ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

A FRAMEWORK FOR A NATION-WIDE ELECTRONIC HEALTH VAULT WITH A SECURE MULTI-CLOUD HYBRID MODEL

M.Sc. THESIS

Halil Emre GÖNEN

Department of Electronics and Communication Engineering

Biomedical Engineering Programme

Thesis Advisor: Dr. Serkan TÜRKELİ

JUNE 2016

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

A FRAMEWORK FOR A NATION-WIDE ELECTRONIC HEALTH VAULT WITH A SECURE MULTI-CLOUD HYBRID MODEL

M.Sc. THESIS

Halil Emre GÖNEN
(504131408)

Department of Electronics and Communication Engineering

Biomedical Engineering Programme

Thesis Advisor: Dr. Serkan TÜRKELİ

JUNE 2016

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

GÜVENLİ HİBRİT ÇOKLU BULUT YÖNTEMİ İLE TASARLANMIŞ ÜLKE ÇAPINDA ELEKTRONİK SAĞLIK KASASI MODELİ

YÜKSEK LİSANS TEZİ

Halil Emre GÖNEN
(504131408)

Elektronik ve Haberleşme Mühendisliği Anabilim Dalı

Biyomedikal Mühendisliği Programı

Tez Danışmanı: Öğr. Gör. Dr. Serkan TÜRKELİ

HAZİRAN 2016

Halil Emre Gönen, a M.Sc. student of ITU Graduate School of Science Engineering and Technology student ID 504131408, successfully defended the thesis entitled "A Framework for a Nation-Wide Electronic Health Vault with a Secure Multi-Cloud Hybrid Model", which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Dr. Serkan TÜRKELİ** Istanbul Technical University	.....
Jury Members :	**Assoc. Prof Dr. Deniz Turgay ALTILAR** Istanbul Technical University	.....
	Assoc. Prof Dr. Serhat ÖZEKES Üsküdar University	.....

Date of Submission : 2 May 2016
Date of Defense : 8 June 2016

FOREWORD

First and foremost, I offer my sincerest gratitude to my advisor, Dr. Serkan TÜRKELİ, who has guided me and presented me the knowledge to complete this thesis while always believing in me and never giving up on me even though I thought about giving up on myself. This thesis could not have been written without his patience, understanding and knowledge.

I would like to thank my fiancée, Ezgi Güreler, who has always been there for me through thick and thin, throughout the three years I have been in graduate school. She has supported me constantly and never let me lose my faith. I could have not imagined finishing this project without her help or herself.

I am grateful for everything Prof. Dr. İnci Çilesiz has done for me both in undergraduate and graduate school. She is probably the biggest reason I chose Biomedical Engineering as a graduate program among all possibilities.

Also, I would like to express my appreciation to TÜBİTAK and their scholarship programme named "2228-A Son Sınıf Lisans Öğrencileri için Lisansüstü". Their financial as well as spiritual contribution to this project gave me the courage in the first place to achieve what I dreamt of.

I would like to show my heartfelt appreciation to Cihan Subaşı, Ayça Taşkın, Tansu Demirel, Utku Büyükkoca, Ferit Veliev, Gizem Pekküçük, Mehmet Muhittin Maç, Cansu Teker, Sibel Aydın, Simge Ay, Kıvanç Aras and Emre Korkmaz who have aided me through this period and just let me be.

Allow me to extend my thanks to all my friends, who cheered me up anytime I was sad and let me share both my happiness and sorrow with them. I could not ignore their roles.

Last but not least, I want to thank my parents and sister for their unconditional support and belief in me. They have raised me to this day and given me everything I have ever needed.

May 2016

Halil Emre GÖNEN
(IT Security Engineer)

TABLE OF CONTENTS

Page

ABBREVIATIONS

API **:** Application Programming Interface
NIST : National Institute of Standards and Technology
AES : Advanced Encryption System
API : Application Programming Interface
ABE : Attribute Based Encryption
CDA : Clinical Document Architecture
CSP : Cloud Service Provider
DDoS : Distributed Denial of Service
EHR : Electronic Health Record
EMR : Electronic Medical Record
HIE : Health Information Exchange
HIPAA : Health Insurance Portability and Accountability Act
HL7 : Health Level 7
IT : Information Technologies
IaaS : Infrastructure as a Service
ISO : International Organization for Standardization
ICD : International Statistical Classification of Diseases
IPSec : Internet Protocol Secure
LOINC : Logical Observation Identifiers Names and Codes
OTP : One Time Password
PHR : Personal Health Record
PaaS : Platform as a Service
RBAC : Role Based Access Control
RFC : Request for Comment
SSL : Secure Socket Layer
SMS : Short Message Service
SaaS : Software as a Service
StaaS : Storage as a Service
SNOMED : Systematized Nomenclature of Medicine
DICOM : The Digital Imaging and Communications in Medicine
IEEE : The Institute of Electrical and Electronics Engineers
TLS : Transport Layer Security
USB : Universal Serial Bus
VPN : Virtual Private Network
WHO : World Health Organization

LIST OF TABLES

LIST OF FIGURES

Page

A MULTI-CLOUD HYBRID SYSTEM FOR A STATE-WIDE ELECTRONIC HEALTHCARE VAULT

SUMMARY

The main aim of this thesis is to design a nation-wide healthcare cloud and finding the necessary aspects in designing such a system. It intends to increase the overall healthcare quality, generalize the use of standards in healthcare and come up with an approach that could be implemented anywhere regardless of the current infrastructure, while protecting the privacy of patients.

The thesis is split into 9 sections, explaining the concepts developed until today and the framework developed. The first section introduces the topic. The second section describes what cloud computing is and its various types that can be implemented. Here we adopt a multi-cloud hybrid system with an Infrastructure as a service approach. Third part gives details about electronic records that replaced the paper-based records in medical institutions. It introduces what distinguishes different record types from each other and which of them will be used in cloud scenarios. Here we are designing a system that could store both electronic healthcare records and personal healthcare records together. Fourth part explains what Health Information Exchange is, what could be the benefits coming from it and some efforts that have been put until now. Fifth part states the interoperability requirement in order to develop a full integrated systems and standardizations in medical community. Without the standards and interoperability our proposed model can not exist. Sixth part examines different papers written in the literature about healthcare clouds and what distinguishes them from each other. It tries to determine characteristics of each framework and implementation. Seventh part clarifies the methodology followed while writing this thesis. It reveals how the dimensions that helped in developing the framework has been established. Eighth part defines the new framework of the healthcare cloud system, states how the data privacy will be protected against internal and external threats. It explains how confidentiality, integrity and availability is provided in each phases of the data journey. This chapter also gives a high level illustration of the topology that could be used while implementing. Ninth and the final chapter is giving the results of the research and concludes the work done by stating the limitations and possible future works.

ÜLKE ÇAPINDA ELEKTRONİK SAĞLIK HİZMETİ KASASI İÇİN HİBRİT ÇOKLU BULUT SİSTEMİ

ÖZET

Bu tezin amacı ülke çapında bir sağlık bilgi sistemi oluşturarak, sağlık hizmetinin genel kalitesini artırmak, tıbbi alanda kullanılan standartların yaygınlaştırılmasını sağlamak ve o an kullanılan sisteme bağımlı olmaksızın yeni bir sistem geliştirmektir. Tez toplamda dokuz bölümden oluşmaktadır ve her bölüm kendi içerisindeki alt kırılımlarda olası sistemlerde kullanılabilecek detayları barındırmaktadır. İlk bölüm sağlık hizmeti konusuna giriş yapmaktadır.

Günümüzde artık kâğıt üzerine yazılan sağlık raporları gittikçe azalmaktadır. Teknolojinin hayatımızın her alanına daha fazla nüfus etmesiyle, sağlık sektörü de bu eğilimi takip etmektedir. Türkiye'de bir süredir var olan elektronik sağlık kayıtları yakın bir zamanda artık tamamıyla kâğıt raporları ortadan kaldıracaktır. Bunun için hem özel hem de devlet hastaneleri, laboratuvarlar, özel poliklinikler yatırım yapıp hem donanım hem de yazılım satın almaktadır. Ancak Türkiye'deki sağlık kuruluşlarının bu çeşitliliği standardizasyonun da güçleşmesine neden olmaktadır. Her sağlık kuruluşunun aldığı donanımlar evrensel standartlar gereği birbiriyle belli bir seviyede uyumlu da olsalar aynı durum yazılımlar için söz konusu değildir. Türkiye'de aynı Aile Sağlık Merkezi'nde bulunan farklı doktorlar bile hastalarını takip için farklı yazılımlar kullanmaktadır ve bu farklı yazılımların elektronik kayıtları birbirleriyle uyumlu değildir. Bir aile sağlık doktoruna giden hasta başka bir gün herhangi bir nedenden dolayı farklı bir doktora gitse eski kayıtları yazılımlar uyumlu olmadığı için görülememektedir. Özel hastanelere ya da devlet hastanelere giden hastalar için de durum aynıdır. Bu nedenle ortak bir yapının gereği şarttır ve bu ortak yapı ancak diğer tüm yan ögelerin kendisini desteklemesi ve birbiriyle uyumlu bilgiler göndermesiyle var olabilir. Türkiye'deki nüfusun çokluğu ve tıbbi görüntülemenin ilerlemesiyle artık yüksek hacimli 3 boyutlu görüntülerin de gittikçe artması nedeniyle bu merkezi yapıda dosya tutma ve bu veritabanı sunucularını yönetme maliyeti çok fazla olabilir. Ancak dosyaları bulut ortamında saklamak hem maliyetleri düşürecek, hem operasyonel yükü azaltacak hem de gerektiğinde neredeyse sınırsız depolama alanına sahip olma özelliğini getirecektir. Dikkat edilmesi gereken nokta ise burada tek bir bulut hizmeti sağlayıcısının kullanılması, verinin güvenliği açısından tehlike yaratabilmektedir. Bulut ortamı %100 güvenilir bir ortam olmadığı için verilerin sızdırılması gibi bir sorunla karşılaşılabilir. Aynı zamanda herhangi bir sebepten dolayı bulut hizmet sağlayıcısının erişilemez duruma gelmesi bir darboğaz yaratıp tüm sistemi kullanılamaz hale getirecektir. Bu nedenden dolayı bulut sistemi, çoklu bulut sistemi, yani birden fazla bulut hizmeti sağlayıcısından alınarak tasarlanacaktır.

Çoklu bulut sistemi sadece verilerin depolanması için kullanılacağından aynı zamanda bir özel başka bir buluta daha ihtiyaç vardır. Bu da vatandaşların şifre yönetimi ve şifrelenmiş dosyalarının anahtarlarının korunumu için var olacak aynı zamanda şifrelenen dosyaların hangi dizin altında olduğunu da tutacaktır. Böylece ikinci bölümde bahsedilen umuma açık ve özel bulut sistemlerinin birleştirilmesiyle ortaya

hibrit bir model çıkarılmıştır. Aynı zamanda "Altyapının Servis olarak Sunulması" yaklaşımıyla zararlı yazılımların sistemlere bulaşması diğer metotlara karşı en aza indirgenmiştir. "Depolamanın Servis olarak Sunulması" nın tercih edilmemesinin sebebi ise bunun çok fazla özel geliştirilmiş API'ya ihtiyaç duymasından kaynaklanmaktadır.

Altyapının Servis olarak Sunulması yaklaşımını tercih eden bu hibrit çoklu bulut sistemi aynı zamanda depoladığı verileri Shamir'in "Bir Sır Nasıl Paylaşılır" yöntemine uygun olarak şifreli bir şekilde saklanacaktır. Bu yöntem ile parametrenin seçimine bağlı olarak aynı anda hem ana anahtar hem de bir ya da birden fazla buluttaki şifreli veriler çalınsa bile ortaya anlamlı bir veri çıkmayacaktır.

Sistem mümkün olduğunca az darboğaz yaratacak şekilde tasarlandığı için tüm verileri şifrelemek yerine sadece kişisel bilgilerin gizli kalması ve tıbbi verilerin açık halde saklanması tasarlanmıştır. Bunun sonucunda gizlilik yine korunacak ve hangi tıbbi verinin kime ait olduğu anlaşılmayacaktır. Ancak bu hem sunucuların ve tüm altyapının yükünü büyük ölçüde azaltacak hem de yetkili olan araştırmacıların gerekli veriye ulaşmaları için gereken süreyi minimuma indirecektir. Çünkü şifreli bir metinde arama yapılması normal metin aramasından çok daha uzun sürmektedir.

Sistemde verinin gizliliğinin, bütünlüğünün ve ulaşılabilirliğinin, verinin kullanıcı bilgisayarı, ya da mobil cihazından çıkıp buluttaki sunucuda depolanana kadar nasıl sağlanacağı veri yolculuğunda anlatılmıştır. Sistemdeki verilere hastalar yalnızca okuma iznine sahip olarak ulaşabilirken, doktorlar ise kendi uzmanlık alanlarına göre hem sadece okuma hem de okuma-yazma yetkisine sahip olacaklardır. Sistemde bir hasta başka birisine kendi verisini görmesi için kalıcı ya da geçici izinler verebilecektir. Aynı şekilde hastalar bunu doktorlar için de ayarlayabileceklerdir. Burada önemli olan başka bir nokta ise bir hastanın veritabanında arama izni yetkisine sahip olmamasına rağmen doktorun bu yetkide olmasıdır. Bu sayede doktor daha önceden izni olmasa bile hastanın verisini görüntüleme isteği gönderebilir. Bu durumda internet bankacılığına benzer bir senaryo ile kullanıcı doktoru yetkilendirebilmektedir. Tek seferlik şifrelerle doktorlar hasta bilgisine hastanın onayıyla ulaşabilmektedir. Acil durumlarda ise hastalar acil durumda tüm doktorlar tarafından görülebilir olan acil durum notu paylaşabilmektedir. Böylece hiçbir onay olmadan sadece daha önce hastanın kendi paylaştığı bilgiye acil bir durumda erişilebilmektedir.

Bunların dışında bilgisayardan web sunucularına ve web sunucularından diğer suculara olan tüm bağlantılar TLS v1.2 protokolü kullanılarak yapılmaktadır. Bu protokol şu an için en güvenilir bağlantı yöntemlerinden biri olmakla birlikte güvenliğin en sıkı tutulduğu endüstrilerden biri olan bankacılıkta da kullanılmaktadır. Böylece iletişim sırasında araya girmeye çalışan biri iletişim şifreli olacağı için veriler onun için yine anlamsız olacaktır. Kullanıcıların kendi anahtarları devlet tarafından kurulan bir organizasyon tarafından yönetilen bir özel bulut sisteminde donanım güvenlik modüllerinde saklanacak ve böylece güvenlik artırılmış olacaktır.

Depolanan veri ise daha önceden bahsedildiği gibi kimlik bilgileri ve tıbbi bilgiler olarak kategorilendirildiği ve kimlik bilgileri şifreli saklandığı için güvenlik açısından herhangi bir probleme neden olmayacaktır.

Çalışma sırasında bu modelin tasarlanabilmesi için izlenen metodolojide ilk önce sağlık ve bilgi teknolojileri sektöründe 3 uzmanla röportaj yapılmıştır. Bunların sonucunda ülke çapında böyle bir sistem tasarlarken hangi boyutların ele alınacağı çıkarmıştır ve bunlar stratejik, teknik, tıbbi, ekonomik, sosyal ve etik & yasal olmak üzere 6 boyut olarak tespit edilmiştir. Daha sonra İTÜ'nün elektronik kaynaklarından bu konuyla ilgili belirlenen 39 tane makale için içerik analizi yapılmıştır. Bu

makalelerin anahtar kelimeleri belirlenmiştir. Son olarak da bu anahtar kelimelerle bir anket hazırlanıp "Bilgi Teknolojileri" ve "Sağlık Sektöründe" yeni çalışmaya başlamış ya da uzun süredir çalışan toplam 23 kişiye anket yapılmıştır. Ankette katılımcılardan istenen anahtar kelimeleri röportajlarda bulunan 6 boyutla eşleştirmeleridir.

Buradan çıkan sonuçlar ise tezin son kısmında verilmiştir. Buna göre literatürde teknik boyutla ilgili çok fazla çalışma olmasına rağmen diğer boyutlar buna rağmen biraz göz ardı edilmiştir. Teknoloji aynı olmasına rağmen finansal sektörde kullanılan yöntemlerin sağlık sektörüne geçirilememesinin sebebi teknik problemler değil stratejik, sosyal ve ekonomik açılardan eksiklikler olmasıdır.

Bu tez teknik olarak detaylı bir iskelet verip, diğer boyutlardaki problemler için çözüm önerileri getirmiştir. Tüm boyutlar için detaylı bir çözüm önerisi sunmak bu tezin kapsamında yer almamaktadır. Ancak getirilen çözüm önerileri ilerideki çalışmalar için bir yol haritası oluşturabilir.

1. INTRODUCTION

With the fast development and widespread use of technology medical institutions saw the benefits in adapting the technology and began to store the healthcare records of patients electronically. When shared and used collaboratively between medical institutions, what electronic medical records offer in the end is, better decision making and a means of characterizing diseases and their root causes through analytics with the help of searching and flexible handling mechanisms. However, they are still unable to benefit from it fully, because the records that are stored in one facility cannot be easily shared even between different branches of the same institution. Data sharing comes with its limitations when talking about data privacy because concepts such as data confidentiality, integrity and authenticity come into play. Since data privacy is a must in doctor patient relationships, it has been a challenge to implement a fully working scenario. Cloud Computing and Cloud Storage have been introduced to the game with the mobile devices and have gained popularity heavily because it introduced concepts like on-demand resource scaling, accessibility and most importantly security.

The use of cloud computing and storage is thought to be beneficial for the healthcare industry and enabling the medical personnel to access the healthcare records when in need will increase the quality of healthcare service as it reduces the financial cost, time loss and provides crucial information about the patient's medical background that could save his/her life. (Richardson, Abramson, & Kaushal, 2012)

The main aim of this thesis is to design a nation-wide healthcare cloud and finding the necessary aspects in designing such a system. It intends to increase the overall healthcare quality, generalize the use of standards in healthcare and come up with an approach that could be implemented anywhere regardless of the current infrastructure, while protecting the privacy of patients.

In order to utilize cloud services in healthcare industry, a lot of researches have been done in recent years, which have been examined in the following chapters. In this thesis, I am proposing a novel model to implement a hybrid multi-cloud model that

adopts Infrastructure as a service model and stores the health information encrypted and allows searching the database without compromising the anonymity of the patients. The thesis is split into 9 sections, each explaining the concepts developed until today and the framework developed. The first section introduced the topic. The second section describes what cloud computing is and its various types that can be implemented. Third part gives details about electronic records that replaced the paper-based records in medical institutions. It introduces what distinguishes different record types from each other and which of them will be used in cloud scenarios. Fourth part explains what Health Information Exchange is, what could be the benefits coming from it and some efforts that have been put until now. Fifth part states the interoperability requirement in order to develop a full integrated systems and standardizations in medical community. Sixth part examines different papers written in the literature about healthcare clouds and what distinguishes them from each other. It tries to determine characteristics of each framework and implementation. Seventh part clarifies the methodology followed while writing this thesis. It reveals the dimensions that helped in developing the framework. Eighth part defines the new framework of the healthcare cloud system, states how the data privacy will be protected against internal and external threats and gives high level illustrations of the topology that could be used while implementing. Ninth and the final chapter is giving the results of the research and concludes the work done by specifying the limitations and possible future works.

2. CLOUD COMPUTING AND STORAGE

Cloud Computing has been defined by the National Institute of Standards and Technology (NIST) as "…a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction." On many researches five essential characteristic of cloud computing stands out.

On-demand self-service: A customer can purchase individually the computing capability needed without coming into contact with a human being.

Broad network access. Any capability provided by the cloud is accessed through either a private or a public network by any platform using standard procedures.

Resource pooling. The resources of Cloud Service Providers (CSPs) are not separated based on the clients or functions. On the contrary there are pools of resources supplied such as processing, storage, bandwidth or memory and they are assigned and detached dynamically from the reserves clients use. This brings a degree of independence as neither the consumer nor the clients that uses the resources truly knows where the data is processed or stored.

Rapid elasticity. The capabilities can be increased and decreased at any moment depending on the demand. From the consumer's perspective there is no limitation to the resources and he/she can claim as much resource as desired.

Measured service. There is an automation of resource usage in the cloud system, where it can monitor, control and report the usage by measuring it based on the service type. (e.g., storage, processing, bandwidth, and active user accounts).

These five characteristics are general to all cloud systems but I would like to add two more characteristics named security and availability, which are of utmost importance to any healthcare cloud system.

Security. The resources used should be able to protect the data at all times. That means data entry, data transmission, data storage and data viewing processes should protect the data confidentiality, integrity and authenticity. We will discuss how this will be realized in the model section.

Availability. All of the resources should be available at all times. If authorized medical personnel wants to access a patient's record, a downtime of any length and duration is not acceptable in a healthcare cloud system. We will also explain how we can decrease the chances of a downtime in our model.

Classification of cloud services by NIST has been widely accepted and used in the industry. There are mainly three service models, namely software as a service (SaaS), platform as a service (PaaS) and infrastructure as a service (IaaS). Besides these three service models recently a new model called Storage as a service (StaaS) also is being offered. Figure 2.1 shows a hierarchical structure of all possible deployment scenarios of cloud computing.

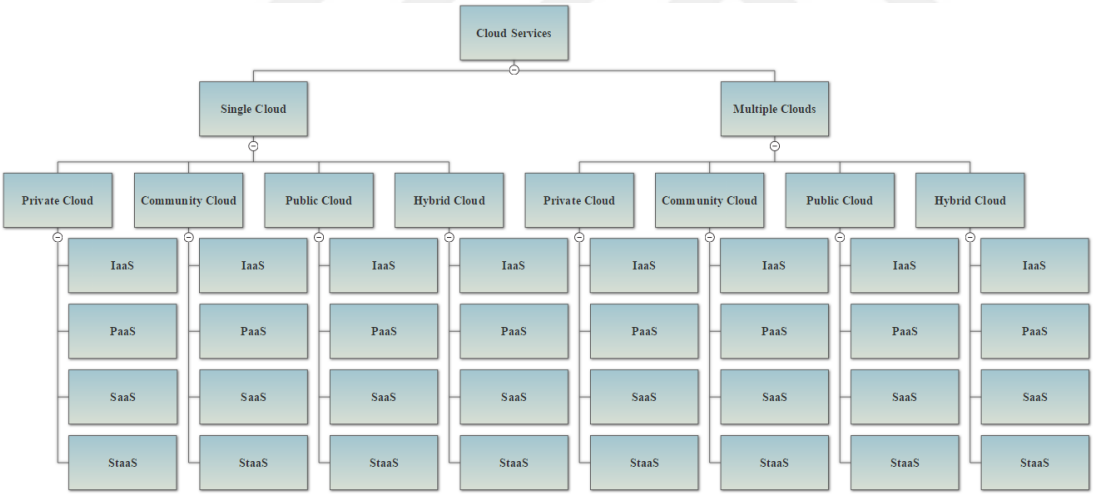

Figure 2.1: Cloud hierarchy with a matrix structure.

SaaS

Software as a service allows the clients to use the cloud service provider's applications and programs running on the cloud infrastructure by means of the Internet. A web based user interface or a thin client may be used to access the applications. The SaaS does not offer an opportunity to its clients for them to build an application or a software. The customer only uses the software offered do not own any of its right. This

service usually adapts a pay as you go model, where the client is only charged in terms of usage.

PaaS

In this service model, the client has a software or an application that it is offering to its own customers and the client needs a framework on which this application can perform its task. This could include an integrated development environment, operating system and the resources of the platform. This model does not allow its customers to manage the infrastructure but only the software developed on these platforms.

IaaS

Infrastructure as a service, as can be understood from its name, denotes everything there is to computing, e.g. network, storage, memory, processor etc. The CSP offers these resources as virtualized systems through either web based user interfaces and/or graphical user interfaces. It is worthy to mention that CSP still has the responsibility of the actual physical resources.

StaaS

Storage as a service is a new phenomenon offered by cloud service providers. The benefits are the same with other services, which includes no investment to be made on hardware, no overhead costs, no technical expertise to manage the technology. However, it requires a proprietary API to create, retrieve, update and delete data, which could cause some interoperability issues with the organization's applications.

When deploying a cloud computing infrastructure, one of the below four models could be adopted according to the desired security and availability level. Figure 2.1 shows a general overview of hierarchical structure of possible cloud deployments.

Private cloud

The cloud is only reserved for a single organization. The infrastructure can be managed by either the organization, which may or may not own the underlying physical infrastructure, or a third party firm. The cloud, likewise, may or may not be located at

the organization's site. It is almost equivalent of having an intranet with an internet access.

Public cloud

The physical infrastructure of the cloud is located somewhere off-site to the customers, open to public and is owned by the CSP. The resources are shared among all customers of the CSP proportional to their resource demands. The customers pay for the services and resources they use and not for the actual physical hardware.

Community cloud

The physical infrastructure is shared by numerous organizations/persons who have something in common. They may share the same interests, mission, security policy etc. The cloud might be managed by any of the organizations or a third party. A good example might be a community of commercial banks or medical institutions.

Hybrid cloud

Hybrid cloud is basically the combination of the two or more aforementioned cloud deployment models. The clouds retain their unique properties internally but they are able to communicate with each other via a standardized or proprietary technology.

3. ELECTRONIC RECORDS

The transformation from paper-based medical records to electronic medical records was being expected since 1990 (Sane, 1990). Electronic Medical records have long started to replace the paper based records and they have been dominating the record-keeping for a while. As the internet uses its power to connect every single machine and the data stored on it, it became almost a necessity to make better use of the medical records. (Kalra, 2006) Electronic medical records are always used in the same context with electronic health records and patient health records but there are clear distinctions between them. Below could be found the definition for each of them.

3.1. Electronic Medical Records

Electronic medical records (EMRs) are electronic correspondents of the paper charts used in medical institutions. They contain notes and information collected by a medical personnel and created each time for each new appointment in order to diagnose and treat the patient, unless it is a follow-up appointment. Nowadays paper charts are being replaced by their digital counterparts, since they enable doctors to track the data over time, observe the patient's progress and improve healthcare quality.

3.2. Electronic Health Records

Electronic health records (EHRs) transcend the general concept of electronic medical records and try to build a standard for record-keeping. They intend to contain any health related information such as medical history, diagnoses, allergies, radiology images etc. gathered from all medical institutions in order to ease the burden on decision-making and automate and streamline provider workflow. EHRs are designed to simplify the Health Information Exchange and they should have standards such as HL7, EN 13606, DICOM, openEHR etc.

Figure 3.1 shows a commercial electronic healthcare record developed by a firm that is used in hospitals to ease managing, tracking and notifying the patients.

Figure 3.1: A commercial electronic healthcare record dashboard.

3.3. Personal Health Records

Personal health records (PHRs) basically comprises of the same components as EHRs—diagnoses, medications, immunizations, family medical histories, and provider contact information—however are designed to be created, accessed, and ***managed by patients***. PHRs are used by patients to keep and manage their health information in a private, secure, and confidential environment. The basis of PHR entries can come from multiple sources such as clinicians, home monitoring devices, and patients themselves.

No matter the type, it is essential for any health system integration to exchange the records. There are a lot of factors affecting medical record exchange from trust among hospitals and perceived benefits of the system to physicians' acceptance and technological and legal challenges (Hsieh, 2015) (Chang, Hwang, Hung, Kuo, & Yen, 2009).

4. HEALTH INFORMATION EXCHANGE

Electronic Medical records have long started to replace the paper based records and they have been dominating the record-keeping for a while. As the internet uses its power to connect every single machine and the data stored on it, it became almost a necessity to make better use of the medical records.

Health Information Exchange (HIE) enables healthcare professionals and patients to access and securely share patient's medical records electronically resulting in a tremendous increase in agility, safety, cost and quality of patient care.

Even today a significant amount of medical records is kept in filing cabinets and shelfs at different medical institutes or even by patients themselves. If those paper based records are tried to share, there is a huge possibility of records being overlooked, missed or outdated. Furthermore, it will increase the time of the delivery immensely. However, if the patient information can be shared timely and adequately, the decision maker will be able to avoid readmission, medical errors and improve the quality of diagnoses by decreasing duplicate testing. We can classify HIE into three categories.

Directed Exchange gives the ability to send and receive secure information electronically between medical institutions to sustain coordinated patient care. This information could be laboratory results, patient referrals, images etc. They are sent over the Internet encrypted and authenticated.

Query-based Exchange gives the ability to find or request information on a patient from other medical institutions. This type of exchange is generally unplanned unlike Directed Exchange. For instance, emergency room physicians can utilize this type of exchange to display medications the patients use or problem lists, which could lead to the adjustment of the treatment to avoid any adverse effects.

Consumer Mediated Exchange gives patients the ability to gather and manage the use of their health information among providers. They can identify and correct wrong

health information, supply additional health information or track and monitor their own health status.

The complicated procedures of diagnosis, treatment and prevention of disease, injury and other physical and mental damages in human beings together constitute the definition of healthcare. The healthcare industry on the other side, which is one of the biggest and quickest-growing portion of Turkey's economy, is the accumulation of consumption of products and services by patients. The effectiveness, defining how good the cure is, the efficiency, a combination of the time it takes and side-effects of the cure determines the quality of health problem detection, solution identification and medical resource allocation determines the quality of healthcare (Yang, Li, & Niu, 2015).

Figure 4.1 illustrates a good example of the advantages of adopting a centralized Healthcare Information Exchange structure. Each advantage and their explanations can be found below.

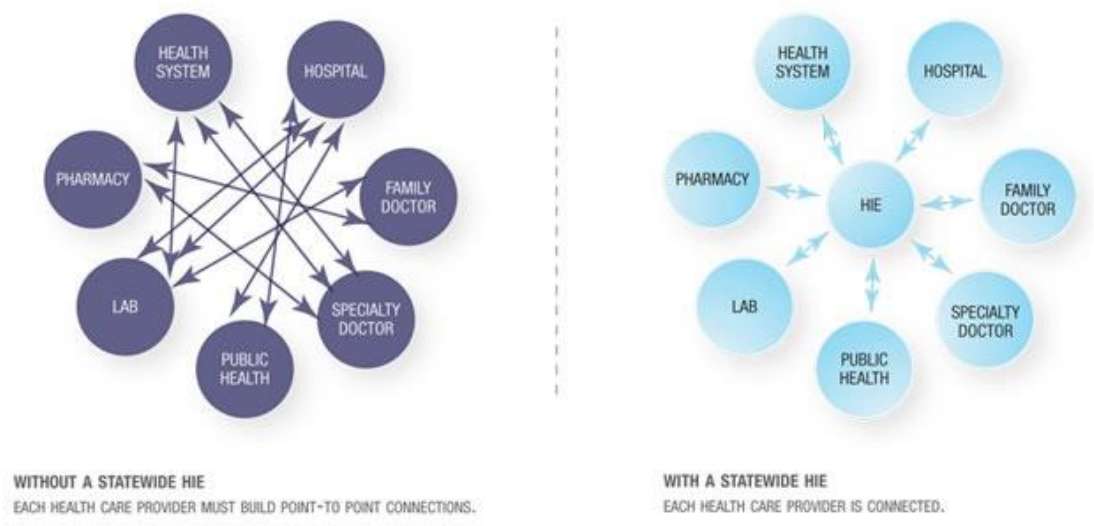

Figure 4.1: Relations in Healthcare with and without a nationwide HIE.

In order to increase the efficiency and effectiveness it is widely accepted that exchanging patients' health data among medical institutions is a must. (Vest, 2008) In Turkey, the case is totally opposite however. There is a highly fragmented "market" in terms of medical institutions with the rapid rise of private hospitals. Every hospital keeps their own clinical records making it very difficult or even impossible to share it

when seeking care at another facility. However, the branches of these private hospitals share the patients' medical information between each other by using electronic information exchange. According to a research in the United States, more than 100 organizations facilitate HIEs among provider organizations (Adler-Milstein, Bates, & Jha, 2013).

- **Health information and data**: The system stores everything that could be in a paper chart and more such as lab results, medication lists, diagnosis, ICD-10 codes.

- **Results management**: Since EHRs are stored electronically; it makes it easier to view lab results, radiology reports, X-ray images, which prevent duplication of any tests.

- **Order entry**: It prohibits prescription forgery and provides the authenticity of the prescription electronically via digitally signature.

- **Decision support**: Health Information Exchange can enable cross checking drug interactions, help diagnosing the patient and offer possible treatment options, because doctors have access to all medical data with anonymity. This in turn allows utilizing evidence-based clinical support tools.

- **Electronic communications and connectivity**: Standardizing electronical medical records and building a platform that allows patients, doctors and hospitals to interact with each other is what health information exchange is all about. Streamlining the workflow to enable meaningful communication can only be done through interoperability.

- **Patient support**: Patients could contribute to their existing electronic healthcare records with their health data taken from smart watches, phones or medical sensors. This would help the doctors in diagnosing the patient, since they would have a huge amount of sample about their lifestyles, fitness activities and diets.

- **Administrative processes**: Patients also can manage and schedule their appointments through the system. The insurance coverage can be checked

online and other doctors in the future would have a better understanding of the patient history.

- **Reporting and population health management**: Any disease outbreak trends, treatment numbers and demographic statistics can be queried from the system, thanks to huge searchable database.

5. INTEROPERABILITY AND MEDICAL STANDARDS

Healthcare Information Exchange is not a new concept. The U.S. Congress passed a legislation called Health Insurance Portability and Accountability Act (HIPAA) in order to establish national standards for electronic storage and transmission of health data, in 1996. European Union Action Plan for a European eHealth Area was announced in 2004, which was aiming to protect interoperability of eHealth systems that employ electronic health records of patients. (Chang, Hwang, Hung, Kuo, & Yen, 2009)

Furthermore, since every institution performs its own tests and scanning on the patient because they do not have the access to the same exact test that has been performed in another institution, cost of healthcare has risen dramatically. 16% of the gross domestic product in the USA is spent on the healthcare costs (Gibbs, Gilreath, Kimbrough, & Vila, 2010). In 2014, as shown by Figure 5.1, the total healthcare expenditures in Turkey by government and private sector adds up roughly to 95 billion of Turkish Lira, which accounts for 5.4% of the gross domestic product.

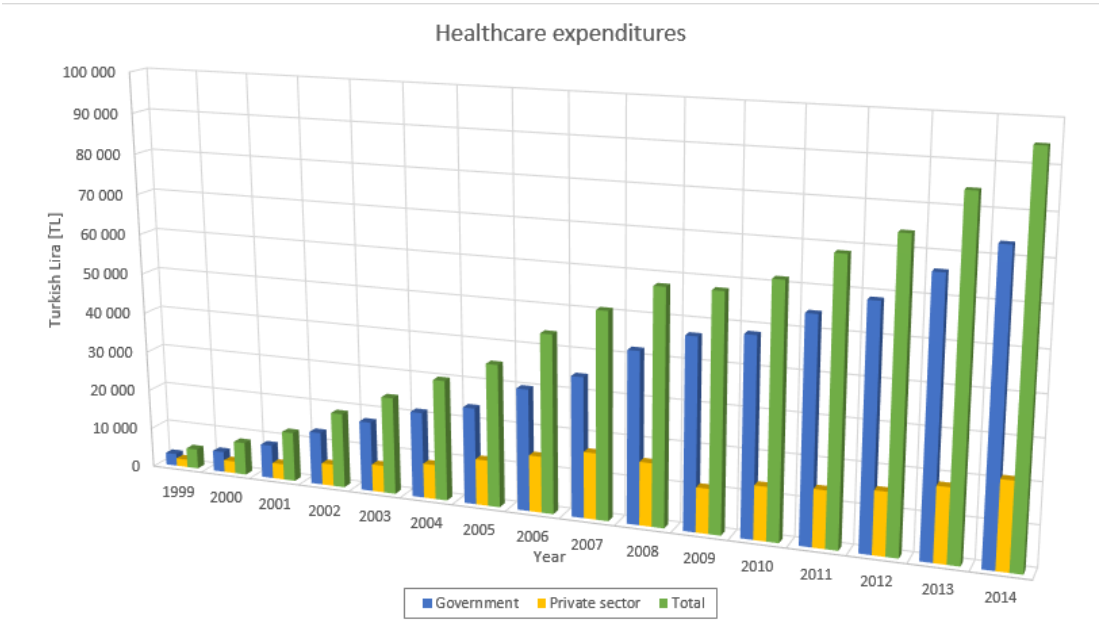

Figure 5.1: Healthcare expenditure in Turkey between 1999-2014.

5.1. Interoperability

The standardization of electronic medical/health records is necessary in order to achieve a seamless integration between every institution's information systems. Information exchange can be accomplished so long as interoperability among systems are maintained. For healthcare industry, we may talk about interoperability on six different levels according to the conceptual interoperability model. (Robkin, Weininger, Preciado, & Goldman, 2015)

5.1.1. Technical interoperability

Technical interoperability is the 1st level in the conceptual interoperability model and is used to connect systems/devices with each other employing low-level network communication protocols that enable the exchange of bits or bytes. In order to achieve technical interoperability, engineers should design a system that utilizes basically the same 0-1 system on different electronic devices. The interoperability on this level provides a basis for the communication protocols between the systems. It results in a technical structure capable of storing and transmitting the data chunks in bit format, which is a standard for packet switching networks.

5.1.2. Syntactic interoperability

Syntactic Interoperability is the 2nd level in the conceptual interoperability model and provides a shared understanding about the format of the data exchanged even though it may not know the true meaning of the data. True meaning of the data refers to the right form and order. Syntactic interoperability is critical for the operation of file systems, since it comes up with a data representation standard for an information that should be monitored and stored. Even though the seamless integration between a USB and a hard disk is taken for granted, it is this level of interoperability that makes it so. For instance, in healthcare systems, a good example for this would be a heart rate monitoring device. The heart rate data is represented in 16-bit format and the monitoring system should be able to interpret the encoding model of the data. Engineers should design the system towards both a shared understanding of basic communication protocols and the following data encoding models in order to achieve interoperability on this level.

5.1.3. Semantic interoperability

Semantically interoperability is the 3rd level in the conceptual interoperability model and provides an understanding of the content of data. The characteristic of systems on this level is knowing the meaning of the data exchanged. A thermometer can measure the temperature in Fahrenheit and send the data to another system for monitoring. While a syntactic interoperable device responsible of displaying the temperature may display 68 °C, a semantic interoperable device would know that it is actually 20 °C, because it knows the meaning of the data and not just the data. A healthcare application for this might be a heart rate monitor. Heart rate sensor can read and send signals with a 15 second period, but the monitor could convert this to a 60 second period, which is a more common way to measure heart rate. One of the biggest advantages of semantic interoperable systems is that even if for instance heart rate sensor is detached and sends -1 as a heart rate value, the monitoring system should be able to interpret this as an error and shows an error message instead of -4 for the value. In order to achieve semantic interoperability engineers must design the systems, so that not only the data but the information carried is understood. Mapping every relevant data to a concept according to its meaning and processing it in a reasonable manner can only be achieved, when systems are interoperable on the semantic level. Implementation of semantic interoperability results in longitudinal / historic records which enables EHR systems. (Robkin, Weininger, Preciado, & Goldman, 2015)

5.1.4. Pragmatic interoperability

Pragmatic Interoperability is the 4th level in the conceptual interoperability model. Pragmatic interoperable systems share an understanding of the context of the data exchanged and the associated information can be interpreted for intended purposes. This means that hidden expectations of a user can be comprehended and the behavior can be changed accordingly. For instance, adaptive video streaming is a good example of this. The supplier can understand the bandwidth requirements of the user and adjust the video resolution accordingly which results in a seamless experience for the user. The same logic can be applied to healthcare systems. In a system where the patient's heart rate is monitored and fed into a system, whose purpose is to produce alerts about the heart rate. The system can recognize, who is viewing the data and if it senses for example that the system is a mobile app that is limited by more constraints than the

usual computers, it could only send the last samples of heart rates that consequently produces better response times. In order to achieve interoperability on this level, engineers should work towards the same goal creating the same workflow and methods for distinct systems.

5.1.5. Dynamic interoperability

Dynamic Interoperability is the 5th level on the conceptual interoperability model. Dynamic interoperable systems are based on a state-model. For a given system there are clearly defined states and even if the inputs going into the system are the same a change in the system's state would change the system outputs. In order for more than two systems to act on the corresponding states, they have to understand both the content exchange and other system's current state. The contents of the data exchanged is the deciding factor in determining the right state. A good example to dynamic interoperability can be the safety belt mechanisms in the new cars. Safety belt has two states, either fastened or loose. The speed of the car could decide the outputs of the safety belt system. If the car is stationary, no warning is given regardless of the safety belt's state. However, pushing the throttle would accelerate the car if the seatbelt is fastened but would not have any effect on the speed if the belt is loose. Consequently, the state change in of the systems, changes the outputs of another. In order to achieve interoperability on this level, engineers designing the systems should consider utilizing distinct systems with deterministic interactions depending on each other's state. Implementation of dynamic interoperability results in predictable and deterministic systems interacting with each other according to their dynamic states.

5.1.6. Conceptual interoperability

Conceptual Interoperability is the 6th and the highest interoperability level in the pyramid. The standards of the systems in this level is so well-defined that it will allow any vendor to create systems/devices capable of understanding data models, concepts and states of other systems/devices they are interacting with. The assumption in this level is that regardless of the vendor, the implementation of the systems is functionally identical. The shared conceptual model allows engineers to design fully integrated systems assembling different processes and products of different vendors into one single system comprised of dynamically interoperable devices/applications.

Figure 5.2 shows the layered approach of interoperability in the shape of a pyramid. As the area of the layers get smaller going up the pyramid, both the availability and the simplicity of such systems are also decreasing. Defining layers and then identifying the standards have always been the key in IT systems design. Much like the network protocols used today, which adopts a 7-layer approach in exchanging the data, in order to build interoperable systems, their interfaces have to be designed considering the standards that describe their characteristics on all six levels. There is no need to have a single standard that describes all six levels mentioned, on the contrary it is more desirable to have a system where a combination of different standards can be employed. Luckily, health care industry has lots of standards developed until today, which is one of its weakness and strength at the same time.

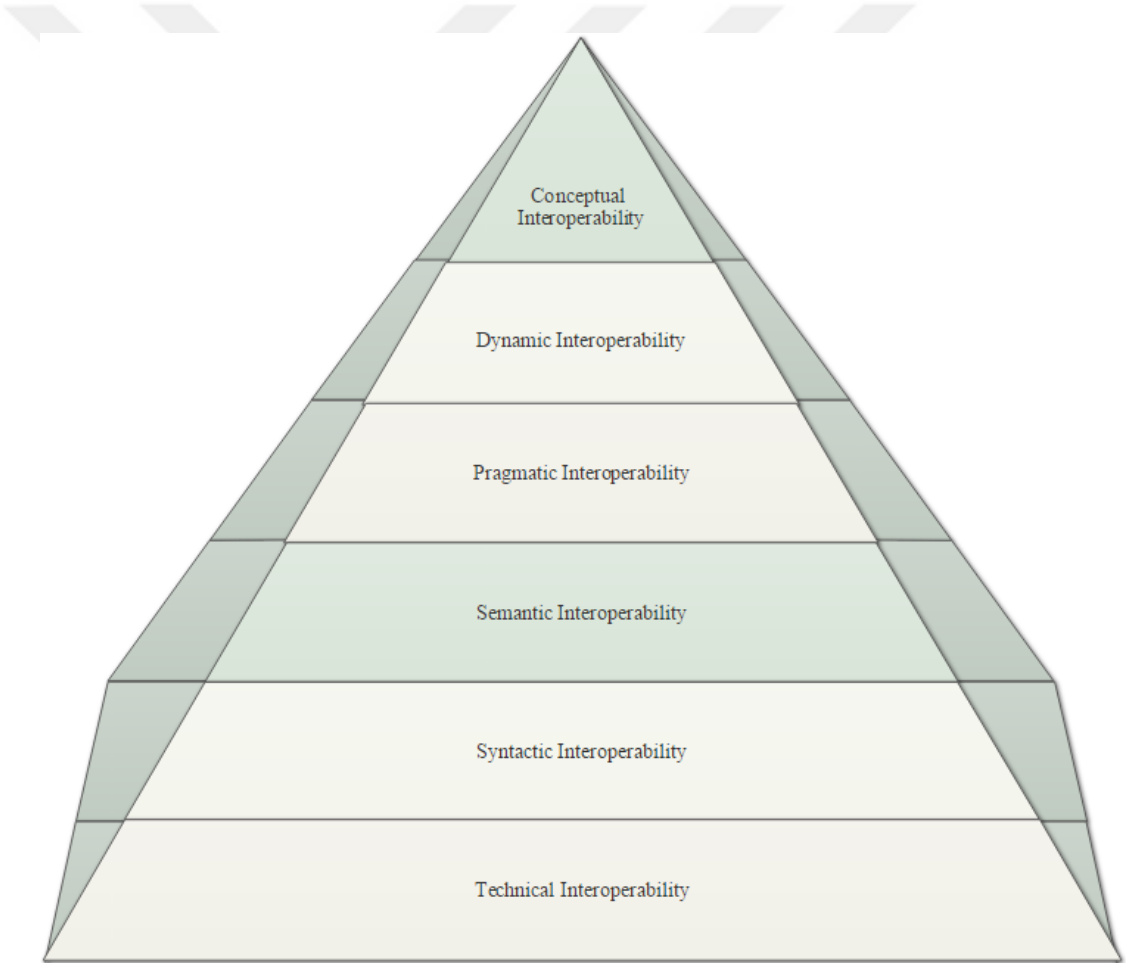

Figure 5.2: Pyramid of Interoperability.

5.2.List of Standards

The standards that will be explained below does not cover all of the standards that has been developed in the medical community but it shows an excellent picture of different

perspectives and spans the most used ones as mentioned in the literature. (Eichelberg, Aden, Riesmeier, Dogac, & Laleci, 2005)

5.2.1. openEHR

An EU research project called Good European started in 1992, which turned into openEHR and it is currently maintained by the openEHR non-profit organization. The introduction of archetypes is the most different characteristic of openEHR standard. Like meta-modeling, clinical information statements are designed in a double layer concept. The first layer only consists of a few components, which can be thought as a simple meta-model. Domain specific notions such as clinical observations are represented by archetypes utilizing the elements of this meta-model. This is done by assembling and naming elements from the meta-model, connecting them and putting constraints on them. Components of archetypes are also linked to other semantic data standards besides naming. Archetype Definition Language (ADL) introduced by openEHR can be used to generate archetypes.

5.2.2. EN 13606

EN 13606 is a communication standard for medical information in electronic medical records and focuses on interfaces for data exchange and structured data packaging for communication. Central databases, applications and software pieces can exchange information between them and health records can either be transferred as a whole or in chunks. The data representation depends on openEHR framework mentioned above.

5.2.3. ISO/IEEE 11073

The ISO 10073 family of standards defines protocols and data formats for transmission between electronic medical devices. It undertakes mobile devices that are used in acute care settings, and is therefore designed with the specific goals:

- Real-time interoperable plug-and-play devices
- Simple implementation of protocol stacks
- Resource-efficient message processing
- Handling of frequent network configuration changes

The set of standards is combined of:

- Defining terms and services that will be utilized in the communication protocol through an object-oriented data model (Domain Information Model (DIM), ISO 1173-10201)
- Identifying transferred items with a set of numeric codes which is a part of standardized nomenclature (ISO 11073-10101).
- Restricting the nomenclature and data models Application profiles, which restrict the nomenclature and data model to specific communication needs.

5.2.4. LOINC

LOINC is an abbreviation for Logical Observation Identifiers Names and Codes. The system is developed to help with the naming and coding of clinical observations. The system is published and has a publicly accessible database which is maintained by Regenstrief Institute (Indianapolis, USA). The information below should be encoded in each observation:

- Observed subject
- Observed property / measurement metric
- Time stamp
- System: kind of sample
- Scale: quantitative, ordinal, nominal or textual

Unlike ICD coding system (International Statistical Classification of Diseases and Related Health Problems.) that covers and encodes acutal diagnoses, laboratory results are especially best expressed with LOINC structure. Standards such as Health Level 7 or Clinical Document Architecture also uses the LOINC coding system to encode health data.

5.2.5. Snomed CT

Snomed CT is an abbreviation for Systematized Nomenclature of Medicine – Clinical Terms and a terminology standard comprising of medical concepts that try to achieve semantic interoperability. A numeric, distinctive code made up of six to eighteen

figures is assigned to each concept. ach concept is assigned a numeric, unique code consisting of six to eighteen digits. The number assigned to “Diabetes mellitus” for instance is 73211009.

Snomed CT is formed by an acyclic graph that consists of concepts represented as nodes and connections between nodes. Specialization/generalization relationship between two concepts is indicated by a connection between two nodes. For example, diabetes mellitus in generalized to disorder of endocrine system, which in turn is a specialization of diabetes, which is in turn a specialization of a metabolic disease. Standards such as Health Level 7 or Clinical Document Architecture also use Snomed CT coding system to provide semantic interoperability.

5.2.6. Health level 7 (HL7)

The most widely used group of standards for communication of clinical information is developed by Health Level 7; a non-profit organization founded in 1987. These standards contain:

- Message protocols (HL7 v2.x, v3)
- Conceptual standards (e.g. HL7 RIM)
- Document standards (e.g. HL7 CDA,)
- Application standards (e.g. HL7 Clinical Context Object Workgroup CCOW)

Message protocols in HL7 are designed to be generated by events. An event in clinical work could be a trigger event (such as a patient admission). A request message sent to another system is generated by a trigger event, which consequently leads to gathering of data in order to reply to the request. The data is assembled according to EDI standards to form a reply message. HL7 version 2 is one of the most widely implemented standard and exists in different subversions extending from 2.1 up to 2.6, which are backward compatible. Textual delimiters are used as part of the encoding but not XML. Even though for many tasks in clinical work processes, message exchange is defined by the HL7 v2, there is no consensus on a base data model, which results in lack of definition and semantics of the data. Although this allows for great

flexibility on one side, on the other side following HL7 v2 standards exactly would not promise interoperability without further mutual agreements.

To improve on top of HL 7 v2, the Reference Information Model (RIM) is introduced. Semantic connections between data entities and concepts are shown and communicated through this model by means of message exchanging. As mentioned previously, medical data standards such as SNOMED CT or LOINC is used to define the data explicitly.

The Clinical Context Object Workgroup (CCOW) is a standard that allows visual integration of clinical applications and tries to achieve a unified view on clinical data that is located in different component interfaces. Linking the context in different applications is how CCOW unifies access to patient's data and functionality. To put it more simply selection of a specific patient in one application triggers selection of the same patient in all other applications via single sign-on mechanism.

5.2.7. Clinical document architecture (CDA)

Assembling clinical information into documents for exchanging purposes is defined by the Clinical Document Architecture which uses XML-Markup-based document standards. Its structural components depend on data categories of the Reference Information Model of HL7 v3.

Interoperability is fragmented into three separate levels of machine readability and processability for CDA documents. CDA documents consisting of a header and a body may contain formatted text and on the first level this is the only requirement for simple transmission of data (Syntactical Interoperability). The document body is built in compliance with RIM into chunks of observations. On the last level every data field is semantically encoded on top of the existing structure, so that it will deliver a document that could be fully processed without human interaction (Neuhaus, Polze, & Chowdhuryy, 2011).

5.2.8. DICOM

Digital Imaging and Communications in Medicine (DICOM) is a standard for producing, storing, displaying, processing, sending, retrieving, querying or printing of medical imaging as well as managing related workflows. It contains a file

format definition and a network communications protocol. TCP/IP is used to communicate between systems. Any two entities can exchange DICOM files if they are capable of receiving image and patient data. The copyright of the standard is held by The National Electrical Manufacturers Association (NEMA), because the development of the standard was DICOM Standards Committee; whose members are also partly members of NEMA.

5.2.9. ICD-10

ICD-10 stands for The International Classification of Diseases 10th version, which is a standard for describing and coding mortality and morbidity incidents, implemented by most World Health Organization (WHO) member states. The change is considered a requirement because of the inadequate and outdated coding offered by ICD-9, and the need for global consistency. The revisions of ICD will continue to happen because the diseases, their root causes and their effects are changing. Diagnostic codes have amplified from 14000 to 68000 with the transition from the 9th version to the 10th. ICD-10 has the capability to describe the circumstances of injury, and contains all imaginable and uncommon injuries such as; problems in relationships with in-laws and prolonged stay in weightless environment.

ICD-10 requires a higher level of specificity in clinical documentation as displayed in Figure 5.3. Doctors will need to document the diseases with severity, laterality and encounter sequence.

Figure 5.3: Representation of disease encoding in ICD-10.

6. LITERATURE REVIEW

Exchanging healthcare information through cloud is a popular research subject, because unlike exchanging medical records between individual servers and individual organizations, it provides a collaborative platform to store and edit all the records. (Mathew, 2013)

One of the researches have focused on the image encryption in the cloud system. It developed a concept, in which images were reduced to pixels and then encrypted using Paillier Cryptosystem compared to text files that were encrypted by Advance Encryption Standard (AES). (Aiswarya, Divya, Sangeetha, & Vaidehi, 2013) Considering how images are reshaping the diagnosing process in the medical world, it is reasonable to focus on image encryption and data storage but even though the paper uses a cloud system, it does not come up with a unified cloud system. Furthermore, it uses an asymmetric cryptography, which has a lower performance compared with symmetric cryptography. Another article puts forward a method to encrypt the images using AES-128, however healthcare cloud is out of their scope. (Radhadevi & Kalpana, 2012)

A different paper took the image sharing on cloud to the next step and developed an engine for lossless and adaptive engine to compress and store 3D images. (Castiglione, et al., 2014) Looking at the usage trends of 3D images, it is safe to assume that they will play an even bigger role in the future at diagnosing and treating the patient. Considering their big size, it is a robust and sizeable solution to store these 3D images on the cloud. Nevertheless, their proposal in securing the data and sharing it with other peers does not scale well to an increase in the network since they rely on a peer to peer network and it would be very hard to handle bandwidth limitations and manual password or key sharing.

Cipher text Policy Attribute Based Encryption (ABE) is used in a framework that is structured upon a cloud system which tries to manage the data created by medical wireless sensor networks (Lounis, Hadjidj, Bouabdallah, & Challal, 2015). The article

aims to use the data gathered from all sensors in case of a medical emergency both as proactive and reactive precautions. Their solution does not focus on the medical data but rather health data. Moreover, they are using asymmetric encryption to secure the data with ABE which is inefficient in encrypting and decrypting the data and has major challenges such as key coordination and attribute revocation. They try to overcome this difficulty by using symmetric encryption in the process also that results in a more complex solution.

A cloud based framework for Health Care System has been designed in order to apply cluster techniques for diagnosis in one of the related articles. They defined key segments of healthcare as patient, doctor, diagnosis and symptoms and focused on these aspects by creating a hybrid cloud that aims to help data mining tools. (Parekh & B., 2015) Nonetheless they have only mentioned the challenges that could be faced in a cloud environment such as security and accessibility and did not provide a solution regarding these challenges.

One of the most comprehensive work in healthcare cloud has been put forward in 2014 covering the data security in cloud, search through cipher text and electronic health record standards. (Yang, Li, & Niu, 2015) Their framework consists of three spheres, namely user sphere, joint sphere, recipient sphere. The article focuses on standardizations of electronic health records and based on this develops an encryption technique that only encrypts as less data as possible without compromising anonymity, resulting in a more efficient storage model. With the help of their vertical partitioning model they achieve to decrease search times in cipher texts nearly to plain text search times levels. Despite the vertical partition model though, instead of using multiple clouds, the model is based on just one cloud service provider. In addition, it is not exactly clear what kind of cloud service is deployed or how users and medical staff will interact with the system or who will be in charge of medical and health related data entry.

A real-world application of an electronic based records on a cloud system has been studied in Kenya for maternal and child health exclusively. (Haskew, et al., 2015) However, their application is narrow in two contexts. First of all, the cloud is not generalized for all patients but rather targeted a very limited profile that consists of only pregnant women. Secondly the cloud system is used only by a small number of

institutions which enabled key distribution and building IPSec VPN tunnels possible. Sizing this to a state-wide model, it would be practically impossible to build and maintain the infrastructure.

Even though the subject is not particularly about healthcare, one of the papers did a research about the governmental use of cloud in US and its possible security risks. The paper presented known cloud challenges from a governmental perspective and states that all stakeholders, which means not only the government but also the citizens, should have the right to voice their opinions about the decision making processes. (Paquette, Jaeger, & Wilson, 2010)

Another paper proposed exchanging the medical records of a patient over the internet with xml language using a chart like medical record. The article proposed that every hospital should have a gateway and a web server to complete the transaction. (Liu, Long, Li, Tsai, & Kuo, 2001) This was a pioneering idea at that time, using electronic based records and sharing them on a common platform but it is outdated now, considering the fact that the world's first DDoS attack was launched around the same time. (Dennis, 2016) Internet has become very unsecure to casually publish data right now and security measures have to be taken to keep the data safe. A simple web server as suggested will not meet the demands of today's massive internet traffic.

A hybrid healthcare cloud concept is put forward in an article, concentrating on the data flow between private and public cloud (Marcu & Popescu, 2014). Their solution includes a role based access control system and encrypted data transmission between the servers and cloud. The model states that data itself is not stored on the public cloud and instead it works as a tool for indexing. However, a public cloud is always thought as an untrusted zone, therefore giving a server on the cloud an access authorization could mean giving access authorization to anyone, who is in charge of the server. Furthermore, their solution does not scale well to a nation-wide solution because, they are suggesting IPSec VPN tunnels between private and public clouds, which is impractical due to the massive number of hospitals, family healthcare centers, laboratories in a country.

The only article known to the author about a health information system in Turkey undertakes the infrastructure of the newly deployed system in 2008 (Köse, et al.,

2008). The article presents the action plans taken by the Ministry of Health in Turkey and the aims of the system that was trying to be created. The paper mentions the use of HL7 and CDA but also a standard developed by the Ministry called National Health Data Dictionary. It goes into the details of the National Health Data Dictionary and articulates the components of the proposed health information system, including pharmacists, family physicians, hospitals, decision support systems etc. Even though the article states that it is expected to collect 90% of patient data from the institutions in 2009, that was not the case. Moreover, the article does not explain any security-related concepts or how the integration between the systems will be performed.

7. METHODOLOGY

7.1.Interview

Interviews and focus groups are the most used methods particularly in Healthcare Research (Britten, 2007). Interview can be defined as a verbal conversation between two people with the objective of collecting relevant information about the subject. The qualitative interview's main aim is to describe these subjects based on the interviewee's experience and understand the meaning of responses.

According to an article, the interviews are especially useful for getting the story behind a participant's experiences and as a follow-up to certain respondents. Furthermore, the interviewer can pursue in-depth information around the topic. (McNamara, 1999)

There are three basic types of research interviews: structured, semi structured and unstructured. Structured interviews are, in principle, verbally administered surveys with a list of prearranged questions are asked that has either little or no variation, resulting in no room for follow-up questions to responses that permit further elaboration. Consequently, even though they are comparatively quick and easy to oversee and useful for situations in which clarification of concepts are required. It is not about what one interviewee particularly say or how it is said but rather the cumulative responses to questions. They would not provide, by their very design, depth about any topic.

Unstructured interviews, on the other hand, do not impose any preconceived theories or ideas and are administered with little or no organization. An unstructured interview may start with casual conversation leading to the topic at hand such as "Can you tell me about your experience of visiting the hospital?" and will shape according to the responses of the interviewee. Unlike structured interviews, they are very time-consuming and difficult to oversee. Unstructured interviews are usually very time-consuming and can be difficult to manage, and to participate in, since there is not a road map about what to talk about. It is advised to use this kind of interview method,

when there is little known about the subject and the depth contributed by the interviewee would help structure, categorize and classify the subject. Often, different people from different professions are selected to capture distinguished perspectives.

Semi-structured interviews are a combination of the two types mentioned. Even though it comprises of several key questions that assists in discovering the areas to be explored, it also enables both sides to track an opinion or a response elaborately. The flexibility provided by this approach lets the researcher examine the concepts or information that had not been previously thought relevant to this subject. In healthcare this method is often used, because it provides some guidance to the interviewee. (Gill, Stewart, Treasure, & Chadwick, 2008)

I have conducted three semi-structured interviews during this thesis from different professions. One of them is an IT Security Unit manager from one of the biggest banks in Turkey, one of them is a director of IT department in one of the most prestigious hospital chains in Turkey and last person is an IT specialist in healthcare industry.

7.2.Content Analysis

Content Analysis is both qualitative and quantitative research method that has been first used in 1950 in a study of mass communications (Berelson, 1952). As a research method content analysis is a systematic and objective approach to describe and quantify phenomena. It could also be described as analyzing of documents. It enables the researcher to enhance understanding of the data by testing theoretical problems. One of the main goals of content analysis is to compress words into content related classes (Elo & Kynga, 2008). Words, phrases and expressions share the same meaning, when classified into groups.

One of the most vital subject regarding the content analysis is if the analysis will be performed on manifest or latent content. Manifest content refers to observable expressions and evident expressions that appear on the text. It analyzes accountable data pursuing a quantitative approach. Alternatively, latent analysis refers to hidden meanings and relationships between words and phrases (Rossi, Serralvo, & João, 2014). Content analysis can be utilized either in an inductive or deductive manner

besides assessing qualitative or quantitative data. It is up to the researcher and the research topic to choose a path.

The categories of the analyzed components are derived inductively, if there is not structured information and sufficient knowledge about the phenomenon. (Lauri & Kynga, 2005). Deductive content analysis on the other hand, is used when the structure is clear and the former knowledge about the concept is enough to test the theory. While an inductive approach is more creative and introduces a general model or a general concept based on one or more specific fragmented applications, a deductive approach does the exact opposite and takes a general concept or a theory that has been in the literature and tries to achieve either prove the theory or apply it on a small scale, which means that it goes from a general to a specific statement (White & Marsh, 2006).

I have chosen to use a quantitative method in content analysis with an inductive approach, because a consolidated and centralized framework for healthcare information exchange has not been fully applied in anywhere and small-scale applications guide the way through a nationwide system.

In order to come with a general model for a Nationwide Healthcare Cloud, a total of 39 articles have been chosen by searching keywords like “Healthcare Information Exchange”, “Electronic Records” and “Healthcare Cloud”. Articles related to cloud systems, healthcare integration and electronic medical records have been reviewed. Content analysis along with a survey performed on the professionals both in IT and Healthcare Industry, is used to fully develop a model and clusters are generated as dimension of the model.

8. FRAMEWORK OF HEALTH VAULT

The content of a health cloud system is three-fold: a medical cloud for sharing electronic health records (EHRs) across facilities in different hospitals; a care-cloud so that wireless patient monitoring devices can allow for the monitoring of blood pressure, heart rate, and glucose, to name a few, and enable a patient's health data to be transmitted between different locations; and a wellness cloud that uses open data and cloud platforms to encourage value-added service providers to develop various innovative applications, thereby allowing people to obtain health-related information at any time in order to enhance self-health management. (Hu & Bai, 2014)

The healthcare industry and commercial banking have much in common. Both are service providers in a low-margin, highly-fragmented, capital-intensive, politically sensitive, commercially challenging and technologically complex industries that are among the country's most heavily regulated (Morrissette, Burgdorfer, & Shields, 2014). However, the difference between the two industries' IT adoption is astounding. Looking at these similarities first and then discussing why the adoption rate of Information Technologies is so different will shape our model.

Both industries are low margin because they do not earn that much profit from a single customer or from a single procedure, as they do in retail or tech companies. They both are however billion dollar industries because of their giant customer profile. They both are highly fragmented because the demand is too much and comes from everywhere. One central or two or three giants cannot meet the demands of the clients, therefore there are a lot of institutions that offers what the other cannot. Both of the industries are capital intensive, as the devices that hospitals use costs millions of dollars and banks by their very definition must have capital. Both of them are politically sensitive, as they have to adjust the current trends in the politics. It is commercially challenging, because there are a lot of rivals in their industry and a single mistake can make them lose their customers. Technological advancement takes its toll heavily on these industries, since they rely on technology to keep up even with their day to day jobs

such as drawing money from an ATM or monitoring the patient's status via sensors. They have to adapt to new technologies, otherwise they will become useless. Finally, they are both heavily regulated by governments in terms of data storage, share and accounting.

It is strange to see even with this much similarity, one of the industries has adopted IT very rapidly and the other one very slowly and the key difference here is the standardization. Medical community does not have a one-size-fit-all standard even for data collection let alone data storage and communication. As mentioned before even standards need complimentary standards to fully capture the essentials of information. This is mostly due to the fact that financial data, which consist of numbers, get along well with informatics. However, healthcare industry is not doomed about this and they are picking up the pace. It is not about adoption anymore even though the adoption rate of electronic health records is not at the desired levels, it is about consolidating the data and accessing it whenever needed from wherever needed according to the security roles that have been established.

In order to cope with ever increasing demands of the patients I have developed a model of a patient-centric and secure public healthcare multi-cloud using Infrastructure as a Service approach, considering the similarities with the financial industry and the biggest reason for this is the fact that security is essential for both parties, since the healthcare data requires just as much and maybe even more privacy than financial data.

Besides the technical aspects that will be explained below, some other assumptions should also be given. The model developed should be owned by the government and they should involve in every step of the implementation. Their persistence on this subject will increase the chances of a sustainable model. They should put the patient to the center, as the model does and define the data owner as patients. They should lead the way in using standardization of electronic records and assure interoperability. It is of utmost importance integrating different systems seamlessly, which requires interoperability at least on the 3rd level. The system should decrease the overhead costs and the workload. The government, as the owner of the system, should define the responsible authorities and organizations very clearly. No one would want to participate if no one is aware of who is in charge of the processes and who is the one to blame, if anything goes wrong, which is a culturally significant trait of Turkish

people. Lastly, if each and every statement above can be realized, which is not that hard to do, the user resistance to a new technology and to a new concept would fade away and user's acceptance would grow exponentially as in the case of every social medium. People would want to be a part of the system, because their family and friends are also parts of the system.

First, let us take a look at what we should focus on regarding the privacy of the data. An electronic medical record can be thought as a building block for the electronic health record. These individual records submitted by different institutions should carry the three most important security characteristics, namely confidentiality, integrity and availability.

Confidentiality refers to keeping the privacy of the data. Data and its contents should not be viewed by anyone who has not any authorization. Integrity means that the content of the data should not be modified in any way. The authenticity should be preserved, unless it should be changed by an authorized person. Being able to access the data whenever needed is defined as availability, which means that all systems should be functioning correctly that store, process and communicates the data. Last but not least, even though it could be somewhat counted as part of integrity, non-repudiation also plays a big role not just in technology but in general. It means, both parties of a transaction cannot deny that they were in fact a part of the receiving/sending process which can be realized by logging.

These concepts should be enforced on data entry, data transmission, data storage and data recall (viewing/printing). Let us examine each phase separately and build a model in which we can enforce the security concepts.

8.1. Data Entry

Data entry may be the most important part of this whole system because if it goes wrong, there is no meaning in storing, encrypting and accessing the data.

Comparing this model with the ones used in financial institutions is not that hard. Authentication scheme is the same as financial institutions, where if you try to access to your bank accounts via a web browser, you go through a two-factor authentication

process with your mobile phone. The difference is that in our framework another person can also access to your information but again only with your authorization.

Data entry is the starting point in creating an electronic healthcare record. In our model it is done via a web user interface from anywhere. Much like today's paper based patient records, the electronic medical records and healthcare records can only be created/edited and viewed by authorized personnel in these organizations. There is a three-layer protection scheme in our model. First one is role base access policy, second one is unique user/password combinations and the last one is one time passwords.

Role based access policy (RBAC) is initially created so that each citizen will be able to access his/her full information. Patients will be given read-only policy access and doctors will be granted with read-write policy. A citizen will grant approval to another person by adding him to the allowed lists using his Citizenship Identifier. The person can get granular due to the nature of electronic records and give only permission for some medical categories and not for others. For instance, the person can allow access to dermatology related medical data but not for radiology. These permissions can be given permanently or for a specified period of time.

The granularity of the model comes from the standards of electronic medical records. Electronic medical records will follow standards such as HL7 and DICOM. This is currently the case for many Turkish medical institutions but not for all of them. Even though the standards are used, the current system still operates without categorizing the information. The patient information is just a sum of many records without any partition. Radiology background and dermatology background can be found in the same text box in the electronic records.

Right now, there are 43 specializations for doctors in Turkey (Pamukkale University, 2016) and the current system does not differentiate different medical records provided by different medical specialists, even though it can prevent e-prescription frauds by not allowing certain medications given by unauthorized staff. Our model suggests employing these different categories, which means that neurology category can only be edited by neurology specialist and not others. This way access permission to a doctor can be as granular as a specialization not the entire patient history.

E-prescriptions can be ordered by doctors through this system and pharmacists can see and grant the required drugs and medicines to the patients. Our model suggests that the e-prescriptions and electronic records modified or created by doctors should be digitally signed, which will improve data integrity and authenticity. Figure 8.1 displays how a user experience can be for the patients in the healthcare vault system.

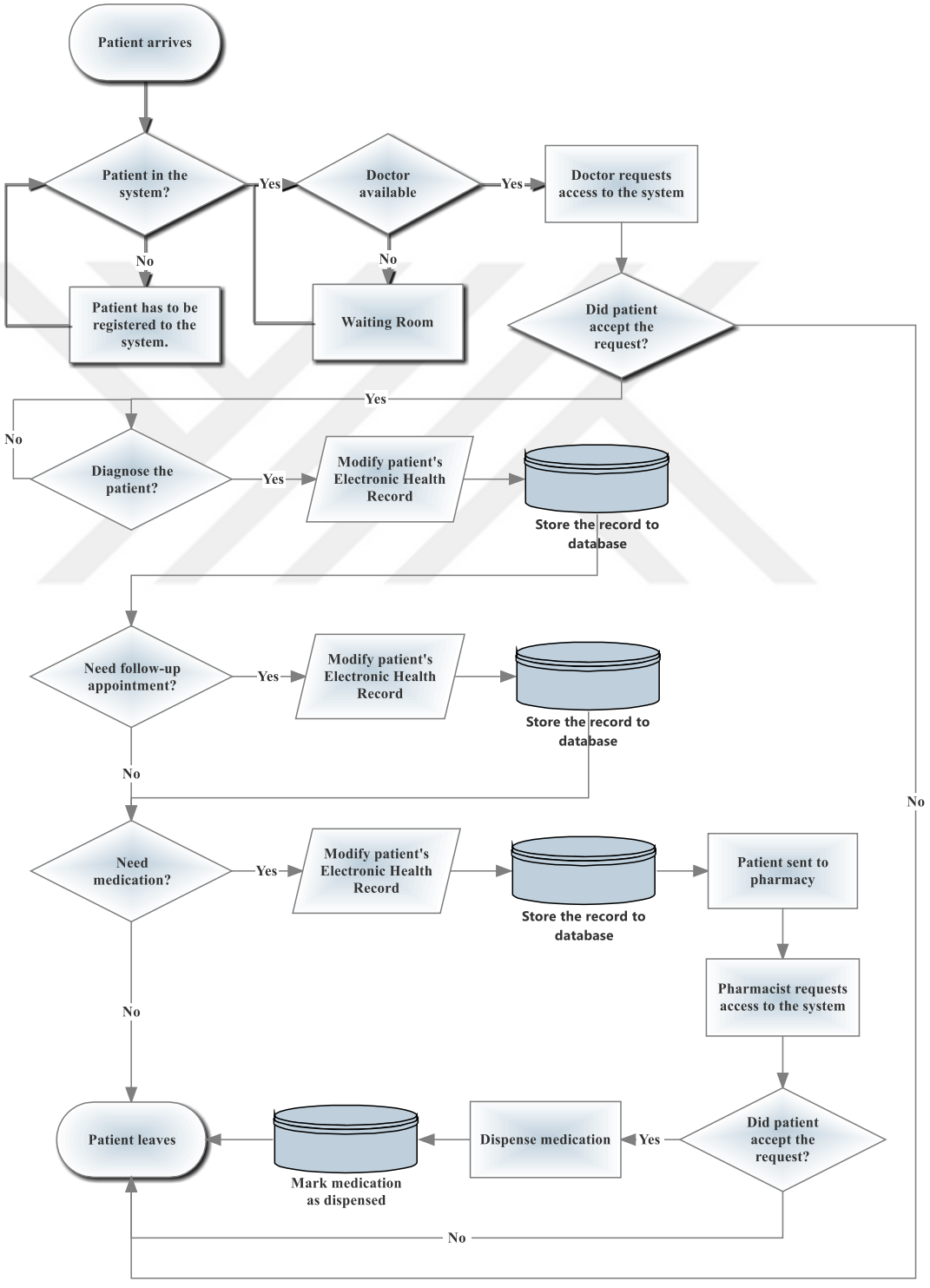

Figure 8.1: Patient Experience of the Healthcare Vault.

There is already an initiation of utilizing digital signature by the Health Ministry of Turkey and some of the doctors are using it on local systems but the total integration is yet to be completed.

Another vital contribution of the cloud is to integrate it with appointment scheduling. Patients can schedule appointments with the doctors and doctors can be notified prior to the visit. Patients can share their medical history with the doctors and doctors would have the option to take a look at the record, which would help him/her in the diagnosing process.

8.1.1. Patient access

The first step is the user/password combination for everyone just like with online banking. They have to first access to the system by using their unique user-password combination and afterwards there will be a push notification or an SMS OTP sent to their mobile phone if they are accessing their own data. SMS OTP will be used in non-smart phones and it will be matched with their SIM-Cards. Application OTP or push notifications will be sent to those who own a smartphone and it will be matched with their device-id again just like online banking. It is a very straight-forward process and used heavily in our everyday lives.

Note that, patients will only be able see their own data and the profiles that are shared with them and they do not have any authorization for requesting access to any other profiles.

All of profile viewing activities will be logged and will have a timestamp for auditing purposes. If a person accesses some other patient's profile through either permanent or temporary permissions, a notice will be sent to said patient notifying who's viewing his records but it will not ask for permission again.

8.1.2. Doctor access

The model described above about two-factor authentication where if a patient tries to view his or someone else's data is very much possible and a part of our lives. Doctors' access though is what distinguishes our model from the current online banking procedures. In online banking, your accounts can only be viewed by you and no one else. In our state-wide healthcare vault, it can also be seen by doctors.

The permissions for the doctors can be done through interactive selection, because doctors and their institution will already be in the system, so the patients should not need any identifiers to allow access to their data. Again just like before these permissions can be either permanent or temporary. The same logic about granularity will also apply for read-write policy and doctors will be given as granular access possible if needed. Finally, if the patient did not allow a doctor to see its medical data beforehand, the doctor could ask for permission.

Doctors will be able to see the profiles that are shared with them and they will also have the authorization to request other patient's data. In this case, once they are logged-on to their accounts, once they try to access a patient's data, OTP procedure will kick in and the user will get again either a SMS OTP or an APP OTP specifying who is trying to access his/her medical records and from which medical institution. If the patient approves the access request, the medical personnel can go ahead with his/her task. This can be done very easily considering the developments in mobile applications. It can even be implemented in applications that have been launched by the Turkish government such as e-government or e-pulse. Surely, we rely on the assumption that the patient has at least a mobile phone, but then again this is a system that relies heavily on information technologies.

As described for both scenarios, a strict role based access model, user/password combinations and OTPs will be used to enhance security. The option to establish a trust relation with a person, which will mean that once the person logs in to the system, he/she will have the right to see all records that have been shared with him/her will dispose of the redundancy of approving the same person's request over and over again if the person is trusted. Furthermore, it will also optimize the resource usages, since every request/response will create an overhead on the system.

Besides read-write and read-only rights, only doctors will be able to search for a patient and their search history will also be logged besides access logs. In order to avoid abuses and DDoS attacks if a doctor searches for more than 10 patients in 5 minutes, he will get a timeout for half an hour. If the abuse goes on, the timeout duration will keep on increasing. However, this will not impeach their ability to search medical data because the threshold will only be set for accessing identifier data and not the medical data.

8.1.3. Emergency access

Maybe, one of the biggest concern of a cloud system is availability. Because when there are so much components, flows, systems and most importantly humans involved, there is always the possibility that something might just go wrong.

Emergency access in this sense refers to the fact the patient might not be in a state where he is capable of approving requests from his/her mobile phone for a doctor that he did not previously build a trust relation with.

Currently, in real emergency situations where the patient is not conscious and the patient has to be treated right away, a doctor has neither the resources nor the time to look at the patient's medical history and allergies etc. They have to either stop the bleeding, give a needle shot or even operate on him to get his body functions to work properly. They have to apply the standard procedures and cannot predict if a conventional method would be even worse on him. The patient may not even have his ID with him for the doctor to search and find him in the database.

Our model would present the doctors with the patient's history, if of course the patient has something to identify himself/herself with. In this case there are two possible scenarios that will be chosen by the patient. The patient could put up an emergency section for all doctors to see. Only he will have the right to create/edit this section and he could give notices for what to be careful about. The patient should bear in mind that this will be always available to any doctor, so any private information in the patient's eyes should not be shared. The other option is to assign three emergency contacts for his profile. These three people will have the permission to authorize the doctor's request in an emergency.

This of course will be the patient's choice and bears the long-lasting question of is it the security-over-availability or availability-over-security.

8.2. Data Transmission

Transferring data securely from one server to another is mostly done by data encryption, which is the most conventional way to guarantee data security in both healthcare and financial systems (Morrissette, Burgdorfer, & Shields, 2014). The

security is established by using symmetric cryptography, whose shared secret is established during the handshakes.

The most common technique is to use Secure Socket Layer (SSL), Transport Security Layer (TLS) or IPSec Virtual Private Network (IPSec VPN). All of them are cryptographic protocols that provide secure communications over unsecure mediums. There are different versions for both SSL and TLS and some of them became obsolete by now, however the newest version are still widely used in applications such as web browsing, voice over IP, instant messaging and e-mail.

8.2.1. SSL v3.0

SSL v3.0 protocol was released in 1996 but first began with the creation of SSL v1.0 developed by Netscape. Version 1.0 was never released and version 2.0 had a number of security flaws due to its weak algorithm, thus leading to the release of SSL 3.0. Some major improvements of SSL 3.0 over SSL 2.0 are:

- Separation of the transport of data from the message layer
- Employing full 128 bits of keying material
- The ability to use certificate hierarchy, which provides a depth greater than two certificates by enabling both client and the server side to send chains of certificates
- Employing a standardized key exchange protocol, allowing Diffie-Hellman and Fortezza key exchanges as well as non-RSA certificates.
- Granting record compression and decompression

However, SSL v3 was compromised in 2014. The SSLv3 key exchange was found to be vulnerable to man in the middle attacks when renegotiation. (Internet Engineering Task Force, 2016)

8.2.2. TLS 1.0

This protocol was first defined in RFC 2246 in January of 1999. There were enhancements form SSL v3 but the differences between them were not big. However,

it was still enough to cause interoperability issues with SSL v3. Considerable distinctions between SSL v3.0 and TLS 1.0 are the followings (Internet Engineering Task Force, 2016):

- Key derivation functions are different.
- Message authentication codes are different. TLS v1 uses a hash mechanism on top of normal authentication codes.
- TLS has a different ending message and more alerts.
- TLS needs Digital Signature Standard and Diffie Hellman key exchange support to work properly

8.2.3. TLS 1.1

This protocol was defined 7 years after the original protocol was released and is an update to TLS 1.0 (Internet Engineering Task Force, 2016). The improvements include:

- Explicit Vector replaced the implicit initialization vector to protect the data against cipher block chaining attacks.
- Cipher block chaining attacks are mitigated by changing the way padded errors are handled with the help of bad_record_mac alert rather than the decryption_failed alert.
- Protocol parameters are defined by IANA registries.
- Unexpected connection drops do not disturb the session integrity.

8.2.4. TLS 1.2

TLS v1.2 protocol was represented in RFC 5246 in August of 2008 (Network Working Group, 2016). It is based on both TLS v1.1 and 1.2 but includes improvements, some of which can be found below:

- Pseudorandom function is replaced with a more secure mechanism that utilizes cipher-suites.
- A single hash was started. The MD5/SHA-1 combination in the digitally-signed element was replaced with a single hash. Signed elements include a field explicitly specifying the hash algorithm used.

- A considerable amount of hashing and signature algorithms have been eliminated both on the client and server's side.
- TLS Extensions definition and AES Cipher Suites were merged in.
- Tighter checking of EncryptedPreMasterSecret version numbers.
- Some attack vectors have been eliminated and the protocol is hardened.

8.2.5. IPSec VPN

IPSec VPN stands for “Internet Protocol Secure Virtual Private Network” and is a protocol that runs at Layer 3 of the OSI model. It offers data confidentiality, integrity, data origin authentication and replay protection of each message by encrypting and signing every message with agreed standards. Many “Request for Comments” (RFCs) are combined to create this protocol and it basically has two different applications. First application method is through Authentication Header (AH), which does not provide data authentication. The second application on the other hand, called Encapsulating Security Payload (ESP), provides both authentication and confidentiality. Two endpoints generate an IPSec Security Association (IPSec SA) via dynamically established keys that use standards such as AES and SHA. Dynamic establishment of keys are done according to Internet Key Exchange v1 (IKEv1) and IKEv2 protocols, which is out of this paper’s scope. Some of the characteristics of IPSec Protocol are as follows (Internet Engineering Task Force, 2016):

- Public Key Infrastructure is not required; pre-shared keys can be used to generate other keys.
- Since it is a protocol that is running on L3, any application (Layer 7) or transport (Layer 4) protocol will be protected.
- Applications will not even know that IPSec is used because they are higher on the OSI model.
- AHs are used in transport mode for secure session between endpoints and ESPs are used in tunnel mode for secure connections between gateways.
- TLS is preferred to IPSec, since it is a newer protocol that has less vulnerable spots.
- Since it adds other headers to the TCP/IP packets, it is not suitable for big datagrams.

- It is only authenticating the network and the data origin and not the actual application or the user.
- Configuring IPSec tunnels are complex and time-consuming, if the IT personnel is not experienced.

In our model data security during transmission is provided by TLS v1.2. All up-to-date browsers have been supporting it for quite some time. The reason we chose TLS v1.2 is its superiority over other protocols. IPSec VPN is not a feasible option, since this system will be used by everybody and creating that many SAs is nearly impossible.

Figure 8.2 depicts a high level overview of the system. Patients and doctors from either the same or different medical institution will be able to access to the system. A multi cloud secure proxy server will be the first hop on the topology. This will serve as both a proxy web server and a load balancer. Thus it will provide a more secure connection mitigating some of the vulnerabilities that would be exposed if the users talked directly with the web servers. Furthermore, it will also distribute the load on the servers evenly.

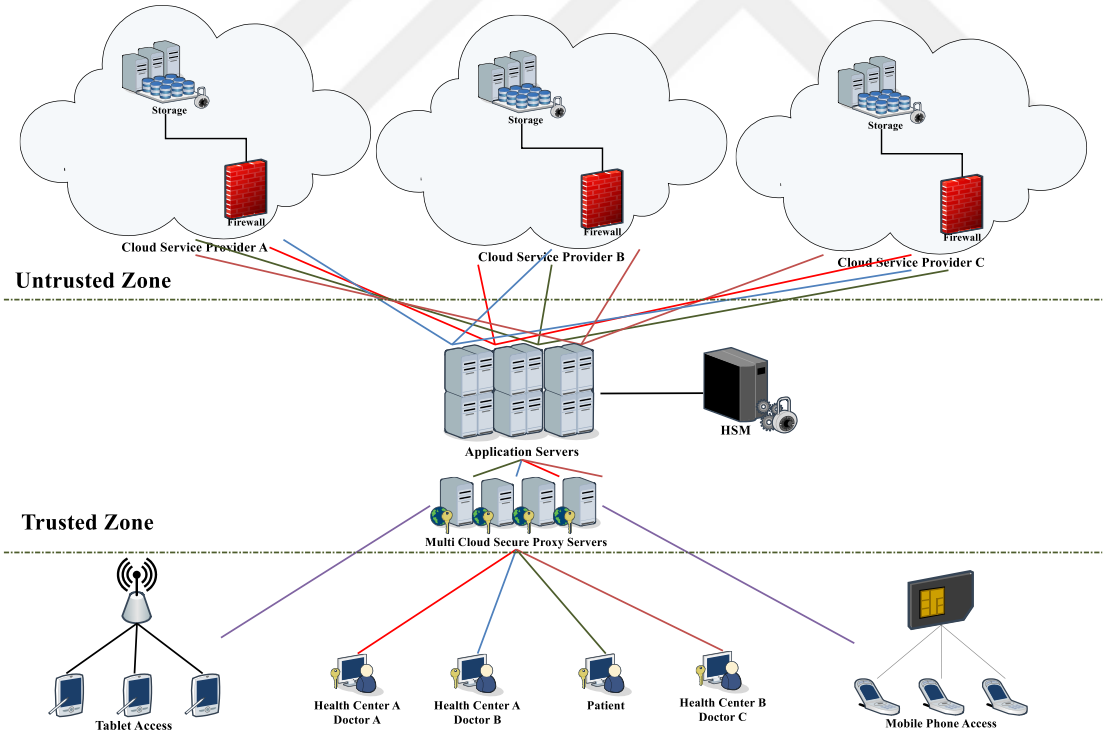

Figure 8.2: Hybrid multi-cloud topology with a trusted zone.

The connections from tablets, mobile phones and PCs will be established using TLSv1.2. It will prevent any eavesdropping and man in the middle attacks. Multi-cloud proxy servers will communicate with the application servers also using TLS v1.2 even

if the layer is shown as a trusted zone. Trusted zone only means that the communication is not open to public but we still assume that anybody in the intranet could be curious. Application servers will process the data and do the encryption/decryption processes via HSM. The data then will be stored encrypted to the database server.

8.3. Data Storage

Data storage may be the second most critical part of the model. Cloud storage has its advantages such as unlimited resources and availability but the biggest challenge has always been security. In our model security in storage is provided by a two-fold model. Firstly, the data will be encrypted and stored this way, which will protect the privacy and integrity not just from outside threats but also threats from within. Even if somebody would be able to break in to the database and access the files, they will be encrypted and will make no sense unless somebody also has the master key to decrypt it. Secondly, the encrypted data will be split into chunks and stored on different clouds based on an "m of n split" method. Thus, even if a cloud is totally hacked and the master key is compromised, the data will not make sense unless m fragments of data have been captured.

The first protection layer of data storage is encryption. We have talked about data encryption during transmission with well-established protocols. Data will also be encrypted using a key derivation function. First, a symmetric key will be generated using Advanced Encryption System (AES) and using a key derivation function another unique key will be generated with the credentials provided from the patient. During the log-in process this key will be compared with the key produced by the client just like it is done on mobile banking applications. There is also going to be timeout period to prevent replay attacks. It is fairly simple to establish such a system considering they are widely used by mobile banking applications. The difference will be another person will also be able to login to the system to access the patient's profile. This will be done through the patient's mobile phone. Either an OTP or a push notification request will be approved by the patient to authorize the doctor.

An EHR can be divided into three partitions. First one is the Real Identifier, attributes that could clearly identify a person, e.g. Name, Phone Number, Citizenship OD etc.

Second one is the Pseudo Identifier, attributes that could be used to identify the patient when combined, e.g. address, birth of date, educational background etc. Last one is the Medical Data, information that is only related to medical procedures, diagnosis, treatments etc. (Yang, Li, & Niu, 2015) Our model suggests that only medical data is unencrypted and the reference table between the medical data and both identifiers should be encrypted. This way the encryption/decryption process will only be performed on identifiers, whose table sizes are very low compared to huge amounts of medical data that could be filled with images, laboratory results and patient history. This model would protect the anonymity of the patients with still encrypted data and would make searching the database easier, since searching on cipher-text is a complex and time-consuming process.

The last protection layer is to split the encrypted data. The main goal of splitting the data is to decrease the chances of privacy violation. The medical data will be stored on each cloud, which will increase the availability and shorten the response times of database servers. Splitting the encrypted data according to m of n principle (similar to Shamir's Secret Sharing) however, will guarantee that even in a worst-case scenario where the key is compromised and the database server of a cloud is compromised, unless m-1 clouds is also hacked, they would not be able to get a meaningful data (Shamir, 1979). Shamir's method is based on the fact knowing two points on a slope would give you all the points on the slope but knowing one would not amount to anything. (Wagner, 2016) The same logic can be applied for polynomials of n^{th} degree that requires at least m points to solve the equation.

The algorithm below is designed to employ Shamir's method in a multi-cloud system.

F_{ui} : File i uploaded by user u.

Dig(F_{ui}) : Digitally signed file

Enc(Dig(F_{ui})) : Encrypted digitally signed file.

SEnc(Dig(F_{ui}))$_j$: jth particular chunks of the encrypted file.

R_k : Large random number.

T : Time interval to change access path.

Algorithm 1

Begin

Apply digital signature on F_{ui};
Store the hash of F_{ui} in trusted zone;
Perform AES-256 bit encryption in Fui;
Split Dig(Enc(F_{ui})) into n chunk according to Shamir's Principle;
Initialize j to 1;
while j<n do:
if Chunk number equals cloud number
Do not store chunkj on the cloud;
else
Store chunkj on the cloud;
Increment j by one;
End while
while 1 do:
Initialize p to 1;
while p<n do:
Generate a large random number R_p;
Increment p by one;
end while
Initialize k to 1;
while k<n do:
Store Enc(R_k) in private cloud;
Increment k by one;
end while
Initialize m to 1;
while m<n do:
if R_m not exists then
make directory R_m;
Increment m by one;
End if
end while
Initialize l to 1;
while l<n do:
move SEnc(Dig(F_{ui}))$_l$ into R_l;
Increment l by one;
End while
Delay loop with time T;
End while

End

8.4. Data Recall

Data recall is the last step on the data chain. It entails gathering encrypted and plaintext data from multiple clouds and presenting it to the user via a web interface. Since only the identifiers are encrypted, reversing the process is not that difficult.

The encrypted data is recalled from multiple clouds, combined and decrypted with the secret key. Application server feeds this data to the web server with a TLS connection. The user can modify/view/print the data according to the authorization profile, he/she has been given.

Data entry and recall can be seen as the same thing. The reason they have been separately undertaken is data entry is discussed from data access perspective and data recall is discussed from data decryption perspective.

9. FINDINGS AND CONCLUSION

This thesis developed a novel model to implement a nation-wide secure healthcare cloud system. In order to structure the model properly, first three semi-structured interviews have been conducted and then a content analysis has been performed on 39 articles, as can be seen on the Table A.1, have been reviewed and their keywords have been analyzed. The keywords have been matched with the dimensions revealed in the interview phase with a survey conducted on professionals working in IT or Healthcare sector.

The first interview has been conducted with an IT Security Unit Manager working in a financial institution with a +20 years' experience. The general concept was how to design a secure cloud system meant to protect data of the users and which aspects should be considered. The structure of the interview allowed him to come up with his own perspective. He mostly approached the topic from a technical perspective, like how the servers should be located, if they should adopt a three-tier structure, which algorithms should be used in encryption/decryption processes and how the key management can be done with so many users. Since the system at hand also resembles a financial institution's online procedures and processes, it was not that hard to find a solution that could govern the general public. The challenge was to protect the data against threats from within and from outside. With a well-written web-API, a good-designed topology and up-to-date softwares along with firewalls and IPSs, it is easier to protect the data against external attacks. However, a database administrator within the cloud can transfer the patient files to his memory stick for instance and walk out. It is easier to prevent this in an organizational structure rather than in a cloud. This is why the patient's information should be stored encrypted. The second perspective he was interested in was the legal/ethical issues that might arise. This is probably due to heavy regulations he is facing from the Banking Regulation and Supervision Agency in Turkey. Recent Distributed Denial of Service (DDoS) attacks against financial institutions opened up a debate about how to filter attack traffic with the agency (Hürriyet Haber, 2016). The traffic could be re-routed through anti DDoS services

outside the country however, this was verbally forbidden by an auditor from the agency. He was interested in solutions that would be accepted by regulators and the public.

The second interview was a conducted with the IT director of a prestigious hospital franchise in Turkey with +15 years' experience in healthcare and a +6 years' previous experience in financial institutions. The conversation started out the same but took a different turn. His approach started out from technical perspective also but then lead to a high level outlook. He was more concerned about the government's approach and their strategy in implementing these technologies. He stated the trend for adaptation electronic and centralized IT systems has started at least 10 years ago, however lack of strategic depth has been a main concern for everyone in healthcare industry. A clear vision about a centralized system was absent and every medical institution had to come up with their own approach about the IT infrastructure and its requirements. The government should promote the adoption and lead the way for the private sector in healthcare instead of other way around. He was also worried about the social acceptance of an IT system in healthcare. The opinion of public, both the patient and the doctor was a deciding factor in such a transition. His views were about how the system is represented. The system should be able to provide an ease of usage for the participants no matter how complicated the system has been designed.

The third interviewee was an IT professional that has been in healthcare industry for +7 years. His area of interest was informatics and worked mostly with doctors designing apparatus and medical devices. His take on a secure cloud system was based on a medical perspective. The discussion was mostly about standardization and interoperability of medical data and medical devices. The main focus was use-cases and real-world scenarios of how a doctor might diagnose the patient and express his findings so that any doctor looking at the records would understand the exact same thing, even if he had not seen the patient himself. The second biggest concept was the financial burden of such a system. The dialog became about the cost of the transition for all including parties and the overhead expenditures. He expressed also his opinions about the operational costs not financially but as a human resource. The only way to have a sustainable nation-wide system is by decreasing the workload of the humans as much as possible.

Having done these three interviews has offered me a structure that can be composed of six dimensions when designing such a system namely; strategic, technical, medical, economic, social and legal & ethical dimension.

Strategic Dimension: Strategic dimension entails deciding who is the stakeholder, who is the data owner and who is responsible from the system. The stakeholders in our model are everybody who's in the medical community. Pharmacists, doctors, hospitals, patients, laboratories etc. They were always immutable parts of the health system and our model is not an exception. The data owner is the patient. Their health and other identity related information belongs to them and only they can give permission to anyone who wants access. Lastly, the responsible organization is the government. Their job is to educate the people about the system and do campaigns so that even if the system initially lacks something, it could be a sustainable system

Technical Dimension: Technical dimension involves building an infrastructure, determining the electronic record standards and managing the secret keys used in the encryption. Government as the responsible body have to come up with the topologies, protocols and choose the cloud service providers.

Medical Dimension: Medical dimension contains anything that is related to the actual medical procedures. It is a required dimension because in the end the whole system is about increasing the quality of healthcare services. People serving their country as doctors, clinicians, specialists and the organization such as laboratories, hospitals are composing the medical dimension.

Economical Dimension: Economical dimension is always a limiting aspect in system designs. No system can be designed with unlimited resources and therefore it is a necessity to look at the economic burden of any system. Even though it is a limitation, the constraints that it manifests always results in optimized systems.

Social Dimension: Social dimension should be considered because after all the system relies on people. Patients and doctors are what constitutes the system. Without their positive take on the adoption and contributions, the system itself will be useless.

Ethical & Legal Dimension: The government is responsible from the legal aspects of the system. Even though data is stored in the cloud and they do not have to deal with

storage they are the ones who should protect the privacy of the patients and of course the constitution. Ethical questions can always arise when privacy is a subject. Apart from a legal perspective the answers to ethical questions will always shape according to the culture of the nation.

A content analysis is then performed on 39 related articles examining their keywords. As mentioned in the methodology section, the reason we have chosen a quantitative inductive content analysis method is because there has not been a full governing model or concept about a nation-wide healthcare system, let alone healthcare cloud.

With the keywords from the articles, after the repetitions have been left out, and the dimensions obtained from the semi-structured interviews a survey has been prepared. The survey has included 144 keywords and the participants are wanted to categorize the keywords to the given dimensions according to their own understanding. A keyword could have been put into multiple categories, if the participant saw it fit.

The survey has been done by 23 people, 15 of whom were IT professionals and the remaining 8 were Healthcare professionals. All of the 23 attendees have at least an undergraduate degree with an average of 4 years work experience. Table 9.1 shows the general results of the survey. The summation of keywords that belong to a dimension is bigger than 144, since a keyword can be in multiple categories.

It is immediately seen from Figure 9.1 that technical dimension of healthcare systems has been researched heavily. Medical dimension and ethical & legal dimension follow the technical dimension with a huge gap and rest of the dimensions are pretty much evenly distributed.

This clearly indicates why a collaborative approach towards a unified system has not happened yet. The technical difficulties are not the main reason for not adapting such a model like in financial industry. There are surely numerous solutions to technical difficulties that might arise. The real reason is that other dimensions are ignored.

6 keywords from the survey have been selected to be demonstrated with percentages and each with their corresponding dimensions. Economic assessment is vital in determining the system’s efficiency. It should be compared with the current infrastructure to evaluate how much benefit would it bring to implement the system.

Since the proposed method is based on cloud, which is very cost-effective even in IT solutions that is itself cost-effective. Cloud computing term has been widely accepted as a technical dimension, which is quite understandable. It is only a medium, a tool in serving people and nothing else. Privacy is righteously rated as an ethical/legal matter. It is very essential to protect the privacy of the patients and generally people do not want to be a part of anything that is accepted as an ethical grey area. User resistance has been rated as a social dimension; healthcare itself has been labeled as a medical dimension and finally risk management is assessed as a strategic dimension. These keywords have helped finding the borders of each dimension and finding solutions regarding the issues faced in each one.

Table 9.1: Survey classification result with sample keywords.

Keywords	Strategic	Technical	Medical	Economical	Social	Ethical & Legal
Economic Assessment	43,5%	8,7%	8,7%	**91,3%**	21,7%	8,7%
Cloud Computing	17,4%	**91,3%**	8,7%	17,4%	0,0%	21,7%
Privacy	8,7%	21,7%	17,4%	4,3%	39,1%	**95,7%**
User Resistance	17,4%	13,0%	4,3%	0,0%	**65,2%**	26,1%
Healthcare	26,1%	21,7%	**87,0%**	34,8%	65,2%	34,8%
Risk Management	**73,9%**	30,4%	17,4%	30,4%	30,4%	39,1%

Figures 9.2, 9.3, 9.4, 9.5 and 9.6 show the word trees for all dimension. The word tree for technical dimension is too long, therefore it is not practical to draw it on such a tree.

To start with everything that is related to privacy, security and exchange is interpreted as ethical & legal dimension. That is why the model should propose a solution towards security and privacy. Even though technical precautions are taken, it is necessary to reflect this to the end user. This could only be done via an organization, which takes the responsibility for this task and anything that might go wrong. Again when compared with a financial institution, if the data is stolen they do not have the chance to deny this and they have to live with the consequences. The same thing should also be valid in this case. Just as a bank would compensate its clients for a damage, the governmental organization should compensate its citizens for any harm.

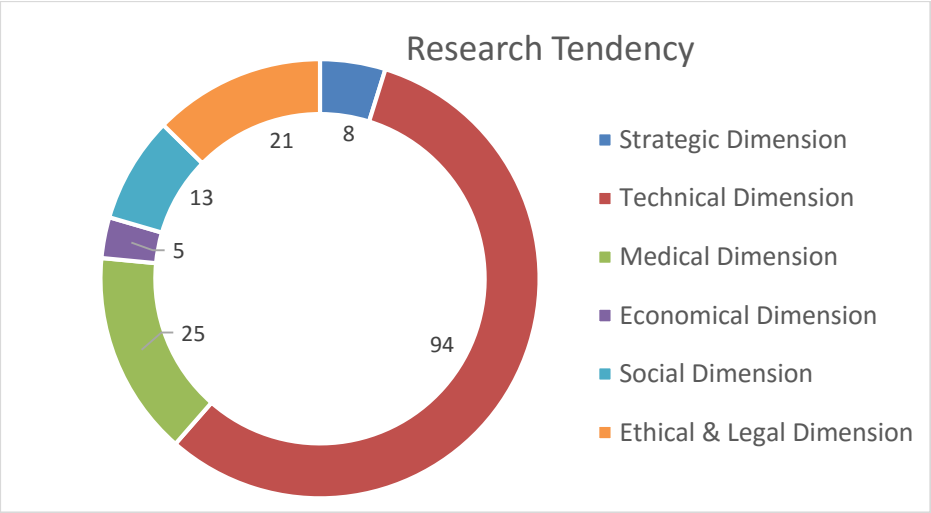

Figure 9.1: A doughnut chart of keyword-dimension classification.

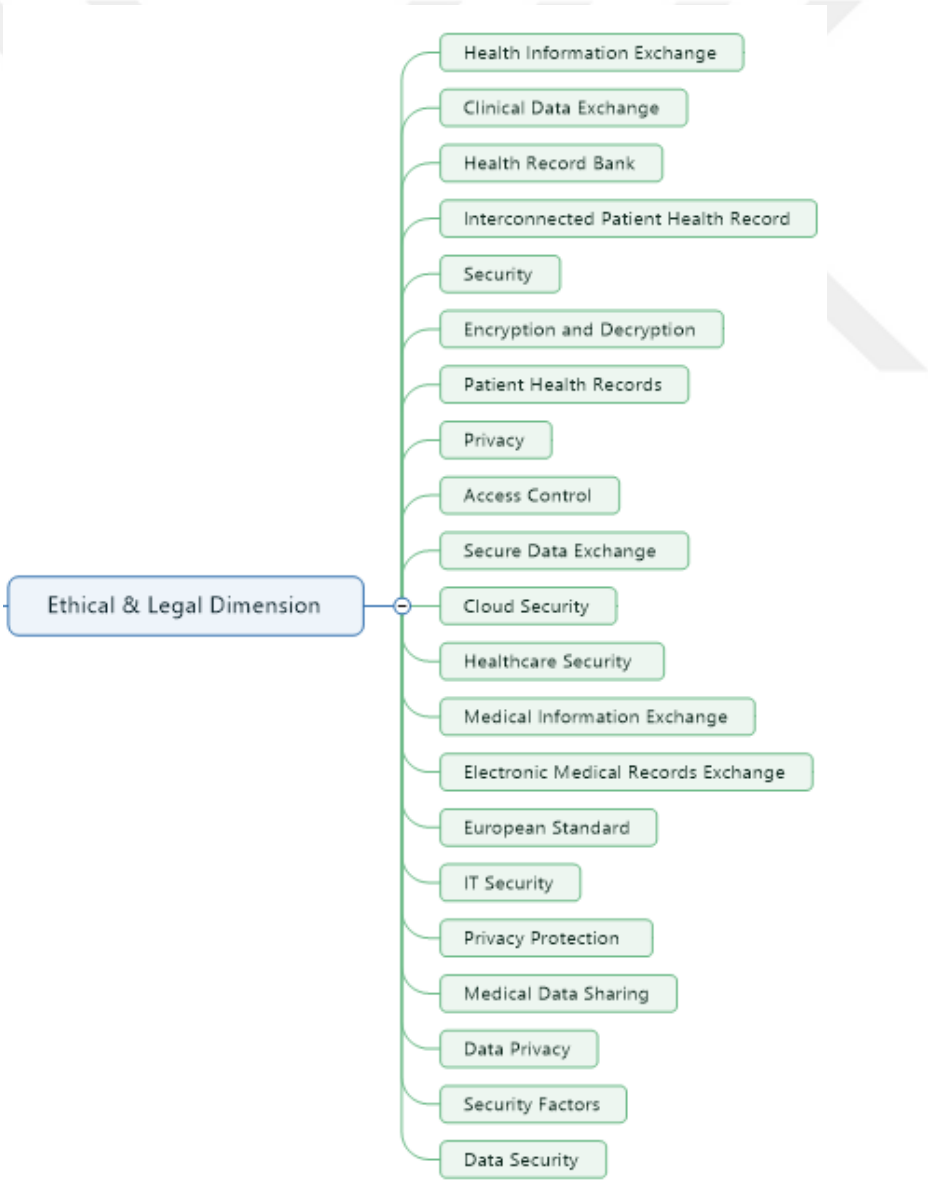

Figure 9.2: Ethical & Legal Dimension Keyword Tree.

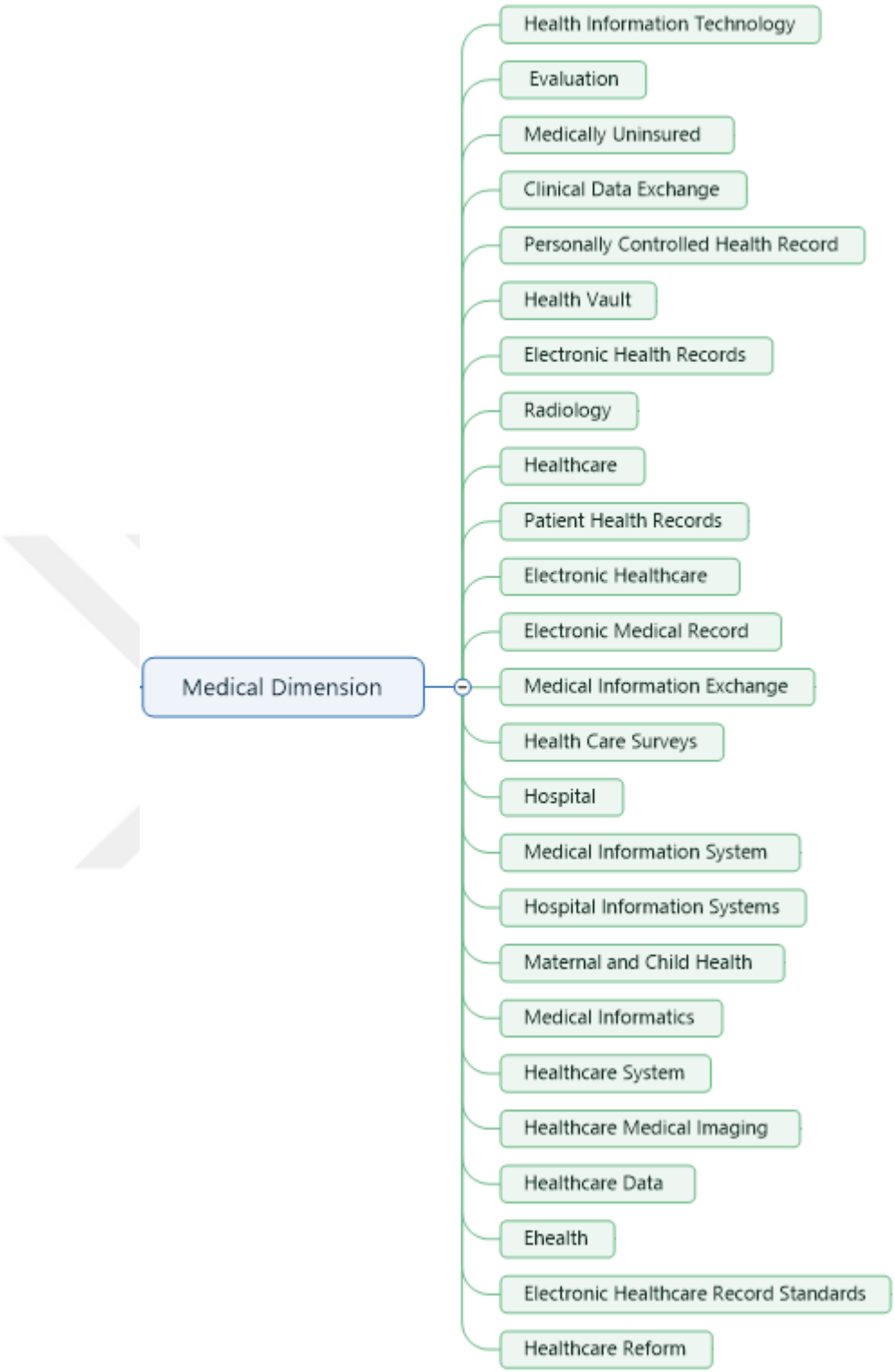

Figure 9.3: Medical Dimension Keyword Tree.

The keywords on medical dimension are mostly words that has health or medical in it. This shows that prioritizing healthcare in such a system is crucial. People should feel that the system's main aim is to increase the quality of healthcare by a patient-centric structure and not to cut costs or make profit.

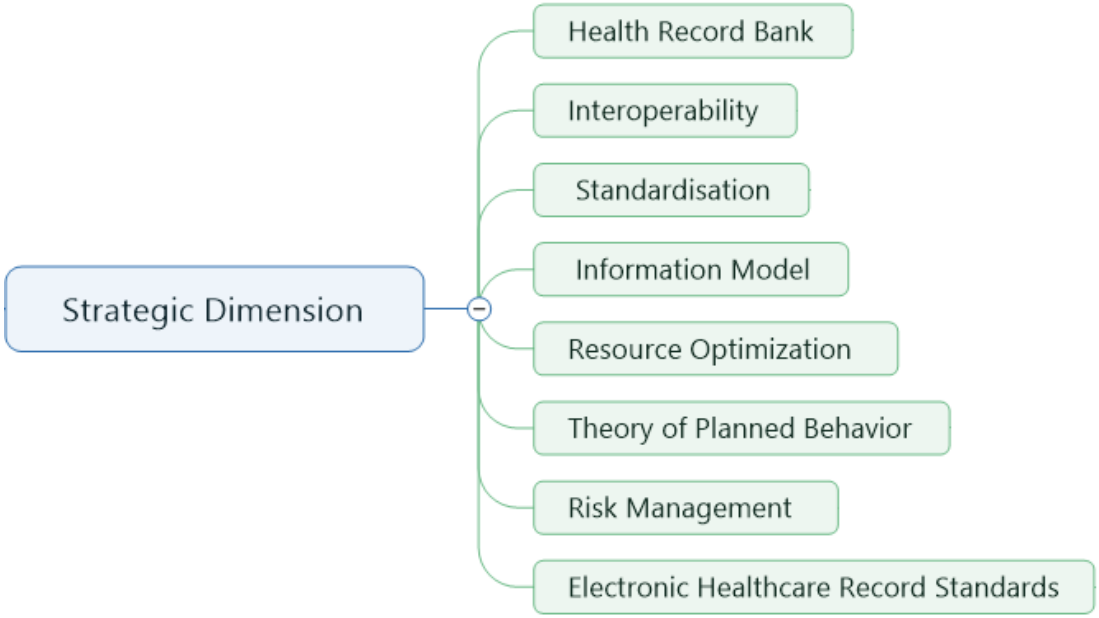

Figure 9.4: Strategic Dimension Keyword Tree.

Strategic dimension is associated with information model, interoperability and standardization. If the goal is to have a unified system, then standardization and interoperability is definitely the key. Government should impose standards and a conceptually interoperable model to medical institutions. They should have a good strategy of how they might manage the risks and come up with an optimized system.

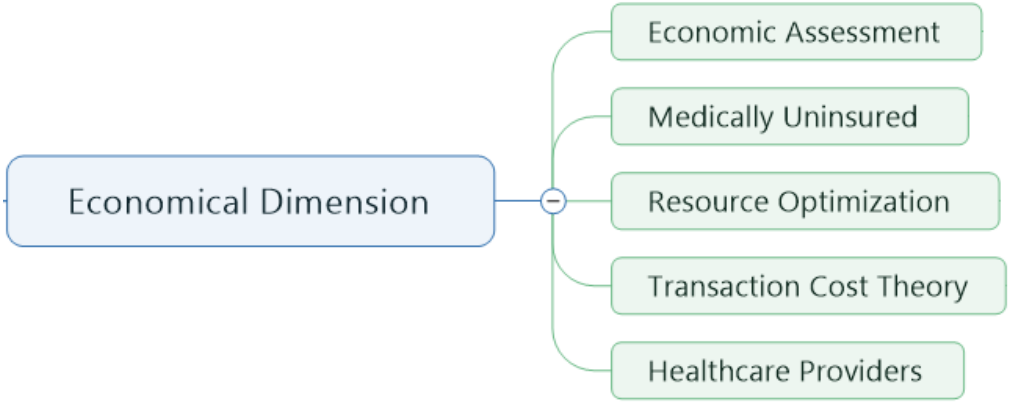

Figure 9.5: Economical Dimension Keyword Tree.

Keywords like transaction cost theory or resource optimization are classified as economical dimension. An evaluation from a financial perspective should be done and the number of cloud service providers should be chosen accordingly. The transaction costs should be optimized.

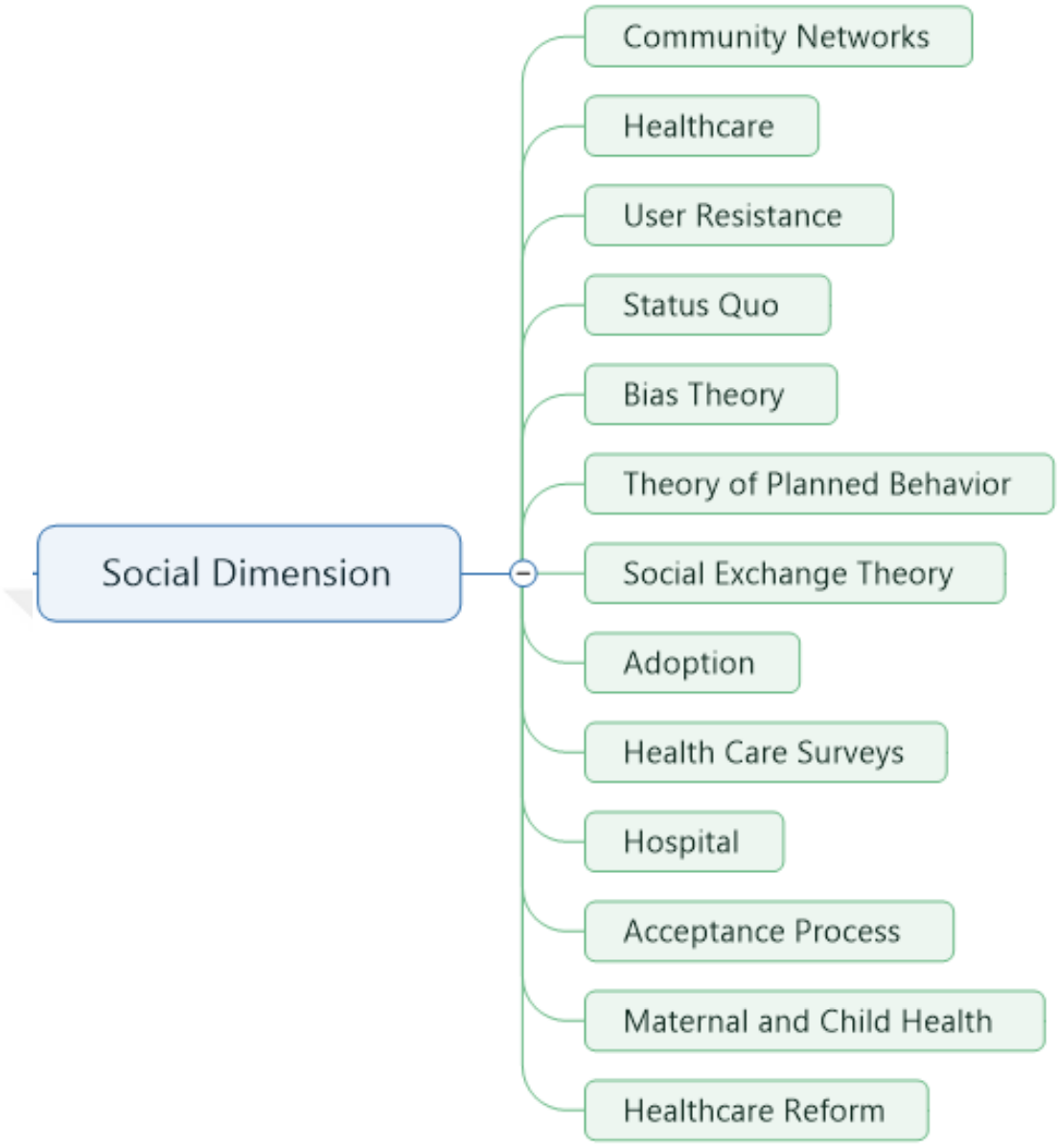

Figure 9.6: Social Dimension Keyword Tree.

Lastly, acceptance process, adoption, status quo keywords are categorized as social dimension. The user should not resist to a change like this and should be encouraged to ease the acceptance process. The users should be given an incentive to utilize the system.

Instead of building a word tree for technical dimension like others, I have created a word cloud in Figure 9.7, with the keywords that has been categorized as technical dimension. Since the solution to technical dimension is a hybrid cloud, a word cloud seemed more appropriate. It shows that there are many concepts and models researched in the literature aiming for a single goal: Improving Healthcare.

Figure 9.7: Technical Dimension Word Cloud.

In order to realize Table 9.2 gives a summary of solution proposals to the issues faced for a unified framework for a healthcare cloud as mentioned in the model. All of these dimensions should be considered when designing a unified, centralized health information system.

Table 9.2: Proposed solution table to corresponding dimensions.

Strategic	Technical	Medical	Economical	Social	Ethical & Legal
Standardization of Electronic Healthcare Records, Interoperability	Hybrid Healthcare Multi-Cloud	Prioritize Patient	Decreasing the overhead costs and workload	Encouraging User Acceptance	Clear Definition of Responsible Authorities

After implementing this system; health information and data are going to be stored in a distributed environment, whose storage is managed by cloud service providers and key management and application servers are managed by a governmental organization. Since EHRs stored electronically, accessing health related information will be easier and duplication of tests will be prevented. E-prescriptions can be managed and forgery is prohibited by digital signature and a centralized system. Patient's drug history can be clearly seen and drug abuses can be minimized. With a huge amount of data, we can call it big data, decision support systems can be utilized in diagnosing the patient. Treatment methods can be optimized and personalized. Furthermore, it would offer a tremendous database for any doctor willing to do a research, collect statistics and contribute their area of interest.

Standardizing electronic records will help medical staff at understanding the patient's condition fully and there would not be time wasted because of doctor referrals. Patients can share in the system anything they want to share such as their nutrition intake, fitness activities, heart rates during cardiovascular activities etc. They can track their condition, which would increase the chances of an early diagnosis. Patients can also manage their appointments through the system and reminders can be set such as: "Tomorrow at 10:00 AM blood sugar test, do not eat anything beforehand." Keeping track of disease trends, statistics is very beneficial for a country, since they can prevent for instance an epidemic turning into a pandemic.

This thesis proposed a novel model for a nation-wide healthcare information exchange using a hybrid multi-cloud structure with Infrastructure as a service model. It came up with six dimensions that should be considered for such a system. It mainly gave a technical model but also mentioned other dimensions and what kind of a strategy should be followed without going into depth.

This model is developed examining the system in Turkey mostly, however the ideas can be applied universally. The author acknowledges however that least developed countries can not adopt such a system because of technological constraints. It is also known to the author that transition to such a system is would not happen over a day, however the shorter the transition period, the better the system can function.

Because the scope of the research is extensive, the model can not be simulated or tested in real life as it is normally the case in narrow applications. The results however have been shown to the three experts and they verified that it would be an applicable model.

As a future work, the new identity cards with secure chips that are distributed can be examined and their use in system access can be another research topic instead of the proposed Application OTP and SMS OTP procedures. The identity cards can also be employed in emergency access, where the emergency notes can be stored. However it would require a card reader.

REFERENCES

Adler-Milstein, J., Bates, D. W., & Jha, A. K. (2013). Operational health information exchanges show substantial growth, but long-term funding remains a concern. *Health Affairs*.

Aiswarya, R., Divya, R., Sangeetha, D., & Vaidehi, V. (2013). Harnessing Healthcare Data Security in Cloud. *International Conference on Recent Trends in Information Technology*, (pp. 482-488).

Berelson, B. (1952). Content Analysis in Communication Research. *The ANNALS of the American Academy of Political and Social Science*, 197-198.

Britten, N. (2007). Qualitative interviews in healthcare. In C. Pope, & N. Mays, *Qualitative Research in Health Care.* Oxford: Blackwell Publishing Ltd.

Castiglione, A., Pizzolante, R., Santis, A. D., Carpentieri, B., Castiglione, A., & Palmieri, F. (2014). Cloud-based adaptive compression and secure management services for 3D healthcare data. *Future Generation Computer Systems*, 120-134.

Chang, I.-C., Hwang, H.-G., Hung, M.-C., Kuo, K.-M., & Yen, D. C. (2009). Factors affecting cross-hospital exchange of Electronic Medical Records. *Information & Management*.

Dennis, M. A. (2016, 04 23). *Denial of Service Attack.* Retrieved from Encyclopædia Britannica: http://global.britannica.com/topic/denial-of-service-attack

Eichelberg, M., Aden, T., Riesmeier, J., Dogac, A., & Laleci, G. B. (2005). A survey and analysis of electronic healthcare record standards. *Journal ACM Computing Surveys*.

Elo, S., & Kynga, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 107-115.

Gibbs, M., Gilreath, H., Kimbrough, M., & Vila, J. (2010). *Medical Data Exchange: A New Approach to Healthcare Interoperability.* Cisco.

Gill, P., Stewart, K., Treasure, E., & Chadwick, B. (2008). Methods of data collection in qualitative research: interviews and focus groups. *British Dental Journal*, 291-295.

Haskew, J., Røa, G., Saito, K., Turner, K., Odhiambo, G., Wamae, A., . . . Sugishita, T. (2015). Implementation of a cloud-based

electronicmedical record for maternal and child health in rural Kenya. *International Journal of Medical Informatics*, 349-354.

Hsieh, P.-J. (2015). Physicians' acceptance of electronic medical record sex change: An extension of the decomposed TPB model with institutional trust and perceived risk. *International Journal of Medical Informatics*.

Hu, Y., & Bai, G. (2014). A Systematic Literature Review of Cloud Computing in eHealth. *Health Informatics-An International Journal*, 11-20.

Hürriyet Haber. (2016, 04 30). *Bankalara siber saldırı bugün de devam ediyor*. Retrieved from Hürriyet: http://www.hurriyet.com.tr/hackerlar-iste-boyle-saldiriyor-40031701

Internet Engineering Task Force. (2016, 04 30). *Deprecating Secure Sockets Layer Version 3.0*. Retrieved from IETF: https://tools.ietf.org/html/rfc7568

Internet Engineering Task Force. (2016, 04 30). *IP Security (IPsec) and Internet Key Exchange (IKE) Document Roadmap*. Retrieved from IETF: https://tools.ietf.org/html/rfc6071

Internet Engineering Task Force. (2016, 04 30). *The TLS Protocol Version 1.0*. Retrieved from IETF: https://www.ietf.org/rfc/rfc2246.txt

Internet Engineering Task Force. (2016, 04 30). *The Transport Layer Security (TLS) Protocol Version 1.1*. Retrieved from IETF: https://www.ietf.org/rfc/rfc4346

Kalra, D. (2006). Electronic Health Record Standards. *IMIA Yearbook of Medical Informatics*.

Köse, İ., Akpınar, N., Gürel, M., Arslan, Y., Özer, H., Yurt, N., . . . Dogac, A. (2008). Turkey's National Health Information. *Proceedings of the eChallanges Conference*, (pp. 170-177). Stockholm.

Lauri, S., & Kynga, H. (2005). Developing Nursing Theories. *Dark Oy*.

Liu, C.-T., Long, A.-G., Li, Y.-C., Tsai, K.-C., & Kuo, H.-S. (2001). Sharing patient care records over the World Wide Web. *International Journal of Medical Informatics*, 189-205.

Lounis, A., Hadjidj, A., Bouabdallah, A., & Challal, Y. (2015). Healing on the cloud: Secure cloud architecture for medical wireless sensor networks. *Future Generation Computer Systems*, 266-277.

Marcu, R., & Popescu, D. (2014). Security solution for healthcare hybrid cloud platform. *Proceedings of the 18th International Conference on System Theory, Control and Computing*, (pp. 225-230). Sinaia.

Mathew, S. (2013). Cloud Computing: A New Foundation Towards Health Care. *International Journal of Innovative Technology and Exploring Engineering*.

McNamara, C. (1999). *Field Guide to Consulting and Organizational Development.* Authenticity Consulting, LLC. Retrieved from http://managementhelp.org/businessresearch/interviews.htm

Morrissette, S., Burgdorfer, J., & Shields, J. (2014). Parallels between Consolidation of the Commercial Banking and Hospital. *International Journal of Business and Social Science*, 199-206.

Network Working Group. (2016, 04 30). *The Transport Layer Security (TLS) Protocol Version 1.2*. Retrieved from IETF: https://tools.ietf.org/html/rfc5246

Neuhaus, C., Polze, A., & Chowdhuryy, M. M. (2011). *Survey on Healthcare IT Systems: Standards, Regulations and Security.* Potsdam: Universitätsverlag Potsdam.

Pamukkale University. (2016, 04 26). *Tıpta Uzmanlık Dalları*. Retrieved from Pamukkale University: http://www.pau.edu.tr/mse/default.aspx/sayfa/tipta-uzmanlik-dallari

Paquette, S., Jaeger, P. T., & Wilson, S. C. (2010). Identifying the security risks associated with governmental use of cloud computing. *Government Information Quarterly*, 245-253.

Parekh, M., & B., S. (2015). Designing a Cloud based Framework for HealthCare System and applying Clustering techniques for Region Wise Diagnosis. *2nd International Symposium on Big Data and Cloud Computing* (pp. 537 – 542). Elsevier B.V.

Radhadevi, P., & Kalpana, P. (2012). Secure Image Encryption Using AES. *International Journal of Research in Engineering and Technology*.

Richardson, J. E., Abramson, E. L., & Kaushal, R. (2012). The value of health information exchange. *Journal of Healthcare Leadership*, 17-23.

Robkin, M., Weininger, S., Preciado, B., & Goldman, J. (2015). Levels of Conceptual Interoperability Model for Healthcare Framework for Safe Medical Device Interoperability. *IEEE Symposium on Product Compliance Engineering* (pp. 1-8). IGI Global.

Rossi, G. B., Serralvo, F. A., & João, B. N. (2014). Content Analysis. *Brazilian Journal of Marketing*, 39-48.

Sane, D. M. (1990). The electronic medical record is closer than you think. *Computers in healthcare*.

Shamir, A. (1979). How to share a secret. *Communications of the ACM*, 612-613.

Vest, J. R. (2008). Health Information Exchange and Healthcare Utilization. *Journal of Medical Systems*.

Wagner, D. (2016, 04 27). *Cryptography Lecture*. Retrieved from Berkeley University:https://www.cs.berkeley.edu/~daw/teaching/cs276-s04/22.pdf

White, M. D., & Marsh, E. E. (2006). Content Analysis: A Flexible Methodology. *Library Trends*, 22-45.

Yang, J.-J., Li, J.-Q., & Niu, Y. (2015). A hybrid solution for privacy preserving medical data sharing in the cloud environment. *Future Generation Computer Systems*.

APPENDICES

APPENDIX A: Articles

APPENDIX A

Table A.1: Articles used in content analysis with keywords, authors and years.

Name	Keywords	Author	Year
The Value of Health Information Exchange	Health Information Exchange, interoperable Systems, Health Information Technology, Economic Assessment, Evaluation	Joshua E Richardson, Erika L Abramson, Rainu Kaushal	2012
Health Information Exchange and Healthcare Utilization	Informatics, Medically Uninsured, Information Dissemination, Community Networks	Joshua R. Vest	2008
Usage and Effect of Health Information Exchange	Health Information Exchange, Clinical Data Exchange, Personally Controlled Health Record, Direct Project, Health Record Bank, interconnected Patient Health Record, Health Vault	Robert S. Rudin, Aneesa Motala, Caroline L. Goldzweig, and Paul G. Shekelle	2014
Electronic Health Record Standards	Electronic Health Records, Interoperability, Standardization, Information Model	D. Kalra	2006
Cloud Computing: A New Foundation Towards Health Care	Cloud Computing, Electronic Medical Report, IaaS, PaaS, SaaS.	Saju Mathew	2013
Security Solution for Healthcare Hybrid Cloud Platform	Healthcare Hybrid Cloud, Security, Radiology	Roxana Marcu, Dan Popescu	2014
Secure Image Encryption Using AES	Security, Image Processing, AES, Encryption and Decryption	P. Radhadevi, P. Kalpana	2012
Synthetic Hardware Performance Analysis in Virtualized Cloud Environment for Healthcare Organization	Virtualization, Cloud Computing, Security, Resource Optimization, Resource Consistency, Load Testing, Healthcare	Chee-Heng Tan & Ying-Wah Teh	2013
Electronic Health Record (EHR) and Cloud Security: The Current Issues	Cloud Computing, Electronic Health Record, Security	Emmanuel Kusi Achampong	2013

Table A.1 (continued) : Articles used in content analysis with keywords, authors and years.

Name	Keywords	Author	Year
Harnessing Healthcare Data Security in Cloud	Cloud, Patient Health Records, Electronic Health Records, Privacy, Encryption, Decryption, Security.	R. Aiswarya, R. Divya, Ms D. Sangeetha, Dr V. Vaidehi	2013
Design of Secure Access Control Scheme for Personal Health Record-Based Cloud Healthcare Service	Personal Health Records, Cloud Computing, Access Control, Key Management, Bilinear Pairing	Chia-Hui Liu, Fong-Qi Lin, Chin-Sheng Chen and Tzer-Shyong Chen	2014
Privacy Preserving Secure Data Exchange in Mobile P2P Cloud Healthcare Environment	P2P Cloud, Secure Data Exchange, Pairing-Based Cryptography, Anonymous Authentication, Security Attacks	Sk. Md. Mizanur Rahman, Md. Mehedi Masud, M. Anwar Hossain, Abdulhameed Alelaiwi, Mohammad Mehedi Hassan, Atif Alamri	2015
RACS: A Case for Cloud Storage Diversity	Cloud Computing, Cloud Storage, Distributed Systems, Vendor Lock-in, Erasure Codes, Fault Tolerance	Hussam Abu-Libdeh, Lonnie Princehouse, Hakim Weatherspoon	2010
Secured Proxy Based Collaboration in Multi Cloud	Cloud Computing, Cloud Collaboration, Cloud Mashups, Cloud Proxies, Cloud Security.	Maitshaphrang Lyngdoh Mawnai, A.Selvakumar	2015
Collaborative and Secure Sharing of Healthcare Data in Multi-Clouds	Cloud Computing, Healthcare Security, Privacy	Benjamin Fabian, Tatiana Ermakova, Philipp Junghanns	2014
Enhanced Security for Multi-Cloud Storage Using Cryptographic Data Splitting with Dynamic Approach Healthcare	Cryptographic Data Splitting, Multi-Cloud Storage, Public Cloud, Private Cloud	Balasaraswathi V.R., Manikandan.S	2014
Professionals' Use of Health Clouds: Integrating Technology Acceptance and Status Quo Bias Perspectives	Health Cloud, Dual Factor Theory, User Resistance, Technology Acceptance Model, Status Quo, Bias Theory, Theory of Planned Behavior	Pi-Jung Hsieh	2015

Table A.1 (continued) : Articles used in content analysis with keywords, authors and years.

Name	Keywords	Author	Year
Factors Affecting Cross-Hospital Exchange of Electronic Medical Records	Electronic Data Interchange, Electronic Healthcare, Electronic Health Record, Electronic Medical Record, Social Exchange, Theory Transaction Cost Theory	I-Chiu Chang, Hsin-Ginn Hwang, Ming-Chien Hung, Kuang-Ming Kuo, David C. Yen	2009
Building A Generic Architecture for Medical Information Exchange Among Healthcare Providers	Medical Information Exchange, Healthcare Providers, Health Information Network	Yu-Chuan Li, Hsu-Sung Kuo, Wen-Shan Jian, Dah-Dian Tang, Chien-Tsai Liu, Li Liu, Chien-Yeh Hsu, Yong-Kok Tan, Chung-Hong Hu	2001
Electronic Health Record Adoption and Health Information Exchange Among Hospitals in New York State	Adoption, Electronic Health Record, Health Care, Health Care Surveys, Health Information Exchange, Hospital	Erika L. Abramson, Sandra Mcginnis, Alison Edwards, Dayna M. Maniccia, Jean Moore and Rainu Kaushal	2011
Physicians' Acceptance of Electronic Medical Record Sex Change: An Extension of the Decomposed TPB Model with Institutional Trust and Perceived Risk	Health Information Technology, Acceptance Process, Electronic Medical Records, Electronic Medical Records Exchange	Pi-Junghsieh	2014
Dependable and Secure Computing in Medical Information Systems	Security Reliability Access Control Medical Information System Medical Device Network	Junbeom Hur A, Kyungtae Kang	2012
Sharing Patient Care Records Over the World Wide Web	Electronic Patient Record, Hospital Information Systems, Referral Systems, Xml	Chien-Tsai Liu, Ann-Ging Long, Yu-Chuan Li, Kuo-Ching Tsai, Hsu-Sung Kuo	2001
The Standard 'Healthcare Information Systems Architecture' and The DHE Middleware	Health Information Systems Architecture, DHE Middleware, European Standard	Fabrizio Massimo Ferrara	1998

Table A.1 (continued) : Articles used in content analysis with keywords, authors and years.

Name	Keywords	Author	Year
Identifying The Security Risks Associated with Governmental Use of Cloud Computing	Cloud Computing, Risk Management, It Security, It Governance, Grid Computing, Governmental Computing	Scott Paquette, Paul T. Jaeger, Susan C. Wilson	2010
Implementation of A Cloud-Based Electronic Medical Record for Maternal and Child Health in Rural Kenya	Electronic Medical Record, Maternal and Child Health, Resource-Constrained Settings, Data Verification, Medical Informatics	Kenyajohn Haskew, Gunnar Røa, Kaori Saito, Kenrick Turner, George Odhiambo, Annah Wamaee, Shahnaaz Sharif, Tomohiko Sugishita	2015
A Hybrid Solution for Privacy Preserving Medical Data Sharing in The Cloud Environment	Privacy Protection, Cloud Storage, Integrity Check, Medical Data Sharing	Ji-Jiang Yanga, Jian-Qiang Li, Yu Niu	2015
State-of-The-Art Survey On Cloud Computing Security Challenges, Approaches and Solutions	Cloud Computing, Storage Security, Cloud Storage, Data Privacy, Cloud Security	Farrukh Shahzad	2014
Designing A Cloud Based Framework for Healthcare System and Applying Clustering Techniques for Region Wise Diagnosis.	Healthcare System, Data Mining, Cloud Computing, Open Stack, Cloud Foundry	Maulik Parekh, Saleena B.	2015
Cloud-Based Adaptive Compression and Secure Management Services for 3D Healthcare Data	SaaS, Cloud Healthcare, Medical Imaging, Adaptive Compression, Lossless Compression, Big Data	Arcangelo Castiglione, Raffaele Pizzolante, Alfredo De Santis, Bruno Carpentieri, Aniello Castiglione, Francesco Palmieri	2015
Outsourcing High-Dimensional Healthcare Data to Cloud with Personalized Privacy Preservation	Healthcare Data, Privacy, Hybrid Cloud	Wei Wang, Lei Chen, Qian Zhang	2015

Table A.1 (continued) : Articles used in content analysis with keywords, authors and years.

Name	Keywords	Author	Year
Security Issues Over Some Cloud Models	Cloud Computing, Security Factors	Passent M. El-Kafrawya, Azza A. Abdoa, Amr. F. Shawish	2015
Security in Cloud Computing: Opportunities and Challenges	Cloud Computing, Multi-Tenancy, Security Virtualization, Web Services	Mazhar Ali, Samee U. Khan, Athanasios V. Vasilakos	2015
Cloud Computing Security: From Single to Multi-Clouds	Cloud Computing, Single Cloud, Multi-Clouds, Cloud Storage, Data integrity, Data intrusion, Service Availability.	Mohammed A. Alzain, Eric Pardede, Ben Soh, James A. Thom	2012
Healing On the Cloud: Secure Cloud Architecture for Medical Wireless Sensor Networks	Wireless Sensor Networks, Healthcare Cloud Computing, Attribute Based Encryption, Emergency Access Control	Ahmed Lounis, Abdelkrim Hadjidj, Abdelmadjid Bouabdallah, Yacine Challal	2015
Integration of Cloud Computing and Internet of Things: A Survey	Cloud Computing, Internet of Things, Ubiquitous Networks, Cloud of Things, Pervasive Applications, Smart City	Alessio Botta, Walter De Donato, Valerio Persico, Antonio Pescapé	2015
Data Security in The World of Cloud Computing	Internet, Cloud Computing, Data Security, Distributed Computing, Service Providers, Virtual Environment	L. M. Kaufman	2009
A Survey and Analysis of Electronic Healthcare Record Standards	E-health, Electronic Healthcare Record Standards, Interoperability	Marco Eıchelberg, Thomas Aden, and Jörg Rıesmeıer, Asuman Dogac and Gokce B. Lalecı	2011
Parallels Between Consolidation of the Commercial Banking and Hospital Industries	Healthcare Consolidation, Banking Industry Consolidation, Healthcare Reform, Mergers & Acquisitions	Stephen Morrissette, James Burgdorfer and Jordan Shields	2013

CURRICULUM VITAE

Name Surname **:** Halil Emre GÖNEN

Place and Date of Birth **:** Istanbul / 30.07.1990

E-Mail **:** albusemreg@gmail.com

EDUCATION:

- **High School** **:** 2009, Istanbul Erkek Lisesi
- **B.Sc.** **:** 2013, ITU, Electrical Electronics Faculty, Electronics Engineering
- **B.Sc.** **:** 2014, ITU, Management Faculty, Industrial Engineering

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2013-2016 TUBITAK Graduate Programme Scholarship.
- 2014-still 2 years' experience as an IT Security Engineer.

PUBLICATIONS, PRESENTATIONS AND PATENTS:

- Çalımlı O., Türkeli S., Eken, E. G., **Gönen H. E.,** 2014: Mining Level of Control in medical organizations. *Studies in Health Technology and Informatics*, 328-332
- **Gönen H. E.**, Yanık S., 2014: District Planning for Family Healthcare Centers. *XII. International Logistics and Supply Chain Congress*, 291-300