will be more efficient and there will be more time to see more patients.

(4) What new changes do you hope to see in the hospital registration process in the future?
Interviewee (patients) answer:
No. 2 patient: Easy online pre-registration and the hospitals have more autonomous online registration machines to avoid a long time queuing. Different hospitals' EMR can be used universally.
Interviewee (physicians) answer:
No. 1 physician: Our official system can use electronic medical records from other hospitals, which can improve my work efficiency.
No. 2 physician: We can communicate with the patient's previous doctor through the patient's registration information, get more comprehensive information and improve the diagnosis rate.

(5) Do you accept the use of Internet applications to improve the efficiency of hospital registration?
Interviewee (patients) answer:
No. 2 patient: If the official can avoid too complicated operations, I would like to use the APP.
No. 2 patient: I am very happy to use this kind of app, it can save me time.
No. 2 patient: I think if this app doesn't leak my personal privacy, I would use it.
Interviewee (physician) answer:
No. 1 physician: If that APP has a comprehensive medical database and can help me improve my work efficiency, I will consider using it.

4.3 Use Case Diagram

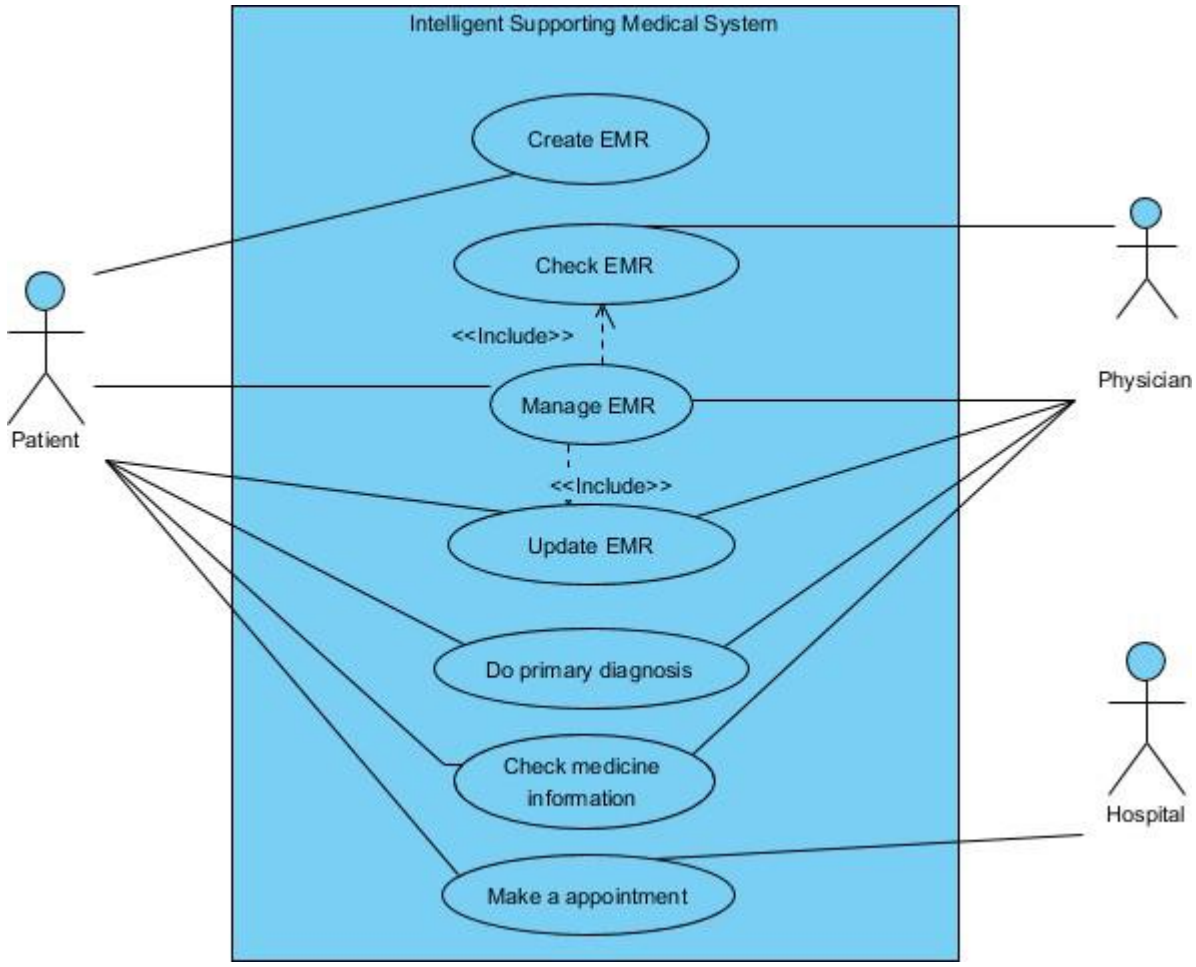

Figure 4.6 Use Case Diagram

This diagram shows use cases inside our system scope. Among these use cases, we selected checking EMR, updating EMR, doing primary diagnosis and making appointments as our target requirements to fulfil. We took checking medicine information as a supportive function to provide better user experience for patients. Specific description is showed in the next chapter.

5. System Design Description

Our system is a web-based application based on 3-tier data architecture. 2-tier data architecture has advantage in speed, the server may have data integrity issue because it cannot respond to multiple request at the same time. In our case, the number of end user is tremendous. It is impossible to use single server to connect every client in one place.

3-tier data architecture is one layer more than 2-tier architecture. It separates the presentation layer into one layer on end device and one layer on the server, and database server as a single layer. The application layer between the end device and database is to responsible for sending and responding request. It also processes functional logic, constraint and rules before passing data to each side. The advantage of this data architecture is improving program-data independence and supporting multiple

views of the data.

This flowchart shows a overview of how each component in our system work together to present functionalities.

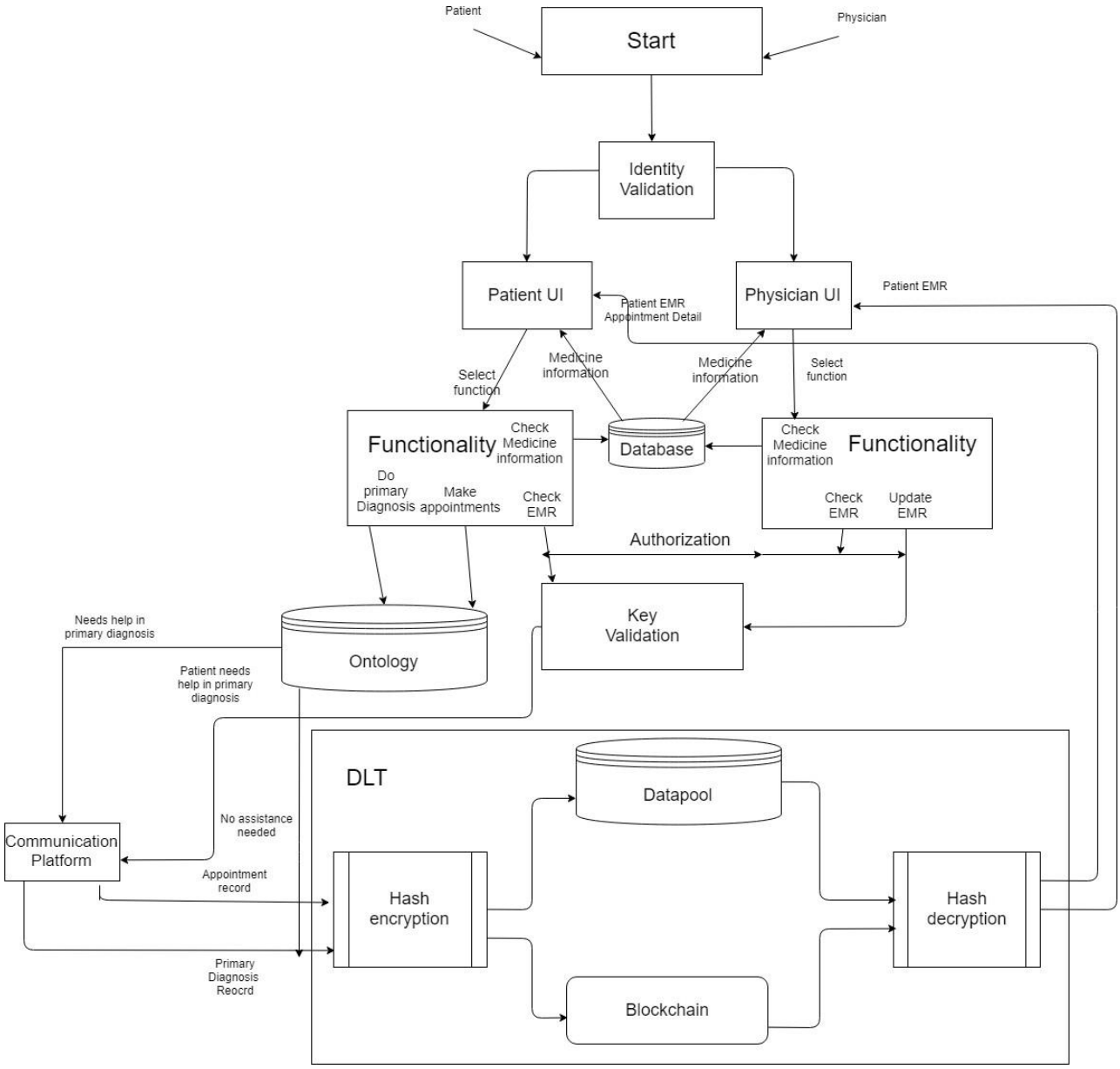

Figure 5.1 System Flowchart

6. Solution

There are three databases in the system. Relational database is used for store medicine data. Ontology is used for store disease related data for primary diagnosis, and the datapool is the database with the highest security used for store EMR.

6.1 Ontology

The current life we are living in is a life of information-driven science. Well organized information will help people to solve problems. With the human beings' higher demands for medical services, medical information system related various analysis and data services which based on medical information technology and network technology. Also, with the keep growing up population, countries like China, and cities like Shanghai are facing to a big problem that do not talk about directly see the doctor, even queues are difficult. Patients do not know which department they should go without queue for the outpatient. Or, they do not know what is the problem with their body without a primary diagnosis. Therefore, our system stores symptom, department and hospital information data.

Ontology is like a relevance tree which use for recording the hierarchies and relationships of the system. Along with the digital transformation of hospitals, and increasing of attention of health and social care, the demand of current medical care services is keep increasing. In addition, shareable clinic information between different hospitals are meaningful to medical care services and medical research. Ontology basically achieves the pattern matching and the knowledge sharing based on a unified semantic level.

Ontology is one of the data model, and it is a formal description of knowledge as a set of concept within a domain and the relationships that hold between them. In the SMedicalS project, we have measured and considered ontology is the most appropriate data model can support patient. The general public is not all have the same education background, some people are still not able to use the Internet and relative technologies, and they either do not know the medical language. In general, ontology in the SMedicalS stores basic knowledge of symptoms, and information of top third-class hospitals in Shanghai. In order to store these data, ontology can provide a better data management, since ontology ensure a common understanding of information and that they make explicit domain assumptions. Therefore, patients can start from enter a simple description of symptoms to find out the suitable hospital to book an appointment, and do the primary outpatient.

On the other hand, another characteristic of ontology is it based information linking approaches, also known as a keyword search (Munir & Anjum 2017). All data in the SMedicalS ontology are interact between each other, and it uses a keyword search mechanism for ontology navigation from the first level to the lower level and keep going until user find the result. So, in other words, the ontology function is like a complete 'brain' with all knowledge and information. Moreover, this brain can be extent if needed.

Ontology is a complex concept with various definition, and it has been used in many areas. Ontology refers to “a representational artifact, comprising a taxonomy as proper part, whose representation are intended to designate some combination of universals, defined classes, and certain relations between them” (Arp, Smith, Spear & Smith, p.1). To simplified, an ontology model introduces vocabulary relevant to domain which includes names for classes and relationships. For example, figure 6.1 shows a hierarchical structure includes 4 levels from a wide range to a detail. We use ontology is because it based on distributed ledger chain network, and combined with distributed multi-dimensional entity authentication system which includes distributed multisource exchange protocol, distributed data collaboration, distributed process, and distributed storage.

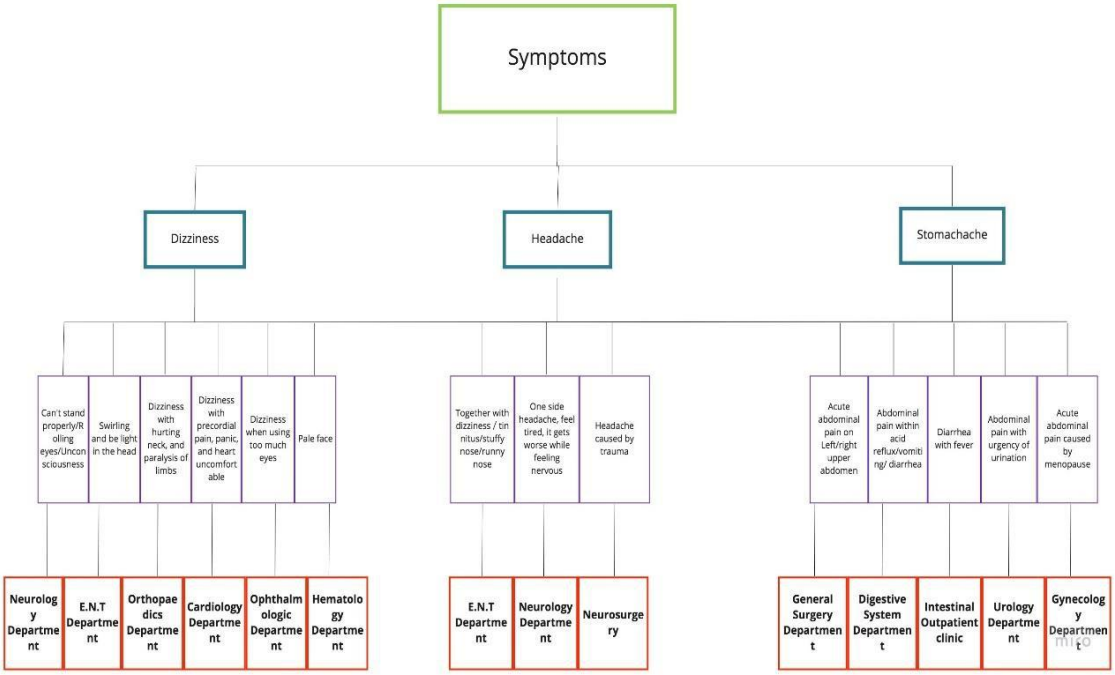

Figure 6.1 Symptoms part

The first level is a generic term, and it relates to different classifications. Figure 6.1 shows only three symptoms as different classifications. Level 3 is a more detail description of the previous level. For instance, if a patient says she/he has a headache. Then we should move to next question to ask the patient, could she/he tells more, or describe more about the headache. After that, the system will lead the patient to the correct department.

In the medical industry, hospitals have their own procedures for medical care, and the information systems of hospitals have been developed and will keep developing to replace the traditional operation process. However, these procedures are not related to each other. Thus, our system uses ontology to relate the third-class hospitals in Shanghai (Figure 6.2) in order to provide a more convenient medical care system, and help patients are able to find the nearest and the most suitable hospital for them. As a result of huge data, Figure 6.2 shows some hospitals of the third-class hospitals in Shanghai (the system stores all hospitals and related data). Also, according to the lower level, it is shows that hospitals’ addresses and telephone numbers are located.

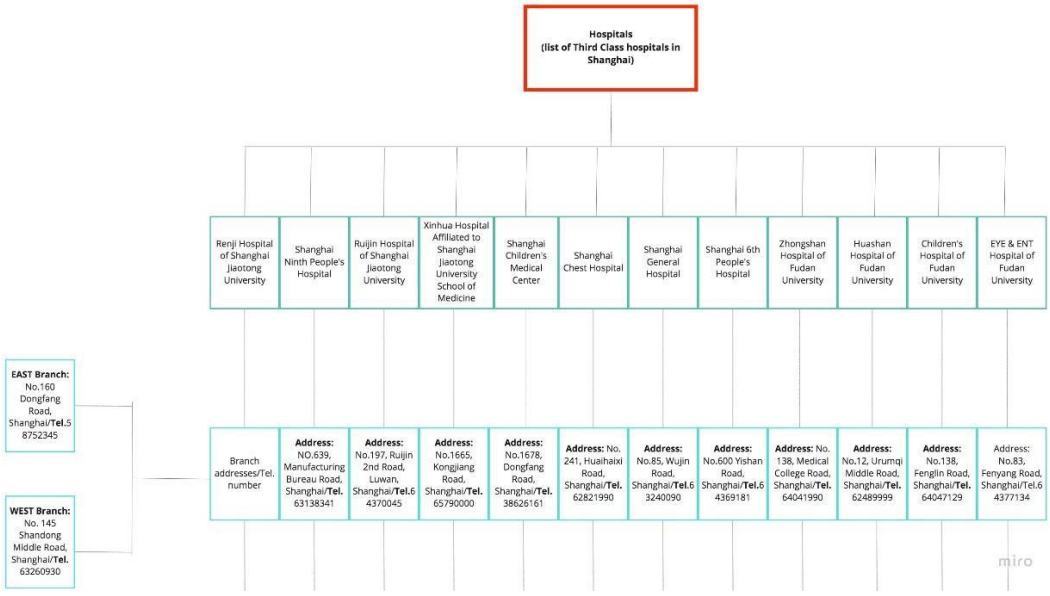

Figure 6.2: Third-Class hospitals in Shanghai

Figure 6.1 and Figure 6.2 are not two separate ontology, and they actually are linked and part of each other (shows in figure 6.3). Rather than saying that ontology is a user interface, ontology will not be shown on the screen. The ontology can be seen as a backend of the system.

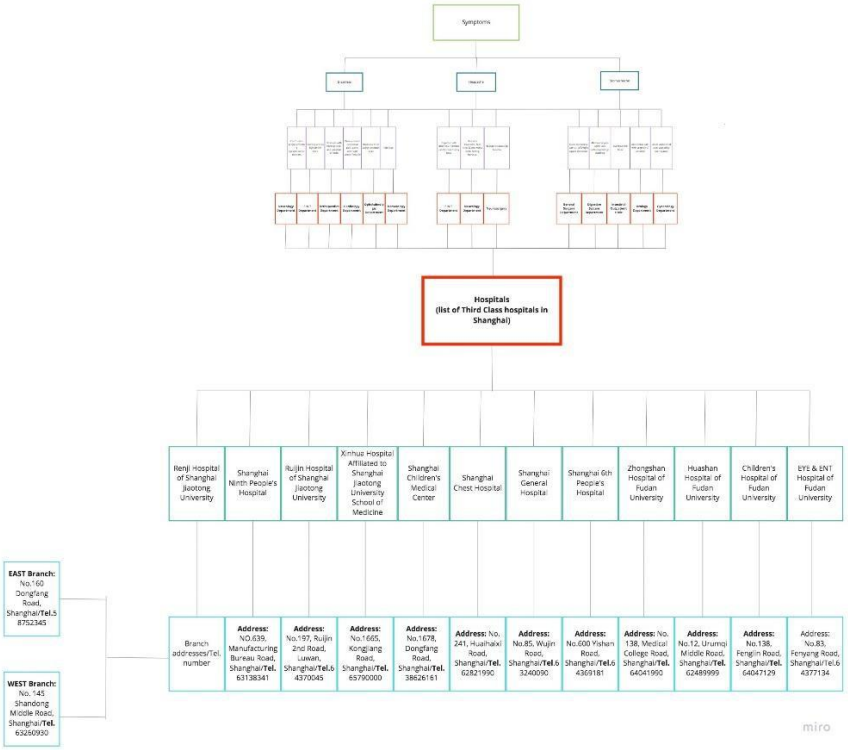

Figure 6.3 Service ontology

The service ontology (figure 6.3) make all information integrated operate, and bind all the information to the ontology. It enables the platform can grasp the concept or other data from the corresponding ontology when parsing the request, in order to efficiently support semantic self-description as well as the interoperability of requests. For patients, they are able to click the options keys on the user interface

to go through the whole process step by step based on this backend.

6.2 Adpotion of DLT

A distributed ledger is a technology that enabling users share information across multiple sites, geographical locations, and different organizations. Each participant in the network can obtain a true and unique copy of the information. Any changes in the ledger will be reflected in all copies.

6.2.1 DLT Performance

A distributed ledger is a technology that can share information across multiple sites, geographical locations, and different organizations. Each participant in the network can obtain a true and unique copy of the information. Any changes in the ledger will be reflected in all copies.
In this ledger, security and accuracy of the stored data are controlled by the use of public and private keys and signatures to control the access rights of the books, thereby achieving encryption protection. According to the consensus rules already reached in the network, the records in the ledger can be updated by one, some or all participants.

Its characteristics are as follows:
Infinite increase: Each page of the ledger is equivalent to a block in the blockchain. The block contains accounting information. If you compare to the blog, you would see the article which is stored in a block, and the entire blog forms a blockchain.

The encryption is smooth: the blocks filled with information are encrypted, and time-stamped to ensure that the blocks are linked in chronological order to form the final general ledger, just like every article in my article has a release date.

Decentralization: If users think that the server-side has changed their personal information without authorization, this is indeed possible on the Internet, but it can't be done in the world of blockchain, because the server with data is not only in the supplier’s hands, but in the hands of each of our users. Each person is a node, which is jointly maintained by us, so it is called decentralized distributed ledger. This distributed ledger is the latest and most desirable form of current accounting.

The advantages of its data are as follows:
Data Permissions: This permission not only specifies the source of the data but also specifies the ownership of the data (accuracy, changes, lifecycle management, etc.) and the location of the final authoritative version of the data.
Data accuracy: Accuracy is a key feature of the data, meaning that the data value records of any object are correct, can represent the correct value, form and content are consistent with the description object.
Data Access Control: The blockchain solution can track public and private information separately, including detailed information about the data itself, transactions corresponding to the data, and those with data update rights.

6.2.2 DLT Functions

1) Decentralization

Distributed is more stable, more risk-resistant, less expensive, fairer, more transparent, and simpler than centralized.

It is this seemingly simple distributed thinking that has enabled us to develop rapidly in technologies such as artificial intelligence, big data, cloud computing, Internet of Things, blockchain, etc., especially for the Internet of Things, distributed. The idea is extremely important.

As the performance of smart devices (single nodes) around us becomes more powerful (performance such as increased computing power and storage capacity), distributed network performance linking nodes is becoming more powerful. The existing centralized networks will be surpassed by distributed networks in terms of performance, maintenance costs, number of users, and so on. For example, the current mainstream technologies GFS and HDFS used in cloud storage, in the case of increased demand, generally take measures to support their services through continuous hardware investment. On the other hand, every user, every box, is the data acquisition party and the data provider. It is both the operator and the user and constitutes a huge point-to-point storage network.

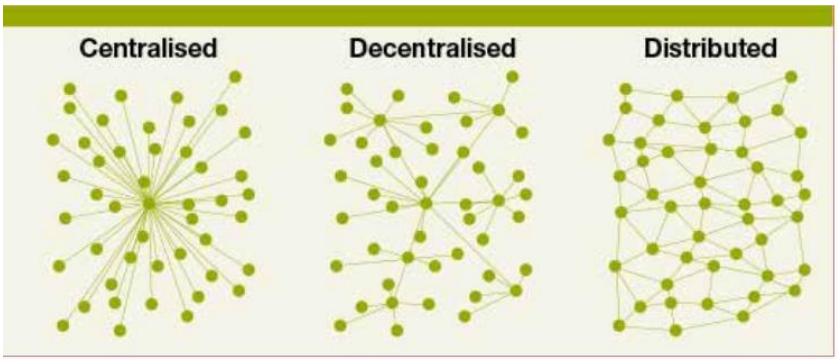

Figure 6.4 Three different kinds of mode

In the above figure,distributed networks can avoid the time-consuming reconciliation processes used by centralized and decentralized databases.

However, after decentralization, we may encounter the problem similar to bitcoin double-spending, that is, the problem of data being tampered with, as shown below:

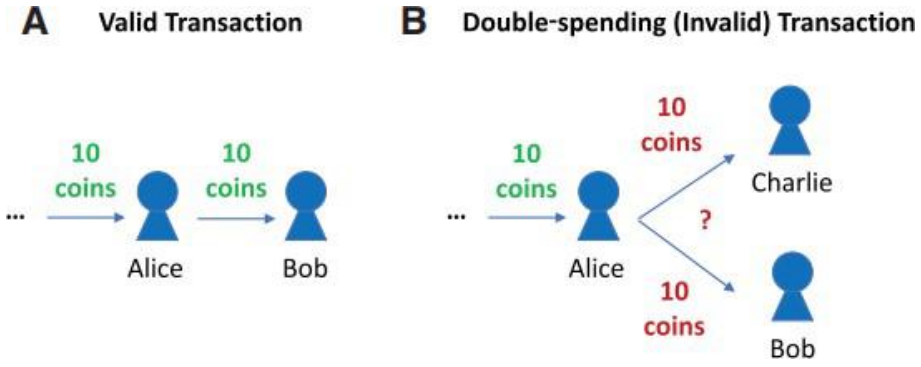

Figure 6.5 Three different kinds of mode

In this figure, A is a valid transaction, and B is an invalid Double-spending transaction. The problem in the figure is that in A we have already given Bob 10 coins to A, but how can Bob know if Alice is using it in B? Because without an official transfer of the bank (that is the center), the same coin was given to Charlie before giving him 10 coins?
This question involves the user's ability to modify the data in a decentralized environment, but in the DLT's decentralization technology and the solution.

For example, mathematical common sense tells us that "1+1=2" compares "1+1=2" to "accounting content," and that everyone in all humans is a "distributed ledger." Under the known mathematical logic, if someone wants to refer to the deer as saying "1+1=3" for the horse, then it is necessary to change the logical consensus of all human beings, but it is quite difficult.

Comparing the World Cup live broadcast of the World Cup to "distributed billing", the game process is "bookkeeping content", and each viewer watching the live broadcast is a "book." The audience is spread all over the world, so the whole audience is combined into a “distributed ledger”. If someone wants to tamper with the results of the game, they have to ask the audience to answer.
The following figure shows the working process of the DLT's decentralization:

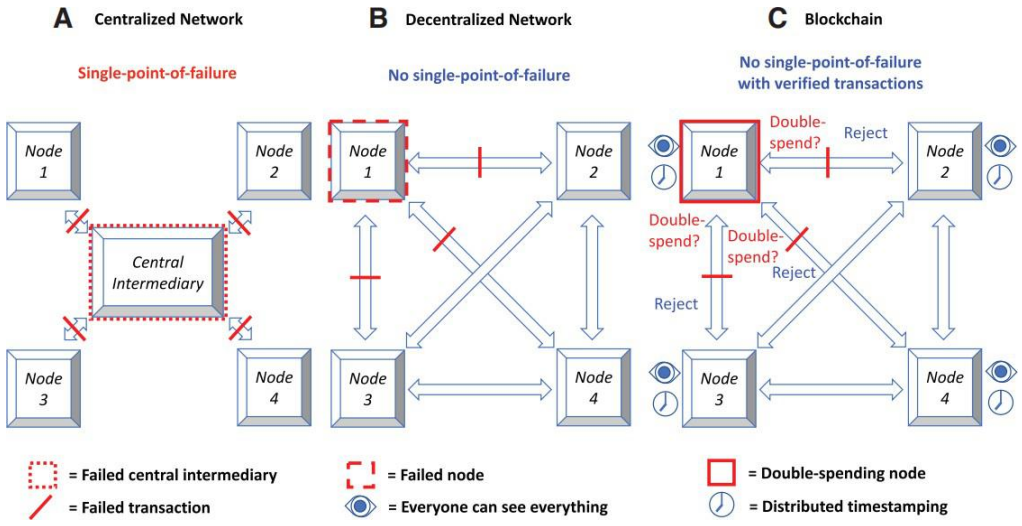

Figure 6.6 Three different kinds of mode

Compared to distributed network topology, A is the topology of a centralized network, which creates a single point of failure (central intermediary). If the central intermediary is down or attacked, the entire network stops working. B is a distributed network topology that does not contain a single point of failure. If Node 1 encounters a shutdown or attack, the rest of the network will still function. C is the topology of the blockchain. If "everyone can see everything" and there is a distributed timestamp mechanism, the problem of repeated consumption can be solved on a distributed network.

2) Hash encryption

Hash algorithms are often designed to generate text of the same length, and the length of the text generated by the encryption algorithm is related to the length of the plain text body.

For example, let's have two paragraphs: "Microsoft" and "Google."

The results obtained by using some hash algorithm are: "140864078AECA1C7C35B4BEB33C53C34" and "8B36E9207C24C76E6719268E49201D94", and the results obtained by using some encryption algorithm are "Njdsptpgu" and "Hpphmf" respectively. As you can see, the hashed results have the same length, and the encrypted results are of different lengths. In fact, if you use the same hash algorithm, no matter how long your input is, the resulting result length is a constant, and the encryption algorithm is often proportional to the length of the plaintext.

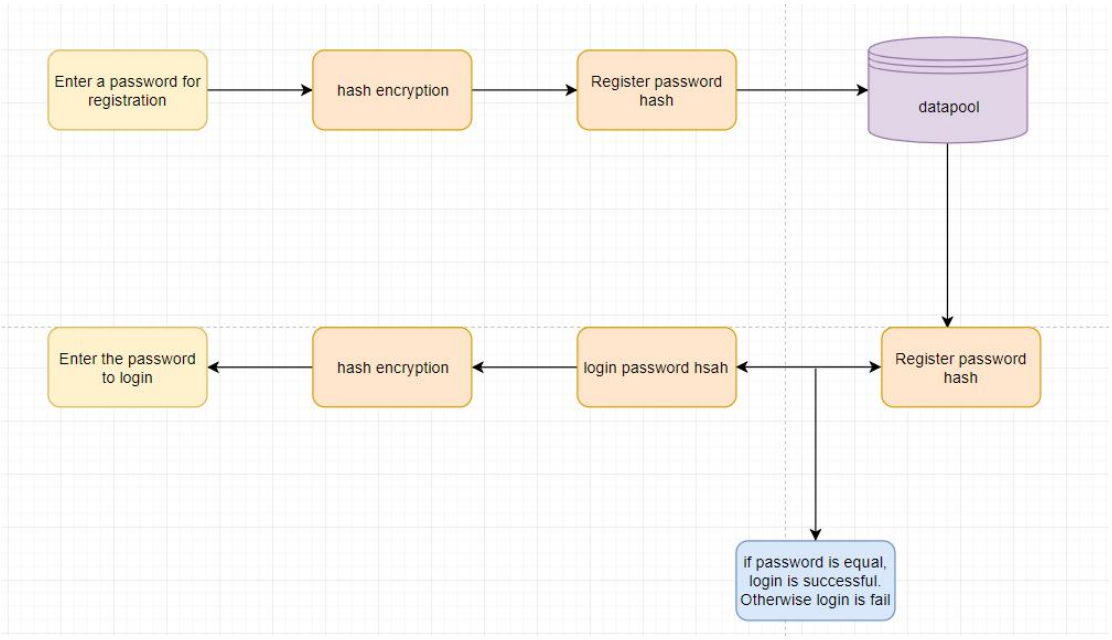

Figure 6.7 The hash algorithm work process

The hash algorithm is irreversible, and the encryption algorithm is reversible.

The irreversibility here has two meanings. One is "given a hash result R, there is no way to convert E to the original target text S", and the second is "given the hash result R, even if you know the hash result of a piece of text S. For R, it cannot be asserted that the original target text is S". In fact, hashing is impossible to reversible, because if it is reversible, then hashing is the most powerful compression method in the world - it can compress files of any size into a fixed size.

The encryption is different. Given the encrypted ciphertext R, there is a way to convert the R determination into the plaintext S before encryption.

Here we briefly introduce the difference between the two from the intuitive level. After the rigorous description of the two from the mathematical point of view, readers would know why there are these two differences.

The hash algorithm is a many-to-one mapping. Given a target text S, H can uniquely map it to R, and for all S, R has the same length. Since there is a many-to-one mapping, there is no inverse mapping for H.

The hash algorithm is extremely sensitive to changes in the input. Even if the input has very small changes, such as 100 million characters changed to one character, the result should be completely different, as shown below:

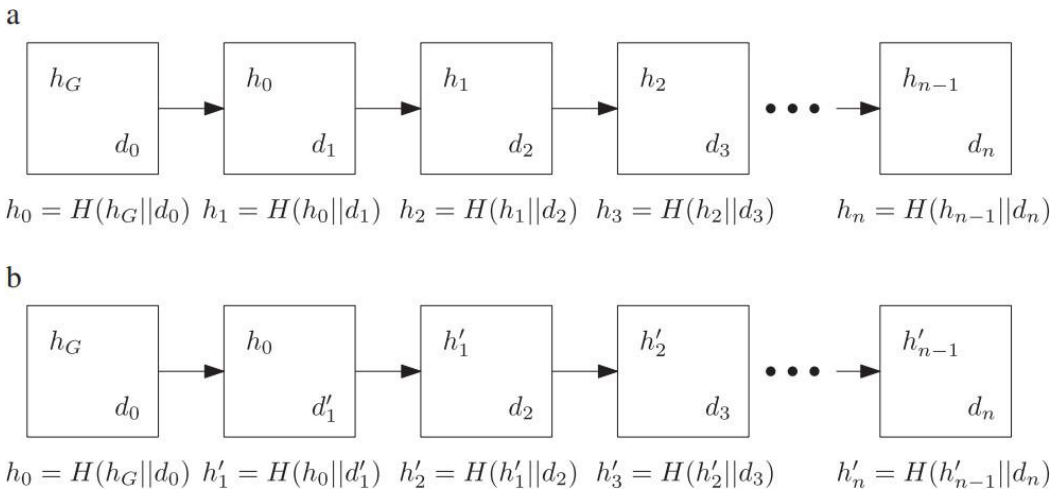

Figure 6.8

In the above figure, the data block in a row is connected by a one-way hash function (h indicates data, and d1 in row b indicates data blocks that have been changed). We can see that row b is the same as row a. The hash algorithm, but since the second data block has changed in row b, the subsequent data blocks have changed.

6.2.3 How does DLT work in our system?

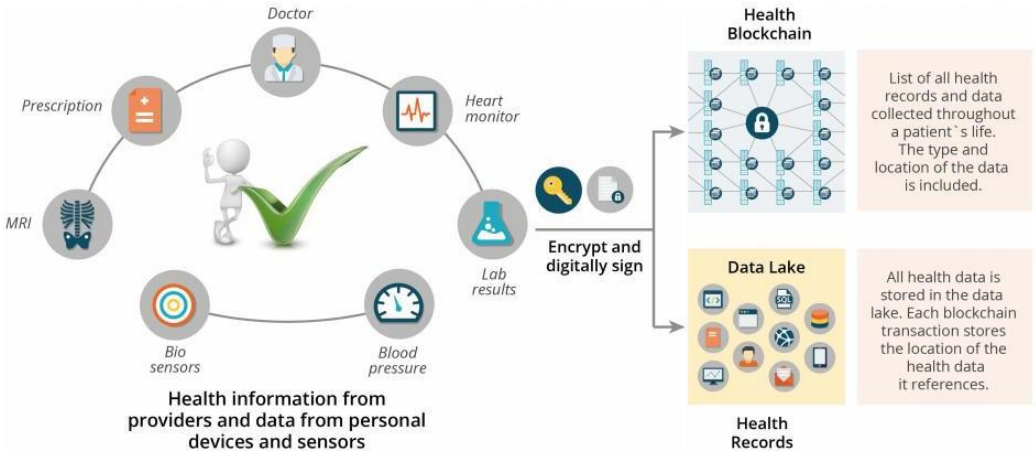

Figure 6.9 An example of DLT working with health record

In the system we designed, all collected medical data will be encrypted and stored in DLT's data pool, which is the data warehouse. This data pool is characterized by high scalability and can store data in different formats, such as image files and text files. In addition, the data pool contains a large amount of medical data, so it can be used for medical data analysis such as interactive query, text analysis and mechanical learning, it can also help the development of various medical research, such as determining genetic markers. All encrypted information in the data pool needs digital signatures, which can ensure the confidentiality and authenticity of the data provided.

When the system starts to work, if users create EMR in the system and upload their medical information, such as lab results, MRI, Heart monitor and other data, the system will require users to create their digital signatures to verify the authenticity of documents or images uploaded by users. Every time they upload medical data, it will be encrypted and sent to the data pool for storage, then it will be saved to the data pool already registered in DLT, and DLT will only work through user-specific authentication. After that, the user will get the system reverse, and can see that the uploaded health data has been added to personal DLT. Users can also use mobile terminals to digitally sign their medical data and encrypt the uploaded data, for example, use smart phones or smart bracelets to open APP and then import the health data collected by smart terminals into the system and perform encryption and digital signature. The user will have full rights to access and control how his medical data is shared in the system. At the same time, the user can give access rights to some users (doctors or other users). The which means user can specify who can query the data and modify his data. However, in order to avoid some malicious data modification, users can check through the mobile dashboard application which users have the right to access his DLT as well as to grant and cancel access rights to other users. In addition, users can also check the audit log to see who has accessed his DLT and what data has been viewed.

User access control rights will not be limited to 'all-or-nothing', but will adopt a flexible way. users can set only specific types of data to be accessed according to their own requirements, specify access time and who can access them, and users can modify these rights at any time. These access control texts created by users will also be stored in DLT, thus ensuring that only users with proprietary authentication can modify these settings. In this way, privacy and security are improved, and users can independently set which data can be shared and accessed.

When user A obtains the access rights of other user B, A can query the medical data of user B through DLT and compare the authenticity of the data through digital signature. Therefore, the designed system can be classified and analyzed by collecting health data shared by users. At the same time, the identity authentication function of the system will adopt the authentication system method of financial institutions or regulatory agencies, because this method not only provides passwords but also provides intelligent tokens, so the security of the identity identification system will be greatly strengthened.

In this system, the user has only unilateral control over his data and the rights obtained from the system, which means that the user cannot directly obtain specific medical data from the system or other medical databases inside the system for private use. In DLT's de-centralized environment and digital signature can ensure that some malicious users can impersonate or destroy, because malicious users will display in the system if they forge digital signatures or control most of the network resources. Similarly, malicious users cannot obtain any encrypted data from public books.

6.2.4 Significance of applying DLT

The extensibility and encryption of DLT can solve the interoperability problem of the current medical system and enable medical institutions and personnel to share electronic data safely. Medical personnel can use DLT-based medical IT system to exchange and integrate data between API and DLT. In this way, the operation of the system will be simpler and the problem of data integration between different systems will be eliminated.

DLT will be used based on reliable commercial hardware, because such hardware can provide effective calculation and thus reduce costs. This method will improve the benefits of some medical research, such as genome research. And DLT's hardware capability can also be analyzed with medical researchers to speed up the research process of new drugs and treatment methods.

DLT can also centralize users, doctors and medical institutions on a data sharing platform, so that doctors can timely obtain accurate and comprehensive medical data of patients. DLT's data structure and its data pool support the import of various medical data sources, such as users' mobile terminals, EMR, documents and graphics, etc. Therefore, the applicability of DLT will be greatly reflected in the medical system.

Decentralized Management	Patient-managed health care records: [Patient] becomes the platform, owning and controlling access totheir healthcare data. This removes all obstacles to patients acquiring copies of their healthcare records or transferring them to another healthcare provider.
Immutable Audit Trail	Unalterable patient records: The data are stored in the private blockchain cloud. Blockchain may guar-antee medical data cannot be changed by anybody including physicians and patients himself/herself internally and natively
Data Provenance	Source-verifiable medical records: Records are signed by source, allows legitimacy of records to be veri-fied (and false records to be plausibly denied).
Robustness/Availabilit	Reduced risk of patient recordkeeping: Because data is stored on a decentralized network, there is nosingle institution that can be robbed or hacked to obtain a large number of patient records.
Security/Privacy	Increased safety of medical records: Data is encrypted in the blockchain and can only be decrypted with the patient's private key. Even if the network is infiltrated by a malicious party, there is no prac- tical way to read patient data.

Figure 6.10 Improve EMR management

6.3 Relational Database

6.3.1 Medicine information checking

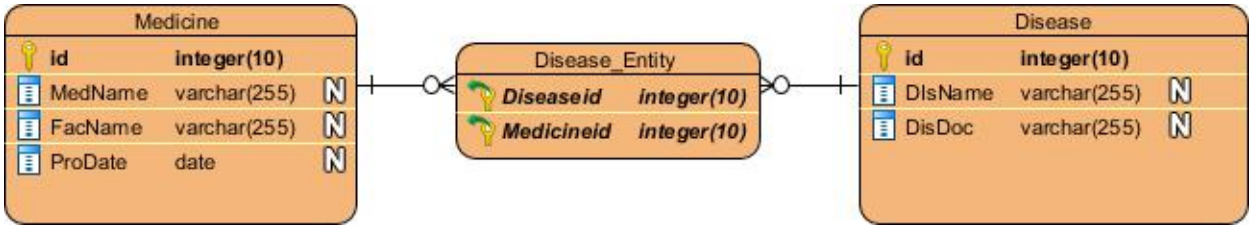

Figure 6.11 Entity Relationship Diagram

This relational database is designed to enable patient and physician to check medicine information. It is usually happened that patient is not familiar with the medicine in the doctor's prescription. This database contains key information such as medicine name, provider name and medicine description. It can help patients know more about the medicine they eat, and avoid deadly accident like consuming Cephalosporin and Alcoho at the same time.

6.3.2 Suggested implementation

In this design, we hope to extract the different types of medicine names in the Chinese clinical trial database, and related symptoms and hospital units to import our system. In the database. In this session we only need to extract information simply without data analysis, so we choose Python crawler technology as a method of data mining.

Python's crawler technology is a technique for crawling information on the web. The following figure shows the basic flow of crawling:

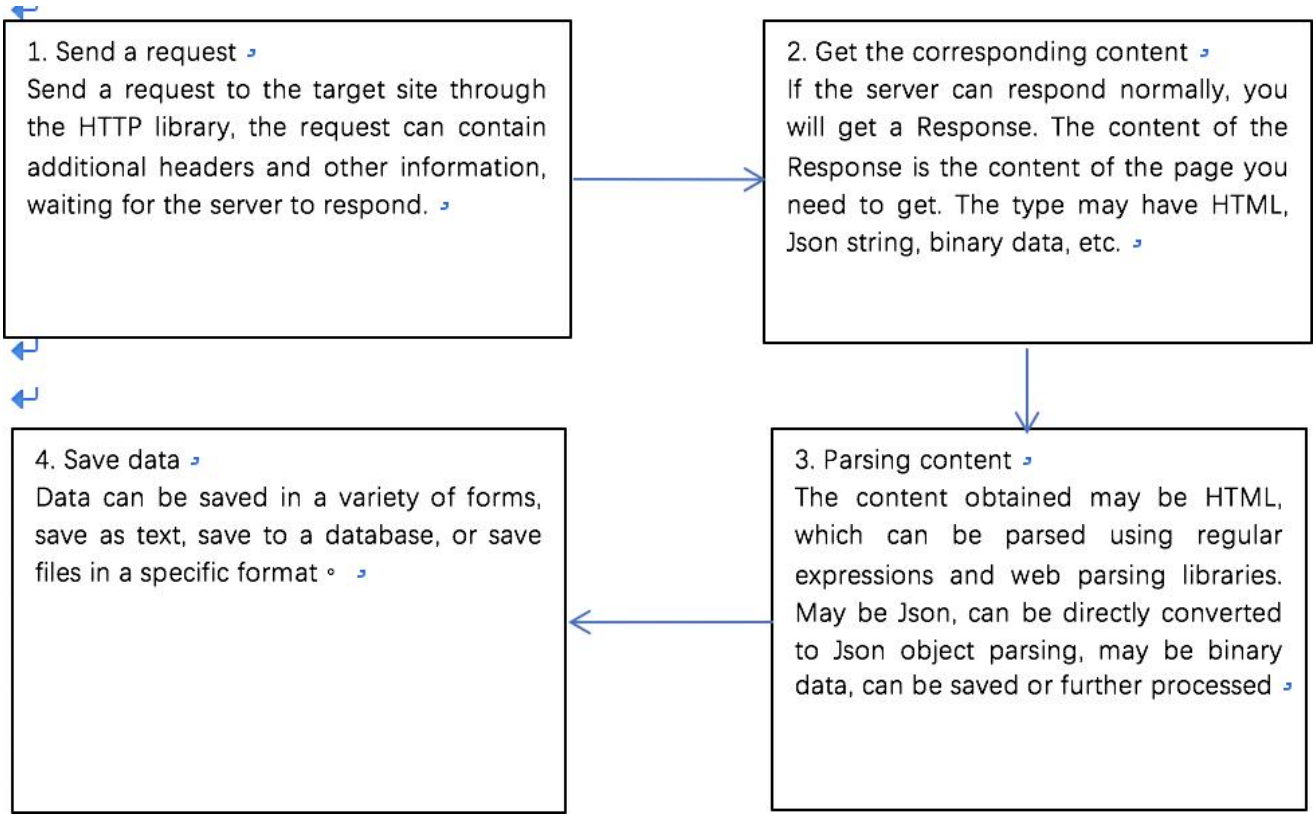

Figure 6.12

The Python Data Analysis Library or panda was created to solve data analysis tasks. Pandas incorporates a large number of libraries and some standard data models to provide the tools needed to efficiently manipulate large data sets. Pandas provides a number of functions and methods that enable us to process data quickly and easily.

Use the read_html method in the pandas library to quickly crawl common tabular data in web pages. We will see some forms on the China Clinical Trial Database page, such as:

登记号	试验题目	药物名称	适应症	试验状态	试验分期	申办单位	试验机构	登记日期
CTR20150825	帕拉米韦氯化钠注射液治疗儿童流行性感冒的临床研究	查看	儿童流行性感冒	进行中（招募中）	4期	广州南新制药有限公司/广州南新制药有限公司/湖南有色凯铂生物药业有限公司/军事医学科学院毒物药物研究...	首都医科大学附属北京儿童医院	2016-02-16
CTR20150508	柴胡鼻腔喷雾剂对普通感冒引起发热有效性和安全性研究	查看	普通感冒引起的发热	进行中（尚未招募）	3期	南阳利欣药业有限公司/复旦大学/	首都医科大学附属北京中医医院	2015-07-28
CTR20150165	感冒康宁颗粒Ⅱ期临床试验	查看	清热解毒，祛邪透表，散风止痛。用于风热感冒，症见恶风发热、口干咽痛、鼻塞、流涕、头晕头痛、肢体酸痛、...	已完成	2期	大连珍奥药业有限公司/	首都医科大学附属北京中医医院	2015-06-17
CTR20140866	帕拉米韦注射液治疗重度甲、乙型流行性感冒的临床研究	查看	重度甲、乙型流行性感冒	进行中（招募中）	3期	天津泰普药品科技发展有限公司/	卫生部北京医院	2014-12-30
CTR20140865	帕拉米韦注射液治疗轻中度甲、乙型流行性感冒临床研究	查看	轻中度流行性感冒	进行中（招募中）	3期	天津泰普药品科技发展有限公司/	卫生部北京医院	2014-12-30
CTR20140345	研究病毒无忧软胶囊治疗普通感冒的安全性和有效性研究	查看	风热感冒	进行中（招募中）	3期	亚宝药业集团股份有限公司/	天津中医药大学第一附属医院	2014-05-13
CTR20140249	柴胡鼻腔喷雾剂对普通感冒引起发热有效性和安全性研究	查看	普通感冒（急性上呼吸道感染）引起的发热	进行中（尚未招募）	2期	南阳利欣药业有限公司/复旦大学/	首都医科大学附属北京中医医院	2014-03-20
CTR20140146	参果老年感冒颗粒Ⅱ期临床试验	查看	普通感冒（气虚感冒）	已完成	2期	陕西天方科技有限责任公司/	中国中医科学院西苑医院	2014-04-02
CTR20132887	探索评价感通颗粒治疗普通感冒（感冒风热证）	查看	普通感冒（感冒风热证）	已完成	2期	深圳市北科联药业科技有限公司/	四川大学华西医院	2014-05-14
CTR20132866	治疗小儿急性上呼吸道感染（感冒风热证）发热	查看	小儿急性上呼吸道感染（感冒风热证）发热	进行中（招募中）	2期	深圳市北科联药业科技有限公司/	江苏省中医院	2014-07-03

共 57 条 1 2 3 下一页 跳转到 页 跳转 20

Figure 6.13

The form web page structure of the table type is roughly as follows:

```
<table class="..." id="...">
    <thead>
    <tr>
    <th>...</th>
    </tr>
    </thead>
    <tbody>
        <tr>
            <td>...</td>
        </tr>
        <tr>...</tr>
        <tr>...</tr>
        <tr>...</tr>
        <tr>...</tr>
        ...
        <tr>...</tr>
        <tr>...</tr>
        <tr>...</tr>
        <tr>...</tr>
    </tbody>
</table>
```

Figure 6.14

Briefly explain the meaning of several labels that appear above:

```
    <table>       : Definition form
<thead>  : Define the header of the table
<tbody>  : Define the body of the form
    <tr> : Define the row of the table
<th>     : Define the header of the table
<td>     : Define table unit
```

Figure 6.15

Fast crawl

Using the table in the Chinese clinical trial database page above as an example to take a simple data capture process:

```
    import pandas as pd
import csv

for i in range(1,3):  # Crawl all 3 pages of data
        url='https://db.yaozh.com/linchuangshiyan?comprehensivesearchcontent=%E6%84%9F%E5%86%92& %(str(i))
        tb = pd.read_html(url)[2]
        tb.to_csv(r'1.csv', mode='a', encoding='utf_8_sig', header=1, index=0)
        print('No'+str(i)+' page has got')
```

Figure 6.16

Result sample:

适应症	申办单位	试验机构	登记日期
儿童流行性感冒	广州南新制药有限公司/广州南新制药有限公司/	首都医科大学附属北京儿童医院	2015/12/31
普通感冒引起的发热	南阳利欣药业有限公司/复旦大学/	首都医科大学附属北京中医医院	2016/2/16
清热解毒	大连珍奥药业有限公司/	首都医科大学附属北京中医医院	2015/7/28
重度甲乙型流行性感冒	天津泰普药品科技发展有限公司/	卫生部北京医院	2015/6/17
轻中度流行性感冒	天津泰普药品科技发展有限公司/	卫生部北京医院	2014/12/30
风热感冒	亚宝药业集团股份有限公司/	天津中医药大学第一附属医院	2014/12/30
普通感冒（急性上呼吸道感染）	南阳利欣药业有限公司/复旦大学/	首都医科大学附属北京中医医院	2014/5/13
普通感冒（气虚感冒）	陕西天方科技有限责任公司	中国中医科学院西苑医院	2014/3/20
普通感冒（感冒风热症）	深圳市北科联药业科技有限公司/	四川大学华西医院	2014/4/2
小儿急性上呼吸道感染发热	深圳市北科联药业科技有限公司/	江苏省中医院	2014/5/14

Figure 6.17

Store to MySQL

Next, we can save the results to a local csv file or to a MySQL database. First, we need to create a table for storing data in the database, named here named_medicine. Sample code would show as below:

```
    import pymysql

def generate_mysql():
        conn = pymysql.connect(
                host='localhost',   # Local server
                user='root',
                password='******',  # Your database password
                port=3306,          # Default port
                charset = 'utf8',
                db = 'medical')
        cursor = conn.cursor()

        sql = 'CREATE TABLE IF NOT EXISTS listed_medicine2 # listed_company
is a table to be created in the 'medical' database for storing data
        cursor.execute(sql)
        conn.close()

generate_mysql()
```

Figure 6.18

The above code defines the generate_mysql() function to generate a table of listed_company in the medical database in MySQL.Then we can write data to this table.

Sample code as below:

```
    import pymysql
from sqlalchemy import create_engine

def write_to_sql(tbl, db = 'medical'):
    engine =
create_engine('mysql+pymysql://root:******@localhost:3306/{0}?charset=utf8'.f
ormat(db))
    # db = 'medical' means stored in the medical database, the * behind the
root is the password
    try:
        tbl.to_sql('listed_medicine',con =
engine,if_exists='append',index=False)

        except Exception as e:
            print(e)
```

Figure 6.19

7. User Interface Design

7.1 Patient and Physician User Interface

First of all, the target audience of our system is the general public. Therefore, we decided to make the user interface convenient, extensible, user friendly, and as secure as possible. Seven user interfaces are listed below:

Figure 7.1: Homepage

Depends on user's role, the interface will be different. This is the homepage patient will see after he/she successfully login.

Figure 7.2: Selection page in Chinese

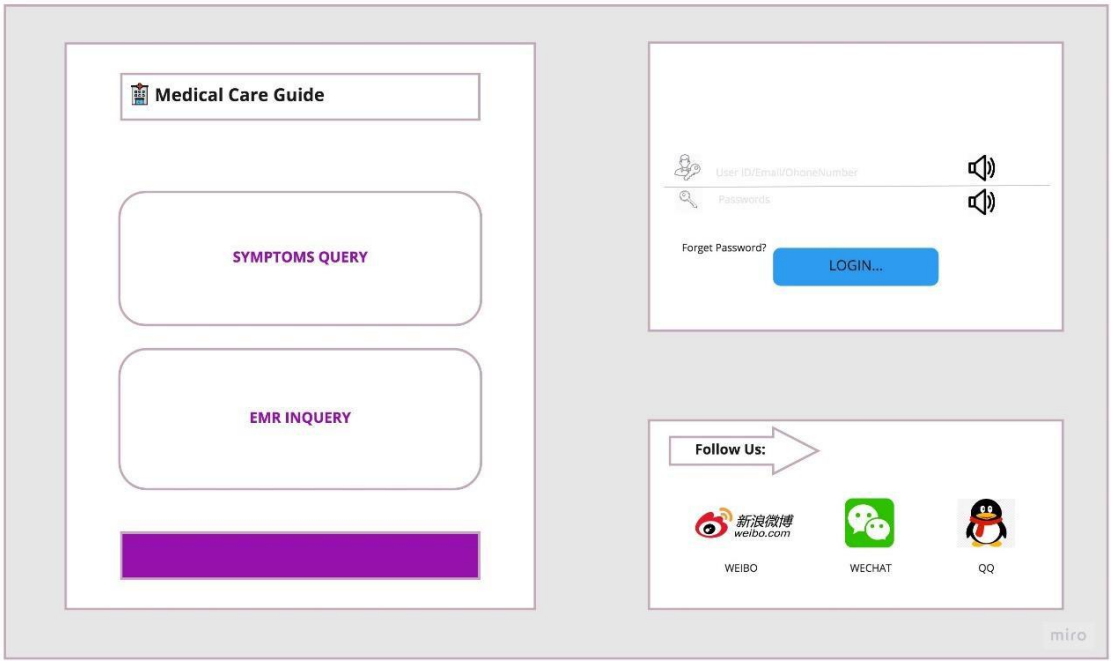

Figure 7.3: Selection page in English (in Web translation feature)

These user interfaces are not the final version of the whole design, but they show how the system going to run step by step. The top three famous social media platforms have been listed in Figure 7.2 and 7.3. The only difference between these two figures is Figure 7.2 is written in Chinese, and Figure 7.3 is written in English. Users can follow any of those social media platforms to get the latest announcements, news, and hospitals' information. Users also can leave comments and feedbacks to help us improve the system.

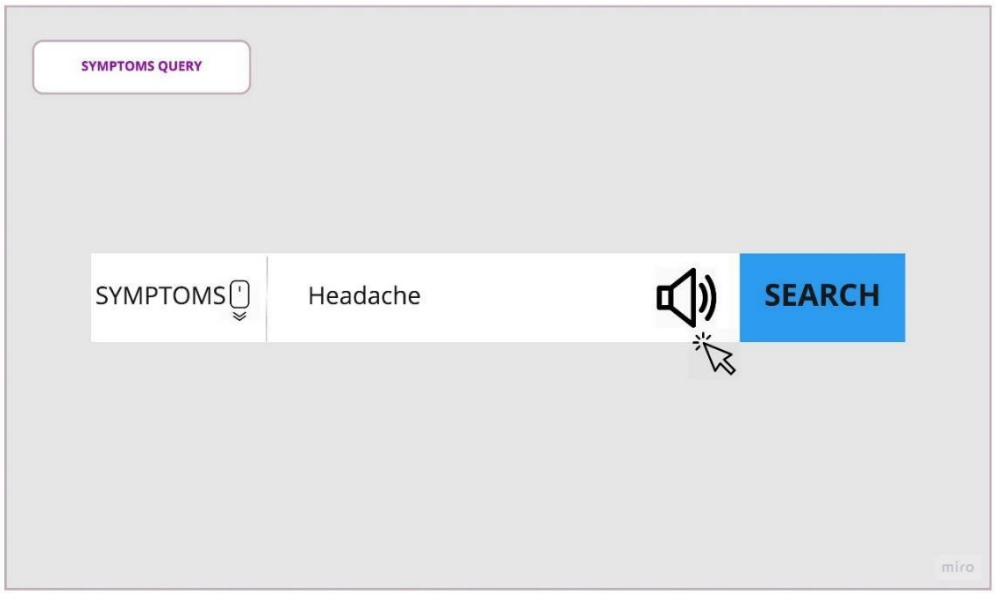

Figure 7.4: Checking Symptoms

By clicking 'symptoms query' in Figure 7.4, users will move to the next page which is showing in Figure 7.5. If users fine with typing the key words of their symptoms in the box, they will directly move to next step. If not, they can easily click the little trumpet icon to move to a speech recognition page (in Figure 7.6). Figure 7.6 simply shows a sample scenario when user needs that page. The only thing that users need for the speech recognition feature is a microphone to record their voice into the system, and start the conversation with our intelligent system. Same as Figure 7.6, voice dictation feature also requires users have a microphone. Then, the system will valid the user's voice to login to the system.

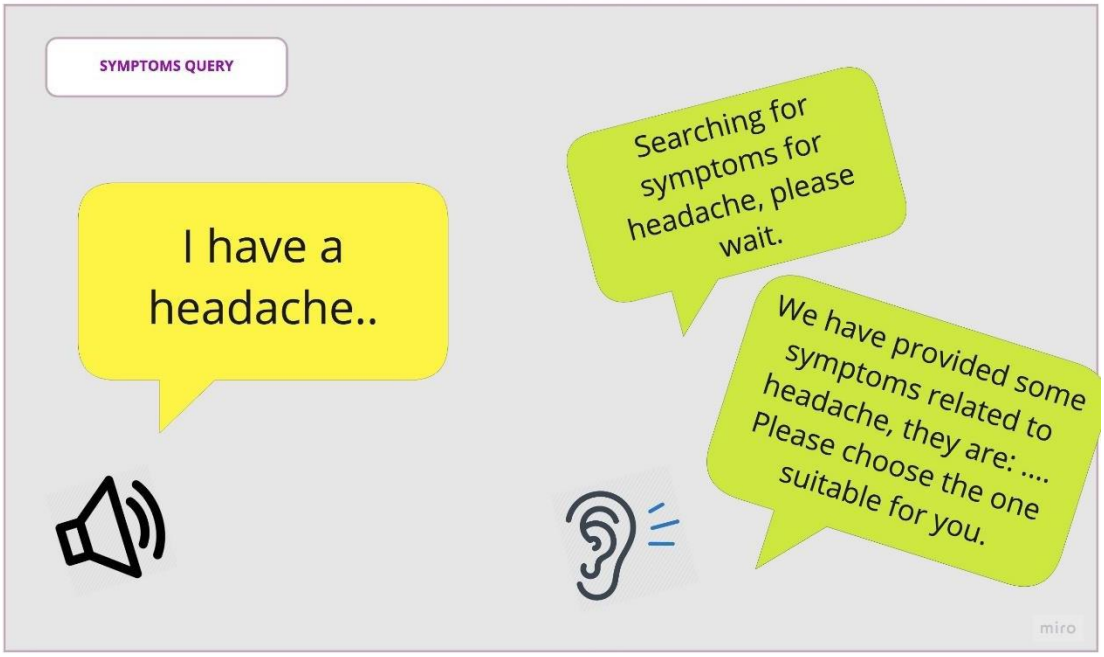

Figure 7.5: Speech Recognition

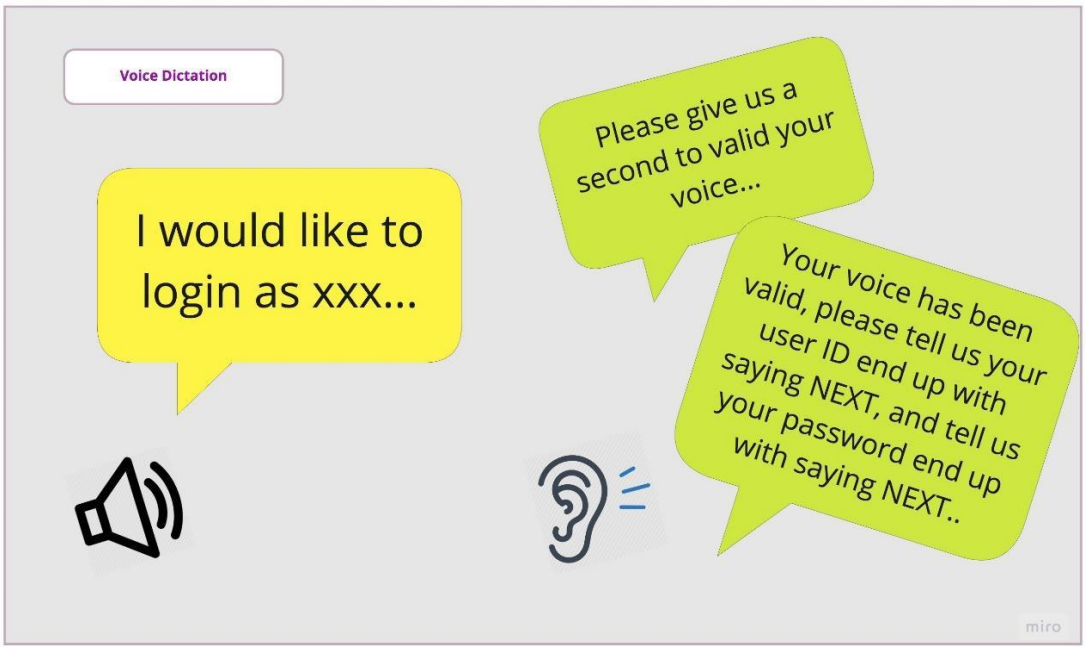

Figure 7.6: Voice Dictation

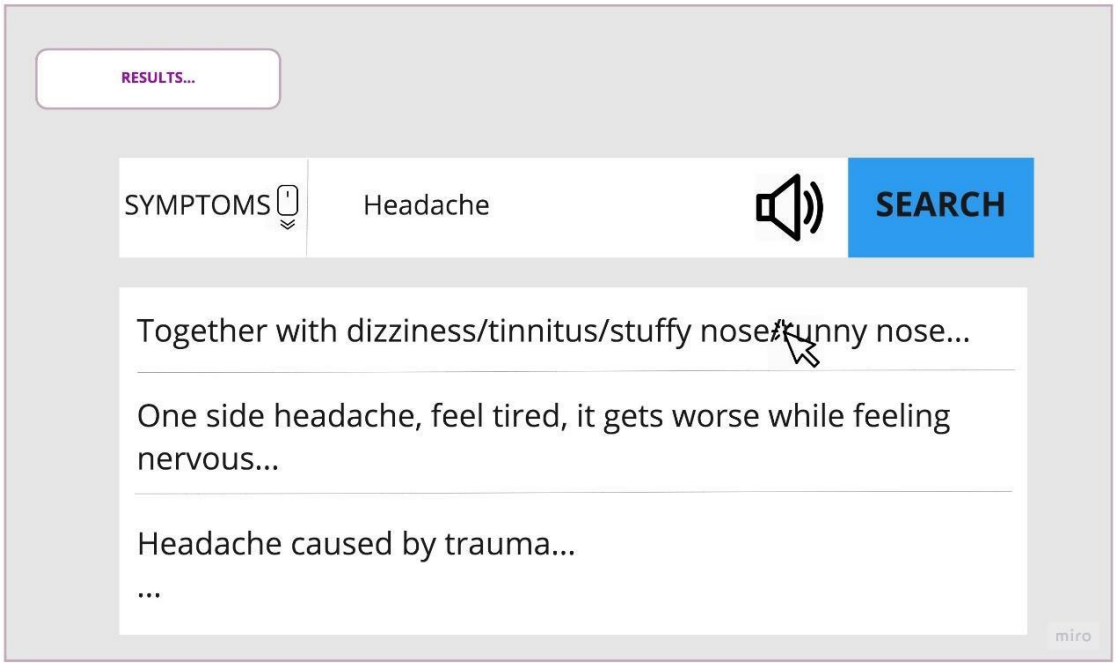

Figure 7.7: Searching Results

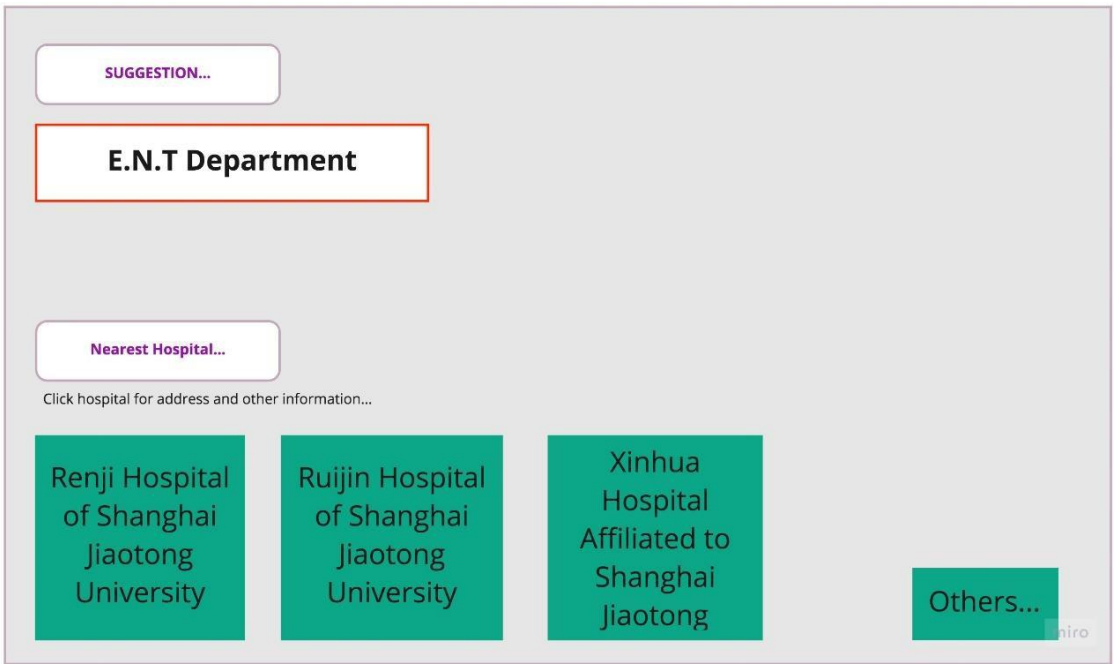

Figure 7.8: Hospital Suggestion

When a user is struggling with having headache, the user can enter 'headache' as a key word into the box shows in Figure 7.7, and click 'Search'. As a result, searching results will pop up after that (Figure 7.8). It is because these user interfaces are creating for people who are not familiar with medical proper nouns. The symptoms will keep as simple as possible. For example, a patient considers she has headache, and together with dizziness and runny nose. Then, she should choose the first symptoms which shows in Figure 7.7 to move to the next page. Figure 7.8 shows the suggested department. This

helps patient to go to the correct department without confusions. It is because hospitals in Shanghai are not small, and have many different departments at same building. Also, the system will provide the nearest hospitals that include the target department based on patient's current location. However, if the patient does not like the suggestion, she can click 'Others…' to take a look on other related hospitals. If the patient happy with the suggestion, she can click the hospital's name for hospital's address and telephone number. Until now, all data and information present on the interfaces are supported by the ontology.

7.2 Supportive tool for disabilities

Medical care industry is a high risky industry. There are many diseases, and the conditions are ever-changing. At the same time, any medical behaviour is closely related to people's life safety and physical health, not only patients, but also physicians. In order to provide more efficiency and healthy service environment for both patients and physicians, and to create a user-friendly system, our new design system contains many features to service different patients.

The target audience of our system is general public. It is because medical care services face to a wide range of people including men, women, and children. Features like web translation, speech recognition, and voice dictation will be included.

As a result of our new designed system is initially set in Shanghai, China, the default language will be Chinese simplified. However, Shanghai is China's largest and most populous city, and according to the research, it is believed that the officially registered foreigners has over 150000 people, including approximately 31500 Japanese, 21000 Americans, and 20700 Koreans (Shanghai Population Research, 2019). Web translation will main focus on translate from Chinese simplified to English by right clicking the computer mouse, a pop-up window will show on the screen, and the English translation option key will be present.

There are no too much differences between speech recognition and voice dictation. Basically, they are interchangeably used. Speech recognition involves recording spoken words by using a microphone. The audio is then converted into a set of words stored digitally, and shown on the screen. For our system, people from foreign countries are able to use this features with both Chinese and English. Furthermore, it supports real-time streaming. Therefore, users do not have to record everything themselves before using the system. For example, if an American who is living in Shanghai has a headache, and he wants to use the system to help him make an appointment for outpatient with a third-class hospital in Shanghai. The situation is this American can only speak English, therefore, he uses speech recognition to record the key word 'headache' in the system, and clicks search. Then he will directly go to a new page with English search results.

On the other hand, voice dictation works by scanning the aspects of speech that differ between individuals. It is because everyone has a unique way of speaking to them. For example, an elderly man with a quavering voice must have spoken differently than a young man who has full of energy. Nevertheless, another factor is the old generation cannot use keyboards or smartphones to type words properly in China. McCarthy (2018) has an article which considers “up until now, people without an understanding of Chinese characters and the romanised writing system pinyin have been locked out online”. The new system cannot search anything without the Internet, therefore, we have this voice dictation for domestic Chinese people who cannot write Pinyin or Chinese characters, to help them use the Internet and use the system. Moreover, if the old patient is a re-visiting patient, the system will remember his/her voice to confirm the identity of the speaker by using the voice dictation technology. Then, even this old patient cannot type anything on the digital devices, he/she can still use the system with his/her own account and EHR.

8. Dry Run

After all things set up, time to show the sequence of how our users can use the system.
We raised 4 functions in the system as examples.

Function 1: Login the system

Patient and physician inserted wrong identity (ID, password) and an error message popped up.

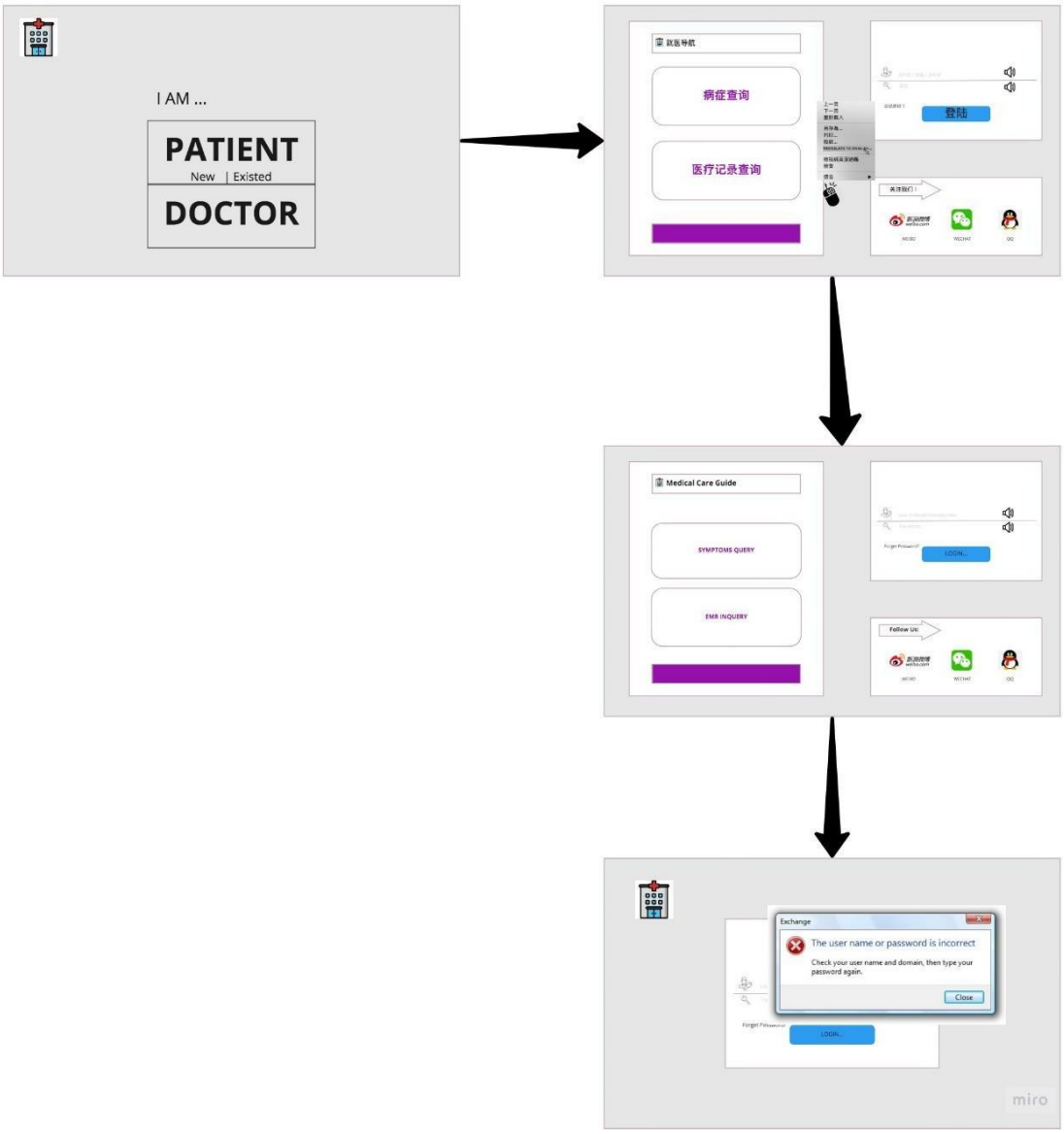

Figure 8.1: Fail to Login

Function 2: Do primary diagnosis

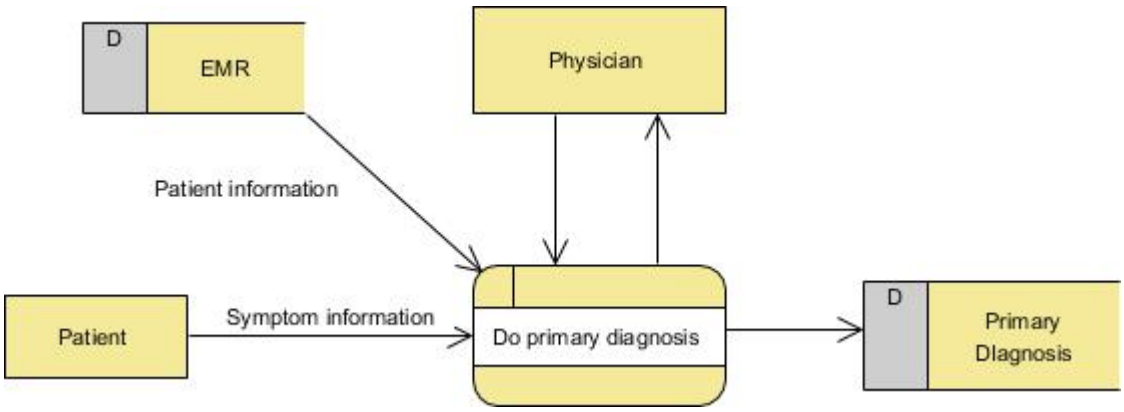

Figure 8.2

Patient login in the system successfully.

Patient select the symptom and get the recommendation of hospital and department.

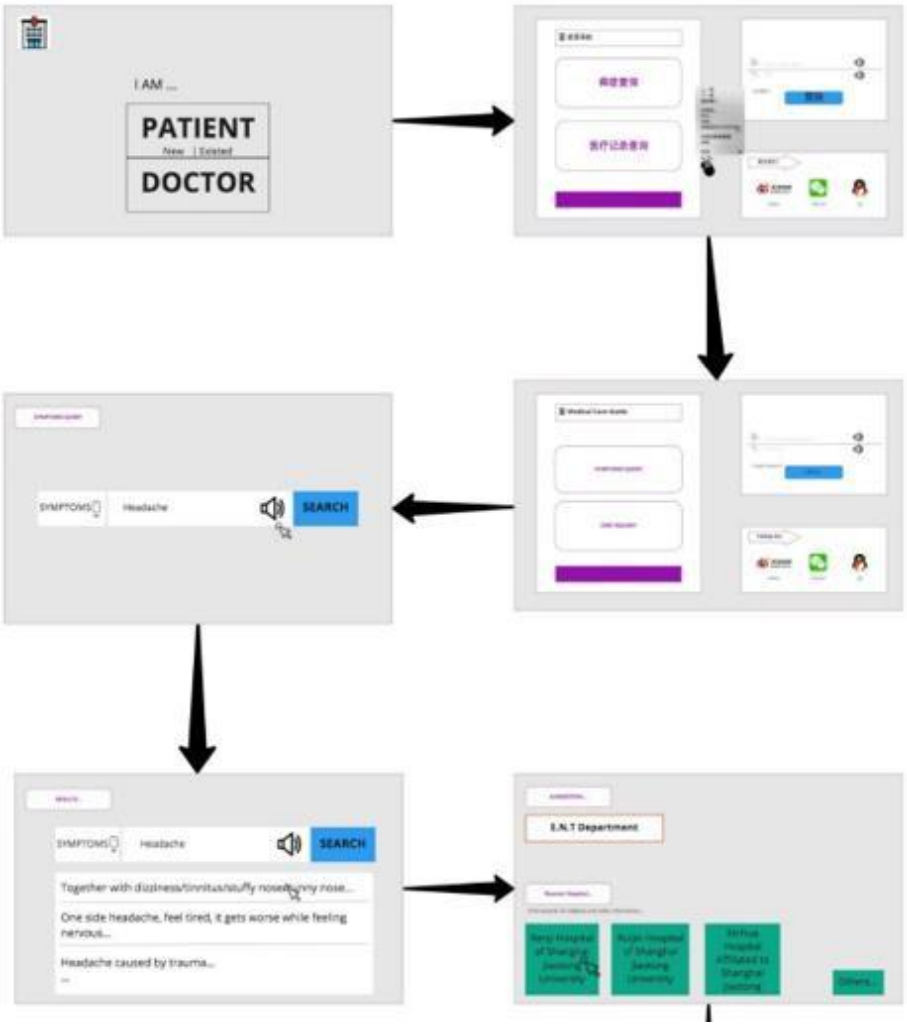

Figure 8.3

Function 3: Make an appointment

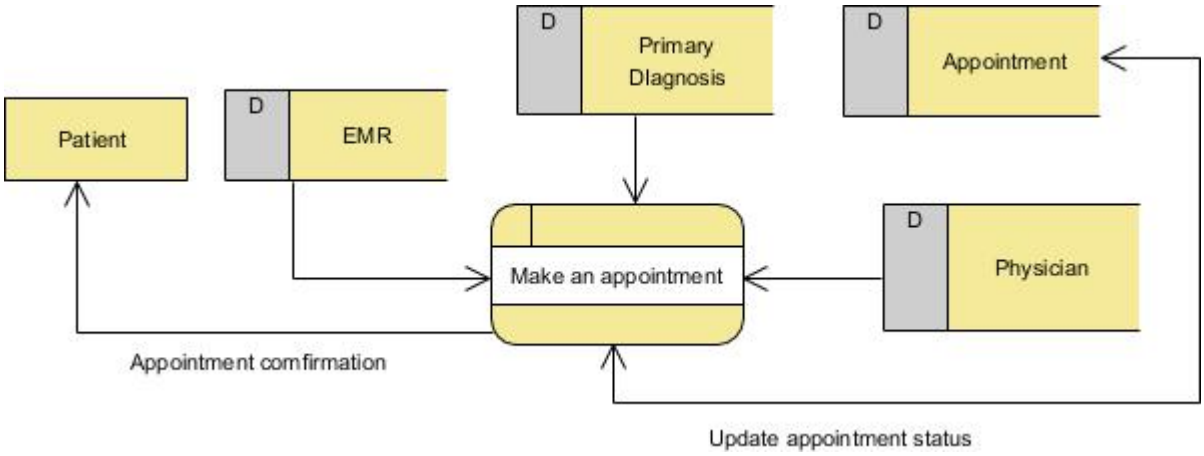

Figure 8.4

Patient login in the system successfully.
Patient select the symptom and get the recommendation of hospital and department.
Patient make an appoinement successfully and get confirmation from the system.

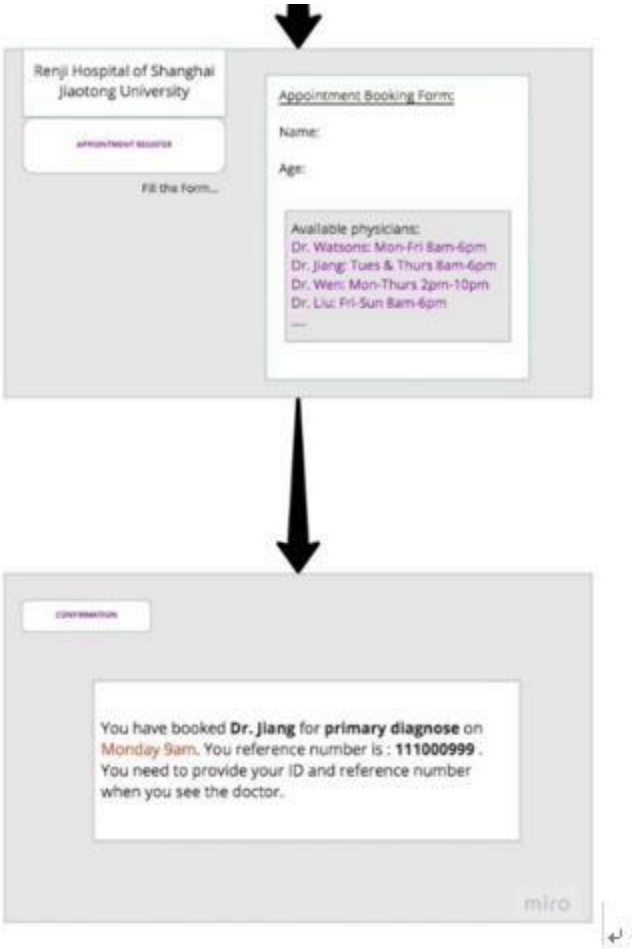

Figure 8.5

Function 4: Phycisian check and update patient's EMR

Physician login the system successfully.

Physician has the right digital siginature and access the patient's EMR successfully.

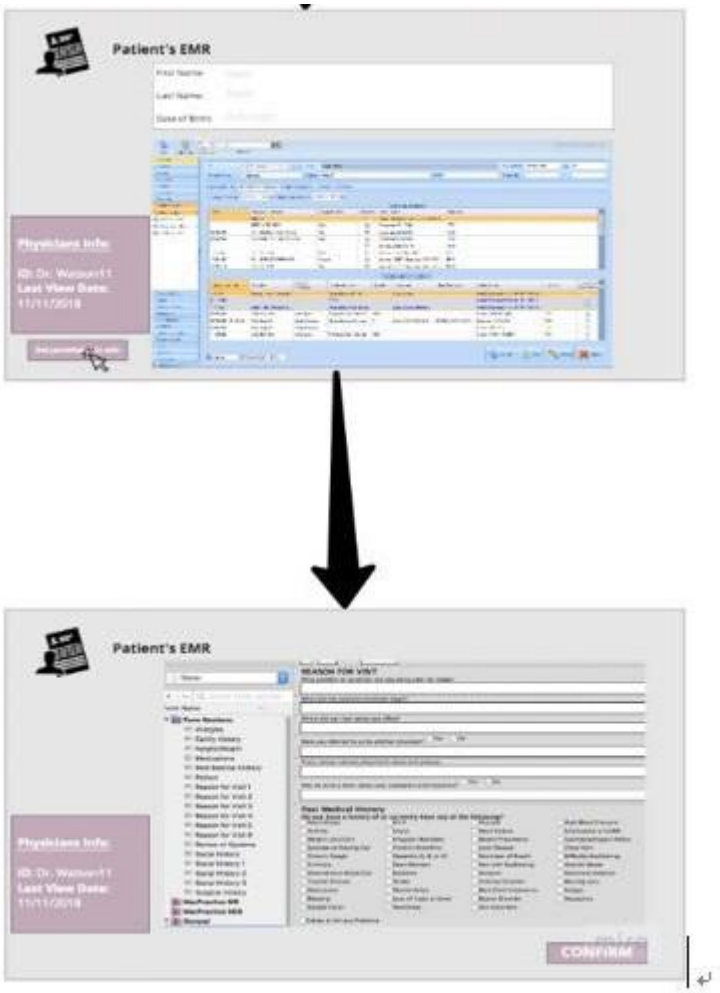

Figure 8.6

Function 5: Warning message for unauthorized access to EMR

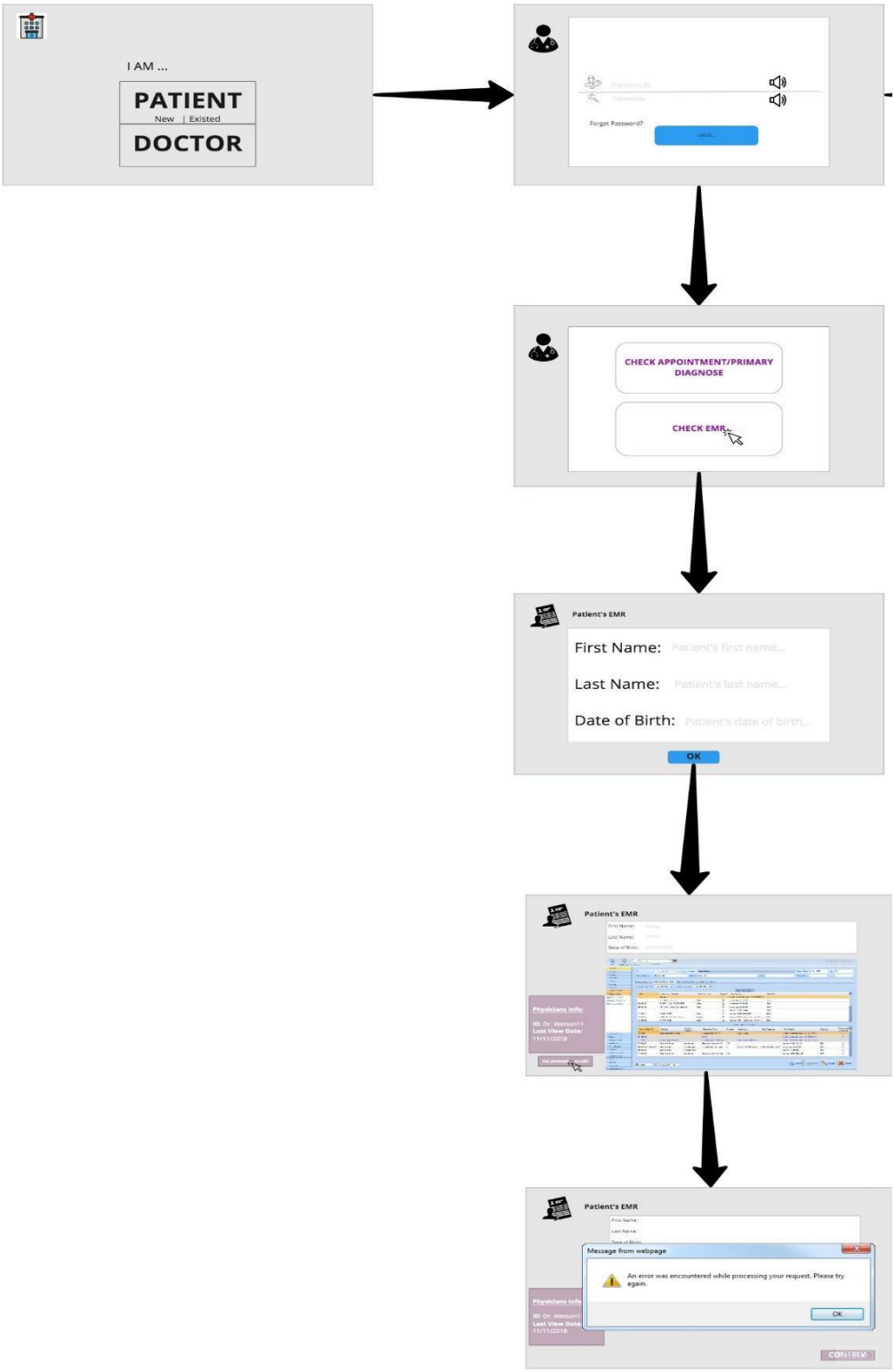

Figure 8.7

If system senses unmatched digital sigiture in the validation. A warining message will pop up in the patient's user interface if unauthorized access happened to his/her EMR.

9. Conclusion

This research project shows us a picture of how ICT can support data management. The real difficulty lies in how to collaborate each component in the system together to achieve expected goal. Design is not only collecting functions and putting them together, but also identifying their priorities and how can they work together. Especially facing with new concepts as Ontology and Distributed Ledger Technology, it is essential to distinguish them with their responsibilities in the system.
The next stage of this research can be located in how to implement this system in the real world. More limitations will need to concern such as physical environment limitation and policy limitation.

10. Reference

Babich, V. & Hilary, G. 2018, 'Distributed ledgers and operations: What operations management researchers should know about blockchain technology', *Forthcoming in Manufacturing & Service Operations Management*.

Brogan, J., Baskaran, I. & Ramachandran, N. 2018, 'Authenticating health activity data using distributed ledger technologies', *Computational and Structural Biotechnology Journal*, vol. 16, pp. 257-66.

Cunningham, J. & Ainsworth, J. 2018, 'Enabling patient control of personal electronic health records through distributed ledger technology', *Stud Health Technol Inform*, vol. 245, pp. 45-8.

Chris Kikel, 'Difference Between Voice Recognition and Speech Recognition', *Total Voice Tecnology*. https://www.totalvoicetech.com/difference-between-voice-recognition-and-speech-recognition/

Hamilton, B.R. 2011, *Electronic Health Records*,2nd edn, McGraw-Hill, New York.

Hollenbeck, S.M., Bomar, J.D., Wenger, D.R. & Yaszay, B. 2017, 'Electronic Medical Record Adoption: The Effect on Efficiency, Completeness, and Accuracy in an Academic Orthopaedic Practice', *Journal of Pediatric Orthopedics*, vol. 37, no. 6, pp. 424-8.

Harris, C.M. and Lazuta, G. 2017, *It's About Patient Care*, McGraw-Hill, United States of America.

Kamran Munir, M. Sheraz Anjum 2018, 'The use of ontologies for effective knowledge modeling and information retrieval', Applied *Computing and Informatics*, No.14, pp. 116-126.

Kuo, T.-T., Kim, H.-E. & Ohno-Machado, L. 2017, 'Blockchain distributed ledger technologies for biomedical and health care applications', *Journal of the American Medical Informatics Association*, vol. 24, no. 6, pp. 1211-20.

Linn, L.A. & Koo, M.B. 2016, 'Blockchain for health data and its potential use in health it and health care related research', *ONC/NIST Use of Blockchain for Healthcare and Research Workshop*. Gaithersburg, Maryland, United States: ONC/NIST.

Porfírio, J., Jacquinet, M. & Carrilho, T. 2013, 'The Impact of ICT and Online Social Networks on Health and Social Services', in Anonymous *Handbook of Research on ICTs and Management Systems for Improving Efficiency in Healthcare and Social Care*, IGI Global, pp. 1226,

viewed May 30, 2019, <https://www-igi-global-com.ezproxy.lib.uts.edu.au/chapter/impact-ict-online-social-networks/78078>.

Report, J. 2014, *ASSESSMENT OF ICT INFRASTRUCTURE IN HEALTHCARE PROVISION IN REFERRAL HOSPITALS: A CASE STUDY OF MUHIMBILI NATIONAL HOSPITAL*, viewedMay 30, 2019.

Robert ARP, Barry Smith, Andrew D. Spear, 2015, *Building Ontologies with Basic Formal Ontology*, The MIT Press, Cambridge. https://ebookcentral.proquest.com/lib/uts/reader.action?docID=3433795

Simone McCarthy: 'The smart technology turning China's illiterate late bloomers into digital natives', *South China Morning Post*, 2nd Sep 2018, https://www.scmp.com/news/china/society/article/2162413/smart-technology-turning-chinas-illiterate-late-bloomers-digital

World Population Review, Shanghai Population 2019, http://worldpopulationreview.com/world-cities/shanghai-population/

2018, 'Market Operation Situation and Development Prospects', *2018 China Electronic Medical Record Industry Analysis Report*, viewed May 30, 2019, < http://baogao.chinabaogao.com/yiliaoqixie/368745368745.html>