DEVELOPMENT OF A CLOUD BASED PRIVACY MONITORING FRAMEWORK FOR THE HEALTH SECTOR

by

Manqoba V. Shabalala

(200704185)

A dissertation submitted in fulfillment of the requirements for the degree of

Master of Science in Computer Science

Faculty of Science and Agriculture

Department of Computer Science

University of Zululand

KwaDlangezwa

RSA

Supervisor: Mr Paul Tarwireyi

Co- Supervisor: Prof M.O Adigun

2014

DECLARATION

I, Manqoba V Shabalala declare that, the work contained in this dissertation has not been previously submitted in whole or in part, to meet requirements for any other degree or professional qualification at this or any other higher education institution. To the best of my knowledge and belief, the dissertation contains no material previously published or written by another person except where due reference is made.

Signature: _____

Date: _____

DEDICATION

I dedicate this work to the Lord almighty, who gave me strength and patience throughout the process of this study. Who has been my eternal rock and source of refuge.

ACKNOWLEDGEMENT

I would like to sincerely thank Prof. M.O Adigun for his unwavering support, advice, guidance and for affording me the opportunity to make this work a reality.

The success of this thesis is attributed to the extensive support and assistance from my supervisors Mr Paul Tarwireyi and the Co-Supervisor Prof M.O Adigun, I would like to particularly express my grateful gratitude and sincere appreciation to Mr. Paul Tarwireyi for his guidance, valuable advice, supervision, encouragement and kindness to me throughout this study.

To Telkom my sponsor, I cannot thank you enough for your patronage.

ABSTRACT

Cloud computing is growing in popularity due to its ability to offer dynamically scalable resources provisioned as services regardless of user or location device. However, moving data to the cloud means that the control of the data is more in the hands of the cloud provider rather than the data owner. This is a great challenge that continues to hinder cloud computing from successfully achieving its potential. This is due to the fact that with cloud computing, the storage and processing of private information is done on remote machines that are not owned or even managed by the cloud consumers. This brings about significant security and data privacy concerns that impede the broader adoption of cloud computing, which compromises the vision of cloud computing as a new IT procurement model.

In an attempt to address the aforementioned challenge, a privacy monitoring framework for the cloud computing environment was developed in this work. The design science methodology for information system was followed. The framework was evaluated using an experimental method. The evaluation of the framework mainly focused on the metrics that evaluate the satisfaction of the users' goals. The quantitative evaluation aspect entailed the usability test and questionnaires to get results. From questionnaires the statistical data was found and analyzed. The results reported in this study show that the developed privacy monitoring framework could help cloud customers monitor the privacy of their personally identifiable information in the cloud. The framework employs the developed informative event and access logs analyser which enables customers to track and comprehend how their data is handled in the cloud.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF PUBLICATIONS

Parts of the work presented in this dissertation have been published in the following conference and journal paper(s):

Shabalala, M.V., Tarwireyi, P., Adigun, M.O., 2014. Addressing Privacy in Cloud Computing Environment, in: Nungu, Amos, Pehrson, Bjorn, Sansa-Otim, Julianne. (Eds.), e-Infrastructure and e-Services for Developing Countries, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Springer International Publishing.

Shabalala, M.V., Tarwireyi, P., Adigun, M.O., 2014. Privacy monitoring framework for enhancing transparency in cloud computing, in: 2014 IEEE 6th International Conference on Adaptive Science Technology (ICAST). Presented at the 2014 IEEE 6th International Conference on Adaptive Science Technology (ICAST), pp. 1–7. doi:10.1109/ICASTECH.2014.7068093

Chapter 1

INTRODUCTION

1.1 Overview

Cloud computing has become the most intriguing enticing technology of today due to its ability to offer dynamically scalable resources provisioned as services over the internet. Cloud Computing offers services that are cost effective, flexible and easy to use. However, despite these advantages and the strong interest in cloud computing, data privacy is a key concern that hinders the adoption of the cloud (Doelitzscher, Reich, & Sulistio, 2010a; Henze, Grossfengels, Koprowski, & Wehrle, 2013; Huang & Du, 2013a; Kun, Abraham, & Yuliang, 2013). As a result, there is a need for mechanisms to allay users' privacy concerns, so that cloud computing can fully reach its potential (Pearson, Shen, & Mowbray, 2009).

The concept of privacy is a broad issue which varies widely among different jurisdictions, countries and culturally, but which is shaped most of all by legal interpretation and public expectations. Legally, privacy is a constitutional right that every citizen is entitled to. It states that an individual has a right to be left alone in any setting (Gellert & Gutwirth, 2013; Pearson & Charlesworth, 2009). From the consumer perspective, privacy refers to the right of the consumer to have his or her personally identifiable information protected and appropriately used to meet the expectations of the consumer regarding its usage and protection from unauthorised third parties or individuals (Mohamed & Ahmad, 2012; Warso, 2013).

In an organisational context, privacy entails the application of laws, policies, standards and principles such as fair information practices which represent widely-accepted concepts

concerning fair information practice. In a nutshell, privacy is about the accountability of organisations to data subjects, as well as the transparency of an organisation's practice around personal information. In an electronic world, these principles provide guidelines for the collection and use of collected data so as to protect consumer data. An emphasis is also put on openness and transparency that must be in operation during the process of collection and processing of Personally Identifiable Information (PII) (Langheinrich, 2001).

Personal information privacy normally refers to the ability of the person to control how their personal information is handled throughout the data Lifecycle; how it is collected, used, accessed, processed, stored, disclosed and destroyed (Bélanger & Crossler, 2011; Kun et al., 2013; Omran, Bokma, & Al-Maati, 2008). However, there is no general agreement with regards to what constitutes personal information. For the purposes of this research work we will utilize the definition embraced by the Organisation for Economic Cooperation and Development (OECD) that says, any information relating to an identifiable individual (Regard, 2013).

Cloud computing presents a situation whereby the data processing and storage of an organisation's delicate data are carried out on remote severs that are not overseen, or owned by the organisation. All that the consumer can see is a virtual infrastructure based on top of conceivably non-trusted physical hardware or operating environment (Itani, Kayssi, & Chehab, 2009). This raises privacy concerns since there is no attestation to the protection of the uprightness and concealment of the cloud client's personally identifiable information while it is being stored and processed in the cloud.

As a result of this, cloud computing has broadened the attack surface to include new opportunities for hackers like social engineers who may target cloud employees in order to steal private information. One example of such an attack is the following embarrassing email sent by

salesforce.com to its customers in 2007 notifying its customers of a data breach:

> Dear Salesforce.com Customer,
> As salesforce.com's community approaches one million subscribers, it has become an increasingly appealing target for phishers. We learned that a salesforce.com employee had been the victim of a phishing scam that allowed a salesforce.com customer contact list to be copied. …. Information in the contact list included first and last names, company names, email addresses, telephone numbers of salesforce.com customers, and related administrative data belonging to salesforce.com. As a result of this, a small number of our customers began receiving bogus emails that looked like salesforce.com invoices, but were not—they were also phishes. Unfortunately, a very small number of our customers who were contacted had end users that revealed their passwords to the phisher. Our support and security teams have been working with the small group of affected customers to enhance their security and with law enforcement authorities and industry experts in an effort to trace

The privacy concerns are not only true for personally identifiable information, but also for any data that requires confidentiality and integrity, such as business and trade secrets, health records and any other data protected under the intellectual property law (Neff, 2011). Report from the Institute for Health Technology Transformation shows that 20 percent of healthcare organisations have suffered a security breach, 804 breaches having occurred in more than 500 patient records between 2009 and 2013 (Pearson, Logan, Reis, Taveras, & Koerner, 2014). Improper disclosure of private information such as financial or health information could result in discrimination, loss of business and much more harm or even worse.

As a result, privacy has received increasing attention from consumers, companies, researchers and legislators (Coen-Porisini, Colombo, & Sicari, 2010). Legislative acts, such as the Health Insurance Portability and Accountability Act (HIPAA) for healthcare and the Gramm Leach Bliley Act (GLBA) for financial institutions, require enterprises to safeguard the handling of such data (Ni, Trombetta, Bertino, & Lobo, 2007). Therefore, there is a need for a mechanism that would enable cloud consumers to have an insight into how their data is handled and stored by the cloud service provider and who has access to it.

The purpose of the envisaged mechanism would be to ensure that there is transparency in the manner in which data is protected, because it is often easy for cloud service providers to claim that their services are private and secure without providing relevant or appropriate attestation to substantiate this claim. This usually misleads consumers to the point where they get to read about the data-breach on some news blog or other that their sensitive data has been compromised.

1.2 Preliminary Background

Business frameworks and procedures have become more perplexing and sophisticated; organizations are gathering an expanding measure of data, especially PII. Subsequently, organizations are attempting to keep pace with their requirements for capacity in a way that minimizes costs. Cloud computing is a model for enabling convenient, on demand network access to a shared pool of configurable resources e.g. (Network, servers, storage, application and services) that can be quickly provisioned and discharged with negligible administration exertion of cloud provider interaction (Jansen & Grance, 2011).

While providing adaptable solution for complex information technology needs, cloud computing poses additional privacy challenges to those using it. Cloud provider may end up processing data without complying with requirements of privacy laws and regulations such as the privacy and

data protection bill of South Africa. This requires that when the cloud consumer has personal information stored and processed in the cloud, the customer must be satisfied that the cloud service provider has implemented the proper technical security controls and organizational measures to secure customer's personal information against unauthorized access to personal information which may result in violation of user privacy (South African Law Reform Commission, 2005; Parliamentary Monitoring Group, 2006).

Without the necessary knowledge of the physical location of the cloud provider's server or of how the processing of personal data is configured, cloud customers consume cloud services without any information about the processes involved. This brings about the risk of data loss when external attacks successfully gain entry to the cloud provider storage and processing centres. It is also important to protect the data while data is being transferred to cloud data centres.

Furthermore, the privacy and data protection bill of South Africa states that the cloud customer must ensure that the cloud service provider maintains an information security safeguard as required by the law. The cloud service provider must also distinguish all sensibly predictable inside and outside dangers to client's personal information in its custody which may violate the privacy of the customer (South African Law Reform Commission, 2005).

In Europe, data subjects or organizations that use cloud services to process and store data containing PII, health records as well financial and credit records remain accountable to the privacy requirements of such data to comply with the European laws of data privacy, a) Health Insurance Portability and Accountability Act (HIPAA), b) Payment Card Industry Data Security Standard (PCI DSS). There are several European laws which aim to protect personally identifiable data, For example the article 8 of the European Convention on Human Rights

(ECHR) provides a right to respect for one's privacy in his home or any other environment (Glott, Husmann, Sadeghi, & Schunter, 2011).

European data protection directive also promotes data privacy through the establishment of a comprehensive data protection system that is used as a way of preserving data from unauthorized disclosure. The directive (95/46/EC) takes into account limited data collection. The OECD privacy principle mandates several principles for example, constrained gathering of information, approval to gather information either by law or by educated assent of the person whose information are constantly transformed, right to rectification and erasure and also the need of sensible security shields for the gathered information.

Numerous nations have received information insurance laws that take after the European Union model the OECD model or the APEC model. Under these laws the information controller stays in charge of the accumulation and processing of private information, actually when outsiders transform the information. The information controller is obliged to guarantee that any outsider processing information for its benefit takes sufficient specialized and authoritative efforts to establish safety to protect the information (Reed, Rezek, Simmonds, 2011).

At the point when information is transferred to a cloud, the obligation regarding securing the information regularly remains with the information authority or caretaker of that information, regardless of the fact that in a few circumstances, this obligation may be imparted to others. When the information gatherer depends on an outsider to store or process data, the overseer of the information is liable for any misfortune, or harm, or abuse of the information. It is judicious, and may be lawfully obliged, that the information overseer and the cloud provider enter into a written legitimate assertion that plainly characterizes the parts, desires of the gatherings, and the responsibility of each party.

Since Outsourcing data does not exculpate the data collector and the data subject from their legal obligations and liabilities concerning the outsourced information, this implies that the cloud client must have the capacity to control and grasp what happens to the information in the cloud and which efforts to establish safety are utilized to preserve its state of privacy (Glott et al., 2011).

A wide range of approaches exists to ensure data privacy, for example the best effort approach (Glott et al., 2011), Third-Party Audit (Glott et al., 2011), cloud provider self-service web portals and publications (Nemati & Van Dyke, 2009), Privacy-aware Role Based Access Control (Ni et al., 2007) and Service Level Agreements (SLAs). The best effort approach is the most popular. With the best effort approach operators promise to do their best to safeguard the information but do not give any guarantees (Glott et al., 2011). This is common for free services today. If executed, self-assessment is prepared and carried out by the cloud service provider based on arbitrary frameworks; for the most part it concentrates on the documentation of security policies (PricewaterhouseCoopers, 2010).

An improvement to this approach is a Third-Party Audit (TPA). This approach is the one most commonly used by cloud service providers. Cloud service providers are validated by an independent organisation to verify if they meet standards such as ISO27001 or SAS70 as well as COBIT or NIST SP800-5 (Glott et al., 2011). TPA checks the trustworthiness of information on the cloud for the benefit of the clients, and it gives the sensible path to the clients to evaluate the safety of cloud storage service and help them gain faith in it.

A TPA report assures cloud customers that the CSP followed the standards agreed upon at the time of certification. This approach is the commonly used today, it only ensures compliance at that point in time and due to its spot check approach non-compliance incidents may be missed.

The use of cloud provider self-service web portals and publications is also another approach used to enable cloud customer assess the transparency of the cloud provider. With this approach the cloud provider publishes their policies, procedures, and contractual service-level guarantees as well as the best practices and the use of standards for the cloud customer to look at. Research has shown that trust is all that is needed to have positive impact on e-commerce usage by reducing concerns, which in turn improve disclosure, reduces the demand for legislation and reduces the perceived risk (Nemati & Van Dyke, 2009).

This approach, however, does not provide information specific to each user, which makes it easier for the cloud service provider to claim to have done a lot of work to protect customer data, when this in fact may not be the case. The lack of user specific evidence makes it hard for the user to believe claims by the cloud service provider.

Privacy-aware Role Based Access Control (P-RBAC) extends the well-known RBAC model by providing full backing for communicating exceptionally perplexing privacy-related policy (Ni et al., 2007). Privacy-aware role based access control, refers to access control policies which take into account additional information relevant for controlled access to private data. It does not provide necessary information about the geographic location of the data and the security safeguards applied to it.

The Service Level Agreements (SLAs) defines the provider's commitments, however they regularly do exclude client driven checking of SLA execution or money related changes for non-performance that protect cloud users (PricewaterhouseCoopers, 2012).

The aforementioned solutions lack the capability to provide cloud customers with transparency so that the customer may have an understanding of how the data is stored and who accesses it. They deny cloud customers the liberty to choose which privacy mechanism they want to employ

to secure their sensitive private information. So there is a need to provide a solution which will provide transparency to cloud customers and to enable them to monitor their own private information without relying too much on third party audit reports which are received after a period of time has passed.

Since outsourcing data to the cloud service provider for storage and processing does not absolve the data subject from the legal obligation regarding the outsourced data. The data subject must have the capacity to control and fathom what happens to the information in the cloud and find out the efforts to establish safety to be utilized. In this manner the most extreme transparency in regards to the processes inside the cloud is obliged to empower the client to complete his legitimate commitment (Glott et al., 2011).

1.3 Statement of the Problem

The shared nature of the cloud storage infrastructure and the fact that when the data is stored in the cloud, the control of the data is more on the hands of the cloud provider rather than the data owner is a challenge that continues to hinder cloud computing from successfully reaching its capability. If not monitored closely, public cloud services can place sensitive data at risk, jeopardizing enterprise data privacy just as much as they do end users' privacy. From this stems the need for the means to allow the data owner to monitor what is happening to her/his data. Current solutions such as the best effort approach (Glott et al., 2011), Third-Party Audit (Glott et al., 2011), cloud provider self-service web portals and publications (Nemati & Van Dyke, 2009), Privacy-aware Role Based Access Control (Ni et al., 2007) and Service Level Agreements (SLAs) lack the capability to enable cloud customers to retrace in detail what happens to their data, where it is stored, who accessed it and what levels of security and privacy are applied to it.

1.4 Research Questions

1. What are the problems with the current privacy protection mechanism for cloud customers?
2. How do the existing privacy and compliance regulations affect the implementation of cloud technologies?
3. How can we design and validate a cloud based privacy monitoring framework for the health sector?

1.5 Research Goal and Objectives

1.5.1 Research Goal

The goal of this research was to develop a privacy monitoring framework to enable the users to monitor the privacy of their personally identifying information while they are being stored and processed in the cloud.

1.5.2 Research Objectives

In achieving the goal of this research work, the following objectives were developed

1) To investigate existing mechanisms and techniques which are used for monitoring privacy.
2) To assess the impact of cloud regulations towards its implementation of cloud technologies.
3) To develop cloud privacy monitoring framework that would assist in ensuring user privacy as required by current regulatory requirements
4) To evaluate the performance of the developed framework using appropriate metrics.

1.6 Rationale of the Study

Despite the unlimited benefits cloud offers, privacy and other security issues are still the key hurdle to business moving into the cloud. Businesses are unease with the idea of giving third party total control of their critical data, since cloud consumers can upload their private information such as shipping addresses, billing addresses, payment option information, and contact information and not have to re-enter it every time a purchase is made. Once this is done consumers are left in the dark on how their important data are treated thereafter. All this requires some form of personal information exchange which introduces the issue of user personal data privacy.

According to a security survey by Martens, Teuteberg, & Gräuler (2011), the lack of transparency in security control and practices employed by the Cloud Service Providers (CSPs) is a barrier to cloud adoption by enterprises. The current mechanism in place lacks the capacity to ensure cloud customers that their private information is well taken care of and those proper privacy mechanisms are implemented to secure their private information during storage and when performing some business transactions. We believe that with the proper monitoring mechanism in place, organizations may take full advantage of cloud and cut on operating cost and to invest more on job creation by venturing on other economic empowering activities and this will reduce crime.

The next section describes the research methodology used to systematically answer the research questions.

1.7 Research Methodology

In this section, the research methodology that was followed to answer the research questions raised in this work and ultimately achieve the research goal is presented. The design science methodology provides guidelines which can help computer science researchers to conduct evaluate and design research.

The Design Science Research Methodology (DSRM) provides a structured approach to conduct research. In this structural approach, the following steps are taken:

i. Development of a specific problem and provide justification for such solution approach.

ii. Establishment of state of the art of existing solution approach to know the way forward.

iii. Design and development of an artificial solution which can be a model or framework.

iv. Validation and demonstration of the efficacy of the artefact to solve the problem.

1.7.1 Literature Survey

The aim of this research method was to investigate existing research works that have been done in protecting and monitoring the privacy of the cloud consumer. Assessments of the current security and privacy regulation have been carried out to determine the impact they have on the implementation of cloud technologies. Furthermore, analysis of the security and privacy risks associated with storing and processing personal information in the cloud was carried out. Chapter 3 will report the result of this step with a comprehensive comparison of existing approaches followed by an evaluation of their strengths and shortcomings.

It is crucial to understand the scope, purpose, and limitation of the currently existing privacy preserving and monitoring approaches so that we can reuse their features to proactively allow individuals to monitor the privacy of their data. The outcome of this method will be used to craft

a solution which will enable the cloud consumer to retrace what happens to their data, where it is stored, who accessed it and what levels of security and privacy safeguards are applied to it, also to provide as much transparency to the customer as possible and do away with the "black box" approach of doing things that cloud computing is well known for. In the literature review section which is fully discussed in Chapter 3, research will be carried out to review the current state of the art so that the researcher can use the knowledge gained from the survey to craft a privacy monitoring framework.

1.7.2 Modelling

The shortcomings of existing approaches to privacy monitoring in the cloud environment as identified in the literature have given the direction to the design of the privacy monitoring framework. Based on the knowledge that was gained from the literature, the design criteria and the privacy framework were developed. Using a scenario, the proposed solution is explored and illustrated through a running prototype.

1.7.3 Prototyping

Validation of the result to ensure its applicability in the real world is a very important step of the design science methodology (Hevner, March, Park, & Ram, 2004a).

A prototype of the proposed privacy monitoring framework was implemented and evaluated as a proof of concept which demonstrates the utility and applicability of the proposed framework. Experiments were then conducted to evaluate the performance of the prototype using metrics such as: scalability, user comprehension, user control and user awareness.

1.8 Dissertation Structure

The rest of this dissertation is structured as follows:

Chapter 2 discusses the overview of cloud computing, deployment and the service model. This includes the privacy and security challenges as well as the data Lifecycle of cloud data. In addition, the basic structural definition of cloud computing, its elements that distinguish it from the utility and the likes of grid computing are discussed.

Chapter 3 presents literature survey in privacy monitoring approaches, law and regulation for cloud computing environment. Existing approaches were analysed and compared to identify their strengths and shortcomings as well as the reuse possibilities for the solution.

Chapter 4 presented the basic structural definition of privacy monitoring framework and its elements as well as the design requirements which forms the basis of the framework.

Chapter 5, one of the most significant tasks in design science (Hevner et al., 2004a) is the validation of the result to ensure its applicability in the real world scenario. Chapter 5 also presents the design, implementation and evaluation of the prototype.

Chapter 6 summarizes the thesis with emphasis on its advanced contributions in the domain of privacy in a cloud computing environment. The limitations of this thesis are reviewed here to derive future research directions that may be tackled by interested readers.

Chapter 2

BACKGROUND

2.1 Introduction

Cloud computing offers over the Internet services that lower IT capital expenditure and reduce business operating costs. These services offer on demand capacity with self service provisioning on pay per use basis for greater flexibility and agility (Sabahi, 2011; Shaikh & Haider, 2011). Cloud resources can be quickly deployed and effortlessly scaled with all procedures and applications. As a result, cloud computing helps organizations enhance service delivery, streamline IT management and best adjust IT services to element business necessities. Cloud computing can also simultaneously support core business functions and provide capacity for new and innovative services (IBM, 2010).

Section 2.2 provides a brief discussion of cloud computing and demonstration of the cloud computing architecture. In section 2.3, basic security and privacy concept are discussed, followed by section 2.4 where data security and privacy challenges are outlined. An overview of privacy legislations and regulations is provided in section 2.5.

2.2 Fundamentals of Cloud Computing

Cloud computing has been defined by different scholars in different ways, but National Institute of Standards and Technology (NIST) a widely acceptable institution in the world for their remarkable work in information technology defines cloud computing as:

"A model for enabling 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 cloud provider interaction. Cloud computing technologies can be implemented in a wide variety of architectures, under different service and deployment models, and can coexist with other technologies and software design approaches"(Jansen & Grance, 2011).

Cloud computing services are internet based and can rapidly re-deployed and effortlessly scaled with all processes, applications and services provisioned on demand paying little respect to the client area or gadget. This results in radical improvements in service delivery by the organizations taking advantage of cloud technology. All that the business need to do, is to rent an IT service and pay for the duration of the service used as opposed to a traditional approach whereby the company needs to first make capital investments on the computing hardware and software resources to perform such service in-house and have to train its employees to effectively use such resources (Kafouris, 2011).

Cloud resources are pooled together to service consumers using the multi-tenancy model to request for service of their interest at that point in time via virtual resource assignment to them as per individual needs (Espadas et al., 2013).

2.2.1 Characteristics of Cloud Computing

Cloud computing is characterized by five essential characteristics which distinguish it from the utility and grid technologies that preceded it (Reed, Rezek, Simmonds, 2011) (Cloud Security Alliance, 2009; Pauley, 2010; Carroll, van der Merwe, Kotze, 2011):

- *On-demand self-service.* Cloud service customers can order and manage cloud services without any assistance from the service providers, IT, human personal or service catalogues. All what the service provider need to do is provide a web portal and a management interface. The provisioning and de-provisioning of services and associated resources are automated by the cloud service provider. Typically, cloud customers are charged a monthly subscription or on a pay-for-what-you-use scenario
- *Measured service.* Cloud computing resource/service usage is constantly measured and controlled, supporting the optimization of resource usage, usage reporting to the customer, and enabling the cloud service provider to bill its customers based on the utilized service using what is known as pay-as-you-go business models which simply means the more you utilize the higher the bill.
- *Ubiquitous network access.* Cloud services are accessed via the network (usually the Internet) anywhere, using standard mechanisms and protocols that promote use by heterogeneous thin or thick client platforms such as smartphones, tablets, laptops, and office computers.
- *Rapid elasticity.* Cloud Resource/service can be scaled up and down rapidly and elastically, in some cases automatically. Providing the cloud users with an illusion that the capabilities available for provisioning are unlimited and they can therefore be appropriated in any quantity at any time.
- *Resource pooling.* Computing resources used to provide the cloud service are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned

2.2.2 Cloud Deployment Models

Cloud computing offers three main different deployment models, these deployment models depict where the cloud is located and for what purpose as well as who owns and manage the infrastructure (Mell, 2011). Table 2-1 summarizes the various features of cloud deployment models.

- ***Public Cloud***

The Public cloud infrastructure is available to anyone with internet access since it is provided as a service over the internet and users are charged on a pay-per-use basis or by subscription. No upfront payments needed for hardware and software purchases (Baca, 2010; Council, 2011; Infrastructure, 2010; Mell & Grance, 2009; Bhuller, 2009; VMware Inc, 2009).

- ***Private Cloud***

The private cloud is the type of the cloud that is owned and solely used by a single organization, "therefore" provides a better security. This has similar services to that a Public cloud except that, it gives the owner greater flexibility and control since it resides on premise (Baca, 2010; Council, 2011; Infrastructure, 2010; Mell & Grance, 2009; Bhuller, 2009; VMware Inc, 2009). However, upfront capital invested is required to purchase and maintain all the software and infrastructure.

- ***Community Cloud***

In this type of cloud deployment, the infrastructure of the cloud is shared by several organizations with same requirement to realize their individual goals whilst taking advantage of cloud benefits. It can be managed by a third party and reside off premise (Armbrust et al., 2010).

- ***Hybrid Cloud***

This is the type of cloud deployment where the infrastructure is composed of multiple clouds, Private, Public, or Community cloud, thus the organization has to provide, maintain and manage some computing resources in-house and outsource some (Reed, Rezek, Simmonds, 2011). It provides some form of control to business critical application and whilst taking advantage of public cloud desirable features of scalability and low cost (Woolf, 2013).

Table 0-1 Summary of the various features of cloud deployment models

Deployment Model	Infrastructure Located At	Infrastructure Owned By	Managed By Infrastructure
Public Cloud	Off-premises	Third party provider	Third party provider
Private Cloud	On-premises	Organization	Organization
Community Cloud	Off-premises Or on-premises	Several Organizations	Third party or Organization
Hybrid Cloud	Both off-premises An on-premises	Both the organization And Third party Provider	Both Third party provider and Organization

2.2.3 Cloud Computing Service Models

Cloud computing offers different service models which together have come to be known as the SPI model (Pearson, 2009). SPI denotes: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a service (IaaS). However, there can be other forms of service models provided such as database as service and many other more. That is why cloud service models are sometimes referred to as XaasS, where X is any kind of service that the service provider might decide to offer.

- ***Cloud Software as a Service (SaaS)***

This provides consumers the capability to access provider's applications running on the cloud via ubiquitous interfaces. However the consumer does not control the underlying cloud infrastructure (operating system, storage, and network).

- ***Cloud Platform as a Service***

Platform as a Service is a service which provides consumers with a platform such as operating system, applications, services as well as development frameworks to develop software applications created by using programming languages supported by the provider. It helps to relieve system developers from dealing with platform administration and help the developers to collaboratively write the code and deploy their applications on the cloud (Mell & Grance, 2009).

- ***Cloud Infrastructure as a Service***

Provides the consumer with computing resources to run software, process and store data, as well as to make use of CSP networks and other fundamental computing resources (Reed, Rezek, Simmonds, 2011). This redundant and shared computing infrastructure is accessible through web browsing technologies.

2.3 Security and Privacy

Data security and privacy challenges are the most cited substantial obstacles to the broader adoption of cloud computing technology across the globe. This owes to the fact that cloud computing acts as a big black box, making it close to impossible for its client to know what is going on within.

2.3.1 Overview of Privacy and security

Despite the benefits that the cloud presents, cloud computing technology is faced with a variety of legal and technological challenges. Security and privacy are amongst the major challenges as identified in the literature (Doelitzscher, Reich, & Sulistio, 2010b; Huang & Du, 2013b; Mather, Kumaraswamy, & Latif, 2009; Pearson, 2009). These challenges are attributed to the lack of proper security control policies and weaknesses in security safeguards in cloud deployments (Prasad et al., 2011a). The rampant accidental and deliberate data breaches have become synonymous with the cloud and have led to a widespread recognition of the privacy risks in cloud deployments.

Fraudulent hackers have possessed the capacity to break into database systems of payment card processors, doctor's facilities, online retailers, schools, money related establishments, and even government agencies to increase extensive scale access to client individual data for the reasons of extortion.

About 1.7 million top secret NSA documents were leaked to several media outlets including the guardian and Washington post over the past year, by Edward Snowden the former NSA contractor. These documents also revealed governments of the United Kingdom, Australia, Canada, and New Zealand spying operational details. It also revealed how the USA government uses internet programs for surveillance operations to spy on domestic and foreign internet traffic, email and phone use (Lewis, 2014).

This is a downright violation of the right of individuals to be free from unwarranted surveillance. This is the main case of the decade, which is reshaping how we have been looking at privacy. Privacy is a foundation for good democracy and mass surveillance is a threat to democracy and

the rule of law. Mass surveillance erodes essential pre-requisite to the exercise of individual freedom which in turn weakens the constitutional foundations on which democracy have been traditionally based. This goes to show that much work still needs to be done by internet companies to protect user privacy.

In 2012, more than 6.5 million LinkedIn customers' passwords were stolen from its site, these passwords were hashed with unsalted SHA1 and within 24hrs, close to 70% of these passwords had been cracked. The hackers later on posted the passwords on the Russian hacker forum for all to see (Perlroth, 2012).

South Africa has been no exception to these attacks by hackers; the SecurID token system seed-key warehouse was breached in March 2011 where the keys for the 2-Factor authentication system were stolen in order to gain secure access to corporate and government records using the replicated hardware tokens. Sony lost an estimated 77 million records (Swart, Grobler, & Irwin, 2013) containing personally identifiable information of PlayStation users. Several lawsuits and complaints have been launched against Sony and it is now facing a hard time trying to repair its image and regain trust from its market (Spiegel, 2011).

Adobe Systems also revealed that in October 2013 their corporate database was hacked and some 130 million user records were stolen (Goodin, 2013). Many of these leaked files included email addresses and password that were reported to have been belonging to American military and government officials from the Department of Homeland Security (Weitzeekorn, 2012). This then triggered a flood of legal actions.

In January 2014 GMR Transcription, an Orange County, CA-based transcription and translation company inadvertent exposed people's medical data, which it shared with a third party data

service provider called Fedtrans. The Federal Trade Commission (FTC) is an independent agency of the United States Government responsible for promotion of consumer protection and the elimination and prevention of anticompetitive business practices. FTC faults GMR for failure to adhere to Section 5 of the FTC Act, which prohibits unfair or deceptive acts or practices in or affecting commerce.

The FTC findings in this case where that the GMR failed to carry out its duties when contracting with Fedtrans data service provider which was to perform due diligence and to sufficiently check that their service provider, executed sensible and proper efforts to establish safety to ensure individual data in sound and transcript documents on Fedtrans' system and machines utilized by Fedtrans' typists and obliging that the these files be securely stored and transferred in a securely manner to their typist (Solove, 2014).

A lot of effort has been channelled to combat diligent worries regarding the nature of cloud computing that are hindering force towards its adoption and that will in the end trade off the vision of cloud computing as another IT acquisition model (Chow et al., 2009).

A wide range of legal and regulatory frameworks has also been developed with an aim of improving the security and privacy in cloud computing. These frameworks on their own also pose a number of challenges, which includes the viability of legal regimes which impose obligations based on the location of data, and the establishment consent from the data subject for collection, usage and redistribution of personally identifiable information, accountability and openness.

Cloud computing providers can store data in any of its data centres around the globe. Moreover, the cloud consumer does not necessarily know the location. To alleviate security and privacy

concerns, this transparency is mandatory, because it plays a major role in determining if the information is not in contravention with compliance, regulatory frameworks and that it is not violating any user privacy (Manan, Mubarak, Isa, & Khattak, 2011). Personal information describes the intimate facts that relate to an individual. This information is very sensitive and when disclosed it could result in harm to the individual whose privacy has been breached. This kind of information is normally referred to as personally identifiable information. PII can be easily traced back to the particular individual. These may include things like.

- Social Security Number.
- Name and Identity number.
- Credit Card Number or Consumer purchase history.
- Phone number and Address Name.
- Passwords.

However, the Federal Trade Commission (FTC) has effectively redefined PII to include what was initially considered machine data such as the device IP address and Universally Unique Identifier (UUID) on the, 2010 Protecting Consumer Privacy in an Era of Rapid Change report, when it said that:

"…the proposed framework is not limited to those who collect personally identifiable information ("PII"). Rather, it applies to those commercial entities that collect data that can be reasonably linked to a specific consumer, computer, or other device."(Dennedy, Finneran, 2014)

Some elements of personal information are more sensitive than the other and therefore such information requires an extra level of protection and a higher duty of care. This sensitive information as considered by some laws and regulations includes but not limited to:

- Information on health conditions

- Religious beliefs, Racial or ethnic origin
- Financial Records
- Trade union membership
- Sexual preferences.

Data privacy is rather a relation between four key areas, firstly, data gathering /collection and distribution of personally identifiable information, secondly technology used to store, manage, and protect personally identifiable information, thirdly the legal and political issues surrounding the personally identifiable information, and lastly privacy expectation of the individual.

Personal data should be overseen as a feature of the information utilized by the framework. It ought to be overseen from the time the data is imagined through its last mien. Assurance of individual data ought to consider the effect of the cloud on each of the accompanying stages as definite in Figure 2.1

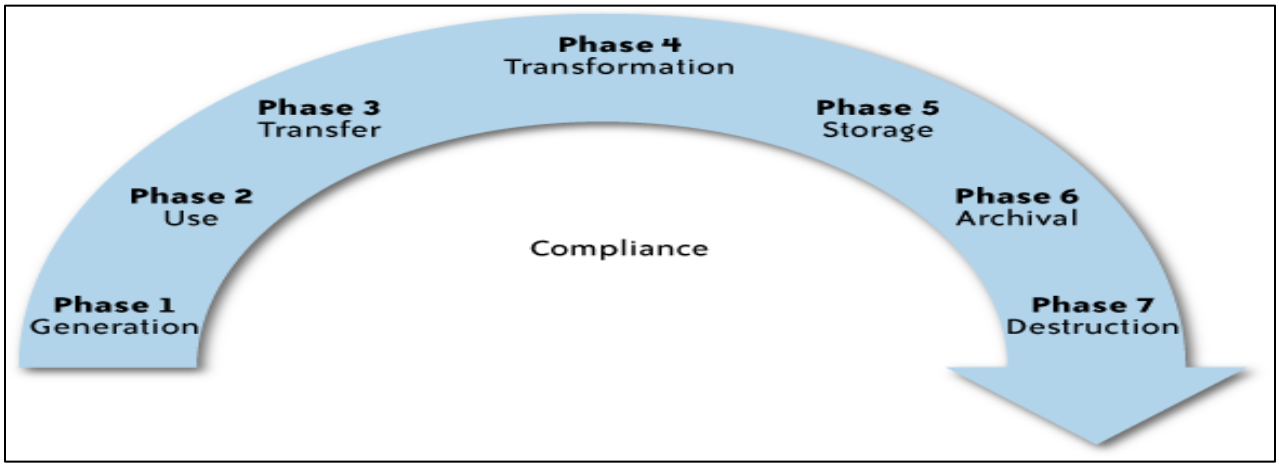

Figure 0.1 KPMG data life cycle (Programming4us, 2010).

The phases are as follows:

Generation

Data Generator is the creation or procurement of the data required to help business, operational or administrative necessities. Throughout this phase, it is crucial to ensure that there is a governance structure in place to ensure the management and protection of PII during its life cycle when stored or processed in a cloud computing environment.

1) **Use**

The generated data can be used either within or outside the collecting organization, e.g. in a public cloud. It must also be ensured that the data is appropriately used in consistent with the purpose for which it was initially collected for and in accordance with the commitments the collecting organization made to the data subjects.

2) **Transfer**

When transferring the generated information over the network into the cloud it must be protected appropriately and the most commonly used approach of achieving this is via encryption as is required by various laws when transferring PII. The data must first be encrypted before being sent over any communication medium. Access control mechanisms must also be put in place in order to control who gets access to PII

3) **Transformation**

The integrity of the data and use limitations must be maintained at all times when data are being processed in the cloud computing environment. Privacy protection methods must also be applied during this phase to ensure the continued preservation of PII. This must be done in compliance with legislative, legal and regulatory privacy requirements.

4) Storage

Data storage identifies with the maintenance of information in a suitable way to help business, operational or administrative prerequisites. Appropriate controls over access to PII must be established, so that only individuals with valid credentials can be able to access it. The CSP must also provide ways to maintain the data integrity, availability, and confidentiality to ensure data is not lost or accessed by unauthorized personnel. During this phase CSP must be on the lookout for the new legislative, legal, regulatory and privacy requirements.

5) Archival

Data retention period must be clearly specified in the contract between the cloud customer and the CSP, the cloud service provider must comply with these specific requirements that dictate how long the PII should be stored and archived. The cloud service provider must also clearly state the type of media in which information will be stored in, if whether the device is portable or not. Since portable device are susceptible to loss who controls the device must be clearly stated - who owns them, are they a cloud service provider or a third party?

A data recovery mechanism is required to be in place for each cloud consumers so that in the event of data loss or system failure, his data can still be recovered by the cloud service provider as required by the privacy industry standard.

6) Destruction

Information discarding includes the suitable evacuation or filing of information that is no more needed to meet statutory necessities, help business or operational prerequisites. The devastation of information should be carried out in a secure manner that would guarantee no reconstruction of data can be carried out. Data destruction policies are not applicable at the end of the data

lifecycle in cloud computing context because of cloud computing characteristic of resource pooling and elasticity, which means the physical disk cannot be destroyed since it is shared by multiple tenants. However media sanitization and personal information deletion must be carried out.

2.4 Privacy and Security Challenges

Cloud computing services depend mainly on virtualization, web application and provide ubiquitous network access to its customers, since cloud services are accessed via the network using the standard network protocols. This network is mainly internet which is well known for being untrustworthy (Grobauer, Walloschek, & Stöcker, 2011). Web application implements session handling and many session handling implementations are vulnerable to session ridding and session hijacking while HTTP by nature is a stateless protocol whereas web application itself requires some notion of session state.

The skilled hacker may successfully escape from the virtualised environment and inflict harm on the cloud infrastructure by altering some configuration settings and gaining access to restricted system. The on demand cloud characteristic is achieved through the management interface. The interface is accessible to cloud service users and is vulnerable, given that authorized access to management interface is bound to take place in cloud systems.

This is contrary to traditional systems, in which the management interface is only accessible to limited system administrators. The management interface is also realised using web application so it often shares the vulnerability of web application, and it can also suffer from denial of service attacks since the most common authentication method is that of password and username which locks the user out after several unsuccessful attempts.

Resource sharing by cloud consumer presents a situation whereby bits of data can be recovered from the last consumer who was using the resource. This is owing to the cloud computing characteristic of resource pooling, which entails that cloud resources are relocated to the next available users once the one who was using it is done with it. In case of adversary attacks and breaking through the encryption of a database operated by a CSP in a multi-tenant service, chances are that attacker might have the capacity to take the information of handfuls or several diverse business buyers put away on that database.

Cryptographic techniques are generally used by cloud computing service providers to solve security related problems. Whoever, cryptanalysis advances day by day and they can easily render a well-known cryptographic algorithm insecure when a flaw is discovered in the algorithm. This flaw is used by the perpetrators to turn what used to be a strong encryption into a weak encryption or, no encryption at all, since more methods of breaking cryptographic mechanisms are often discovered.

The issue of poor key management is also a very troubling challenge. A study conducted by the European network and information security agency revealed that cloud computing infrastructure requires management and storage of many different kinds of keys as a result of many virtual machines (Grobauer, Walloschek, Stocker, 2011).

These vulnerabilities in core cloud computing technologies are the main contributing factor to customers' reluctance to move into the cloud. Figure 2.2 depicts the cloud security challenges/issues from a survey conducted by IDC in august 2008.

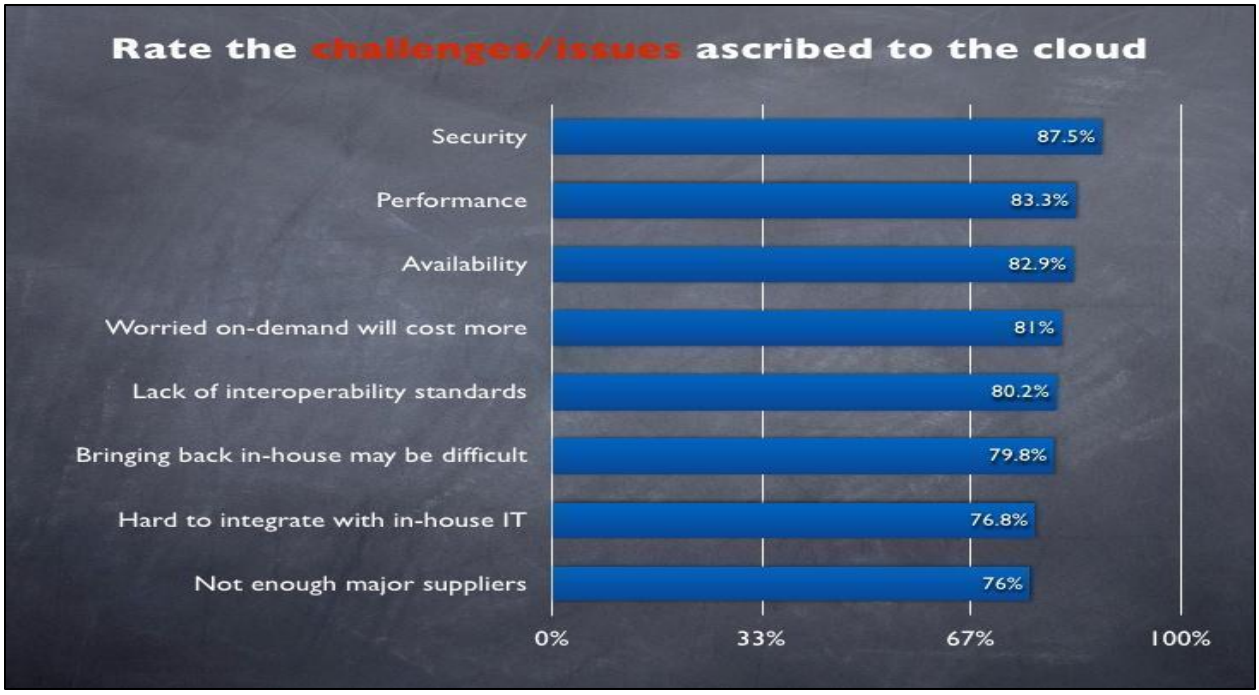

Figure 0.2 Cloud Security Challenges/Issues (Emerson Network Power, 2010)

Most customers are still reluctant to release their data and services into the cloud, mostly because of lack of user control, unauthorised secondary usage, multi-tenancy, inadequate monitoring and compliance.

2.4.1 Lack of user control

It is a cloud customer legal obligation to control and comprehend what happens to the data outsourced to any data processing and storage system, and to ensure that the proper security measures are put in place to protect private information (Glott et al., 2011; Reed, Rezek, Simmonds, 2011). However, with cloud computing there is no mechanism in place to enable cloud customer to fulfil their legal obligation (Pearson, Siani, Yee, 2013). Since applications, data and resources are located within the cloud computing service provider, cloud consumers have to rely on the cloud provider to ensure data security and privacy. User identity management and access control rules are handled by the cloud service provider, leaving the cloud customer

with no control whatsoever over data outsourced to the CSP.

Furthermore, consumers have to rely on the cloud service provider to ensure monitoring, repairing of service/resources and ensuring resource availability as well as the management of security policies and enforcements. Thereof since the consumer data is processed in the cloud on machines they do not own or have control over. The cloud service provider may not even comply with a request for deletion of data once the cloud consumer no longer requires the cloud service provider services, so the data owner is also stripped of control over the destruction of the data.

2.4.2 Unauthorized Secondary Usage

A lot of revenue can be generated from selling private and confidential data due to the sensitive nature of many of these pieces of information. Any piece of information that can be directly or indirectly linked to the identity of a specific person, therefore this information can be sold by the cloud service provider for advertising and marketing purposes.

This information is very valuable to retail and insurance companies since it enable them to target the right consumers for their products, and generate much needed revenue. This practice is against the law since it violated the very same privacy laws. Identity theft is the most common hazard presented by the cloud paradigm, where customer personally identifiable information may be used to conduct illegal activities online (Mowbray, Pearson, & Shen, 2012).

2.4.3 Trans-border Data Flow and Data Proliferation

Cloud service providers have data centres all over the world, and the data backups are stored all over in case of natural disaster taking place in one country. This data proliferation is not controlled or managed by the data subjects, and such data are subjected to the laws of the country

where the copies are stored or processed (Subashini & Kavitha, 2010).

The data owner has no control over this kind of setting, the court can also grant a subpoena to access such information and disclosing it to unauthorized third parties. Any data put away in the cloud inevitably winds up on a physical machine possessed by a specific organization or individual spotted in a particular nation, and it might be liable to the laws of that nation where the physical machine is placed.

Accordingly the area of information in the cloud has noteworthy consequences for the security and secrecy securities of data and on the commitments of the individuals who prepare and/or store the data. If your data is located under U.S, then the Patriotic Act will have an impact on your data since this act allows the FBI to access any business record that is stored on the service providers data centre given under the possession of the court order, so the cloud users' privacy can't be protected (Warren & Brandeis, 1890; Lovalls, 2010). Besides the Patriotic Act, Snowden has revealed that the US has always had back door access to data. Prism is the name of the program used by NSA to collect material, including search history, the content of emails, file transfers and live chats to the systems of Google, Facebook, Apple and other US internet giants (MacAskill, 2013).

2.4.4 Multi-tenancy

Multi-tenancy is an architecture in which a single instance of the software serves dozens or hundreds of users/customers at the same time (Leinwand, Yim and Allan, 2014; Takabi, Joshi, & Ahn, 2010). This presents data recovery vulnerabilities in data storage resource since it might be possible to recover data which was written by the previous user. If there is a data breach by a skilled hacker all the different PII and business customers' data stored on that disk will be at risk.

An attacker can legitimately be on the same physical machine as the target, since cloud tenants share the same physical infrastructure. As they have opposing goals and requirements, catering for each and every tenants need is impossible and this leads to an even lower trust in the cloud. The tenants are logically isolated, but physically integrated; tenants are application instances of the various organisations which all form a list of cloud customers.

2.4.5 Inadequate Monitoring, Compliance and Audit

Cloud service providers do not provide its customer's permission to conduct their own compliance audits with regards to the manner in which the cloud service provider processes customer's data (Massonet et al., 2011). This makes the customer's previous investments in security certification to be at risk, especially in the case whereby the cloud service provider doesn't provide its customer with the relevant evidence of their compliance with the relevant requirements as required by privacy regulation and data processing standards (Pearson, 2012). Furthermore, it makes it difficult for the cloud customer to evaluate how cloud computing affects compliance with internal security policies.

A wide range of legal and industrial regulatory frameworks have been developed to improve security and privacy in cloud computing. However, the harmonization of such framework is needed to effectively achieve the high level of security and privacy in the cloud environment. Complying with these frameworks is prominent to organizations utilizing cloud services since they serve as an attestation to proper data management and increase client base when result of compliance audit are published. Compliance has a huge effect on the notoriety of a business and it helps building client trust.

Most of these frameworks require that productive applicable rules be established for

accountability and transparency helping an abnormal state of security and security in information insurance guidelines and development of rupture warning administrations to cover cloud computing providers.

A. Industrial Regulatory Frameworks

Industrial regulatory frameworks includes, but not limited to: Health Insurance Portability and Accountability Act, Payment Card Industry (PCI) Data Security Standards (DSS) and Gramleachbliley Act (GLBA), ISO/IEC 29100:2011.

I. Health Insurance Portability and Accountability Act (HIPAA)

HIPAA was enacted by the U.S. Congress in 1996; it affects other countries and their citizens that do health care/insurance business with U.S. HIPAA regulates the management, collection, storing and distribution of health information records on clinics, hospitals, doctor's offices and even on insurance companies to mitigate the risk of privacy violation which may result from it (US Health & Human Services Dept, 2003). The benefits that come with the research conducted using data harvested from patient records must be weighed against the outcome of the unauthorized access and disclosure of such information to third parties whose interests might be malicious ones (Canadian Institutes of Health Research, 2003).

This regulation seeks to strike a balance between the needs of the advance public health and the individual right to privacy. Furthermore, it specifies the technical and administrators safeguards required for business process and information system used to process the health identifiable information.

II. Payment Card Industry (PCI) Data Security Standards (DSS)

PCI is a comprehensive standard for enhancing the information security controls surrounding

payment card transactions. It also helps to alleviate the risk of associated credit card information fraud by putting policies in place which stipulates that the cardholder data should be protected during storing or when transmitted and that a secure network should be maintained at all times by the organization dealing with such data.

A thorough set of security prerequisites has been set that all dealers' point-of-sale vendors and money related establishments must conform to keeping in mind the end goal to hold their entitlement to utilize the payment system. These requirements state that the program for maintaining a system security and a vulnerability management should be implemented as well and the access control methods to limit access to credit card information. All the cardholder data should be encrypted at rest and in transit (Ngugi & Dardick, 2010).

Furthermore the requirements state that the network and system monitoring mechanism should be put in place and tested every regularly. For any business which make use of such payment card information must comply with this standard so that it may not suffer restrictions on their use of payment card services.

III. Gramleachbliley Act (GLBA)

A standard developed for financial institutions with an aim to protect consumer privacy, it compels the financial institution to provide its customers with the details on what information is collected about them and who it is shared with as well as the security safeguards put in place to protect such delicate information. The Requirements include:

- Physical Access controls on systems containing customer data
- Use of encryption
- Monitoring for abuse, attacks, and intrusion as well as Incident response plans

The nature of physical access violation compliance of the business determines the damage level of enforcement, and other factor that the regulators may take into account. However, this varies from business to business (Kendrick et al., 2013; Restrepo, 2005).

IV. ISO/IEC 29100:2011

(ISO/IEC JTC 1 and the ISO and IEC, 2013) The International Standards Organization/ International Electro-technical Commission (IEC), describes privacy safeguarding that must be taken into consideration when dealing with PII either on personal capacity or as an organization. This standard stresses that privacy controls are mostly required when using or designing information and communication technology systems to store or processes PII. This standard focuses on personally identifiable information. In section 4.5.1, the standard says controllers should be aware of all legal and regulatory requirements. The organizations that process PII are required to document their internal and external privacy policies. It further stresses that security risk management process must be put in place so that privacy controls can be reviewed and reassessed over time (section 5.11).

B. Privacy legislations and regulations

The data stored by the cloud computing service provider is subject to varying laws, than the data stored in servers owned by the organisation, for instance, the Government can compel the cloud computing service provider to handover information in their storage, rather than when the data is stored in the house (Lovalls, 2010).

Certain laws impose some data security requirement on PII that the data owners and processors have to adhere to especially when PII is outsourced to the external entity for storage and processing. When IT services are outsourced to the cloud, the data consumer has no way of

ensuring the compliance with compliance laws since the data is stored on remote servers that consumer does not own. Usually, when data breaches take place, data owners are not notified of the breach immediately, which leaves them in the immediate danger of having their private data used to commit fraudulent acts.

I. State Laws

Data breach notification requires notification of the breach (hacking, lost data, unauthorized access) to be made to all affected individuals when sensitive information is lost or exposed in that manner that can provoke identity theft. These notifications of data breach need to be made in reasonable amount of time after the breach take place. However, in certain circumstances this action may be delayed due to investigations that the government might deem necessary to carry out in relation to the breach (Lovalls, 2010).

Many states impose data security requirements on entities operating on the state or store data of state residents. The contract between the cloud computing service provider and the cloud customer must explicitly specify where the data will be stored and what security standards the cloud provider will adhere to and who is going to be liable for the financial implications of a data breach if it happens.

Although there is currently no specific data protection legislation in forced in South Africa, the following laws are relevant:

II. Electronic communication and transaction (ECT)

ECT originally enacted in 2002 provides principles set out to protect the electronically collecting personal information. Protection of Personal Information Bill is an amendment to this act (SA Justice Dept, 2009), ECT principles are aimed at protecting private personal information and

include an obligation to obtain express permission for the collection and processing of personal information, disclosure in writing of the specific purpose for which personal information is requested, record keeping, non-disclosure obligations and requirements to destroy obsolete personal information (The presidency, Government Gazette No.23708, 2002).

In case of electronic, unsolicited messages intended for direct marketing such as SMS's and spam email the ECT Act prescribes that the sender must provide an option for the consumer to opt out or cancel the subscription altogether and be removed from the mailing list and should the consumer wish to know how the marketer got hold of their personal information the sending party should disclose the source from which the consumer's personal information was obtained.

III. Protection of Personal Information Bill (POPI)

POPI Bill is a comprehensive South African data privacy law which gives expression to the right to privacy provided for in the Constitution, which entails right to protection against an unlawful collection, retention, dissemination and use of anyone's personal information, the main focus of this bill is to regulate the manner in which personal information may be processed, which largely in line with the international standards that prescribe the minimum threshold requirements for the lawful processing of personal information.

It was passed in the South African parliament on the 22nd august 2013 (Government Gazette No. 37067, 2013). This Bill applies to all public and private bodies (defined jointly as a "responsible party" in the POPI Bill). It legislates the computerized or non-computerized processing of individual data in South Africa by or for a responsible party, paying little respect to whether the party is a permanent resident of South Africa.

2.5 Privacy Regulation Principles

Fair Information Practices are a set of universally perceived practices for tending to the protection of data about people. Information privacy is a subset of privacy. FIPs are vital in light of the fact that they give the underlying strategy to numerous national laws tending to protection and information assurance matters. The core principles of privacy addressed by these principles are:

Notice/Transparency/Awareness- ensures no secrete data collection, it stresses that the data collector must provide the consumers with clear and conspicuous notice of their information practices before any information is collected from them and let them know how they collect it; provides information about the collection of personal data to allow users to make an informed choice.

Consent/Choice – data collectors should give individuals a choice as to how they wished their information to be used, including the secondary usage of data beyond the reason it was initially collected for.

Access - allows individuals reasonable access to review the information that has been collected about them and ensure that it's accurate and up to date.

Information Protection - requires organizations to take reasonable steps to protect the security, quality and integrity of the personal information they collect.

Integrity/Security - Information collectors must ensure that there are proper security safeguard put in place to protect the collected data.

Enforcement - holds organizations accountable for complying with FIPPs

Whereas the Generally Accepted Privacy Principles (GAPP) framework was developed from a

business perspective to protect the privacy of personal information it is used by Certified Public Accountants (CPAs) in the United States and chartered accountants (CAs) in Canada, both in industry and in public practice, to guide and assist the organizations they serve in implementing privacy programs that addresses privacy risks and obligations, and business opportunities. GAPP, previously known as the AICPA/CICA Privacy Framework, it realizes its single privacy objective by employing 10 privacy principles, namely Management, Notice, Choice and consent, collection, access, disclosure, quality, monitoring, enforcement, use retention and disposal.

Unlike GAPP, the Asia-Pacific Economic Cooperation (APEC) forum was established in 1989 to further enhance economic growth and prosperity for the region and to strengthen the Asia-Pacific community. The APEC Privacy Framework is aimed at encouraging the development of appropriate information privacy protection mechanisms that will ensure the free flow of information in the Asia Pacific region. This Framework seeks to balance information privacy with business needs and commercial interests while paying special attention to the cultural diversity that exist within the members of these economies. It also give clear direction and bearing to organizations in APEC economies on regular privacy issues and the effect of privacy issues upon the way authentic businesses are conducted.

The comparison in table 2-2 shows various privacy laws and regulations that every cloud service provider should adhere to. These privacy regulations discussed in this section have much in common in that they all adhere to the principles of the Fair Information Practice and Directive (95/46/EC). As depicted in the comparison table almost all these regulation provides consent to access personal information that is outsourced to the cloud service provider (Privacy Rights Clearinghouse, 2005).

Table 0-2 Various privacy laws and regulation

	Fair Information Practice	GLBA	PCI- DSS	GAAP	HIPAA	Directive (95/46/EC)	APEC	OECD Guidelines
Access	√			√	√	√	√	√
Notice	√	√		√	√	√	√	√
Consent	√	√		√	√	√		√
Security	√	√	√	√	√	√	√	√
Enforcement	√	√	√	√	√	√	√	√
Integrity	√	√	√	√	√	√	√	√

The cloud service providers are required by law to sign the non-disclosure agreements with its customers, this agreement must reflect the customer's needs for data protection and operational details that will be continuously reviewed. The cloud service provider making use of third party resources services to store and process data is required to select and monitor that outsourced third party provider is in compliance with the privacy laws of the country where the data originates.

A cloud service provider is required by the industry's best practice to publish on their portals/websites third party service reports as well as their internal audit reports. Upon implementation of the cloud computing technology the cloud service provider is required to secure its components (hypervisors, router, operating systems, etc.) with security measures that are relevant to particular industry standards such as COBIT and ISO207001 and document such

information security baselines in such a way that they can be used for auditing purposes. It is also required to segregate the data of each tenant from the other data subjects and it must be encrypted upon storage.

Timely revocation of access right to any cloud employee upon the detection of status change is required as well as timely cloud customer notification upon any privacy policy and information security changes. Privacy policies need to be in accordance/alignment with industry standards.

Cloud computing introduces a lot of uncertainty with regards to privacy compliance regulations as far as far as the privacy compliance regulations are concerned. Currently there is no privacy regulation that makes direct reference to cloud computing this is mainly because the majority of these privacy compliance were introduced before the concept of cloud computing became this most popular. Most of these privacy regulations adhere to the safe harbour agreement which was established for safe transfer of data between USA and Europe, adherence to this regulation obligates the cloud service providers to adhere to other several privacy principles such as fair information principles and directive 95/46/EC and much more.

Privacy Regulations discussed in Chapter 2 have much in common; they all share a common principle. Almost all of them require the processing of personal data to be fair, lawful, and adequate as stated in the data protection directive. The regulation also requires the cloud service provider to provide the data subject with necessary information with regards to the processing of their personal data.

Proper security measures should also be put in place to protect personal data against any form of alteration, accidental loss, unauthorized disclosure or access. The cloud service provider has to design a system that able to comply with the privacy requirements of each country in which their

data centres are located, since privacy laws are not consistent across countries. Data privacy laws are defined by the data storage location and are therefore subject to that country's laws.

2.6 Chapter Summary

This chapter has presented the background for this research work. Fundamental cloud computing concepts, including cloud computing deployment models, cloud computing service deployment model as well as security challenges in cloud computing were introduced. Since security is diverse, the main focus specifically is on confidentiality and privacy. An in-depth security and privacy analyses have been carried out in this chapter, and it was carefully noted that laws and regulations are not very useful if not enforced and we therefore in chapter 3 discuss various ways in which this privacy protecting laws can be enforced.

Chapter 3

STATE-OF-THE-ART ANALYSIS

3.1 Introduction

This chapter reviews the state-of-the-art in privacy monitoring mechanisms and further assesses the impact the current privacy compliance and regulation have on implementation of the cloud technologies. A lot of attention has been channelled to coming up with mechanism to mitigate security and privacy threats that continue to hinder cloud computing technology from reaching its success capability. Partial solutions have been proposed which range from superficial to academic. However privacy remains a challenge (Lemoudden, Bouazza, Ouahidi, & Bourget, 2013).

Approaches that were analysed and reviewed in this research are being discussed in the following sections. Section 3.2 discusses third party based privacy monitoring approaches, in section 3.3 network based solutions are critically analysed. Section 3.4 discusses privacy by encryption solutions. Section 3.5 discusses privacy by computation solution and section 3.6, privacy by design is discussed respectively. Lastly anonymization techniques are discussed in section 3.7. From these discussions, the strength and weaknesses of these approaches are established. The categorization of different type of privacy enhancing technologies was carried out, in order to assess the solutions, compliance to privacy regulation and to assess if they instil trust back into cloud consumers. In addition to ascertain that cryptography satisfies the cloud security requirements in providing data confidentiality and integrity and the impact they have

toward the implementation of cloud computing various encryption techniques have been evaluated and reviewed.

It is from this gained knowledge that a solution approach was systematically formulated that advocates the use of informative events and access logs to enable the cloud customers to retrace in details who accesses to their personal private information, operations performed on the information, by whom and the security safeguards applied to it.

3.2 Privacy Monitoring Techniques

The most commonly used approaches to provide privacy and data security are, Best Effort Approach, Third Party Audit, Privacy as a Service, Trusted Third Party Audit and Privacy Manager.

3.2.1 The Best effort approach

Security and privacy are attributed to the lack of proper security control policies and weaknesses in security safeguards in cloud deployments (Prasad et al., 2011b). To tackle the privacy challenge, the best effort is proposed by Glott et al. (2011). This is done through the conduct of self-assessments to determine if the systems are compliant with privacy regulations and standards. These assessments are based on arbitrary frameworks which generally focus on the documentation of security policies.

However, this approach gives no guarantees of securing the data that is in cloud service provider's custody. It only promises to set proper security safeguard to preserve data privacy without giving any guarantees, and should something goes wrong it could not be held liable since

it would claim it did all that it could to protect the data. The Santos, Gummadi, & Rodrigues, (2009) and Pearson, (2013) contends that service providers should be able to provide attestation on why they claim their services are private and secure in the most transparent manner. These authors' highlights that the promises of security are usually not enough, customers need to be shown convincing evidence that security measures are in place to ensure security.

3.2.2 Third-Party Audit (TPA)

TPA performs tests of the subject matter to form an opinion or a report on the matter of assertions. This report discloses whether the cloud service provider security safeguards meet the security standards and also to assess the effectiveness of its control over the collection, use, retention, and disclosure of personally identifiable information.

Glott et al. (2011) proposed Third-party audit as an alternative to best effort approach. When employing this approach, a trusted third party organization which has the right combination of knowledge and expertise in the field is recruited to validate if the cloud service centres are satisfying the well-defined privacy standards such as ISO27001, HIPAA etc. (Glott et al., 2011). The major benefit of this approach is that it assures customers that the organization followed the standards agree upon at the time of certification so that the cloud consumers can then be ascertained that indeed the cloud service provider complied with the relevant standards and regulations. TPA serves to enhance consumers trust and provide a means to verify CSP compliance with applicable data privacy laws and regulatory compliance. Consequently, this enables the cloud customer to remain compliant with domestic and international laws and regulations. Furthermore, TPA helps in obtaining external feedback that helps to build up CSP privacy scorecard which helps CSP to secure more clients and gain competitive advantage over

its competitors.

However, According to Rüdiger et al, (2011) employing this approach only ensures compliance at that point in time. Therefore it is prone to miss areas of non-compliance that by accident, were not checked due to the spot-check approach. It also leaves user data privacy exposed to the third party auditor which introduces another challenge.

Wang, Chow, Wang, Ren, & Lou, (2013) tried to address the shortcoming of leaving users' data privacy exposed to the third party auditor by developing a privacy preserving auditing protocol, which enable the auditor to audit users' data without having knowledge of the data contents. The authors also proposed batch public auditing protocols which cater for various auditing requirements for multiple different users that can simultaneously be performed by a TPA using multiple algorithms to achieve different goals

However, these solutions leave a lot to be desired. They ignore the user control aspect which is mostly responsible for building trust between the cloud customer and cloud service provider, hence they still implement the black box approach whereby the cloud service provider keeps the cloud customers in the dark with regards to the whereabouts of their data, therefore, the cloud customer cannot obtain insight of the operations performed on the outsourced data (Pearson, 2013; Glott et al., 2011). As a result Privacy as a Service come into play to compensate for this approach shortcoming, this approach tries puts to put the clients in control of their data.

3.2.3 Privacy as a Service

Itani et al., (2009) proposed the Privacy as a service, Privacy as a service is a set of security protocols developed to ensure legal and privacy compliance of cloud customer private data in

cloud computing architectures. This approach makes use of the tamper-proof capabilities of cryptographic coprocessors to secure the storage and processing of confidential data of the cloud users in the cloud computing environment. Just like the third TPA approach this solution relies on the third party to ensure that the agreement to protect the outsourced cloud consumer data by the cloud service provider is being adhered to.

These coprocessors are owned by the private third party, who signs them using a secret key and then supply them to the cloud service provider, this helps to protect the data even when the adversary has direct physical access to the storage and processing infrastructure. The advantage of this approach is that it provides maximum user control to cloud customers for them to be able to manage their critical aspects of their data; it achieves this by using user-configurable software which enables registered cloud customers to share the secret encryption key with the trusted third party. As depicted in Figure 3.1, Privacy as a service architecture is composed of the CSP, Cloud Customer and Trusted Third Party, this architecture is designed to enable communication between the trusted third party and coprocessors to share the encryption key over the secure mutually agreed upon public/private key pair channel.

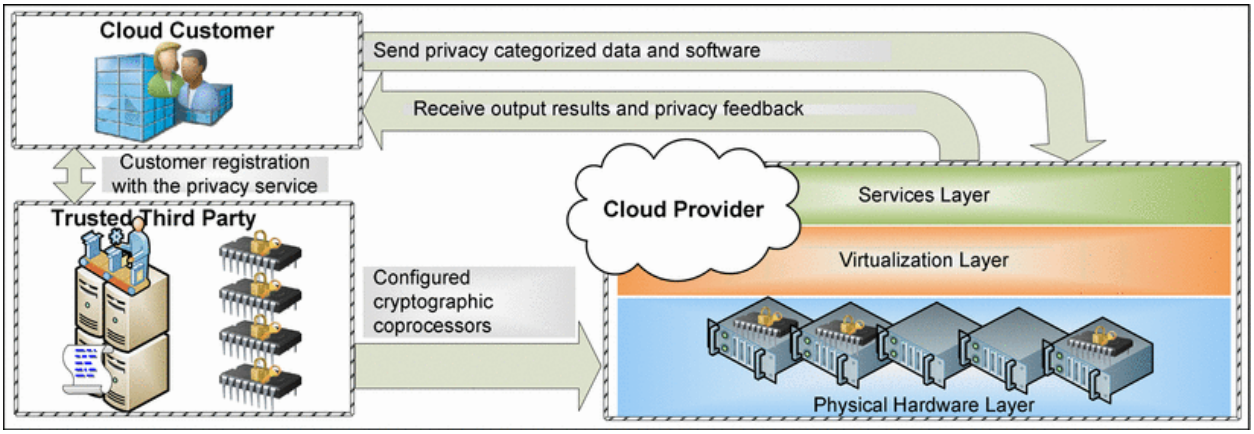

Figure 0.1 PasS Cloud-based System Model (Itani et al., 2009).

(Allison & Capretz, 2011) proposed the used of the Trusted Third Party (TTP) to ensure that the interest of all parties involved is taken into consideration. (Abadi, 2004) found that resorting to trusted third parties may not be always practical, as it typically results in deployment difficulties, communication overhead, and other additional costs which makes the process even more expensive.

Moreover, well-founded trust is scarce in large-scale distributed systems, and so are reliable TTP. This solution therefore inherits all the shortcomings of the TPA and furthermore, co-processors are generally a very expensive hardware and using such service is resultant costly for cloud customers, and defeat the purpose of cost cutting that cloud customers initially adopted cloud to obtain (Vimercati, Foresti, & Samarati, 2012).

3.2.4 Trusted Platform Module (TPM)

Trusted Computing Group, 2010 and MIMOS, 2012 developed standards for hardware enabled trusted computing, TPM is a hardware security component was then formulated and built into many computers and computer-based products(Abadi, 2004; MIMOS National R&D Centre in ICT, 2012; Trusted Computing Group, 2010). According to Morris, (2011) the chip allows for machine authentication, hardware encryption, signing, secure key storage and attestation. TPM monitors, software as it is loaded and provides secure reports on exactly what is running on the machine.

The major advantage of using this chip is that it generates evidence based confidence that every claim made by the cloud service provider concerning the security of its infrastructure is legitimate. Having gained this kind of trust, the customer feels at ease to outsource any kind of information while it reaps an even more lucrative pay-out. According to Kleyman, (2012) this

chip does not protect against attacks that exploit security vulnerabilities introduced by the programming bugs. Furthermore, TPM does not protect the hardware from its owner, but only for its own, leaving the door wide open for insider threat such as rogue administrators.

Halderman et al., (2008) reported that they were able to overcome the disk encryption of TrueCrypt, FileVault, and BitLocker with cold boot attacks. TPM also lacks the capability to provide user control, it also does not provide situational awareness, and changes to security and privacy that are in best interest of the organization /user. To enable the user to express preferences, concerning the treatment of their personal information, including the degree and type of encryption to be used to secure the private information while it's stored in the cloud the Client Based privacy manager was proposed.

3.2.5 A Privacy Manager for Cloud Computing

S. Pearson et al., (2009) proposed a Client-Based Privacy Manager; this manager helps users to remotely manage the privacy of their data in the cloud. Privacy Manager provides the obfuscation feature which allows for data obfuscated to the degree specified by the data owner, obfuscation is a form of data encryption, which requires no decryption key to reverse.

The same approach of encrypting data before handing it over to the CSP was also successfully employed by Total Claims Capture and Control (TC3) (Jain, Payal, 2012; Zhou, Zhang, Xie, Qian, & Zhou, 2010), a healthcare company with access to sensitive patient records and health care claims, when moving their HIPAA compliant application to the cloud service.

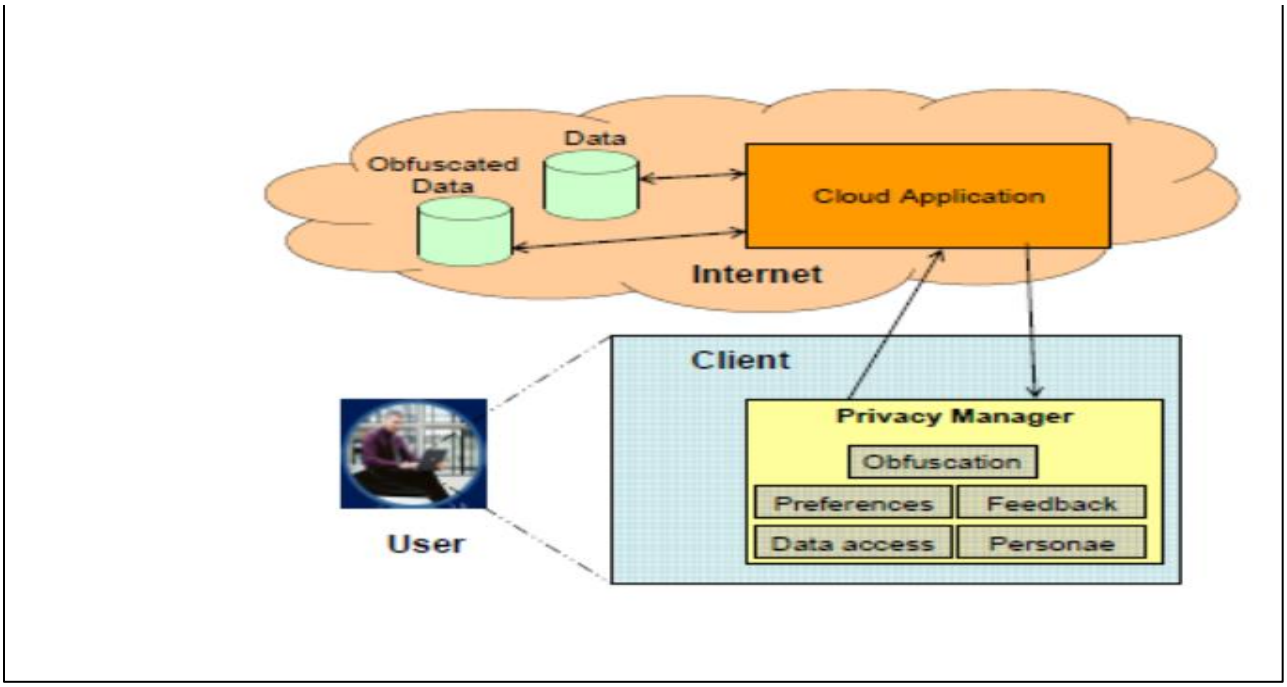

Figure 0.2 Client-Based Privacy Manager (S. Pearson et al., 2009).

As depicted in Figure 3.2, user preferences are set via the privacy manager, they form a privacy policy which is first bound to the obfuscated data and together transmitted to the cloud remote servers. The major drawback of this approach is that it does not allow the customer to know where his data is located since the location of data has a direct impact on the laws applied to it and who has physical access to his data.

Knowing data located, is key for knowing which laws, standards, and regulations must be consented to. Certain geographical locations might expose Electronic Protected Health Information (EPHI) to international law that stipulates who can have access to data even though it's contradictory to HIPAA and HITECH laws. This research work therefore proposes a solution that builds on top of the strength of this approach and others.

3.3 Network Based Approaches

An emerging class of network based services that promote collaboration when storing and processing the data on untrusted servers, namely Secure Untrusted Data Repository and Group Collaboration use Untrusted Cloud Resources.

3.3.1 Secure Untrusted Data Repository (SUNDR)

Li et al., (2003) developed an efficient fork-linearizable storage protocol based on Secure Untrusted Data Repository (SUNDR), which reduces communication complexity. SUNDR is a network file system that guarantees the integrity of the data when stored on the server that is controlled by malicious parties. The main benefit of this approach is that it provides the verification of the integrity of the retrieved block of data from the clients. However, SUNDR protocols are blocking as result they may block in an event whereby the client crashes even when the storage server is correct. Moreover, SUNDR does not address the reliability and storage issues. SUNDR uses public-key cryptography algorithms that are in orders of magnitude slower than their symmetric key equivalents (Naor, Shenhav, & Wool, 2005). As a result, they introduce performance degradation issues which make them the least preferable option. A more efficient fork-linearizable storage protocol is implemented by Cachin, Shraer, & Shelat, (2007) which reduces communication complexity and is also based on SUNDR.

3.3.2 SPORC: Group Collaboration uses Untrusted Cloud Resources

To overcome the challenges posed by the untrustworthy cloud providers, SPORC framework was proposed by Halderman et al., (2008). This generic framework is meant for building a collaborative system which is based on untrusted servers (Feldman, Zeller, Freedman, & Felten,

2010).

According to (Pappas, Kemerlis, Zavou, Polychronakis, & Keromytis, 2012; Stefanov, van Dijk, Juels, & Oprea, 2012) SPORC framework closely monitors the server so that it cannot deviate from its normal operation and only operates on encrypted data. However, the server could introduce other undesired noises to degrade the system performance or launch replay attach or even worse block the communication between clients, since all this framework guarantee is that the server cannot corrupt the shared state by using sophisticated algorithm which in turn imposes significant overhead on the client as a result of it using checkpoint to reduce the saved client's state.

3.4 Privacy by Encrypting

Nowadays, access to applications and servers hosted in the cloud is done in a browser, which introduces vulnerabilities to insider threats as well as hackers.

The most preferred route to enforce privacy is through data encryption. Implementing encryption for the cloud data gives guarantees that the data will remain safe during the course of its storage on cloud servers in the event of a hacker hacking into cloud or malware deployment in the cloud that may exploit a vulnerability to access critical data or in the event of physical disk theft (Reed, Rezek, Simmonds, 2011). There exists a wide range of encryption approaches that one can employ to encrypt the data.

3.4.1 Randomized Encryption (RND)

Ronald L. Rivest, (1983) proposed RND which achieves greater privacy by encrypting data using randomly chosen sets of cipher text that corresponds to the data. The major drawback of this

process is that it increases the original data size and subsequently the required bandwidth to transmit it. Furthermore, RND does not allow any computation to be efficiently performed on the plaintext, which helps to protect the privacy of the data while it's being transformed.

3.4.2 Deterministic Encryption (DE)

Bellare, Boldyreva, & O'Neill, (2007) conducted a study of deterministic public-key encryption. According to this study, DE allows for fast searching on encrypted data. Furthermore, the deterministic encryption is length-preserving, which ideal for securing legacy code or in bandwidth-critical applications. However, DE provides a weaker privacy guarantee since it permits logarithmic time search on encrypted data, as opposed to RND which only provides linear search time search, meaning the DE enables the server to check plain texts for equality by checking for equality of cipher text in the whole database

3.4.3 Order-Preserving Symmetric Encryption (OPE)

The concept of order-preserving symmetric encryption (OPE) was proposed by Agrawal, Kiernan, Srikant, & Xu, (2004) in the database community. According to Boldyreva, Chenette, Lee, & O'Neill, (2009) Order-preserving symmetric encryption function preserves numerical ordering of the plaintext. It allows efficient range queries to be performed on encrypted data. It achieve this by enabling the remote untrusted database server to index the private data it receives, in encrypted form, in a data structure that allows effective reach.

This kind of encryption, OPE is even more relaxed in the sense that it enables inequality checks and sorting operations, but has a property which ensures that the distribution of cipher text is independent of the encrypted data values and also pseudorandom (Xiao & Yen, 2012). However, it exposes sensitive data to inference attacks with order information when data are used together

with unencrypted columns in the database.

3.4.4 Homomorphic Encryption (HOM)

R. Rivest, (1978) initially introduced the notion of Homomorphic encryption originally then known as Privacy Encryption. The fully Homomorphic encryption scheme enables one to perform complex operations over encrypted data without having or knowing the decryption key. According to Guan, Tsai, & Zhuang, (2013) response to these operations is normally obtained in an encrypted form which is the cipher text of the result of operations performed on the plaintext. Xiang, Yu, & Zhu, (2012) contends that there is no way one can know that the cloud did not perform the correct computation, nor if the value ones receive after decrypting the final ciphertext is correct.

Ciphertext in the ciphers is much larger than the plaintexts, which result in performance degradation since the computations on large ciphertext are typically slower than if the computation is performed on the plaintext. Encrypting inputs and decrypting the output contribute to performance degradation considerably.

3.4.5 Transparent Data Encryption (TDE)

TDE is the technology used by Microsoft SQL Server 2008 to encrypt database contents (Pasha & Gafoor, 2011). One of the benefits of TDE is that, it offers encryption at a column, table, and the table space level (Mattsson, 2005). This makes TDE one of the most highly configurable ways to encrypt database content, though some of these configuration options come with a performance price. TDE secures information "at rest", meaning the data and log files. It gives the capability to follow numerous laws, regulations, and rules secured in different commercial enterprises. This empowers programming designers to scramble information by utilizing AES

and 3DES encryption algorithms without changing existing applications. According to Zhang, Liu, Chen, & Li, (2009) TDE encryption keys are often kept in a database "wallet", which is it often a file on a disk. The concern is that hackers may attack the virtual disk in the cloud, and from there get into the wallet and through that to the data. This encryption approach does not protect the data from authenticated, authorized database users, including the DBA.

3.4.6 CryptDB

Popa & Redfield, (2011) proposed CryptDB, which enables queries to be performed on the encrypted database managed by the cloud provider and provides provable confidentiality in the face of privacy threats. According to (Glott et al., 2011; Popa & Redfield, 2011) it only grants access to the legitimate data owner, and this is achieved by chaining the user password with the encryption key, thus the data can only be decrypted by using the password of the users who hold the legitimate password. Research shows that that CryptDB has low overhead, reducing throughput by 14.5% for phpBB, a web forum application, and by 26% of queries from TPC-C, compared to unmodified MySQL (Popa & Redfield, 2011).

Data encryption serves as a protection of the data. However, this approach has its downfalls since it requires encryption keys which in turn introduce the new challenge of key management; key management is a significant issue because eventually a trusted third party is needed to manage the encryption keys for cloud users, the effectiveness of the encryption solution ultimately depends on protecting the key because if the key is exposed, the data being protected by the key is, essentially, exposed. Encryption is CPU intensive and it requires frequent key changes, especially in the case whereby an administrator with access to a key leaves an organization, the key should be changed.

3.5 Privacy by Computation

Lindell & Pinkas, (2014) discuss the secure computational algorithms which have been used to perform computation on an untrusted server. The author emphasise that these algorithms enable cloud users to use insecure cloud infrastructure to compute without revealing the exact input for the computation. This is achieved by implementing the Yaos protocol which provides basic techniques to perform such computation without revealing the actual expression.

Yaos protocol uses gates to manupulate the expression of a computational problem in terms of logical circuits(Yao, 1982). The input obtained from each gate is randomly encrypted and the resulting output is decrypted to provide the user with the usable answer. The encryption of the input and decryption of the final output is carried out at the client end (Yao, 1982). The major drawback of this approach is that it is far too complex to implement. It therefore still resides in the realm of theory.

3.6 Privacy by Design

(Langheinrich, 2001) proposed the Privacy by design framework. This framework is characterized by proactive approach rather than reactive measures. According to the authors this approach does not wait for privacy risks to materialize. It anticipates and prevents privacy invasive events before they even occur. Furthermore, it offers remedies for resolving privacy infractions once they have occurred, by preventing them from occurring in the first place(Davies & Langheinrich, 2013).

Unlike other system whereby once the system has been design completely, privacy is then bolt in. Privacy by design embeds privacy into the ICT systems design and architecture of and

business practices, enabling it to prevent privacy invasive events from happening. Privacy by design framework keeps component parts of IT systems and operations of business practices visible and transparent, to users and providers alike, while ensuring end to end security amongst the communicating parties; in doing so it protects and respects the interests of all information owners.

However, Privacy by design principles remain vague and leave many questions unanswered concerning their application when developing systems, very little past experience exists pertaining designing and implementing systems with privacy in mind, these processes are therefore prone to error, since they success lies in the hands of inexperienced engineers which could on its own increase cost and consume quite a considerable amount of time

3.7 Anonymization Techniques

Anomymization techniques are widely used to protect the identity of the data subject this is achieved through data minimization which is the process of removing the PII of the data subject and reveal least necessary data needed for communication thereby providing the much desired user anonymity or at least pseudonymity. Several techniques have been used to achieve user anonymity and the common goal of these technologies is to hide the correlation between the input and output data in order to protect the PII of the end user.

These technologies include the FreeHaven Project (Dingledine, Freedman, & Molnar, 2001), Platform for Privacy Preferences (Reagle & Cranor, 1999), Bugnosis and Pretty Good Privacy (PGP) (Zajac, 1994). Dingledine, Freedman, & Molnar, (2001) developed an anonymous storage system, it focuses on achieving data persistence. Freehaven achieves Anonymity via a

community of servers which are known as servnets. Servnets are responsible for the processing and storage of the actual document parts. Each document part has a pseudonym, so they can be easily identified within the network and maintains a reputation index. Publishers and are not necessarily part of the servnet network, therefore only a small part of the P2P community actually manages the securely stored data. Employing P2P means that the initial path must be pre-constructed and encrypted using asymmetric cryptography. However, pre-constructed paths must be periodically probed, assessing whether all the included peers are still online before they can be used or the else the message may end up being lost. The pre-constructed of paths lead to system overhead since a heavy burden is put on the source peer.

The World Wide Web Consortium's Platform for Privacy Preferences (P3P), enable its users to describe their preferences on how they want their data to be handled or dealt with by the receiving party, it also allows the sites to describe their data handling policies (Mark S. Ackerman and Donald T. Davis, 2003). However, this technology uses technical privacy and contractual terms that are hard to be understood by the novice or non-technical users.

Bugnosis (Alsaid , Martin , Arabian, 2003) detects Web bugs that are used by the third parties to gather information about user behavior without user consent, these bugs also trick Web browsers into assisting with this surveillance. Web bugs are invisible third party images added to a Web page so that the third party receives notice that their page is being viewed. These bugs trick the web browser into claiming that an image is required as part of a Web page in order for them to be able gather personal data which is no longer based only on the individual's dealings with a Web page.

Upon the detection of these bugs by the Bugnosis internet explorer add-on, the add-on then informs the web browser operator of their presence. Bugnosis lacks web browser dynamics; it is

mainly designed for the use in the internet explorer web browsers constricting user web browsing choice to one browser. It also relies upon an internal database regarding certain businesses and their privacy practices without relevant updates to this database it can be rendered useless obsolete.

Pretty Good Privacy (PGP) (Morris, 2011) software is free and it runs on multiple platforms. PGP provides a highly secure, reliable electronic mail and file transfer, and high quality data encryption, data compression and digital signatures. PGP achieve this by combining several existing public-key encryption algorithms and protocols into a single package. Zimmermann, (1991) developed the software as a reaction to a questionable measure in Senate Bill 266 that would have obliged all encryption strategies to incorporate a secondary passage for law enforcement. PGP is based on extremely secure algorithms such as MD5, RSA, Diffie-Hellman and IDEA. The main weakness in this approach is that it uses the public system, because there is no secure way to know that the public key really belongs to the right correspondent.

3.8 Chapter Summary

In general, this research work observe that most of the existing mechanisms are restricted in using cryptographic techniques to protect data confidentiality and integrity in the cloud, due to the notion that many cloud computing security and privacy requirements are solvable through this technique. However, in the face of ever advancing cryptanalysis mechanism and perpetually growing requirements for privacy of data in a number of jurisdictions, there is a need for a privacy solution which will compensate for these shortcomings.

Furthermore the current existing approaches fail to cover the holistic picture of cloud services

across all three layers of the cloud stack. In conclusion, the state of the art analysis has identified the shortcomings that these solutions lack the capacity to provide cloud consumers the means to control and comprehend what happens to the data in the cloud and which efforts to establish safety are deployed to preserve its private nature. The issues derived from this chapter have served as the requirements for developing the privacy framework which will be introduced in the next Chapter.

Chapter 4

DESIGN AND DEVELOPMENT OF A PRIVACY MONITORING FRAMEWORK

4.1 Introduction

The lack of transparency in cloud security controls and privacy measures is widely seen as an obstacle that is hindering the broad adoption of the cloud computing technology (Pearson, 2009). Researchers are advocating for a mechanism that would enable cloud consumers' to have an insight into how their data is stored and processed by the cloud service provider and who else have access to it.

This research work, proposes a solution that builds on top of the strengths of the previous works to address the privacy issues in the cloud. To achieve this end in view privacy manager approach was adopted and to compensate for its drawbacks the informative event and access log analyser was developed. It enables cloud customers to retrace in details what happened to their data, where they are stored, who accesses them and the security safeguards applied to it. The conceptual solution put forward by this research work will attempt to solve the privacy issues that continue to hinder broader cloud computing adoption.

Section 4.1 presents the domain specific scenario which highlights a number of challenges relating to cloud service provisioning from real-world instances that may lead to privacy violations in the cloud computing environment, Section 4.2 addresses the design requirements which form the basis of the framework. Section 4.3 presents, the proposed Privacy Monitoring Framework and provide details of how the design requirements are realised, through the use of

the information events and access logs as a means to answer the main research question, component interaction is also described in this section. Section 4.4 provides the summary of the chapter.

The framework proposed in this research seeks to protect the data privacy and issue alerts on the events of suspicious behaviour patterns being detected in order to ensure privacy through the promotion of transparency.

4.1 Domain Specific Usage Scenario

In order to explain this work in a practical setting, MediClinic health enterprise scenario was chosen for demonstration purposes since it is one of the three largest private hospital groups in the southern region of the African continent. MediClinic unified its hospitals from three continents, electronic health records helps them to improve coordination and integration of services among care providers.

We consider a health enterprise whereby the health enterprise MediClinic stores its electronic records in the cloud computing environment. These health related information on individuals must conform to national interoperability standards such as HIPAA. Health records can be created, managed and consulted by authorized clinicians and staff, across different organization.

Electronic health records contain existing conditions and progress notes, past medical history, patient demographics, immunizations, medications, as well as laboratory data and radiology reports. Cloud enables the health organization to collaborate and share the

much needed information which includes some diagnostic results from its patients in order to derive a cure for a common illness that troubles them.

However, security is one of the major issues of cloud computing. Being completely centred on Internet makes it susceptible to unanticipated assaults. Electronic health records can be stolen from the service provider's system by hackers and such data breach may take time to be detected since they are mostly performed by skilled hackers. This kind of attack may result into catastrophic effect to individuals whose information is being maintained. This unwanted exposure of individually identifiable health information, and PII may result in emotional distress, blackmail, humiliation or loss of self-esteem. In the worst case scenario, private insurance company may have access to this information, and with this kind of information at their disposal they may begin to modify their actuarial tables depending on their patient private leaked information. This may eventually lead to widespread illegal financial behavioural modification, since the client may have to be asked to pay higher premiums than what they are already paying as a result of evidence emanating from these records of their clients having minor injuries while playing contact sports such as surfing, horse riding, skiing, sky diving, etc.

This information empowers the private insurance providers to give their clients the option of totally stopping the activity, paying higher premiums, or worse risk not being covered while doing it. Some patients may never be able to get medical insurance since insurance institutes may discriminate them based on the information recorded on these stolen files. MediClinic may lose its reputation and furthermore be forced by the law to pay monetary penalties as a result of this incidence and may further lose its license to collect and store personally identifiable information.

Patients need to be assured that their medical records are well protected when they are stored in the cloud. They also want an audit function which will enable them to know each and every individual who accessed their medical record. These capabilities can afford these patients the opportunity to monitor the privacy of their personal information against misuse, loss or unauthorised access, modification or unauthorised disclosure by the cloud personnel. This can also allow for openness about privacy practices, this includes but not limited to informing data subjects when problems threatening their data arise and provide clarity on what has been done to circumvent or remedy any real or potential damage.

4.2 Design Criteria's

From the issues outlined in the above scenario and the literature review in Chapter 3, the following design criteria were extracted to be considered when designing a privacy monitoring framework for the cloud computing environment.

1. **Automate and Manage data Collection, Reporting and Alerting**- The proposed framework must facilitate compliance with privacy laws, regulations and standards, whether it's HIPAA, PCI-DSS, GLBA or any other government regulations. A detection mechanism is therefore mandatory should any non-compliance be detected. The proposed framework must automatically identify important privacy violating events and alerts appropriate individuals. The aim of this requirement is to enable the cloud customer to remain compliant with domestic as well international laws and regulations. The

framework must also help CSP customers to keep track of what their business partners with direct access to confidential data are doing.

2. **User-driven definition of events and specifications**- The framework must be able to dynamically define and provide specifications of the conditions that will trigger alerts. User preference functionality is therefore needed for the specifications of events which are of interest to the data subject. Events of interest are registered and triggered by applying simple filters or complex regular expression driven filters that pick out the customer entries of interest. This framework must also provide a mechanism that will enable these events to be manipulated and fired at runtime. This functionality must enable users to clearly specify which events are interested in receiving alerts on.
3. **Limit data access to authorized users**- The framework should prevent unauthorized users from accessing confidential data. The solution should likewise implement distinctive levels of access rights focused around data owner's specifications, as a way of supporting collaboration with internal (colleagues) and external clients, for example, business specialists, promoting agencies, industrial configuration firms, legal counsel, and different sorts of business accomplices. For instance, a few business accomplices need to have the capacity to add, while others have to have the capacity to both read and compose documents, and different business accomplices require just perusing records, not altering or making them. The framework must protect its customers from insider threats by keeping track of what their administrators are doing. Administrators pose much greater risks to data security, since they have the ability to create and modify permissions, privileges and access to any device(Nkosi, Tarwireyi, & Adigun, 2013). The data security best practice that this framework is putting forth clients and giving them the capacity to

set lapse dates on the records. Lapse dates empower approved clients to access data within sensible measure of time, while guaranteeing that sensitive information is not put away on servers inconclusively.

4. **Encryption of data in transit and at rest**- The solution should apply FIPS 140 (Easter & French, 2014) validated encryption algorithms such as AES 128-bit or AES 256-bit encryption since they are efficient in both software and hardware and they are the current approved standards for strong encryption.

4.2.1 Basic Framework Assumptions

These are the basic assumptions that were made towards the development of this framework

1. A cloud service provider has mechanisms that provide its customer with some logging facilities within its resources, since the log files record all events carried on the tenant data, can therefore be prune for a single tenant for monitoring purposes.
2. A legally binding contractual obligation has been established between the cloud customer and cloud computing provider.
3. Event logs are protected from unauthorized access and modification. So that the events and time stamps recorded on them can be relied on and to be able to consider time zone differences to determine the location of the data and for auditing purposes.
4. The client application can read from cloud log files.
5. The cloud customer(s) is only able to see the logs or to receive events pertaining to their data only.
6. The cloud customer can be able to request for response action and for it to be executed it must be authorized by more than one administrator.

7. User's rights and their contextual information to use these rights have been aggregated across the different data centre that the cloud service provider owns and stored on the single repository. To help streamline the reporting and analysis of access rights that represent the biggest privacy violation risks, since some users may have excessive privileges which poses a great threat to cloud customer's privacy.

4.3 Privacy Monitoring Framework

The framework presented below addresses the limitation identified in the knowledge base as discussed in the preceding chapter. The proposed framework illustrated in Figure 4.1 intends to realize this by using an information events and access logs analyser component which would enable cloud customers to retrace in detail what happened to their data, where they are stored and who accesses them and the security safeguards applied to it while it is in custody of the cloud service provider.

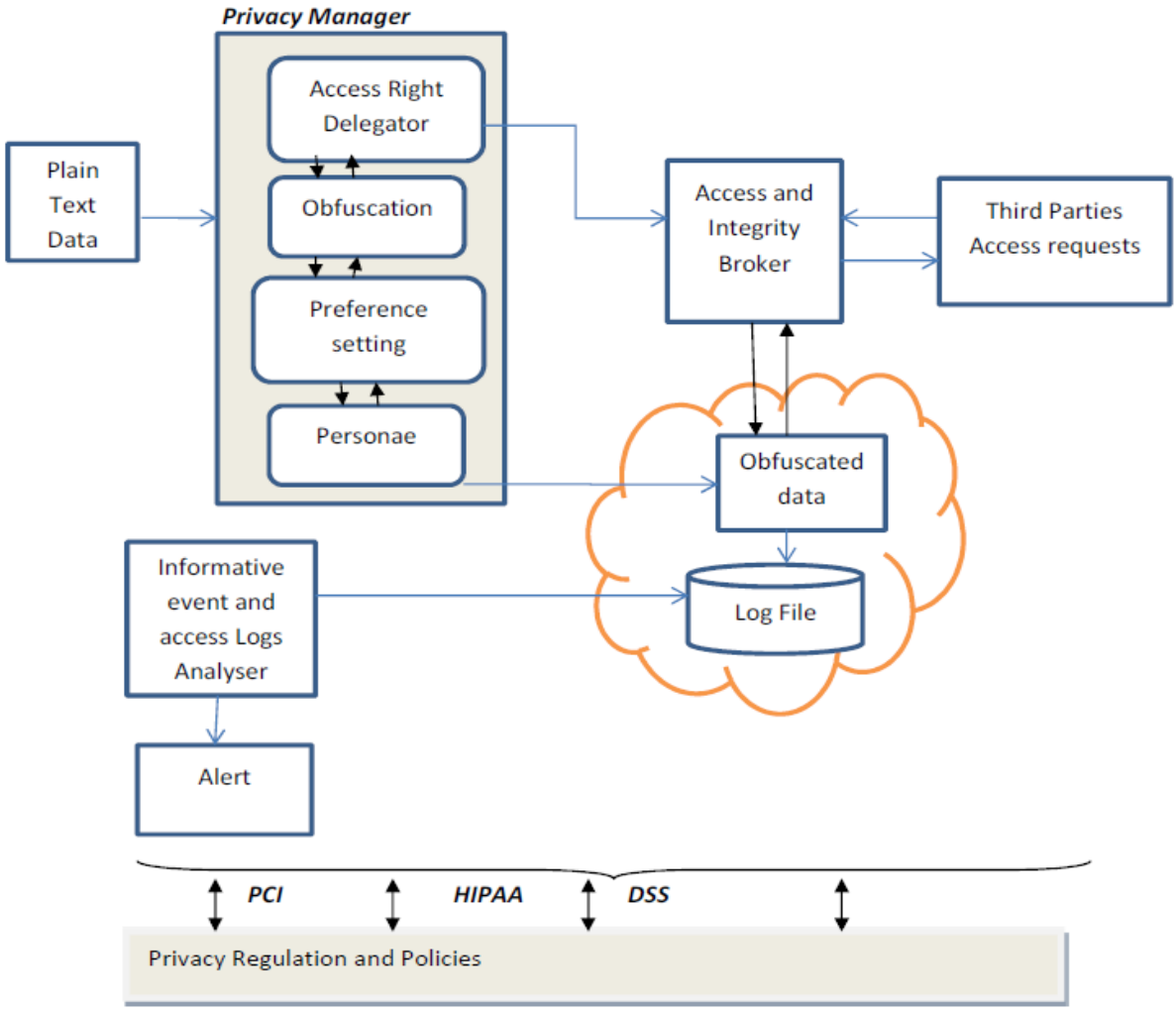

Figure 0.1 Privacy Monitoring Framework

Monitoring the privacy of the outsourced data plays a critical role in determining whether the privacy of user data stored in the cloud has been violated or not. Since the data owner by law remains responsible for the compliance with all the principles of data privacy regarding the outsourced data (Glott et al., 2011; South African Law Reform Commission, 2005; Reed, Rezek, Simmonds, 2011). The data subject must therefore be able to control and comprehend what happens to the data in the cloud. Monitoring gives answers to both parties with regards to data handling issues with regards to determining which party is responsible in the event of data loss.

A customer perspective:

- If something goes wrong, how will I know? (Detection)
- How can I tell if it is my fault or cloud provider's fault?
- If it is the clouds fault, how can I convince the provider?

A Provider's perspective:

- If something goes wrong, how will I know? (Detection)
- How can I tell if it's my fault or cloud customers fault?
- If it's the customer's fault, how can I convince the customer?

These monitoring attributes help to resolve the accountability issues in case of a data breach by providing attestation to reliably link the responsible party (customer or provider) with the security fault. It is through this functionality that the customer can reliably detect violations, and can hold the cloud provider responsible by proving that proper security safeguard were not in place to protect PII provided that is the case.

In order to realize the design criterion's the information events and access log analyser component was implemented which gave the capacity to users to retrace in detail what happened to their data:

(1) Where the data is stored

(2) Who accesses it and

(3) What security safeguards are applied to protect their sensitive information?

The **informative events and access log analyser component**, retrieves specific logs using user preference parameters/ filters for analysis. These log files are then analysed for action events. An

event record is only regarded as actionable if and only if the event record indicates a strong likelihood of malicious activity, personal data revealing. Log entries encompass all event data that occurred on the personal data of each user which then form privacy evidence needed to warrant a privacy alert detailing the privacy violation detected.

These event logs contain information, such as the time that the events occurred, the type of the event, data category accessed, the unique ID of the user. Furthermore, this component enables the discovery of the indicators of possible security incidents which can help to minimize the number and severity of security incidents. These collected data logs are directly associated with a user and the related privacy policy to produce privacy evidence as illustrated in Figure 4.2, this helps in making the parameterized query to retrieve a relevant log file of the user in question easy.

The information helps the data subject to have the necessary understanding and knowledge with regards to the handling of their data, it enables the data subject:

1. To comprehend to how their personal data are handled
2. Empowers the data subject to know who is processing their personal data and for what purpose
3. Make the data subject understand the limits of processing transparency
4. To comprehend the limit on objecting to processing
5. To be truly informed when giving consent to the processing.

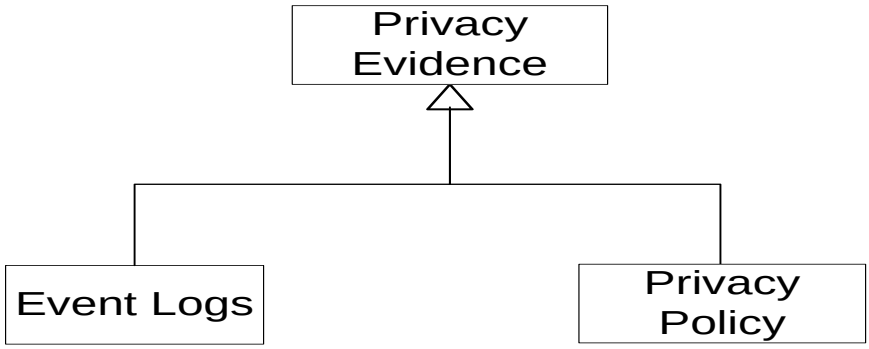

Figure 0.2 Privacy Evidence Creation Process

Some of the network attributes that are stored on the event logs that help with the provision of location of the data and for some auditing purposes are time stamps and time zone. These event logs are protected from unauthorized access and modification as required by the data protection laws. Such knowledge is crucial since the data are subjected to the rules and regulations of the country where the data centre is located.

The alert component in the framework responds to critical events that are recorded in the log files, actionable events always lead to an immediate alert trigger. Also, the system will trigger if there is an attempt to modify data, or there have been failures, errors, status changes, access and administration events, and other unusual events in the system.

User driven definition of events and specifications criteria is met through the information events and access logs analyser component. This is responsible for analysing all log events based on the privacy preferences set by the data owner by also taking privacy policy into consideration in ensuring that the processing of information is in accordance with the relevant privacy standards and regulation. It is through this assessment that the relevant course of action is taken, which might range from notifying the relevant parties or suspending the operation until the issue is resolved. The framework proposed in Figure 4.1 addresses the limitation identified in the knowledge based as discussed above and at length in the preceding chapter.

4.3.1 Framework Analysis

Monitoring is a complex task that cannot be achieved by a single component; several components need to work together. The approach taken to address the monitoring challenges is the use of the event-based model, which is a commonly used pattern for inter-object communications. This model is progressively being utilized in the context of web services (Bertino, Martino, Paci, & Squicciarini, 2010). The detailed description of the privacy monitoring framework components is given below.

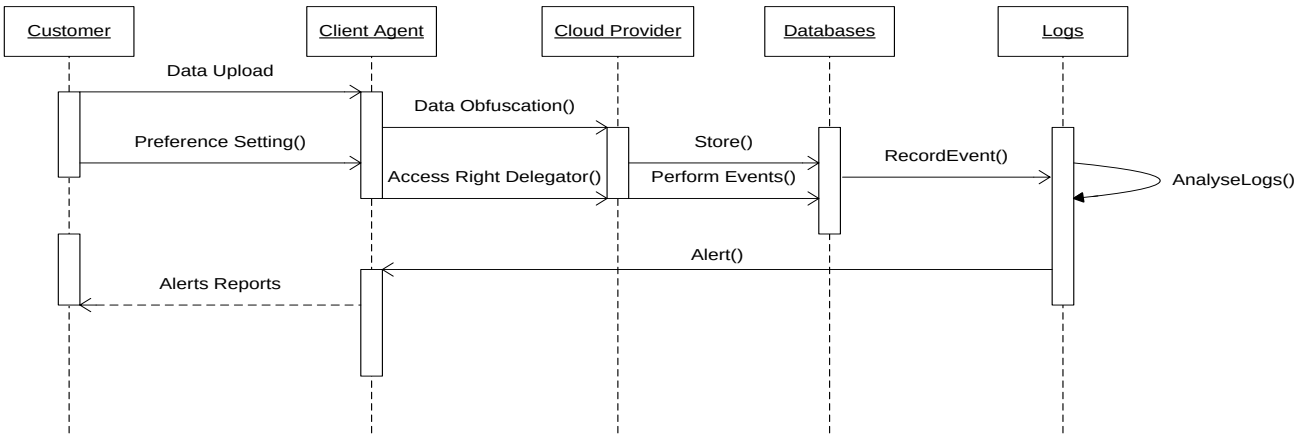

Figure 0.3 Privacy monitoring sequence diagram

The above sequence diagram shown in Figure 4.3 describes a transaction from capturing events from privacy violation, the customer agent first start by encrypting data prior to sending it to the cloud service provider. The customer agent first start by encrypting data prior sending it to the cloud service provider using the obfuscation privacy manager features. The data subject then

issues access delegation rights to the cloud service provider so that it will be able to handle data utilization request from the requester, the requester in this case can be anyone that has a personal or business relationship with the customer who has outscored data to the cloud, the customer could act as a requester as well. The data subject make use of Preference setting module to set preferences regarding the handling of their private and non-private information that is stored in the cloud, since privacy is a subjective matter and what is considered private to one person is not private to the other, From the preferences a policy is formed and that resultant policy is sent together with the data to the cloud and the preference cryptographically bound to it.

The integrity broker provides assurance that the data has not been tempered with. The essential prerequisite to be met is that when the cloud consumer wishes to recover information from the CSP, the information ought to be promptly accessible and in the same state as the initially outsourced, with no unauthorized single bit error, and if any changes were made it's only those who were made or permitted by the data owner.

The select/specify option allows consumers to receive alert by using collection preference setting component embedded on the client agent. And the alert is triggered once such events are recoded on the log file. A detail discussion at each component and the role each play in the framework developed in this research work is provided in this section.

4.3.1.1 Access Rights Delegator

The main functionality of this component is to grant access rights to business partners to be able to access the outsourced data. It specifies who can access and edit which aspect of the data stored in the cloud. This component supports collaboration with internal (colleagues) and external users, such as business consultants, advertising agencies, industrial design firms, legal counsel,

and other business stakeholders. It is through this component that the control is establish of who can view only or view and edit based on the presented credentials. Once the entity has been granted access right the credentials are stored in the database where each time the access request is received the credentials presented are compared to those stored on the database.

4.3.1.2 Preference setting

This component is there to create and apply simple filters or complex regular expression driven filters that pick out the customer entries of interest, since on many level logs files can contain towering amounts of uninteresting, hard to decipher events, burying more useful information (Grimes, 2010).

This component enables the cloud customers to explicitly configure the system according to their personal preferences. To spare the user from privacy technical jargons and to save time pre-configured settings are already in place for the user to choose from. It empowers the customer to control which events should he receive notifications for. P3P and PRIME have also followed the same approach (Shen & Pearson, 2011). The preferences then form the policy and this resultant policy is together sent to the cloud with the personally identifiable information obfuscated with the degree that the customer deems appropriate.

4.3.1.3 Personae

This feature enables users to choose the manner in which they would like to interact with the cloud agent, by enabling the user to act anonymously, partially or fully disclose their identity. This helps to protect the identity of the users if they wish to operate on the anonymity ground with the system. A pseudoname is used instead of user personal identity in case the user chose to act anonymously. The client's decision of persona gives a basic interface to a possibly complex

set of information use preferences imparted to the service provider via the preference feature, and may also determine which data items are to be obfuscated(S. Pearson et al., 2009).

4.3.1.4 Access and integrity broker

The cloud consumer and the cloud service provider have a need for integrity affirmation. The fundamental necessity to be met is that when the consumer wishes to recover information from a service provider, the information ought to be promptly accessible and in the same state as the initially outsourced, with not one unapproved single bit error. This critical feature also enables the service provider to demonstrate to its consumer that the data stored with it is indeed the same data that the consumer placed into the cloud and if any changes were made, it's only those which were made or permitted by the data owner and other authorized business partners.

Maintaining appropriate access into customer data is of a paramount importance to the service provider, but again the same data must be made available to for access to authorized parties, so the access broker sole functionality is to ensure that business partners which make legitimate request are granted access to encrypted data

By using the obfuscation feature the customer is able to encrypt data prior to sending it out to the cloud and issues access delegation to the cloud access control broker that will handle data utilization request from the requester (third party organization, partners etc.) The requester in this case can be any party that has a personal or business relationship with the cloud consumer who has outsourced data to the cloud, the data owner can act as a requester as well.

4.3.1.5 Informative Events and access logs analyser

The information events and access logs analyser component retrieves specific logs using user preference parameters/ filters for analysis and forensic investigation. If an actionable event is identified, an appropriate individual is notified. Furthermore, it enables the discovery of the indicators of possible security incidents which can help to minimize the number and severity of security incidents.

4.3.1.6 Notifier(Alert)

This component is responsible for alerting the user of useful actionable critical events. Critical events always lead to an immediate alert and a responsive investigation by the administrator. Alerting component is then triggered only if there is an attempt to modify data, or there have been failures, errors, status changes, access and administration events, and other unusual events in the environment. Alert notification is crucial where personal data has been compromised The individual may start to change passwords and inform institutions such as bank and police for damage control in an event where this information has been compromised and the necessary alert has been received.

4.4 Chapter Summary

This chapter presented a privacy monitoring framework which aims to help cloud customers to able to control and comprehend what happens to the information they entrusted to the cloud service provider this knowledge gap is bridged successfully by a solution approach which employs informative events and access logs analyser as a means of monitoring the operation carried out on the outsourced data, since privacy can no longer be assured solely by compliance

with regulatory frameworks. In order to achieve the end goal, a detailed description of the framework and component interaction has been presented in this research work. The next chapter presents a detailed implementation and experimental results to prove that the framework overcomes the shortcomings identified in the literature. The feasibility and validation of the result of the proposed framework will be carried out to demonstrate its applicability in the real world since this is a prominent step of the design science methodology.

Chapter 5

FRAMEWORK IMPLEMENTATION AND PERFORMANCE EVALUATION

5.1 Introduction

Conceptually, the design and implementation of a cloud privacy monitoring framework belongs to design science. In the nomenclature of design science, a framework that can be used to ensure privacy in the cloud was constructed. This framework can be instantiated into a prototype that provides a “proof by construction" of the feasibility of the designed framework.

This chapter validates the framework by means of proof-of-concept prototyping and experimentation. Section 5.2 presents the design of the implementation of the framework. Then, the functionalities provided by the prototype are explained in subsection 5.2.1. Section 5.3 introduces the underlying technologies and tools that are used to build this prototype. Experimental setup is presented in Subsection 5.3.2.1 and Section 5.4 present both the qualitative and quantitative experimental results. Subsequently in Subsection 5.4.2.3 the findings of this study are discussed. Section 5.6 concludes the chapter.

5.2 Implementation Design

This section presents the prototype design of the privacy monitoring framework in the form of Unified Modeling Language (UML), the modelling language for software design. This design details the use case diagram, activity diagram, class diagram and entity relationship diagram.

5.2.1 Use Case Modelling

Taking into consideration the design requirements and the assumption established in the previous chapter, a use case scenario is painted through which the privacy monitoring framework can be evaluated.

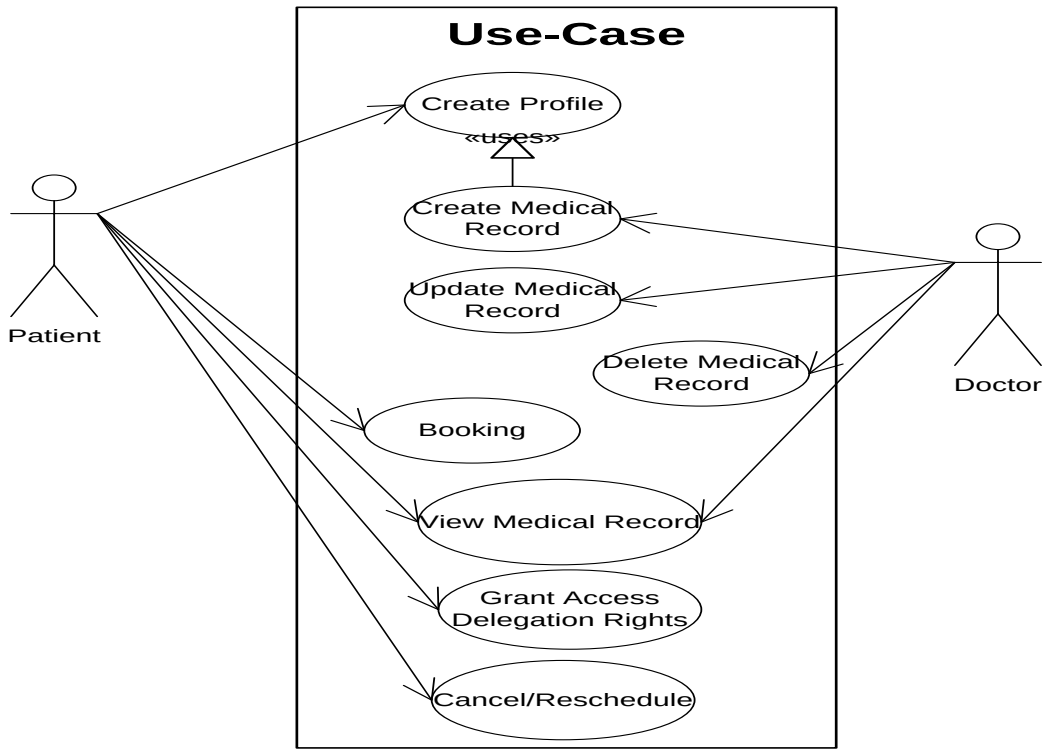

Figure 0.1 Use Case Diagram

Figure 5.1 illustrates the use case diagram for the prototype whereby the two main actors, the patient and the doctor carry out different functions.

5.2.1.1 Patient

The patient makes use of the system to register so that the system will be able to uniquely identify each patient amongst others, patient bookings and appointment with the doctor, based on examination and diagnosis outcome. The medical record is created giving personal detail of the patient and treatment. Moreover the patient is able to view personal medical record generated by the doctor. The patients has an option of choosing whether or not to grant access rights to their medical information to other parties like medical insurance company and other medical specialist who might like to make use of it. He is also permitted to cancel and reschedule the appointment.

5.2.1.2 Doctor

The doctor examines diagnoses and treats the patient. All these transactions are recorded in the medical record of which the medical doctor creates for every patient, given that none exists before. The doctor can also alter, delete or view medical record and personally identifiable information (PII) should the need arises. The doctor can also share this information with the consulting doctor or junior doctor and as per patient preference and access delegation rights allows.

The set of operations that can be performed on the system are read (R) and write (W) and they can be carried out only by the authorized individuals for different purpose which might be diagnosis (d), prescription (p), personal details (pd). Table 5-1 explains the actions, purposes in the medical history as well as the data-categories.

5.2.1.3 Logs

Event records are recorded and analyzed for any anomalies and if their violation of privacy or suspicious behavior pattern is detected the data subject is notified, all the events of interest which are explicitly defined by the user through the privacy preferences are observed in the log files and reported accordingly. Table 5.1 presents the authorization matrix of the prototype.

Table 0-1 Authorization matrix

Data Category	Doctor	Patient
Medical History	R (d, p) W (d, p)	R (d, p)
Patient personal Details	R (pd)	R (pd) W (pd)
Operative Reports	R (d, p) W (d, p)	
Laboratory Reports	R (d, p) W (d, p)	
Progress Summary	R (d, p) W (d, p)	

5.2.1.3.1 Log features

The privacy log is created whenever access is granted to some information. Logging is done via collating data from various database sources. The log contains various attributes of interest which relate to the privacy of the stored data. Privacy related attributes are extracted from the log file. These attributes include event time, user id, source IP address, destination IP address as well as purpose and data categories. Moreover the login ID and the number of records accessed are derived from the database access logs.

- Event_Time: The time that the events occur
- Id: The unique id of the system users
- Command Argument: the SQL statements that are parsed
- src_ip: the IP address of the client machine that have initiated a request
- dest_ip: the IP address of the machine that has the data stored
- dest_port: the port number of where the requests are processed
- src_port: the source machine port number

Regular expression techniques were used to extract and analyse system generated event log data and generates necessary alerts. Regular expressions are patterns used to match character combinations in strings. Regex and XML were used to turn log files into searchable data, which makes it easy for the system to pinpoint the exact log entry which violate privacy rules, find the exact time at which the corresponding event had happened, who initiated the activity and also, the location from where the activity originated. The algorithm outlined here explains the log filtering and analysis process.

1. Initialize string pattern
2. Initialize search string
3. Open log file for reading
4. Create an xml writer object instance
5. Initialize xml writer
6. Compile a pattern instance
7. Read each line in the file, use a pattern instance to obtain a matcher instance
8. Use the matcher to find the matches of the pattern of a pattern in a text

a. Pattern p = Pattern.compile(stringpattern);

b. Matcher m = p. Matcher (searchstring);

c. boolean match = m.matches( );

9. if (match.find()) then

10. (int i= 1; i<match.group.count;i++)

11. Parse and write to xml log file

12. Send appropriate alert(s)

13. end

14. End if

5.2.2 Monitoring Activity Diagram

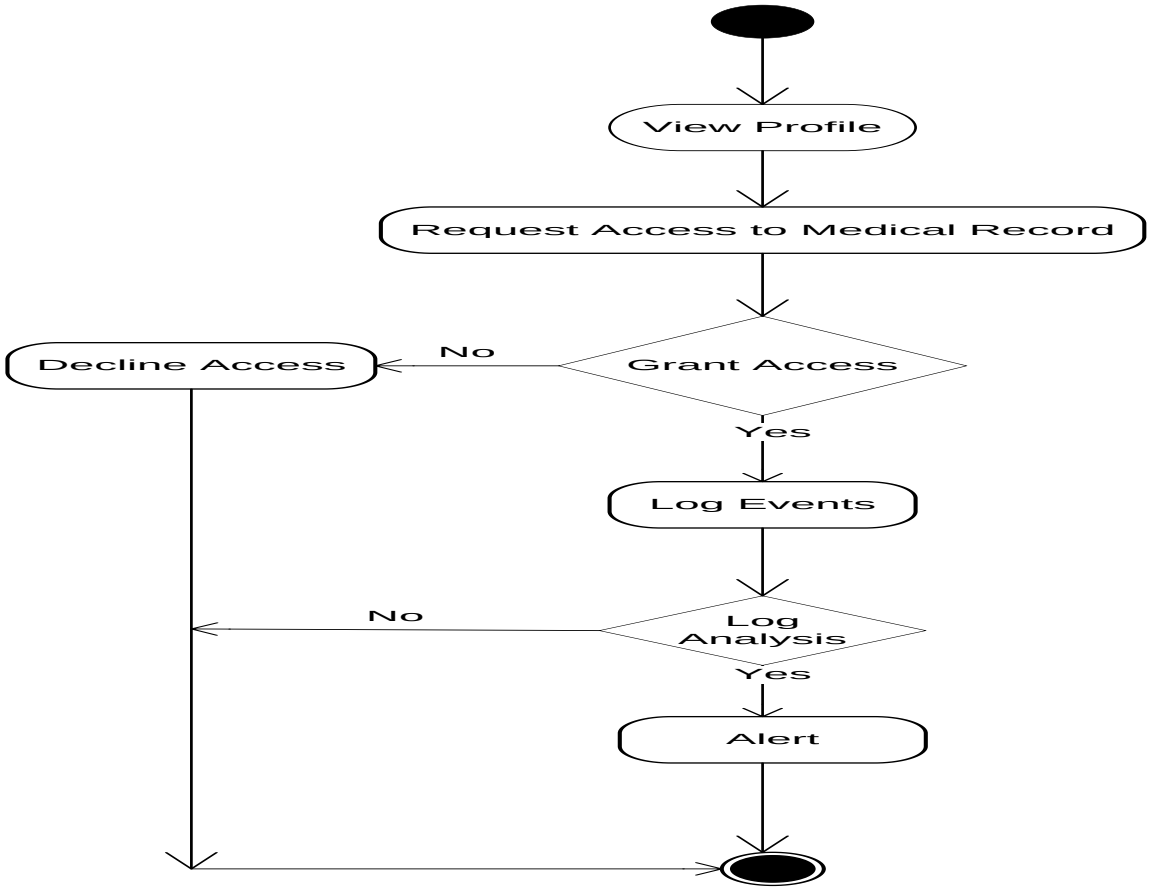

Figure 0.2 Activity Diagram

5.2.2.1 Booking

Appointment booking module allows users to book a new appointment, cancel an upcoming appointment and reschedule upcoming or missed appointment for the patient. The patient is only allowed to cancel their appointment a maximum of three hours times before the appointment time.

5.2.2.2 Databases

The database stores the health records and personally identifiable information together with the access rights delegated in a tokenized format. Security of personal health records from health care providers is very important as imposed by HIPAA and HITECH in their data security and privacy requirements. To secure this sensitive and private information obfuscation is used. The PII is obfuscated before being stored in the database.

5.2.3 Class Diagram

`

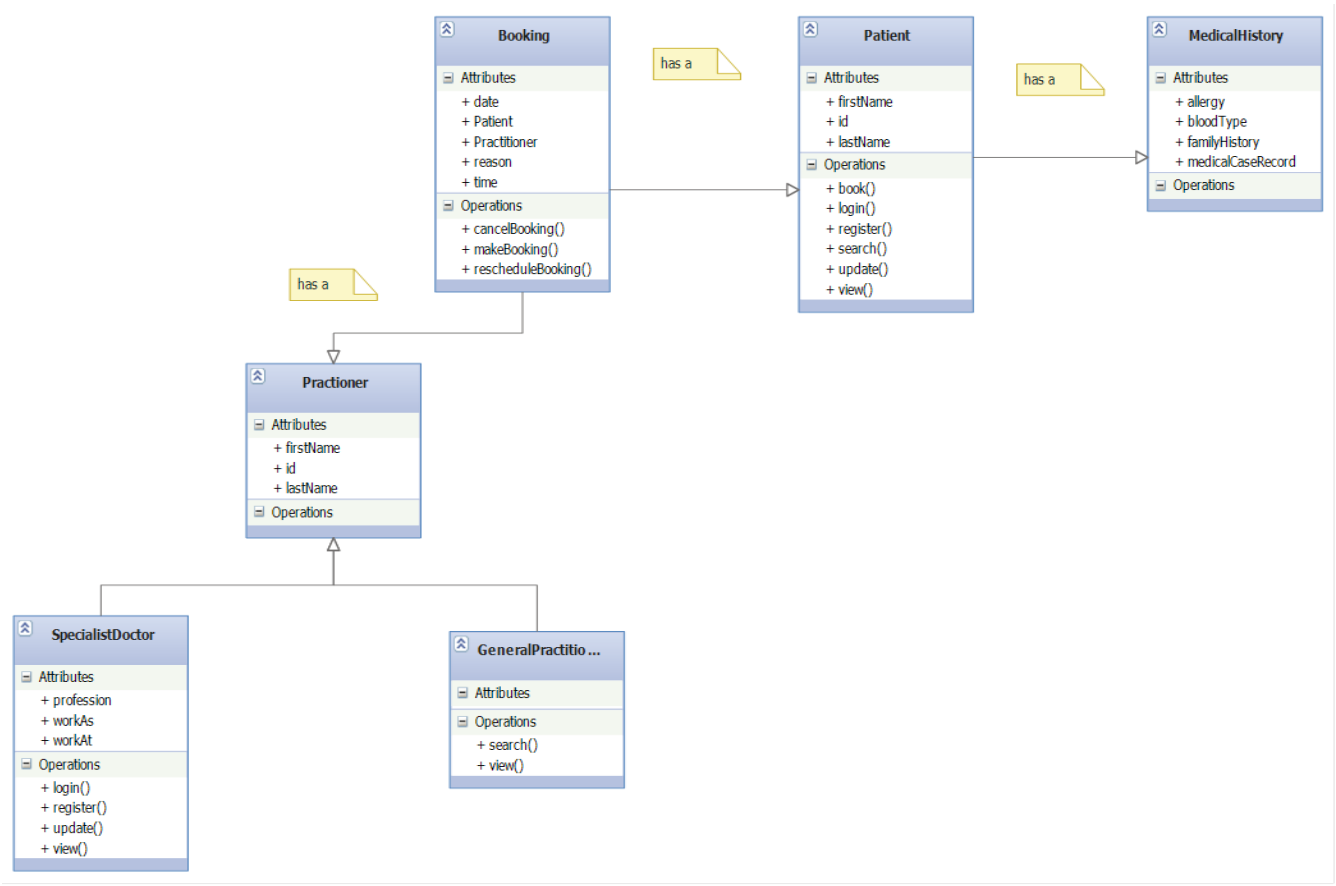

Figure 0.3 Class Diagram

Figure 5.3 presents the classes for the patient, doctors and booking from the medical point of view, with the medical class which specifies any previously reported allergies, intolerances and adverse reactions, which forms the entire patient's medical history; it also specifies various triggers, related to drugs, diet or environment as well as specifying the administered Immunization and vaccinations medication received for solving these problems in the past.

5.3 Implementation of a Privacy Monitoring Framework

The implementation details of how the privacy monitoring framework was realized in this research work are presented in this subsection. Open source community has been developing a wide range of cloud computing platforms lately, all with a common goal of achieving the requirements of an Infrastructure as a service (IaaS). In order to successfully achieve this goal they provide computing resources such as memory, CPU, Network bandwidth as well disk space to their multiple subscribers.

The hypervisors running on the back end enable the middleware application in the creation of Virtual Machines which in turn enables the emulation of physical computers, each equipped with its own CPU, network resources, memory and disk. There are several solutions for open source cloud computing platform focusing on different areas and ranging from hardware resource outsourcing to user services providing services.

Each solution represents a different vision about cloud architecture and implementation. Moreover, each approach has an implication that directly affects its business model. For the selection of a suitable cloud computing platform, this work explored three different cloud

platforms mainstreams namely, CloudStack, OpenStack, and Eucalyptus. The study mainly focused on the following features:

Scalability – Personal information needs to be instantly accessible to the authorized parties wishing to make use of it through a variety of devices across the globe, a cloud service provider needs to able to handle a large scale of workload with concurrent frequent readers and writers to a database record.

Architectural Modularity- Modularization of different communicating components will make the contribution and troubleshooting to be more specific and manageable. Modular design enables easy integration with some legacy or third-party technologies.

Openness and Compatibility – Compatibility with other existing cloud service provider is essential since this will enable the seamless transition from one cloud service provider to the next should any unhappiness with the service offered by the cloud service provider arises.

From the criteria outlined above OpenStack was chosen, since its main characteristic is object storage "swift" which creates petabytes of storage using a cluster of servers for longer term storage and meets the requirements specified above as opposed to others.

5.3.1 OpenStack Architecture

This section presents the working principles of OpenStack. Amongst other OpenStack components are OpenStack Compute as well as OpenStack Glance Imaging Service. OpenStack Compute manages and provision networks of virtual machines for the greater scalability of OpenStack computing platform while OpenStack Imaging Service register, discover and manages massive libraries of server images (Crago et al., 2011).

OpenStack in its entirety is a collection of open source technology projects that provides an operating platform for orchestrating clouds in a massive scale (Wen, Gu, Li, Gao, & Zhang, 2012). Openstack presents the best storage solution to organization with varying storage needs motivated by both performance and price requirements. OpenStack supports both Object Storage and Block Storage, with many deployment options for each, depending on the use case (Von Laszewski, Diaz, Wang, & Fox, 2012). Openstack presented the best solution as per this research work is concerned since it deals with highly sensitive health record information which requires optimum privacy and confidentiality measures. Openstack enabled this study to narrow its focus into the object storage component as well as dashboard an identity, in order to monitor the operations and events carried out on this outsourced massively private and confidential data. The conceptual architecture of OpenStack cloud is presented below:

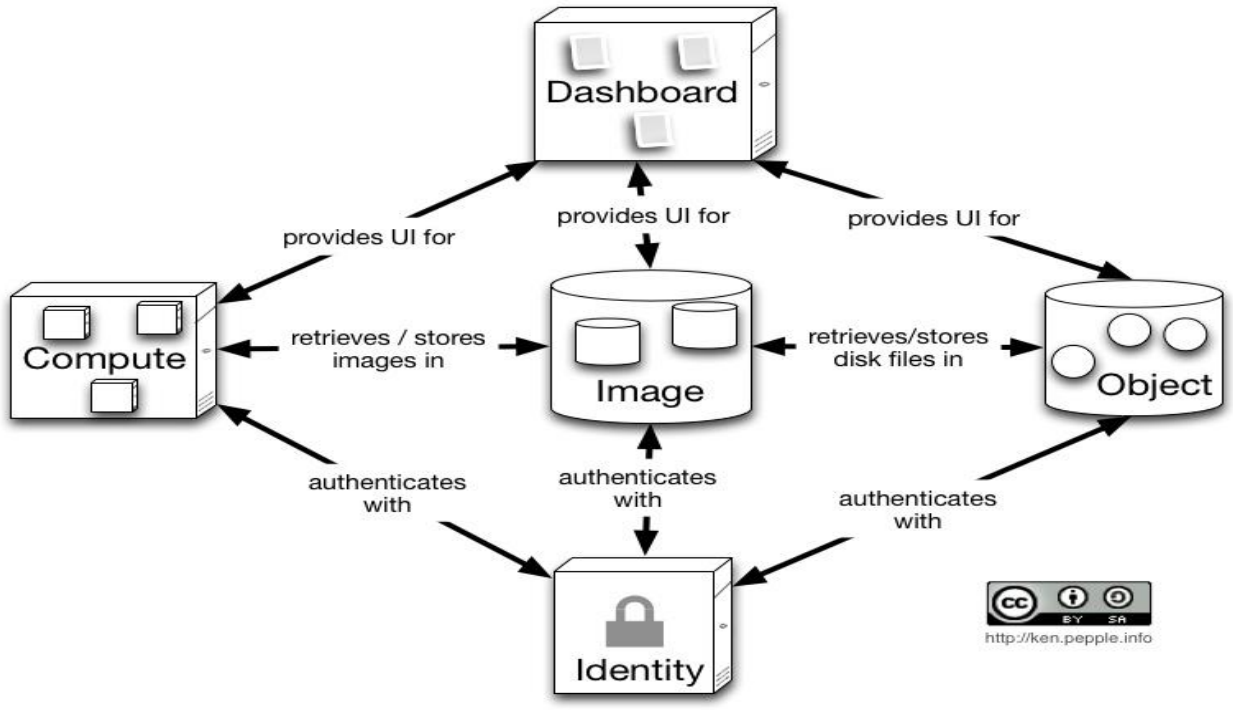

Figure 0.4 Conceptual OpenStack Cloud Architecture (OpenStack, 2013).

Dashboard provides a web application module that enables the end-user and administrator to interface with the services offered by the OpenStack cloud computing platform as shown in the Figure 5.4.

5.3.2 Implementation Environment

Two Intel (R) Core (TM) machines that support the virtualization technology machines were used, one of the machines has i3 CPU with a processing speed of 2.13 GHZ and a RAM size of 4.00 GB. The second machine has as i5 CPU with a processing speed of 320GHZ and RAM size of 3.00GB. With the aid of VirtualBox virtualization software, Folsom, OpenStack cloud computing platform latest version was configured and installed, then deployed on the virtual machines that support the virtualization technology to power up the virtual machines.

As a prerequisite for OpenStack installation, Virtual machines had to run Ubuntu 10.04LTS as an operating system. Object Storage (swift) project was then installed on the virtualized environment to emulate a cloud data centre. In order to provide an API for customers, a web monitoring service which can be accessed using a standardized communication protocol was developed and exposed using Netbeans 7.1.1. The web monitoring services monitors the operations carried out on the outsourced data by system users. This web monitoring service was deployed in three machines that were used when interacting with the system. The example installation architecture from Administrator Guide for deployment of OpenStack Object Storage was followed. Keystone and quantum server were added to a virtual machine alongside the object storage node to store user data, with the quantum server running the virtual network. MySQL database was also added to store and log the sequence of events generated. Keystone

provided the necessary authorization and authentication to the cloud resources via the configured web based management interface.

Three network interfaces were used and they were configured to communicate through the same local area network. Each serving to achieve a different purpose, interface 1 served to manage virtual machines. The second interface handled traffic between two virtual machines. Quantum server used Interface 3 to route the traffic between instances to the internet so to that the instances can be publicly reachable. Java programming Language was used to implement the solution. Java Persistence API, Enterprise Java Beans (EJB) and Java Server Pages (JSP) technologies were used.

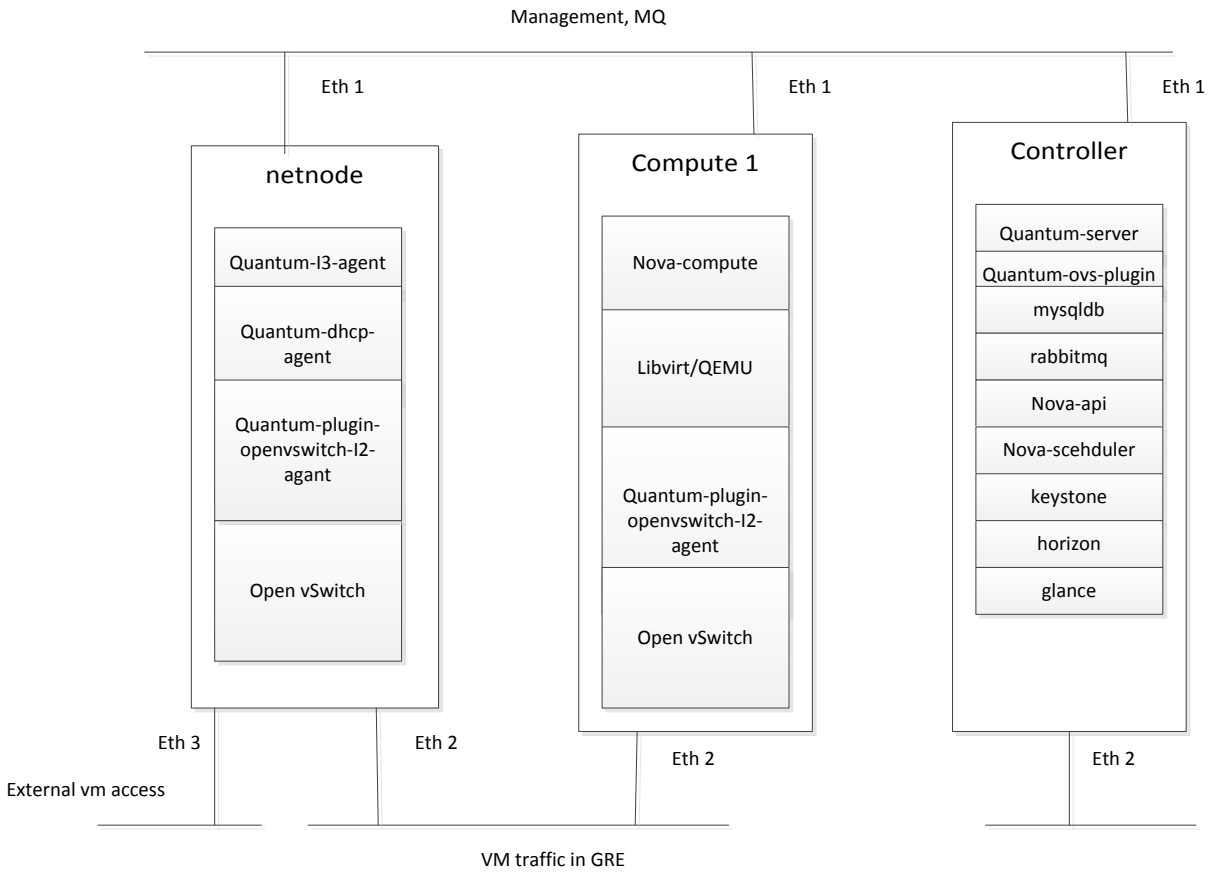

Figure 0.5 Overview of the installation (OpenStack, 2011).

To illustrate the execution of result of our framework, a number of interfaces were designed as shown in Figure 5.5 to enable the users to interact with the online doctor application in assisting them to conduct their health related functions.

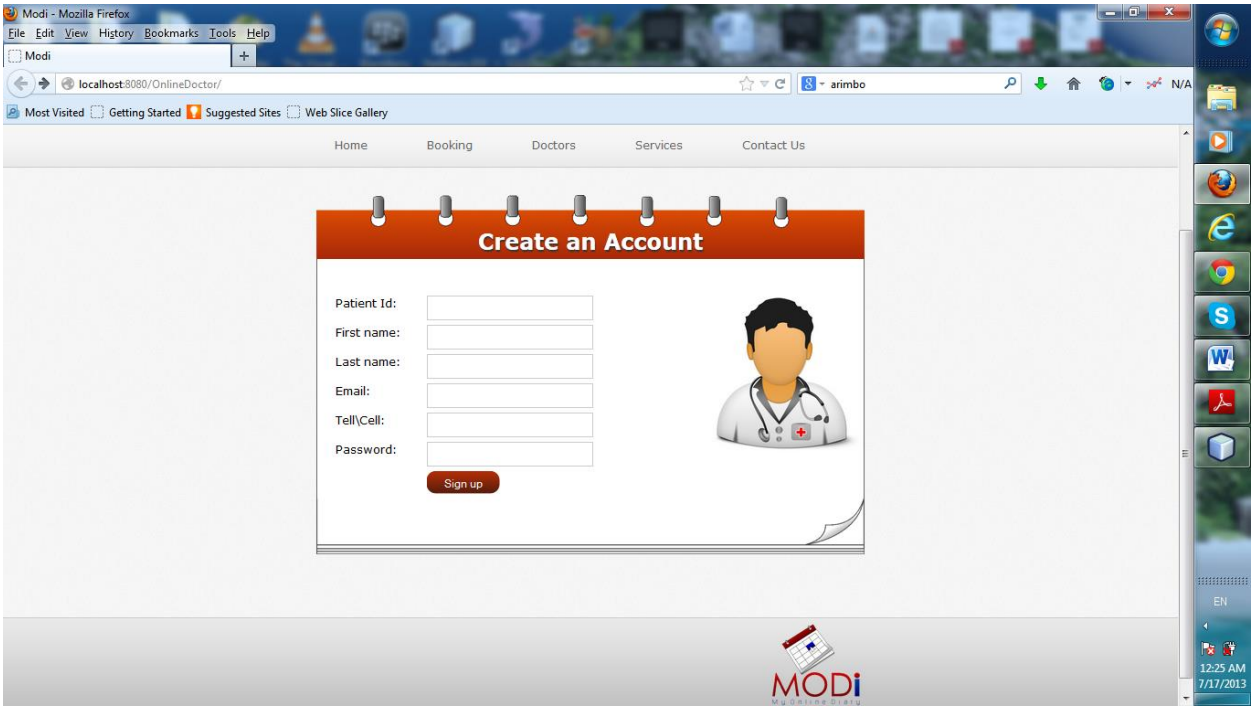

Figure 0.6 Home Page

Figure 5.6 shows the initial home page which serves as a starting point of the online doctor on health. It enables the user to register before making use of the system and enable the registered users to enter their credentials and to be logged into the system.

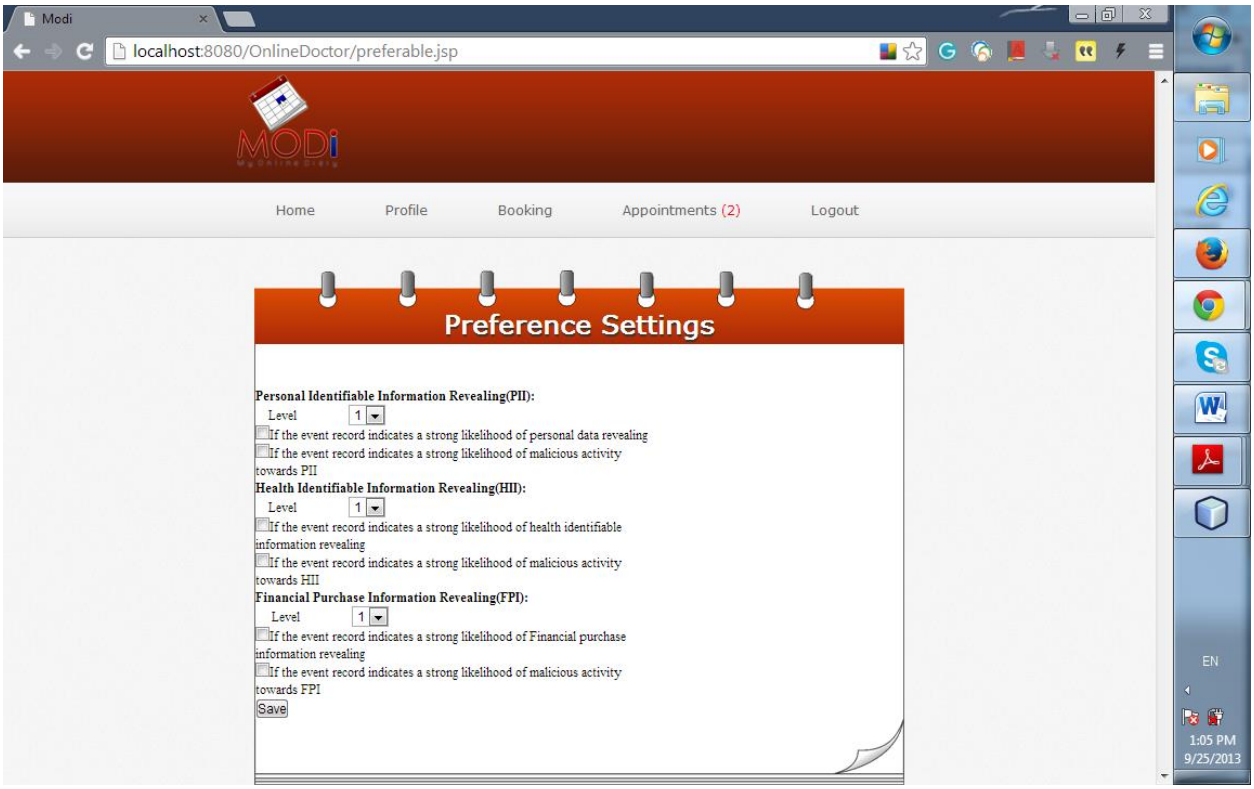

Figure 0.7 Preference Setting Interface

The above interface in Figure 5.7 enables the users to set privacy preferences and clicking the relevant check box that appeal to their preference for the security measures that need to be applied to the personally identifiable information and health identifiable information. The references made here dictate which kind of privacy notification he/she receives. The client sets the access right delegation of the interface below.

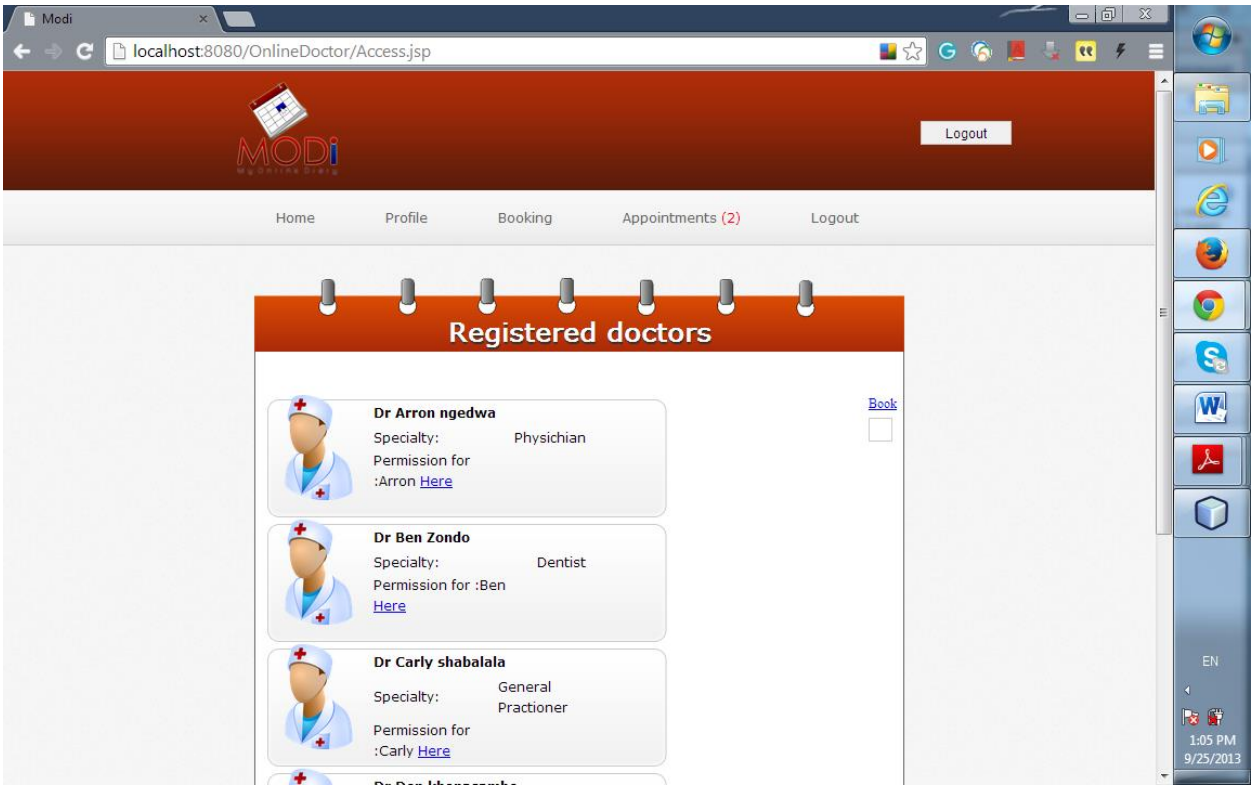

Figure 0.8 Access Rights Delegation Interface

Figure 5.8 is an interface that shows the list of medical practitioners that the user may wish to delegate access right to. The user may choose to give total access right to the practitioner according to their area of specialisation. The user may choose to give the viewing rights or restrict editing rights to medical record.

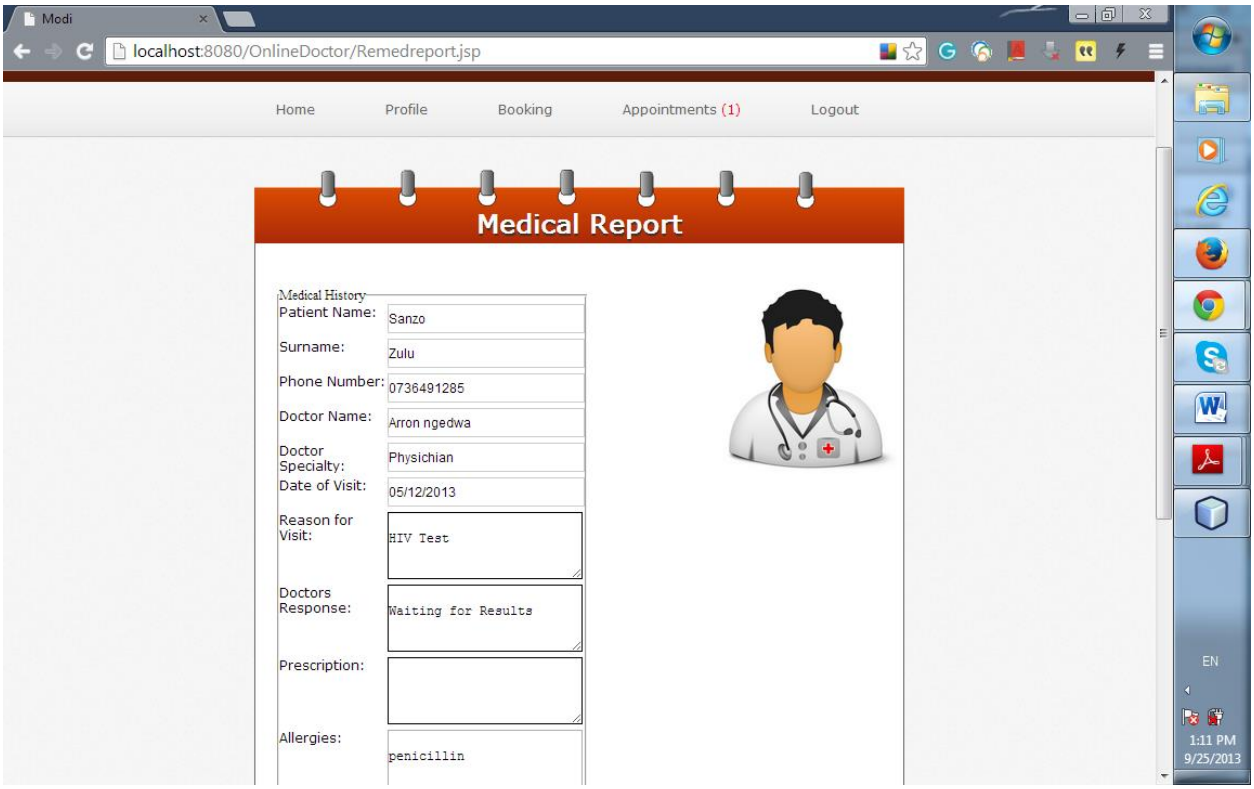

Figure 0.9 Medical Record Interface

The interface in Figure 5.9 shows the medical record which contains information such as age, medical conditions, medications, zip code, hospital, physician, health care, insurance, and other data points. This kind of information is strictly private and confidential.

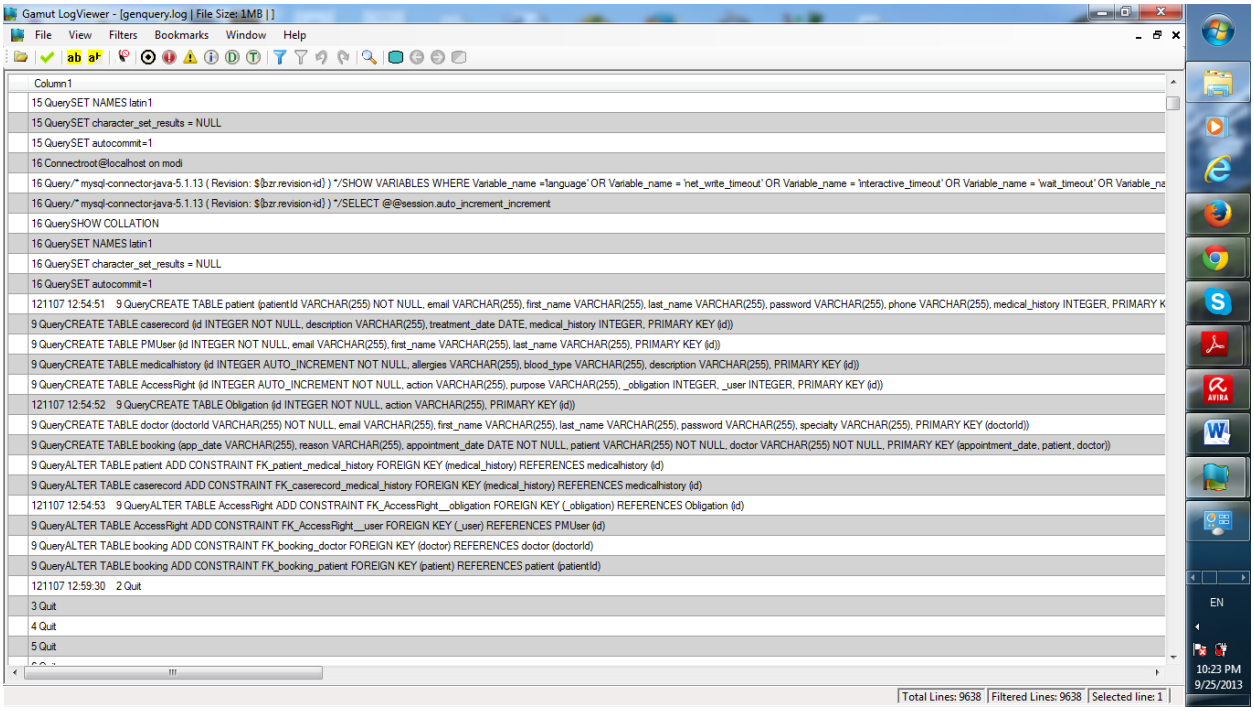

Figure 0.10 Generated Logs Events

Figure 5.10 is the interface that shows activities and events which took place on the stored data, this information is made to be useful through the application of filters and get analyzed for the privacy violation actions and alert the necessary party if there is a need for such action.

5.3.2.1 Experimental Environment

To evaluate the performance of the implemented solution in supporting the privacy monitoring in the cloud environment, an online doctor application was implemented which manages patient's health records. The web service client invokes the service and capture performance measures. Two machines were used for the experiment, the first computer acted as a server and the other computer as a client. The two machines run on Windows 7 professional, 32-bit operating system.

5.3.2.2 Jmeter

JMeter is a testing framework for Apache Software Foundation designed to be used as a load test tool for analyzing and quantifying the performance of system applications among other capabilities (Beat, 2003). It was used to simulate a heavy load of requests to the web service to measure response time and throughput performance under different loads. Figure 5.11 shows the interface of this tool which is able to run on all major operating systems including Linux, UNIX and Windows. JMeter is a 100% Java application and runs correctly on any system provided it is Java Virtual Machine (JMV) compliant.

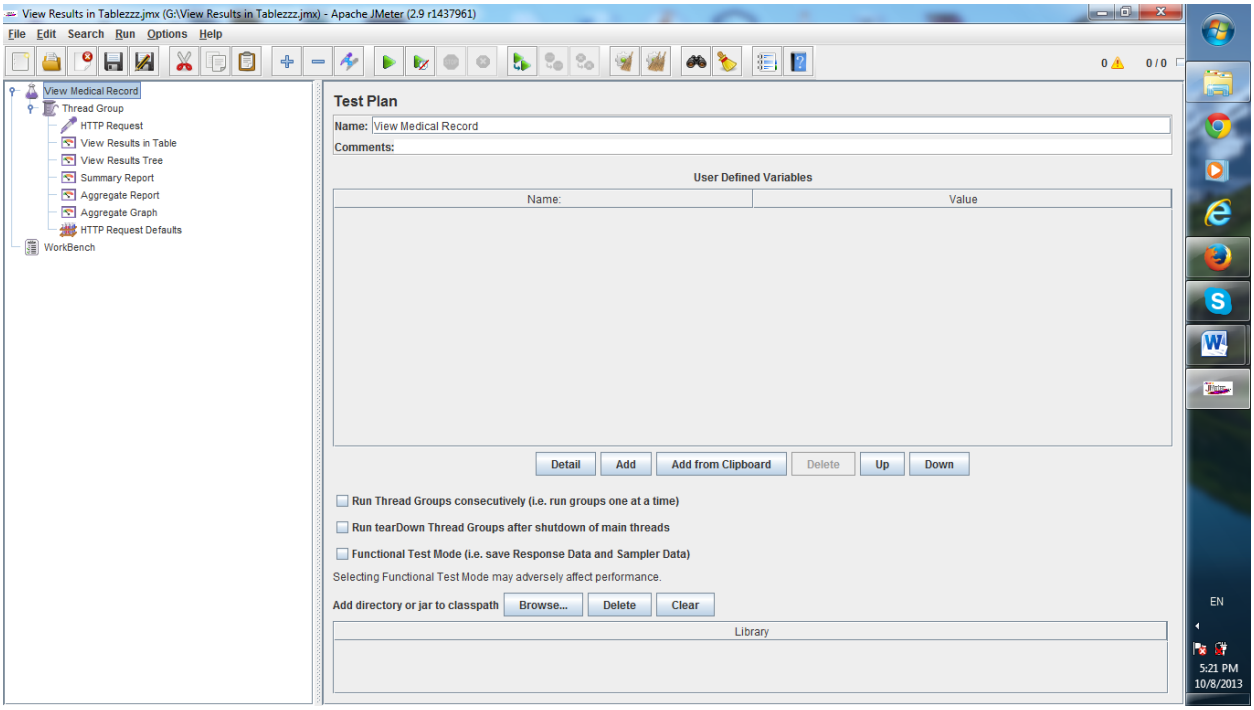

Figure 0.11 Jmeter's load Generator interface

5.4 Performance Evaluation

For the performance evaluation of the proposed solution, both the qualitative and quantitative evaluation were used. The evaluation of the solution focuses on the metrics that evaluate the satisfaction of the goals of the system user.

5.4.1 Quantitative Evaluation

Under the quantitative evaluation aspect we consider the following metrics scalability, throughput and response time.

5.4.1.1 Scalability

In order to determine the effect of the increasing load, the scalability experiments were carried out, observing closely the response time and throughput, since scalability is cloud services, most prominent attribute which helps business handle the exponential demand of the service to its consumer without having to invest in the new infrastructure or increase infrastructure maintenance budget.

5.4.1.2 Throughput

This experiment was conducted in order to measure the time (T) it takes to process transactions (Tr) over a given period. The equation to calculate throughput is as follows:

$$Throughput = 1/n \sum_{k=0}^{n} \mathrm{TrT}$$

Where n is the number of request and k represent the request at index, Table 5-2 show the dataset obtained via Jmeter load generator.

Experimental design

To test the effectiveness of the developed framework experiments were conducted to evaluate the throughput scalability of the solution. To achieve this, we observe the effect of increasing medical history requests ranging from 500 requests up to 5000 to obtain the framework throughput. The data gathered from this experiment is shown in the table 5-2 below.

Table 0-2 Data Gathered for Throughput Experiment

No of Request	Throughput(Tps)
500	4.952903
1000	6.798244
1500	9.798244
2000	15.23856
2500	22.61318
3000	27.63182
3500	28.93182
4000	29.20395
4500	29.40224

Figure 5.12 chart below shows the throughput of the presented solution when increasing the number of active clients during the test. Despite the increasing in the load from 10 to 500 concurrent clients requests, it is observed that the graph doesn't change it fluctuates slights and but stays at the same average level regardless of the increased in the load level. From this graph

it can be deduce that transaction processing time, gradually increases as a number of requests significantly climb from 1000 to 5000 as expected.

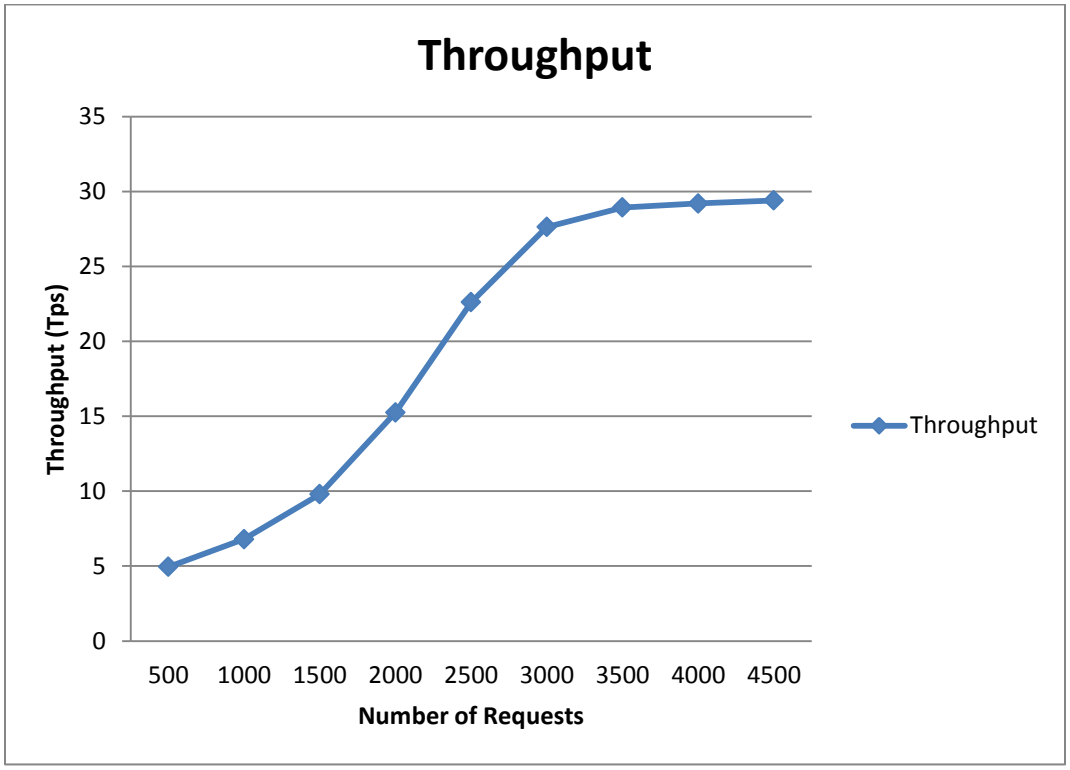

Figure 0.12 Throughput

5.4.1.3 Response time

This experiment was conducted to investigate the response quality of response time with respect to increase in the number of requests. Response time is the amount of time it takes for the system to return the output as a way of determining the systems capability to withstand increasing load. It was calculated as follows:

Average Response Time= $1/n \sum_{k=0}^{n} \binom{n}{k} Tr - Ts)$k ,

Where Tr in this case denotes the time at which the desired response was received, and Ts denotes the time the request was sent out while k denotes the request at index

5.4.1.3.1 Experiment

This experiment was conducted to determine the response time quality of the solution as the number of service requests increases.

a. Experimental design

In conducting this experiment, concurrent medical history requests ranging from 500 requests up to 5000 to obtain the average response time were issued. The experiment was randomized and 25 runs performed with the average being the final response time.

Table 0-3 Data gathered for the Average Response Time Experiment

No of Requests	Average Response Time (ms)
500	3
1000	4
1500	7
2000	9
2500	16
3000	28
3500	36
4000	48
4500	53
5000	80

Table 5-3 show the dataset generated by the JMeter request simulator. In this experiment the load level is ramped up from 500 to 5000 requests. It is observed that the graph is showing steady, gradual increase in response time and as from 2000 concurrent users, the linear increase in the response time graph throughout the test is observed. The framework sustains a high level of load moderately well considering the amount of time needed for decrypting data for each and every transaction.

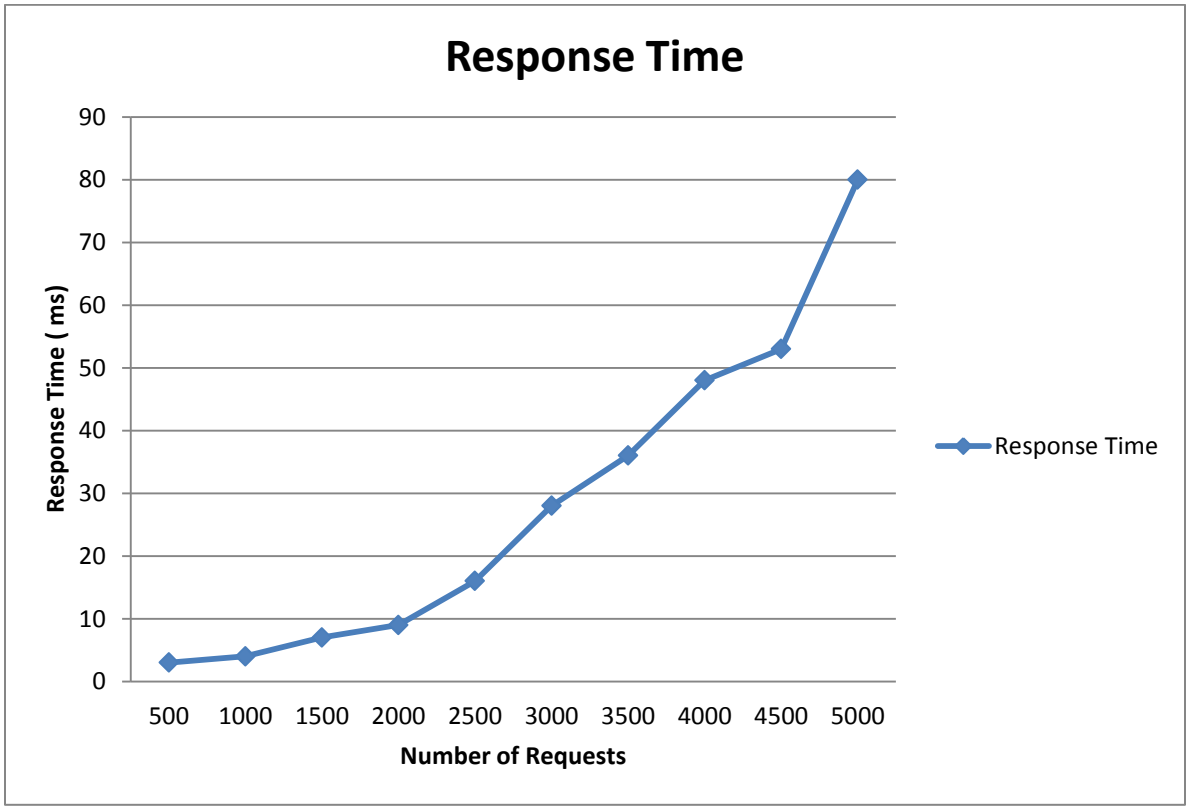

Figure 0.13 Average Response Time

Figure 5.13 shows that as the number of users on the system increases, the response time starts to increase as well. It can therefore deduce that the demand for the same resource is exceeding availability of the resource causing a graceful performance degradation this largely is due to the

amount of effort put into securing the request and the desired response given back by the system. Since security is the main priority, there is a significant tradeoff between the security and response time, the system has to first perform the queries on obfuscated data and then map back the result into the user readable format and secure the message layer in the process. The observed result show that the privacy monitoring framework trades-off response time for optimal data security.

5.4.2 Qualitative Evaluation

Since our solution involves the handling of sensitive personal health data therefore it must conform to privacy principles since it governs the use of personal data. These principles include lawfulness and fairness, proportionality, purpose specification, data quality, openness, transparency and accountability.

These privacy principles have been widely accepted by various countries and have incorporated them in their legislation, reaching the consensus is in terms of economic and geographic correspondence (Langheinrich, 2001, 2002).

Openness: Data subject shall have transparent policies with regard to the storage and processing of personal information. Online Doctor System Users are informed when their information is edited, viewed or deleted by means of an email notifications which is achieved through the informative event and access log component which enabled the user to control and comprehend what happens to their data while it's stored on the service provider's servers.

Users give informed consent to all processing of data, through the click through user agreement which requires the users to click the agree button by which the user is agreeing to the collection and processing of his personal identifying information.

Proportionality and purpose specification: This means that only relevant information must be subject to storage and processing in the cloud service provider's servers and it must be made use only for the intended legitimate purpose of collection (i.e A patient's name, medical record number, identity number, and other data fields that directly link a patient to their data) in this case. Our system collects only minimal information necessary to perform the relevant system evaluations. The collected information is only used within the scope of what it was collected for as permitted under applicable laws i.e. POPI bill, ECT act.

Lawfulness and Fairness: The personal data of the data subject must be fairly processed in accordance with privacy regulations. This must be done with the data owner's consent, and compliance with the purpose and HIPAA principle of storing and processing of the electronic health records as well as data protection laws such as POPI bill, ECT act. The solution implementation was guided by these privacy principles and regulations.

Data quality: Personal data must be as correct and as accurate as possible and not retained beyond the period in which it was intended. The Cloud service provider must allow only the legitimate data subjects to examine and modify some data attributes of them where necessary. And this information should not be retained beyond the period for which it was intended. The integrity broker provides assurance that the data has not been tempered with. The basic requirement that this functionality meets is that when the data is retrieved from the cloud service provider, the data is in the same state as the originally outsourced, with not one unauthorized single bit error alteration.

Security measures: Protection against loss and unlawful processing of sensitive data. Sensitive health data is obfuscated before being uploaded to the cloud service. This helps to secure data against hackers and identity thieves who might hack into the system.

Usually typical business queries can be executed on the obfuscated data stored in the cloud, and results from the queries can then be mapped back to a readable format once an appropriate key known only to the client is used to de-obfuscated the data, the key is managed by the client and the cloud service provider doesn't have knowledge it. This brings about a solution to the constant challenge of secondary data usage by the cloud service provider without the consent of the data owner.

Accountability: The data cloud service provider is liable to comply with all the obligations set out by the Madrid resolution, and it must have measures in place to demonstrate such compliance to the data subject and the regulatory authorities. The events are recorded in the log file and automatic alert is sent to the data subject instantly when an actionable event is encountered.

Trans-border Transfer: Exchange of personal data to other countries is permitted only if that country offers adequate privacy protection regulations. Data location has a significant effect with regards to its security and privacy laws applied to it. For instance, the data located in America would be subjected to USA Patriot act, which allows the court to issue orders to force disclosure of personal data to the federal bureau of investigation (FBI).

Once served with this order the service is required by the law to disclose every bit of data the CSP might have in its custody regarding the person or organization in question. It is therefore recommended that the data subject knows where the cloud provider's data centers are located.

5.4.2.1 Privacy Impact Assessment (PIA)

Privacy impact assessment was used to further evaluate our solution, and to achieve this a privacy impact assessment guide was used. It enables us to evaluate if our solution meets legislative privacy requirements when monitoring the health and personally identifiable information of the data subject since our solution handles both as illustrated in Table 5-4. Privacy Impact Assessments (PIAs) are particularly used to identify the potential privacy risks of new or redesigned federal government programs or services.

PIAs also help eliminate or reduce those risks to an acceptable level. Furthermore PIA is a procedure which helps assess privacy risks to individuals in the collection, use and disclosure of information. It helps to assess legal and data protection compliance (Government of Canada, 2010). PIA has helped to minimize privacy impacts by analyzing the possible privacy impacts on individuals' privacy.

Table 0-4 Privacy Impact Assessment Criteria

QUESTION	***Yes***	***No***	***N/A***
Are privacy policies or procedures in place to ensure that:			
There is a business purpose for all personal information	✓		
Individual consent is obtained whenever possible	✓		
Individuals are duly informed of the purpose and authority for the	✓		

collection			
Information about personal information collected is readily available to individuals	✓		
Personal information correction and annotation are available when required	✓		
Physical records are appropriately stored and managed to maintain privacy			✓
Are privacy controls in place in the organization?			
Need-to-know policies and procedures for personal information access	✓		
Physical security and access controls			✓
IT security and access controls	✓		
Waste management controls for personal Information			✓
Records management & disposition schedules			✓

The legitimate and business purpose behind the collection of personally identifiable information in this case was to evaluate the framework and to prove its effectiveness in the real world situation. Individual's consent, which provides the primary basis for the collection, use and

disclosure of personal information for this project were obtained through the click through user agreement which requires the users to click the agree button in order for them to continue making use of our solution. Evidently our solution flexibly affords the best privacy protection to personally identifiable information as attested by the number of requirements that the solution meets as shown in the Table 5-4.

This click-through agreement serves as consent to collect, use and disclosure of personal information in line with the privacy standards and regulation in this case POPI bill and ECT act as well as HIPAA standard. The collected information is readily available for individuals to update and view if necessary. The assessment demonstrates that the framework satisfies protection security necessities by tending to and determining expansive privacy concerns from different angles. The Common Criteria security evaluation method could have been used also. However, it was not used because it mainly focuses on the generic security functionalities and their assurance. Furthermore, the process of certification is a complex procedure involving several independent actors, due to the time limit this course of evaluation couldn't be pursued.

5.4.2.2 Usability Test Method Results

Privacy principles, derived from privacy legislation, influence the Human-Computer Interaction (HCI) (Angulo, Fischer-Hübner, 2013). HCI is the investigation of mental courses of action and conduct as they relate to clients interacting with computers (and other technical devices). Interface design is important for privacy on the grounds that the users interact with the product or service through the interface, so all the privacy features must be represented in the interface design. Table 5-6 shows how the requirements were met.

Online Doctor is implemented such that its users know who is accessing their data and what reasons, and this is achieved through giving the user the **control** to their health and personally identifiable information, and to **comprehend** the reason for processing of this information and the comprehension of all the processes that are applied to their personal data and for them to be **aware** of the options they have towards getting the necessary transparency in the activities carried out on their data.

Online Doctor is built such that it will enable its users to have the necessary **understanding and knowledge** with regards to the handling of their data:

- Comprehend to know how their personal data are handled
- Know who is processing their personal data and for what purpose
- Understand the limits of processing transparency
- Comprehend the limitation on objects to process
- Be truly informed when giving consent to the processing
- Understand when a contract is being formed and its implications

The Online Doctor gives its user the **control and power** they need for the manipulation of their medical information:

- Control how their medical records are handled
- Control who has access to their medical
- Be able to examine and update personal data

The Online Doctor is implemented such that its user's **conscious** of the activities carried out on their personal and identifiable health information:

- Aware of the transparency options that the application offers
- Get the necessary information when the personal data and medical information is being processed
- Be aware of what happens to personal data once the retention period has expired

Table 0-5 HCI Requirements, and Design Solutions

HCI	**Solution**
Control	Preference setting, Access and integrity broker, Alert
Consciousness	Preference setting, Alert
Comprehension	The alert component is responsible for alerting the user of the useful actionable critical events. Alerting will be triggered only if there are any status changes.
Consent	Click through user agreement

5.4.2.2.1 The System Usability Scale (SUS)

The degree to which the product can be utilized by specified users to accomplish detailed objectives with effectiveness, efficiency, and satisfaction in a specified context of use is known as Usability as defined by the International Organization for Standardization (ISO) (Scholtz, 2013). This section presents the results of the usability test of the proposed solution, derived from using the System Usability Scale (SUS); SUS is a simple ten-item scale giving a global view of subjective assessments of usability.

It covers a variety of aspects of system usability, such as the need for support, training, and complexity, and thus has a high level of face validity for measuring the usability of a system (Brooke, 1996). The system usability scale is generally used after the participants have had a chance to interact with the system being evaluated, without thinking about their experience for a

long time the respondents were asked to record their immediate response to the questionnaires provided for as required by the SU

The purpose of the testing was to determine if the design concepts in the proposed solution were successful in constructing a privacy monitoring system that (1) users can use, (2) users can understand, and (3) users can trust with secure and private information, since the privacy principles require that users **understand** the transparency options, are **aware** of when they can be used, and are able to **control** how their PD is handled. These requirements are related to mental processes and human behaviour, and HCI techniques are available to satisfy these requirements.

i) User Comprehension

To evaluate if users have the necessary **understanding and knowledge** with regards to the handling of their data via the email based notification, which aims to furnish data subjection with an understanding, comprehension to know how their personal data are handled (1) Know who is processing their personal data and for what purpose, (2) Understand the limits of processing transparency, (3) Comprehend the limitation on objects to process. The questionnaire included a question asking the participants if they understand the purpose of various interface elements, terminology, and procedures with regards to the handling of their data. The questionnaire containing six questions is attached as Appendix C.

This test involved laboratory sessions where volunteers interacted with the software system on a computer and answered questions concerning the features and performance that they experienced while utilizing the system. The participants in this study were generally students, with a range of technology skills (Amateur, intermediate and expert users). Participants were asked to indicate their agreement to the statement “I understand the purpose of various interface elements,

terminology, and procedures" on a scale of: 1 ="strongly disagree", 2 ="disagree", 3 ="undecided", 4 ="agree" and 5="strongly agree". Their response was then recorded to yield the following results.

Table 0-6 Statistical Result, Frequency and Percentage for User Comprehension

Statistical Result				Frequency and Percentage for user comprehension			
N Valid	Missing	Mean	Std Dev	Value Label	Value	Frequency	Valid Percentage
60	0	4.50	0.79	Strongly Disagree	1	1	1.67
				Disagree	2	1	1.67
				Undecided	3	2	3.33
				Agree	4	19	31.67
				Strongly Agree	5	37	61.67
					Total	60	100

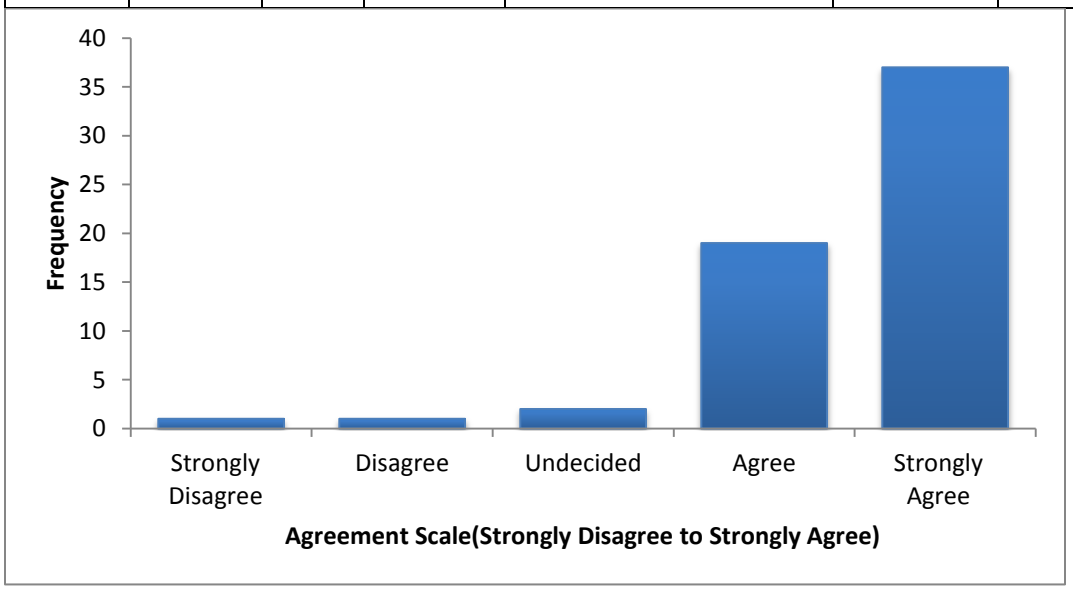

Figure 0.14 Frequency Distribution for User Comprehension

The average rating was 4.50 (SD=0. 79) as shown in Table 5-6, which represents a rating of "Understandable "which is with an agreement with the statement. A frequency count of the

ratings is shown in Table 5-6, and it can be seen that the most common ratings were 4 or 5, but these were offset by a fair number of low ratings (1 or 2 or 3) as illustrated in Figure 5.14. Table 4 also indicates that only 61.67 percent of the concepts and functions were understandable enough for the participants to have the necessary understanding and knowledge with regards to the handling of their data.

ii) User Awareness (Consciousness)

Participants were also asked to indicate their agreement to the statement "I found privacy breach notification to be very informative" on a scale of: 1- 5. In order to evaluate if users are conscious of the activities carried out on their personal and identifiable health information that is to (1) get the necessary information when the personal data and medical information is being processed i.e. Viewed, edited and/or deleted. The participant's response was then recorded respectively to yield the following results.

Table 0-7 Statistical Result, Frequency and Percentage for User Awareness

Statistical Result				Frequency and Percentage for User Awareness			
N Valid	Missing	Mean	Std Dev	Value Label	Value	Frequency	Valid Percentage
60	0	4.58	0.62	Undecided	3	4	6.67
				Agree	4	17	28.33
				Strongly Agree	5	39	65
				Total		60	100

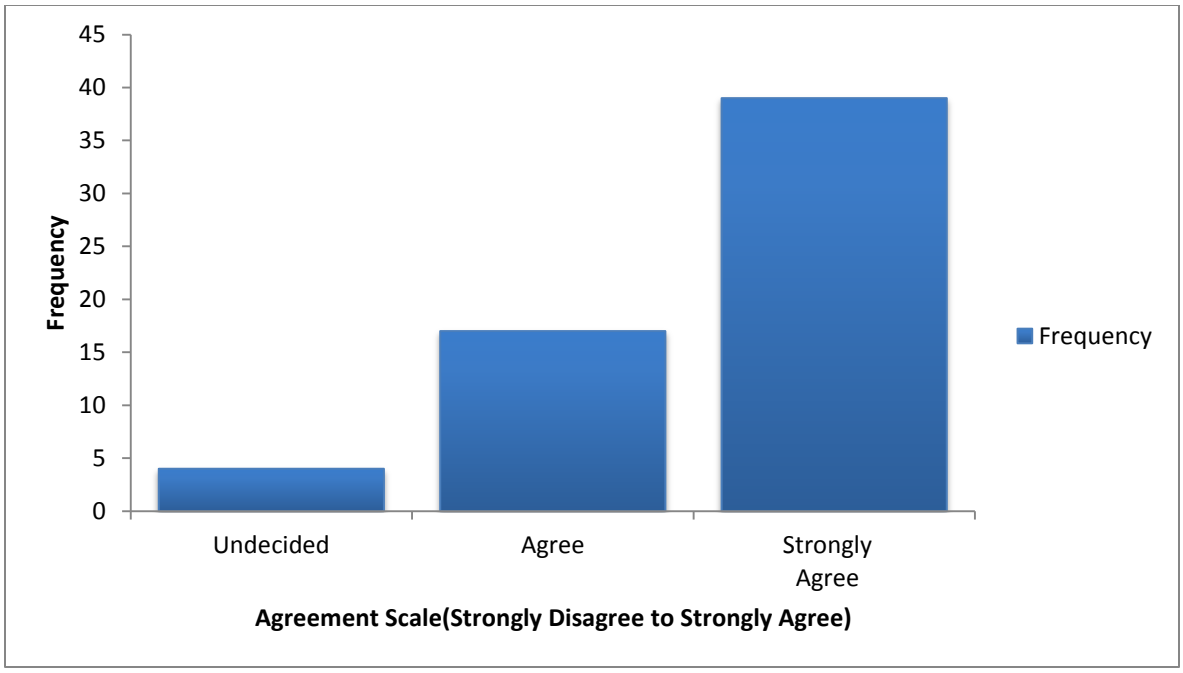

Figure 0.15 Frequency Distribution Scale for User Awareness

The average rating was 4.58 (SD =0. 62) as shown in Table 5-7, which again represents a highly positive result. A count of the number of people who agreed or strongly agreed (ratings of 4 or 5) indicated that 93% agreed with the statement as illustrated in Figure 5.15. Thus, indicating that the email based privacy notifications were informative enough to keep users conscious of the activities carried out on their personal and identifiable health information.

iii) User Control

To evaluate user-control participants were asked to rate the ease of performing the different user centric tasks. They were asked to use a scale of 1 to 5, in order to establish if the solution allow the participants to, (1) control how their medical records are handled, (2) control who has access to their medical information, (3) Be able to examine and update personal data.

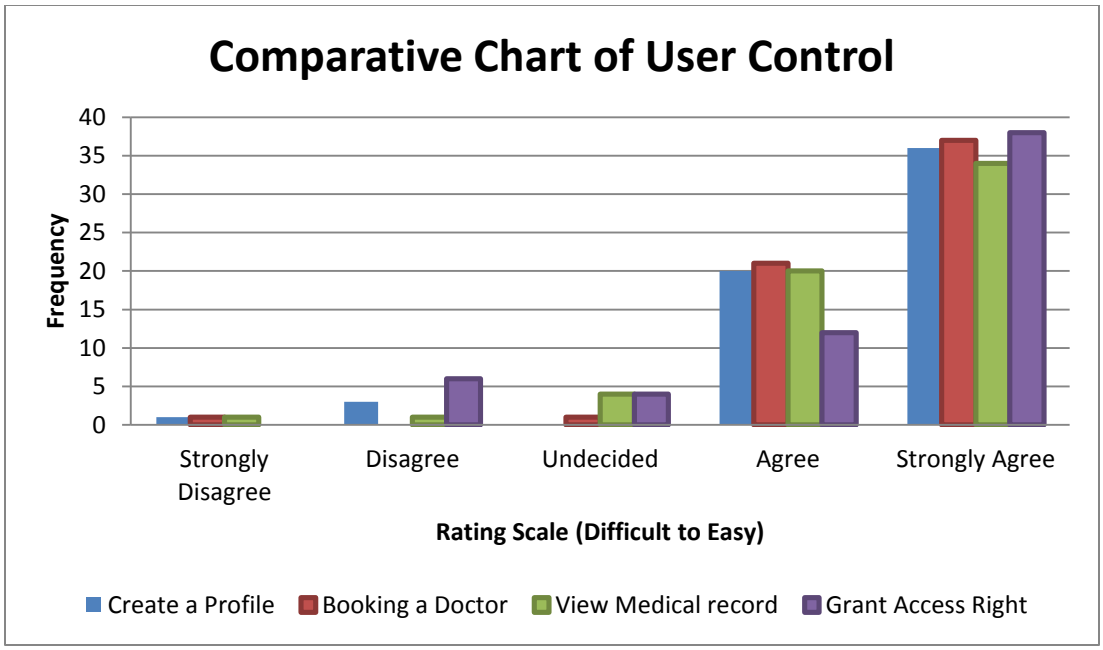

Figure 0.16 Comparative Frequency Distribution Scale for User Control

Figure 5.16 illustrates the comparative frequency distribution scale of ease of understanding control features of the solution presented ranging from very difficult to understand and very easy to understand .As evident above the majority of participants understand the main control features of the solution as expected. These results indicate that the solution developed worked very well, but that there was room for improvement, with the creation of a profile with a standard deviation of 0.87, booking a doctor the standard deviation of 0.70 and medical record the standard deviation of 0.83 as well as grants of access right the standard deviation of 0.99 respectively as shown in Table 5-10 attached as Appendices A. The findings indicate that our solution control features are understandable and therefore easy to be used by the targeted end users to gain control of their outsourced data.

Table 0-8 Summary Percentage for User Comprehension

Function	Percentage of People Understanding

Create a Profile	93
Book a Doctor	97
View Medical Record	90
Grant Access Rights	83

Table 5-8 shows the percentage of participants who understand and accurately carried out all the major functionalities of the service functions. This is an important finding that indicates that the prototype is understandable and easy to use by the targeted end users. In terms of the percentage of users who understand each function, 90% of users indicated that they understand the functionality of viewing medical record and 97% percentages understand the Book a Doctor function very well and a slight variation when granting access right was noted as well when compared to carrying out the different major service functionalities and lastly 93% also indicated that they faced no difficulties in creating their profile as a primary essential step in order for the users to be able to make use of the solution.

iv) Ease of Use

Table 0-9 Statistical Result, Frequency and Percentage for Ease of use

Statistical Result				Frequency and Percentage for Ease of use			
N Valid	Missing	Mean	Std Dev	Value Label	Value	Frequency	Valid Percentage
60	0	4.43	0.79	Strongly Disagree	1	1	1.67
				Disagree	2	1	1.67
				Undecided	3	2	3.33
				Agree	4	23	38.33

				Strongly Agree	5	33	55
					Total	60	100

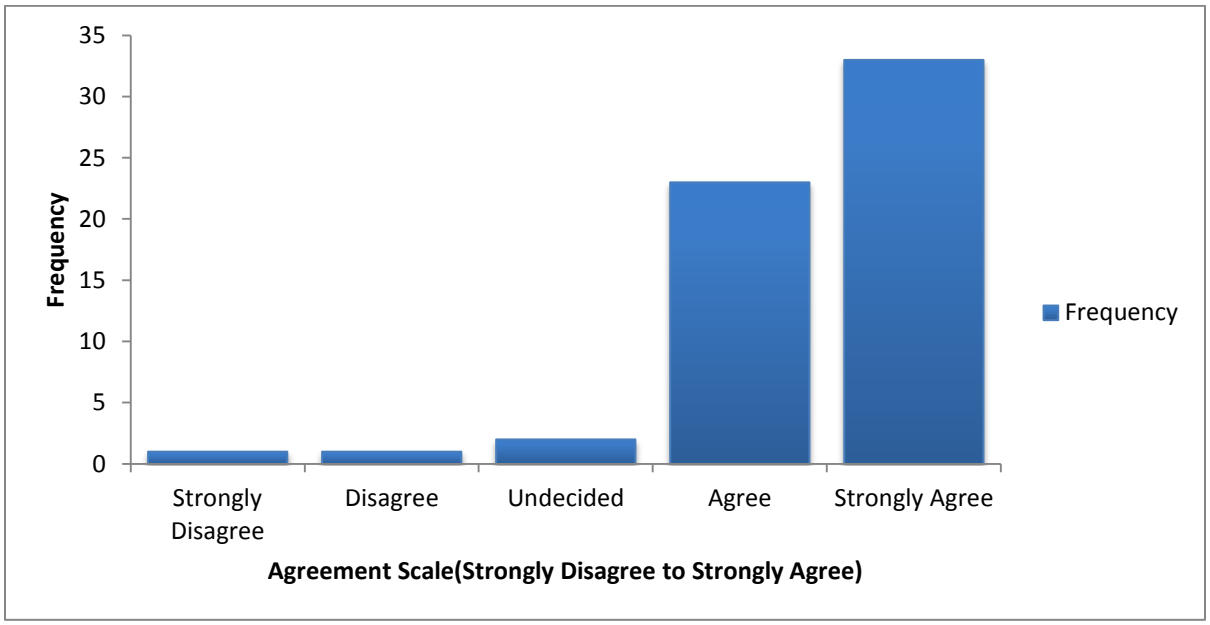

Figure 0.17 Frequency Distribution scale for Ease of use

Participants were asked to rate the ease of using the services with a rating scale from 1 to 5 with 1 being "Strongly Disagree", 3 = Undecided, 5 = Strongly Agree. The average ease of use rating was 4.43 (SD =, 79) as shown in Table 5-9, which represents a rating of "very easy". A frequency count of the ratings is shown in Figure 5.17, and it can be seen that the most common ratings were 4 or 5.

5.4.2.3 Discussion

This study reported on the effectiveness of using informative events and access log analyser as a means of monitoring the operation carried out on the outsourced data. Privacy can no longer be assured solely by compliance with regulatory frameworks that are put in place by the regulatory

bodies, with the aim of enabling cloud customer to effectively monitor the privacy of their data while it is kept in the custody of the cloud service provider.

The core of this research work is under the informative events and access log analyzer component which retrieves and filters event logs for analysis. To achieve this the regular expression technique was used to extract and analyse system generated event log. The log files were then analysed for action events, to find information such as the event time, who initiated and the activity carried out. To simplify the process, the Regex and XML were used to turn log files into searchable data. Log entries encompass all event data that occur on the personal data of each user which then form privacy evidence that are needed to warrant a privacy alert detailing the privacy violation detected. A series of experiments were designed and implemented to prove its applicability in the real world as required by design science methodology. The performance evaluation was measured based on the following performance metrics: scalability, user comprehension, user control and user awareness.

From the results, it is clear that the framework yield good throughput, but the solution experienced some polynomial time complexity since it trade-off response time for optimal data security as a result of time spent on encrypting and decrypting data for each transaction.

For the purpose of validating the technical feasibility of the privacy monitoring framework proposed in this research work, the privacy impact assessment guide was used, PIA is a well-established standard for identify the potential privacy risks of new or redesigned federal government programs or services. PIA enabled the evaluation of the solution to determine if it meets the legislative privacy requirements when monitoring the health and personally identifiable information of the data subject. The assessment showed that the framework satisfies

protection assurance necessities by tackling and resolving broad privacy concerns from different angles.

The usability tests were conducted, the evaluation of the results was based on user-control participants where consumers were asked to rate the ease of performing of the different user centric tasks. Findings of questionnaire were analysed for each criterion and these were explained in both tabular and graphical representations. Validation of the research work is based on available data obtained from questionnaires.

We constructed the survey questionnaire so that it was as precise as possible, used consistent terminology. Our goal was to have respondents answer all questions we therefore, used close-ended response categories as often as we could. Selecting a response makes it easier for most respondents to answer a question. Providing for a unique response assures those who might feel "boxed in" by response selections that do not meet their special needs. We limited the length of the survey so that most respondents could answer all the items within 15 to 30 minutes.

To calculate the SUS score, the score contributions from each item was assumed. Each item's score contribution ranged from 0 to 4. The odd items score contribution is the scale position minus 1 and for the even items 2,4,6,8 and 10; the contribution is 5 minus the scale position. The sum of the scores is then multiplied by 2.5 to obtain the overall value of SU. The obtained average SUS score from all 60 participants is 80. A SUS score above a 68 is considered above average and anything below 68 is below average. The score of 80 that the system obtained indicates that users are likely to be recommending our solution to a friend. An 80 SUS score is interpreted as a grade of a B. SUS scores have a range of 0 to 100. This clearly shows that the most of the users are satisfied with the solution regarding usability but there are some users who

were less satisfied due to some problems they experienced which mean that there are some issues regarding usability that needs to be improved to enhance user interest to the solution.

Data privacy is a very personal thing, and each customer has individual requirements, which presents a challenge for cloud providers to address with standardized offerings. The availability of a wide range of popular toolsets together with appropriate access privileges is one possible approach to address this challenge.

The results given above indicate that the solution control features are understandable and therefore easy to be used by the targeted end users to gain control of their outsourced data. Effective utilization of software functions and features is more probable when clients feel in control of the system and have appropriate flexibility available to tailor the system to meet their needs.

5.6 Chapter Summary

This chapter presented the results for privacy monitoring framework. The experiments on scalability, user comprehension, user awareness, ease of use and user control were conducted. The informative event and access log analyser was employed, which enabled users to retrace in details what actions were carried out in the data, where it is stored and who accesses it. The purpose of conducting these experiments was to determine if the design concepts of the privacy monitoring framework were successful in constructing solution that (1) users can easily use, (2) can help users comprehend how their personal data is handled, and (3) users can control and trust with personal information. This goal was achieved by monitoring user's activities and when an actionable event is detected an appropriate alert was sent with an aims to furnish data subjection with an understanding as to know how their personal data are handled.

Chapter 6

CONCLUSION AND FUTURE WORK

6.1 Summary

The lack of proper privacy and security mechanisms to monitor the sensitive information entrusted to cloud service providers by its consumers has been the barrier to broader adoption of cloud computing technology, as reported by a number of surveys conducted in this area of study. Despite many cloud computing services properties, such as low entry cost, elasticity is perfectly suited to support a variety of business operations and lower the operating costs, However, privacy has remained the most difficult proposition when a business is considering to adopt cloud computing. This owes to the fact that with cloud computing, the storage and processing of private information are done on remote machines that are not owned or overseen by the clients. All that the client can see is a virtual infrastructure built on top of possibly non-trusted physical hardware or operating environments.

The purpose of this research work was to develop a privacy monitoring framework for cloud computing environment, to allay users' privacy concerns and to allow for broader adoption in order for cloud computing to fully achieving its potential. The developed privacy monitoring framework has the potential to help cloud customers to comprehend what happens to their data while stored in the cloud, by employing an informative event which would enable users to trace in details what happens to the data, where they are stored and who accesses it. The framework realized this by using an informative event and an access logs analyzer, which enabled cloud

customer to monitor the privacy of their data while it is kept in the custody of the cloud service provider.

One distinct thing that differentiates our work from others is that, other authors approached solutions from the cloud provider context perspective by allowing the use of cloud or third party to be in full control of consumers data without considering how the consumers can monitor their data. In this work we considered the user control aspect which is mostly responsible for building trust between the cloud customer and cloud service provider.

6.2 Results

Cloud computing offers over the Internet services that lower IT capital Expenditures and reduces business operating costs. These services offer on demand capacity with self service provisioning on pay per use basis for greater flexibility and agility. However, given all these benefits cloud computing is still hounded by the escalating data privacy and security concerns that continue to impede widespread adoption of cloud computing. These concerns emanate from the lack of knowledge of the physical location of cloud service provider's servers where the cloud consumer data is stored and processed and what privacy and security safeguards are put into place to protect the data outsourced.

Cloud computing provides a relatively easy and cost effective way of processing data, however, the securing and controlling of personal information in the cloud is a huge challenge

In Section 1.4, three distinct research questions for the dissertation were defined. This section provides highlight of how these questions were answered.

RQ1: What are the problems with the current privacy protection mechanism for cloud customers?

The state-of-the-art in privacy protection mechanism for cloud computing has been discussed in Chapter 3. This study discovered that most of these approaches have focused on providing data privacy monitoring mechanism from the context of cloud providers which may not really provide adequate private data security. This study also identified the lack of means of empowering cloud consumers with tools of verifying if the security measures being put in place to protect their data. Though the safeguard of data by cloud providers is very paramount and highly imperative, the issue of how and who access this data for processing and storage is also important to data owners. As the result of this investigation, several ideas were synthesized and a privacy monitoring framework for cloud computing environment to address the aforementioned shortcomings was crafted.

RQ2: How do the existing privacy and compliance regulations affect the implementation of cloud technologies?

Assessments of the current security and privacy regulation have been carried out to determine the impact they have on the implementation of the cloud technologies in Chapter 2. A review of the published literature revealed that the majority of the privacy regulations were made long before the term cloud computing materialized and therefore they these regulations don't provide stipulations on how organizations should implement cloud computing. Complying with the plethora of security and privacy regulation exist within distinctive nations is a challenge to organizations implementing cloud computing technologies. However, organizations are legally

obliged to comply with privacy regulation for the protection of personal data. Therefore, organizations must put into place satisfactory system of measures and procedures to achieve this.

RQ3: How can we design and validate a privacy monitoring framework for the cloud environment?

Chapter 4 has proposed the Privacy Monitoring Framework; consisting of five major component, informative events and access logs analyzer, access right delegator, personae, preference setting and Alerting component. The crux of work is under the informative events and access log analyzer component. This component retrieves specific logs using user preference parameters/ filters for analysis. These log files are then analyzed for action events. An event record is only regarded as actionable if and only if the event record indicates a strong likelihood of malicious activity. Log entries encompass all event data that occur on the personal data of each user which then form privacy evidence that are needed to warrant a privacy alert detailing the privacy violation detected. The information helps the data subject to have the necessary understanding and knowledge with regards to the handling of their data, it enables the data subject to retrace what happens to their data, where it is stored, who accessed it and what levels of security and privacy safeguards are applied to it to guarantee as much transparency to the consumers. The validation of the privacy monitoring framework, to prove its applicability in the real world as required by design science methodology was achieved through an online doctor "proof of-concept" prototype that has been implemented using well-known standardized languages and widely adopted tools. Applying the Online doctor prototypes helped us to validate the applicability of the privacy monitoring framework in the cloud computing domain. The quantitative evaluation aspect entailed the usability test and questionnaires to get results. The results obtained from the usability test, questionnaires are explained and analyzed in chapter 5.

From questionnaires the statistical data was found and analyzed. This data was then used to evaluate the overall satisfaction of the users for the solution presented. The results reveal that consumers' data privacy transparency is highly important and will be a motivating factor to cloud adoption by enterprises. The findings also indicate that our solution meets the expectations of ordinary cloud service consumers as reported in Chapter 5.

Analysis of the results obtained from the implementation revealed that our privacy monitoring framework provides for better user comprehension, user awareness, scalability, and ease of use as well as user control due to the use of informative event and access log analyser. In essence, the evaluation concluded that the privacy monitoring framework as proposed in this research work provides the required transparency that enable users comprehend how their personal data is handled.

6.3 Evaluation

The development of the framework conceptually belongs to design science. Therefore, it is absolutely necessary to evaluate the research result against the requirements for an effective design science research project. This research work used the list of criteria for evaluation introduced by Hevner, March, Park, & Ram, (2004) as criteria for the evaluation:

Design as an Artifact: This research work delivers the following viable artefacts:

The privacy monitoring framework including:

– The Privacy Evidence Creation Process UML model

– The informative events and access log analyser component

Problem Relevance: In Section 1.3 above, the problem definition of this research work was identified as the lack of capacity to provide cloud consumers with the means to control and comprehend what happens to the data while it is stored and processed in the cloud.

The development of the privacy monitoring framework, which comprises the Privacy Evidence Creation Process UML model and the informative events and access log analyser component, is totally in line with the problem definition.

Design Evaluation: The privacy monitoring framework has been evaluated using an experimental method. More specifically for the performance evaluation of the solution, both qualitative and quantitative evaluations were used. The evaluation of the solution focuses on the metrics that evaluate the satisfaction of the users' goals. The quantitative evaluation aspect entailed the usability test and questionnaires to get results. From questionnaires the statistical data was found and analysed. The evaluation of the results was based on user-control participants where consumers are asked to rate the ease of performing of the different user centric tasks. The privacy impact assessment guide was used to validate the technical feasibility of the proposed solution, mainly because PIA is the well-established standard for recognizing the potential privacy risks of new or overhauled federal government projects or administrations. Under the quantitative evaluation aspect the scalability experiment were carried out, in order to determine the effect of the increasing load, while observing closely the response time and throughput.

6.4 Limitations and Future work

Although the framework has been proven to enable scalability, user comprehension, and user control as well as user awareness it has some limitations which need to be addressed for future

enhancements. Our framework advocate for user preference to be taken into consideration, as means to ensure user control. However, giving user's choice has limitations. The choice is often an annoyance or even a disservice to individuals. The user's choice can render some system inefficient, since some system requires a rapid, constant update of personal information asking for users' consent each time becomes impractical and could annoy most individuals, for instance the United State credit reporting system which is updated four times a day on average. Credit reporting systems are good for nothing if individuals selectively dictate what to include and exclude (Cate, 2006).

Most people are often ill-equipped to make the right privacy choice regarding the collection and use of their personal information, let alone understanding the risks it could bring or the benefits that may be lost if information is not available. Usually, they will go out of their way to avoid making these choices. The choice is no 'one size fit all' solution in some instances; it contradicts with the things such as national security, freedom of communication, law enforcement and many more. There is a need for future work to investigate on this matter and find the solutions that are less reliant on user's choice.

The overall result was based on the simple scenario which can be seen as the approximation of the reality. The participants in this study were generally students, with a range of technology skills (amateur, intermediate and expert users). These different levels of technology skills raised different sets of needs and expectations regarding the user control and user satisfaction. Students are not qualified medical practitioners who have relevant experience dealing with electronic health record systems. Therefore, for future work, we would like to observe the behavior of the solution in real world deployment. Furthermore in the future work we would further investigate how we can improve scalability of our solution since we experience some polynomial time

complexity. This is due to the fact that since this solution tradeoff response time for optimal data security as a result, most of the system time is spent on encrypting and decrypting data for each transaction. Since the scope of this research was privacy, the study concentrated only on the privacy aspects of the system, however, a comprehensive Common Criteria evaluation could have demonstrated that the set of system components evaluated in this work together in a secure way within the assumed environment. It is vital to look at all the elements of security that need to be in place to ensure that the whole infrastructure is secure (Powell, 2003). Lastly, even though all data is important, organizations often have limited security resources. This necessitates better data identification and classification so that enterprises can concentrate their efforts on securing the data that matters most. Future work lies in formulating better data identification and classification mechanisms which can inform the preference setting component presented in this work and possibly different levels of security to be applied on the different data categories.

Bibliography

Abadi, M. (2004). Trusted Computing, Trusted Third Parties, and Verified Communications. In *In SEC2004: 19th IFIP International Information Security Conference*.

Adil Alsaid , David Martin , Saudi Arabian, M. A. (2003). *Privacy Enhancing Technologies*. (R. Dingledine & P. Syverson, Eds.) (Vol. 2482). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/3-540-36467-6

Agrawal, R., Kiernan, J., Srikant, R., & Xu, Y. (2004). Order preserving encryption for numeric data. In *Proceedings of the 2004 ACM SIGMOD international conference on Management of data - SIGMOD '04* (p. 563). New York, New York, USA: ACM Press. doi:10.1145/1007568.1007632

Allison, D. S., & Capretz, M. A. M. (2011). Furthering the growth of cloud computing by providing privacy as a service. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6868 LNCS, pp. 64–78). doi:10.1007/978-3-642-23447-7_7

Armbrust, M., Stoica, I., Zaharia, M., Fox, A., Griffith, R., Joseph, A. D., … Rabkin, A. (2010). A view of cloud computing. *Communications of the ACM*. doi:10.1145/1721654.1721672

Baca, S. (2010). Cloud Computing: What it is and what it can do for you. Retrieved November 02, 2012, from http://www.globalknowledge.be/content/files/documents/386696/386787

Beat, B. (2003). Apache JMeter. *Group*.

Bélanger, F., & Crossler, R. E. (2011). Privacy in the digital age: a review of information privacy research in information systems. *MIS Quarterly*, *35*(4), 1017–1042. Retrieved from http://dl.acm.org/citation.cfm?id=2208940.2208951

Bellare, M., Boldyreva, A., & O'Neill, A. (2007). Deterministic and Efficiently Searchable Encryption. In *Proceedings of the International Cryptology Conference on Advances in Cryptology (CRYPTO)* (pp. 535–552). doi:10.1007/978-3-540-74143-5_30

Bertino, E., Martino, L. D., Paci, F., & Squicciarini, A. C. (2010). *Security for web services and service-oriented architectures*. *Security for Web Services and Service-Oriented Architectures* (pp. 1–226). doi:10.1007/978-3-540-87742-4

Boldyreva, A., Chenette, N., Lee, Y., & O'Neill, A. (2009). Order-preserving symmetric encryption. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 5479 LNCS, pp. 224–241). doi:10.1007/978-3-642-01001-9_13

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability Evaluation in Industry*, *189*, 194. doi:10.1002/hbm.20701

Cachin, C., Shraer, A., & Shelat, A. (2007). Efficient fork-linearizable access to untrusted shared memory. *PODC 07 Proceedings of the Twentysixth Annual ACM Symposium on Principles of Distributed Computing*, 129–138. doi:10.1145/1281100.1281121

Canadian Institutes of Health Research. (2003). Secondary Use of Personal Information in Health Research: Case Studies, November 2002 - CIHR. Retrieved from http://www.cihr-irsc.gc.ca/e/1475.html

Chow, R., Golle, P., Jakobsson, M., Shi, E., Staddon, J., Masuoka, R., & Molina, J. (2009). Controlling Data in the Cloud : Outsourcing Computation without Outsourcing Control. *Security*, 85–90. doi:10.1145/1655008.1655020

Coen-Porisini, A., Colombo, P., & Sicari, S. (2010). *Software Engineering for Secure Systems*. (H. Mouratidis, Ed.). IGI Global. doi:10.4018/978-1-61520-837-1

Council, C. S. C. (2011). Cloud Computing Use Cases Version 1.0. Retrieved June 31, 2012, from http://www.cloudstandardscustomercouncil.org/use-cases/CloudComputingUseCases.pdf

Crago, S., Dunn, K., Eads, P., Hochstein, L., Kang, D. I., Kang, M., … Walters, J. P. (2011). Heterogeneous cloud computing. In *Proceedings - IEEE International Conference on Cluster Computing, ICCC* (pp. 378–385). doi:10.1109/CLUSTER.2011.49

Daniella Kafouris. (2011). Deloitte talks about maintaining privacy and security in the cloud | Deloitte SA Blog. Retrieved February 25, 2012, from http://deloitteblog.co.za.www102.cpt1.host-h.net/2011/10/11/deloitte-talks-about-maintaining-privacy-and-security-in-the-Cloud/

Davies, N., & Langheinrich, M. (2013). Privacy By Design [From the Editor in Chief]. *IEEE Pervasive Computing*, *12*(2), 2–4. doi:10.1109/MPRV.2013.34

De Capitani Di Vimercati, S., Foresti, S., & Samarati, P. (2012). Managing and accessing data in the cloud: Privacy risks and approaches. In *7th International Conference on Risks and Security of Internet and Systems, CRiSIS 2012*. doi:10.1109/CRISIS.2012.6378956

Dingledine, R., Freedman, M. J., & Molnar, D. (2001). The Free Haven Project: Distributed Anonymous Storage Service. In *Designing Privacy Enhancing Technologies* (pp. 67–95). doi:doi: 10.1007/3-540-44702-4_5

Doelitzscher, F., Reich, C., & Sulistio, A. (2010a). Designing cloud services adhering to government privacy laws. In *Proceedings - 10th IEEE International Conference on Computer and Information Technology, CIT-2010, 7th IEEE International Conference on Embedded Software and Systems, ICESS-2010, ScalCom-2010* (pp. 930–935). doi:10.1109/CIT.2010.172

Doelitzscher, F., Reich, C., & Sulistio, A. (2010b). Designing cloud services adhering to government privacy laws. In *Proceedings - 10th IEEE International Conference on Computer and Information Technology, CIT-2010, 7th IEEE International Conference on Embedded Software and Systems, ICESS-2010, ScalCom-2010* (pp. 930–935). doi:10.1109/CIT.2010.172

Easter, R. J., & French, C. (2014). Annex A: Approved Security Functions for FIPS PUB 140-2, Security Requirements for Cryptographic Modules. Retrieved October 22, 2014, from http://csrc.nist.gov/publications/fips/fips140-2/fips1402annexa.pdf

Espadas, J., Molina, A., Jiménez, G., Molina, M., Ramírez, R., & Concha, D. (2013). A tenant-based resource allocation model for scaling Software-as-a-Service applications over cloud computing infrastructures. *Future Generation Computer Systems*, *29*(1), 273–286. doi:10.1016/j.future.2011.10.013

Feldman, A. J., Zeller, W. P., Freedman, M. J., & Felten, E. W. (2010). SPORC : Group Collaboration using Untrusted Cloud Resources. *System*, *34*, 1–14. Retrieved from

http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:SPORC+:+Group+Collaboration+using+Untrusted+Cloud+Resources#0

Gellert, R., & Gutwirth, S. (2013). The legal construction of privacy and data protection. *Computer Law & Security Review*, *29*(5), 522–530. doi:10.1016/j.clsr.2013.07.005

Glott, R., Husmann, E., Sadeghi, A. R., & Schunter, M. (2011). Trustworthy clouds underpinning the future internet. *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6656*, 209–221. doi:10.1007/978-3-642-20898-0_15

Goodin, D. (2013). How an epic blunder by Adobe could strengthen hand of password crackers | Ars Technica. Retrieved December 12, 2013, from http://arstechnica.com/security/2013/11/how-an-epic-blunder-by-adobe-could-strengthen-hand-of-password-crackers/

Government of Canada, T. B. of C. S. (2010, March 29). Directive on Privacy Impact Assessment. Retrieved from http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?section=text&id=18308

Grimes, R. (2010). Finding gold in your log files | InfoWorld. Retrieved February 22, 2013, from http://www.infoworld.com/article/2627500/data-mining/finding-gold-in-your-log-files.html

Grobauer, B., Walloschek, T., & Stöcker, E. (2011). Understanding cloud computing vulnerabilities. *IEEE Security and Privacy*, *9*, 50–57. doi:10.1109/MSP.2010.115

Guan, D. J., Tsai, C.-Y., & Zhuang, E. S. (2013). Detect Zero by Using Symmetric Homomorphic Encryption. In *2013 Eighth Asia Joint Conference on Information Security* (pp. 1–7). IEEE. doi:10.1109/ASIAJCIS.2013.8

H Regard. (2013). OECD guidelines governing the protection of privacy and transborder flows of personal data. Retrieved March 20, 2014, from http://www.oecd.org/sti/ieconomy/2013-oecd-privacy-guidelines.pdf

Halderman, J. A., Schoen, S. D., Heninger, N., Clarkson, W., Paul, W., Calandrino, J. a., … Felten, E. W. (2008). Lest We Remember: Cold Boot Attacks on Encryption Keys. *USENIX Security Symposium*, 1–16. doi:10.1145/1506409.1506429

Henze, M., Grossfengels, M., Koprowski, M., & Wehrle, K. (2013). Towards data handling requirements-aware cloud computing. In *Proceedings of the International Conference on Cloud Computing Technology and Science, CloudCom* (Vol. 2, pp. 266–269). doi:10.1109/CloudCom.2013.145

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004a). Design Science in Information Systems Research. *MIS Quarterly*, *28*, 75–105. doi:10.2307/25148625

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004b). Design Science in Information Systems Research. *MIS Quarterly*, *28*, 75–105. doi:10.2307/25148625

Huang, X., & Du, X. (2013a). Efficiently secure data privacy on hybrid cloud. In *IEEE International Conference on Communications* (pp. 1936–1940). doi:10.1109/ICC.2013.6654806

Huang, X., & Du, X. (2013b). Efficiently secure data privacy on hybrid cloud. In *IEEE International Conference on Communications* (pp. 1936–1940). doi:10.1109/ICC.2013.6654806

IBM. (2010). Strategies for assessing cloud security. Retrieved August 14, 2012, from http://public.dhe.ibm.com/common/ssi/ecm/en/sew03022usen/SEW03022USEN.PDF

Infrastructure, C. for the P. of N. (2010). INFORMATION SECURITY BRIEFING 01/2010 CLOUD COMPUTING. Retrieved October 21, 2012, from http://www.cpni.gov.uk/Documents/Publications/2010/2010007-ISB_cloud_computing.pdf

ISO/IEC JTC 1 and the ISO and IEC. (2013). Publicly Available Standards. Retrieved September 04, 2014, from http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html

Itani, W., Kayssi, A., & Chehab, A. (2009). Privacy as a service: Privacy-aware data storage and processing in cloud computing architectures. In *8th IEEE International Symposium on Dependable, Autonomic and Secure Computing, DASC 2009* (pp. 711–716). doi:10.1109/DASC.2009.139

Jain, Payal, and J. J. (2012). CRACKS IN THE CLOUD. doi:ISSN: 2249-3905

Jansen, W., & Grance, T. (2011). Guidelines on Security and Privacy in Public Cloud Computing. *Director*, *144*, 800–144. doi:10.3233/GOV-2011-0271

Julio Angulo, Simone Fischer-Hübner, J. S. P. (2013). D:C-7.1 General HCI principles and guidelines. Retrieved October 24, 2013, from http://www.a4cloud.eu/sites/default/files/D37.1 General HCI principles and guidelines.pdf

Kleyman, B. (2012). Weighing the pros and cons of the Trusted Computing Platform. Retrieved September 15, 2013, from http://searchdatacenter.techtarget.com/tip/Weighing-the-pros-and-cons-of-the-Trusted-Computing-Platform

Kun, Z., Abraham, A., & Yuliang, S. (2013). Data combination privacy preservation adjusting mechanism for software as a service. In *Proceedings - 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013* (pp. 2007–2012). doi:10.1109/SMC.2013.344

Langheinrich, M. (2001). Privacy by design—principles of privacy-aware ubiquitous systems. *Ubicomp 2001: Ubiquitous Computing*, 273–291. doi:10.1007/3-540-45427-6_23

Langheinrich, M. (2002). A privacy awareness system for ubiquitous computing environments. *UbiComp 2002: Ubiquitous Computing*, 237–245. doi:10.1007/3-540-45809-3_19

Leinwand, Tim Yim and Allan, S. (2014). Multi-tenant or multi-instance cloud – why not both? — Tech News and Analysis. Retrieved April 27, 2014, from https://gigaom.com/2014/01/26/multi-tenant-or-multi-instance-cloud-lets-do-both/

Lemoudden, M., Ben Bouazza, N., El Ouahidi, B., & Bourget, D. (2013). A survey of cloud computing security overview of attack vectors and defense mechanisms. *Journal of Theoretical and Applied Information Technology*, *54*, 325–330.

Lewis, P. (2014). Edward Snowden | US news | The Guardian. Retrieved March 11, 2014, from http://www.theguardian.com/us-news/edward-snowden

Lindell, Y., & Pinkas, B. (2014). An Efficient Protocol for Secure Two-Party Computation in the Presence of Malicious Adversaries. *Journal of Cryptology*. doi:10.1007/s00145-014-9177-x

MacAskill, G. G. and E. (2013). NSA Prism program taps in to user data of Apple, Google and others | US news | The Guardian. Retrieved October 22, 2014, from http://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data

Manan, J.-L. A., Mubarak, M. F., Isa, M. A. M., & Khattak, Z. A. (2011). Security, trust and privacy: a new direction for pervasive computing, 56–60. Retrieved from http://dl.acm.org/citation.cfm?id=2028299.2028313

Mark S. Ackerman and Donald T. Davis, J. (2003). Privacy and Security Issues in E-Commerce. Retrieved October 22, 2012, from http://web.eecs.umich.edu/~ackerm/pub/03e05/EC-privacy.ackerman.pdf

Martens, B., Teuteberg, F., & Gräuler, M. (2011). Design and Implementation of a Community Platform for the Evaluation and Selection of Cloud Computing Services: A Market Analysis. *Proceedings of the 19th European Conference on Information Systems*, *Paper 215*, 1–7. Retrieved from http://aisel.aisnet.org/ecis2011/215

Massonet, P., Naqvi, S., Ponsard, C., Latanicki, J., Rochwerger, B., & Villari, M. (2011). A monitoring and audit logging architecture for data location compliance in federated cloud infrastructures. In *IEEE International Symposium on Parallel and Distributed Processing Workshops and Phd Forum* (pp. 1510–1517). doi:10.1109/IPDPS.2011.304

Mather, T., Kumaraswamy, S., & Latif, S. (2009). *Cloud Security and Privacy*. (M. Loukides, Ed.)*Governance An International Journal Of Policy And Administration* (p. 336). O'Reilly. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Cloud+Security+and+Privacy#0

Mattsson, U. T. (2005). A practical implementation of transparent encryption and separation of duties in enterprise databases: Protection against external and internal attacks on databases. In *Proceedings - Seventh IEEE International Conference on E-Commerce Technology, CEC 2005* (Vol. 2005, pp. 559–565). doi:10.1109/ICECT.2005.9

Mell, P., & Grance, T. (2009). The NIST Definition of Cloud Computing. *National Institute of Standards and Technology*, *53*, 50. doi:10.1080/1047621920040218

Michelle Finneran Dennedy, Thomas R. Finneran, J. F. (2014). Privacy Engineering and Data Governance. Retrieved May 12, 2014, from http://www.tdan.com/view-articles/17236

MIMOS National R&D Centre in ICT. (2012). Mi-Trust - MIMOS. Retrieved October 27, 2012, from http://www.mimos.my/technology-innovations/mi-trust/

Mohamed, N., & Ahmad, I. H. (2012). Information privacy concerns, antecedents and privacy measure use in social networking sites: Evidence from Malaysia. *Computers in Human Behavior*, *28*(6), 2366–2375. doi:10.1016/j.chb.2012.07.008

Morris, T. (2011). *Encyclopedia of Cryptography and Security*. (H. C. A. van Tilborg & S. Jajodia, Eds.). Boston, MA: Springer US. doi:10.1007/978-1-4419-5906-5

Mowbray, M., Pearson, S., & Shen, Y. (2012). Enhancing privacy in cloud computing via policy-based obfuscation. In *Journal of Supercomputing* (Vol. 61, pp. 267–291). doi:10.1007/s11227-010-0425-z

Naor, D., Shenhav, A., & Wool, A. (2005). Toward Securing Untrusted Storage Without Public-Key Operations. Retrieved October 22, 2013, from https://www.eng.tau.ac.il/~yash/sss06-naor.pdf

Neff, T. (2011). Cloud Computing's Not-So-Silver Lining | Compliance Week. Retrieved March 20, 2014, from http://www.complianceweek.com/news/news-bulletin/cloud-computings-not-so-silver-lining

Nemati, H. R., & Van Dyke, T. (2009). Do Privacy Statements Really Work? The Effect of Privacy Statements and Fair Information Practices on Trust and Perceived Risk in E-Commerce. *International Journal of Information Security and Privacy*, *3*(1), 45–64. doi:10.4018/jisp.2009010104

Ngugi, B., & Dardick, G. (2010). Security and Privacy Assurance in Advancing Technologies. In H. Nemati (Ed.), (pp. 2011–2013). IGI Global. doi:10.4018/978-1-60960-200-0

Ni, Q., Trombetta, A., Bertino, E., & Lobo, J. (2007). Privacy-aware role based access control. *Proceedings of the 12th ACM Symposium on Access Control Models and Technologies - SACMAT '07*, 41. doi:10.1145/1266840.1266848

Nkosi, L., Tarwireyi, P., & Adigun, M. O. (2013). Detecting a malicious insider in the cloud environment using sequential rule mining. In *IEEE International Conference on Adaptive Science and Technology, ICAST*. doi:10.1109/ICASTech.2013.6707505

OECD. (2013). OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. Retrieved October 22, 2014, from http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm

Omran, E., Bokma, A., & Al-Maati, S. A. (2008). Chain ontology based: A model for protecting personal information privacy. *2008 Third International Conference on Digital Information Management*. doi:10.1109/ICDIM.2008.4746767

OpenStack. (2013). OpenStack Manuals, Open-Stack Compute Administration Guide. Retrieved November 12, 2013, from https://launchpad.net/openstack-manuals

Pappas, V., Kemerlis, V., Zavou, A., Polychronakis, M., & Keromytis, A. D. (2012). CloudFence: Enabling Users to Audit the Use of their Cloud-Resident Data. Retrieved from http://academiccommons.columbia.edu/catalog/ac:145374

Parliamentary Monitoring Group, S. A. (2006). Consumer Protection Draft Bill, 2006 (Part 1) | Parliamentary Monitoring Group | Parliament of South Africa monitored. Retrieved October 23, 2011, from http://www.pmg.org.za/policy-documents/2006/03/13/consumer-protection-draft-bill-2006-part-1

Pasha, A., & Gafoor, A. (2011). Transparent Data Encryption- Solution for Security of Database Contents. *International Journal of Advanced Computer Science and Applications*, *2*, 25–28.

Pearson, J., Logan, C., Reis, D., Taveras, L., & Koerner, D. (2014). Answers to Healthcare Leaders ' Cloud Questions.

Pearson, S. (2009). Taking account of privacy when designing cloud computing services. *Current*, 44–52. doi:10.1109/CLOUD.2009.5071532

Pearson, S. (2013). Privacy, security and trust in cloud computing. *Privacy and Security for Cloud Computing*, 3–42. doi:10.1007/978-1-4471-4189-1

Pearson, S., & Charlesworth, A. (2009). Accountability as a way forward for privacy protection in the cloud. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 5931 LNCS, pp. 131–144). doi:10.1007/978-3-642-10665-1_12

Pearson, S., Shen, Y., & Mowbray, M. (2009). A privacy manager for cloud computing. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 5931 LNCS, pp. 90–106). doi:10.1007/978-3-642-10665-1_9

Pearson, Siani, Yee, G. (Eds. . (2013). Privacy and Security for Cloud Computing. doi:ISBN 978-1-4471-4189-1

Perlroth, N. (2012). LinkedIn Breach Exposes Light Security Even at Data Companies - NYTimes.com. Retrieved April 28, 2013, from http://www.nytimes.com/2012/06/11/technology/linkedin-breach-exposes-light-security-even-at-data-companies.html?pagewanted=all&_r=1&

Popa, R., & Redfield, C. (2011). Cryptdb: protecting confidentiality with encrypted query processing. *Proceedings of the …*, 85–100. doi:10.1145/2043556.2043566

Powell, D. (2003). Understanding the Common Criteria evaluation. Retrieved August 05, 2014, from http://intosaiitaudit.org/intoit_articles/18p32top35.pdf

Prasad, P., Ojha, B., Shahi, R. R., Lal, R., Vaish, A., & Goel, U. (2011a). 3 dimensional security in cloud computing. *2011 3rd International Conference on Computer Research and Development*, *3*, 198–201. doi:10.1109/ICCRD.2011.5764279

Prasad, P., Ojha, B., Shahi, R. R., Lal, R., Vaish, A., & Goel, U. (2011b). 3 dimensional security in cloud computing. *2011 3rd International Conference on Computer Research and Development*, *3*, 198–201. doi:10.1109/ICCRD.2011.5764279

PricewaterhouseCoopers. (2012). The role of internal audit in assuring data security and privacy: PwC. Retrieved September 17, 2012, from http://www.pwc.com/us/en/risk-assurance-services/publications/internal-audit-assuring-data-security-privacy.jhtml

Privacy Rights Clearinghouse. (2005). A Review of the Fair Information Principles: The Foundation of Privacy Public Policy | Privacy Rights Clearinghouse. Retrieved October 20, 2014, from https://www.privacyrights.org/content/review-fair-information-principles-foundation-privacy-public-policy

Programming4us. (2010). Cloud Security and Privacy : What Is the Data Life Cycle? - Microsoft Certification Examples, exercises, practises, tutorials, solutions about Programming. Retrieved January 22, 2014, from http://mscerts.programming4.us/programming/cloud security and privacy what is the data life cycle.aspx#3Ad9l3cAVKdFeYKw.99

R. Rivest, L. A. (1978). On data banks and privacy homomorphisms. *In Foundation of Secure Computations, Academic Press*.

Reagle, J., & Cranor, L. F. (1999). The platform for privacy preferences. *Communications of the ACM*. doi:10.1145/293411.293455

Reed, A., Rezek, C., Simmonds, P. (2011). Security guidance for critical areas of focus in cloud computing v3.0. Retrieved October 20, 2014, from https://cloudsecurityalliance.org/guidance/csaguide.v3.0.pdf

Restrepo, M. J. (2005). ARTICLE: The Convergence of Commercial and Investment Banking Under the Gramm-Leach-Bliley Act: Revisiting Old Risks and Facing New Problems. Retrieved from https://litigation-essentials.lexisnexis.com/webcd/app?action=DocumentDisplay&crawlid=1&doctype=cite&

docid=11+Law+%26+Bus.+Rev.+Am.+269&srctype=smi&srcid=3B15&key=3ebc228a24f8773f6d46cc8813161f9c

Ronald L. Rivest, A. T. S. (1983). *Advances in Cryptology*. (D. Chaum, R. L. Rivest, & A. T. Sherman, Eds.). Boston, MA: Springer US. doi:10.1007/978-1-4757-0602-4

S. Bennett Mans Bhuller, R. C. (2009). Architectural Strategies for Cloud Computing. Retrieved October 22, 2014, from http://www.oracle.com/us/ciocentral/architecrl-strategies-for-cc-396213.pdf

Sabahi, F. (2011). Cloud computing security threats and responses. *2011 IEEE 3rd International Conference on Communication Software and Networks*, 245–249. doi:10.1109/ICCSN.2011.6014715

Santos, N., Gummadi, K. P., & Rodrigues, R. (2009). Towards Trusted Cloud Computing. *Computing*, *10*, 3. doi:10.1016/j.istr.2005.05.004

Scholtz, J. (2013). Usability Evaluation. Retrieved October 22, 2014, from http://notification.etisalat.com.eg/etisalat/templates.backup.16082011/582/Usability%2520Evaluation_rev1%5B1%5D.pdf

Shaikh, F. B. F., & Haider, S. (2011). Security threats in cloud computing. *2011 International Conference for Internet Technology and Secured Transactions*, 214–219. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6148380

Shen, Y., & Pearson, S. (2011). Privacy Enhancing Technologies : A Review Abstract : Privacy Enhancing Technologies : A Review. *Development*, *2739*, 282–287. Retrieved from http://www.springerlink.com/openurl.asp?genre=article&id=W19UAWJ7V3H9EGQ0

Solove, D. (2014). SafeGov.org - Duties When Contracting With Data Service Providers. Retrieved May 22, 2014, from http://safegov.org/2014/2/18/duties-when-contracting-with-data-service-providers

South African Law Reform Commission. (2005). Privacy and data protection. Retrieved June 03, 2012, from http://www.justice.gov.za/salrc/dpapers/dp109.pdf

Spiegel, R. (2011). Sony Tallies $171M in Data Breach Losses... and Counting. Retrieved November 15, 2012, from http://www.ecommercetimes.com/story/72520.html

Stefanov, E., van Dijk, M., Juels, A., & Oprea, A. (2012). Iris. In *Proceedings of the 28th Annual Computer Security Applications Conference on - ACSAC '12* (p. 229). New York, New York, USA: ACM Press. doi:10.1145/2420950.2420985

Subashini, S., & Kavitha, V. (2010). A survey on security issues in service delivery models of cloud computing. *Journal of Network and Computer Applications*, *34*, 1–11. doi:10.1016/j.jnca.2010.07.006

Swart, I. P., Grobler, M. M., & Irwin, B. (2013). Visualization of a data leak. In *2013 Information Security for South Africa* (pp. 1–8). IEEE. doi:10.1109/ISSA.2013.6641046

Takabi, H., Joshi, J. B. D., & Ahn, G.-J. (2010). Security and Privacy Challenges in Cloud Computing Environments. *IEEE Security & Privacy Magazine*. doi:10.1109/MSP.2010.186

Trusted Computing Group. (2010). Cloud Computing and Security – A Natural Match. Retrieved October 27, 2012, from http://www.trustedcomputinggroup.org/files/resource_files/3C58110D-1A4B-B294-D060C83C1723209D/Cloud Computing and Security Whitepaper_April.23.2010.pdf

US Health & Human Services Dept. (2003). HIPAA Privacy Rule and Public Health Guidance from CDC and the U.S. Department of Health and Human Services. Retrieved October 20, 2014, from http://www.cdc.gov/mmwr/pdf/other/m2e411.pdf

VMware Inc. (2009). Eight Key Ingredients for Building an Internal Cloud. Retrieved March 05, 2011, from http://www.vmware.com/files/pdf/cloud/eight-key-ingredients-building-internal-cloud.pdf

Von Laszewski, G., Diaz, J., Wang, F., & Fox, G. C. (2012). Comparison of multiple cloud frameworks. In *Proceedings - 2012 IEEE 5th International Conference on Cloud Computing, CLOUD 2012* (pp. 734–741). doi:10.1109/CLOUD.2012.104

Wang, C., Chow, S. S. M., Wang, Q., Ren, K., & Lou, W. (2013). Privacy-preserving public auditing for secure cloud storage. *IEEE Transactions on Computers*, *62*, 362–375. doi:10.1109/TC.2011.245

Warren, S. D., & Brandeis, L. D. (1890). The Right to Privacy. *Harvard Law Review*, *4*, 193–220. doi:10.2307/1321160

Warso, Z. (2013). There's more to it than data protection – Fundamental rights, privacy and the personal/household exemption in the digital age. *Computer Law & Security Review*, *29*(5), 491–500. doi:10.1016/j.clsr.2013.07.002

Weitzeekorn, B. (2012). Adobe Data Breach Exposes Military Passwords - Technology & science - Tech and gadgets - TechNewsDaily - msnbc.com. Retrieved January 03, 2012, from http://www.nbcnews.com/id/49826676/ns/technology_and_science-tech_and_gadgets/t/adobe-data-breach-exposes-military-passwords/#.VEWPZCKUeSo

Wen, X., Gu, G., Li, Q., Gao, Y., & Zhang, X. (2012). Comparison of open-source cloud management platforms: OpenStack and OpenNebula. In *Proceedings - 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2012* (pp. 2457–2461). doi:10.1109/FSKD.2012.6234218

Woolf, B. (2013, April 1). Best practices using infrastructure as a service in IBM PureApplication System. Retrieved from http://www.ibm.com/developerworks/cloud/library/cl-aim1301-bestpractices-iaas-pureapp/

Xiang, G., Yu, B., & Zhu, P. (2012). A algorithm of fully homomorphic encryption. In *2012 9th International Conference on Fuzzy Systems and Knowledge Discovery* (pp. 2030–2033). IEEE. doi:10.1109/FSKD.2012.6234023

Xiao, L., & Yen, I. L. (2012). Security analysis for order preserving encryption schemes. In *2012 46th Annual Conference on Information Sciences and Systems, CISS 2012*. doi:10.1109/CISS.2012.6310814

Yao, A. C. (1982). Protocols for secure computations. *23rd Annual Symposium on Foundations of Computer Science (sfcs 1982)*. doi:10.1109/SFCS.1982.38

Zajac, B. (1994). Pretty good privacy. *Computer Fraud & Security Bulletin*. doi:10.1016/0142-0496(94)90185-6

Zhang, X., Liu, F., Chen, T., & Li, H. (2009). Research and application of the transparent data encpryption in intranet data leakage prevention. In *CIS 2009 - 2009 International Conference on Computational Intelligence and Security* (Vol. 2, pp. 376–379). doi:10.1109/CIS.2009.107

Zhou, M., Zhang, R., Xie, W., Qian, W., & Zhou, A. (2010). Security and Privacy in Cloud Computing: A Survey. *2010 Sixth International Conference on Semantics, Knowledge and Grids*, 105–112. doi:10.1109/SKG.2010.19

Zimmermann, P. (1991). PGP, Pretty Good Privacy. Retrieved July 17, 2012, from http://www.livinginternet.com/i/is_crypt_pgp.htm

Kendrick, Herbert S, and John J Kendrick Jr. (2003), State Financial Institutions. Texas Transaction Guide--Legal Forms, 2013.

Appendices

Appendix A: Statistical Result, Frequency and Percentage for User Control

Table 5-10 Statistical Result, Frequency and Percentage for User Control

Function	**Statistical Result**				**Frequency and Percentage of People Understanding**			
Create a Profile	**N Valid**	**Missing**	**Mean**	**Std Dev**	**Value Label**	**Value**	**Frequency**	**Valid Percentage**
	60	0	4.45	0.87	Very Difficult	1	1	1.67
					Somewhat Difficult	2	3	5
					Understandable	4	20	33.33
					Very Easy to Understand	5	36	60
					Total		60	100
Book a Doctor	**N Valid**	**Missing**	**Mean**	**Std Dev**	**Value Label**	**Value**	**Frequency**	**Valid Percentage**
	60	0	4.55	0.70	Very Difficult	1	1	1.67
					Undecided	3	1	1.67
					Understandable	4	21	35
					Very Easy to	5	37	61.67

					Understand			
					Total		60	100
View Medical Record	**N Valid**	**Missing**	**Mean**	**Std Dev**	**Value Label**	**Value**	**Frequency**	**Valid Percentage**
	60	0	4.42	0.83	Very Difficult	1	1	1.67
					Somewhat Difficult	2	1	1.67
					Undecided	3	4	6.67
					Understandable	4	20	33.33
					Very Easy to Understand	5	34	56.67
					Total		60	100
Grant Access Rights	**N Valid**	**Missing**	**Mean**	**Std Dev**	**Value Label**	**Value**	**Frequency**	**Valid Percentage**
	60	0	4.37	0.99	Very Difficult	2	6	10
					Undecided	3	4	6.67
					Understandable	4	12	20
					Very Easy to Understand	5	38	63.33
					Total		60	100

Appendix B: Consent Form

INFORMATION AND INFORMED CONSENT FORM

RESEARCHER'S DETAILS	
Title of the research project	**Development of a Privacy Monitoring Framework for the Cloud Computing Environment**
Principal investigator	**Manqoba Victor Shabalala**
Contact telephone number (research office)	**035 902 6012**

A. DECLARATION BY OR ON BEHALF OF THE PARTICIPANT		Initials
I, the participant and the undersigned	(full names)	

A.1. HEREBY CONFIRM AS FOLLOW:		Initials
I, the participant was invited to participate in the above-mentioned research project		
that is being undertaken by	Manqoba Victor Shabalala	
from	Department of Computer Science	
of the University of Zululand		

A.2 THE FOLLOWING ASPECTS HAVE BEEN EXPLAINED TO ME (THE PARTICIPANT)		Initials
Aim	The shared nature of the cloud storage infrastructure and the fact that when the data is stored in the cloud, the control of the data is more on the hands of the cloud provider rather than the data owner is of a great challenge that continues to hinder cloud computing from successfully reaching its successful capability. The investigator is studying how to enables the data owner to stay in control over their data, thereby providing the required transparency to comprehend how personal data is handled in the cloud. The information will be used for research purposes.	
Risks	I understand that there are no risks involved in participating in this process.	
Confidentiality	I am fully aware that my identity will known and only be visible to me and the researcher. It will not be revealed in any discussion, description or scientific publications.	
Access to findings	I am also aware that any new information that develops during this process would be shared as follows: In a *dissertation, journal or conference article*.	

Voluntary participation / refusal / discontinuation	My participation is voluntary	Y	N	
	My decision whether or not to participate will in no way affect my present or future career/employment/lifestyle	T	F	
No pressure was exerted on me to consent to participate and I understand that I have the right to withdraw at any stage without penalization.				
Participation in this study will not result in any additional cost to me.				

I HEREBY VOLUNTARILY CONSENT TO PARTICIPATE IN THE ABOVE-MENTIONED RESEARCH PROJECT

Signed at: (place) on the (date) of (month) 20 (year)

Signature	Signature of witness:	
	Full name(s) of witness:	

Appendix C: Usability Questionnaire

QUESTIONNAIRE ON ONLINE DOCTOR (e-MEDICAL HEALTH SYSTEM) BASED ON A PRIVACY MONITORING FRAMEWORK

Kindly answer the following questions. This information is required for the research purpose only.

The information will be treated as confidential.

Full Names: _____ (Optional)

Student Number_____ (Optional)

Do you consider yourself?

- □ **Amateur**
- □ **Intermediate**
- □ **Expect**

Questionnaire

1. How satisfied are you with the system?
    a. Unsatisfied
    b. Somewhat satisfied
    c. Undecided
    d. Satisfied
    e. Very satisfied
2. This site was simple to navigate?
    a. Strongly disagree
    b. Disagree
    c. Undecided
    d. Agree
    e. Strongly agree
3. Rate the ease of performing the task (Create Profile, Booking, view medical record and grant access rights)
    a. Very difficult
    b. Somewhat difficult
    c. Neither easy nor difficult

    d. Easy
    e. Very easy
4. Do you understand the purpose of various interface elements, terminology, and procedures?
    a. Very confusing
    b. Somewhat Confusing
    c. Undecided
    d. Understandable
    e. Very easy to understand
5. Overall, this site worked very well technically?
    a. Strongly disagree
    b. Disagree
    c. Undecided
    d. Agree
    e. Strongly agree
6. Did you find your privacy breach notification to be informative enough?
    a. Not informative
    b. Somewhat Confusing
    c. Undecided
    d. Somewhat informative
    e. Very informative