

Master's Thesis
Computer Science
Thesis no: MCS-2013-18
December 2013



A Hybrid Cloud Approach for Sharing Health Information in Chronic Disease Self-Management

Cong Peng

School of Computing
Blekinge Institute of Technology
SE – 371 79 Karlskrona
Sweden

This thesis is submitted to the School of Computing at Blekinge Institute of Technology in partial fulfillment of the requirements for the degree of Master of Science in Computer Science. The thesis is equivalent to 20 weeks of full time studies.

Contact Information:

Author(s):

Cong Peng

E-mail: pengcong888@gmail.com

University advisor(s):

Prof. Guohua Bai

School of Computing

School of Computing
Blekinge Institute of Technology
SE – 371 79 Karlskrona
Sweden

Internet : www.bth.se/com
Phone : +46 455 38 50 00
Fax : +46 455 38 50 57

ABSTRACT

Context. Health information sharing improves the performance of patient self-management when dealing with challenging chronic disease care. Cloud computing has the potential to provide a more imaginative long-term solution compared with traditional systems. However, there is a need for identifying a suitable way to share patient health information via cloud.

Objectives. This study aims to identify what health information is suitable and valuable to share from a type 2 diabetes patient when multiple stakeholders are involved for different purposes, and find out a promising and achievable cloud based solution which enables patients to share the health information what and where they want to share.

Methods. To get a clear and deep understanding on the subject area, and identify available knowledge and information on relevant researches, a literature review was performed. And then, a prototype on the case of type 2 diabetes is implemented to prove the feasibility of the proposed solution after analyzing the knowledge acquired from literatures. Finally, professionals and patient were interviewed to evaluate and improve the proposed solution.

Results. A hybrid cloud solution is identified as a suitable way to enable patient to share health information for promoting the treatment of chronic disease.

Conclusions. Based on the research with type 2 diabetes, it was concluded that most records in daily life such as physiologic measurements, non-physiologic measurements and lifestyle are valuable for the treatment of chronic diseases. It was also concluded that hybrid cloud is suitable and achievable for sharing patient-recorded health information among trusted and semi-trusted stakeholders. Moreover, anonymous and patient opt-in consent model are suitable when sharing to semi-trusted stakeholders.

Keywords: health information sharing, hybrid cloud, self-management, chronic disease

ACKNOWLEDGEMENT

This thesis would not be completed without the sincere help and selfless contributions from many people. I would like to express my sincere thanks to them.

First of all, I'm very thankful to my supervisor Dr. Guohua Bai for his valuable and patient guidance throughout the whole process of the thesis work. Thanks to Yan Hu for her advice and discussion when I encountered problems. I'm also grateful to all the teachers who taught me lessons and partners who studied with me for helping me progress. Thanks to my supervisor in Zhejiang University of Technology (ZJUT), Dr. Jianping Ma, who gave me guidance and encouraged me to study further in Blekinge Institute of Technology (BTH). I'm really grateful to BTH and ZJUT for this precious study opportunity.

Special thanks to the interviewees and associates who providing me their valuable opinions. I'm appreciated that they spent their limited time to discuss with me and share their knowledge for contributions.

At the end, I would like to express my sincere thanks to my parents and my girlfriend for their continuous support and encouragement. And thanks to my friends and roommates for the great time in Sweden.

LIST OF CONTENTS

A HYBRID CLOUD APPROACH FOR SHARING HEALTH INFORMATION IN CHRONIC DISEASE SELF-MANAGEMENT.....	I
ABSTRACT	I
ACKNOWLEDGEMENT	II
LIST OF CONTENTS.....	III
LIST OF FIGURES.....	V
LIST OF TABLES.....	VI
ACRONYMS.....	VII
1 INTRODUCTION	1
1.1 BACKGROUND.....	1
1.2 RELATED WORK	2
1.3 AIMS AND OBJECTIVES	3
1.4 RESEARCH QUESTIONS	4
1.5 RESEARCH METHODOLOGY	4
1.5.1 Literature Review.....	5
1.5.2 Prototype	5
1.5.3 Interview	5
1.6 EXPECTED OUTCOME.....	6
1.7 THESIS STRUCTURE	6
2 THEORETICAL WORK.....	7
2.1 EHEALTH	7
2.1.1 Electronic Medical Record (EMR)	7
2.1.2 Electronic Health Record (EHR)	8
2.1.3 Personal Health Record (PHR)	8
2.1.4 Home TeleHealth	9
2.1.5 Health Information Exchange (HIE).....	9
2.1.6 eHealth Interoperability	9
2.1.7 CCR and CDA/CCD	10
2.2 CHRONIC DISEASE SELF-MANAGEMENT.....	10
2.2.1 Challenges Brought by Chronic Diseases	10
2.2.2 Self-Management.....	11
2.2.3 Type 2 Diabetes	11
2.3 HEALTH INFORMATION FOR SHARING.....	12
2.3.1 Sharing Health Data to Promote Treatment.....	12
2.3.2 Privacy and Security.....	13
2.3.3 World Wide Security and Privacy Standards in Healthcare.....	13
2.3.4 Sweden Policy.....	14
2.3.5 Patient Information Sharing in Blekinge	15
2.3.6 Sharable Patient-Recorded Health Information and Data Segmentation.....	16
2.3.7 Conceptual Scenario of Health Information Sharing	16
2.4 CLOUD COMPUTING IN EHEALTH.....	18
2.4.1 Benefits of Adopting Cloud Computing in Sharing Health Information	18
2.4.2 Practical solutions or Research Projects with Cloud Computing	19
2.5 CLOUD COMPUTING DEPLOYMENT MODELS	20
2.5.1 Public Cloud	20
2.5.2 Private Cloud.....	21
2.5.3 Community Cloud	22
2.5.4 Hybrid Cloud	23

2.6	CLOUD SERVICE PLATFORMS.....	24
2.6.1	Amazon Web Services	24
2.6.2	OpenStack.....	26
2.7	SELECTION OF CLOUD DEPLOYMENT MODEL	27
2.7.1	Security and Privacy.....	27
2.7.2	Scalability and Capability	28
2.7.3	Customization	28
2.7.4	Cost.....	29
2.7.5	Legal.....	29
2.7.6	Summary	30
3	PROTOTYPE DESIGN	31
3.1	SYSTEM MODEL.....	31
3.2	SYSTEM DESIGN.....	32
3.2.1	Health Record Structure	32
3.2.2	Conceptual Structure of Private Cloud Platform.....	34
3.2.3	Conceptual Structure of Public Cloud Platform.....	34
3.3	SCENARIO OF CONCEPTUAL SYSTEM	35
3.4	PROTOTYPE SIMULATION.....	37
3.4.1	Health Data Recording and Sharing	37
3.4.2	Shared Content Comparison.....	39
4	EMPIRICAL WORK	42
4.1	INTERVIEWS	42
4.1.1	Purpose.....	42
4.1.2	Design.....	42
4.1.3	Interviewee.....	42
4.1.4	Recording and Transcribing.....	43
5	EVALUATION AND VALIDATION	44
5.1	EXPERT OPINIONS	44
5.1.1	Analysis of Feedback 1	44
5.1.2	Analysis of Feedback 2	45
5.1.3	Analysis of Feedback 3	46
5.2	OPINIONS FROM POTENTIAL USER	47
5.2.1	Analysis of Feedback	47
5.3	DISCUSSION	47
5.3.1	Current Status of the Research Area	47
5.3.2	Contributions	48
5.3.3	Future Trends	49
5.4	QUALITATIVE VALIDITY	49
5.4.1	Credibility.....	49
5.4.2	Transferability	50
5.4.3	Dependability.....	50
5.4.4	Confirmability.....	50
6	EPILOGUE	51
6.1	CONCLUSION	51
6.1.1	Answers to Research Questions	51
6.2	FUTURE WORK	52
	REFERENCES	54
	APPENDIX	59
A.	INTERVIEW TRANSCRIPTS.....	59
	Interview with Mark Braunstein	59
	Interview with Andr éCavalleiro.....	60
	Interview with Marian Lenehan.....	62
	Interview with TED.....	63

LIST OF FIGURES

Figure 1 Research Approach	4
Figure 2 Conceptual Scenario of Sharing Health Information	17
Figure 3 An example of AWS utilization [72]	25
Figure 4 OpenStack Overview [75]	26
Figure 5 Radar Chart of Clouds	30
Figure 6 System Structure	32
Figure 7 Conceptual Structure of Private Cloud Platform.....	34
Figure 8 Conceptual Structure of Public Cloud Platform.....	35
Figure 9 Scenario of Private Cloud Communication.....	36
Figure 10 Scenario of Public Cloud Communication.....	36
Figure 11 Interface of Recording and Sharing to Private Cloud	37
Figure 12 Interface of Choosing and Sharing to Public Cloud.....	38
Figure 13 Workflow of Sharing.....	39
Figure 14 Swift Administration on Horizon Graphical Console Interface	39
Figure 15 Health Record Shared to Private Cloud	40
Figure 16 AWS S3 Administration on Horizon Graphical Console Interface.....	40
Figure 17 Health Record Personal Information Shared to Public Cloud.....	41
Figure 18 Health Record Optional Shared to Public Cloud.....	41

LIST OF TABLES

Table 1 Research Methods.....6

Table 2 Comparison of Clouds30

Table 3 Health Record Content33

Table 4 Interviewees.....43

ACRONYMS

AWS: Amazon Web Services
CCR: Continuity of Care Record
CCD: Continuity of Care Document
CDA: Clinical Document Architecture
CDSS: Clinical Decision Support System
EMR: Electronic Medical Record
EHR: Electronic Health Record
HITECH: Health Information Technology for Economic and Clinical Health Act (U.S.)
HIPAA: Health Insurance Portability and Accountability Act (U.S.)
IaaS: Infrastructure as a Service
ICT: Information and Communication Technologies
NIST: National Institute of Standards and Technology (U.S.)
NPÖ: Nationell Patient översikt
PaaS: Platform as a Service
PHR: Personal Health Record
SaaS: Software as a Service
SLA: Service Level Agreement
WHO: World Health Organization
VPN: Virtual Private Networks
XACML: eXtensible Access Control Markup Language

1 INTRODUCTION

This chapter will describe the background of the problems brought by chronic disease, potential solution, and some related works. Then this chapter will continue with the aim of this research, research questions, selected research methodologies and expected outcome.

1.1 Background

Chronic diseases, such as diabetes, chronic obstructive pulmonary disease and cardiovascular disease, have become a main challenge to health sectors around the world. According to the report by WHO [25], almost 1 in 10 adult has diabetes in the world, and over 90% of all diabetes cases are type 2 diabetes. The incidence and prevalence of chronic diseases is continuing to increase [1]. This challenge asks for a new eHealth approach. Today patients suffered from chronic diseases have to be checked and monitored from time to time, and result in high cost, time consuming and inconvenient for patients [1]. Meanwhile most of the cases for chronic diseases do not need urgent medical diagnosis and treatment. It is therefore more appropriate to adopt effective approach to trace and control the patients' conditions via the healthcare services such as physiological signals monitoring and recording [2] at the home environment.

Thanks to the development of information and communication technology (ICT), the feasibility of home based eHealth has been highly raised [2]. Many physiological signals can be measured by individuals in their living environments during daily activities [3]. It would be efficient if various aspects health parameters could be recorded by the patients themselves at homes and shared to the certain healthcare provider. This leads to the eHealth applications' providing favorable long-term monitoring and healthcare. The home based eHealth enables independent living at home and improves the quality of life for individuals. For instance, patient suffered from diabetes could record clinical parameters such as blood sugar level, cholesterol, triglycerides by digital health devices. Simultaneously, the effect of environmental factors such as nutrition, stress and exercise should also be taken into account [4]. Doctor could provide some clinical suggestions by monitoring the patient through all these daily recorded health data, which was shared by patient.

In addition, the process of this long-term monitoring will generate massive samples of patient self-recording health data. Consider the huge number of patients, there is a big volume of data to be mined. Most of patients do not have much knowledge on healthcare, they may not have the ability to catch "signals" from their daily health data. So if they share their health data to third party web based applications for analysis, they can receive some useful feedback to help them to have a clearer understanding of their state of health or some suggestions. So, if the patients volunteer to share their dairy health data to other healthcare research institutions or service providers, it will benefit the research of the chronic disease, and give improved and appropriate treatment to the patients as well. For instance, Huang et al. [5] proposed a model of chronic diseases prognosis and diagnosis, integrating data mining and case-based reasoning, the knowledge can be discovered from these shared large volumes of data.

Due to the high demand of workload for uploading and sharing these huge amount patient dairy health data, the traditional server-client computing system is no longer suitable. Cloud computing provides a better way to achieve sharing these data to hospital and third-party research or healthcare institutions. According to the definition of U.S. National Institute of Standards and Technology (NIST) [6], "a model that can provide distributed, rapidly provisioned and configurable computing resources (such as servers, storage, applications,

networks and other services), which are on-demand, rapid elastic and measured, to whom have network connections”.

Because of the obvious scalability, flexibility and availability at low cost of cloud services, there is a rapid trend of adopting cloud computing among enterprises or health-related area last few years [7]. Sabah Mohammed et al. [8] proposed a distributed, modular and scalable system, which called Health Cloud Exchange (HCX), allows for sharing different Continuity of Care Record (CCR) Electronic Health Records (EHRs) and discovering related healthcare services to be used by client programs. Carlos Oberdan Rolim et al. proposed a cloud based solution for patients’ vital data collection in [9], which automates the collecting process by using wireless sensor networks. It makes the data collection real-time, and enables the data to be processed by expert systems or medical staff in the cloud. To take the idea further, cloud computing allows the construction of web-based applications, which implement all the needed medical workflows.

However, there are barriers for sharing health information through cloud. The patient health information is very sensitive since it includes personal information, probably the details of medical history, symptoms, treatment, associated diseases or even the family health history [4]. Besides, privacy and security are always the barriers for widely adopting cloud computing in some area. It brings many new challenges for data security and access control when users outsourcing sensitive health data for sharing on cloud servers [10]. Moreover, the threats to privacy and security may not just be brought from the using of cloud computing, the involving of third party institutions brings more complicated problems. The certain doctor in hospital, who is trusted and authorized, absolutely has the access to the health information including personal information. But patients certainly do not want their personal information or some other sensitive part to be shared to other one who is not fully trusted or not authorized. The mechanism of selecting health information for sharing should obey the patient’s willing and relevant regulations, not just the usefulness of health information. Either in the European Union (EU) and the United States (US), there are regulations to address the security and privacy issues in the sharing of electronic health information, such as ISO/TS 18308 standard and Health Insurance Portability and Accountability Act (HIPAA). In Sweden, The Patient Data Act (2008:355) [11], governs all aspects of patient records using.

Another issue is the selection of cloud deployment model. Cloud systems can be divided into four categories from the view of deployment model: public cloud, private cloud, community cloud and hybrid cloud [6]. Each deployment model has different organizational structure and different provisioning location. It results in different utilization style, security level and other characteristics offered by these four deployment models. There are many crucial requirements on sharing health information among groups of different credibility, such as privacy, accessibility and utility. Adopting an inappropriate cloud deployment model will bring problems.

This study will try to find out a solution to address these issues. Since it is impossible to study all kinds of chronic diseases during a limited period of time, and the pattern of health data sharing of many chronic diseases is similar, therefore this study will take the frequently-occurring disease, type 2 diabetes, as a case to start with.

1.2 Related Work

There have been many researches related to this area. M. Hussain et al. proposed a Smart Clinical Decision Support System (Smart CDSS) [12]. It gathers data from diverse modalities, such as sensors, patient profile information, social media, clinical knowledge

bases and medical expertise, with cloud architecture to generate standards-based recommendations to patients. But this CDSS lacks of some important health data from patient daily life when adopting it to chronic disease patient, such as physiological measurements, diet or daily activities.

For exchange health data between eHealth applications and health information systems, A. S. Radwan et al. [13] proposed a cloud based solution which emphasizes standardizing a data exchange mechanism and interoperability between health information systems and eHealth applications. In addition, this system allows patient to create and store PHR, and allows healthcare professionals to access to PHR. In this system, Information Systems (HIS) and third party eHealth applications were combined in the same environment, it will probably bring potential problems. Another approach is a web based data management application proposed by SS Deo et al. [4]. It is used to collect, manage and analyze medical information of patients who suffered from diabetes. This system allows both authorized healthcare providers and patients have access to data for updating and analysis.

To meet the increasing requirements of enabling chronic disease patients to manage and improve their healthcare delivery, a distributed service application on supporting chronic disease self-management was proposed by Ali Sunyaev et al. [14]. The proposed application will keep tracking blood glucose and other related health measurements, allow reviewing history records by displaying graphs over a period or comparing several periods. To enhance the usability and interoperability, this system is able to integrate with the Microsoft HealthVault and the closed Google Health. But this application lacks of information sharing, it is just for personal using, therefore it has no concern on security and privacy.

For collecting patients' vital data, C. O. Rolim et al. proposed a cloud based solution in [9], which automates the collecting process by using wireless sensor networks. It makes the data collection real-time, and enables the data to be processed by expert systems or medical staff in the cloud. L. Fan et al. [15] identified the concerns on service integration, large scale deployment, and security, integrity and confidentiality of sensitive medical data. For addressing these challenges, they proposed the cloud based solution Data Capture and Auto Identification Reference (DACAR). But in this solution, there is no concern on what actually be shared and what is the role of hospital.

1.3 Aims and Objectives

The aims of this thesis are to identify a suitable form of health information sharing for chronic disease self-management, figure out the problems of using cloud computing to achieve health data sharing among multi-groups, and find out a suitable solution to address these problems. The solution should enable patient to share the health information what and where they want to share, and improve the quality and expandability of chronic disease self-management. To prove the feasibility and practicality of the notion, the research will start with the self-management of type 2 diabetes, and then a prototype will be implemented and evaluated.

The aims will be achieved by addressing the following objectives:

- Find out what have been done in health information sharing in similar application scenario (either based on cloud computing or other techniques) through literature review and online search.
- Identify a form to record and share health information for the treatment of type 2 diabetes and guarantees privacy and security in the case of self-management.

- Identify the challenges when using different deployment models of cloud computing to share sensitive data among multiple groups which have different credibility.
- Find out a suitable and feasible solution by prototype design to address the challenges when using cloud computing to share sensitive data among multiple groups which have different credibility.

1.4 Research Questions

RQ1: What form of patient self-management health information sharing guarantees privacy and security when utilizing cloud computing among trusted and semi-trusted stakeholders?

RQ2: Which cloud deployment model is suitable for sharing patient-recorded health information among trusted and semi-trusted stakeholders?

1.5 Research Methodology

For figuring out the challenges and answering our research questions, literature review, prototype and interview will be performed. Figure 1 shows the anticipated research process and relevant knowledge gained by certain work. The details will be described as following.

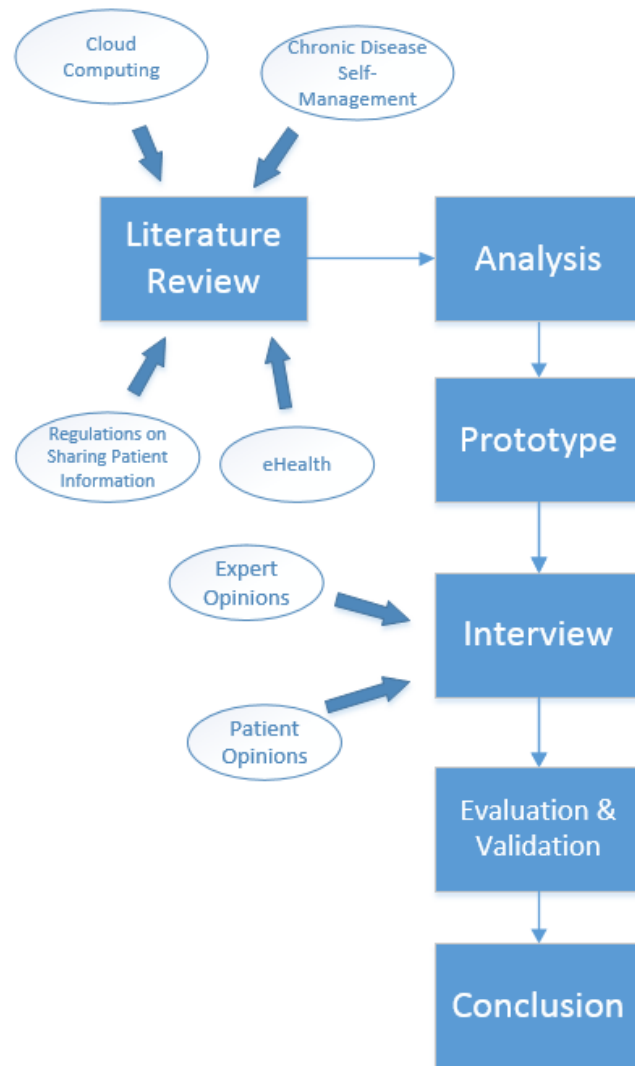


Figure 1 Research Approach

1.5.1 Literature Review

At the beginning of the research, a clear and deep understanding on the subject area is needed. Furthermore, identify the available information, which relevant to the particular research domain and research questions, is also necessary. A literature review is a suitable research method to do this. It is commonly used by researchers to explore the existing knowledge on a subject area. As Levy and Ellis defined in [16], an effective literature review is a process to “collect, know, comprehend, apply, analyze, synthesize, and evaluate quality literature in order to provide a firm foundation to a topic and research method”. Through this process, various aspects of knowledge can be acquired, include the development situation of chronic disease self-management in home based eHealth, Swedish regulations on sharing patient information, the development situation of cloud computing and so on.

In addition, there are many relevant researches already proposed and published, information and ideas from existing researches therefore are valuable to be referenced, such as what have been done in health data sharing in similar application scenario either based on cloud computing or other techniques, how different categories of cloud computing suited in different application areas. The research protocol defined by Kitchenham [17] is a good way to follow. A literature review can probably give an answer to RQ1 and partly answer to RQ2.

To search for the past and current researches and other supporting materials, the publications through the online databases, which are provided by the library of Blekinge Institute of Technology (BTH), will be searched. The helpful databases include ACM Digital Library, IEEE Xplore, SpringerLink, Google Scholar and other online sources.

1.5.2 Prototype

On the basis of the knowledge gained through analyzing the information of literature review, there will be a theoretical cloud-based method for sharing patient recorded health information. There is a need to test the feasibility of the proposed conceptual solution.

Prototype development is defined as the process of preparing a device, technique or system that demonstrates the feasibility of a solution to a problem [18]. It has been widely used in systems development and research. In this study, a prototype, which can achieve patient recorded health information sharing via hybrid cloud, is implemented to prove the feasibility. Cloud services provided by Amazon, Google or other enterprises, and open source cloud platform such as Openstack, Eucalyptus, could be used to support the development of prototype. The sharing health data format can be designed based on popular standard such as Continuity of Care Document (CCD) or CCR.

1.5.3 Interview

Besides, for supporting and evaluating the proposed solution, there is a need to get more detailed and practical understanding and expert opinions from professionals in eHealth, cloud computing and healthcare. The expectation and opinion from patient who suffered from chronic disease are also necessary. For achieving this, a qualitative research method, interview, will be performed. Interview is a useful method for data collection with better response rate and closer judgment of people's experience, opinion, desire and feelings [19].

The interview participants could be found in university, research community, social network such as LinkedIn, PatientsLikeMe and Google+, and hospital. The selection of interviewees will be restricted to meet the criteria which include highly relevant to the area, years of experience and dependability. Semi-structured interview will be used, which contains closed questions for gathering anticipated information and open questions for gathering unforeseen information [20]. The interview could support to answer RQ1 and RQ2.

Research Questions	Research Methods
RQ1	Literature Review, Interview
RQ2	Literature Review, Prototype, Interview

Table 1 Research Methods

1.6 Expected Outcome

The main expected outcome of this thesis is a suitable cloud-based solution which can be used to share patient recorded health information in the application scenario of chronic disease self-management. In particular, the following outcomes will be conducted through the process of this research:

- Knowledge of suitable and valuable health information for sharing in self-management of type 2 diabetes under relevant regulations.
- Properties and adaptabilities of different categories of cloud computing deployment model in the application of sharing chronic disease patient-recorded health.
- A prototype of the cloud-based solution which suit for sharing patient-recorded health information in the application scenario of chronic disease self-management.

1.7 Thesis Structure

Chapter 1 Introduction: This chapter consists of the background of the problems brought by chronic disease, the potential solution, as well as the aim of this thesis, the research questions, the selected research methodologies and expected outcome.

Chapter 2 Theoretical Work: This chapter consists of related concepts in eHealth, introduction of chronic disease self-management and health information sharing. In addition, the description of cloud computing, the introduction of two cloud computing platforms, the description and comparison of the four categories of cloud deployment model will also be included.

Chapter 3 Prototype Design: This chapter consists of the design of the prototype, and the presentation of the vertical prototype with the key functions.

Chapter 4 Empirical Work: This chapter introduces the plan of interview, describe how the interview is organized, the selection of interviewees, and how the interview is conducted.

Chapter 5 Evaluation and Validation: This chapter consists of the analysis of feedbacks and evaluation from interviewed experts and patient, as well as the discussion and validation of this research.

Chapter 6 Epilogue: This chapter consists of the conclusion of the research study, the answers to research questions, and future works.

2 THEORETICAL WORK

This chapter will start with some concepts in eHealth, introduce the chronic disease self-management and health information sharing for improving the treatment of chronic disease. Then describe the cloud computing and its related work on supporting healthcare, introduce two cloud computing platforms. And four categories of cloud deployment model will be described and compared.

2.1 eHealth

eHealth is a relatively recent area which is supported by ICT to provide healthcare services and medical informatics [21] [22]. There exists different terms associated with this area. Today, the commonly used term is “eHealth”. The World Health Organization (WHO) defines eHealth as “the cost-effective and secure use of information and communications technologies in support of health and health-related fields, including healthcare services, health surveillance, health education, knowledge and research” [23].

The goal of eHealth is to improve the cooperation and coordination of healthcare to achieve improving the quality of care and reducing the cost of care at the same time. The high benefits and broad prospect accelerated the rapid development of eHealth. The main and deeply researched eHealth applications around the world include [22][24][25]:

- Electronic Medical Record (EMR)
- Electronic Health Record (EHR)
- Personal Health Records (PHR)
- Chronic Disease Management
- Home TeleHealth
- Electronic Prescribing
- Telemedicine
- Consumer Health Informatics
- Health knowledge management
- Virtual healthcare teams
- mobile-Health or mHealth
- Medical research eHealth Grids
- Healthcare Information Systems

2.1.1 Electronic Medical Record (EMR)

EMR is defined in [24] as “an individual electronic record of health related information in the medical process that can be created, gathered, managed, and consulted by authorized clinicians and staff within one healthcare organization”. It was designed under the requirements and workflow of healthcare from hospitals, to meet the service and management of hospital. The content of EMR includes diagnoses, medical histories, lab tests,

images and prescriptions [26]. It is the digital version of patient paper chart, aims to improve the ability of observation and treatment.

However, there exist multiple standards of EMR, and lacks of nationally recognized interoperability standards. It will restrict the usage of EMR in the medical information exchange among healthcare organizations.

2.1.2 Electronic Health Record (EHR)

EHR is defined in [24] as “an individual electronic record of health-related information that adopts nationally recognized interoperability standards and that can be created, managed, and consulted by authorized clinicians and staff across more than one healthcare organization”. It will digitally store the healthcare information of an individual, which includes data of observations, treatments, laboratory tests, imaging reports, therapies, drugs administered, patient identification information, legal permissions, insurance benefit and so on, for a lifetime for the purposes of supporting patient care, education, and research [27]. And EHR is managed and utilized by authorized care providers.

Compare with EMR, EHR is the parts of the EMR which are appropriately shared with stakeholders outside the hospital. It focuses on the efforts of employing the most comprehensive information available in multiple locations and systems throughout the health industry [24]. EHR is a more efficient way of organizing and sharing healthcare information exceed the scope of one healthcare organization or single health encounter for healthcare providers. It can present a comprehensive, longitudinal, patient focused view of a patient’s health history, which is recorded by various healthcare providers include hospitals, long-term care facilities, home based care, laboratories, pharmacies, and other authorized healthcare services. Due to the depth and breadth of data, an EHR thus can offer a perspective on changes in health and medical conditions overtime. It is especially useful in recording healthcare of chronic disease patients.

2.1.3 Personal Health Record (PHR)

Recently, there is a trend in healthcare delivery that led to a shift from EHR systems, which is controlled by healthcare providers, to PHR systems, which is controlled by patients themselves [28]. PHR is defined as “an individual electronic record of health related information that can be acquired from diverse sources and be controlled, managed and shared by the owner with adopting nationally recognized interoperability standards” in [24].

PHR systems allow individuals to import their health records which may include medical history, laboratory and imaging results, list of medical problems, and medication history from EHR systems. The health measurements uploaded from individuals’ home based electronic health devices or other daily health data collected by individuals themselves will also contribute a lot, especially on people living with chronic diseases. [28] Thus it provides a tool for collecting and sharing individual’s own health information. The information that a PHR contains is under the control of the individual, this is the key aspect of PHR distinct from EMR and EHR. This enables and encourages individuals to become more engaged in their health and care, such as allowing patients to create, manage and control their PHRs through the Internet. It will provide a well-rounded and accurate picture of individuals health, lead to more informed healthcare decisions, improved personal health status, and cost reduced and quality improved healthcare [24][28].

There are PHR services available such as Microsoft HealthVault [29] and open platform Indivo [30]. However, there are barriers for adopting and using PHR by individuals, for instances, lacking of interoperability and integration with EMR or EHR, friendly and convenience of user interface, and trust of users [28].

2.1.4 Home TeleHealth

The modern term Home TeleHealth is described as “the services provided by home healthcare providers which use the ICT to link home-based patients to one or more outside sources of healthcare information and services” [31]. In brief, it is the application of eHealth to the home environment.

The purposes of Home TeleHealth are to coordinate the care of patients who live with chronic conditions, to avoid unnecessary long term in-hospital healthcare, to provide new methods of early intervention and better long term treatment of these patients [32]. According to A. Darkins et al., although the Home TeleHealth unit rose additional costs, a solution of enterprise-wide Home TeleHealth is still a cost-effective way to manage chronic disease patients [32].

2.1.5 Health Information Exchange (HIE)

HIE is the process of sharing health related information among different organizations while guaranteeing the confidentiality, privacy and security of the shared health information [24]. It can gather information from health research labs, disease communities, organizations via the utilization of electronic health record system and so on, to support coordinated healthcare by utilizing EHRs. Thus HIE can promote healthcare of patients by providing key information through utilizing PHRs, as well as support research, emergency case and so on. Exchange of health information may be a key strategy to improve the quality and efficiency of healthcare [33].

However, HIE requires nationally recognized standards that can make different systems of organizations understand the meaning of the shared information. The process of which can be regarded as interoperable exchange, and the nationally recognized standards can be regarded as interoperability standards [24].

2.1.6 eHealth Interoperability

eHealth interoperability is the ability of diverse information systems, components and software applications to work together and exchange data meaningfully [34] [35].

Usually, the management of chronic disease involves many healthcare organizations or related health service providers, which may deploy different healthcare information systems. Interoperability is a key issue to provide efficient and convenient healthcare services in an environment where multiple healthcare systems are deployed. To make the health data sharing accurately and effectively among different healthcare information systems and applications, there need to be some common standards. [33]

Health Level Seven (HL7) International is the main international healthcare informatics interoperability standards organization. Its name derives from the seventh layer of the Open

Systems Interconnection Reference model – the application layer. HL7 and its members provide a framework and related standards to exchange, integrate, share, and retrieve electronic health information. HL7 standards support clinical practices to manage, deliver, and evaluate healthcare services, and are recognized as one of the most commonly used interoperability standards in the world. [36]

Other widely deployed interoperability standards include Standard for Systematized Nomenclature of Medicine-Clinical Terms (SNOMED CT), which is the most comprehensive, structured clinical nomenclature [33]. Digital Imaging and Communications in Medicine (DICOM) is a messaging standard used for exchanging of medical images and related information [37].

2.1.7 CCR and CDA/CCD

The electronic health data is transferred from computer system to computer system, whether it is structured or free text. It's necessary to make the computer systems understand what each section of data represents [33]. So the health data should be wrapped in some standard electronic document format when transferring and parsing.

The Continuity of Care Record (CCR) is a standard of health record that be developed together by American Society for Testing and Materials (ASTM) International and other sponsors [38]. While HL7 developed The Clinical Document Architecture (CDA) standard, which specified the structure and semantics of clinical documents [39]. Both the CCR and the CDA are eXtensible Markup Language (XML)-based, and are used for the purpose of health information exchange between healthcare providers and patients.

Both the CCR and CDA have the similar ultimate goal of improving healthcare performance, and they provide similar usability [40]. However, they are still two different formats, which is ineffective for interoperability. That is the reason why Continuity of Care Document (CCD) is created. The CCD is a joint effort of HL7 International and ASTM International, which is a guide for implementing the sharing of CCR data while utilizing the HL7 v3 CDA Release 2 [39]. The CCD is built on the CDA elements, while the data itself is defined by CCR [41]. The CCD establishes a rich set of templates to represent the typical sections of a health record, and these templates are expressed by the constraints of the CDA. These templates, which include problems, vital signs, medications, immunizations and so on, can be reused in other CDA documents [39]. This will support the interoperability in eHealth in a wide range [39].

2.2 Chronic Disease Self-Management

2.2.1 Challenges Brought by Chronic Diseases

Chronic diseases have become the main threats to health all over the world. The incidence and prevalence of chronic diseases continue to increase [1]. According to WHO, “chronic diseases, such as diabetes, heart disease, chronic respiratory diseases and hypertension, brought about 63% of all deaths, are becoming the main causes of death around the world so far. Over 36 million people died from chronic diseases in 2008, nine million of them were under the age of 60” [42]. The greatest price is being paid by patients who suffering chronic diseases.

Chronic diseases are the diseases influence patients in long last duration. The patients suffered from chronic diseases have to be monitored from time to time. That's very high cost, time consuming and inconvenient for most of patients, especially the limitations to patients' daily activities and inconvenience of aged people. The spending on medical care is continuously growing due to chronic diseases. The challenge is even more complex due to the limited health care resources [1]. There are increasing burdens of resource and cost on chronic disease healthcare. These factors drive better ways to address these barriers.

2.2.2 Self-Management

Most of the chronic diseases usually do not need immediate medical diagnosis and treatment. For the patients with cardiopathy or hypertension, it is adequate to trace and control the patients' conditions via the healthcare services such as physiological signals monitoring and recording [2]. Moreover, lifestyles, diet and metabolism are the main factors of causing most common chronic diseases. Most conditions of chronic diseases therefore might be avoided by changing behaviors, such as quitting smoking, keeping healthy diets, or increasing physical activities [43].

For instance, diabetes is a disease that patients need to be monitored continuously for a number of parameters, such as blood sugar level, blood pressure and so on. Simultaneously, the effect of environmental factors such as nutrition, stress and exercise should also be taken into account when managing this disease [4]. It would be beneficial and efficient if these various aspects of parameters could be recorded by the patients themselves at homes.

Thanks to the development of ICT, the feasibility of home based eHealth has been highly raised [2]. Most of the health condition monitoring and basic treatment can be handled by patients themselves in their living environments during daily activities [3]. There are already some commercial or research-based solutions for the mentioned way of self-management. For instance, HealthVault [29], launched by Microsoft, is a web-based PHR system to store and maintain health and fitness information. There are a lot of specific third-party applications, such as blood pressure management tool or medical image viewer, and hundreds of devices, such as blood glucose meter, blood pressure monitor, cooperated with this platform to record health data and manage health.

The home based eHealth self-management enables independent living at home and improves the quality of life for individuals. Doctor could provide some clinical suggestions by monitoring the patient through all these patient shared daily health data, as well as reply feedback of medical examination report which the patient undergone. This leads to the eHealth applications' providing favorable long-term monitoring and healthcare by linking patients to the sharing of health information and enabling chronic diseases self-management activities [44].

2.2.3 Type 2 Diabetes

This study will start with the self-management of type 2 diabetes to prove the feasibility and practicality of the notion, and then try to apply the pattern to other chronic diseases.

According to the report by WHO [25], almost 1 in 10 adult has diabetes in the world, measured by elevated fasting blood glucose ($\geq 126\text{mg/dl}$). People with diabetes have increased risk of heart disease and stroke, and are 10 times more likely to need a lower limb

amputation than people who do not have diabetes. It is also one of the leading causes of renal failure, visual impairment and blindness.

According to the information gathered by C. Nordqvist [46], lacking of enough proper functioned insulin production, the cells in the body have improper response to insulin, or both, are the reasons cause diabetes. There are three types of diabetes: type 1 diabetes, type 2 diabetes and gestational diabetes. Over 90% of all cases of diabetes worldwide are type 2 diabetes. Most type 2 diabetes patients will have the condition of prediabetes first. The blood glucose level of a person had prediabetes is higher than normal, but not high enough for a diabetes diagnosis. As one kind of chronic disease, type 2 diabetes will get worse gradually. In addition, obesity causes the body to release chemicals which will destabilize the body's cardiovascular and metabolic systems, this will lead to a much higher risk of developing to type 2 diabetes. [46]

People could control their symptoms of type 2 diabetes by keeping a healthy diet, losing weight, doing a plenty of exercises, and monitoring blood glucose level [46]. Therefore, patient can self-manage at home by daily recording blood glucose level parameter, diet, insulin intake, physical exercise, feet checking, psychological status (such as mood and pressure) and living habits (such as smoking and alcohol abuse), and then sharing to healthcare providers.

2.3 Health Information for Sharing

2.3.1 Sharing Health Data to Promote Treatment

The treatment procedure for patients suffered from chronic diseases involves continuous monitoring and recording of health related vital signs. The process of this long-term monitoring will generate massive samples of patient self-recording health data. Synchronizing these patient data to the certain healthcare provider is an obvious benefit and convenience for the management and treatment of patients with chronic diseases, such as type 2 diabetes.

In addition, consider the huge number of patients, there are a big volume of data to be mined. If patients volunteer to share their dairy health data to help the research of their diseases, make other healthcare research institutions or healthcare service providers get the authorized access to these massive samples of patient recording data, it will benefit the research of the chronic disease, and give improved and more appropriate treatment to the patient in return. For instance, as Huang et al. mentioned in [5], knowledge can be discovered from these large volumes of data. They proposed a model of chronic diseases prognosis and diagnosis integrating data mining and case-based reasoning. This model will also be convenient for chronic diseases knowledge creating, organizing and sharing.

Moreover, most of patients do not have much knowledge on healthcare, they may not have the ability to catch some “signals” from their daily health data. So if they share their health data to some trustful third party web-based applications for analysis, they can receive useful feedbacks to help them to have a clearer understanding of their health status and some suggestions. M. Hussain et al. proposed a Smart Clinical Decision Support System (Smart CDSS) [12] which can generate appropriate medical recommendations to users by collecting patient activities data, social media data, and medical history. Once there are more and more shared long-term health records, patients can find people who have similar conditions. If one patient has good progress, others will benefit from his shared records of treatment and lifestyle.

2.3.2 Privacy and Security

The patient health information includes personal information, details of medical history, symptoms, treatment, associated diseases or even the family health history [4]. It is very sensitive, therefore the adopting of sharing electronic health information with other providers will bring new risks to secure health information [47].

In addition, sharing health information to third party organizations brings more complicated problems. The hospital's certain doctor, who is trusted and authorized, absolutely has the access to the health information include personal information. But patients certainly do not want their personal information or some other sensitive part to be shared to other one who is not fully trusted or not authorized. The mechanism of selecting health information for sharing should obey the willingness of patients and relevant regulations, not just the usefulness of health information. Therefore it needs to consider privacy and security when sharing health information.

Privacy

The health information, although is shared, is owned by patients. It is important to ensure that only the patient or the authorized representative can authorize exactly who can view the shared health information and for what purposes [33].

All the information generated by a patient is not at the same sensitive level. In other words, there are some parts of information is that the patient willing to share for some purposes, and some parts not. It has to deal with the problem of how to logically group the health information to share for a patient's different purposes under different conditions.

Data segmentation provides a method to protect specific sections of health information while giving choices to patients, and abiding requirements of legal. Technical consideration and definition of sensitive information have to be addressed when segmenting data. [48]

Security

Security is assurance that only authorized persons or entities can access to patients' data. Employees illegitimate access and theft is one of the most frequent reason of data leakage [33], as well as innocent disclosure because of system problems. Another issue is unauthorized access and malicious attacks from outside [28].

Providers have to ensure that the shared health data will not leak. So the encryption, identification and access control of patient health data are not optional [33].

2.3.3 World Wide Security and Privacy Standards in Healthcare

The Health Insurance Portability and Accountability Act (HIPAA) is a regulatory process for securing electronic health records, and was enacted by the United States Congress and signed by President Bill Clinton in 1996 [49]. It ensures health providers, practitioners and middlemen to keep the radar on what information is protected and how it is protected [50].

For addressing the privacy and security concerns in the electronic transmission of health information, HIPAA was expanded by the Health Information Technology for Economic and Clinical Health Act (HITECH) [51]. The Privacy Rule of HIPAA provides protections for

patients' identifiable health information, which is held by covered entities and their associates, and gives a range of rights to patients with respects to their information. It also balances the health information disclosure permits of the needs for patient care and other important purposes at the same time. [52] The Security Rule of HIPAA aims to guarantee the confidentiality of patient electronic health information by specifying various safeguard methods for covered entities and their associates [52].

The policies on the creation of privacy legislation in the European Union (EU) are different to the United States (US) [53]. The approach to legislation of EU favors the participation among businesses and governments, which is opposed to the self-regulation approach of US.

In the last few years, there are various researches on the security issues of healthcare information systems. The European Advanced Informatics in Medicine/Secure Environment for Information Systems in MEDicine (AIM/SEISMED) project addresses a range of security issues within healthcare, and provides practical guidelines for secure healthcare establishment [54]. ISO/TS 18308 standard defines security and privacy issues for processing, managing and communicating EHR [55]. The Working Group 4 of International Medical Informatics Association (IMIA) was set up to investigate the issues of data protection and security in EHR network systems and common security solutions for communicating patient data [53].

2.3.4 Sweden Policy

In Sweden, healthcare is regulated by various Acts. The Health and Medical Service Act regulates healthcare in general. Its goal is to assure the good health and care of entire population on equal terms. In addition, there are five specific acts effect electronic health records in particular: The Freedom of the Press Act, The Secrecy Act, The Patient Register Act, The Act on Healthcare Records and The Personal Data Act. In 2008, the new law Patient Data Act published. [56]

The Patient Data Act (2008:355) [11], which was draft by Ministry of Health, governs all aspects of patient records usage, especially on how personal data and medical records should be kept and protected in healthcare sectors. It regulates the confidentiality, cohesive record, access right of patient data, and rights of patient.

The provisioning of the Patient Data Act replaces the previous one in health law and health records law. The purpose of Patient Data Act is to organize information management, enhance patient safety and privacy, and promote cost efficiency in healthcare. For example, a patient can have the right to influence who and in what context his health record can be read. Appropriate technical and organizational measures should be implemented to protect the patient data by the providers [56]. For protecting the patients privacy, the Patient Data Act states that if a patient record, a copy or photocopy of the document has been released to anyone, it must be documented in the patient record who has received the document, the transcript or the copy, and when that has been disclosed. The electronic access to a patient's data for health purposes should be restricted, unless the patient agrees to or the patient's consent cannot be obtained and the information probably have an impact on the care when in urgently need. When in the scenario of health record exchange, a patient should be informed and have the right of objection before the data is made available to other healthcare providers through record exchanging.

2.3.5 Patient Information Sharing in Blekinge

National Patient Overview (Nationell Patientöversikt - NPÖ) is a national electronic service in Sweden that enables medical professionals to share patient healthcare records with other healthcare providers through computer. It is created by the county councils in collaborated with Carelink. The purposes of NPÖ are to facilitate cooperation among different healthcare providers in Sweden, and to give patient the access to own their healthcare information. Moreover, another objective is to establish common services in healthcare IT infrastructure in Sweden. [57]

For guaranteeing privacy of patient health records, only the healthcare professionals who have a patient relationship can access to the patient's information through the NPÖ. Security around the usage of NPÖ is very high. Health professionals must be competent and have valid e-service identifications to access the NPÖ. Patient needs to be asked and agree before anyone takes note from the information. Every time when someone takes note of the information from patient record, it will be recorded. Therefore it is possible to retrospectively track who has received the information, where and when it happened. [57]

NPÖ does not replace the former patient record. It just makes the information available to more healthcare providers. If patients do not want some information appear in the NPÖ, they can block the records and decide who can get access to these blocked information. In Blekinge County Council, the following information from the records will be shown: [57]

- The name of patient, date of birth, address, telephone, contacts information to associates or representatives if needed. Information in 3 years is presented.
- Attention signal, such as allergy to drugs or foods, treatment with blood thinning drugs or the certain treatment that the patient does not want. Information from 1997 is presented.
- Information includes notes and summaries of care, visits or stays. Information in 3 years is presented.
- Diagnoses back from 1997.
- Drugs which the patient was prescribed. Information in 3 years is presented.
- Results of examinations and tests, such as x-rays, blood tests and urine tests. Information in 3 years is presented. (Results of X-ray referrals and test results are not displayed at the moment.)

For the convenience of treatment in different places, a patient can give consent to the so-called cohesive record. Cohesive record means that healthcare providers share medical records to each other no matter where they have sought healthcare. In other words, it means that if a patient receive healthcare in a new place, the new healthcare provider can receive the patient's health record which contains the past medical test results, medications and treatments, under the condition of both the new and the former healthcare providers are connected to the cohesive record. Cohesive record makes it possible to read patient records, which contains a complete picture of a patient's health history, directly in electronic form, and makes the treatment easier, safer and faster. To strengthen patient privacy, safety and security, the cohesive record will be regulated by Patient Data Act. [57]

2.3.6 Sharable Patient-Recorded Health Information and Data Segmentation

Based on the analysis of CCD template and PatientsLikeMe, it is identified that the followings are the personal health information which can be recorded by patients themselves at home:

- Personal information: besides identification information, age and gender are also related.
- Personal health status: diagnosed diseases, smoke history, and other health related information.
- Physiologic measurements: Blood Glucose, Blood Pressure, Pulse and so on.
- Medication dosage: amounts of injected insulin per day, or other medicines the patient is taking.
- General health measurements: weight, height, cholesterol or other related measurements.
- Non-physiologic records: mood, pain, fatigue, insomnia, blurry vision or some other related symptoms.
- Life style records: diet, physical activity and so on.
- Family health history: diseases and health situations occurred to immediate family members.
- Patient notes: free texts beyond what can be expressed by information above.
- Related metadata: the date and time of observation, the range of symptoms, the devices used in measuring and so on.

Some information will be more useful if recorded with related metadata. For instance, blood glucose can be recorded combine with time, such as Fasting Plasma Glucose (FPG) recording in the morning before breakfast, Post Prandial Glucose (PPG) recording two hours after eating. The mood can be divided into depressed mood, anxious mood or others. The severity of mood, pain or other symptoms can be recorded as a range from 0 to 5 for instance.

2.3.7 Conceptual Scenario of Health Information Sharing

The figure below depicts a conceptual scenario that a patient self-manages himself at home, records and shares his health information to a hospital and third party healthcare service providers or research institution for the purposes of monitoring, receiving healthcare services, improving treatment and contributing to researches.

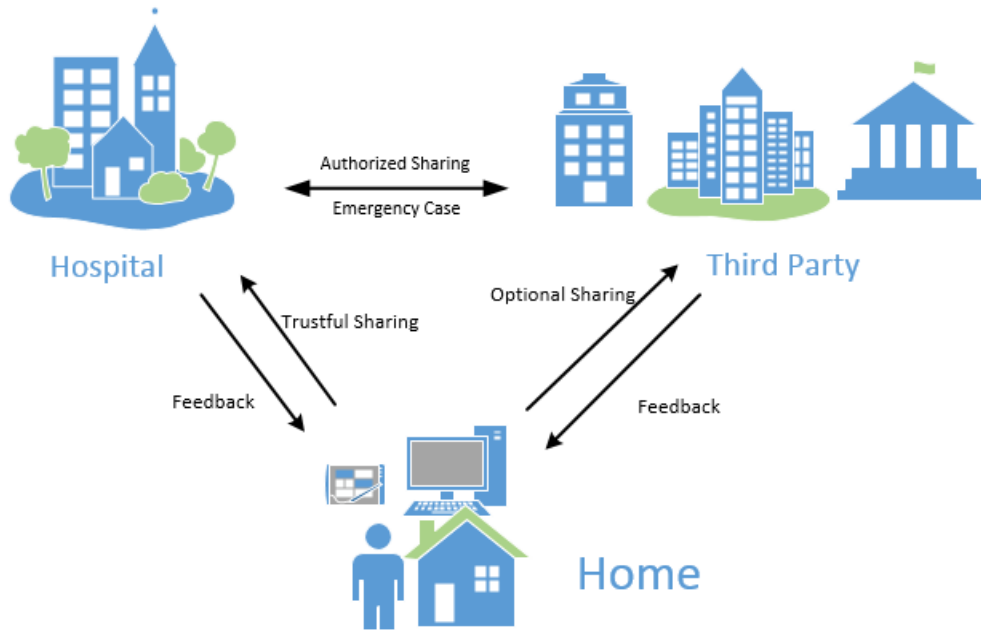


Figure 2 Conceptual Scenario of Sharing Health Information

Share to Hospital

When sharing to the patient's doctor, the more comprehensive and detailed health information the better. And the doctor also needs a complete record including personal information from patient. There is no need to consider which part to be included or not. But the process of treatment probably involve nurses or other medical personnel, they do not need to get the whole information about the patient. Except emergency cases, it's better to make the health information with fine-grained data access control mechanism to provide patient-centric privacy control over their health information. In addition, the shared health information is also valuable to be integrated into electronic records in hospital.

Share to Third Parties

Privacy and security are patients' greatest concerns [28]. When the sensitive personal health information is shared to the third parties who were not fully trusted, the main concern will become whether the patients have the actual rights of controlling their personal health information when being accessed or shared. There was a serious incident [58] of health information leakage. An employee stole the sensitive personal health information of 26.5 million military veterans from the database of U.S. Department of Veterans Affairs without authorization, which contains their social security numbers and health problems. Patients claim full control over their health data by only revealing non sensitive information when sharing.

To enable patient's sensitive health information being appropriately shared for various purposes, the implementation and management of disclosure policies should originate from the patient's willingness, the law or an organization [33]. There are different patient consent models [59]:

- No consent: health information automatically included and patients cannot opt out.
- Opt-out: health information is included in default, patients are able to opt health information out completely.

- Opt-out with exceptions: health information is included in default, patients are able to opt health information out completely or only allow selected information to be included.
- Opt-in: patients must actively express consent to choose health information to be included, once if they opted in, then their health information will be all included.
- Opt-in with restrictions: no patient health information is included in default, but patients are able to allow selected data to be included

Personal information is inappropriate to be shared to a third party, so the identity can be instead of a tracking ID. Since there is no need to share the whole health information to the third parties usually, it will be adequate that just make the shared health information meet the requirement of specific purpose. And a survey shown that 59% patients prefer the opt-in consent model [33]. In short, patient should be given the right to choose what information to share and whom to share. The opt-in with restrictions therefore is the appropriate consent model in the conceptual health information scenario.

Share between Hospital and Third Parties

There have been a lot of researches on the information exchange between healthcare organizations already. When the health information needed to be shared to another organization, it is necessary to inform the patient with reasons. Then the action of sharing should be authorized by the patient before the transferring of health data. In addition, a third party organization can ask the hospital for a patient's health record in emergency case. Every time when a health record is shared in emergency case, it needs to be recorded for verification.

2.4 Cloud Computing in eHealth

Several benefits will be brought by adopting cloud computing in the current healthcare ecosystem, such as health information management system, laboratory information system, pharmacy information system and so on [51]. Researches have proposed that cloud could also stimulate new forms of healthcare.

2.4.1 Benefits of Adopting Cloud Computing in Sharing Health Information

The process of the long-term monitoring will generate massive samples of patient self-recording health data. Consider the huge number of patients, it will generate a big volume of data. Due to the high demand of workload for uploading and sharing these huge amount patient dairy health data, the traditional server-client computing system is no longer suitable. Cloud computing provides a better way to achieve sharing these data to hospital and third party research or healthcare organizations.

According to the definition by NIST, cloud computing is “a model can provide distributed, rapidly provisioned and configurable computing resources (such as servers, storage, applications, networks and other services), which are on-demand, rapid elastic and measured, to whom have network connections” [6]. Usually, there are three layers of service model, which are Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS).

Because of the obvious scalability, flexibility and availability at relatively low cost of cloud services, there is a rapid trend of adopting cloud computing in enterprises and health-related area since last few years [7]. In the eHealth area, the PHR providers are increasingly willing to migrate their PHR services into cloud for the sake of lowering costs and improving the performance of services [60].

Due to the huge amount patient dairy health data, cloud storage service provides a better way to achieve the sharing of these data among hospital and third party research or healthcare organizations. When the data becomes available in the cloud, it can be processed by remote services, such as medical systems in hospital, CDSS, expert systems, or distributed to other medical personnel [9]. These features give more reasons to adopt cloud computing in sharing health information.

2.4.2 Practical solutions or Research Projects with Cloud Computing

Indivo [30] is an open source, open standards Personally Controlled Health Record (PCHR) with open APIs. It provides a web based interface, and is built to public. Indivo's implementation of the PCHR concept focuses on complete transparency and high security. The software of Indivo allows other institutions to create and administer PCHR infrastructure which meets the requirements of the HIPAA Privacy and Security Rules.

HealthVault is a PHR system from Microsoft, and hosted on the Windows Azure Platform public cloud [29]. It is a web based platform to store and maintain health and fitness information. There are a lot of specific third-party applications cooperate with this platform, such as for diabetes or lab tests. The security feature of HealthVault is similar or stronger than the requirements of the HIPAA Privacy and Security Rules [61]. Microsoft uses database-level encryption and many other security features to address the security problems.

A cloud based solution for collecting patients' vital data is proposed in [9]. It automates the collecting process by sensors attached to legacy medical equipment which are inter-connected to exchange service.

An ongoing research project Data Capture and Auto Identification Reference (DACAR) aims to implement a large scale and secure cloud service platform to gather and process health data [15]. It is deployed in a private cloud [61]. A prototype had been implemented with components of single point of contact for authorization, data buckets for reliable storage, and information sharing policy for access control. But DACAR did not support legacy medical records which are not based on atomic attributes then. Health Cloud Exchange (HCX) [8] is a distributed and modular system, allows patients to share CCR based EHRs and use healthcare services within a federated private cloud.

Besides personal health record system, cloud based hospital information system is also announced [33]. The company Meddserve [62] in the UK provides cloud computing service MeddCloud to hospital. The MeddCloud hosts the company's clinic applications to provide software as services.

Other related research projects include a Smart Clinical Decision Support System (Smart CDSS) [12]. It gathers data from diverse modalities, such as sensors, user profile information, social media, clinical knowledge bases, and medical experts within a hybrid cloud architecture to generate standards-based personalized recommendations.

2.5 Cloud Computing Deployment Models

Cloud systems can be divided into four categories from the view of deployment model: public cloud, community cloud, private cloud and hybrid cloud [6]. Each deployment model has different organizational structure and different provisioning location. It results in differences on utilization style, security level and other characteristics offered by these four deployment models. There are many crucial requirements on sharing health information among groups with different credibility, such as privacy, accessibility and utility. Adopting an inappropriate cloud deployment model will bring problems.

2.5.1 Public Cloud

The infrastructure or services of public cloud are open provisioned for public use. The definition of public cloud is the closest one to the traditional sense of cloud computing. As defined by NIST, “A public cloud is usually provided, managed, and operated by an enterprise, academic organization, government agency, or some combination of these. The cloud service providers make resources, such as computing capability, storage, application and other resources, available to the general customers through the connections of Internet.” Public cloud services are usually offered by a pay-per-usage charging model, or free in some cases. [6]

Public cloud services have been widely used by many companies today. There are service providers like Amazon, Google and Microsoft who provide infrastructures and various applications as services, which can be accessed via the Internet by customers. Or dependent cloud service provider such as Dropbox, who aggregates cloud infrastructure from Amazon to provide application and storage.

The main benefits of adopting a public cloud include [6], [63], [64]:

- Automatic deployment: customers don't have to worry about complicated and time-consuming server deployment when using public cloud.
- Inexpensive setup: costs of hardware, application and bandwidth are covered by public cloud providers.
- Ultimate scalability: the resources of public cloud are available on demand, the vast pools of resources can respond seamlessly to peak workload by expanding virtual infrastructure to customer.
- Cost effective: public cloud services offer a pay-per-usage charging model, customers only need to pay for what they've used. And the centralized management of hardware and software resources cost less on the same capabilities.
- Reliability: public cloud providers will perform data redundancy and setup active-backup servers in datacenters. Big providers will deploy multiple datacenters in different places for disaster tolerance. There will be no single point of failure and the services can keep running when suffering natural disaster.
- Flexibility: there are different layers of service provided by public cloud, such as IaaS, PaaS and SaaS. Various services of each layer are plug-and-play for many kinds of requirements, such as database service, message queue service and media transcoding service.
- Powerful management: virtual servers are obviously easier to manage compared with physical servers. Public cloud providers always provide detailed monitoring on all the services.

- Long term maintenance: customers don't need to worry about time consuming and costly server maintenance and update of software and hardware when using services from public cloud.

There are so many benefits to adopt public cloud. It seems appropriate to share health information through a public cloud. However, there have been some incidents aroused on public cloud services. Privacy information in the cloud were exposed with famous enterprises including Google, Facebook and Twitter in 2010 [65]. This incident made people think that if it is safe enough to store sensitive data outside in public cloud.

Disadvantages of adopting a public cloud include [54], [6], [63], [64]:

- Data lock-in: once customers deployed their services or business to a public cloud, their programs and data will be locked in the cloud. In case the provider increases the price or changes service agreement, it will not easily for customers to extract their programs and data from one provider to another.
- Security issue: once the data is stored in a public cloud, the actual control of data is out of customers. Famous cloud servers are often the targets of malicious attacks. Although the encryption is performed, the danger of data exposure is still a worry
- Privacy concern: privacy always concerns people to put data in online servers. The identity and access management provided by public cloud providers may not meet the high requirements in some services such as healthcare and financial.
- Customization: although public cloud providers deliver a set of services, it's impossible to satisfy every special requirement from customers. Customers have no control of what services they need.
- Accessibility: the base of public cloud is the Internet connection. Once some problems occur to the Internet connection between customers and public cloud, customers are unable to access to all the services.
- Long term reliability: when customers deployed their services or business to the public cloud, the control of keep running of services is under the cloud provider. If the public cloud provider goes bankruptcy or other problems occur, which will stop the cloud services from running, customers will suffer the loss.

2.5.2 Private Cloud

Private cloud is designed to offer the same features of public cloud, at the same time removes a number of objections including control over enterprise and customer data, worry about security, and issues relevant to regulatory compliance [63]. According to the definition of NIST, a private cloud is "the cloud infrastructure or services privately provisioned for exclusive using, managing and operating by a single organization. It is usually owned, managed, and operated by the organization, sometimes provided by a third party private cloud provider". [6]

Different to public cloud services, private cloud is hosted behind a corporate firewall. Therefore it gives private cloud the same security level as the organization's internal local area network, and makes it capable of delivering an organization's in-house cloud computing services remotely. And the network access is limited to a certain authorized group of users. A private cloud offers the potential to achieve better security over cloud based assets [66].

Private cloud is considered more beneficial than public cloud by some researchers. Because there are many advantages for adopting a private cloud [63][10][6]:

- Security: a private cloud is hosted behind a firewall, it gives more control over the data, and only offers access to a limited group of users, therefore external security concerns will be much less.
- Privacy: data sharing only in a limited environment, customized privacy policy can be implemented for special requirements.
- Customization: although public cloud providers offer various services, they cannot meet every requirement. Private cloud gives the capability to customize services for particular demands.
- Support legacy resources: some legacy software and hardware can be utilized as well under a private cloud by modification. Public cloud is barely possible to support legacy resources.
- Service Level Agreement (SLA): private cloud is behind a firewall other than a remote datacenter hosted by others, so the SLA will be stable when authorized personnel using services from a private cloud.
- No data lock-in: private cloud is hosted by an organization itself, the resources from infrastructure to data is controlled by the organization.

It seems that private cloud is more beneficial than public cloud in highly regulated industries. But adopting a private cloud requires significant level of IT ability, and high engagement degree of business environment virtualizing. There are some important challenges, which private cloud customers may have to face, includes [6] [63]:

- Costs on initial setup: either cost on financial or human resource is very high. The deployment of a private cloud requires constructing hardware and software, configuring networks, organizing security policies and so on. The process of setup is also time-consuming.
- Scalability: since a private cloud is operated by one organization, the scalability will not be large. The computing capability cannot compare with a public cloud under the same costs.
- Long-term maintenance: it requires a specific team to maintain a private cloud instead of passing the responsibility to a third party cloud provider. The later expanding is also need to be operated by the organization self.
- Potential risk for data loss by natural disaster: unlike public cloud, a private cloud usually located in one place, once a fire or natural disaster happened, the data cannot be restored.

2.5.3 Community Cloud

According to the definition of NIST, community cloud is “the cloud infrastructure or services provisioned for exclusive using, managing and operating by a specific community of organizations that have shared concerns. It is usually owned, managed, and operated by one or more organizations in that community, sometimes provided by a third party provider”. [6]

In this deployment model, cloud infrastructure is shared among several organizations with the same policy and compliance considerations. It allows multiple independent entities to gain the benefits of a shared nonpublic cloud while avoiding security and regulatory concerns which might not be addressed when using a public cloud [65]. It also helps to reduce costs further as it is shared by a larger group when comparing with a private cloud.

A community cloud is similar to a public cloud only with limited users. The benefits of deploying a community cloud include [6][63]:

- Shared data: the data will be easily shared with involved organizations, because all the data is in the same cloud. And there is no concern about the data being out of control.
- Security: community cloud is usually behind a firewall, the access is limited to a group of users which ensures that there are fewer external security concerns.
- Customization: similar to private cloud, a community cloud gives the capability to customize services for special demands.
- Relatively low cost: the cost of setting up and maintaining a community cloud are spread over the involved organizations. It's lower than a private cloud but higher than a public cloud.

Although community cloud owns some advantages of public cloud and private cloud, there are barriers for adopting a community cloud in some cases, it includes [6][63]:

- Hard to set up: because of the involving of different organizations, it may be hard to organize investment and management.
- Long-term maintenance: the organizations will in essence be self-serving. Maintenance requires a specific team to maintain the community cloud, the responsibility will be evaded.
- Accessibility: the involved organizations are probably located in different places, once some problems occur to the Internet connection between one organization and the community cloud, this organization is unable to access to all the services.
- Higher potential risk for data loss due to natural disaster: a community cloud is mostly located in one datacenter, once a fire or natural disaster happened, the data cannot be restored.
- Accountability: incidents such as data leakage will happen with a higher possibility since multiple members use the same cloud. And once some problems occur, it is also hard to be held accountable.

2.5.4 Hybrid Cloud

A hybrid cloud is a cloud computing environment which an organization utilizes and manages both internal cloud resources and external cloud resources. The NIST defines hybrid cloud as “the cloud infrastructure and services are a composition of two or more distinct cloud infrastructures and services which are different entities, but are combined together by techniques”. [6]

This deployment model can take advantage of cost efficiency and scalability offered by public cloud without hosting critical applications and data on a third party public cloud. Hybrid cloud can also be used for handling cloud bursting. For instance, when the existing private cloud is not able to handle workload overflow, it can require a fallback option to support the peak load. Therefore, the workload migration between public and private cloud will be transparent to users. [67][68]

By having some combined benefits of public cloud and private cloud, hybrid clouds are thought as an effective solution in some cases. The benefits of deploying hybrid cloud include [69][70]:

- Security: sensitive data and services can be kept and executed in private cloud behind a firewall to offer access to a limited group of people.
- Effective scalability: insensitive data and services can be migrated to public cloud, which offers a pay-per-usage charging model and available on demand.
- Customization: hybrid cloud still keeps the capability to customize services for particular demands.
- Expandability: uncritical utilizations can be expanded to public cloud ultimately when in a peak workload.
- Relatively low cost: part of the cost is reduced by combining public cloud with private cloud. The total cost will be lower than a pure private cloud for the same capacity.

The combining of public cloud and private cloud not just brings benefits. There are still barriers for deploying a hybrid cloud, it includes [69][70][10]:

- Integration: the network connection between a public cloud and a private cloud is complex. And to make the services work together with two clouds is not easy.
- Management: the management over separating sensitive data and services from public cloud, and migrating uncritical missions to public cloud at the same time is obviously complicated.
- Higher potential risk for data loss due to natural disaster: the critical data and services are still kept in a private cloud located in one place. Once a fire or natural disaster happened, a hybrid cloud will suffer almost the same damage to a pure private cloud.

2.6 Cloud Service Platforms

There are a lot of cloud services providers available around the world, such as public cloud providers like Amazon, Rackspace, Microsoft and so on, open source cloud platforms like OpenStack, Eucalyptus, CloudStack and so on. The top-1 public cloud provider and open source platform will be introduced in this section.

2.6.1 Amazon Web Services

Amazon Web Services (AWS) is a cloud computing platform, which offers a collection of remote computing services, provided by Amazon [71]. AWS is one of the earliest and most widely adopted public cloud service providers.

The most common services provided by AWS include [71] [72] [73]:

- Elastic Compute Cloud (EC2): EC2 provides the ability to run multiple virtual servers on demand, scale computing resources up and down flexibly. The environment of EC2 is built on top of the open source Xen hypervisor. The service offers APIs to start and stop server instances, apply access and networking permissions, and manage server images.
- Simple Storage Service (S3): S3 is scalable, distributed and reliable object storage system which offers secure online storage space for any kind of data. It makes the data accessible to any other authorized applications or individuals from anywhere on the Web. Its storage API is deliberately simple and offers complete control over how the data is represented in service. For addressing latency, minimizing costs, or

regulatory requirements, AWS provides an option to allow customers to choose the regions for the storage of data. It assures that data will not leave the customer selected regions. The current available choices include: US Standard, US West (Oregon), US West (Northern California), EU (Ireland), Asia Pacific (Singapore), Asia Pacific (Tokyo), Asia Pacific (Sydney), South America (Sao Paulo), and GovCloud (US) Regions.

- Simple Queue Service (SQS): SQS provides the ability of distributed web application to communicate reliably without having to build or maintain messaging system. It is implemented as a distributed application within Amazon to delivery short messages between any computers or systems which access to the Internet. The SQS API allows sending and receiving messages and controlling the full life cycle.
- Simple DB (SDB): SDB provides scalable, indexed, non-maintenance storage of small piece textual information in a simple database structure. It can replace traditional relational database when the application relies on a relatively simple database. And the data will be stored securely, redundantly and safely in SDB.
- Identity and Access Management (IAM): IAM enables customer of AWS to centrally manage users, security credentials such as access keys and x.509 security certificate, and permissions of controlling which AWS resources users can access. With IAM, customer can control what data in the AWS system users can access and how they access it. In addition, customer can restrict users' AWS access based on organizational groups permissions, and make sure users access AWS resources only from corporate network based on networking controls.
- Virtual Private Cloud (VPC): VPC enables users to launch AWS resources into a virtual network that the users have defined. It isolates from other virtual networks in the AWS cloud logically. Users can select the VPC's IP address range, create subnets, and configure route tables, network gateways, and security settings. VPC can be accessed from the Internet using SSH or Remote Desktop.

The figure below shows the architecture of a web search application, which is powered by some common services of AWS previously mentioned.

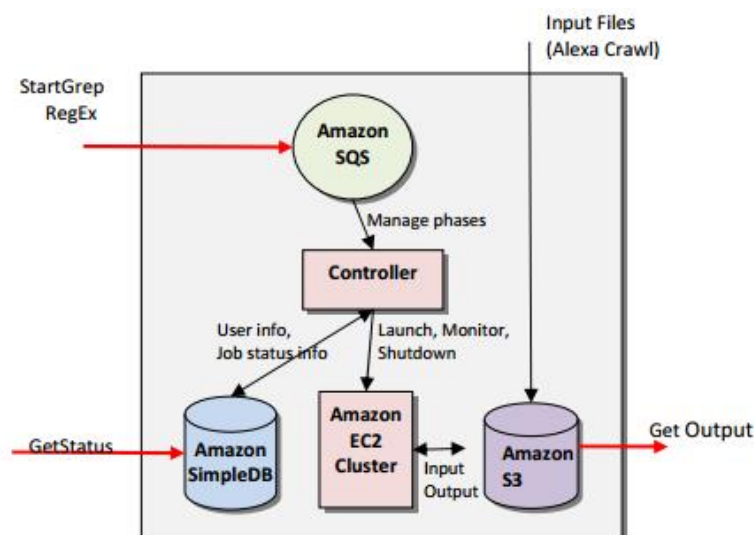


Figure 3 An example of AWS utilization [71]

2.6.2 OpenStack

OpenStack is a cloud computing platform for public and private clouds, which is free and open source released under the Apache 2.0 License. It is jointly launched by Rackspace and NASA in 2010. OpenStack provides APIs, command line interface and dashboard to control computing, storage, and networking resources pools. It enables users to utilize this platform integrating with legacy systems and third party technologies to deliver and manage large scale virtual machines networks and storage to offer on-demand and ubiquitous computing resources. And all the computing resources are accessible for developers to build cloud applications via APIs. [74] [75]

The figure below shows a simple structure of OpenStack.

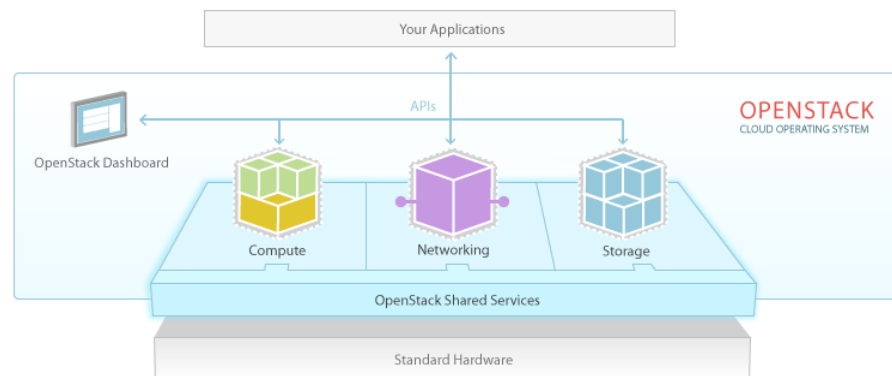


Figure 4 OpenStack Overview [74]

As shown in the figure, OpenStack is composed of different components, it includes: [79] [75]

- **Compute (Nova):** Nova is designed to provide, manage and automate pools of virtual machine resources and it can work with widely available virtualization technologies. The hypervisor technology of Nova so far can be chosen between Xen and KVM.
- **Object Storage (Swift):** Swift is a distributed, redundant and scalable object storage service which uses standard server clusters to provide storage for petabyte level of data. It can be directly integrated into applications by APIs, or used for data archiving and backups.
- **Block Storage (Cinder):** Cinder allows Nova instances to expand storage by connecting to block devices. It is appropriate to be utilized in the application scenarios where require high performance, for instances, database storage, expandable file system, and raw block level storage.
- **Networking (Neutron):** Neutron is used for managing networks and IP address in OpenStack. It provides flexible networking models such as flat networks or Virtual Local Area Networks (VLAN) for the requirements of different applications and user groups. Additional network services, such as firewall, load balancing or Virtual Private Networks (VPN), are also allowed to be deployed and managed by Neutron.
- **Identity Service (Keystone):** Keystone is used to centrally manage users which the OpenStack services they can access. It supports standard username and password credentials for authentication, as well as token-based systems or similar authentication way of AWS.
- **Dashboard (Horizon):** Horizon provides a browser-based graphical interface for administrators and users to access, manage and monitor OpenStack resources.

- Image Service (Glance): Glance is used to manage and provide virtual disks and server images. The Glance provides REST APIs to pull information of disk images and performing the images to new servers. Glance allows uploading various formats of images, including: VHD, VDI, VMDK and Raw.

2.7 Selection of Cloud Deployment Model

In the four deployment models, each model has its own characteristics. Deploying a community cloud involves multiple organizations, it will bring problems in this application scenario. Third party organizations and hospital have different properties and duties in nature. They have their own different interests other than common shared mission and compliance. The administration and management will be complicated, and it will be confused of accountability and responsibility when problems occur. Therefore community cloud deployment model will not be considered in the selection.

Since patient health information is very sensitive, security and privacy will be patients' greatest concerns [76], especially when these sensitive information is shared through cloud to third parties, who are not fully trusted. Other factors such as scalability and capability, customization, costs of setup and maintenance, and legal issue are also critical when choosing cloud deployment model in this application scenario. For the other three cloud deployment models, public cloud, private cloud and hybrid cloud, the selection will be conducted through analyzing and comparing on the 5 mentioned factors.

2.7.1 Security and Privacy

Security and privacy of patient health information is one of the key factors to build the trust of patients when sharing their health information. If the involved stakeholders lack of trust due to perceived risks, it may affect patients' willingness to share necessary health information. [47] The cloud security is multifaceted and layered [65]. Physical security such as natural disaster and human actions prevention, malicious attacks defense, encryption, identity and access management are all the potential factors which will affect the security and privacy protection of cloud.

As a public cloud is hosted by a third party enterprise, so the health data will be stored in the servers which are not fully trustable and out of control by either patients or hospital. It will increase the possibility of potential sensitive information leakage. The third party storage servers are often the targets of malicious attacks, it will be even worse due to the high value of the sensitive health information. Besides the inflexibility, the public cloud built-in mechanisms of encryption, identity and access management also cannot be fully trusted due to the same problems mentioned. In 2010, several cloud privacy information exposures occurred with a number of cloud based services, including Facebook, Twitter, and Google [65]. Because of these facts, patients are probably not willing to let their entire health data stored in public cloud servers.

Usually, a private cloud is hosted by the organization in-house. The health data will be stored in the servers which are operated by healthcare service providers. Although the disaster tolerance is not good as a public cloud, the control over the data is held in hand by healthcare service providers. Because of the firewall and invisible target, the threats of outside malicious attack will be much less. In addition, it is more flexible and reliable to implement fine-grained access control mechanism to protect the privacy of health data. Access control mechanism, such as XACML based policy [77], requires a secure and reliable server, which is more preferable to be hosted in a private cloud.

A hybrid cloud, which is combined by private cloud and public cloud, hosts sensitive data and workflow in its private cloud part, can provide the same security and privacy guarantee as private cloud.

Therefore, private cloud and hybrid cloud perform much better than pure public cloud in this important aspect.

2.7.2 Scalability and Capability

Another important reason for adopting cloud computing in health information sharing is the scalability and powerful capabilities.

The resources of public cloud are available on demand, the vast pools of resources can respond seamlessly to peak workload by expanding virtual infrastructure [6]. The broad network access of public cloud enables the easily sharing of insensitive health information, provides wide access to authorized healthcare research institutions and service providers to obtain the data they need. Due to the benefits of distributed massive computer clusters, ubiquity and virtualization, a public cloud can provide high performance services with low requirements on user-end computer infrastructure. And the capability of development platform, rather than economics, is an important factors to adopt a public cloud [12]. Third party eHealth service providers can develop various services, such as historical health analysis, health data mining [78] and CDSS [12], on a public cloud platform.

Theoretically, a private cloud can provide the same features as a public cloud. But it asks high investment. Because of the private cloud's own characteristics of smaller scale and limited access, it's unrealistic to make a private cloud to be a platform which needs to host various third party services [63]. This narrows the potential value of massive health information. And also the limited access will restrict the convenience of sharing insensitive health information.

If a hybrid cloud hosts shared health information and insensitive services in its public cloud part, it can provide the approximate scalability and capabilities to a public cloud.

Therefore, public cloud performs better than hybrid cloud, and hybrid cloud performs better than pure private cloud in this important aspect.

2.7.3 Customization

There are legacy systems used by hospitals, such as various clinical information systems include EMR, Radiology Information System (RIS) and Laboratory Information System (LIS), and other non-clinical information systems [51]. The deployed cloud model needs to be integrated into these systems without impact on the previous workflow. In addition, the specific requirements of healthcare data security and privacy ask for high demands on authentication, authorization and traceability.

Although public cloud providers deliver a set of various services, it is hardly to satisfy the special requirements from hospital legacy systems and workflow. When deployed a pure public cloud, the hospital can only utilize the services provided by public cloud providers. The expectations of customizing either on the services to meet the requirements of hospital, or the access control policies on patient data seem impossible.

A private cloud gives the capability to customize services for particular demands. The hospital can modify the cloud system to be coordinated with legacy systems, or even migrate the legacy systems to a private cloud. The key component for keeping confidentiality of patient data also can be customized to meet the requirements.

A hybrid cloud has the ability to keep the capability of customization, at the same time to migrate some appropriate services to the public cloud part. It makes a hybrid cloud even more flexible than a private cloud.

Therefore, private cloud performs better than public cloud, and hybrid cloud performs slightly better than pure private cloud in this important aspect.

2.7.4 Cost

Public cloud offers a pay-per-usage charging model, the hospital only needs to pay for what they have used. The initial setup costs of hardware, software and bandwidth are covered by public cloud providers. The costs of peak workload and prospected expanding will be more effective due to the elastic of public cloud. Moreover, the spending of time consuming and costly long-term maintenance, and the update of software and hardware, is no longer bottomless. These spending will be covered by public cloud providers as well.

The support of legacy software and hardware somehow will reduce part of expenditure on initial setup. However, the legacy hardware just cut probably a small piece of the total demand. The investment on new hardware and human resources will be very high. Moreover, the private cloud's limited resources do not qualify it to be a long-term solution [12]. The charging on expanding, updating and long-term maintenance will be a burden to the hospital.

Although a hybrid cloud combined capacity of private cloud with the on-demand capacity of public cloud, the investment of adopting a hybrid cloud is approximate to a private cloud [79]. But the migration of insensitive data and services to the public cloud part will reduce some burden compare with a pure private cloud.

Therefore, public cloud performs better than hybrid cloud, and hybrid cloud performs better than pure private cloud in this important aspect.

2.7.5 Legal

At present, a majority of big public cloud providers are from the US. It leads to the fact that most datacenters are located in the US. Because of the specialty of health data, it might cause some potential legal risks for other countries to deploy a public cloud in healthcare [69]. Different countries have different laws and regulations on managing patient data, some nations do not allow sensitive health data to be transferred out of border [68]. Although big public cloud providers, such as Amazon, may offer options to allow customers to choose the regions for the storage of data, the choices of region are still limited.

A private cloud will be hosted by hospital in-house, so there will be no concern about the violation of laws or other regulations.

As for a hybrid cloud, the sensitive data can be kept in the private cloud part while utilizing the public cloud part to process insensitive data.

Therefore, private cloud and hybrid cloud perform better than pure public cloud in the legal factor.

2.7.6 Summary

Based on the analysis and comparison of the 5 key factors, it can be summarized that public cloud is less appropriate than the other two clouds in 3 factors, which are security and privacy, customization and legal. Private cloud is less appropriate than the other two clouds in 2 factors, which are scalability and cost. By combining private cloud with public cloud, hybrid cloud seems to be the most appropriate cloud deployment model for the objective of sharing patient-generated health data to hospital and third parties at the same time. Hybrid cloud has 3 factors as good as private cloud, which are better than public cloud; and 2 factors better than private cloud, which are just lightly worse than public cloud.

More importantly, a pure public cloud is an inappropriate choice for the security and privacy concerns, as well as the potential restrictions from regulations of patient data. A pure private cloud may not be a high cost effective choice either. Besides the cost, the reason of restricting benefits from third party organizations makes a pure private cloud inappropriate.

The table below shows the quantitative comparison of the 3 kinds of cloud in 5 factors.

	Security and Privacy	Scalability and Capability	Customization	Cost	Legal
Public Cloud	1	2	1	2	1
Private Cloud	2	1	2	1	2
Hybrid Cloud	2	1.5	2	1.5	2

Table 2 Comparison of Clouds

For a clearer view of the quantitative comparison, a radar chart is drawn in below.

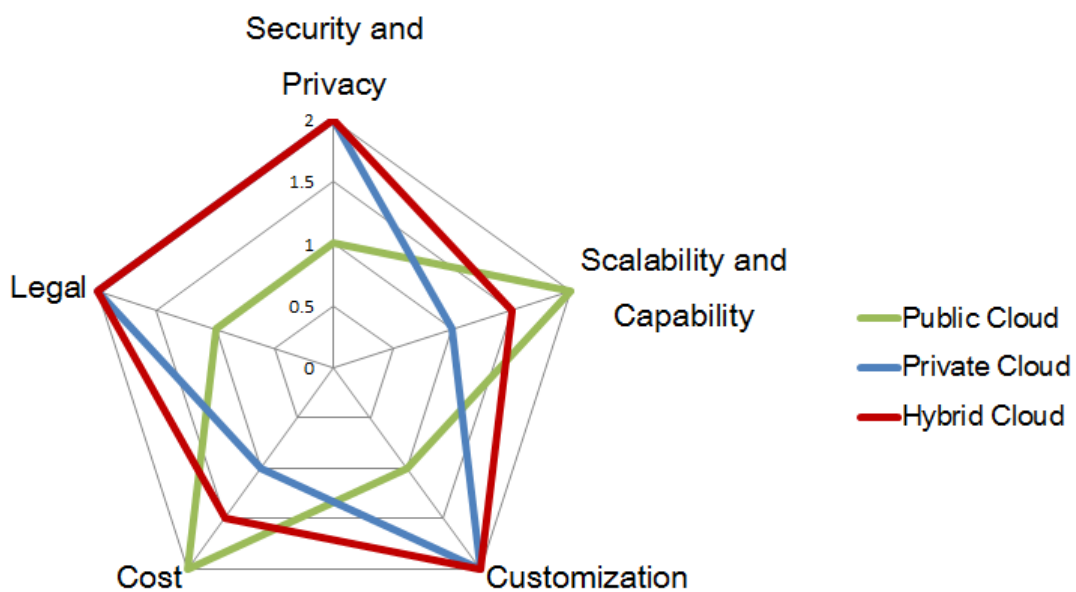


Figure 5 Radar Chart of Clouds

3 **PROTOTYPE DESIGN**

A prototype is defined as an early model or demonstration of a system which is built to test the feasibility of a concept, solution or process. It has been widely used in system development and research. [18] There are various types of prototype include [80]:

- Concept prototype: used to analyze system approaches when defining conceptual stages.
- Feasibility prototype: used to determine feasibility of various solutions to prove concept for specific issues.
- Horizontal prototype: used to clarify scope and requirements or demonstrate outer layer of human interface
- Vertical prototype: used to early test key components of a working, though incomplete, system.
- Functional storyboard: used to demonstrate the typical order in which information is presented.

In this chapter, the whole structure of the hybrid cloud platform will be designed and depicted for a clear demonstration of the expected purpose. Then an implemented vertical prototype will be presented, which is though incomplete, with the key functions of recording health information, sharing to private cloud and optional sharing to public cloud.

3.1 **System Model**

The model of this system consists of roles, objects, health records and procedures. They are described as followings.

- 1) **Roles:** A role in this system refers to an individual or an organization which receive benefits, provide resources or supply services [15]. This system involves different roles, such as patients, hospitals, third party research institutions and eHealth service providers. Each role has different degrees of credibility. The role of patients acts as the principal component to launch the main processes of this system. The role of hospitals locates in a private cloud by being considered as trustful. The role of third party research institutions and eHealth service providers locates in a public cloud by being considered as non-fully trustful.
- 2) **Objects:** An object refers to a person or an entity acts in this system. It includes patient, directly authorized hospital personnel, indirectly authorized hospital personnel, and service. A patient is an object of the patients role, who record and share his health data at home. A directly authorized hospital personnel object is the doctor who responsible for the patient. An indirectly authorized hospital personnel object is the other cooperative doctor or nurse who authorized by a directly authorized hospital personnel object under the permission of patient. In addition, each service entity either in private cloud or public cloud is regarded as an object. For instances, an EHR system service, which may integrate the patient-generated data, is an object in hospital's private cloud. As well as a data mining system, which analyze big data shared by patients, is an object provided by a third party organization in public cloud.
- 3) **Health Records:** A health record is the patient-recorded health data which transfer among different objects for different purposes under the authorization of patient object.

A patient's health record includes his personal information, vital physiologic measurements and other information recorded by patient. Since the patient has the right to choose what to share in his health record, the data in health record need to be lose-coupled and anonymous in some cases. And in case of emergency, the anonymous data should be able to be tracked to the entire version.

- 4) **Procedures:** Due to the different credibility degree of each role, there are different health record transferring procedures. The hospital is considered as trustful, so the health record sharing between a patient's home and a private cloud of hospital is a trustful sharing procedure. The whole health record will be shared in the trustful sharing procedure. A public cloud of third parties is considered as non-fully trustful, and some services just require a part of health record, so this sharing procedure should be treated differently. And which service entity objects in public cloud can get access to the shared health records are chosen by patient objects. Therefore the optional sharing procedure enables patients to share what and where they are willing to. Because of the relatively low cost and high capability offered by public cloud, it will be efficient that the hospital private cloud utilize the public cloud in non-critical applications. The non-critical utilization procedure expands the capability of a private cloud without involving sensitive personal information. In addition, there has to be an emergency case procedure, which can track to the certain patient in case of emergency cases.

Figure 6 shows the structure of this system.

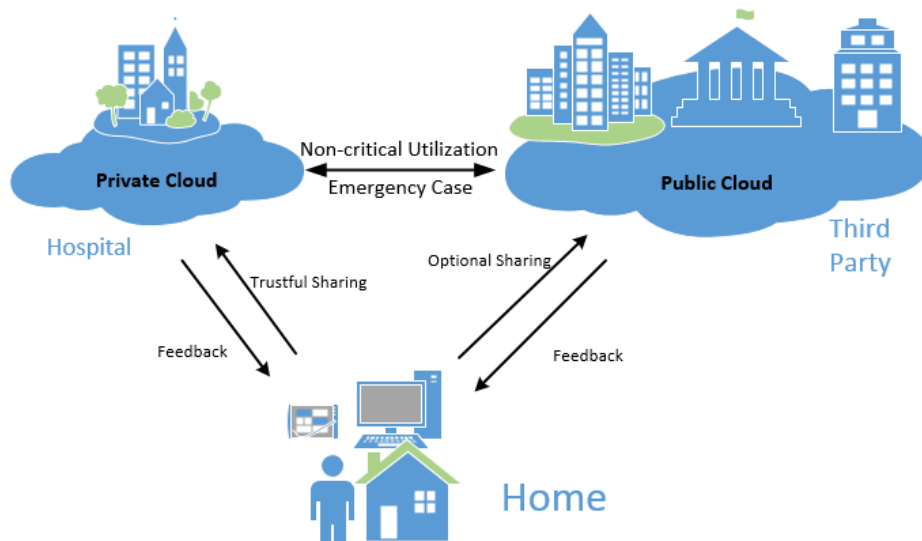


Figure 6 System Structure

3.2 System Design

3.2.1 Health Record Structure

Based on what has been discussed in previous sections, the structure of health record should achieve the features include:

- Patient opt-in with restrictions consent model.
- Lose-coupled and sectionalized data.
- Emergency case tracking ID.
- Interoperable data carrier.

Table 3 shows each section and content of the patient-recorded health record, which is inspired and referenced from the popular online eHealth service PatientsLikeMe and widely utilized health record standard CCD.

Record Section	Content	
Record target*	Personal information, Tracking ID	
Personal health status*	Other diagnosed diseases, smoke history, or other health related information	Related metadata such as the date and time of observation, the range of symptoms, the devices used in measuring
Physiologic measurements*	Blood Glucose, Blood Pressure, Pulse, and so on	
Medication dosage**	Amounts of injected insulin per day, and other medicines the patient is taking	
General health measurements	Weight, height, cholesterol or other related measurements	
Non-physiologic records**	Mood, pain, fatigue, insomnia, blurry vision or other related symptoms	
Life style records	Diet, physical activity and so on	
Family health history*	Diseases and health situations occurred to immediate families	
Patient notes	Free texts beyond what can be expressed by information above	

Table 3 Health Record Content

*The sections with * are referenced from the CCD template.*

*The sections with ** are referenced from PatientsLikeMe user charts.*

The Personal information of the Record target section will be removed for the purpose of privacy protecting when sharing to a public cloud. The Tracking ID is a randomly-generated global-unique ID which can be used to track the identity of patient in emergency case. Each section in the record can be opted-in by the willingness of patient when sharing to a public cloud.

The health record will be contained in an XML file, the format of which is inherited from HL7-CCD.

3.2.2 Conceptual Structure of Private Cloud Platform

The conceptual structure of private cloud platform in this system is briefly depicted in Figure 7.

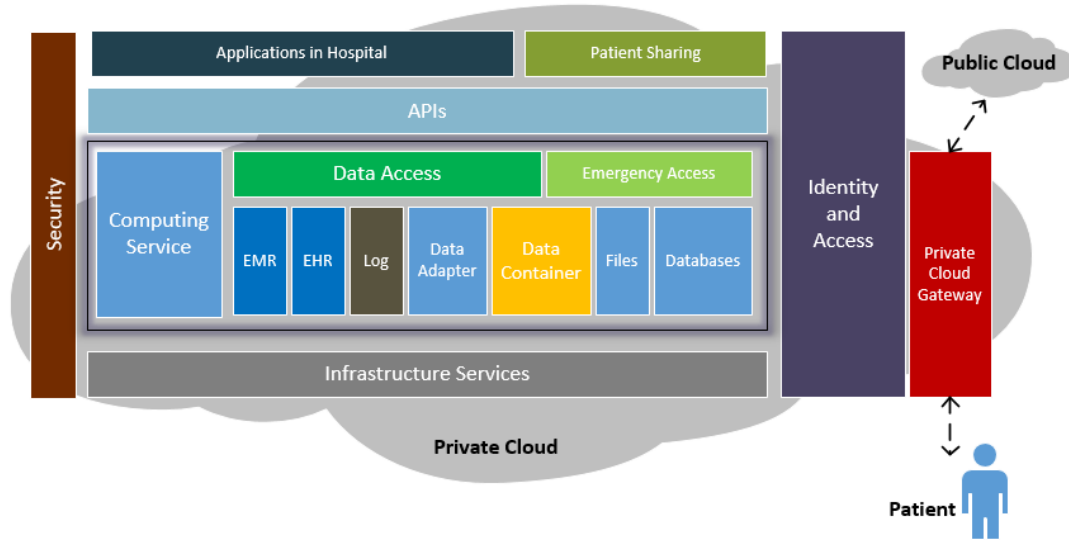


Figure 7 Conceptual Structure of Private Cloud Platform

The initial prototype of this private cloud platform is based on OpenStack Cloud System. By utilizing of private cloud infrastructure services, the health records shared by patients will be stored in assigned data containers of OpenStack Swift. The in-relation doctor can get the access to the shared health records in data containers after being authorized by the owners of these data. The access control can be achieved by OpenStack Keystone and other modifications. Other cooperative doctor or nurse can get access to part of or the whole health record through authorizing by the in-relation doctor under the permission of patient. Role-Based Access Control (RBAC) [77] and Attribute-Based Encryption (ABE) [28] are probably able to support this kind of fine-grained access control requirement. An EHR system can be hosted in this private cloud, therefore the shared health records will be integrated into EHRs. The data or images of previous medical examinations can be integrated into EMR and EHR, or sent to patients by requests as well. The computing service of the private cloud provides capability to operate healthcare applications such as clinical decision support system by utilizing patient's health information. When receiving emergency request, the emergency access component can get access to health records by tracking a unique tracking ID and then send back the required data. All the accesses to data and computing will be monitored and recorded by the log component. It can detect insecurity operations and analyze usage of resources [15].

3.2.3 Conceptual Structure of Public Cloud Platform

The conceptual structure of public cloud platform in this system is briefly depicted in Figure 8.

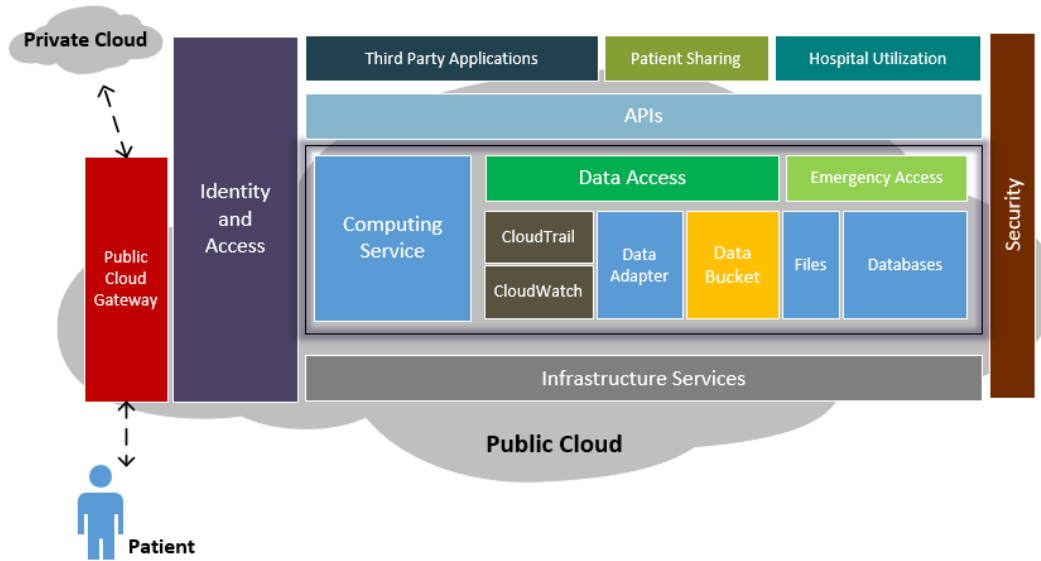


Figure 8 Conceptual Structure of Public Cloud Platform

The initial prototype of this public platform will be based on Amazon Web Services. By utilizing of public cloud infrastructure services, the health records shared by patients will be stored in specific data buckets of Amazon S3. The conceptual structure of public cloud platform gives the possibility to provide IaaS, PaaS or both to third party research institutions and eHealth service providers, which should be certificated as trusted organizations [15]. As they developed based on this platform, they can provide SaaS to patients for better healthcare, or gathering data for research.

The third party applications, either research programs or eHealth services, can get the access to the shared health records in different data buckets only when being authorized by the owners of these data. The access control can be achieved by AWS IAM and further modifications. For higher demand of access control, XACML-based mechanism [77] is a good choice. The computing service of public cloud provides high performance computing resources to these third party applications. The private cloud of hospital can expand or migrating non-critical computing tasks and insensitive data to this public cloud platform. The emergency access component can access to health records, send alarm to data owner and hospital by tracking the unique tracking ID once in emergency situation. The AWS CloudWatch service and CloudTrail service can be used to detect insecurity operations and analyze [15] by monitoring and recording all the resources usage and API calls.

3.3 Scenario of Conceptual System

The working scenario of this system will be described as two parts, which are the private cloud communication scenario and the public cloud communication scenario.

The scenario of the private cloud communication is shown in the figure below.

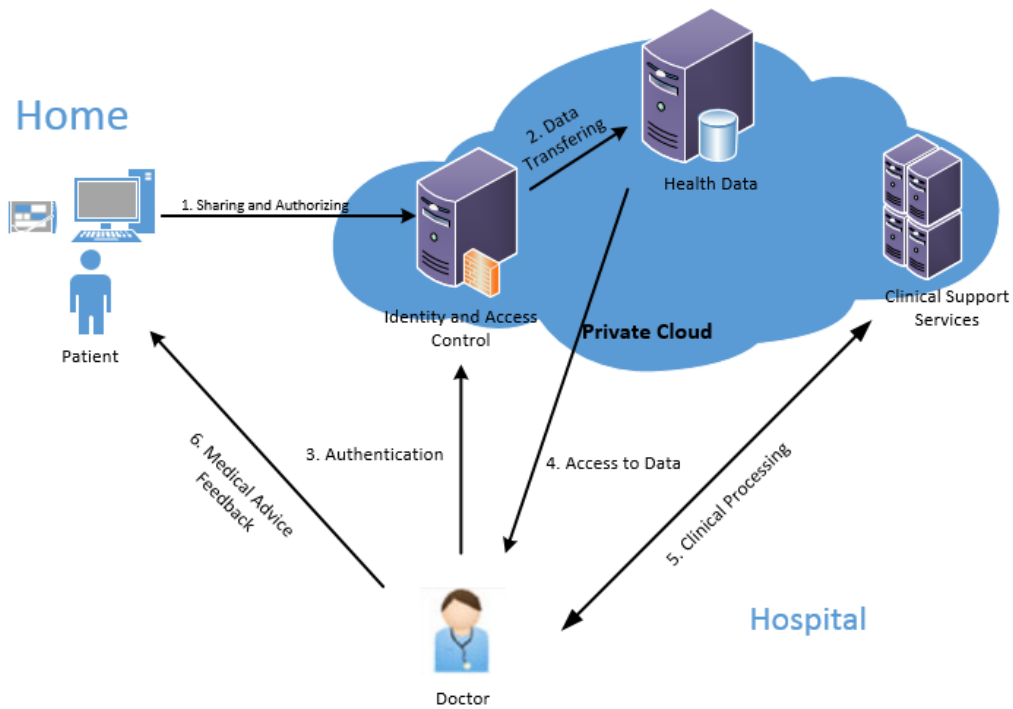


Figure 9 Scenario of Private Cloud Communication

A patient suffering from one or more chronic diseases self-manages and records health information at home. After inputting all the data into the health record, the patient chooses the target doctor in the hospital and shares to the private cloud. Then the health record will be transferred and stored in his container in Swift storage. When the doctor wants to get the access to the health record, he has to be authenticated that he has the right to access. When the doctor received the health record, he can send medical advice back to patient immediately or after the analyzing with clinical support services.

The scenario of the public cloud communication is shown in the figure below.

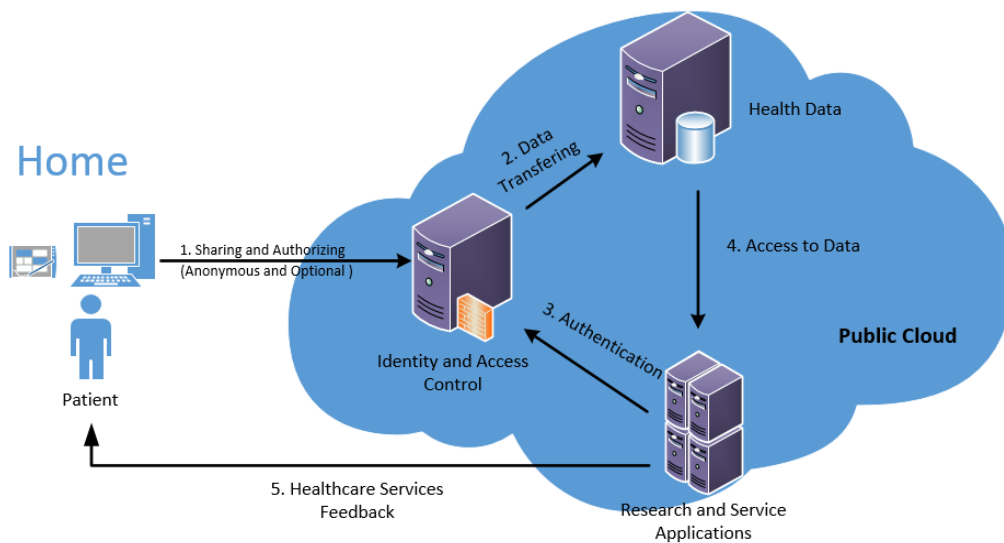


Figure 10 Scenario of Public Cloud Communication

After inputting all the data into the health record, the patient chooses a target application or research institution in the cloud platform, and then chooses which parts of data to share. When the sharing procedure is performed, the anonymous health record fragment will be transferred and stored in a particular bucket in AWS S3 storage. Any registered research or service application in the cloud platform want to get access to the health records, it has to be authenticated that it has the access right, which has been authorized when the patient was sharing. When the application received the health record, it can send certain results back to the patient. And research programs, such as clinical data mining, can utilize the authorized health data to perform researches.

3.4 Prototype Simulation

On the preliminary stage of the whole hybrid cloud platform, it is necessary to establish the connections among all the roles involved for exploration. Therefore an initial prototype is implemented, which is composed of a simple patient client, an OpenStack system with all components hosted in virtual machine, and Amazon Web Services. The patient client has been implemented with the functions of recording health information, sharing to private cloud and optional sharing to public cloud. The structure of health data inherits from HL7-CCD. The prototype is programmed by Java, with AWS SDK, open source libraries dom4j [81] for XML parsing and Joss [82] for OpenStack Swift API.

3.4.1 Health Data Recording and Sharing

Since this version of prototype is used to test the key functions of the incomplete working system, the interface of patient client is kept in simple. One of the interfaces is shown in Figure 11. This interface classifies some health record sections, which are defined in section 3.2.1. After inputting health data, patients can save the record in local as an CCD similar XML file, upload the record to private cloud, or press the “Sharing Options” button to share to public cloud, which will be performed in another interface.

Health Data Upload to Hospital

Vital Signs:
 Blood Glucose Level (mg/dL) Pulse (per min)
 Blood Pressure (mmHg larger/small)

Medication Dosage:
 Insulin (unit/0.01ml)

General Health Measurements:
 Weight (kg)

Non-Physiologic Records:
 Mood (ex: 2 happy) Pain (range 0-5)

Life Style Records:
 Physical Activity (quantity/activity) Diet (quantity/food)

Notes:

Date ex: Oct 14, 2013

Figure 11 Interface of Recording and Sharing to Private Cloud

The “Sharing Options” is shown in Figure 12. In this interface, patients can choose which health record sections to share based on their willingness. Patients can also choose where they want to share or what service they want to receive, such as diabetes research lab or clinical decision support system.

Share to Public Cloud

Choose what to share

- ☐ Personal Health Status
- ☒ Vital Signs
- ☒ Medication Dosage
- ☒ General Health Measurements
- ☐ Non-Physiologic Records
- ☒ Life Style Records
- ☒ Family Health History
- ☐ Patient Notes

Choose Where to share

- ☒ Diabetes Research Lab A
- ☐ Clinical Decision Support System
- ☐ Tendency Chart Application
- ☐ Health Status Application
- ☐ Others...

Share Cancel

Figure 12 Interface of Choosing and Sharing to Public Cloud

Figure 13 shows the workflow of sharing. When a patient decides to share to the private cloud, the health record will be uploaded to a container in OpenStack Swift, which belongs to this patient. In this prototype, the identification and access management of data container utilize the default Keystone without further customization. When patient decides to share to the public cloud, the chosen health record sections will be uploaded to a bucket in AWS S3, which belongs to the third party organization the patient chose. The identification and access management of data bucket utilize the default AWS IAM without further customization.

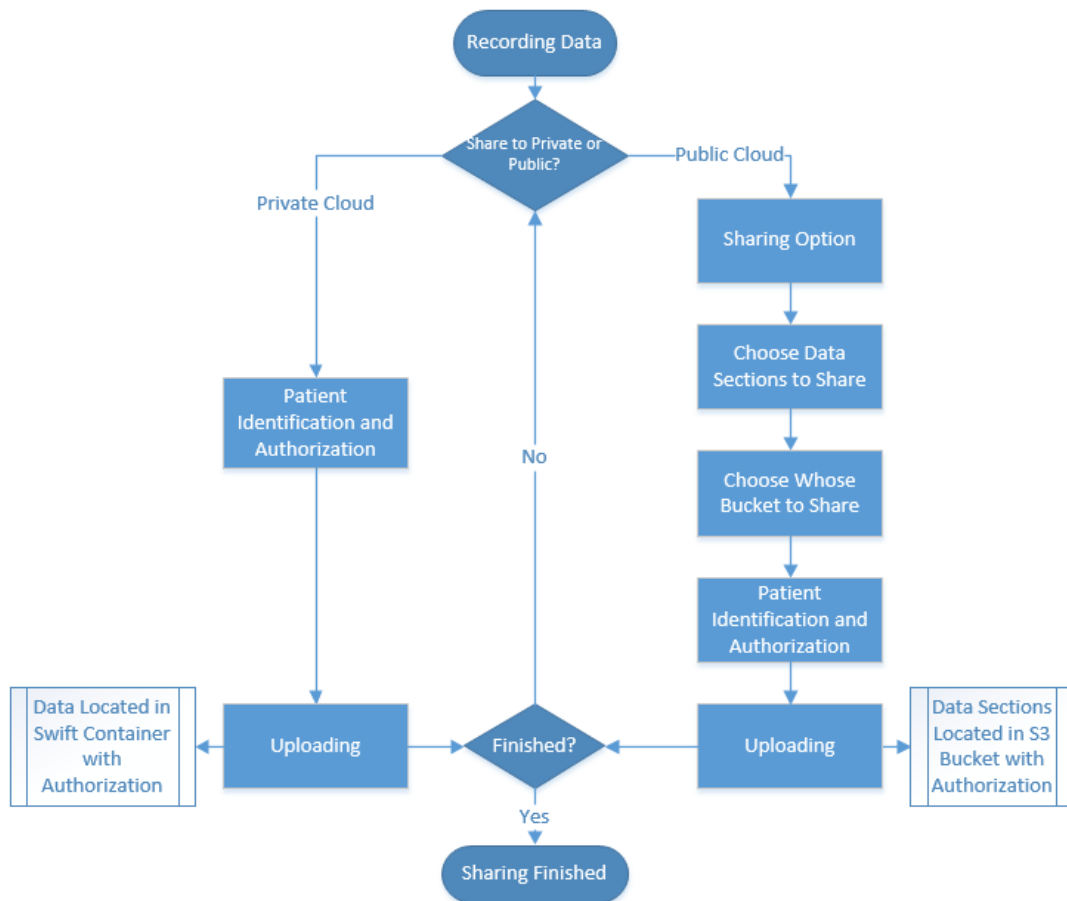


Figure 13 Workflow of Sharing

3.4.2 Shared Content Comparison

Figure 14 shows the Swift containers and objects in the Horizon graphical console interface. The content in the red box is a container on the private cloud. The name of container is the patient tracking ID, which is uniquely assigned to each patient. The content in the blue box is the file name of the uploaded health record.

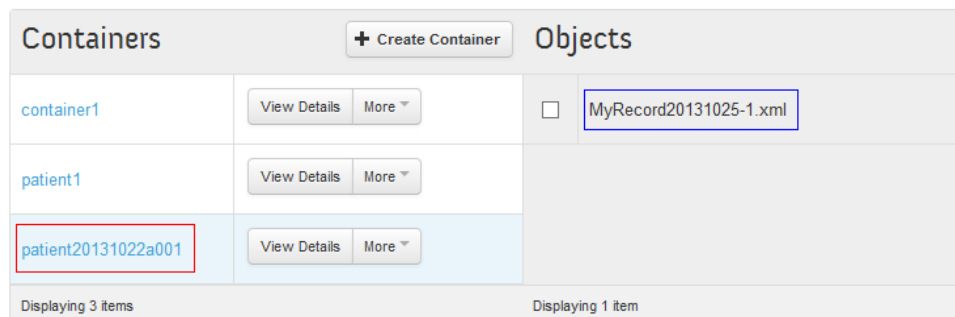


Figure 14 Swift Administration on Horizon Graphical Console Interface

The health record shared to the private cloud is complete and non-anonymous, it contains personal information such as name, gender and birthday, as well as patient tracking ID. An example is shown in the figure below.

```
<ClinicalDocument xmlns="urn:hl7-org:v3" xmlns:voc="urn:hl7-org:v3/voc" xmlns:x
  <typeId root="2.16.840.1.113883.1.3" extension="POCD_HD000040"/>
  <templateId root="2.16.840.1.113883.10.20.1"/>
  <!-- CCD v1.0 Templates Root -->
  <id root="db734647-fc99-424c-a864-7e3cda82e703"/>
  <code code="34133-9" codeSystem="2.16.840.1.113883.6.1" displayName="Summariz
  <title>Good Health Clinic Continuity of Care Document</title>
  <effectiveTime value="20000407130000+0500"/>
  <confidentialityCode code="N" codeSystem="2.16.840.1.113883.5.25"/>
  <languageCode code="en-US"/>
  <recordTarget>
    <patientRole>
      <id extension="996-756-495" root="2.16.840.1.113883.19.5"/>
      <patient>
        <name>
          <given>Cong</given>
          <family>Peng</family>
        </name>
        <administrativeGenderCode code="M" codeSystem="2.16.840.1.113883.5.1"/>
        <birthTime value="19880813"/>
      </patient>
      <patientTrackingID>20131025a001</patientTrackingID>
      <!--this patientTrackingID is used to track patient in emergency case-->
    </patientRole>
  </recordTarget>
```

Figure 15 Health Record Shared to Private Cloud

Figure 16 shows the AWS S3 bucket and objects on the AWS graphical console interface. The content in the red box is the bucket on the public cloud. The name of each bucket is the organization name or service name, and combined with the patient tracking ID. The content in the blue box is the file name of the anonymous optional shared health record sections.

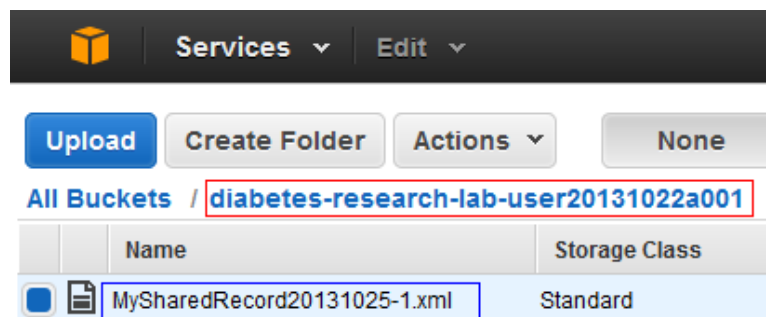


Figure 16 AWS S3 Administration on Horizon Graphical Console Interface

The health record sections, which optional shared to public cloud, are anonymous. It will not contain any personal information except the patient tracking ID. An example is shown in Figure 17.

```

<ClinicalDocument xmlns="urn:hl7-org:v3" xmlns:voc="urn:hl7-org:v3/voc" xmlns:x
  <typeId root="2.16.840.1.113883.1.3" extension="POCD_HD000040"/>
  <templateId root="2.16.840.1.113883.10.20.1"/>
  <!-- CCD v1.0 Templates Root -->
  <id root="db734647-fc99-424c-a864-7e3cda82e703"/>
  <code code="34133-9" codeSystem="2.16.840.1.113883.6.1" displayName="Summariz
  <title>Good Health Clinic Continuity of Care Document</title>
  <effectiveTime value="20000407130000+0500"/>
  <confidentialityCode code="N" codeSystem="2.16.840.1.113883.5.25"/>
  <languageCode code="en-US"/>
  <recordTarget>
    <patientRole>
      <patientTrackingID>20131025a001</patientTrackingID>
      <!--this patientTrackingID is used to track patient in emergency case-->
    </patientRole>
  </recordTarget>

```

Figure 17 Health Record Personal Information Shared to Public Cloud

The health record optional shared to public cloud is probably incomplete. It only contains the health record sections which were chosen by patient when sharing. An example is shown in Figure 18. The content in the red box is the section which the patient didn't share, and the content in the blue box is the section which the patient shared.

```

<componentPersonalHealthStatus>
  <!--Information is not shared.-->
</componentPersonalHealthStatus>
<componentVitalSigns>
  <section>
    <templateId root="2.16.840.1.113883.10.20.1.16"/>
    <!-- Vital signs section template -->
    <code code="8716-3" codeSystem="2.16.840.1.113883.6.1"/>
    <title>Vital Signs</title>
    <text>
      <table border="1" width="100%">
        <thead>
          <tr>
            <th align="right">Date / Time:</th>
            <th>Nov 14, 1999</th>
            <th>April 7, 2000</th>
            <th>Oct 22, 2013</th>
          </tr>
        </thead>
        <tbody>
          <tr>
            <th align="left">Blood Glucose</th>
            <td>260 mg/dL</td>
            <td>280 mg/dL</td>
            <td>280 mg/dL</td>
          </tr>

```

Figure 18 Health Record Optional Shared to Public Cloud

4 EMPIRICAL WORK

4.1 Interviews

Interview is a qualitative research method and can be used to collect data with better response rate and closer judgment of people's experience, opinion, desire and feelings [19]. Interview has many advantages. The flexibility of interview enables researcher to get more knowledge below the surface. The possibility of getting in-depth information from interviewees makes it differ from other data collection methods [83].

Interview can be conducted either in structured, unstructured or semi-structured ways [84]. A structured interview is composed of closed questions and appropriate to new interviewers. An unstructured interview will be composed of open questions, which allow interviewer to get extended answers. So combining closed questions and open questions composes a semi-structured interview.

This chapter will introduce the plan of interview, describe how the interview is organized, the selection of interviewees, and how the interview is conducted. The content of the analyzed expert opinions and patient opinion will be presented in the next chapter.

4.1.1 Purpose

The main purpose of this interview is to have detailed, practical understanding, and expert opinions from professionals on eHealth, cloud computing and healthcare to evaluate and improve the proposed solution and the prototype. The expectation and opinion from patient who suffered from chronic disease are also necessary to gather.

4.1.2 Design

For gathering anticipated information and exploring unforeseen and rich information from interviewees, a semi-structured interview will be conducted [20]. To grab comprehensive and detailed information, and to encourage interviewees to extend their answers, follow-up questions will be asked throughout the predetermined interview [83]. For the respect of confidential and privacy of interviewees, especially the patient, the questions need to be carefully formulated and organized.

Finally, 10 groups of questions are prepared for the interview of expert in eHealth and cloud computing; 10 groups of questions are prepared for the interview of expert in healthcare; and 11 groups of questions are prepared for the interview of patient. Most groups of questions are composed by a main question and one or more sub-questions. Each questionnaire contains closed questions and open questions.

4.1.3 Interviewee

The interviewees are planned as three groups. One group is familiar with eHealth and cloud computing, one group is familiar with healthcare, and one group suffered from chronic disease. After searching on relevant communities, social networks, and with the help of

supervisor, four interviewees were contacted. The introduction of interviewees shows in the tables below. For the respect of privacy, the name of patient will be instead of its initial.

Name	Roles	Organization	Perspective
Mark Braunstein	Professor of the Practice, College of Computing	Georgia Institute of Technology	eHealth, Cloud Computing
André Cavalleiro	Healthcare IT Manager	FIDI	eHealth, Cloud Computing
Marian Lenehan	Former nurse, Former nursing teacher	None	Healthcare
TED	Patient suffered from Type 2 Diabetes	None	Patient

Table 4 Interviewees

4.1.4 Recording and Transcribing

The data of interview was collected in different ways. For interviewees' clearer understanding and sufficient time of preparation, the description of thesis and the interview questions will be sent to interviewees before the interview. Dr. Mark Braunstein helped me improve the expression of the solutions description and questions for a more proper understanding. The three experts sent their opinions for the questions via written format at first. This is for their sufficient preparation and organized answers. For these qualitative questions, professionals need time to search information from different sources and organize their best answers. So the video call is considered as an inappropriate method for gathering detailed and reliable expert opinions. Then some unclear parts were discussed with each expert via video call or email. The interview with TED was conducted through video call, the data was recorded by taking notes and hints.

5 EVALUATION AND VALIDATION

In this chapter, the analyzed feedbacks from interviewed experts and patient will be presented to evaluate the proposed solution of this research. And then the research will be validated through the four criteria: credibility, transferability, dependability and confirmability.

5.1 Expert Opinions

5.1.1 Analysis of Feedback 1

Dr. Mark Braunstein is a professor of the practice, school of interactive computing in Georgia Institute of Technology. He teaches health informatics and is involved in research on its applications to the problems of healthcare delivery in the US.

According to Dr. Braunstein, the major barriers for adopting cloud computing in eHealth is discomfort with sensitive and protected health information being stored off site and out of the control of healthcare provider organizations. However, many contemporary Health IT tools (such as personal health records and even some newer electronic medical record systems) are cloud-based. The management of chronic disease has long been unsatisfactory because patients are outside of the purview of their healthcare providers except when they are in the clinic. This can lead to unnecessary visits which wastes time and resources or to visits that happen long after they were needed which can lead to the development of complications requiring expensive or even hospital-based care. Patients sharing self-management health data through the cloud brings a more continuous flow of information from patients to their care providers. It could potentially allow for a more continuous approach to care management avoiding both of these potential problems. But the barriers here is that, patients with chronic diseases face the challenge of complying with possible complex medication regimens and other instructions about their diet and exercise programs and recording of information about their compliance, clinical status and symptoms. And the technologies to support these activities have yet to achieve a level of simplicity and usability that is sufficient for their routine use by all patients.

In his opinion, the factor of security and privacy of the health information stored in the cloud will affect the selection of the cloud computing deployment model for supporting health information sharing most. This confirms the previous analysis for not selecting the public cloud model. He figured that many technologies exist today but none of them is completely secure. The current best approaches use the same encryption technologies that are used for other sensitive transactions (such as financial transactions) done over the Internet. Technologies to secure health information are available and have been developed for other Internet-based applications. Privacy is a much more difficult problem and one for which no really satisfactory solutions exist. Health data is complex and what is needed is a simple, transparent means for patients to indicate what data they wish to share and for what purposes and by whom it can be accesses. Ideally patients should be able to edit or even retract these permissions at any point in time.

The system model is considered as correct to Dr. Braunstein so long as it can handle multiple instances of each role. For example, a patient may be sharing data with more than one hospital, research institution or service provider and their sharing preferences may differ among the instances within each role. About the problem of patients can choose the data sections they want to share and share for specific purposes, he thought that so long as there is enough patient participation, the gaps that exist in individual patient datasets because of

reticence to share may not be a problem. However, if virtually all patients don't share particular data items, an incomplete picture of care may not be available for research or aggregate reporting purposes. It may be important to offer patients reasons why they should share any items they initially indicate they don't want to share or at least encourage them to consider sharing those items for specific purposes that might benefit them or other patients with their conditions. Although it wasn't be asked, he pointed out it is critical that patients are able to understand the terms and concepts offered for them to choose. This will be a critical part of the project to find a good solution to address.

Dr. Braunstein believes that open source technologies should be utilized wherever they can for providing more secure solutions. He thought this is an interesting and worthwhile project but the identified problems will require careful consideration and meaningful solutions. At the end, he pointed out that the actual patient facing interface will need to be much simpler and will need to offer the information and explanations. He suggested in both written and verbal (such as video) forms. He was not convinced that all patients can understand any of the terms used in the "Chose What to Share" column of choices in the patient interface. This is also a critical task needs to be addressed.

5.1.2 Analysis of Feedback 2

André Cavalleiro, with more than 15 years of experience in healthcare IT, used to be a Senior Project & Solution Manager at Siemens, and now works as Healthcare IT Manager at a non-profit organization, FIDI, which privately held and one of the biggest teleradiology service provider in Brazil. He is in charge of the IT operations over 54 units (which produces all types of medical images) and 4 reading centers in cloud.

According to André Cavalleiro, there are "natural barrier" against adopting cloud computing in health information exchange since it's a relatively new technology. He identified four factors in technical that may affect the hybrid cloud based solution in health information sharing:

- Interoperability: plenty of different systems supposed to exchange the same patient health information.
- Technology: distribution, access and utilization.
- Privacy & Security: protection against intrusion, unauthorized use, data corruption, damage, fraud and safeguard.
- Standardization: existing and new rules that will rule the act of sharing the patient health information.

He thought that some human factors will also affect the sharing of data. Some basic technical skills are necessary for the patient to start the process of sharing. And the patients who will self-manage and share their information, either to the private or public cloud, must be aware on how and when to do it.

On the issue of cloud computing deployment model selection, he mentioned that legal and technical issues may affect the selection of the cloud model. Since in the cloud, data is being archived somewhere by a business associate (e.g. cloud provider, healthcare facility) which is co-responsible for the privacy and security of the patient health information. How to get a "Business Associate Agreement" which clearly defines the shared responsibilities among the parties is not that easy. Lack of this Agreement may be accepted in developing countries/regions, but not in some other places, for instance in the US. On the other hand, technical issues (not legal ones) may be a barrier in third world or developing countries to adopt the model.

Besides data encryption when transferring, storing and archiving the PHI, Cavalleiro figured that products like Porticor can be used to reinforce security. Nevertheless, some of the cloud providers like Amazon Web Services are HIPAA and HITECH compliant. The transfer part of the process can be done through a VPN. He agreed with the 3 categories of roles in the proposed solution, and the locating each role in different place in the hybrid cloud seems good for dealing with privacy and security. To his opinion, what health data to share depends on what the other party is expecting to receive, and the use of the shared health data should be provided to patients. Cavalleiro is an AWS S3 and Glacier user. There are 500 million medical imaging examinations (or 25TB of data) archived in the cloud, and he is satisfied with their performance so far. Therefore the cloud provided an efficient way to handle health data.

5.1.3 Analysis of Feedback 3

Marian Lenehan used to be a nurse in a private hospital, and then taught nursing in the last ten years. Recently she found health informatics is fascinating, and started to pursue an MSc in health informatics.

Marian's father lived with type 1 diabetes for over 30 years and found it very difficult to manage. As far as she knew, type 1 diabetes was usually associated with leanness. More recently, obesity has given rise to additional cases of type 2 diabetes. This phenomenon seems to be associated with behavior change in relation to diet and exercise. Therefore, besides blood glucose level, lifestyle data is also valuable to share.

Marian knows a little bit of cloud computing, she agreed that information in the cloud seems to be the future direction for healthcare IT. However, she thought it will be a huge task while assuring the individual about the issues of privacy and security of their personal healthcare information. The proposed solution sounds like a good idea to her. The use of informatics in laboratory work is very crucial. She mentioned that she watched a TED (Technology, Entertainment, Design) talk where someone recommended printing the lab results for the patient rather than the doctor, in order to incorporate feedback and health advice. The attractive layout and additional information made the document very patient centric as it should be. The widely sharing of health data brings on the big data empowers the care as mentioned in the solution. As far as she knew, the person owns the data in the US, but the provider owns the media that produced it. In the UK the Department of Health owns the record. In her opinion, the concerns about data privacy and security might be obstacles in the proposed solution, as the same thing holding healthcare informatics back.

On the patient interface of the prototype system, she figured that lacking of medical knowledge might be a barrier for patient to use it. And data is getting bigger, it's hard to handle so much data. Automatic devices need to be utilized. All we can do is insist of secure platforms for storage and use. Therefore she thought the data sections and patient opt-in consent model is good for data privacy and security. Patients need to be guided with detailed explanation then they can understand each item. Since Marian hasn't been in hospital for a long time, so she is not sure how is the process of transferring medical record in an emergency case. And she thought health informatics needs to urgently address the need for coordinated care in relation to chronic diseases management. Based on her experience, it is worrying that the data recorded by patients may be unreliable. She thought automatic way will be better. And recently she found a wireless device for constant monitoring of blood glucose level.

At the end, she believes that mobile health may be a good place to go further. She had watched a TED talk about a mobile phone data collection programme. Interoperability is

another key in her opinion. She thought it's tragic that we are not taking full advantage of the interoperability of computer systems.

5.2 Opinions from Potential User

5.2.1 Analysis of Feedback

TED is a 52 years old retired car dealer with type 2 diabetes under control. His type 2 diabetes is diagnosed in 2005. He also lives with major depressive disorder and hypertension. He thinks he can be doing better and better.

In response to the question of his experience on diseases self-management, TED said he is using PatientsLikeMe to manage himself recently. He highly praised this kind of applications, it organized his treatment and symptoms, and the good thing is he had a chance to know what is working for others. In his opinion, the tests and the treatments, or some side effects probably are worth to record and share. On questions about his opinions to privacy issue, he said he is willing to share and trust the privacy agreement provided by big company, and he usually won't leave too much information about himself. When being asked questions about sharing via cloud computing, he expressed that he is not quite sure what cloud computing really is. But he thought sharing is fine, and got a little concern about if he could handle with that, because he was got confused when he used PatientsLikeMe at first. Therefore, according to TED's response and the popularity of PatientsLikeMe, patients are willing to share their health information and experience to help similar patients and also benefit themselves in return. Concerns are the privacy of data and their skills.

TED told that the proposed solution probably be very beneficial when being asked his opinions. He also mentioned that there might be some risks that the information will be misused by third parties. Anonymous is good, but he concerned the results might be delivered to a wrong person. And he thought the function of enabling share what and where they want to share is very important to his feeling. About the initial patient client interface, he mentioned that it will be better if every item and step be attached detailed explain. At the end, he expected automatic medical device could be worked with health information sharing. He figured that the participation of physicians and nurses is probably helpful, and boosts confidence for patients.

5.3 Discussion

5.3.1 Current Status of the Research Area

Patient health information sharing is always a critical problem either in normal healthcare or coordinated care, especially when dealing with chronic diseases. Most of the current healthcare systems are hospital centered model, in which doctors make treatment decisions, and patients barely participate in their own healthcare. Healthcare resource shortage is always a big problem when dealing with chronic diseases, especially in developing countries. There have been many researches trying to address the problem, such as the mentioned Smart CDSS [12], DACAR project [9] and web based data management application [4]. These researches have the potentials to reduce the burden. However, there are no essence changes, and the key factor, which is patients, did not play the vital role in managing their health as they should play. Some other proposed researches might have drawbacks such as confusing the roles in the system or less concern on security and privacy.

The status of aged population and growing chronic disease patients calls for a new patient centered model, which requires a continuing collaboration and connection among patients, different healthcare providers, communities and research institutions. It requires active participation of patients, and widely accessible patient information. The sharing of health information is the key of this kind of new healthcare model. The mentioned research of health self-management service [14] supports diabetes patients to track their blood glucose level. Although the provided services are relatively basic, it encourages patients to maintain their healthcare actively, and integrates with other PHR service to enhance the usability and interoperability.

The main obstacle of encouraging patients to share their health data is the privacy and security of their health data. The key in this process is giving patients the control over their own data. And regulations on patient data sharing need to keep up with the pace of development. The policy difference between nations is another barrier to make the health information sharing and care cooperation worldwide.

5.3.2 Contributions

The main contribution of this research is a potential hybrid cloud platform which has the abilities to address the problem of sharing long-term home monitored chronic disease patient health information, and to extend the value of the shared large volume health information to promote the treatment. Since chronic diseases will influence patients in a long last duration, many health conditions such as physiological signals need to be monitored, and most conditions of chronic diseases might be avoided by changing lifestyles and dietary, thus the continuous long-term monitoring, recording and sharing by patients is a better way to deal. The proposed system enables patients self-manage and share their health information to hospital private cloud at their homes, this makes chronic disease patients have independent livings and improves the quality of their lives.

In addition, the comprehensive long-term health data of patients has great potential benefits to provide further healthcare out of hospital and high value samples for researches. The proposed potential hybrid cloud platform enables patients to share their health information to a public cloud service platform, to receive various health related services, which are unlikely provided by hospitals, and to make contributions to researches on their diseases, which are probably beneficial to their healthcare in the future. Besides, the proposed system enables patients to record and share comprehensive health related data, and at the same time guarantees patients' greatest concerns, the security and privacy. By analyzing some regulations on protecting patient data either in Sweden or outside Sweden, one of the major ideas is obeying the willingness of patients. The proposed system therefore obeys the same idea to give patients choices. All the health information is anonymous when sharing to a public cloud service platform. Patients can choose the services or research organizations to share all by themselves. And the health record is split apart into sections, patients can only share the required sections to certain service program, or the sections they want. The interviewed experts agreed with the proposed sharing method as well.

Currently, due to the hardware limitation, inadequate skills and limited time, only a vertical prototype is implemented. The communication between the hospital and third parties is not yet established. But the key point of this research is to make the health information available from patients to healthcare providers, and there have been a lot of researches on the information exchange between healthcare organizations already. However, the key functions of recording and optional sharing from patient to private or public cloud have been achieved. As planned, the proposed hybrid cloud system has the potential of becoming a platform for

the purposes of enabling self-management of chronic disease patients, promoting the monitoring and treatment of chronic diseases, and contributing to healthcare researches.

5.3.3 Future Trends

An optimistic point is that this kind of cloud based solution is driving force for good interoperability and no patient data locked in one product. There seems to be a vast array of EMR, EHR and PHR products, which patients may be got locked in to one product and then find it is not compatible with a better or updated product from the previous system provider. In a cloud based solution like the proposed one, data will be contained in a widely identified international standard. And all the data will have an entire backup in private cloud and in patient client devices. It gives patients their own controls over their data. However, to make the changes of interoperability and transferability at an international level, it requires worldwide organizations to promote.

Either in accordance with the trend of development, the opinions from interviewees, or discussions with professionals in this area, automation and mobilization are the future of health information sharing and self-management. The more and more utilization of automatic devices for monitoring and recording will significantly reduce human errors and the impact from patients' lacking of medical knowledge. So many people around the world have mobile devices such as smart phone and tablet. These mobile devices enable patient to record, share and check feedbacks anywhere, and have the potential to discover a new variety of health data. Mobilization and cloud computing are compatible in nature.

5.4 Qualitative Validity

According to William M.K. Trochim [85], there are four criteria to validate a qualitative research, credibility, transferability, dependability and confirmability.

5.4.1 Credibility

Believable or credible from the perspective of the participant is a key criterion to the results of qualitative research, since the purpose of a qualitative research is to describe or understand the phenomena of interest from the participant's eyes [85].

After evolved from certain related literatures, the preliminary idea had been discussed with supervisor and senior to confirm the potential feasibility. All the concepts, definitions, statistics and other related knowledge in the process of theoretical work are extracted from published literatures, book chapters, official websites or other trusted sources, after synthesis and analysis. The selection of cloud deployment model and the design of prototype are based on the acquired knowledge, credible information and analysis. And the feasibility has been proved by the implemented vertical prototype. The positive opinions from the uninfluenced interviewees made the results more convinced. However, the results of patient interview are not well enough due to the limited respondents, unfavorable communication, and patient's privacy concern or insufficient knowledge.

5.4.2 Transferability

Generally, if a research cannot be transferred or referenced to other researches or applied in practice, the value of this research is relatively low. Transferability is primarily the responsibility of a qualitative research, it refers to the degree of the research results can be generalized or transferred to other contexts or settings [85]. According to William M.K. Trochim, clearly and comprehensively describing of the research context and the assumptions can improve transferability.

Although this research started with type 2 diabetes patient self-management information sharing, the method of data segmentation can be formulated to other chronic diseases. The research context and settings are clearly described from concepts to analysis. Therefore the modeling of system and the design of prototype can be easily generalized and evolved to other similar research contexts. The implementation of the vertical prototype is based on open source software and open API. The initial version of health record format is slightly modified from CCD, a widely accepted health record standard. It can be easily enhanced into practical application.

5.4.3 Dependability

Another criterion is dependability. The ever-changing context occurs in research is need to be accounted by researcher. Describing the changes occurred in the setting, and explaining how the changes will affect the research approach are the responsibilities to the researcher. [85]

The research was primary started with general chronic diseases self-management. Then it was found impossible during a limited period of time and also unnecessary to deal with so many different kinds of chronic diseases before long. Therefore the research turned to start with type 2 diabetes self-management information sharing. In the set application scenario, the information sharing of most chronic patient self-management are similar, so the impact of this changed context is very small. The host of OpenStack turned from physical machines to virtual machine is another change. Consider that this is an implementation of a prototype, and all the functions are trouble-free running, so this change of setting is dependable as well.

5.4.4 Confirmability

Confirmability is the degree of research results could be confirmed by others, it is under the assumption that unique perspective is brought by researcher to that study [85].

After deeply studied related literatures, analyzed the proposed solution, and tested the implemented prototype, then the idea of this research, the proposed solution, and the prototype were shared with interviewees, who were invited to evaluate and improve this research. Although a few questions might not be answered at great large due to some interviewees' limited time, the entire interview was well conducted. The interviewed experts praised the proposed solution, and provided many valuable advices for further research. In addition, the proposed solution was also agreed by other interview candidates, who had discussed with author about this research but had no enough time to be interviewed.

6 EPILOGUE

6.1 Conclusion

The objectives of this study are to address the problems brought by the purpose of sharing patient self-management health information to hospital and third parties for better treatment. The increasing chronic diseases care, which is costly and time consuming, is a major challenge in healthcare. Self-management is becoming a more and more important method for chronic patients to deal with their diseases by continuously monitoring and recording health vital signs and other related data. Therefore patient-recorded health information sharing is the key process to promote the treatment of chronic diseases, either sharing to hospital or other healthcare related organizations. This relatively new requirement brings challenges as well. One problem here is that, among the plenty of data generated in patient's daily life, what data is valuable to record and share, and what data is suitable to share to the non-fully trustful organizations.

The process of the long-term monitoring huge number of patients will generate a big volume of data. The traditional server-client computing system is no longer suitable to deal with the high demand workload of sharing these huge amount patient dairy health data. So the scalable and flexible cloud computing is recommended to handle this problem in this research. Another problem is the appropriate selection of cloud deployment model, since the sensitive health data is going to be shared with multiple stakeholders who have different properties and credibility.

For addressing these problems, author analyzed the information and knowledge gained from the process of literature review, then implemented a vertical prototype of the conceptual hybrid cloud solution, and performed an interview to gather expert opinions to evaluate and improve the proposed solution. As a result, the prototype, which is developed with OpenStack and AWS, successfully worked with the functions of recording health information, sharing to the private cloud and optional sharing to the public cloud. The way of health data segmentation and the proposed hybrid solution are agreed by interviewees. However, according to the analysis of interview and literature review, privacy and security, interoperability, regulations and medical knowledge of patients are the issues need to be addressed steadily.

Through researching particularly with type 2 diabetes, the proposed solution is capable of dealing with other chronic diseases just with slight modification, since the method of segmenting valuable and suitable health data is similar. Therefore the proposed hybrid cloud solution achieved the objective of sharing patient recorded health information in the application scenario of chronic disease self-management. In addition, the modeling of system and the design of prototype can be easily generalized and evolved to meet other similar purposes, since the utilized software, API and standard are open.

6.1.1 Answers to Research Questions

RQ1: What form of patient self-management health information sharing guarantees privacy and security when utilizing cloud computing among trusted and semi-trusted stakeholders?

By analyzing the information and knowledge gained from the process of literature review and some policies and standards on patient privacy protecting, it was identified that physiologic measurements such as blood glucose, blood pressure, medication dosage such as insulin injection, non-physiologic measurements such as mood and pain, lifestyle and so on

(Table 3) are valuable to record with metadata and share for long-term monitoring and better treatment. It was also found that the whole health record including personal information are suitable and should be shared to certain doctors in hospital, who are trusted, under the willing of patient. And it is inappropriate to make personal information appeared in the health record share to third party organizations who are semi-trusted. In the application scenario with cloud computing, there is no need and should not to share the whole record. To guarantee privacy and security of patient data in cloud among trusted and semi-trusted stakeholders, anonymous sharing and patient opt-in with restrictions consent model are the better choices. Patients have the right to control their health information, to share what they are willing to share for certain purposes. And it may be important to offer patients the reasons why they should share any items they initially indicate they don't want to share or encourage them to consider sharing those items for specific purposes that might benefit them or other patients.

RQ2: Which cloud deployment model is suitable for sharing patient-recorded health information among trusted and semi-trusted stakeholders?

After discussed the definitions, characteristics, advantages and disadvantages of the four cloud deployment models, the community cloud was considered as inappropriate. It is because the third party organizations and hospital have their own different interests other than common shared mission and compliance in this application scenario. Then the public cloud, private cloud and hybrid cloud were analyzed and compared on five factors (section 2.7.1 – 2.7.5). It was summarized that hybrid cloud is a better choice in this application scenario due to the combined characteristics of public cloud and private cloud. A hybrid cloud based prototype was implemented to prove the feasibility and suitability. In addition, the expert opinions from the interview also approved that the hybrid cloud deployment model is suitable for sharing patient-recorded health information among trusted and semi-trusted stakeholders

6.2 Future Work

The total plan of the proposed solution is too broad, it is nearly impossible for a master thesis work to complete the whole system. Although the prototype achieved the main functions of recording health information, sharing to private cloud and optional sharing to public cloud, there are still works remained to achieve an actual usable hybrid cloud based health information sharing solution. The traditional healthcare model has its limitations to deal with growing chronic diseases. An open platform with plenty of third party healthcare services will be a great complement to the existing healthcare systems for a better use of shared health data. For achieving an interoperable solution, common medical coding system standards and health record standards such as CDA or CCD are worth to implement in the solution for interoperable electronic health information exchange.

Patients own their health data, hence the control over their data should be given to patients. Ideally, patients should be able to edit or even retract these permissions at any point in time. The expectation of privacy protected in cloud requires a strict mechanism of authorization, authentication, and access control. Fine-grained access control might be a good choice to protect patient privacy. Either in the transferring process or storing, the encryption of health data is an essential feature to guarantee the security of data. Database level encryption seems a good way to ensure health data security in the cloud.

Health data is complex but what patients needed is a much simpler way. Expertise from chronic disease healthcare is therefore required to guide the process of health data segmentation, data object selection and health data recording. And it is critical to offer the information and explanations in both written and verbal forms for patients to understand the terms and concepts. Besides the works mentioned, more requirements and opinions are needed from chronic patients. Due to the time limitation, some dedicated experts in the related areas have not been interviewed.

Look to the future, the integrating of automatic health devices and personal mobile devices is a vital task of health information sharing and chronic disease self-management. The automatic way of health data collection is not only to facilitate patients to record massive data, but also to reduce errors due to lacking of medical knowledge. And the support of mobile devices not just releases patients from a restricted place, it also brings a flexible way of collecting data, or even a new variety of health data.

REFERENCES

- [1] G. Pare, M. Jaana, and C. Sicotte, "Systematic Review of Home Telemonitoring for Chronic Diseases: The Evidence Base," *J. Am. Med. Inform. Assoc. JAMIA*, vol. 14, no. 3, pp. 269–277, 2007.
- [2] C.-C. Lin, R.-G. Lee, and C.-C. Hsiao, "A pervasive health monitoring service system based on ubiquitous network technology," *Int. J. Med. Inf.*, vol. 77, no. 7, pp. 461–469, Jul. 2008.
- [3] C.-M. Chen, "Web-based remote human pulse monitoring system with intelligent data analysis for home health care," *Expert Syst. Appl.*, vol. 38, no. 3, pp. 2011–2019, Mar. 2011.
- [4] S. S. Deo, D. N. Deobagkar, and D. D. Deobagkar, "Design and development of a web-based application for diabetes patient data management," *Inform. Prim. Care*, vol. 13, no. 1, pp. 35–41, 2005.
- [5] M.-J. Huang, M.-Y. Chen, and S.-C. Lee, "Integrating data mining with case-based reasoning for chronic diseases prognosis and diagnosis," *Expert Syst. Appl.*, vol. 32, no. 3, pp. 856–867, Apr. 2007.
- [6] P. Mell and T. Grance, "The NIST Definition of Cloud Computing," National Institute of Standards and Technology, Special Publication 800-145, 2011.
- [7] S. Q. Ren and K. M. M. Aung, "PPDS: Privacy Preserved Data Sharing Scheme for Cloud Storage," *Int. J. Adv. Comput. Technol.*, vol. 4, no. 16, pp. 493–499, Sep. 2012.
- [8] S. Mohammed, D. Servos, and J. Fiaidhi, "HCX: A Distributed OSGi Based Web Interaction System for Sharing Health Records in the Cloud," 2010, pp. 102–107.
- [9] C. O. Rolim, F. L. Koch, C. B. Westphall, J. Werner, A. Fracalossi, and G. S. Salvador, "A Cloud Computing Solution for Patient's Data Collection in Health Care Institutions," in *2010 Second International Conference on eHealth, Telemedicine, and Social Medicine (ETELEMED)*, 10-16 Feb. 2010, 2010, pp. 95–9.
- [10] C. Ray and U. Ganguly, "An approach for data privacy in hybrid cloud environment," in *2011 2nd International Conference on Computer and Communication Technology (ICCT)*, 2011, pp. 316–320.
- [11] "Patientdatalag (2008:355)," 2008. [Online]. Available: <http://www.notisum.se/rnp/sls/lag/20080355.HTM>. [Accessed: 04-Oct-2013].
- [12] M. Hussain, A. M. Khattak, W. A. Khan, I. Fatima, M. B. Amin, Z. Pervez, R. Batool, M. A. Saleem, M. Afzal, M. Faheem, M. H. Saddiqi, S. Y. Lee, and K. Latif, "Cloud-based Smart CDSS for chronic diseases," *Health Technol.*, vol. 3, no. 2, pp. 153–175, Jun. 2013.
- [13] A. S. Radwan, A. A. Abdel-Hamid, and Y. Hanafy, "Cloud-based service for secure Electronic Medical Record exchange," in *2012 22nd International Conference on Computer Theory and Applications (ICCTA)*, 2012, pp. 94–103.
- [14] A. Sunyaev and D. Chorneyi, "Supporting chronic disease care quality: Design and implementation of a health service and its integration with electronic health records," *J Data Inf. Qual.*, vol. 3, no. 2, pp. 3:1–3:21, May 2012.
- [15] L. Fan, W. Buchanan, C. Thummler, O. Lo, A. Khedim, O. Uthmani, A. Lawson, and D. Bell, "DACAR Platform for eHealth Services Cloud," in *2011 IEEE International Conference on Cloud Computing (CLOUD)*, 2011, pp. 219–226.
- [16] Y. Levy and T. J. Ellis, "A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research," *Informing Sci. J.*, no. 9, pp. 181–212, 2006.
- [17] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering – A systematic literature review," *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, Jan. 2009.

- [18] A. Sarkissian, "Definition of Prototype Development | eHow," *eHow*, 18-May-2013. [Online]. Available: http://www.ehow.com/about_5103906_definition-prototype-development.html. [Accessed: 18-May-2013].
- [19] J. W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, Second Edition. SAGE, 2003.
- [20] Hove, S.E., Anda B., "Experiences from Conducting Semi -Structured Interviews in Empirical Software Engineerin g Research," presented at the Proceedings - 11th International Software Metrics Symposium, 2005, pp. 203–212.
- [21] G. Eysenbach, "What is e-health?," *J. Med. Internet Res.*, vol. 3, no. 2, Jun. 2001.
- [22] "eHealth," *Wikipedia, the free encyclopedia*. 04-Sep-2013.
- [23] "WHO | E-Health," *WHO*. [Online]. Available: <http://www.who.int/trade/glossary/story021/en/>. [Accessed: 16-Sep-2013].
- [24] "Defining Key Health Information Technology Terms," The National Alliance for Health Information Technology, Apr. 2008.
- [25] "Ehealth General," *WikEhealth*. [Online]. Available: <http://www.wikehealth.org/>. [Accessed: 11-Aug-2013].
- [26] J. A. Akinyele, C. U. Lehmann, M. D. Green, M. W. Pagano, Z. N. J. Peterson, and A. D. Rubin, "Self-Protecting Electronic Medical Records Using Attribute-Based Encryption."
- [27] M. Eichelberg, T. Aden, J. Riesmeier, A. Dogac, and G. B. Laleci, "A survey and analysis of Electronic Healthcare Record standards," *ACM Comput Surv*, vol. 37, no. 4, pp. 277–315, 2005.
- [28] C. Wang, X. Liu, and W. Li, "Implementing a Personal Health Record Cloud Platform Using Ciphertext-Policy Attribute-Based Encryption," in *2012 4th International Conference on Intelligent Networking and Collaborative Systems (INCoS)*, 2012, pp. 8–14.
- [29] "Microsoft HealthVault," *Microsoft HealthVault*. [Online]. Available: <http://www.HealthVault.com>. [Accessed: 22-Sep-2013].
- [30] B. Adida, A. Sanyal, S. Zabak, I. S. Kohane, and K. D. Mandl, "Indivo X: Developing a Fully Substitutable Personally Controlled Health Record Platform," *AMIA. Annu. Symp. Proc.*, vol. 2010, pp. 6–10, 2010.
- [31] S. Koch, "Home telehealth—Current state and future trends," *Int. J. Med. Inf.*, vol. 75, no. 8, pp. 565–576, Aug. 2006.
- [32] A. Darkins, P. Ryan, R. Kobb, L. Foster, E. Edmonson, B. Wakefield, and A. E. Lancaster, "Care Coordination/Home Telehealth: The Systematic Implementation of Health Informatics, Home Telehealth, and Disease Management to Support the Care of Veteran Patients with Chronic Conditions," *Telemed. E-Health*, vol. 14, no. 10, pp. 1118–1126, Dec. 2008.
- [33] M. L. Braunstein, *Health Informatics in the Cloud*. Springer New York, 2013.
- [34] "Interoperability," *Wikipedia, the free encyclopedia*. 21-Sep-2013.
- [35] M.E. Mansoor and R. Majeed, "Achieving Interoperability among Healthcare Organizations," Master Thesis, Blekinge Institute of Technology, 2010.
- [36] "Health Level Seven International." [Online]. Available: <http://www.hl7.org>. [Accessed: 23-Sep-2013].
- [37] S. Bhartiya and D. Mehrotra, "Exploring Interoperability Approaches and Challenges in Healthcare Data Exchange," in *Smart Health*, D. Zeng, C. C. Yang, V. S. Tseng, C. Xing, H. Chen, F.-Y. Wang, and X. Zheng, Eds. Springer Berlin Heidelberg, 2013, pp. 52–65.
- [38] E31 Committee, "Specification for Continuity of Care Record (CCR)," ASTM International, 2012.
- [39] "HL7/ASTM Implementation Guide for CDA® R2 -Continuity of Care Document (CCD®) Release 1." [Online]. Available: http://www.hl7.org/implement/standards/product_brief.cfm?product_id=6. [Accessed: 12-Aug-2013].

- [40] J. M. Ferranti, R. C. Musser, K. Kawamoto, and W. E. Hammond, "The Clinical Document Architecture and the Continuity of Care Record: A Critical Analysis," *J. Am. Med. Inform. Assoc. JAMIA*, vol. 13, no. 3, pp. 245–252, 2006.
- [41] "Continuity of Care Document (CCD): Changing The Landscape of Healthcare Information Exchange," Corepoint Health, White Paper, 2011.
- [42] "WHO | Chronic diseases," *WHO*. [Online]. Available: http://www.who.int/topics/chronic_diseases/en/. [Accessed: 21-Sep-2013].
- [43] G. Danaei, E. L. Ding, D. Mozaffarian, B. Taylor, J. Rehm, C. J. L. Murray, and M. Ezzati, "The preventable causes of death in the United States: comparative risk assessment of dietary, lifestyle, and metabolic risk factors," *PLoS Med.*, vol. 6, no. 4, p. e1000058, Apr. 2009.
- [44] M. Solomon, "Information technology to support self-management in chronic care: a systematic review," *Dis. Manag. Health Outcomes*, vol. 16, no. 6, pp. 391–401, 2008.
- [45] "World Health Statistics 2013 - A Wealth of Information on Global Public Health," World Health Organization, 2013.
- [46] C. Nordqvist, "All About Diabetes." [Online]. Available: <http://www.medicalnewstoday.com/info/diabetes/>. [Accessed: 08-Sep-2013].
- [47] The Office of the National Coordinator for HIT, "Guide to Privacy and Security of Health Information." .
- [48] M. M. Goldstein, "Data Segmentation in Electronic Health Information Exchange: Policy Considerations and analysis," Department of Health Policy, School of Public Health and Health Services, The George Washington University Medical Center, White Paper.
- [49] "Health Insurance Portability and Accountability Act," *Wikipedia, the free encyclopedia*. 13-Oct-2013.
- [50] "Integrating Privacy & Security Into Your Practice," *HealthIT.gov*. [Online]. Available: <http://www.healthit.gov/providers-professionals/ehr-privacy-security/practice-integration>. [Accessed: 10-Oct-2013].
- [51] A. Bahga and V. K. Madiseti, "A Cloud-based Approach for Interoperable Electronic Health Records (EHRs)," *IEEE J. Biomed. Health Inform.*, vol. 17, no. 5, pp. 894–906, 2013.
- [52] "Understanding Health Information Privacy," *U.S. Department of Health & Human Services*. [Online]. Available: <http://www.hhs.gov/ocr/privacy/hipaa/understanding/index.html>. [Accessed: 16-Oct-2013].
- [53] G. Gavrilov and V. Trajkovik, "Security and Privacy Issues and Requirements for Healthcare Cloud Computing," presented at the ICT Innovations 2012, Web proceedings, pp. 143–152.
- [54] R. Zhang and L. Liu, "Security Models and Requirements for Healthcare Application Clouds," in *2010 IEEE 3rd International Conference on Cloud Computing (CLOUD)*, 2010, pp. 268 –275.
- [55] "Health informatics -- Requirements for an electronic health record architecture," ISO/TC 215, ISO 18308:2011, Apr. 2011.
- [56] E. A. Brox, "Information Security in Distributed Health Information Systems in Scandinavia: A Comparative Study of External Conditions and Solutions for Exchange and Sharing of Sensitive Health Information in Denmark, Norway and Sweden," Norwegian University of Science and Technology, 2006.
- [57] 1177.se, "Patientjournalen." [Online]. Available: <http://www.1177.se/Blekinge/Regler-och-rattigheter/Patientjournalen/>. [Accessed: 13-Oct-2013].
- [58] J. Foreman, "At risk of exposure," *Los Angeles Times*, 26-Jun-2006.
- [59] J. Daniel, M. M. Goldstein, and A. L. Rein, "Consumer Consent Options for Electronic Health Information Exchange: Policy Considerations and Analysis," Office of the National Coordinator for Health IT, Washington, DC, White Paper, Mar. 2010.
- [60] M. Li, S. Yu, K. Ren, and W. Lou, "Securing Personal Health Records in Cloud Computing: Patient-Centric and Fine-Grained Data Access Control in Multi-owner

- Settings,” in *Security and Privacy in Communication Networks*, vol. 50, S. Jajodia and J. Zhou, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 89–106.
- [61] E. Ekonomou, L. Fan, W. Buchanan, and C. Thuemmler, “An integrated cloud-based healthcare infrastructure,” 2011, pp. 532–536.
- [62] “meddserve,” *Meddserve - Cloud health IT service provider*. [Online]. Available: <http://www.meddserve.com/>. [Accessed: 21-Oct-2013].
- [63] “Compare the Cloud.” [Online]. Available: <http://www.comparethecloud.net/>. [Accessed: 26-Sep-2013].
- [64] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, “A view of cloud computing,” *Commun ACM*, vol. 53, no. 4, pp. 50–58, 2010.
- [65] J. R. Winkler, *Securing the cloud cloud computer security techniques and tactics*. Burlington, MA: Elsevier, 2011.
- [66] H. Lühr, A.-R. Sadeghi, and M. Winandy, “Securing the e-health cloud,” in *Proceedings of the 1st ACM International Health Informatics Symposium*, New York, NY, USA, 2010, pp. 220–229.
- [67] “Hybrid Cloud,” *Searchcloudcomputings*. [Online]. Available: <http://searchcloudcomputing.techtarget.com/definition/hybrid-cloud>. [Accessed: 03-Oct-2013].
- [68] “The 4 Primary Cloud Deployment Models,” *CloudTweaks*. [Online]. Available: <http://www.cloudtweaks.com/2012/07/the-4-primary-cloud-deployment-models/>. [Accessed: 10-Oct-2013].
- [69] Z. Mahmood, “Data Location and Security Issues in Cloud Computing,” in *2011 International Conference on Emerging Intelligent Data and Web Technologies (EIDWT)*, 2011, pp. 49–54.
- [70] “Hybrid Cloud Guide,” *ComparetheCloud.net*. [Online]. Available: <http://www.comparethecloud.net/cloud-computing/cloud-deployment-models/hybrid-cloud/>. [Accessed: 03-Oct-2013].
- [71] “Amazon Web Services,” *Amazon Web Services*. [Online]. Available: <http://aws.amazon.com/>. [Accessed: 08-Oct-2013].
- [72] “Cloud computing with Amazon Web Services, Part 1: Introduction,” 29-Jul-2008. [Online]. Available: <http://www.ibm.com/developerworks/architecture/library/ar-cloudaws1/>. [Accessed: 08-Nov-2013].
- [73] J. Murty, *Programming Amazon Web Services - S3, EC2, SQS, FPS, and SimpleDB*. Farnham: O’Reilly, 2008.
- [74] “OpenStack,” *OpenStack Open Source Cloud Computing Software*. [Online]. Available: <http://www.openstack.org/>. [Accessed: 12-Oct-2013].
- [75] “OpenStack,” *Wikipedia, the free encyclopedia*. 30-Oct-2013.
- [76] M. Li, S. Yu, Y. Zheng, K. Ren, and W. Lou, “Scalable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute-Based Encryption,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 1, pp. 131–143, Jan. 2013.
- [77] A. A. El-Aziz and A. Kannan, “Access control for healthcare data using extended XACML-SRBAC model,” in *2012 International Conference on Computer Communication and Informatics (ICCCI)*, 2012, pp. 1–4.
- [78] H. Chen, S. Compton, and O. Hsiao, “DiabeticLink: A Health Big Data System for Patient Empowerment and Personalized Healthcare,” in *Smart Health*, D. Zeng, C. C. Yang, V. S. Tseng, C. Xing, H. Chen, F.-Y. Wang, and X. Zheng, Eds. Springer Berlin Heidelberg, 2013, pp. 71–83.
- [79] O. Mazhelis and P. Tyrväinen, “Economic aspects of hybrid cloud infrastructure: User organization perspective,” *Inf. Syst. Front.*, vol. 14, no. 4, pp. 845–869, Sep. 2012.
- [80] “Prototyping: Types of Prototypes,” *Toolbox.com*. [Online]. Available: <http://it.toolbox.com/blogs/enterprise-solutions/prototyping-types-of-prototypes-14927>. [Accessed: 03-Oct-2013].
- [81] “dom4j,” *DOM4J - Making Your Own Site*. [Online]. Available: <http://www.dom4j.org/>. [Accessed: 10-Sep-2013].

- [82]“joss,” *GitHub*. [Online]. Available: <https://github.com/jaswift/joss>. [Accessed: 15-Nov-2013].
- [83]I. Seidman, *Interviewing as qualitative research a guide for researchers in education and the social sciences*, 2nd ed. New York: Teachers College Press, 1998.
- [84]C. Dawson, *Projects in Computing and Information Systems: A Student's Guide*, Second Edition. Pearson Education Canada, 2009.
- [85]W. M. K. Trochim, “Qualitative Validity,” in *The Research Methods Knowledge Base*, 2nd Edition., 2006.

APPENDIX

A. Interview Transcripts

Interview with Mark Braunstein

About you

1. What are your roles and responsibilities in your work?

I teach health informatics and am involved in research on its applications to the problems of health care delivery in the US.

Interview Questions

1. What are the major barriers for adopting cloud computing in eHealth? How will they affect the effectiveness and usefulness of health information exchange?

The major barrier is discomfort with sensitive and protected health information being stored off site and out of the control of healthcare provider organizations. However, many contemporary health IT tools (such as personal health records and even some newer electronic medical record systems) are cloud based.

2. What benefits do you see from patients sharing self-management health data through the cloud? What challenges/barriers do you think there might be?

The management of chronic disease has long been unsatisfactory because patients are outside of the purview of their healthcare providers except when they are in the clinic. This can lead to unnecessary visits which wastes time and resources or to visits that happen long after they were needed which can lead to the development of complications requiring expensive or even hospital-based care. A more continuous flow of information from patients to their care providers could potentially allow for a more continuous approach to care management avoiding both of these potential problems.

Patients with chronic diseases face the challenge of complying with possible complex medication regimens and other instructions about their diet and exercise programs and recording of information about their compliance, clinical status and symptoms. The technologies to support these activities have yet to achieve a level of simplicity and usability that is sufficient for their routine use by all patients.

3. In your opinion, what will most affect the selection of the cloud computing model for supporting health information sharing?

Security and privacy of the health information stored in the cloud.

4. How can data be protected in the process of transferring and storing it in the cloud?

Many technologies exist today but none of them is completely secure. The current best approaches use the same encryption technologies that are used for other sensitive transactions (e.g. financial transactions) done over the Internet.

5. What methods are usually performed to ensure the privacy and security of data stored in the cloud? Can these mechanisms meet the requirements of protecting the privacy of health information contributed by patients in your opinion?

As stated in my prior answer, technologies to secure health information are available and have been developed for other Internet based applications. Privacy is a much more difficult problem and one for which no really satisfactory solutions exist. Health data is complex and what is needed is a simple, transparent means for patients to indicate what data they wish to share and for what purposes and by whom it can be accessed. Ideally patients should be able to edit or even retract these permissions at any point in time.

6. In the proposed prototype system, the roles divide into patients, hospitals, third party research institutions or health related service providers and eHealth service providers, operating within a local home client, a private cloud or a public cloud. Do you have any suggestions on this?

This model seems correct to me so long as it can handle multiple instances of each role. For example, a patient may be sharing data with more than one hospital, research institution or service provider and their sharing preferences may differ among the instances within each role.

7. To facilitate patient indication of the data they wish to share, health data is divided into sections. Patients can choose the sections they want to share or only data items within a section. Is there any potential problem when a lot of incomplete health data sections appear in the cloud?

You didn't ask but it is critical that patients are able to understand the terms and concepts offered for them to choose. So long as there is enough patient participation the gaps that exist in individual patient datasets because of reticence to share may not be a problem. However, if virtually all patients don't share particular data items an incomplete picture of care may not be available for research or aggregate reporting purposes. It may be important to offer patients reasons why they should share any items they initially indicate they don't want to share or at least encourage them to consider sharing those items for specific purposes that might benefit them or other patients with their conditions.

8. In the initial prototype system, the uploaded health data is stored in an object storage system, OpenStack Swift and AWS S3. What is your opinion about this, from the perspective of efficiency and operability for further utilization of data?

No opinion other than I believe open source technologies should be utilized wherever they can provide secure solutions.

9. What do you think about the proposed hybrid cloud system?

I think it is an interesting and worthwhile thesis project but one that will require careful consideration and meaningful solutions to the problems you pose and some of the issues I raise here.

10. Do you have any other suggestions to improve this prototype and the proposed system?

I assume the user interface you show in this document is a prototype. The actual patient facing interface will need to be much simpler and will need to offer the information and explanations I've suggested in both written and verbal (e.g. video) forms. For example, I'm not convinced that all patients can understand any of the terms used in the Chose What to Share column of choices.

Interview with André Cavalleiro

About you

1. What are your roles and responsibilities in your work?

Healthcare IT Manager at FIDI, a not for profit organization, privately held and one of the biggest teleradiology service provider in Brazil, performing about 3 million examinations per year. In this context I am in charge of the IT operations over 54 units (which produces all types of medical images) and 4 reading centers.

Interview Questions

1. What are the major barriers for adopting cloud computing in eHealth? How will they affect the effectiveness and usefulness of health information exchange?

Since cloud computing is relatively new, there is a “natural barrier” against its adoption. Technically speaking, and assuming that your Hybrid Cloud Based solution is a mix of SaaS and IaaS, I do see a couple of things that could affect the exchange of information:

- Interoperability – plenty of different systems supposed to exchange the same patient health information (PHI)
 - Technology – distribution, access and utilization
 - Privacy & Security – protection against intrusion, unauthorized use, data corruption, damage, fraud, safeguard
 - Standardization – existing and new rules that will rule the act of sharing the PHI
2. What benefits do you see from patients sharing self-management health data through the cloud? What challenges/barriers do you think there might be?

I would say that some human factors can negatively affect the sharing of data. The individuals that will be self-managing and sharing their information, either to the private or public cloud, must be aware on how and when to do it. Indeed, some basic (at least) technical skills are necessary for the patient to “pull the trigger” and start the process of sharing.

3. In your opinion, what will most affect the selection of the cloud computing model for supporting health information sharing?

In the cloud model data is being archived somewhere by a business associate (e.g. cloud provider, healthcare facility) which is co-responsible for the privacy and security of the patient health information. However, how to get a “Business Associate Agreement” which clearly defines the shared responsibilities among the parties is a factor. Lack of this Agreement may be accepted in developing countries/regions, but not in the US for instance. In the other hand, technical issues (not legal ones) may be a barrier in 3rd world or developing countries in order to adopt the model. So the legal and technical issues can affect the selection of the cloud model.

4. How can data be protected in the process of transferring and storing it in the cloud?

Encrypting the data seems to be the way to transfer and store/archive the PHI in a safe way.

5. What methods are usually performed to ensure the privacy and security of data stored in the cloud? Can these mechanisms meet the requirements of protecting the privacy of health information contributed by patients in your opinion?

Products like Porticor (<http://www.porticor.com/features/>) can be used to reinforce security, nevertheless some of the cloud storage providers like Amazon Web Services are HIPAA and HITECH compliant. The transfer part of the process can be done through a VPN.

6. In the proposed prototype system, the roles divide into patients, hospitals, third party research institutions or health related service providers and eHealth service providers, operating within a local home client, a private cloud or a public cloud. Do you have any suggestions on this?

In general the solution looks good. The locating of each role in different place in the hybrid cloud seems good for dealing with privacy and security.

7. To facilitate patient indication of the data they wish to share, health data is divided into sections. Patients can choose the sections they want to share or only data items within a section. Is there any potential problem when a lot of incomplete health data sections appear in the cloud?

It depends on what the other party is expecting to receive I guess and the use of the information is being provided.

8. In the initial prototype system, the uploaded health data is stored in an object storage system, OpenStack Swift and AWS S3. What is your opinion about this, from the perspective of efficiency and operability for further utilization of data?

I am an AWS S3 and Glacier user, and so far I have no complaints about it. I am heading to have 500 million medical imaging examinations archived in the cloud or 25TB of data. I am not familiar to OpenStack.

9. What do you think about the proposed hybrid cloud system?

At first sight looks good as I have said.

10. Do you have any other suggestions to improve this prototype and the proposed system?

It is need to mention that there is a need of using a XDS to manage the PHI.

Interview with Marian Lenehan

About you

1. Please briefly describe your experience in healthcare.

I used to be a nurse in a private hospital, and I taught nursing last few years. Recently I find Health Informatics is fascinating, and I start to pursue a MSc in Health Informatics.

Interview Questions

1. When a patient suffered chronic disease (or specifically diabetes type 2), what kind of health data is valuable and suitable to record and share?

Blood glucose level is valuable to record. My father had Type 1 Diabetes for over 30 years and found it very difficult to manage. As far as I know, Type 1 diabetes was usually associated with leanness. More recently, obesity has given rise to additional cases of Diabetes, often Type 2. This phenomenon seems to be associated with behavior change in relation to diet and exercise.

2. Have you heard about cloud computing? What challenges/barriers do you think there might be in sharing self-management health data through cloud computing?

I only know a little bit of cloud computing. I agree that information in the cloud seems to be the future direction for HIT. However, it sounds like a mammoth task to make that happen while assuring the individual about issues of privacy and security of their personal healthcare information.

3. What do you think of a home based solution that chronic disease patient can record and share health information to his doctor for long-term monitoring, upload to a third party platform for receiving health related service, and contribute to healthcare researches?

It sounds like a good idea. The use of informatics in laboratory work is so crucial. I watched a TED talk where someone recommended printing the lab results for the patient, not the doctor, in order to incorporate feedback and health advice. The attractive layout and additional information made the document very client centric - as it should be. As you mentioned in the solution, bring on the big data empowers the care. Healthcare technologies like these are changing patients' lives for the better by making things so much more accessible.

4. What challenges/barriers do you think there might be in the proposed system?

I think concerns about data privacy and security might be obstacles, as what is holding healthcare informatics back. In the US, I believe the person owns the data, but the

provider owns the media that produced it. In the UK the Department of Health owns the record.

5. According to the patient interface of the prototype system, what would be the barriers for patients to record their health information?

Lacking of medical knowledge might be a barrier. And data is getting bigger, we can't insist on controlling it any more. The lights are on automatic pilot, all we can do is insist of secure platforms for storage and use.

6. In the procedure of sharing to a public cloud, health data is divided into different sections to support a patient opt-in consent model. This facilitates patients having the ability to choose what they want to share when sharing data to a certain place. What do you think about this? Is there any potential problem when a lot of incomplete health data sections appear in cloud?

In general it looks good for data privacy and security.

7. In the phase of choosing what to share, do you think can patients understand the meaning of each option? If not, how to improve it?

I'm not sure. More explanation to items will be better.

8. When a patient encounter an emergency case in other place, how is the process of transferring medical record?

I haven't been in hospital for a long time. In your case, I don't know if there are any rules on that. In my opinion, it needs to urgently address the need for coordinated care in relation to chronic diseases management, particularly in the case of diabetes.

9. As your experience, do you have the confidence on the accuracy and reliability of health data recorded by patients themselves? If not, what is your opinion to improve this situation?

It is worrying that the data may be unreliable. Automatic way will be better. Recently I found a wireless device for constant monitoring of blood glucose level, this would be helpful.

10. Do you have any suggestions on this prototype and the proposed system?

I've watched a TED talk about an eMobile phone data collection programme. Mobile Health may be a good place to go further, as so many people around the world have this basic device. Interoperability is another key. It's tragic that we are not taking full advantage of the interoperability of computer systems.

Interview with TED

About you

1. Briefly describe your condition, please.

Diabetes Type 2 diagnosed in 2005, now is under control.

Interview Questions

1. Have you ever used any system for self-management, which you record your health data by yourself? Do you worry about the privacy issue when you use them? Simple or complicated to use them?

I'm using a website PatientsLikeMe recently. Probably be ok because I didn't leave much about myself. I would say it confused me at first, it's getting better and then.

2. Do you think sharing home-monitored health data beneficial or not? If beneficial, what kind of data is valuable and suitable to record and share?

Yes, of course I believe it's beneficial. It organizes my treatment and symptoms, the good thing is I have a chance to know what is working for others. I think my tests and the treatment, or some side effects are probably worth to share.

3. Have you heard about cloud computing? Will you willing to share self-management health data through cloud computing? What will be your concern about this? If not, what are the barriers for you to use?

I'm not quite sure what cloud computing really is. Sharing is fine. I have a little concern about if I could handle with that

4. What do you think of a home based solution that chronic disease patient can record and share health information to your doctor for long-term monitoring, upload to a third party platform for receiving health related service, and contribute to healthcare researches?

It sounds pretty good. My physician will be convenient to monitor.

5. What barriers do you think there might be for you in this proposed system? What should be noticed in this process?

I don't know. If gonna to say, probably be the information abuse by third party.

6. The health data shared to third parties will be anonymous, does this relief your worry about health information leakage? What is your suggestion for protecting your privacy?

Anonymous will be good. But, is there any risk of result misdelivery?

7. In the procedure of sharing to third party platform, health data is structured loose-coupled to perform opt-in consent model, which makes you have the ability to choose what you want to share. How do you think about this?

It's good, you give the choices to patient, good feelings.

8. According to the patient interface of the prototype system, what would be the barriers for you to record health information?

It looks good, might be a little simple. More detailed explanations of each item will be better to me.

9. In the phase of choosing what to share, can you understand the meaning of each option? If not, how could make you understand more clearly?

The same as above, more explanations

10. Besides what you know about this prototype system, what else do you expect from sharing your health information? What will make you willing to use a system like this?

It would be great if the process of recording data becomes automatic, like the medical meters.

11. Do you have any suggestions to improve this prototype and the proposed system?

If physicians or nurses work together with this system, it might be good, and gives confidence to me.