UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING

Electronics and Computer Science

A Security Framework to Protect Data in Cloud Storage

by

Farashazillah Yahya

Thesis for the degree of Doctor of Philosophy

November 2017

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING

Electronics and Computer Science

Doctor of Philosophy

A SECURITY FRAMEWORK TO PROTECT DATA IN CLOUD STORAGE

by Farashazillah Yahya

According to Cisco Global Cloud Index, cloud storage users will store 1.6 Gigabytes data per month by 2019, compared to 992 megabytes data per month in 2014. With this trend, it has been shown that more and more data will reside in cloud storage and it is expected to grow further. As cloud storage is becoming an option for users for keeping their data online, it comes with security concerns for protecting data from threats. This thesis addresses the need to investigate the security factors that will enable efficient security protection for data in cloud storage and the relationships that exist between the different security factors. Consequently, this research has developed a conceptual framework that supports security in cloud storage.

The main contribution of this research is the development of a Cloud Storage Security Framework (CSSF) to support an integrative approach to understanding and evaluating security in cloud storage. The framework enables understanding of the makeup of security in cloud storage and measures the understanding of security in cloud storage. Drawing upon established theories and prior research findings, the framework indicates that security in cloud storage can be determined by nine factors: (1) security policies implementation in cloud storage, security measure that relates to (2) protecting the data accessed in cloud storage; (3) modifications of data stored; (4) accessibility of data stored in cloud storage; (5) non-repudiation to the data stored; (6) authenticity of the original data; (7) reliability of the cloud storage services; (8) accountability of service provision; and (9) auditability of the data accessed and stored in cloud storage.

An example of CSSF application has been demonstrated through the development of a measuring instrument called Security Rating Score (SecRaS) and through a series of experiments, SecRaS has been validated and used in a research scenario. The instrument consists of several items generated using goal-question-metric approach. These potential items were evaluated by a series of experiments; the security experts assessed using content validity ratio while the security practitioners took part in the validation study. The validation study completed two experiments that look into the correlation analyses and internal reliability.

SecRaS instrument was later applied in a research scenario; the validated instrument was distributed and a number of 218 usable responses were received. Using structural equation modelling, the data has revealed a good fit of the measurement analyses and structural model. The key findings were as follow: the relationships between factors were found to have both direct and indirect effects in the result. While establishing the relationship(s) among the factors, the structural model proposes three types of causal relationships in terms of how the security implementation in cloud storage could be affected by the security factors.

This thesis presents a detailed discussion of the CSSF development, confirmation, and application in a research scenario. For security managers, CSSF offers a new paradigm on how stakeholders can make cloud storage security implementation successful in some depth. For security practitioners, the CSSF enables deconstruction of the concept of security in cloud storage into smaller, conceptually distinct and manageable factors to guide the design of security in cloud storage. For researchers, the CSSF provides a common framework in which to conceptualise their research and make it easier to see how the security factors fit into the larger picture.

Table of Contents

List of Tables

List of Figures

List of Equations

Declaration of Authorship

I, Farashazillah Yahya declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

A SECURITY FRAMEWORK TO PROTECT DATA IN CLOUD STORAGE

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:
    - Yahya, F, Walters, RJ and Wills, GB. (2017) Clustering Security Factors for Protecting Data in Cloud Storage using Exploratory Factor Analysis (EFA): An Empirical Study, In 7th International Conference on Cloud Computing and Services Science (CLOSER), Portugal, 24-26 April 2017
    - Yahya, F, Walters, RJ and Wills, GB. (2017) Using Goal-Question-Metric (GQM) Approach to Assess Security in Cloud Storage, In Enterprise Security, Lecture Notes in Computer Science 10131, Chapter 10, Springer Book 2017
    - Yahya, F, Walters, RJ and Wills, GB. (2016) Goal-Driven Components for Cloud Storage Security Framework. In International Conference on

Cyber Security and Protection of Digital Services (Cyber Security), London, 13-14 June 2016

- Yahya, F, Walters, RJ and Wills, GB. (2015) Modelling Threats with Security Requirements in Cloud Storage. International Journal for Information Security Research (IJISR), Volume 5, Issue 2, ISSN: 2042-4639.
- Yahya, F, Walters, RJ and Wills, GB. (2015) Analysing Threats in Cloud Storage. In World Congress on Internet Security (WorldCIS-2015) Conference, Dublin, 19-21 Oct 2015.
- Yahya, F, Chang, V, Walters, RJ and Wills, GB. (2014) Security Challenges in Cloud Storage. In 6th IEEE International Conference on Cloud Computing Technology and Science (Enterprise Security 2014), Singapore, 15-18 Dec 2014.

Signed:..

Date: ....

Published Work

Lists of peer reviewed publications in support of this research.

1. Yahya, F, Walters, RJ and Wills, GB. (2017) Clustering Security Factors for Protecting Data in Cloud Storage using Exploratory Factor Analysis (EFA): An Empirical Study, In 7th International Conference on Cloud Computing and Services Science (CLOSER), Portugal, 24-26 April 2017
2. Yahya, F, Walters, RJ and Wills, GB. (2017) Using Goal-Question-Metric (GQM) Approach to Assess Security in Cloud Storage, In Enterprise Security, Lecture Notes in Computer Science 10131, Chapter 10, Springer Book 2017
3. Yahya, F, Walters, RJ and Wills, GB. (2016) Goal-Driven Components for Cloud Storage Security Framework. In International Conference on Cyber Security and Protection of Digital Services (Cyber Security), London, 13-14 June 2016
4. Yahya, F, Walters, RJ and Wills, GB. (2015) Modelling Threats with Security Requirements in Cloud Storage. International Journal for Information Security Research (IJISR), Volume 5, Issue 2, ISSN: 2042-4639.
5. Yahya, F, Walters, RJ and Wills, GB. (2015) Analysing Threats in Cloud Storage. In World Congress on Internet Security (WorldCIS-2015) Conference, Dublin, 19-21 Oct 2015.
6. Yahya, F, Chang, V, Walters, RJ and Wills, GB. (2014) Security Challenges in Cloud Storage. In 6th IEEE International Conference on Cloud Computing Technology and Science (Enterprise Security 2014), Singapore, 15-18 Dec 2014.

Acknowledgement

When I was in high school, I was ambitionless. I had no dreams at least in the context of a professional career. My father, Yahya decided (on my behalf) that I should pursue a study in computer science. He enrolled me into a computer science diploma programme. Which later I discovered I was quite good at. For that, I thank my father for everything; for initiating my passion and making this endeavour possible.

My late mother, Asmah was my emotional pillar. Unfortunately, she passed on too soon, unable to watch me graduate my bachelor's degree with flying colours and received a full scholarship during my studies. Later, I secured a job at Federal Government of Malaysia, served seven years managing a data centre and was offered another scholarship to pursue PhD studies. This is briefly what brought me to the University of Southampton.

I would like to take this opportunity to thank my funder, the Federal Government of Malaysia for granting a full scholarship and compensating for my family expenses here.

Three years of research has not always been smooth. My appreciation goes to my supervisors, Dr Robert J Walters and Associate Professor Dr Gary B Wills for their never-ending support, for having an open-door policy and offering sage advice. During tough times, they have given me space and needed comfort. This thesis may have never been written in the first place without their impetus and impetuous. Thank you very much.

I would also like to thank security experts and practitioners in the United Kingdom and Malaysia who took part in my research for their respected ideas, inputs and time spent throughout the study.

I would also like to thank several people who have assisted me; Associate Professor Lester Gilbert for providing insights and grasp on statistics, Dr Hasnida Ghazali for explanation on factor analysis and structural equation modelling, Marinah Muhammad for recommending statistics resources, Aida A. Rahman for knowledge sharing on thesis writing, and Quintin Gee for proofreading my thesis. With the oversight of my supervisors, editorial advice has been sought. No changes of intellectual content were made as a result of this advice.

Throughout my period here, I have also been involved with the OPERANDO European project, I would like to thank the people at IT Innovation Centre for the first-hand experience being part of a distinguished research agency. My accomplished work there has inadvertently helped improve the quality of my own research's state of the art.

To my fellow colleagues at Electronic and Software Systems group, thank you for supporting me in many ways. Thank you to Dr Vanissa Wanick, Khairul Hanim Pazim and Watthanasak Jeamwatthanachai for preparing me for the oral examination (Viva Voce). My former colleagues at Ministry of Higher Education for lending innumerable support whenever needed. Most of the days I still feel that I am part of the data centre team. Not forgetting, my best friends, Arlene Andrew, Sucy Leong, Adzlina Ab Rahman and Dr Saidah Sahid. You might not realise your friendship has helped me survive the United Kingdom in an unthinkable way. My support group, the QA HLP JPA batch 2014 friends, Nazihah Shahari, Liyana Malek, Bashirah Fazli and much more. You are all stars. I hope everything goes well in your research too. Also, the Malaysian community in Southampton, thank you for the sisterhood.

My whole family in Malaysia that we have left (physically) for years, I would like to express guilt for the missing years in all family matters. Although we were absent but just so you know you are always in my thoughts. Special recognition goes to my stepmother, Noor Haiyaty, and my brothers Fareez and Faiq for helping me take care of personal matters, and my step brothers and sisters, sorry for missing your weddings and your kids birthdays. Hopefully, we can make it up to you. My in-laws, especially my mother-in-law, thank you for being considerate. I cannot express how desolate it is that I cannot be there during the passing of my father in law two years ago. I thank my brothers and sisters-in-laws for being there replacing the comfort of my husband and me. From the bottom of my heart, thank you.

To every Tom, Dick, and Harry that I have met along the journey, thank you for teaching me something. I have learnt a great deal. I may forget your name but never the lessons.

Last but not least, to my altruistic ally, my loving husband, Amir Mohmad and my bubbly daughter, Maya Zara for their sacrifice and unconditional love. Your understanding that time is envious of us and each second was never wasted but put into the effort to finish this piece of work is a treasured devotion. This thesis is dedicated to both of you.

Definitions and Abbreviations

Chapter	Abbreviation/Symbol	Definition/Description
1 and 2	API	Application Programming Interface
	CSA	Cloud Security Alliance
	CSP(s)	Cloud Storage Provider(s)
	IaaS	Infrastructure as a service
	ICT	Information Communication Technology
	IT	Information Technology
	PaaS	Platform as a service
	SaaS	Software as a service
	RAM	Random-access memory
3	HTML	HyperText Markup Language; a set of codes used to display on a World Wide Web browser page(s)
	REST	Representational state transfer
	STRIDE	A threat classification model developed by Microsoft for computer security threats
	SSH	Secure Shell; a cryptographic network protocol
	SSL Certificate	Digital certificate; A data file that digitally bind a cryptographic key to an organisation
5	α	Alpha; level of significance
	β	Power value
	d	Cohen's measure of effect size
	df	Degree of Freedom; the number of values that are free to vary
	n	Number of items
	ρ	Probability
6	GQM	Goal-Question-Metric
	CVR	Content Validity Ratio
	r	Pearson Correlation
	p	p-value; statistical significance level
	α	Alpha; Cronbach's Alpha index of internal consistency
7	X^2	Chi Square
	AGFI	Adjusted Goodness of Fit
	CFI	Comparative Fit Index
	CMIN	Minimum Discrepancy
	CR	Critical Ratio
	df	Degrees of Freedom; the numbers of knowns minus the number of free parameters; used in many measures of fit.
	GOF	Goodness of Fit
	GFI	Goodness Fit Index
	IFI	Incremental Fit Index
	KMO	Kaiser-Meyer-Olkin; measure of sampling adequacy
	N	Number of participants/sample size
	NPAR	Number of Parameters
	NFI	Normed Fit Index
	P	Probability
	PCLOSE	Close of Fit
	PGFI	Parsimony Goodness-of-Fit-Index
	RFI	Relative Fit Index
	RMR	Root Mean Square
	RMSEA	Root Mean Square of Approximation
	RUD	Rational Unified Process
	SE	Standard Error
	SEM	Structural Equation Modelling
	TLI	Tucker-Lewis Index

To my father and late mother

For raising me to believe anything is possible

To my husband

For making everything possible

And to my daughter

For making everything possibly incredible

Chapter 1: Introduction

1.1 Overview of Research

As the cloud becomes the tool of choice for data storage services, the number of cloud storage providers (CSPs) is also increasing. From these providers, users have a wide selection of services available to move their data into the cloud. However, the security factors and concerns maintaining the security of sensitive data stored therein remains paramount (Zissis and Lekkas 2012).

Cloud storage provides a facility for those users who mainly require highly scalable storage on demand that is accessible globally. Since 2016, each user is expected to store at least 3.3 terabytes of their data in the cloud witnessed by the rising number of users for the commercial leaders in personal cloud storage: Dropbox, Box, and Google Drive (Gartner 2012). Despite the benefits cloud storage may bring to users through convenience and lower cost, it also may bring security concerns. These worries have been identified by various researchers (CSA 2010, 2013a, Sabahi 2011, Shaikh and Haider 2011, GTISC and GTRI 2013) who conclude that security concerns are increasing significantly every year.

Recent reports by Cloud Security Alliance (CSA 2013a) and Georgia Technology Information Security Center (GTISC and GTRI 2013) reveal a trend of insecure APIs, data loss and leakage concerns, as outsiders gain access to unencrypted data. Several simple methods are employed: a user password may be discovered by brute force, and unencrypted local files or folders located in the cloud can easily be accessed. The cloud service itself can also be compromised. In these cases, users and CSPs should implement security measures before the data is stored online.

CSPs have been implementing security measures to secure access to sensitive data in the cloud, such as two-factor authentication, encryption, etc., making access to the data more difficult for attackers (Mather et al. 2009). Then again, an increase in security measures affects the usability of the data and therefore may cause the system to be avoided by users (Honan 2012, Zarandioon et al. 2012, Zhang and Chen 2012, Zissis and Lekkas 2012, Zhao and Yue 2014). Most CSPs are unwilling to reduce the efficiency of access to cloud storage because users expect equally efficient access into secured data as into plain text data. Security protection based on security factors and mitigating its concerns is foreseen as one of the efforts to overcome this issue.

The cloud is an environment where users share the resources and to store their data online. To secure data in the cloud, many different security frameworks have been proposed. The existing security frameworks, industry accepted standards, and best practices have focused on securing the generic cloud implementation. A review of existing studies revealed a lack of insights and guidelines to secure data specifically in cloud storage facilities. There has been less study conducted on synthesising security factors specifically for cloud storage.

Cloud storage has not been well explored. Many researchers have considered cloud storage within a generic cloud computing model. However, since cloud storage by necessity stores data outside the control of the data owner, this has raised specific security concerns. In order to understand security challenges and measures for securing data in cloud storage, the security concerns are abstracted from the threat landscape at a higher level. As the result, security measures can be addressed more effectively by specifying the important security factors associated with respective concerns.

Security measures will try to address the common security goals: protect the confidentiality of data, preserve the integrity of data, and promote the availability of data for authorised use. What is particularly missing in the existing literature is a research to identify the factors associated with the security goals in the context of cloud storage, security concerns affecting these factors, underlying relationships between these identified security factors, and empirical testing based on the data collected.

1.2 Research Questions

This research aims to build an appropriate framework that can help improve the security of cloud storage. To achieve the research aim, the following research was conducted in three phases; framework development and confirmation, instrument development and validation and establishing the relationship(s) of the security factor(s). The first phase includes the development of an appropriate security framework for cloud storage. The research questions asked and answered are:

1. What is an appropriate security framework for cloud storage?
    a. What are the factors of a security framework in cloud storage?
    b. What are the security threats classification to a cloud storage?

The framework is developed by exploring existing security frameworks, addressed by previous research, and recommended by industry-accepted standards. Consequently, the results will identify security factors that can be used as baselines to protect data in cloud storage.

In the second phase, based on the confirmed framework, an instrument was developed to demonstrate the application of the framework. Below are the questions asked and answered:

2. How can stakeholder(s) evaluate the level to which the cloud storage security framework is being followed?
    a. What is a suitable instrument to evaluate security in cloud storage?
    b. How can the instrument be validated?

The final phase involved applying the validated instrument in a research scenario to establish the relationship(s) between the investigated factor(s). The questions asked and answered are:

3. What are the relationship(s) among security factors identified from Security Rating Score (SecRaS)?
    a. What are the relationship(s) among the security factors identified from factor analysis and structural equation modelling?
    b. Which relationship(s) will affect security implementation in cloud storage?

The next section will briefly describe the thesis structure. Figure 1.1 illustrates the chapters in the thesis and how these chapters are also presented to answers the research questions.

1.3 Thesis Structure

This thesis described a research to investigate how data can be protected in cloud storage. Two main aspects set the background for this research: (1) the security frameworks in cloud storage, and (2) the security concerns in cloud storage. Following a critical review of relevant work in cloud computing and cloud storage in Chapter 2, a set of synthesised factors was produced as the foundation of the framework. In Chapter 3, a threat classification was undertaken for security concerns in cloud storage. The threat landscape is used as a basis of concerns in cloud storage environment. Next, the development process and factors of the newly-developed Cloud

Storage Security Framework (CSSF) was discussed in Chapter 4. The framework indicates that securing data in cloud storage can be determined by seven factors. Having developed the framework, subsequently in Chapter 5, 20 security experts from the industry and academia was interviewed to confirm the CSSF's factors suitability within the framework structure. A number of 34 security practitioners were also surveyed to support the expert's recommendations for CSSF. An extension of the framework added two more factor; accountability and auditability making the framework a nine-factor framework. Positive results from this study demonstrate that factors of the CSSF are theoretically sound and that further applications of the CSSF can be explored.

Having established the validity of the framework, Chapter 6 demonstrated an application of CSSF through the development and validation of a measuring instrument called Security Rating Score (SecRaS). Positive results from the validation study support the SecRaS as a reliable measuring tool that can be used in a research scenario and suggest that findings from SecRaS can be used to inform research and design. Next, Chapter 7 demonstrated the application of SecRaS in a research scenario. The validated instrument was distributed to security practitioners in Malaysia. A number of 218 usable responses from the instrument was obtained from SecRaS and analysed using factor analysis and structural equation modelling. Results from this study highlight the security factors and relationship(s). The final parts of the thesis, in Chapter 8, brings together the work undertaken thus far in this research programme. This chapter presents the discussion of the research, the theoretical and practical implications, and limitations of the research. Lastly, this chapter also highlights the contributions of the research and identifies several potential directions for future work in applying and refining the CSSF. The chapter concludes by summarising the motivations, aim, and contributions of this research.

As part of the research, several appendices are included in order to clarify and complete some of the contributions of the thesis. In Appendix A, information related to the confirmation of the framework is detailed. This includes the interview questions and thematic analysis, and the survey questions and statistical results. Appendix B contains the detailed metrics developed using Goal-Question-Metrics for each factor and the SecRaS instrument. Also in Appendix B are the result of the validation study for SecRaS instrument. Finally, in Appendix C and D contains data relevant to the development of the model using exploratory factor analysis, confirmatory factor analysis, and structural equation modelling.

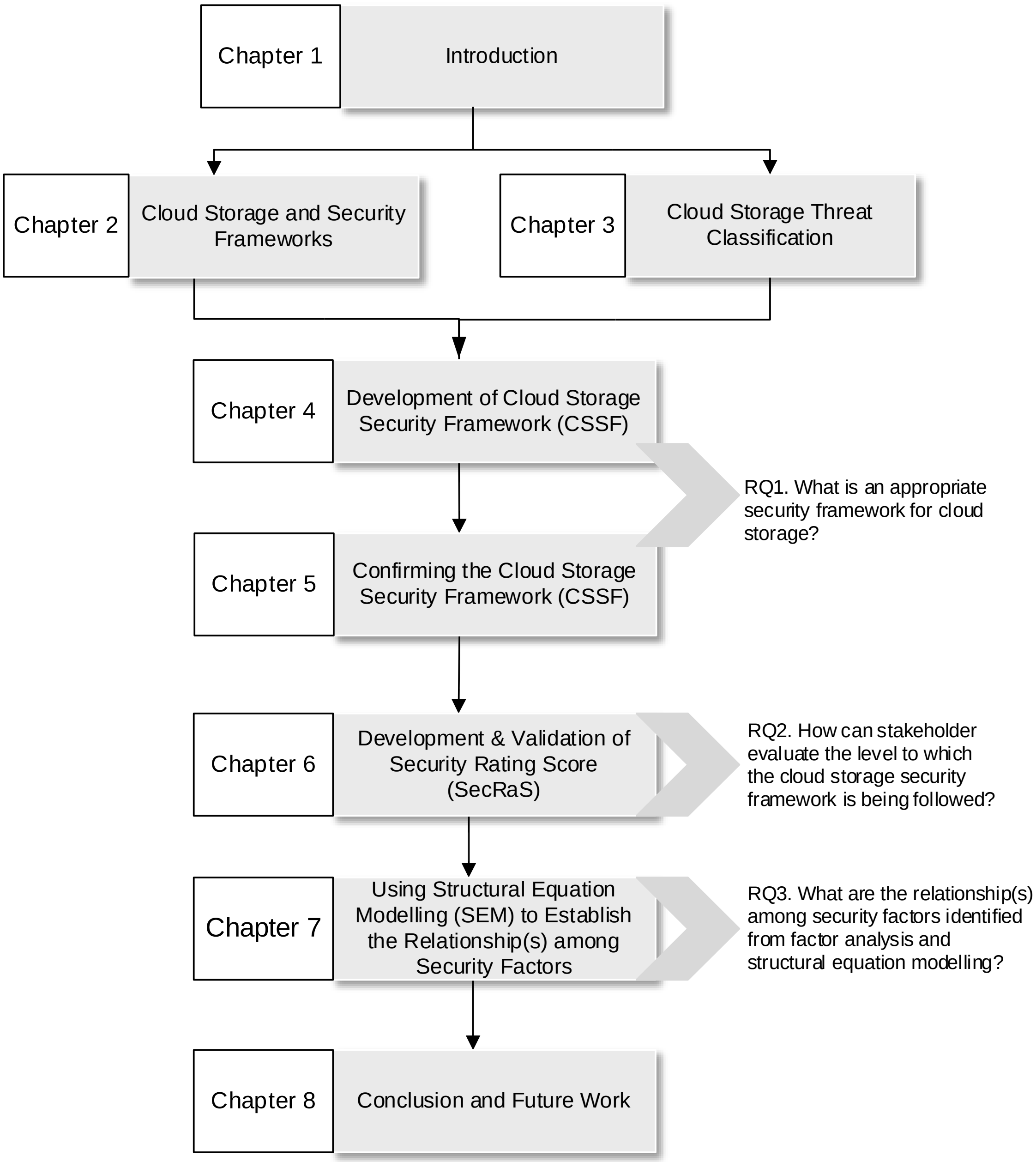

Chapter 1
Introduction
Chapter 2
Cloud Storage and Security Frameworks
Chapter 3
Cloud Storage Threat Classification
Chapter 4
Development of Cloud Storage Security Framework (CSSF)
RQ1. What is an appropriate security framework for cloud storage?
Chapter 5
Confirming the Cloud Storage Security Framework (CSSF)
Chapter 6
Development & Validation of Security Rating Score (SecRaS)
RQ2. How can stakeholder evaluate the level to which the cloud storage security framework is being followed?
Chapter 7
Using Structural Equation Modelling (SEM) to Establish the Relationship(s) among Security Factors
RQ3. What are the relationship(s) among security factors identified from factor analysis and structural equation modelling?
Chapter 8
Conclusion and Future Work

Figure 1.1 Thesis structure

Chapter 2: Cloud Storage and Security Frameworks

This chapter provides the context of this research. The purpose of this chapter is to provide an understanding of cloud computing, specifically cloud storage by discussing its definition and emerging security issues in the cloud. Common security factors for cloud storage are reviewed based on existing literature and industry standards. These goal-driven factors will later be used as a baseline in constructing the security framework.

2.1 Cloud Computing

Cloud computing can be defined as a computing style with scalable computing capabilities that can be delivered ‘as a service’ to users using Internet technologies (Weiss 2007, Mell and Grance 2009, Vaquero et al. 2009). In current technology development, users can easily gain access to applications or take full advantage of their resources in the cloud at a small cost.

According to the NIST, cloud computing has five main characteristics (Mell and Grance 2009).

1. On demand service – users should receive the intended computing services from their service providers.
2. Broad Network Access – the service should be available for access through the network using any type of devices, such as tablets, laptops, and computers.
3. Resource Pooling – cloud resources are pooled and used by different users concurrently (multi-tenancy); the user might choose the location or region for his machines but without being aware of their exact location.
4. Rapid Elasticity – users should be able to acquire suitable computing capabilities. Users may automatically increase and decrease the resources used.
5. Measured Service – users pay only for what they use.

In cloud computing, there are three recognised service models: software-as-a-service (SaaS), platform-as-a-service (PaaS), and infrastructure-as-a-service (IaaS) (Mell and Grance 2009, Weinhardt et al. 2009, Ertaul et al. 2010). These are explained below.

1. SaaS model – service providers offer their applications to users through the network. Users can access the applications using thin clients and browsers or program interfaces designed to communicate with the other applications hosted in the cloud.
2. PaaS model – mainly offered for applications developers as a development environment on which to host and support developers with libraries, services, tools, networks, and storage.
3. IaaS model – provides the infrastructure and resources to host users' machines (virtual machines). IaaS providers offer computing power, storage, networks, and any other supporting resources to host virtual machines (VMs). Each host in the cloud IaaS model is occupied by a number of VMs sharing the resources; the VMs are isolated from each other by the virtualization layer.

IaaS, PaaS, and SaaS signify an abstract level, with IaaS being the lowest abstraction level. IaaS may provide services to PaaS and SaaS, PaaS may provide services to SaaS but not to IaaS, and SaaS may not provide services to PaaS or IaaS (Milenkoski et al. 2013).

2.2 Cloud Storage

Computer storage is where data is held. Storage has been divided into primary storage, which holds data in memory (RAM) and other "built-in" devices such as the processor's L1 cache, and secondary storage, which stores data on hard disks and other devices requiring input/output operations (Rouse 2005). Primary storage provides faster access than secondary storage mainly for the reason of its close location to the processor or the architecture of the storage devices. On the other hand, secondary storage can store more data than primary storage, and this includes external hard disks and USB flash drives (Rouse 2005).

Cloud storage is known as utility storage if delivered through public cloud service providers (Wu et al. 2010). On the other hand, private service providers offer the same scalability, flexibility, and storage mechanism with restrictions or non-public access. Cloud storage runs on a virtualization platform providing end users and applications with a scalable and provisioned virtual storage architecture. Generally, cloud storage is accessed through an API (Wu et al. 2010, Ju et al. 2011).

Cloud storage can be defined as a cloud computing model that stores data on distributed servers and is accessible anywhere through the Internet. A cloud service provider maintains, operates, and manages storage servers that are built using virtualization techniques (Wu et al. 2010).

Cloud storage commonly fits the SaaS level. The stored units themselves are always built as resources accessed by Representational State Transfer (REST) (Dewan and Hansdah 2011).The range of services cover storage protocols like iSCSI (Satran et al. 2004), storage of files as provided by Dropbox (Miller 2013, Dropbox 2016), SpiderOak (SpiderOak 2016), Google Drive (Google 2016a), Microsoft Skydrive (Microsoft 2016a), or professional No-SQL databases such as AWS S3 (Amazon S3 2016), Google Cloud Storage (Google 2016b) or Microsoft Azure (Microsoft 2016b).

2.3 Existing Secure Cloud Storage Technologies

The analysis of existing technologies focuses on secure cloud storage technologies due to the variety of different cloud storage. A report describing existing solutions for secure file-based cloud storage was published by the Fraunhofer Institute for Secure Information Technology (SIT) (Borgmann et al. 2012). SIT evaluates common cloud storage providers and identified that many cloud storage providers mirroring the data locally. Storing data locally ensures availability when network access to the remote data is interrupted. These providers offer server-side versioning: Dropbox (Dropbox 2016), Wuala, which is no more in operation (Covic 2015) and TeamDrive (TeamDrive 2016). Not yet available at the publication date of the report was Google Drive, offering local mirroring and versioning (Google 2016a). Wuala and TeamDrive offer client-based encryption of the data. Integrity guarding techniques like checksums are not available in any of the products evaluated. Neither offers any user the ability to reconstruct unavailable data. The applied security measures to gain confidentiality by Wuala and Teamdrive are furthermore vulnerable.

EncFS (EncFS 2016) and TrueCrypt (TrueCrypt 2016) offers additional security layers such as the possibility to gain confidentiality independent of the location of the data. EncFS represents a file system in user space and requires supplementary implementations on the operating system. Boxcryptor (Boxcryptor 2016) builds on top of EncFS. Boxcryptor has been a dedicated solution for cloud storage but the implementation is unfortunately closed source. Moreover, Boxcryptor does not guarantee accessibility despite the fact that EncFS protects the data against unpermitted changes. On the other hand, an encrypted container is generated by TrueCrypt. Small changes in single files might result in a complete rewrite-operation depending on this container. Afterwards, the rewritten container needs to be transferred to the cloud.

SpiderOak (SpiderOak 2016) represents a classic SaaS provider focussing on the satisfaction of security requirements: locally cached and former versions are held. Remotely lost data can thereby be reconstructed. Additionally, parts have already been released to the open source community. Spideroak has the ability to work with user-owned servers. Strong security measures become obsolete in this scenario. Further examples relying on own servers are Sparkleshare (SparkleShare 2016) and Owncloud (Martini and Choo 2013, OwnCloud 2016). Since these approaches run on ownhosted servers, no security measures need to establish, although Owncloud optionally encrypts the data on the server. Custom-built systems like Duplicity (duplicity 2016) enables skilled users to create encrypted and checksummed archives. These archives are directly storable on various remote stores including Amazon S3. Table 2.1 summarises some of the security measures implemented by existing cloud storage such as Dropbox, Wuala, Teamdrive, Google Drive, SpiderOak, Sparkleshare, Owncloud, Duplicity, EncFS, Truecrypt and Boxcrypt. Suitable specific block stores in public clouds are rarely protected (Graf et al. 2012, Graf 2014). Lower-level data like blocks rely on dedicated storage protocols such as iSCSI (Satran et al. 2004). The overlaying file systems are able to manage confidentiality. Integrity, as well as availability, are achieved by additional techniques like summarising multiple remote volumes using RAIDs.

Table 2.1 Existing Products Providing Secure Cloud Storage Technologies (Graf 2014)

Product	Type	Access	License	Backend	Versioning	Integrity Checks	Encryption
Dropbox	Third-party	REST & File	Closed	Proprietary	✓		
Wuala		REST & File	Closed	Proprietary	✓		✓
Teamdrive		File	Closed	Proprietary & Own	✓		✓
Google Drive		REST & File	Closed	Proprietary	✓		
SpiderOak		File	Partly Open	Proprietary	✓	✓	✓
Sparkleshare		File	Open	Proprietary & Own	✓		✓
Owncloud		REST & File	Open	Own	✓		✓
Duplicity		File	Open	Own	✓	✓	✓
EncFS	Layer	File	Open	Any			✓
Truecrypt		File	Open	Any			✓
Boxcrypt		File	Closed	Any			✓

The next section explains on cloud security frameworks and associated goal-driven security factors in the context of cloud storage.

2.4 Cloud Security Frameworks

The cloud stores a large quantity of data, and executes many (and heavy) computational processes possibly valuable and belonging to many numbers of users. (Zissis and Lekkas 2012). In this section, security technological factors are reviewed based on existing literature and industry standards to discuss the common factors for a security framework.

2.5 Security

The Institute for Security and Open Methodologies (ISECOM) states that security provides protection where a separation is created between the assets and the threat (ISECOM 2001). These separations are generically called security measures and sometimes include changes to the asset or the threat. The degree of resistance to or protection from harm may apply to any vulnerable and valuable asset, such as a person, community, nation, or organisation.

The cloud has enhanced security in many ways. For instance, the cloud provider is trusted to implement better and more recent security technologies and practices than the data owner (Ryan 2013). Conversely, data is stored outside the control of the data owner, which inevitably introduces security issues (Subashini and Kavitha 2011, Srinivasan and Rodrigues 2012, Suntharam et al. 2013).

2.6 Frameworks for Security

A framework is a basic structure underlying a system or concept. It is a broad overview, outline, or skeleton of interlinked items which supports a particular approach to a specific objective and serves as a guide that can be modified as required by adding or deleting items (Firesmith 2004).

IT security frameworks and standards can be helpful in addressing many areas such as encryption, application security, and disaster recovery. An information security framework is a series of documented processes that are used to define policies and procedures around the implementation and ongoing management of information security measures in an enterprise environment (Granneman 2013).

These frameworks are basically a blueprint for building an information security programme to manage risk and reduce vulnerabilities. Information security can utilise these frameworks to define and prioritise the tasks required to build security into an organisation. Frameworks can also be customised to solve specific security problems to meet the required specifications and use (Firesmith 2004, Granneman 2013).

2.7 Cloud Security

The main features of cloud computing include scalability through elastic resource management, on-demand provisioning and no upfront cost by enabling cloud users to use services when they need them and pay only according to what they use. In cloud computing, data is usually held in large data centres, where virtualization technologies enable better resource utilisation by sharing hardware resources whilst still providing separation between cloud users.

Despite its advantages, the cloud also has a number of concerns and measures (Subashini and Kavitha 2011, Srinivasan and Rodrigues 2012, Wang et al. 2012, Ryan 2013) in particular with respect to security, including:

- Data protection: data from one user must be properly segregated from that of another, i.e. it must be stored securely and it must be able to move securely from one location to another. Cloud providers have systems in place to prevent data leaks or access by third parties. Proper separation of duties should ensure that auditing and/or monitoring cannot be defeated, including by privileged users at the cloud provider.
- Identity management: every organisation will have its own identity management system to control access to information and computing resources. Cloud providers either integrate the customer’s identity management system into their own infrastructure, using federation or single-sign-on technology or provide an identity management solution of their own.
- Virtualisation security: virtual machines need to be adequately secured.
- Authorisation: providers should assure customers that they will have regular and predictable access to their data and applications. Sensitive data can be cached on devices, in order to ensure availability in intermittent connectivity situations, but its usage must be governed by the usage control directives.
- Application security: cloud providers should ensure that applications available as a service via the cloud are secure by implementing testing and enforcing

acceptance procedures for outsourced or packaged application code. Moreover, application security measures (application-level firewalls) should be in place in the production environment as well.

- Access control: contextual usage control involves protecting sensitive data according to specific permissions, conditions and obligations, and when contextual conditions are deemed secure, in terms for instance of physical location and Internet connections

Security issues associated with cloud broadly fall into two categories: those faced by cloud providers and those faced by users. In most cases, the providers must ensure that their infrastructure is secure and that their users' data and applications are protected while the users must gain assurance that the provider has taken appropriate security measures to protect their information. To ensure data security, data stored cannot be accessed by unauthorised users, and that data protection is maintained.

2.8 Security Goals

The term “CIA” is a widely recognised meaning of security representing the common security goals of confidentiality, integrity and availability (ISECT 2014a). Below is the definition of security goals in the cloud (CSA 2013b, ISO/IEC 2016):

- Confidentiality – Data is not made available or disclosed to unauthorised individuals, entities, or processes,
- Integrity – Data is accurate and complete, and
- Availability – Data being accessible and usable upon demand by an authorised entity.

This definition is extended in the literature by non-repudiation, authenticity and reliability (CSA 2013b, ISECT 2014a, ISO/IEC 2016).

- Non-repudiation – Ability to prove the occurrence of a claimed event or action and its originating entities,
- Authenticity – Data is original of what it claims to be, and
- Reliability – Ability to provide consistent intended behaviour and results.

2.9 Goal-Driven Security Factors for Cloud Storage

Security is a continuous process involving constant synergy between policies implementation and technical measures (Firesmith 2004, Brock and Goscinski 2010, Takabi et al. 2010, Zissis and Lekkas 2012, Mapp et al. 2014). The definition of the goal-driven security factors in the context of cloud storage are concisely explained by the following definitions supported with existing researches fulfilling the goals.

2.9.1 Realising Security Policies Implementation in Cloud Storage

Achieving security is a challenge not solvable by applying technical solutions only (Ma et al. 2008). Existing research on security policy implementation and its integration into secure service can be leveraged to build a comprehensive security framework in the cloud environment (Takabi et al. 2010). Recent developments show the need for techniques as well as for establishing policies, procedures and controls to protect user data (Firesmith 2004, Brock and Goscinski 2010, Zissis and Lekkas 2012). To add, regulations and technical measures may use different terminologies resulting in different interpretations questioning the same concerns. Therefore, consistent terminology in terms of security in cloud storage must be well-defined in cloud storage policies and procedures.

2.9.2 Ensuring Confidentiality

Confidential data handling in cloud storage is the protection of data by allowing only the intended recipient to access the data. Data should be handled correctly to prevent unauthorised exposure (Firesmith 2004). Brock and Goscinski (2010) has characterised security concerns of clouds by proposing a Cloud Security Framework (CSF) that takes into consideration cloud infrastructure protection to ensure confidentiality. Data can be protected by applying access controls, authentication and authorisation while handling data effectively (Vrable et al. 2012, Mapp et al. 2014, El-Booz et al. 2016). Ensuring confidentiality is performed straightforward: before data are uploaded to the cloud, user access to the data needed to be decided (El-Booz et al. 2016). Security in any system including cloud storage involves primarily ensuring that the right user gets access to only the authorised data in the authorised format at an authorised time and from an authorised location. Identity and access management is of prime importance in this regard (Habiba et al. 2014, Singh and Chatterjee 2015).

2.9.3 Integrity Checks on Remote Data

Integrity is the ability of a provider to detect changes or modifications to an original status of remote data stored in cloud storage. Some techniques implement integrity across a packet header and/or data field by creating a hash across the contents of the packet (Firesmith 2004). Most approaches ensuring confidentiality care about integrity of the data as well: Inconsistent access to the data automatically harms the modification to its correct status. Therefore, having encrypted the data before uploading to the cloud storage safeguards the remote data. Data is therefore equipped with checksums and probes. The status of the data must be checked continuously. The remote location makes incessant checks hard to perform. The “Proof of Retrievability” (POR) (Bowers et al. 2009a) tackles this problem. A Message Authentication Code (MAC) combined with an Error Correction Code (ECC) is applied on the buckets. The MAC detects large errors and is relying on units in the buckets. The ECC protects the bucket against small errors. The number of units for the MAC and the size of the ECC gives an assumption about a possible successful retrieval. This assumption is provided as probability to successfully access data. This technique can be combined with a cloud-of-clouds approach (Bowers et al. 2009b).The “Proof of Data Possession” (PDP) (Ateniese et al. 2007) represents a similar approach offering a probability of possession using sampling. Current approaches working with single clouds are used mainly for synchronisation (Mahajan et al. 2011). Checksums are combined with encryption to guard data (Yao et al. 2010). These approaches use the versioned data by generating a chain of hashes. Other approaches (Wang et al. 2012) combine sampling with erasure codes similar to the POR but are working on single clouds only. The usage of multiple clouds (Abu-Libdeh et al. 2010, Cachin and Haas 2010, Mu et al. 2012) needs sophisticated integrity checks guarding the data against single, faulty clouds. All data receives a version number ensuring consistency additionally to the accessibility.

2.9.4 Maintaining Availability

Availability guarantees the access to the data. Availability on the server-side is hard to be assured from a users’ perspective. The status of the cloud as well as the connectivity stays out of focus of a user. Measures to increase availability include mirroring the data in multiple clouds as well as local caching. Availability is the idea that the data is accessible to all authorised users at all times. Its unavailability may occur in a physical way, as the failure of critical network components, power disruptions, and physical plant disruptions, either malicious or natural (Firesmith 2004, Takabi et al.

2010). Availability can also be impacted in a logical way, in the form of improper addressing or routing, and through the use of Denial-Of-Service attacks, which are the deliberate insertion of unwanted data into the network (Brock and Goscinski 2010). This is often associated with address spoofing, which associates the introduction of unwanted data with a trusted end node. Zissis and Lekkas (2012) recommended a cloud system designed and maintained with important aspects, which include contingency planning for power failures and disaster recovery, is also part of a system availability (Firesmith 2004, Mapp et al. 2014). The availability of the data is nevertheless dictated by cloud storage providers. Ways to overcome this dependability are local caching of the data and/or the usage of multiple clouds (Cachin and Haas 2010). Similar to the idea of RAID, several approaches distribute the data in disjoint clouds (Abu-Libdeh et al. 2010, Cachin and Haas 2010). The "Proof of Retrievability" (POR) (Juels and Jr 2007, Bowers et al. 2009a) or "Proof of Data Possession" (PDP) (Bowers et al. 2009a) generates knowledge about the integrity and thereby indirectly about the availability. These approaches focus on cloud storage only. Similar techniques also exist in the area of P2P storage (Caronni and Waldvogel 2003).

2.9.5 Guaranteeing Non-repudiation to Data

Non-repudiation is to assign attribution, i.e. provenance, to data that a third party could verify and be confident that it cannot be disputed. It can also prevent a recipient from denying data was received. Firesmith (2004) highlights that non-repudiation attempts to provide a comprehensive interaction (e.g., transaction and transmission of data) is prevented from successfully repudiating (i.e. denying) any aspect of the interaction. Non-repudiation thus assumes data integrity so that a party cannot argue that the data was modified. Focusing on concurrent access, some approaches (Shraer et al. 2010) use optimistic, time-stamped writes. Combinations of integrity checks with probabilistic tests on remote data result in higher-level architectures (Kamara and Lauter 2010, Wei et al. 2010). Examples of these architectures are cloud-based file systems (Kamara et al. 2011, Stefanov and Dijk 2012, Vrable et al. 2012) or database systems (AlZain et al. 2011). These approaches use the idea of Merkle-Trees (Merkle 1988). The folder structure leverages from the tree structure in combination with remote integrity checks. Optionally, the task of integrity checks can be delegated to untrusted cloud components (Nepal et al. 2011). In this scenario, encrypting and the computation of checksums are performed by different cloud services.

2.9.6 Preserving Authenticity of Data Originality

The authenticity of data refers to its original conception by its owner or author. Maintaining this relationship of data and network communications is performed with the use of public key encryption and a process called digital signing (Brock and Goscinski 2010, Zissis and Lekkas 2012). To create a digital signature, a hash is created from the data. A hash ensures the data is coming from an authentic source (Mapp et al. 2014). When ownership of a digital signature secret key is bound to a specific user, it demonstrates that the data was sent by a valid user. Thus, authenticating the source of data. Focusing on collaborative use cases, the challenge is to provide suitable key management. The key management should make use of the scalability and availability of the cloud (Cachin and Haas 2010). Challenges are especially the key distribution and flexible access to versatile groups (Popa et al. 2011). Consequently, in preserving authenticity these approaches should also satisfy integrity and availability. Authenticity must include adaptable access rights mapped to different versions.

2.9.7 Reliability of Service Provider

Reliability refers to the ability of a provider to provide a consistent intended service (Brock and Goscinski 2010, Zissis and Lekkas 2012). Operational reliability and flexibility is needed in cloud environments using capabilities (Mapp et al. 2014). The proposed capabilities are functions developed into mechanisms using a capability-based approach. Several different models guarantee a reliable cloud service. Examples include logging and monitoring (Ko et al. 2011a), establishing procedural approaches (Pearson et al. 2012), combinations of sampling, replaying modifications and time-stamping or establishing an entire life cycle using all of these measures (Ko et al. 2011a). Focusing on cloud storage only, modifications must be traceable by a user. Some approaches guard integrity by versioning hashes. These approaches care about reliability as well (Mahajan et al. 2011). Other approaches put the data directly under version control including adjacent metadata (Shraer et al. 2010, Bessani et al. 2011, Stefanov and Dijk 2012).

These goal-driven security factors represent the basic design goals for securing data in cloud storage. The evaluation of current technologies and of current research approaches refers thereby to these design goals. Table 2.2 shows current technologies of current research approaches mapped to the goal-driven security factors described in Section 2.7. Most approaches satisfy more than one goal-driven security factors.

Table 2.2 Summary of goal-driven security factors for cloud storage from existing studies

Name and References	Confidentiality	Integrity	Availability	Non-repudiation	Authenticity	Reliability	Model	Technological Contribution
TOTP and ABP (El-Booz et al. 2016)	✓	✓		✓	✓		SaaS	Time-based one time password (TOTP) and automatic blocker protocol (ABP)
Data Classification for Cloud Storage (Tawalbeh et al. 2015)	✓	✓		✓		✓	SaaS	Confidentiality-based cloud storage framework assures confidentiality and integrity through data classification using TLS, AES and SHA based on the type of classified data.
SCS & DPaaS (Vu et al. 2015)	✓	✓			✓	✓	SaaS	Data Encryption as a Service with full life-cycle of data security management in a multi-cloud environment.
CSSF Capability ID system (Mapp et al. 2014)	✓	✓	✓		✓	✓	SaaS	Capability ID system as a mechanism for e-Health applications
Erasure code (Yao et al. 2013)	✓		✓			✓	PaaS	Storage servers and key servers, multiplicative homomorphic property of the public key encryption algorithm
Privacy-Preserving Public Auditing (Wang et al. 2013)		✓		✓	✓		SaaS	Homomorphic linear authenticator and random masking to enable public auditing
RBAC cloud storage (Zhou et al. 2013)	✓	✓	✓				SaaS	Role-based encryption (RBE) scheme that integrates the cryptographic techniques with RBAC
Bluesky (Vrable et al. 2012)	✓	✓					SaaS	Encrypted block stored in buckets equipped with checksums
μLibCloud (Mu et al. 2012)		✓	✓				SaaS	ECC-guarded writes on multiple clouds
Iris (Stefanov and Dijk 2012)	✓	✓	✓			✓	SaaS	Block-based integrity using on tree-ordered MACs
CloudProof (Popa et al. 2011)	✓		✓				PaaS	Chained hashes over encrypted blocks
DIaas (Nepal et al. 2011)		✓					PaaS	Integrity-Checks performed on untrusted services
CS2 (Kamara et al. 2011)	✓	✓					PaaS	PDP and queries over encrypted data
MCDB (AlZain et al. 2011)		✓	✓				PaaS	Deploying DBMS In a cloud-of-clouds
DepSky (Bessani et al. 2011)	✓	✓	✓			✓	SaaS	Usage of multiple clouds including error detection and encryption
Depot (Mahajan et al. 2011)		✓	✓			✓	SaaS	Weak consistent access using versioned-chained hashes

2.10 International and Industry Standards, Best Practice, and Guidelines

The interest in cloud computing has led an explicit and constant effort to assess the latest trends in security (Honer 2013). Effective governance in cloud computing environments follows from well-developed information security processes as part of the organisation's obligations (Srinivasan and Rodrigues 2012, CSA 2013a). In this section, IT industry standards in relation to promoting security are reviewed.

When the cloud was first introduced, the Cloud Security Alliance (CSA), a non-profit organisation developed and published cloud security best practice (CSA 2009). Almost all major cloud providers (such as Amazon, Oracle, RedHat, and Salesforce) are members of the CSA. Their efforts include identifying the top concerns. CSA conducted a survey of industry experts to compile professional opinion of the vulnerabilities within cloud computing. In the latest edition, experts have identified data loss and breaches, and insecure APIs as the critical concerns to cloud security (CSA, 2013a; 2013c). A compliance standard, called Cloud Control Matrix (CCM), was developed to provide standard security measures that can guide providers and help users in the assessment of the risks associated with a provider (CSA 2013c). The CCM is specifically designed as a control framework with security concepts aligned to CSA guidance in 13 domains. It also describes the relationship with other industry-accepted security standards, regulations, and control frameworks (such as the ISO 27001/27002, ISACA COBIT, PCI, NIST, Jericho Forum and NERC CIP) (CSA 2013c).

The National Institute of Standards and Technology, Security and Privacy Controls for Federal Information Systems and Organisations (NIST SP 800-53 Revision 4) was created to assist organisations in making the appropriate selection of security measures for information systems by introducing security control baselines (NIST 2013a). Security control baselines are used as a starting point for the security control selection process and are based on the security category and associated impact level of information systems determined in accordance with FIPS Publication 199 and FIPS Publication 200 (NIST 2004). The baselines address the security needs of a comprehensive and varied set of constituencies and are developed from several assumptions, including common environmental, operational, and functional considerations. However, the baselines also assume typical concerns facing common information systems (NIST 2013a) but not specifically in the context of a cloud or cloud storage. Moreover, the suggestion of security protections based on categories of impact (low, moderate and high) has also not been included in the latest revision.

The International Organisation of Standardisation (ISO) and the International Electrotechnical Commission (IEC) formed a worldwide standardisation through providing a model to follow in setting up and operating a management system. The standardisation provides guidance on the information security components of cloud computing, recommending and assisting with the implementation of cloud specific information security measure and controls (ISECT 2014a, 2014b). The standard is known as Information Security Management System (ISMS) family of standards. By applying the ISMS family of standards, organisations can develop and implement a framework for managing the security of their information assets such as financial information, intellectual property, and employee details, or information entrusted to them by users or third parties (ISO/IEC 2016). These standards can also be used to prepare for an independent assessment of their ISMS applied to the protection of information (ISO/IEC 2016).

The European Network and Information Security Agency (ENISA) developed an authoritative security reference that listed risks, vulnerabilities, and a survey of related research recommendations. It consists of a report and practical guides designed for managing security in the cloud. In the asset management section, security measures highlight the point that service providers should review user data sensitivity. Another recommendation is that providers request information from users whether deploying their data in the cloud is deemed as sensitive enough to require additional security protection. Service providers are also encouraged to apply appropriate segregation between systems with different classifications (Catteddu and Hogben 2009). The recommendation was only made in general and then only if there are sensitive data.

The United Kingdom Centre for the Protection of National Infrastructure (CPNI) has also provided critical security measures for cyber defence as baselines for high-priority information security measures and controls (CPNI 2014a). They can be applied across an organisation to improve its cyber defence. The Council on Cybersecurity is coordinating the development of these controls. In their guidelines, the 20 controls (and sub-controls) concentrate on technical measures and activities. The main goal is assisting organisations in prioritising efforts to secure against the current and most common attacks. Besides that, comprehensive security should take into account other areas of security such as policy, organisational structure, and physical security. CPNI has added these in their latest guideline publication (CPNI 2014b). However, this guideline has not discussed cloud security in depth but some recommendations can also be applied to the cloud context.

In 2011, the Australian Signals Directorate (ASD) published 35 best practice strategies to mitigate targeted cyber intrusions (ASD 2014a) but it was simplified into four top mitigation strategies in 2012 (ASD 2012) focusing on application whitelisting, patching applications and operating systems, using the latest version, and minimising administrative privileges. The strategies are ranked in order of overall effectiveness and are based on ASD’s analysis of reported security incidents and vulnerabilities.

These are derived from ASD security testing and audits on Australian government networks. At the same time, the top four mitigation strategies are expected to effectively help in achieving a defence-in-depth ICT system. The combination of all four strategies, if correctly implemented, will protect an organisation from low to moderately sophisticated intrusion attempts.

The Australian Government Information Security Manual (ISM) was published in 2014 and is the standard which governs the security of government ICT systems (ASD 2014b). It has 15 security aspects including physical security, personnel security, communications security, information technology security, product security, media security, software security, email security, access control, secure administration, network security, cryptography, cross-domain security, data transfers and content filtering, and working off–site. There is an interesting section on protecting classified information and suggestions on encryption methods to protect confidential, secret and top secret information. The ISM comprises three documents targeting different levels within the organisation, making the ISM accessible to more users and promoting information security awareness in Australian government agencies.

These industry-accepted standards, guidelines and best practice are reviewed and mapped against the goal-driven security factors discussed in the previous section (Firesmith 2004, Brock and Goscinski 2010, Takabi et al. 2010, Zissis and Lekkas 2012, Mapp et al. 2014). The Cloud Control Matrix (CCM) has been used as a reference document to crosscheck with other standards. NIST and ENISA are already included in CCM but are separated in this research in the context of cloud storage as it discusses information security in general. CPNI and ASD are not included in CCM as they provide the latest security guidelines and therefore have been thoroughly analysed. The summary is shown in Table 2.3 where the international and industry standards, guidelines, and best practices are mapped accordingly towards the security factors (Pohlman 2010, CSA 2013c).

Table 2.3 Summary of security goals pointed out in international and industry standards, guidelines, best practice

Organisation Security goals	(CSA 2013c)	(NIST 2013a)	(ISECT 2014b, ISO/IEC 2016)	(ENISA 2009)	(CPNI 2014a)	(ASD 2014b)
Confidentiality	✓	✓	✓	✓	✓	✓
Integrity	✓	✓	✓	✓	✓	✓
Availability	✓	✓	✓	✓	✓	✓
Non-repudiation	✓		✓			
Authenticity	✓		✓	✓	✓	✓
Reliability			✓	✓	✓	✓

2.11 Chapter Summary

Cloud computing is recognised as a scalable user service. In some ways, it has improved security but since cloud storage necessarily puts data outside the control of the data owner, this will raise cloud-specific security issues. Cloud storage that has derived from the cloud computing model is used to store data in the cloud. With the rise of cloud computing, security issues in cloud storage have surrounded users, practitioners and providers. With concerns about data in cloud storage, previous studies were undertaken in the area of cloud storage security; this involved accomplishing security goals (confidentiality, integrity, availability, non-repudiation, authenticity, reliability) by developing frameworks to guide users and cloud storage providers (CSPs). This has encouraged governing bodies and agencies to publish standards, best practice and guidelines that can be used as references by those adopting cloud computing. Cloud Security Alliance, in particular, has been actively developing guidelines and the Cloud Control Matrix is among the important ones that map the controls to other standards protection domains. NIST published NIST 800-53 R4 that presented establishing responsibilities for implementing important controls in the cloud. Previous research has shown that organisations and CSPs have implemented many controls to ensure security and data protection. However, some measures involve many controls that most CSPs are reluctant to impose, as it is likely to decrease the efficiency of accessing the cloud. Applying controls based on security factors and concerns is proposed to protect data efficiently in the cloud. The next chapter discusses concerns to the cloud and threats classification for cloud storage.

Chapter 3: Threat Classification

3.1 Computer Threat

In computer security, a threat is a risk of potential danger to a computer system that might exploit a vulnerability to breach security and therefore cause possible harm (Microsoft 2015). A vulnerability is a weakness which allows an attacker to reduce a system's information assurance. The vulnerability is the intersection of three elements: a system susceptibility or flaw, attacker access to the flaw, and attacker capability to exploit the flaw. Vulnerabilities may lead exploitation of the weaknesses (Microsoft 2015). The weaknesses may be caused by gaining unauthorised access to stored information, by denial of service to authorised users, or by the introduction of false information to mislead users or trigger incorrect system behaviour (called spoofing) (Wang et al. 2010).

Concerns to the cloud include interception, modification of data at rest and in transit, data interruption (deletion), data breach, impersonation, session hijacking, and exposure in the network (Sabahi 2011, Shaikh and Haider 2011, Zissis and Lekkas 2012, CSA 2013a, GTISC and GTRI 2013). Consequently, with these emerging concerns, research has focused on security frameworks in the cloud (Mather et al. 2009).

Several terms are introduced to fully understand threat classification for cloud storage. The term *computer security* means to protect information. It deals with the prevention and detection of unauthorised actions by users of a computer. Lately, it has been extended to protect confidentiality, integrity and availability (Microsoft 2015). A system is a set of interacting or interdependent component parts forming an intricate whole. An asset is any data, device, or other components of the environment that support information-related activities (ISO/IEC JTC 1/SC 27 2004). Assets generally include hardware (e.g., servers and switches), software (e.g., mission critical applications and support systems) and confidential information (ISO/IEC JTC 1/SC 27 2004, ENISA 2009). Assets should be protected from illicit access, use, disclosure, alteration, destruction, and/or theft, resulting in a loss to the organisation (Jones 2005, The Open Group 2009).

Threats analysis techniques have been introduced, such as DREAD and STRIDE, to consider concerns and elicit security factors that address the concerns (Swiderski and

Snyder 2004). A threat model allows security designers to estimate an attacker's capabilities. It might be tempting to skip threat modelling and simply extract the system security factors from industry's best practice or standards. However, these standards merely provide general security guidance. The common standards almost always need some customisation for the target system and additional factors need to be defined (Myagmar et al. 2005).

In this study, three-step threat modelling is used to identify the threats (Myagmar et al. 2005):

1. Characterising the system,
2. Identifying system assets, and
3. Identifying concerns.

The threat modelling process targets software applications since cloud storage provide software-as-a-service to users. Characterising the system involves understanding the system factors and their interconnections, and creating a system model emphasising the main characteristics. Then assets and access points of the system are identified. Identifying concerns creates a threat profile of the system, describing all the potential attacks that need to be mitigated against or accepted as low risk. Although these three steps are common to all types of system, the actual execution differs depending on the type of system. Each of these threat modelling steps is now elaborated in the context of cloud storage.

3.2 Threat Modelling

At the start of the threat modelling process, the security designer needs to understand the system in question. This entails understanding every factor and its interconnections, defining usage scenarios, and identifying assumptions and dependencies.

3.2.1 Characterising the system

A cloud storage scenario can be modelled with three participants: users, cloud instances, and cloud storage provider (Gruschka and Jensen 2010). Every interaction in a cloud scenario involves two interfaces of these participants.

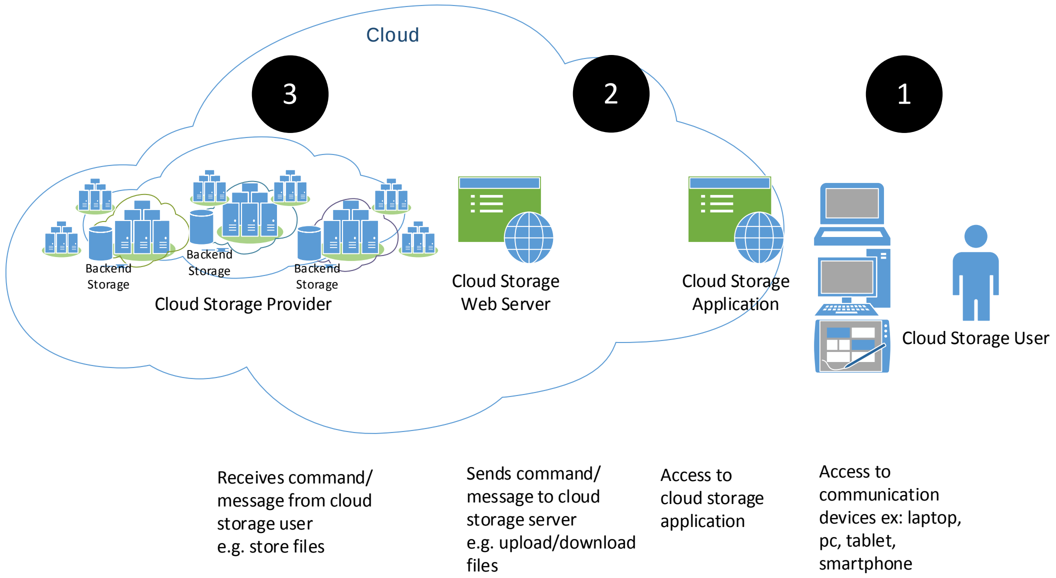

Cloud
3
2
1
Backend Storage
Backend Storage
Backend Storage
Cloud Storage Provider
Cloud Storage Web Server
Cloud Storage Application
Cloud Storage User
Receives command/ message from cloud storage user e.g. store files
Sends command/ message to cloud storage server e.g. upload/download files
Access to cloud storage application
Access to communication devices ex: laptop, pc, tablet, smartphone

Figure 3.1 Cloud Storage Scenario

As illustrated in Figure 3.1 above, interaction examples are a user requesting a service, or a service instance requiring more storage from the cloud storage provider. In the same way, every attack attempt in the cloud scenario can be analysed into a set of interactions within this model.

3.2.2 Identifying system assets

Each of the three participant roles provides a specific kind of interface to the other participants. For instance, the cloud system provides every service instance with a specific interface (API depending on the service model type, IaaS, PaaS, or SaaS) that the service instance can use, i.e. run on. In the same way, a service instance provides its service to a user with a dedicated interface (e.g., website, SSH connection, Web Service). Thus, with the three participants, there are six such interfaces to consider (as shown in Figure 3.2). Below are the descriptions involved in this triangle of attack.

1. Service-to-user

   The first and most prominent attack is a service instance towards a user. The common server-to-client interface is vulnerable to all kinds of attacks that are possible in common client-server architectures. This involves threats like account hijacking from SQL injection or privilege escalation (CSA 2013b).

2. User-to-service

The threats included in user to service attacks involves user programs requesting services from the server. For an example, browser-based attacks for an HTML based service like SSL certificate spoofing (Marlinspike 2009), attacks on browser caches, or Malwares or Phishing attacks (CSA 2013b).

3. Cloud-to-service

The interface between a service instance and a cloud system is a bit more complex. Here, the separation of service instance and cloud provider can be tricky, but in general, the cloud system's attack on the service instance includes all threats that a service instance can run against its hosting cloud system. Examples would be resource exhaustion attacks, triggering the cloud provider to provide more resources or end up in a Denial-of-Service, or attacks on the cloud system hypervisor (CSA 2013b).

4. Service-to-cloud

The other way around, the attack by a service instance against the cloud system is a very sensitive one. It incorporates all kinds of attacks a cloud provider can perform against a service running on it. This may start with availability reductions such as shut down of service instances. It may also cover privacy-related attacks (scanning a service instance's data in the process) or even malicious interference, e.g., tampering with data in the process, injecting additional operations to service instance executions; rootkit (Brumley 1999).

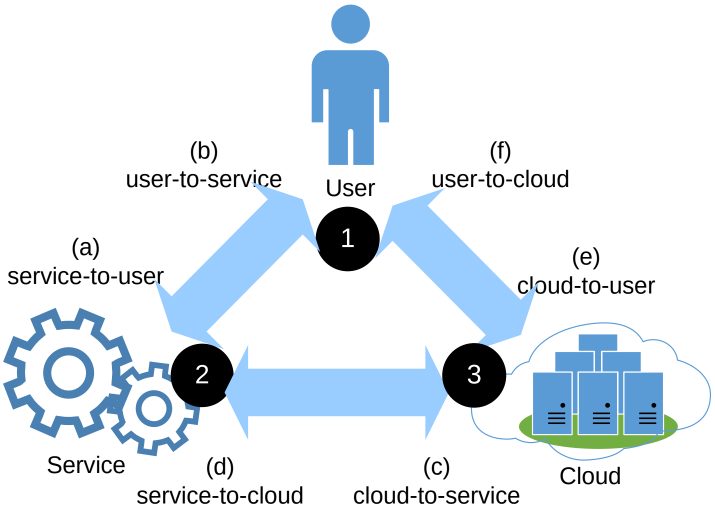


Figure 3.2 Cloud Computing Triangle Attacks

(Modified from Gruschka and Jensen, 2010)

5. Cloud-to-user

   The fifth attack surface of interest is that of the cloud system towards the user. This threat is challenging to define since both usually do not have a real contact point; in common scenarios, there always exists a service in between. However, the cloud system has to provide an interface for controlling its services. That interface, called cloud control, provides cloud users with the ability to add new services, require more storage, delete data in a cloud storage etc. As this is not a service instance, it is discussed as a separate attack, with threats being similar to the ones a common cloud service has to face from a user.

6. User-to-cloud

   The last attack surface is the one provided by a user towards the cloud provider. Considerable attacks may involve phishing-like attempts to trigger a user into manipulating its cloud-provided services, such as presenting the user a fake usage bill of the cloud provider. In general, this threat involves every kind of attack that targets a user and originates or spoofs to originate at the cloud system.

3.2.3 Identifying concerns

After completing the previous steps, specific concerns related to cloud storage are identified. Concerns may come from either inside or outside the system, from authorised users or from unauthorised users, who masquerade as valid users or find ways to bypass security mechanisms (Myagmar et al. 2005). Concerns can also come from human error. The goal of this step is to identify concerns to the system using the information gathered so far.

The threat is the adversary goal, or what an adversary might try to do to a system (Swiderski and Snyder 2004). Sometimes it is described as the capability of an adversary to attack a system. A list of known concerns and vulnerabilities found in similar systems is often the place to start with threat modelling. System specific concerns require deeper analysis of the unique qualities of the system being modelled. The concerns in cloud storage are identified as (Gruschka and Jensen 2010, Shaikh and Haider 2011, CSA 2013b, GTISC and GTRI 2013):

1. Data breach
2. Data leakage and loss
3. Insecure APIs

4. Account hijacking
5. Denial of Service
6. Malicious insiders
7. Abuse of cloud service
8. Inadequate cloud planning
9. Cloud-related malware
10. Closure of cloud service
11. Natural disaster
12. Hardware failure
13. Shared technology vulnerabilities
14. Insufficient due diligence

A security factor can be mapped to security concerns showing the effects of each concern to the security goal a system is acquiring. Figure 3.3 shows the security factor and mapping to concerns according to the CSA Control Matrix (CSA 2013c). Reliability was added as an additional factor and also mapped with relevant concerns.

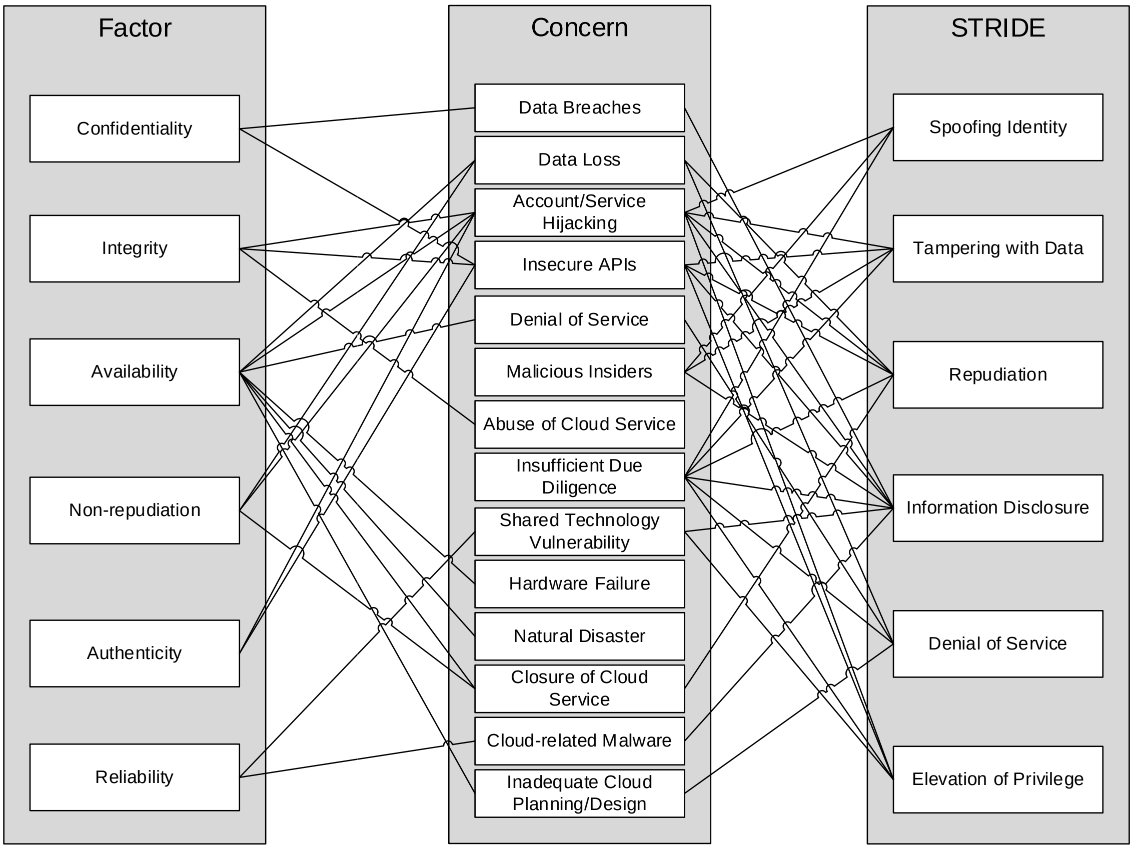


Figure 3.3 Cloud Storage Concerns (CSA 2013b)

STRIDE was used in this study because it fits the output of threat identification (Swiderski and Snyder 2004). The output of a threat identification process is a threat profile for a system, describing all the potential attacks, each of which needs to be mitigated or accepted. In general, threats can be divided into six classes based on their effect (Swiderski and Snyder 2004).

1. Spoofing – Using someone else’s credentials to gain access to otherwise inaccessible assets.
2. Tampering – Changing data to mount an attack.
3. Repudiation – Occurs when a user denies performing an action, but the target of the action has no way to prove otherwise.
4. Information Disclosure – The disclosure of information to a user who does not have permission to see it.
5. Denial of Service – Reducing the ability of valid users to access resources.
6. Elevation of privilege – Occurs when an unprivileged user gains privileged status.

When defining a threat model, security designers are concerned with defining attacks and also prioritising them (Myagmar et al. 2005). Risk assessment is performed to map each threat either into a mitigation mechanism or priority assumptions. The security factors for the system can be defined clearly once the concerns are identified as showed in Table 3.1 and Table 3.2.

Table 3.1 Threat classification with STRIDE threat modelling

Threat Classification	**STRIDE Threat Modelling**					
	S	T	R	I	D	E
Data breach	✓			✓		
Data leakage and loss			✓		✓	
Insecure APIs		✓	✓	✓		✓
Account hijacking		✓	✓	✓		✓
Denial of Service					✓	
Malicious insiders	✓	✓		✓		
Inadequate cloud planning					✓	
Cloud-related malware				✓		
Closure of cloud service			✓			
Shared technology vulnerabilities				✓		✓
Insufficient due diligence	✓	✓	✓	✓	✓	✓

* S – Spoofing identity, T – Tampering with data, R – Repudiation, I – Information disclosure, D – Denial of service, E – Elevation of privilege

Table 3.2 Threat classification with security factors

Threat Classification	**Security Factors**					
	C	I	Av	N	At	R
Data breach	✓					
Data leakage and loss			✓	✓		
Insecure APIs	✓	✓			✓	
Account hijacking		✓	✓	✓	✓	
Denial of Service			✓			
Abuse of cloud service		✓				
Inadequate cloud planning			✓			
Cloud-related malware						✓
Closure of cloud service			✓	✓		
Natural disaster			✓			
Hardware failure			✓			
Shared technology vulnerabilities						✓

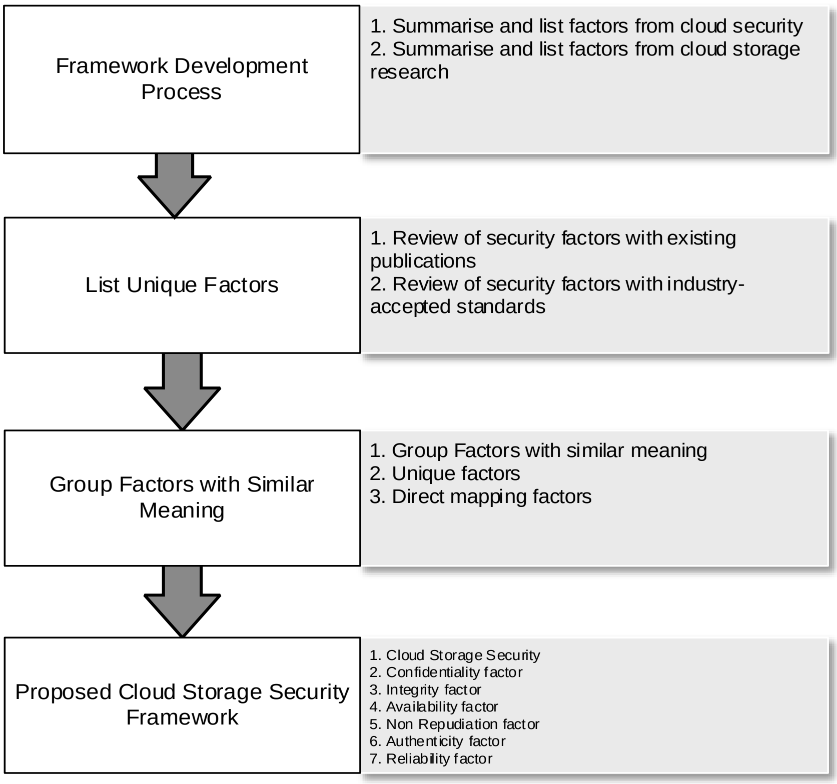
* C – Confidentiality, I – Integrity, Av – Availability, N – Non-repudiation, At – Authenticity, R – Reliability

3.3 Chapter Summary

This chapter discussed threat classification in the cloud. A threat identification approach was adopted to explore concerns in cloud storage. Cloud storage scenario was used to characterise the system and system specific concerns were analysed. Some of the concerns identified are data breach, data leakage and loss, insecure APIs, account hijacking, denial of service, malicious insiders, abuse of cloud service, inadequate cloud planning, cloud-related malware, closure of cloud service, natural disaster, hardware failure, shared technology vulnerabilities, and insufficient due diligence. A risk assessment tool, STRIDE was used to assess the identified concerns. The concerns and security factors are mapped based on existing studies and industry-accepted standards on cloud-related concerns. The next chapter discusses the development of the conceptual cloud storage security framework.

Chapter 4: Development of the Cloud Storage Security Framework (CSSF)

Chapter 2 and 3 discussed the existing cloud storage security frameworks and the threats classifications as this challenge greatly affect the security implementation in cloud storage. This chapter proposes a security framework for cloud storage to overcome the security issues in cloud storage.



Framework Development Process
1. Summarise and list factors from cloud security
2. Summarise and list factors from cloud storage research
List Unique Factors
1. Review of security factors with existing publications
2. Review of security factors with industry-accepted standards
Group Factors with Similar Meaning
1. Group Factors with similar meaning
2. Unique factors
3. Direct mapping factors
Proposed Cloud Storage Security Framework
1. Cloud Storage Security
2. Confidentiality factor
3. Integrity factor
4. Availability factor
5. Non Repudiation factor
6. Authenticity factor
7. Reliability factor

Figure 4.1 Development of the Cloud Storage Security Framework (CSSF) Process

4.1 Framework Development Process

The review of literature in chapter 2 has shown that few empirical studies have been conducted to investigate security frameworks in cloud storage. Moreover, there has been little research that specifically studies security factors and concerns in cloud storage. The framework proposed here is intended to identify the security factors and concerns in a cloud storage context. Development of this framework went through several stages. This section describes the stepwise approach to developing the framework and Figure 4.1 illustrates the process involved during the framework development. The diagram shows four main phases with eight steps taken in the process. The following section explains in detail how cloud storage security framework was developed.

4.2 List Unique Factors

The first step in the framework development is to list all the unique factors derived from the research findings in two areas: (1) cloud security frameworks, and (2) cloud storage. These research findings were the results of Chapter 2 analysis of literature in cloud security frameworks and cloud storage research. Both Table 4.1 and Table 4.2 show summarised factors and sources for those research findings. A total of 13 goal-driven factors with 20 potential characteristics items can be observed from the tables.

Table 4.1 depicts factors and items from cloud security research. There is a total of six factors and 6 character items. The first factor, confidentiality refers to undisclosed data to unauthorised users. Next, integrity refers to unmodified, accuracy and completeness of data. Thirdly, availability ensures data are always accessible. Fourthly, non-repudiation proves the occurrence of data by the recipient. Next, authenticity is data originality. Finally, the reliability is the ability to provide consistent intended behaviour or results.

Table 4.1 List of factors from cloud security research

Security Goals	**Characteristics**	**Sources**
Confidentiality	Data is not made available or disclosed to unauthorised individuals, entities, or processes	(Catteddu and Hogben 2009, CSA 2013b, NIST 2013a, ASD 2014b, CPNI 2014a, ISECT 2014a, ISO/IEC 2016)
Integrity	Data is accurate and complete	(Catteddu and Hogben 2009, CSA 2013b, NIST 2013a, ASD 2014b, CPNI 2014a, ISECT 2014a, ISO/IEC 2016)
Availability	Data being accessible and usable upon demand by an authorised entity	(Catteddu and Hogben 2009, CSA 2013b, NIST 2013a, ASD 2014b, CPNI 2014a, ISECT 2014a, ISO/IEC 2016)
Non-repudiation	Ability to prove the occurrence of a claimed event or action and its originating entities	(CSA 2013b, ISECT 2014a, ISO/IEC 2016)
Authenticity	Data is original of what it claims to be	(Catteddu and Hogben 2009, CSA 2013b, ASD 2014b, CPNI 2014a, ISECT 2014a, ISO/IEC 2016)
Reliability	Ability to provide consistent intended behaviour and results	(Catteddu and Hogben 2009, ASD 2014b, CPNI 2014a, ISECT 2014a, ISO/IEC 2016)
6 factors	6 items	

Table 4.2 shows goal-driven security factors and potential security control items from secure cloud storage research consisting of seven factors and 14 items. The first factor, ensuring confidentiality refers to the identification of cloud storage user and authorisation to access data. Secondly, integrity checks on remote data accurately preserve ownership of data and encryption of data. Maintaining availability represents accessibility to the data and up-to-date available data. Next, guaranteeing non-repudiation to data involves accurate time-stamping of accessed data and assurance with user signature. Preserving authenticity is verifying data based on authentication and synchronising data in the storage. Finally, the reliability of service provider that represents consistency and validity of cloud service.

Table 4.2 List of factors from secure cloud storage research

Security Factors	Items	Sources
Cloud Storage Security	Security policy, security procedure	(Firesmith 2004, Brock and Goscinski 2010, Takabi et al. 2010, Zissis and Lekkas 2012, Mapp et al. 2014)
Ensuring Confidentiality	Identification of cloud storage user, authorisation to access data	(Bessani et al. 2011, Kamara et al. 2011, Popa et al. 2011, Stefanov and Dijk 2012, Vrable et al. 2012, Zhou et al. 2013, Yao et al. 2013, Mapp et al. 2014, Tawalbeh et al. 2015, Vu et al. 2015, El-Booz et al. 2016)
Integrity Checks on Remote Data	Accurate ownership of data, encryption of data	(AlZain et al. 2011, Bessani et al. 2011, Kamara et al. 2011, Mahajan et al. 2011, Nepal et al. 2011, Vrable et al. 2012, Mu et al. 2012, Stefanov and Dijk 2012, Wang et al. 2013, Zhou et al. 2013, Mapp et al. 2014, Vu et al. 2015, Tawalbeh et al. 2015, El-Booz et al. 2016)
Maintaining Availability	Accessible to the data, up-to-date available data	(AlZain et al. 2011, Bessani et al. 2011, Mahajan et al. 2011, Popa et al. 2011, Mu et al. 2012, Stefanov and Dijk 2012, Mapp et al. 2014, Vu et al. 2015)
Guaranteeing Non-repudiation to Data	Accurate time-stamping of accessed data, assurance with user signature	(Wang et al. 2013, Tawalbeh et al. 2015, El-Booz et al. 2016)
Preserving Authenticity	Verified data based on authentication, synchronised data in the storage	(Yao et al. 2013, Mapp et al. 2014, Vu et al. 2015, El-Booz et al. 2016)
Reliability of Service Provider	Consistency of cloud service, valid service	(Bessani et al. 2011, Mahajan et al. 2011, Stefanov and Dijk 2012, Mapp et al. 2014, Tawalbeh et al. 2015, Vu et al. 2015)
7 factors	14 potential items	

4.3 Group Factors with Similar Meaning

This section attempts to merge two research areas using the two unique factor lists from Table 4.1 and Table 4.2:

i. Unique factor list from cloud security; and
ii. Unique factor list from cloud storage

An initial alignment was conducted to discover any patterns between the factors. In Table 4.1, column 1 shows the list of factors from cloud security. On the other hand, Table 4.2, column 1 shows the list of factors for cloud storage security. The result of this step produced a set of two lists containing the summary of findings from cloud security research and; cloud storage research.

Cloud storage security implementation that represents the security policies and procedures was placed as a unique factor of its own, as it will be used to show the policy implementation. A look at the lists in Table 4.1 and Table 4.2 shows some similarities in the meanings even though for cloud security the meaning is more general in the context of cloud compared to in cloud storage, the defined meaning is focussing on cloud and storage.

For an example, the term confidentiality (row 2, Table 4.1), ensuring confidentiality (row 3, Table 4.2) represent the meaning of security measure that relates to protecting the data accessed in the cloud. Similarly, the terms integrity (row 3, Table 4.1) and integrity checks on remote data (row 4, Table 4.2) represent security measure to protect modifications of data stored. These two factors were reclassified as confidentiality and integrity factor respectively for simplicity. Next, the remaining factors and their characteristics/elements will be analysed based on their relevance towards security in cloud storage.

Table 4.3 Initial alignment of factors

1. Security Goals from Generic Cloud Research	2. Security Factors from Cloud Storage
	Cloud Storage Security
Confidentiality	Ensuring Confidentiality
Integrity	Integrity Checks on Remote Data
Availability	Maintaining Availability
Non-repudiation	Guaranteeing Non-repudiation to Data
Authenticity	Preserving Authenticity
Reliability	Reliability of Service Provider

An initial analysis identified some factors that have a direct mapping between the two domain research areas. More importantly, however, the inclusion or exclusion of a factor depends on its relevance to cloud storage. These seven factors have shown supporting evidence towards cloud storage security; but they were identified as more of a reusable security factor and/requirement goal, which all information systems must have. The factors were examined and synthesised by determining whether each of the factors falls into any of these three categories:

i. Unique factor, a factor that can only be found either in column 1: cloud security factors or column 2: cloud storage factors.
ii. Direct mapping factor, a factor that can be found in both columns 1: cloud security factors and column 2: cloud storage factors.
iii. Combination factor, two or more factors were combined to best represent the meaning of the factor.

Following the criteria listed above, one factor was found to be a unique factor and six factors are direct mapping and combination factor as shown in Table 4.3.

4.4 The Proposed Cloud Storage Security Framework

The conclusion from our empirical research, supporting the proposed framework, is that the security framework in cloud storage systems consists of seven security factors and fourteen items as shown in Table 4.4. The factors are security in cloud storage, confidentiality, integrity, availability, non-repudiation, authenticity, and reliability. The concerns are data breach, data leakage and loss, insecure APIs, account hijacking, denial of service, malicious insiders, abuse of cloud service, inadequate cloud planning, cloud-related malware, closure of cloud service, natural disaster, hardware failure, shared technology vulnerabilities, and insufficient due diligence.

The summary is presented in Table 4.4. The table shows the proposed factors and items for cloud storage security framework supported by literature synthesis and industry best practices.

Table 4.4 The Proposed Factors and Items for Cloud Storage Security Framework

Factor	Item
1. **Cloud Storage Security**	1. Cloud Storage Security Policies
	2. Cloud Storage Security Procedures
2. **Confidentiality**	3. Identification of cloud storage user
	4. Authorisation to access data
3. **Integrity**	5. Accurate ownership of data
	6. Encryption of data
4. **Availability**	7. Accessible to the data
	8. Up-to-date available data
5. **Non-repudiation**	9. Accurate time-stamping of accessed data
	10. Assurance with user signature
6. **Authenticity**	11. Verified data based on authentication
	12. Synchronised data in the storage
7. **Reliability**	13. Consistency of cloud service
	14. Valid service
Factors: 7	Items: 14

4.5 Chapter Summary

This chapter presented the proposed security factors for cloud storage. The first step in identifying the security factors involved identifying and reviewing published papers specifically on security frameworks concerned with cloud computing and cloud storage. All security principles, concepts or factors were listed and a process of selecting factors that are relevant in cloud computing was investigated; some factors were included or excluded in the context of cloud storage. The security factors were finalised by grouping and filtering the relevant factors and items that belong to each domain.

The next step involved the same process as in stage one but was examined for industry-accepted standards, guidelines and best practice documents related to cloud security. The controls were mapped to the factors found in stage one. The process started by gathering security factors from organisations such as CSA, NIST, ENISA, CPNI and ASD. This was used to take full benefit of the widely recognised security frameworks in the cloud. The conclusion from our empirical research, supporting the proposed framework, is that the security framework in cloud storage systems consists of seven security factors and fourteen items. The factors are security in cloud storage, confidentiality, integrity, availability, non-repudiation, authenticity, and reliability.

The next chapter confirms the framework that includes the seven factors in an exploratory study. An explanation of the methods is given in the next chapter.

Chapter 5: Confirming the Cloud Storage Security Framework (CSSF)

This chapter outlines the research methods applied in confirming the framework as shown in Figure 5.1. The first section briefly discusses the methods, qualitative and quantitative, with an explanation of the triangulation technique. The next section explains in detail the design of research methods applied in the study that includes interview and questionnaire designs, data collection process and instruments, the pilot test, sample size and analysis.

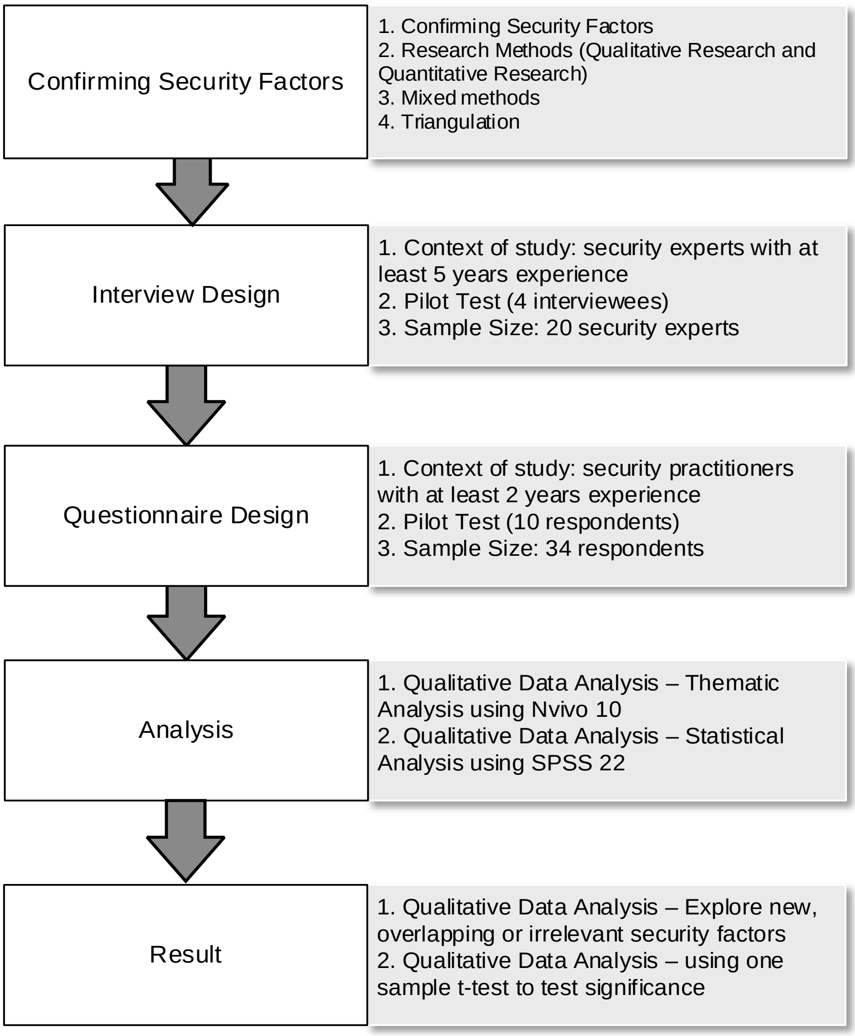

Confirming Security Factors
1. Confirming Security Factors
2. Research Methods (Qualitative Research and Quantitative Research)
3. Mixed methods
4. Triangulation
Interview Design
1. Context of study: security experts with at least 5 years experience
2. Pilot Test (4 interviewees)
3. Sample Size: 20 security experts
Questionnaire Design
1. Context of study: security practitioners with at least 2 years experience
2. Pilot Test (10 respondents)
3. Sample Size: 34 respondents
Analysis
1. Qualitative Data Analysis – Thematic Analysis using Nvivo 10
2. Qualitative Data Analysis – Statistical Analysis using SPSS 22
Result
1. Qualitative Data Analysis – Explore new, overlapping or irrelevant security factors
2. Qualitative Data Analysis – using one sample t-test to test significance

Figure 5.1 Steps involved in Confirming the Framework

5.1 Confirming Security Factors

Confirming a framework will look at how well the factors of the framework represent a concept or domain of content (Rubio et al. 2003, Huang et al. 2010). Confirmation of a framework becomes an important step especially when a new framework is being developed where there is no existing measure that operationalises the concept as the researcher intended (Rubio et al. 2003). A framework that defines security in cloud storage is currently explored and needs to be confirmed.

Using the proposed Cloud Storage Security Framework (CSSF) as a source of reference and guide, seven factors were selected for inclusion in measuring experts and practitioners agreement on the security factors of cloud storage. The seven selected factors are:

1. Cloud Storage Security (CS): The security policies implementation in an organisation that includes policy and procedures in ensuring security measures in cloud storage.
2. Confidentiality (Co): The confidentiality of data held in cloud storage from the stakeholder's viewpoint including identification of cloud storage user and authorisation to access data.
3. Integrity (In): The integrity of data stored in cloud storage from the stakeholder's viewpoint including accurate ownership of data and encryption of data.
4. Availability (Av): Extent of data availability in cloud storage from the stakeholder's viewpoint including accessible to the data and up-to-date available data
5. Non-repudiation (Nr): Level of the non-repudiation of data stored in cloud storage from the stakeholder's viewpoint including accurate time-stamping of accessed data and assurance with user signature.
6. Authenticity (At): The authenticity of data stored and accessed by authorised user in cloud storage from the stakeholder's viewpoint including verified data based on authentication and synchronised data in the storage.
7. Reliability (Re): The reliability of service provided by cloud storage from the stakeholder's viewpoint including consistent and valid cloud service.

5.2 Research Paradigms

Research paradigms play an important role in preparing a foundation for knowledge claims upon which research is based. Paradigms act as assumptions made by researchers as “how they will learn and what they will learn” during their research (Creswell 2003). Creswell and Clark (2011) discussed four paradigms in research:

1. Post-positivist: generally related to quantitative data, through theory verification, determination and critical realism
2. Constructivist: often associated with qualitative data, focused on theory creation and wider understanding
3. Participatory: shaped by political issues, more associated with qualitative approaches
4. Pragmatic: related to consequences of research and multiple methods of data collection

In conducting a research, the big sets of ideas have to be reduced to a small and discrete set of ideas such as those variables for hypothesis testing and research questions, numeric measures of observation has to be developed and theories have to be verified. Normally, researchers start with a theory followed by data collection and end up with either supporting or refuting the theory (Creswell 2003).

This study is based on an epistemology or knowledge claim of post-positivist as assumptions and questions are based on the review of the literature. There is also a discussion around meanings and thoughts, which makes this research post-positivist and not just positivist. The study focus to obtain objectives and generalisable result by using a valid and reliable instrument, a suitable sampling method and an appropriate statistical method to the collected data. Furthermore, the quantitative paradigm is employed in this research because it has been used by many researchers with regards to information system studies (Deluca et al. 2008).

5.3 Research Methodology for Confirming the CSSF

This section outlines the research approach, design and methods applied in this study. The first section briefly discusses the approaches, qualitative and quantitative, with an explanation of the triangulation technique. The next section explains in detail the design of research designs applied in the study that includes interview and survey, data collection process and instruments, the pilot test, sample size and analysis.

5.3.1 Research Approaches

This section provides a brief description of the research approaches deployed in this study to answer the research questions, including a literature review, semi-structured interviews and questionnaire survey. It explains the triangulation techniques used to confirm the proposed security framework.

5.3.2 Qualitative Research

Qualitative approaches are mainly exploratory and used to discover the perceptions of target audiences with reference to specific issues (Sekaran 2000). This will assist in understanding the situation by providing insights into problems for potential research. A qualitative approach is useful when secondary data, such as literature review, are insufficient to develop depth in a research (Fink 2003). The participants are usually few in number.

An interview is considered the most common data collection method in this research which can be undertaken through structured or unstructured questions or a mixture of both (Sekaran 2000). The interviews can be conducted face-to-face, by telephone and online, such as video conferencing (Rubin and Rubin 2005, Bolderston 2012).

5.3.3 Quantitative Research

Quantitative approaches are used to quantify numerical data into usable statistics by surveying a number of participants or simple measurements (Saunders et al. 2009). One of the methods for collecting quantitative data is obtaining answers from a set of relevant questions in a questionnaire. The data gathered are analysed by statistical techniques and results obtained are generalised to the population (Mertens 2010).

A Likert scale is a technique used to measure attitudes that yield reliable correlation between scores and case history (Likert 1932). It is commonly used in a questionnaire to capture the opinions of a subject (Saunders et al. 2009).

5.4 Mixed Methods

Mixed methods is a combination of qualitative and quantitative approaches to data gathering, analysis, interpretation, and presentation (Teddlie and Tashakkori 2010).

This research approach provides more choices, options, and approaches to consider (Creswell and Clark 2011).

One of the factors of any research process is the reliability of the data and findings (Cotten et al. 1999). Reliability involves consistency, dependability and replicability of the results obtained from a study. A mixed methods approach may improve the reliability for the research (Cotten et al. 1999, Mertens 2010, Teddlie and Tashakkori 2010).

Reliability can be achieved through consistency obtained from different procedures such as interviews and questionnaires. Different types of data are collected from different sources enhancing the reliability of the data and results. The purpose is not to attain the same result but rather to agree that, based on the data collection processes, the findings and results are consistent and dependable.

In this study, a mixed methods approach was chosen as different techniques were applied to collect both qualitative and quantitative data.

5.5 Triangulation

Triangulation is a technique which is defined as a combination of two or more methods in a study (Thurmond 2001). The technique can be applied to explore the issues extensively and improve the accuracy of research findings (Fink 2003). The idea is that if multiple sources of information or data produce similar results, the credibility of the findings is enhanced (Fink 2003). The findings from each method can be compared to find similar conclusions (Guion et al. 2011).

In this study, methodological triangulation was applied. Data were collected from two different methods, interviews and questionnaire methods. Subsequently, the results obtained were compared to identify similar decision patterns (Golafshani 2003). By using triangulation and drawing on multiple viewpoints, the researchers is confident that the research is moving towards accuracy and credibility as it devolve into several sources, confirmation, and processes of data collection. Besides, a researcher can overcome the weakness of any single method and any bias inherent in each source of data (Creswell 2003, Creswell and Clark 2011).

The triangulation ensured that the process of refining and verifying the security framework was thoroughly explored. It involved a synthesis process of assessing, combining, and comparing data discovered from a detailed literature review of security frameworks, industry-accepted standards, technical white papers, expert reviews and practitioner surveys as shown in Figure 5.2.

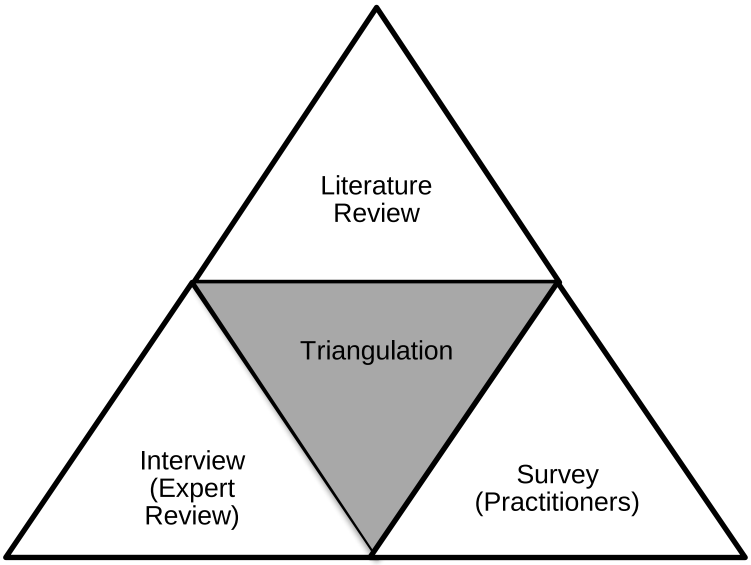


Figure 5.2 Triangulation technique used to confirm the framework

5.6 Research Design

This section describes the two research design utilised in this study to facilitate the confirmation of the findings and to discover possible potential factors for the security framework.

The interview was designed using qualitative approach techniques to confirm the proposed factors and explore more potential factors and their items. Interviews were carried out with selected IT security experts.

By contrast, the questionnaire was chosen so as to make generalisations from a sample of the whole population. The questionnaire was designed using quantitative techniques to collect data from IT practitioners, defined as employees that have worked in an IT environment, to confirm the proposed security quality factors.

5.6.1 Interview Design

A qualitative interview is carried out to explore the knowledge, opinions and values of individuals or groups who are experts in a particular field of knowledge (Fink 2003).

Semi-structured interviews were utilised to collect data from a group of people. This kind of interview was selected due to its advantage of gathering statements regarding the individuals' attitudes and exploring in-depth their experience (Bolderston 2012). It was used to capture their experience regarding:

- Aspects of security factors
- Concerns associated with cloud storage systems
- Current issues of security in cloud storage

The interview questions were developed in the English language. The detailed questions are provided in Appendix A.1. The following subsections briefly describe how the semi-structured interviews were designed.

5.6.1.1 Context of study and the participants

This study was conducted in June 2015 within various information systems roles: IT/Technical (Application Security, Digital Forensics Investigation, and Threat Intelligence), Consultant/Advisory (Consultant, Industry Research/Analyst), Security Analyst/Expert (Information Security), and Security Policy Maker (CIO, CTO, CSO, and Chief Security Information Officer). The experts involved were mainly from industry, government, healthcare, and oil and gas. All of the experts have at least five years' experience in ICT Security.

5.6.1.2 Data collection process and instrument

The interviews with the experts were scheduled over two weeks. The interview had three sections. The first section was demographic information with closed questions. The next section discussed each factor of the study and consisted of 2-3 open-ended questions. The last section explored new factors (if any). The interview featured confirmatory and exploratory questions about the factors making up the metric. Several interviewees are approached at European Information Security Expo, London on 4 June 2015.

A Livescribe2 pen and notebook were used as a tool for recording the interviews. Interviews that were conducted via videoconferencing or the Internet use the Skype application and audio was recorded using the QuickTime application. Before any recording was made, the interviewer sought permission from every interviewee.

5.6.1.3 Pilot Test

A pilot test is an opportunity to try out an instrument well before it is made final. A pilot test will simulate the use of the instrument in its intended setting (Fink 2003).

The pilot session to test the interview questions was carried out with four people; two were IT security experts from Cyber Security Malaysia, and two were computer science researchers from the University of Southampton. The interviewees were asked about the security factors. After the pilot session, it was concluded that discussing each of the security factors individually was better than showing a detailed diagram of the proposed security framework.

5.6.1.4 Determination of sample size

Qualitative studies usually depend on non-probability sampling where participants are chosen based on non-random criteria (Bhattachejee 2012). In expert sampling, participants are chosen based on their knowledge of the area being studied (Bhattachejee 2012). The size of a sample depends on saturation being reached when no new knowledge can be collected (Guest et al. 2006). Saturation is often achieved around twelve interviews (Guest et al. 2006).

5.6.1.5 Qualitative data analysis

Thematic analysis was used to analyse, identify and report the themes within raw data. The themes reflect patterns that exist within the collected data, and the patterns describing the phenomenon. Therefore, it is a method of organising and describing a corpus in a way that helps researchers capture important things to describe their research questions (Braun and Clarke 2006).

The interview questions revolve around security quality factors and their items. Therefore, themes were factors, and the sub-themes address any related items. Nvivo 10 software was used in the qualitative data analysis to split the raw data into themes. Each factor was given a node, each node had its characteristics and its items clustered into “confirmed”, “irrelevant”, “additional” and “overlapped”. The next step was to code and assign data from the transcript to related codes.

Besides, there was a node for the severity factors to which every statement concerning severity factors linked. That is any statement that described the security concerns state severity was coded and assigned to the node of severity factors. The next section will explain the questionnaire design including participant context, data collection and instruments, sample size, reliability test and quantitative data analysis.

5.6.2 Questionnaire design

A survey was chosen to collect information to capture knowledge, attitudes and behaviours. Questionnaires are a data collection tool in which participants are requested to answer various predetermined questions. The detailed questions in the questionnaire are provided in Appendix A.6. There are two types of questionnaire: self-administered and interview-administered. A self-administered survey is one in which the respondents take responsibility for reading and answering the questions, while the communication medium used in interview-administered surveys is either a personal or a telephone interview (Zikmund et al. 2012). The following subsections briefly describe how the questionnaires were designed.

5.6.2.1 Context of the study and the participants

This study was conducted with IT security practitioners, particularly IT/Technical people who have been involved in implementing, applying, or involved with, security measures in cloud systems. The questionnaire was distributed in July 2015. The practitioners involved were primarily from the industry and government sectors. All of the experts had at least two years' experience in IT Security.

5.6.2.2 Data collection process and instrument

The communication medium for data collection was self-administered via the University of Southampton's online application called iSurvey. Subsequently, the questionnaire was distributed to IT security practitioners to capture their opinions of the factors identified and ability to assess the security quality of protecting data properly. The survey was also distributed to participants and speakers at the Information Security Summer School, in Bilbao on 10 July 2015, as well as the Science and Information Conference, in London on 28 July 2015.

The questionnaire was divided into three sections. The first collected demographic data about the participant. The second section measured the practitioners' opinions on the importance of the proposed factors and items. The questionnaire featured five identified determinants on a five-point Likert scale with the following ratings: 'strongly agree' (=1); 'agree' (=2); 'neutral' (=3); 'disagree' (=4) and 'strongly disagree' (=5). The last section checked whether there was any ambiguity in the survey and whether any security factor needed to be added.

5.6.2.3 Pilot survey

A pilot survey was conducted to see if the respondents understood the directions for completing the questionnaire and each of the questions. This included the wording of the questions and clarity on where to mark the responses. Usually, for a pilot, ten or more people are needed to test a questionnaire (Fink 2003). The questionnaire was tested by ten IT security practitioners drawn from the IT Division, Ministry of Higher Education, Malaysia, from Cyber Security Malaysia, and from computer science researchers at the Royal Holloway University of London, and the University of Southampton.

The pilot produced two results: one question had unclear instructions, and some words of the survey were not interpreted in the same way by all respondents. Overall, respondents followed the directions correctly and the whole range of responses was selected. An improvement was subsequently made to the questionnaire before it was distributed online to respondents. Data obtained from this test was also used to calculate the internal consistency using Cronbach's alpha coefficient, as explained in the next subsection.

5.6.2.4 Statistical reliability test

A statistical reliability test was conducted to see if each of the items from the questionnaire was understood similarly by each respondent. Cronbach's alpha coefficient is one of the several analyses that may be used to gauge the reliability and accuracy of research measurements (Cronbach and Shavelson 2004).

A Cronbach's alpha analysis is able to measure the internal consistency; higher values indicate a greater correlation between responses, with a maximum possible value of 1. An alpha value of 0.5 is acceptable to establish the reliability of the internal consistency of the questions in this study.

5.6.2.5 Determination of minimum sample size

Statistical power analysis was used to calculate the minimum sample size for this study. Table 5.1 shows the sample size needed. G*Power software was applied to facilitate the acquisition of the minimum sample size.

Table 5.1 Sample size according to G*Power software

Statistical Test	**Means: Difference from constant (one sample test)**
Tails	Two

Statistical Test	**Means: Difference from constant (one sample test)**
Effect size (d)	0.8
Error probability (α)	0.05
Power (1– β error probability)	0.8
Minimum sample size	15

The parameters identified to determine the minimum sample size are detailed below.

- Effect size, d

There are three effect sizes: small, medium and large (Cohen 1988). This effect size for this exploratory research is large (0.8).

- Type I error, α

The accepted value for this study is 0.05. The 0.05 value is the conventional value for alpha, which is the level of significance.

- Power, 1– β error probability

The accepted value for this study is 0.8. The 0.8 value is the conventional value for power.

5.6.2.6 Quantitative data analysis

The one-sample t-test is a statistical procedure to test whether the mean value of a distribution is significantly different from a constant. In this study, to accept a factor as a reliable factor, the mean rating of a representative question of this factor needs to be significantly lower than 2.5. The rationale behind choosing 2.5 is that this number falls on the ‘somewhat agree’ before the ‘neutral’ point on the five-point Likert scale. An alpha, α was set at 0.05 to conduct the hypothesis test for this statistical test.

The null hypothesis and alternative hypothesis were as follows:

H0: the mean value of a factor is not significantly different from 2.5
H1: the mean value of a factor is significantly lower than 2.5

The factor is accepted as a sound factor for this study if the null hypothesis was rejected (ρ<0.05).

5.7 Ethical Approval

The participants have to give their informed consent before taking part in a survey; they can formally agree to participate after being informed about the risks and benefits of participation, the terms of their participation and their rights as research subjects (Fink 2003). During the data collection process, a consent form with sufficient information (participation information) was given to the interviewees to sign to ensure their agreement to participate in the study in Appendix A.2 and Appendix A.3. Online respondents were also given printed information (participation information) on the first online page before they could start responding to the questionnaire in Appendix A.4.

The University of Southampton Ethics Committee approved the quantitative and qualitative methodologies conducted in this study. Ethics approval was granted under reference number 14962 on 15 May 2015 for both interviews and survey.

5.8 Analysis of Findings and Results of the CSSF Evaluation

This section presents the findings, results of security factors with regard to cloud storage. The first section presents the findings from expert interviews, while the second section describes the results of the survey of practitioners. A summary draws conclusions from both data collected.

5.8.1 Findings of the Interviews

This section presents the findings of the interviews with security experts. The data was collected using semi-structured interviews from 20 security experts in Malaysia and the United Kingdom. The aim of this task was to review the security factors identified by the literature review and to explore other factors. The result is divided into two sections: general demographic information (quantitative data) and findings of qualitative data.

5.8.1.1 Demographic Information

Initially, 25 experts were invited by email to participate in an interview. Only 22 of them responded and two later cancelled their participation. The interviews were then conducted with remaining 20 security experts from different organisations in Malaysia and the United Kingdom. All the interviewees had at least five years’ experience in IT security, so they had the ability to understand and explain the current security situations and trends. The interviews were carried out in June and July 2015. Most of the interviews were performed via videoconferencing on the Internet using Skype. The audio was recorded using the QuickTime recorder application. Some of the interviews were performed face-to-face and this audio was recorded using the Livescribe2 pen. Permission was sought from all the interviewees before any recording was made. Most of the interviewees agreed that their audio could be recorded for transcribing purposes.

Description of the IT experts interviewed in this study is presented in Table 5.2 and Table 5.3.

Table 5.2 Demographic information of experts

Variable		Frequency	%
Country	Malaysia	12	60.0
	UK	8	40.0
Domain	Academic	2	10.0
	Government	2	10.0
	Industry	14	70.0
	Others	2	10.0
Job	Consultant	4	20.0
	IT Technical	10	50.0
	Researcher	2	10.0
	Policy Maker	4	20.0
Experience	5 years	4	20.0
	6 – 10 years	16	80.0
Data Stored in the Cloud	0 – 25 %	6	30.0
	26 – 50%	3	15.0
	51 – 75%	10	50.0
	76 – 100%	1	5.0
Using Cloud Storage	Yes	20	100.0
	No	0	0.0
Type of Cloud Storage	Commercial	12	60.0
	Public Cloud Storage provided by your organisation	1	5.0
	Private Cloud Storage provided by your organisation	7	35.0

Table 5.3 Expert Interviewees involvement with the cloud

Domain	Job Description	Experts
Academic	Information Security Researcher	L, P
Government	Information Security Officer	E
	Cloud System Administrator	H
Industry	Information Security Consultant	A, J, Q, T
	Cloud Systems Implementer	B, S
	IT Technical	F, G, M, N, R
	Cloud System Administrator	I
	Information Security Manager	K, O
Others	Policy Maker	C
	IT Technical	D

5.8.1.2 Qualitative Data

The purpose of the expert interviews was to review the possible security factors identified from the literature and to identify whether there were any further factors. The experts were given six closed questions before starting the interview regarding their domain, and job role, the length of experience in cloud and security and whether they stored their own data in the cloud. The result is shown in Table 5.2. Then, the experts were shown a list of the factors and their opinions were asked on the importance of each of the security factors. The next enquiry was to identify whether there were any other factors not mentioned in the study. The experts were also asked about security concerns happening in the cloud and potential risks of the concerns. Their opinions were analysed and coded to produce the following findings.

5.8.1.3 Security factors

The interviewees answer the closed questions and proceeds to share the opinion on security in the cloud. Some of the interesting quotes are:

Expert O: *"An increasing number of organisations recognise that developing a security framework is more important than designing protections because paying attention to security factors in the early stage potentially saves millions of dollars."*

Expert S: *"You can't protect what you don't specify."*

The interviewees were asked to indicate if they thought the factors were important based on the list of the proposed factors. There was a consensus among the interviewees that all the proposed factors were important: confidentiality, integrity, availability, non-repudiation, authenticity, and reliability.

Overall: Experts A, E, L, O, P, and S mentioned that security policy implementation is most important and confidentiality, integrity and availability are most basic and common security factors.

The interviewees have given a review on the security factors as below:

5.8.1.3.1 Security in Cloud Storage

All agreed that security implementation in cloud storage must be initiated by having a well-defined security policy in place. Below are some of the supporting statements by the experts:

Expert S: *“It is important to have a defined security policy and process”*

Expert A: “A security policy will inform users, staff, and managers, specify mechanisms for security and provide a baseline”

Expert P also mentioned that in some organisation when going for certification and compliance to standards that policy implementation is compulsory. He added: *“Policy is the best compliance tool – legislatively”*

5.8.1.3.2 Confidentiality

All agreed that confidentiality is an important security factor.

Experts F, I, J, R, and T mentioned that identification, authentication, authorisation, access control and encryption are practices and measures to prove confidentiality and integrity. This is shown by some of these quotes:

Expert J: *“Cloud storage should only show documents to authorised users. It should be encrypted whenever necessary, encrypts data in transit and multiple times at rest.”*

Expert R: *"Systems require passwords as part of access control."*

Expert T: *“Most have Single-Sign-On (SSO) in place. Many are considering multi-factor authentication – typically integrated into SSO solution.”*

5.8.1.3.3 Integrity

Eighteen of the twenty agreed that integrity is an important security factor.

Experts N and Q mentioned that integrity involves ensuring data are not leaked or lost in the cloud. Without integrity, data that is received or sent cannot be trusted.

Expert N: “*In many cases, good people are doing bad things by circumnavigating IT controls that have made their systems unusable. They might email a confidential file to their personal email account so they can work on it at home*. *This breaches integrity and leaks information.*”

5.8.1.3.4 Availability

All agreed that availability is an important security factor.

Expert J mentioned that although availability is more like a functional factor of a system, it is important as it ensures the data can always be accessed.

Expert P mentioned that availability is the most important factor, as the user will always expect the system to be accessible, regardless of any concerns.

“Cloud is becoming the only real choice when faced with the pressure of finance, availability, security, ease-of-use, and scalability” (Expert T).

5.8.1.3.5 Non-repudiation

Fourteen of the twenty agreed that non-repudiation is important.

Expert H explained that non-repudiation is an assurance that any party cannot deny sending or receiving the data. This includes obligations for contracts, standards, etc.

Expert P: “*Non-repudiation can reduce fraud and promote the legal enforceability of electronic agreements and transactions.”*

5.8.1.3.6 Authenticity

Fifteen of the twenty agreed that authenticity is an important factor.

Expert M revealed that currently IT systems are also reviewed, based on their quality of delivering authentic data. This involves verifying the source as genuine.

Expert R: *“Solid authentication defends a system against the security risk of impersonation, in which a sender or receiver uses a false identity to access a system. Digital certificates can provide a more secure method of authentication while offering other security benefits as well.”*

5.8.1.3.7 Reliability

All of the experts mentioned that reliability is an important factor.

Experts G and T mentioned that reliability is the quality of measuring whether the system is consistent and valid.

Expert T: *“Multitenant Cloud allows customers to gain huge advantages in innovation by the Cloud provider at a rate of change that traditional IT systems simply can’t hope to provide. Features that may have been requested by one customer are delivered to* *all customers using the platform. Fixes are applied across all customers equally, ensuring consistency in security and reliability for all users."*

5.8.1.4 Additional security factors

The interviewees were asked to suggest other security factors. The answers given were:

5.8.1.4.1 Accountability

Expert A mentioned that accountability is provided on trust basis by having a contract or SLAs with a clear and concise definition of security policy.

Expert K emphasises that accountability and trust is an important factor that must be fulfilled by providers (internal or external). They have made it compulsory for Security Level Agreement in many organisations as an extension to Service Level Agreements (SLAs).

Expert Q mentioned that:

"Accountability involves the processes, policies, and controls necessary to trace actions to their source. These develop trust among systems."

5.8.1.4.2 Auditability

Expert C specified that organisations are providing security policies to ensure work tasks follow guidelines and best practices.

Expert E stated that IT security standards are being implemented to gain confidence. The system must enable assessment, examination and audit to be completed smoothly.

Another important point stated by Expert T:

"A cloud storage should provide full access logs, allowing organisations to see how data is being accessed, shared and used in real time. This audit data is available through APIs to real-time systems allowing organisations to respond to data governance issues and provide full audit logs where required."

The experts explicitly mentioned that there is a difference between non repudiation, accountability and auditability. Non repudiation looks specifically into security measures enabling assurance that a party cannot deny accessing and storing the data in cloud storage. This includes binding keys to the data owners. Accountability involves trust and responsibility of the provider as a whole such as the readiness of Service Level Agreement (SLAs). SLAs will ensure conformance of responsibility among internal and external parties. On the other hand, auditability is define by security assessment such as security audits for cloud storage system. Security audits produce audit assertions, audit logs, reports etc.

5.8.2 Concerns in Cloud Storage

Experts were asked about the following concerns happening in the cloud and to rate their potential risk, according to their experience and knowledge. These are some of the highlights:

Expert T: "*In my experience, seeing many customers every year, their threats are remarkably consistent. While external threats exist, the primary threat to customers is not being in control of their data. By this I mean they do not know where it is, how it is protected, what is shared externally and how externally shared data is being accessed. This leads to data leakage and loss.* "

Expert P: "*When a corporate doesn't know where their data is, they can't protect it. It may not be backed up, it is likely to have no security measures around it, and it may not be meeting corporate data retention policies and may considerably expose the customer to data governance fines if they are in a regulated industry. Important threats should be assessed to imply security measures*."

Expert Q: "*Account hijacking via social hacking. An employee is persuaded to grant access to a system or share information externally. This is not to say that traditional hacking of systems doesn't exist however, IT systems are continuously evolving to meet these threats and social hacking is a far more successful mechanism when data is protected well by traditional IT methods.*"

All of the experts agree that the concerns are important. The concerns are listed below:

- Data breach
- Data leakage and loss
- Insecure APIs
- Malicious insiders

- Denial of service
- Account hijacking
- Abuse of cloud service
- Shared technology vulnerabilities
- Hardware failure
- Cloud-related malware
- Inadequate cloud planning

The interviewees have also rejected the concerns from natural disaster, hardware failure, insufficient due diligence and closure of cloud service. The supporting statements given are:

Expert J: *"Organisations are realising that storing their data in a commercial cloud storage is more secure than on-premise. While this may seem counterintuitive at first glance – many people initially see Cloud as riskier – when educated about the security infrastructure that it has in place, customers quickly realise that it is safer than their own servers. The data centre is equipped with disaster recovery and hardware is always configured with high availability."*

Expert L: *"Security accreditations is included in many data centre. These include ISO27001, HIPAA HITECH, Safe Harbor, PCI DSS 2.0, SSAE16 Type II, SOC1 & SOC2 etc. These are measures to eliminate many concerns."*

5.8.3 Results of the Questionnaires

This section provides the result of the survey. The quantitative data was collected in July and August 2015 using an online questionnaire. Initially, it was distributed to 40 respondents, only 36 of whom responded. Two abandoned the survey without saving it. Overall, 34 were taken as the sample. All of the respondents are currently working in an information or computer technology environment and have at least two years' experience in information security. The aim of the survey was to confirm the updated factors and concerns obtained from the interviews. The results of the survey are divided into three sections. The first section relates to demographic information, while the second and third sections present closed questions on nine factors and eleven concerns in the cloud.

5.8.3.1 Demographic information

Demographic data were collected to determine the respondent's eligibility for the study. Only respondents with at least two years' experience of working in information security were considered. The job role of the respondents varied: IT technical, IT security consultant, researcher, and IT security policy maker. The majority of respondents were working in the IT industry (41.2%), whereas 35.5% were academic researchers in IT security and 23.5% were government IT employees. Most of the respondents were either working in IT Technical (application security, digital forensics, threat intelligence etc.), or security researchers showing more than five years' experience in this domain. All the respondents used cloud storage, especially commercial provisions (such as Dropbox, Box), and other private cloud storage provided by their organisations.

Table 5.4 Demographic information of practitioners

Variable		**Frequency**	**%**
Country	Malaysia	20	58.8
	UK	14	41.2
Domain	Academic	12	35.3
	Government	8	23.5
	Industry	14	41.2
Job	Consultant	9	26.5
	IT Technical	11	32.4
	Policy Maker	3	8.8
	Researcher	11	32.4
Experience	2 – 5 years	8	23.5
	6 – 10 years	25	73.5
	More than 10 years	1	2.9
Data Stored in the Cloud	0 – 25 %	10	29.4
	26 – 50%	6	17.6
	51 – 75%	16	47.1
	76 – 100%	2	5.9
Using Cloud Storage	Yes	34	100.0
	No	0	0.0
Type of Cloud Storage	Commercial	18	52.9
	Public Cloud Storage provided by your organisation	2	5.9
	Private Cloud Storage provided by your organisation	12	35.3
	Others	2	5.9

5.8.3.2 Security factors

The second section of the questionnaire was collected to gather opinions from practitioners on factors obtained in the expert review. This section has 18 questions that covered the nine factors. The responses to these questions were based on a five-point Likert scale, with 1 denoting 'strongly agree', 2 denoting 'agree', 3 denoting 'neutral', 4 denoting 'disagree', and 5 denoting 'strongly disagree'. SPSS was used to analyse the data. The hypothesis was tested for each factor using a one sample t-test with a test value of 2.5. The justification of choosing 2.5 is that this number falls on the 'somewhat agree' before the 'neutral' point on the five-point Likert scale. The proposed factors are considered to affect the security of cloud storage if they have a mean value of less than 2.5.

The factors were considered to be statistically significant if the p-value is less than 0.0028. Bonferroni correction was used for controlling false positive results by dividing alpha (α) by the number of items (*n*) included in the questions = (α/*n*), (0.05/18) = 0.0028. Table 5.5 shows the result of the analysis. This table clearly shows that all the proposed factors are considered to have an effect on the security of a cloud storage as each had a mean value of less than 2.5.

Table 5.5 Analysis of security factors using one sample t-test[a]

Factor		**Mean**	**t**	**Sig. (2-tailed)**
Cloud Storage Security	Security policies are initiated to protect data in cloud storage	1.76	-4.098	<0.001**
	Procedures must be in place to ensure all security measures are followed	1.62	-5.393	<0.001**
Confidentiality	Sensitive data must not reach the wrong person	1.65	-6.426	<0.001**
	Access must be restricted to those authorised to view the data	1.59	-7.152	<0.001**
Integrity	Data must not be changed or altered by unauthorised people	1.79	-4.504	<0.001**
	Data is hidden from those that are not supposed to see it	1.76	-4.098	<0.001**
Availability	A functioning system environment must be correctly maintained	1.62	-5.393	<0.001**
	Keeping up with the latest necessary system upgrades	1.76	-4.217	<0.001**
Non-	Data correctly reflects the object	1.94	-3.545	<0.002**

Factor		Mean	t	Sig. (2-tailed)
repudiation	Owner of an account must not allow other users to use his/her account	1.82	-4.537	<0.001**
Authenticity	A user whose authentication request is approved becomes authorised to access the accounts of that account holder	1.79	-4.504	<0.001**
	Data across the system should be in synch with each other	1.68	-5.698	<0.001**
Reliability	Critical components or functions of a system are duplicated to increase reliability of the system	1.88	-4.265	<0.001**
	Proof of the integrity and origin of data must be provided	1.85	-5.386	<0.001**
Accountability	The data source is trustworthy	1.88	-4.095	<0.001**
	Accountable against data loss or interruptions	1.74	-4.400	<0.001**
Auditability	A source must be able to provide proof of identity	1.82	-5.206	<0.001**
	Data are protected with policies by accredited bodies	1.88	-4.682	<0.001**

[a] *df* =33
** p<0.0028

All the security factors were found to be statistically significant as all the p-values are <0.0028. The overall reliability test of security factors, Cronbach's alpha, is 0.919, indicating that the results are reliable. Table 5.6 shows the results of each factors alpha value.

Table 5.6 Reliability Statistics Test of security factors

Factors	Number of Items	Cronbach's alpha Value
Cloud Storage Security	2	0.840
Confidentiality	2	0.720
Integrity	2	0.767
Availability	2	0.870
Non-repudiation	2	0.759
Authenticity	2	0.896
Reliability	2	0.878
Accountability	2	0.830
Auditability	2	0.980

A Shapiro-Wilk test (p>0.05) (Shapiro and Wilk 1965, Razali and Wah 2011) and a visual inspection of their histograms, normal Q-Q plots, and box plots, showed that the security factors' scores were approximately normally distributed for the respondents.

For respondents data, the assumptions are approximately normally distributed in terms of skewness and kurtosis (Cramer 1998, Cramer and Howitt 2004, Doane and Seward 2011).

5.9 Discussion of Findings and Results

The expert review confirmed the proposed factors as important and identified two additional factors. The factors are confirmed in the survey. The following section discusses the findings from both expert review and results from the survey.

5.9.1 Discussion of Expert Review Findings

A thematic analysis was carried out to examine themes within the interview results. According to the theme coding, the proposed factors are considered important in affecting the security of cloud storage.

Experts mentioned that security is the priority in an organisation. Organisations meet this goal by striving to accomplish the following factors. Confidentiality, Integrity and Availability (CIA) are known as the underlying security factors in information systems. This is consistent with recommendations from industry standards (Catteddu and Hogben 2009, CSA 2013c, NIST 2013b, ASD 2014b, CPNI 2014b). It was also supported by all the experts.

Few of the experts have mentioned that Integrity and Authenticity combine to produce what is known as non-repudiation. One expert considered Non-repudiation and Authenticity as one but other experts said it addressed different issues. This finding is interesting as it is consistent with existing studies on information security objectives and practices conducted with certified information security professionals (Ma et al. 2008). One of the possible reasons behind this issue is the need for a trusted method e.g., creating digital signature that has led to non-repudiation becoming a new security objective. Another possible explanation is the interpretation of terms between experts. This relation could be investigated in future.

The majority of experts agreed with the factors and there was no strong reason to remove any of them. Two additional factors were determined by synthesising the expert's suggestions. These factors are Accountability and Auditability.

Accountability was given emphasis as an important factor by two experts and three more have stated clearly the importance of SLAs, contracts etc. as a medium to build trust between provider and user. This finding is in agreement with a study on accountability and trust in cloud computing (Ko et al. 2011b, Pearson et al. 2012, Na et al. 2013) that highlighted that a cloud system should provide continuous auditing, logging, and monitoring.

Auditability was also highlighted as an important factor by four of the experts. The experts specified that organisations are providing policies to comply with international IT security standards such as ISO27001. This has been supported by existing studies on auditing the integrity of data stored in the cloud (Wang et al. 2011, Singh et al. 2012) which is consistent with the experts opinion that cloud storage should provide audit records to allow users monitor data being accessed, shared and used in real time.

There were other factors mentioned but rejected by the researcher, as they were redundant with the proposed factors. The rejected factors are:

(a) Assurance overlaps with Non-repudiation
(b) Responsibility and Trust overlaps with Accountability
(c) Verification overlaps with Authenticity

Initially, there were fourteen concerns based on existing studies, industry reports and white papers. Discussions with the experts discovered if there were any of the concerns happening in their organisation and its potential risk. A majority of the experts did not agree on the natural disaster, hardware failure, insufficient due diligence, and closure of cloud service in the context of cloud storage. The experts mentioned that these four concerns rarely happen. They have also mentioned, most data centres are equipped with disaster recovery and hardware is always configured with high availability. This is supported by a study that was conducted to identify security issues in SaaS (Subashini and Kavitha 2011). One explanation behind this is cloud storage delivers its service as SaaS, the user has to depend on CSP for proper security measures. CSPs are expected to provide restrictions to keep multiple users' from seeing each other's data.

It has been decided that hardware failure is an important concern as it is reported as the third top concerns in the cloud (CSA 2013a). Only eleven concerns were considered important affecting the security of a cloud storage. Three concerns (natural disaster, insufficient due diligence, and closure of cloud service) were removed before it was asked in the questionnaire following expert's recommendations.

5.9.2 Discussion of Questionnaire Results

All the factors proposed, based on existing studies and suggested in the expert review, were deemed statistically significant. Confidentiality and Availability received the strongest consensus. This shows that although security protections are important, the availability of service and accessibility of data in the cloud is considered important too.

This is in agreement with existing studies that mentioned availability as an important factor (Firesmith 2004, Brock and Goscinski 2010, Takabi et al. 2010, Zissis and Lekkas 2012, Mapp et al. 2014). One of the arguments was on service outages that cause disruption to access to data and revenue generated by storing data. This statement is supported by a study conducted on data availability (Mouratidis et al. 2013). CSPs provide on request and reliable service with highest uptimes. If an organisation's data gets locked-in, service disruption could pose potential financial damage to the organisation and its clients.

The confirmed framework is shown in Table 5.7 showing nine factors and 18 items supported by security experts and security practitioners.

Table 5.7 Confirmed factors for Cloud Storage Security Framework

Factor	Item
1. **Cloud Storage Security**	1. Cloud Storage Security Policies
	2. Cloud Storage Security Procedures
2. **Confidentiality**	3. Identification of cloud storage user
	4. Authorisation to access data
3. **Integrity**	5. Accurate ownership of data
	6. Encryption of data
4. **Availability**	7. Accessible to the data
	8. Up-to-date data available
5. **Non-repudiation**	9. Accurate time-stamping of accessed data
	10. Assurance with user signature
6. **Authenticity**	11. Verified data based on authentication
	12. Synchronised data in the storage
7. **Reliability**	13. Consistency of cloud service
	14. Valid service
8. **Accountability**	15. Trustworthiness of cloud storage services
	16. Accountability of providers and users
9. **Auditability**	17. Proof of data stored in the cloud storage
	18. Accredited policies implied by the cloud storage provider
Factors: 9	Items: 18

5.10 Chapter Summary

A mixed method was used to confirm the proposed security framework for cloud storage. It consisted of qualitative and quantitative approaches. A methodological triangulation technique was chosen to explore the factors extensively and improve the reliability and accuracy of research findings. Semi-structured interviews were conducted with experts, which included open-ended and closed questions. The objective of the interviewing was to confirm the proposed security framework and explore other factors. The questionnaire was used to gather data from IT security practitioners using closed questions with a five-point Likert scale. Both the interviews and survey are carried out for participants in Malaysia and the United Kingdom. The instruments underwent a pilot test to ensure their suitability in the intended setting. A statistical reliability test for the questionnaire was conducted using Cronbach’s alpha coefficient for this study.

This chapter also presented the findings from the interviews and results from the questionnaires. A semi-structured interview was conducted to review the factors and concerns identified earlier and other factors and concerns that were not mentioned in previous studies. The findings from interviews revealed that all the proposed factors are important. In addition, the factors of Accountability and Auditability affect the security of cloud storage. The result of the questionnaires supported the experts’ views. The discussions of results have explained the consensus of the factors and reasons of rejecting three other factors; responsibility, assurance, and verification as it is overlapping with other confirmed factors in this study. On the other hand, based on the interview findings, only ten concerns were rated important by the experts. The rejected concerns were the natural disaster, hardware failure, insufficient due diligence and closure of cloud service. The researcher has considered hardware failure to be rated by practitioners as it is reported as the second highest concerns in the Cloud Computing Vulnerabilities Incidents report. The survey results have shown that hardware failure remains as an unimportant concern. The results from questionnaires revealed that only five of the concerns are deemed as statistically significant (data loss and leakage, insecure APIs, account hijacking, shared technology vulnerabilities, and cloud-related malware).

Chapter 6: Development and Validation of the Security Rating Score (SecRaS) Instrument

In chapter 5, a study confirmed the Cloud Storage Security Framework (CSSF). The results support the factors of CSSF as theoretically sound and based on established research in cloud storage and cloud security. Chapter 6 presents the development and validation process of an instrument, the Security Rating Score (SecRaS), which can be used to measure the level of security in cloud storage. SecRaS is used to answer the second research question: How can stakeholder evaluate the level to which cloud storage security framework (CSSF) is being followed? SecRaS was developed through a stepwise approach using CSSF as a reference and guide. The development process is described in the first section, the approach and methodology in the second, after which a study was carried out to validate the instrument in the next section. Prior the validation study, the instrument underwent a pre-test. The pre-test was carried out to ensure the content validity. After pre-test, the validation study was undertaken to check the validity and reliability between each item in a factor and how they relate to the rating score (SecRaS) as a whole. The following section describes the data analysis and results. Lastly, the final section summarises the chapter.

6.1 Development of the Security Rating Score (SecRaS) Instrument

A reliable measuring instrument was developed and validated. The diagram in Figure 6.1 shows the instrument's development and validation process. The Security Rating Score (SecRaS) instrument was developed to measure the level to which the Cloud Storage Security Framework (CSSF) is being followed (Straub et al. 2004). Even though the instrument is exploratory, the methods implemented for instrument development and validation abide by conventional methodologies for instrument development (Fink 2003, Dwivedi et al. 2006). In this section, the factors selection and operational definitions will be described for each factor. The process also includes generating items for the instrument.

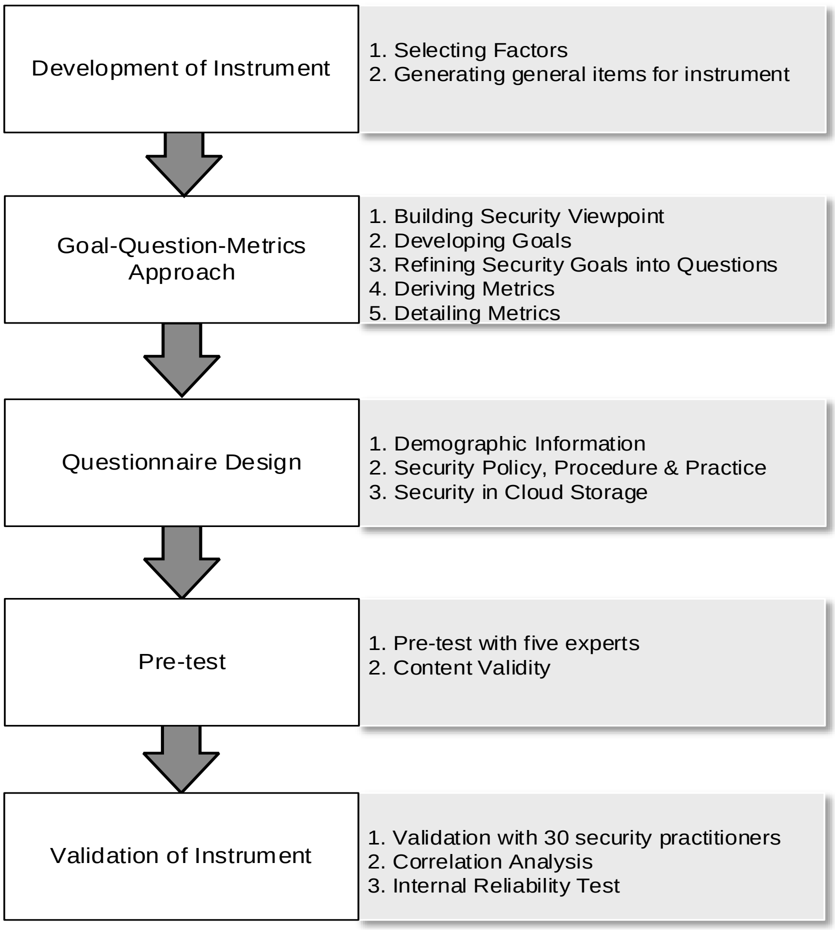

Development of Instrument
1. Selecting Factors
2. Generating general items for instrument
Goal-Question-Metrics Approach
1. Building Security Viewpoint
2. Developing Goals
3. Refining Security Goals into Questions
4. Deriving Metrics
5. Detailing Metrics
Questionnaire Design
1. Demographic Information
2. Security Policy, Procedure & Practice
3. Security in Cloud Storage
Pre-test
1. Pre-test with five experts
2. Content Validity
Validation of Instrument
1. Validation with 30 security practitioners
2. Correlation Analysis
3. Internal Reliability Test

Figure 6.1 The Security Rating Score (SecRaS) Development and Validation Process

6.1.1 Selecting factors and operational definition

Looking at the Cloud Storage Security Framework (CSSF) as a source of reference and guide, nine factors were selected for inclusion in measuring the level of security to protect data in cloud storage; (i) Security in Cloud Storage (ii) Confidentiality, (iii) Integrity, (iv) Availability, (v) Non-repudiation, (vi) Authenticity, (vii) Reliability, (viii) Accountability, and (ix) Auditability.

6.1.2 Generating general items for instrument

The following step is to generate items that represent those factors identified for inclusion in the instrument. The Security Rating Score (SecRaS) instrument was developed to measure the level of security in cloud storage. The following explains each goal-driven factor used to formulate the operational definition used in the instrument. The factors investigate the security of data and protection controls in a cloud storage system. The description for each of the nine factors is presented as follows:

1. Cloud Storage Security (CS): The security policies implementation in an organisation that includes policy, procedures and processes in ensuring security measures in cloud storage.
2. Confidentiality (Co): The confidentiality of data held in cloud storage from the stakeholder’s viewpoint including policy on identity management, access management, authorisation, secure APIs and securing access communication channel (s).
3. Integrity (In): The integrity of data stored in cloud storage from the stakeholder’s viewpoint including policy on encryption, key management, data ownership and stewardship, data deletion, and data protection for sensitive data.
4. Availability (Av): Extent of data availability in cloud storage from the stakeholder’s viewpoint including policy on accessibility, backup and restoring data stored in cloud storage.
5. Non-repudiation (Nr): Level of the non-repudiation of data stored in cloud storage from the stakeholder’s viewpoint including policy on time stamp, bind and validation, and geographical location as authentication factor.
6. Authenticity (At): The authenticity of data stored and accessed by authorised user in cloud storage from the stakeholder’s viewpoint including policy on cryptographic protection, anti-counterfeit and tampering program, and authenticity of communication session.
7. Reliability (Re): The reliability of service provided by cloud storage from the stakeholder’s viewpoint including policy on disaster recovery plan, patch management/maintenance, malicious code program and system monitoring.
8. Accountability (Ac): The Accountability of service provided by cloud storage from the stakeholder’s viewpoint including conformance with external and internal standards/policy, security functionality and security assurance.
9. Auditability (Au): Level of the auditability of data stored and accessed in cloud storage from the stakeholder’s viewpoint including audit policy, audit records and report generation.

With these operational definitions in mind, potential items were developed at this stage (Appendix B) following a Goal-Question-Metric (GQM) approach that will be explained in the next section.

6.2 Research Approach for Developing SecRas Instrument

According to Kassou and Kjiri (2012), there are three approaches that support metrics derivation from goals: GQM (Goal-Question-Metric) approach (Basili et al. 1994), GAM (Goal-Argument-Metric) (Cyra and Górski 2008) and BSC (Balanced Scorecard Framework) (Buglione and Abran 2000). In Table 6.1, a comparison and similarities between GQM, GAM and BSC is shown. Metrics derivations are carried out to generate items in an instrument. GQM approach provides an outline of process that defines goals, refining them into questions and then specifying measurements and finally data to be collected. GAM is a goal-oriented methodology for defining measurement plans. BSC is a framework that look into several factors for describing, implementing and managing strategy at different levels of an organisation by linking objectives, initiatives and measures to an organisation's strategy.

Considering the purpose and the general approach (top down derivation and bottom-up interpretation) GQM and GAM look the same. The differences relate to the way of defining and maintaining the relationship between the measurement goals and the metrics. In GAM, the goals and sub-goals are denoted as claims and then the analysis focuses on classifying which data and which properties of the data (further sub-goals) are needed to fulfil these claims whereas in GQM referring to a goal, several questions are defined in such a way that obtaining the answers to the questions leads to the achievement of the measurement goal then based on the questions, metrics are defined, which provide quantitative information then treated as answers to the questions (Basili et al. 1994).

GQM goals are referred to as a mission, while BSC goals are referred to a certain perspective and a certain particular tier in the organisational pyramid (hierarchy). Besides, GQM can be defined as a technique for deriving quantitative measures from a list of goals while BSC can be viewed as performance management framework that uses a GQM-like technique to derive the indicators (Buglione and Abran 2000).

Assessing the security of cloud storage in organisation begins by defining relatively the security metrics appropriate to the context of its information systems. For that purpose, we propose a GQM approach to produce security metrics for organisation based on its related security indicators (that is created as security questions).

Table 6.1 GQM, GAM and BSC approach (adapted from Buglione and Abran (2000), Kassou and Kjiri (2012))

Measurement Approach Approach Level	GQM	GAM	BSC
Conceptual - Objects	Goal	Claim	Goal
Operational - Assessment	Question	Assertion	Driver
Quantitative - Objective/Subjective	Metric	Metric	Indicator

6.2.1 Goal-Question-Metrics (GQM) Approach

This section will present the Goal-Question-Metric (GQM) and provide an example of its application in security research. GQM approach is based upon the assumption that for an organisation to measure in a purposeful way it must first specify goals for itself and its projects, then it must match those goals to the data that are expected to define those goals operationally, and finally provide a framework for interpreting the data with respect to the stated goals. The GQM paradigm (Basili 1992, 1993, Basili et al. 1994) is based on the notion that all measurement should be goal-oriented i.e. there has to be some rationale and need for collecting measurements and each measurement collected is stated in terms of the major goals. Questions are then derived from the goals and help to refine, articulate, and determine if the goals can be achieved. The metrics or measurements that are collected are then used to answer the questions in a quantifiable manner.

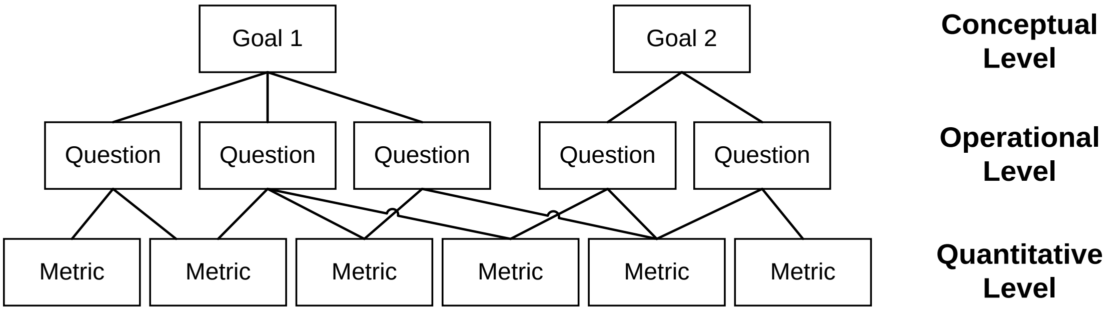


Figure 6.2 GQM hierarchical approach

A GQM Model is a hierarchical structure as presented in Figure 6.2 starting from a goal (specifying purpose of measurement, object/issue to measured, and viewpoint from which measure is taken). GQM defines a measurement model on three levels:

1. Conceptual level (Goal)

A goal is defined for an object, for a range of reasons, with respect to different models of value, from different perspectives and relative to a specific domain.

2. Operational level (Question)

A set of questions is utilized to define models of the object of study and after that concentrates on that object to describe the evaluation or accomplishment of a particular goal.

3. Quantitative level (Metric)

A set of measurements, taking into account the models, associated with every question in order to answer it quantifiably.

6.2.2 Application of GQM

Security goals are a set of conditions that describe properties such as confidentiality, integrity, availability, authenticity, and non-repudiation etc. of the systems security goal (Islam and Falcarin 2011). It is a set of requirements which consider organisation policies, security goals and security policies. The security goals identified are based on the security goal identified in the cloud storage security framework (CSSF). According to (Kassou and Kjiri 2012), there are four steps to build security metrics using GQM as follows.

6.2.2.1 Building Security Viewpoint

For each security goal, we need to define the viewpoint of the security context related to control that is provided by other viewpoints: Infrastructure, Governance etc.

6.2.2.2 Developing Goals

Based on Cloud Security Alliance Control Matrix (CCM) and other controls from the literature, we have identified security goals; confidentiality, integrity, availability, non-repudiation, authenticity, and reliability. Each security goal has relationships to the security of data in cloud storage.

6.2.2.3 Refining Security Goals into questions and deriving metrics

We present in Table 6.2 an example of the security goal of cloud storage security framework (CSSF) that can be described with a set of questions on security measures implementation.

6.2.2.4 Detailing metrics

A set of questions and sub questions follows the key indicator. Examples include, identity management, authentication, access management, access points etc. The four steps is explained in the context of the stakeholders. In this scenario, the stakeholders are IT managers in data centers that manage the cloud storage. IT managers are responsible to maintain an organisation's effectiveness and efficiency by defining, delivering, and supporting strategic plans for implementing secure information technologies i.e. ensuring that security policy and controls are in place and implemented in an organisation. Therefore, security goals in an organisation can be assessed and achieved.

Table 6.2 Applying GQM for Cloud Storage Security Framework (CSSF)

Goal: Assessing the confidentiality of data accessed in cloud storage from the stakeholder’s viewpoint

Goal		Question	Metric
Purpose	Assessing	Are there identity management system? Are there authentication process/solutions? Are there authorisation and restriction to data? Are there detection of security defects in API? Are there encrypted communication channel used when transferring data (i.e. to virtual machines etc.)?	Rating Score*
Factor	Confidentiality		
Object	Data		
Where	Cloud Storage		
Viewpoint	Stakeholder’s		

* Rating Score 1 - No Plan to Implement, 2 - Planning to Implement, 3 - Do not know/Unsure, 4 - Partially Implement, 5 - Fully Implement

Goal: Assessing the integrity of data stored in cloud storage from the stakeholder’s viewpoint

Goal		Question	Metric
Purpose	Assessing	Are there encryption of data at rest (on disk/storage)? Are there data ownership encryption keys? Are data ownership documented? Are there secure data deletion? Are there additional protection for sensitive data?	Rating Score*
Factor	Integrity		
Object	Data		
Where	Cloud Storage		
Viewpoint	Stakeholder’s		

* Rating Score 1 - No Plan to Implement, 2 - Planning to Implement, 3 - Do not know/Unsure, 4 - Partially Implement, 5 - Fully Implement

Goal: Assessing the availability of data stored in cloud storage from the stakeholder’s viewpoint

Goal		Question	Metric
Purpose	Assessing	Are data logically segment for each user? Are there backup or redundancy mechanisms? Are there data recovery mechanisms?	Rating Score*
Factor	Availability		
Object	Data		
Where	Cloud Storage		

Goal		Question	Metric
Viewpoint	Stakeholder’s		

* Rating Score 1 - No Plan to Implement, 2 - Planning to Implement, 3 - Do not know/Unsure, 4 - Partially Implement, 5 - Fully Implement

Goal: Assessing the non-repudiation of data stored in cloud storage from the stakeholder’s viewpoint

Goal		Question	Metric
Purpose	Assessing	Are there synchronised time-service protocol (e.g., NTP etc.)? Are there key management policies binding keys to identifiable owners? Are there restrictions of user data to specific countries or geographic locations?	Rating Score*
Factor	Non-repudiation		
Object	Data		
Where	Cloud Storage		
Viewpoint	Stakeholder’s		

* Rating Score 1 - No Plan to Implement, 2 - Planning to Implement, 3 - Do not know/Unsure, 4 - Partially Implement, 5 - Fully Implement

Goal: Assessing the authenticity of data accessed and stored in cloud storage from the stakeholder’s viewpoint

Goal		Question	Metric
Purpose	Assessing	Are there cryptographic protection mechanisms? Are there anti-counterfeiting policy? Are there session-level protection where needed?	Rating Score*
Factor	Authenticity		
Object	Data		
Where	Cloud Storage		
Viewpoint	Stakeholder’s		

* Rating Score 1 - No Plan to Implement, 2 - Planning to Implement, 3 - Do not know/Unsure, 4 - Partially Implement, 5 - Fully Implement

Goal: Assessing the reliability of service provided by cloud storage from the stakeholder’s viewpoint

Goal		Question	Metric
Purpose	Assessing	Are there multi-failure disaster recovery? Are there Patch Management or System Maintenance policy? Are there malicious code protection mechanisms at entry and exit points? Are there monitoring process or solutions?	Rating Score*
Factor	Reliability		
Object	Service		
Where	Cloud Storage		
Viewpoint	Stakeholder’s		

* Rating Score 1 - No Plan to Implement, 2 - Planning to Implement, 3 - Do not know/Unsure, 4 - Partially Implement, 5 - Fully Implement

Goal: Assessing the accountability of services provided by cloud storage from the stakeholder’s viewpoint

Goal		Question	Metric
Purpose	Assessing	Are there conformance with external standards? Are there conformance with internal standards/policy? Are there clarity of Service Level Agreement/Guarantee (SLAs/SLGs)? Are there penetration tests of cloud service infrastructure regularly as prescribed by industry best practices/guidance?	Rating Score*
Factor	Accountability		
Object	Service provided		
Where	Cloud Storage		
Viewpoint	Stakeholder’s		

* Rating Score 1 - No Plan to Implement, 2 - Planning to Implement, 3 - Do not know/Unsure, 4 - Partially Implement, 5 - Fully Implement

Goal: Assessing the auditability of data accessed and stored in cloud storage from the stakeholder's viewpoint

Goal		Question	Metric
Purpose	Assessing	Are there audit assertions using a structured, industry accepted format (e.g., CloudAudit/CloudTrust, ISACA's Cloud Computing Management Audit, etc.)? Are audit logs reviewed on a regular basis? Are there on-demand audit review? Are the audit log in original content or time?	Rating Score*
Factor	Auditability		
Object	Data		
Where	Cloud Storage		
Viewpoint	Stakeholder's		

* Rating Score 1 - No Plan to Implement, 2 - Planning to Implement, 3 - Do not know/Unsure, 4 - Partially Implement, 5 - Fully Implement

The GQM application above was used as the set of ideas in developing the statements for questionnaire design in the next section.

6.3 Questionnaire Design

The constructs and statements of a questionnaire designed to validate the study were administered in English. The whole questionnaire survey is shown in Appendix B. Table 6.3 presents the set of statements items that has been used to describe each factors.

Table 6.3 Questionnaire Statements

Item Code	Statements/Questions
CS1	My organisation has cloud storage security policy in place
CS2	My organisation has implemented security procedures to comply to industry standards
CS3	My organisation has enforced security processes to support cloud security storage policy/ies
CS4	My organisation has imposed security controls to support cloud storage security policy/ies
Co1	My organisation has identity management policy
Co2	My organisation has user-based authentication process
Co3	My organisation has access management policy
Co4	My organisation has process to specify rights and restrictions for user access to data
Co5	My organisation follows industry standards to build in security for systems
Co6	My organisation has a secure communication channel policy
In1	My organisation has documentation of encryption management practices/guidelines
In2	My organisation encrypt user data at rest (on disk/storage) within the environment

Item Code	Statements/Questions
In3	My organisation leverage encryption to protect virtual machine images during transport between hypervisor instances
In4	My organisation supports data stewardship
In5	My organisation support secure deletion (e.g., degaussing/cryptographic wiping) of archived data
In6	My organisation supports additional protection for user to store sensitive data
Av1	My organisation segments data logically for each user
Av2	My organisation has backup or redundancy mechanisms
Av3	My organisation has data recovery mechanisms
Av4	My organisation verifies data authenticity after restore process
Av5	My organisation documents the restore or redundancy mechanisms
Av6	My organisation has define restore procedure for responding to requests for user data from governments or third parties
Nr1	My organisation has key management policies binding keys to identifiable owners
Nr2	My organisation uses synchronised time-service protocol (e.g., NTP etc.)
Nr3	My organisation uses geographical location as an authentication
Nr4	My organisation has restriction of user data to specific countries or geographic locations
Nr5	My organisation provides the means for authorised individuals to determine the identity of the data producer
Nr6	My organisation support integration of location as an authentication factor
At1	My organisation has cryptographic protection mechanisms
At2	My organisation ensures the origin authentication
At3	My organisation has verification assurances to ensure session authenticity
At4	My organisation has anti-counterfeiting policy
At5	My organisation provide session-level protection where needed
At6	My organisation provides mechanisms to protect the authenticity of communications sessions
Re1	My organisation has multi-failure disaster recovery
Re2	My organisation has system maintenance process/policy
Re3	My organisation has patch management policy/process
Re4	My organisation has continuous monitoring process/solutions
Re5	My organisation has malicious code protection mechanisms at entry and exit points
Re6	My organisation has conducted failover test
Ac1	My organisation has process to conformance with external standards
Ac2	My organisation has mechanisms to put internal security policies in effect
Ac3	My organisation supports transparency and participation to

Item Code	Statements/Questions
	conformance process with internal standards/policy
Ac4	My organisation ensures clarity of Service Level Agreement/Guarantee (SLAs/SLGs)
Ac5	My organisation conducts penetration tests of cloud service infrastructure regularly as prescribed by industry best practices/guidance
Ac6	My organisation has means of remediation for internal enforcement
Au1	My organisation produces audit assertions using a structured, industry accepted format (e.g., Cloud Audit/Cloud Trust, ISACA's Cloud Computing Management Audit, etc.)
Au2	My organisation reviews audit logs on a regular basis
Au3	My organisation has on-demand audit review
Au4	My organisation generates audit report
Au5	My organisation ensures the audit log is in original content or time
Au6	My organisation has a process to audit records for events of interest

An online questionnaire was distributed electronically through emails and posted on security group pages in LinkedIn and Facebook. The online questionnaire comprised four pages and a brief introduction page. The introduction page consisted of three parts: a welcome statement, the description of the Security Rating Score (SecRaS) and consent information. The other four pages covered the different parts of the study, and these are:

Part I: Demographic information: this part included a request for general information such as organisation domain, job role, experience (in ICT security) and estimation of percentage data stored in cloud storage. This part was important to give the researcher an overview of information that may be needed for the purposes of group comparison.

Part II: Security Policy, Procedure & Practice information: this part comprises the participants' organisation's existing security policies in general ICT security policies, cloud storage security policies and whether the security procedures and practices have been implemented.

Part III: Security in Cloud Storage information: this part was designed to obtain the security measures in cloud storage in their organisations. Questions were asked to what extent they agree with the controls/statements associated with factors.

The University of Southampton Ethics Committee approved the quantitative methodologies conducted in this study. Ethics approval was granted under reference number 18945 on 1 February 2016 for the online questionnaire.

6.4 Response Item

iSurvey software was used to generate the questionnaire, with a five-level Likert scale implemented for all statements. Based on the nature of the item that acquiring user's opinion about SecRaS applicability in security in cloud storage, the agreement with the following ratings were used:

1. Strongly agree (rating score=5)
   The highest scale that indicates a total agreement of the application of security factors to protect data in cloud storage. The rating shows an important effect to the items in SecRaS.
2. Agree (rating score= 4)
   The satisfactory scale that indicates an agreement to the application of security factors to protect data in cloud storage. The rating shows a satisfactory effect to the items in SecRaS.
3. Neutral (rating score= 3)
   The medium scale that indicates an undecided agreement to the application of security factors to protect data in cloud storage. The rating shows some effect to the items in SecRaS.
4. Disagree (rating score= 2)
   The low scale that indicates disagreement agreement to the application of security factors to protect data in cloud storage. The rating shows minor effect to the items in SecRaS.
5. Strongly disagree (rating score= 1)
   The lowest scale that indicates total disagreement with the application of security factors to protect data in cloud storage. The rating shows little effect to the items in SecRaS.

6.5 Validity and Reliability

After completing the design of the questionnaire, it was necessary to ensure that the statements in the questionnaire were measuring the factors in the instrument accurately; thus, validity and reliability tests were considered to obtain accurate results from the instrument (Saunders et al. 2009). Instrument validation is essential to ensure that the question is measuring what it is supposed to measure i.e. the factors (Pallant 2013). The validity represents a high degree of confidence that the data collected and findings represent a scientific and truthful investigation. On the other hand, reliability ensures the multiple items are consistent in the same construct and the results of the study are able to be repeated and reliable (Cramer and Howitt 2004). Validity and reliability tests are independent of each other; this means that if the instrument is valid it is not necessarily also reliable, and also that if it is reliable it is not necessarily valid (Field 2013). There are different methods of establishing validity and reliability. The following sections discuss the validity and reliability tests in detail carried out in pre-test and validation study.

6.6 Validating the SecRaS Instrument

A study was designed to validate the Security Rating Score (SecRaS) instrument developed in the previous section. The accuracy of findings and interpretations are based on strong validation of the instruments that are used to collect the data (Straub et al. 2004).

The validation process involves two parts: (a) a pre-test, and (b) a validation study. The following sections describe each of the parts in detail.

6.6.1 Pre-test

A pre-test of the instrument was conducted with five security experts that consist of respondents from IT manager of an IT Department in a government office, security experts and academic researchers. Experts were selected to undertake content validity during pre-test. Content validity is sufficient to be performed with experts that are experienced in the research context (Lynn 1986). The experts were asked to complete the questionnaire to determine whether they could understand the wording of the questions and to suggest improvements. The objectives of the pre-test were to evaluate whether:

1. An item is relevant and adequate in examining the concept being studied,
2. An item’s wording, response format, instructions, instrument length and layout is appropriate, and
3. The instrument as a whole is easy to read and understand.

The pre-test participants suggested that all the items for one construct should be measured in the same direction. The content validity experts also made this comment. Another issue that the respondents commented upon was the length of the questionnaire. Therefore, issues were carefully considered and such changes were made wherever required. For example, bearing in mind the suggestion of the pre-test participants the total length of pages in the questionnaire was reduced from six to four pages. Apart from the changes and a few spelling and typographical errors, the respondents from the pre-test studies supported the content of the questionnaire.

6.6.1.1 Content Validity

This study was validated using content validity. Content validity refers to how accurately the instrument is representative of the construct of the items; this type of validity relies on the knowledge of experts, either in the particular content area or as researchers (Cronbach, 1971; Straub 1989). Content validity is largely a matter of judgment, involving two distinct phases: a preliminary process by the instrument developer to enhance content validity through conceptualisation and domain analysis prior to item generation, and later evaluate the relevance of the instrument's content through expert assessment (Lynn 1986, Polit and Beck 2006). Content validity was established after designing the questionnaire and before conducting the survey. Without undertaking content validity, the instrument validity is questionable (Garver and Mentzer 1999).

Essentially, there are two stages in the process of assessment of content validity: the developmental stage and the judgment quantification stage (Lynn 1986). The developmental stage begins with measurement of the objective of the instrument and identification of the full content domain; this step can be accomplished through a literature review and consulting experts. Then the cognitive measure is to ensure that each item in the instrument is representing appropriately the scope of the content and it is clear that generating many indicators is better than only one or two indicators for each construct; although three indicators are acceptable at minimum, it is better when the construct has four indicators or more (Hair et al. 2014). The next stage is adjustment of the indicators into useable form. In the next step, the indicators need to be refined and revised. If necessary, the last two steps can be justified personally by the main researcher (Lynn 1986). The instrument of the study was constructed and the statements in the questionnaire were informed from previous studies related to the area of study (CSA 2014) and additional statements were devised be the researcher. These statements referred to cloud storage security factors that were not included in previous studies within an academic context.

The second stage of judgment quantification is concerned with two concepts; all item indicators are content valid and the developed instrument is content valid for the research context. This stage is accomplished through justification by experts (Lynn

1986). Quantification of expert judgements was performed amongst five security experts during pre-test. Through this, the researcher gathered valuable insights and suggestions from different researchers' perceptions and could verify whether respondents were able to understand and answer all the questions. The experts were a selection of five security experts from various organisation in UK and Malaysia. The number of experts is hard to decide and there is no a standard number, because it is based on the number of accessible people who gave consent to participate; however, Lynn (1986) advised a minimum of three experts, but indicated that more than 10 was probably unnecessary.

The first step of judgment quantification involved a one and a half hour meeting with two researchers, during which they were asked to identify key issues in relation to which questions and statements could be developed or removed. Through their reading of the questionnaires some questions and comments emerged around ambiguous statements and repeated indicators. Each expert has filled up a response for whether each question or item is necessary for the concept being studied. Their responses were "important", "neither important nor unimportant" or "unimportant". By the end of the meeting significant comments had been received; therefore, appropriate changes were made. Thus a new version of the questionnaire was prepared to present to the next three experts. The following step of the stage were conducted in the same way, but through individually meeting with experts, and amendments to the statements were implemented during each meeting. During the last two meetings there was no significant adjustment. Overall, through content validity, about forty three statements and questions were reformulated, as well as the welcome statement.

6.6.1.2 Results of Content Validity

Responses from all experts were gathered and items indicated "important" by the experts were calculated. A statistical significant level for each factor was estimated based on the content validity ratio (CVR). CVR is a quantitative approach to content validity introduced by Lawshe (1975), the calculation is as follow:

$$CVR = \left(n_e - \frac{N}{2}\right) / \left(\frac{N}{2}\right) \qquad 6.1$$

where n_e is number of experts agreeing that the item is "important" and N is the total number of participating experts in the pre-test study. For a CVR to be considered as important, the level of agreement among experts was greater than 50% i.e. the value must be 0.5 or more to be considered significant at 0.05 and items lower than 0.5 is considered as not significant (Lawshe 1975, Ayre and Scally 2014). Lawshe (1975) reported a table of critical CVR values, where critical CVR is the level of agreement of probability for a given item and for a given alpha (Type I error probability, suggested to be 0.05 using a one-tailed test). The results shows that from a pool of 52 questions, only 43 are significant at the range of 0.8 (highest) to 0.6 (lowest), at 0.05 significant level. These CVR indicate that the items in SecRaS has adequate content validity, which means that the items in SecRaS is measuring the concept being studied.

Table 6.4 Content Validity Ratio for 52 potential items

Factor	Total of Items	Significant Items	CVR item 1	CVR item 2	CVR item 3	CVR item 4	CVR item 5	CVR item 6	Average CVR
CS	4	3	1	1	1	0.2	-	-	0.80
Co	6	6	0.6	0.6	1	1	0.6	1	0.80
In	6	6	0.6	1	0.6	0.6	1	1	0.80
Av	6	4	1	1	1	1	0.2	0.2	0.73
Nr	6	4	1	0.6	1	1	0.2	0.2	0.67
At	6	5	0.6	1	1	0.6	0.6	0.2	0.67
Re	6	5	1	0.6	1	0.6	0.6	0.2	0.67
Ac	6	5	1	0.6	1	0.6	0.6	0.0	0.63
Au	6	5	1	0.6	0.6	0.6	0.6	0.2	0.60
Total	52	43							

Items with CVR value less than 0.5 were removed from the instrument. From a pool of 52 items (see Appendix B.3), the refined instrument now consists of 43 items (see Appendix B.7). From the assessment the experts also provided a number of suggestions. The suggestions were mostly about the wording, the structure of the sentence, and adding example for the items. The suggestions was to ensure that the meaning of the items can be delivered as it should be. The response items (5 – strongly agree to 1 – disagree) were considered as appropriate.

6.6.2 Validation Study

A study was conducted to determine the response rate and learn of any discrepancies within the questions, which included determining whether the format of the questionnaire and questions were suitable. Additionally, the duration required to complete the questionnaire was also established. Next, the refined instrument was distributed to a sample of respondents and an analysis of the responses was conducted to obtain the instrument's reliability. The objectives of the validation study are to investigate:

1. The relationship between items and the factor, and

2. The relationship between each factor and the scale as a whole.

Analysing the relationship is important to look into the intercorrelation between items and factors and explore the suitability of applying factor analysis in the next stage. Factor analysis is suitable if most correlations exist are not too low (r < .30) and not too high (r > .90) (Field 2013, Pallant 2013).

A number of 30 security practitioners were invited to participate in the study. Statisticians mentioned that a sample size of 30 (above about 30) is large enough for a study considering the principles of the Central Limit Theorem (Field 2013). The practitioners were recruited on the basis of their interest in security, their research in security and their experience. The majority of the respondents reported that the questionnaire was easily understandable and required 15-20 minutes for completion. Additionally, the majority of the respondents validated the content of the questionnaires, although minor changes to the final design of the questionnaire were undertaken based upon the received feedback, and a final questionnaire was developed. Changes made to the original instrument include selecting an adequate number of items to represent a factor.

6.6.2.1 Correlations Analysis

Correlations analysis show the value of the correlation coefficient. This can range from -1.00 to 1.00. This value will indicate the strength of relationship between variables. A correlation of 0 indicates no relationship at all, a correlation of 1.00 indicates a perfect positive correlation, and a value of -1.0 indicates a perfect negative correlation. The relationship guideline (Cohen 1988) suggests the following:

1. Small, r = .10 to .29
2. Medium, r = .30 to .49
3. Large, r = .50 to 1.0

These guidelines apply whether or not there is a negative sign on the front of the r value. The negative sign refers only to the direction of the relationship, not the strength (Field 2013, Pallant 2013, Hair et al. 2014).

6.6.2.1.1 Correlation among security factors

The correlation matrix shows the strength of relationship between the factors. The correlation gives us information to decide whether it was reasonable to assume that the factors were not related. When applying correlation between the factors, the correlation matrix retrieved is shown in Table 6.5. The results shows the significant correlations for the factors related to this research.

- Cloud Storage Security (CS) score is significantly correlated to Confidentiality, $r(30)$ = 0.375, Integrity, $r(30)$ = 0.370, Non-repudiation, $r(30)$ = 0.392, Authenticity, $r(30)$ = 0.378, Reliability, $r(30)$ = 0.394, and Auditability, $r(30)$ = 0.396, (all p<0.05). There is also a significant correlation to Availability, $r(30)$ = 0.602 and Accountability, $r(30)$ = 0.518, (both p<0.01).
- Confidentiality (Co) score is significantly correlated to Integrity, $r(30)$ = 0.432, and authenticity, $r(30)$ =0.388, (both p<0.05) and availability, $r(30)$ =0.490, p<0.01.
- Integrity (In) score is significantly correlated to Availability, $r(30)$ = 0.411, Non-repudiation, $r(30)$ =0.467 and Accountability, $r(30)$ =0.396, (all p<0.05).
- Availability (Av) score is significantly correlated to Non-repudiation, $r(30)$ = 0.585, and Accountability, $r(30)$ =0.378, (all p<0.01).
- Non-repudiation (Nr) score is significantly correlated to Authenticity, $r(30)$ = 0.515 p<0.01, Accountability, $r(30)$ = 0.378, and Auditability, $r(30)$ =0.391, (both p<0.05).
- Authenticity (At) score is positively correlated to Accountability, $r(30)$ = 0.429, p<0.05 and Authenticity, $r(30)$ = 0.564, p<0.01.
- Reliability (Re) score is significantly correlated to Auditability, $r(30)$ = 0.415, p<0.05.

Table 6.5 Correlation Matrix for nine factors

Factors	Co	In	Av	Nr	At	Re	Ac	Au
CS. Cloud Storage Security	0.375*	0.370*	0.602**	0.392*	0.378*	0.394*	0.518**	0.396*
Co. Confidentiality	-	0.432*	0.490**	0.221	0.388*	0.270	0.280	0.199
In. Integrity		-	0.411*	0.467*	0.755**	0.307	0.396*	0.540**
Av. Availability			-	0.585**	0.352	0.148	0.519**	0.231
Nr. Non-repudiation				-	0.515**	0.125	0.378*	0.391*
At. Authenticity					-	0.110	0.429*	0.564**
Re. Reliability						-	0.268	0.415*
Ac. Accountability							-	0.179
Au. Auditability								-

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

6.6.2.1.2 Correlations for Cloud Storage Security (CS)

The results from Table 6.6 show the correlations from the cloud storage security implementation factor. The first item, security policy is significantly correlated to security procedures, r(30) =0.85 and security processes, r(30) =0.80, (all p<0.01).

Table 6.6 Correlation for Cloud Storage Security Factor

	CS2	CS3
CS1. My organisation has cloud storage security policy in place	.850**	.801**
CS2. My organisation has implemented security procedures to comply to industry standards	-	.816**
CS3. My organisation has enforced security processes to support cloud storage security policy/ies		-

**. Correlation is significant at the 0.01 level (2-tailed).

6.6.2.1.3 Correlations for Confidentiality (Co)

The results from Table 6.7 show the correlations from the confidentiality factor. The first item, identity management is significantly correlated to user-base authentication, r(30) =0.71, access management, r(30) =0.54, rights and restriction, r(30) = 0.48, standards to build-in security systems, r (30) = 0.48, (all p<0.01), and secure communication channel, r(30) = 0.36, (p<0.05).

Table 6.7 Correlations for Confidentiality Factor

	Co2	Co3	Co4	Co5	Co6
Co1. My organisation has identity management policy	.707**	.537**	.482**	.484**	-.363*
Co2. My organisation has user-based authentication process	-	.576**	.625**	.363*	-.314
CO3. My organisation has access management policy		-	.628**	.546**	-.344
Co4. My organisation has process to specify rights and restrictions for user access to data			-	.661**	-.452*
Co5. My organisation follows industry standards to build in security for systems				-	-.400*
Co6. My organisation has a secure communication channel policy					-

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

6.6.2.1.4 Correlations for Integrity (In)

The integrity factor shows encryption management is significantly correlated to encrypt user data at rest, r(30) = 0.88, leverage encryption on VM, r(30) = 0.60, Data stewardship, $r(30) = 0.70$, secure deletion, $r(30) = 0.70$, and data protection for sensitive data, $r(30) = 0.6$, (all $p<0.01$).

Table 6.8 Correlations for Integrity Factor

	In2	In3	In4	In5	In6
In1. My organisation has documentation of encryption management practices/guidelines	.878**	.592**	.683**	.683**	.611**
In2. My organisation encrypt user data at rest (on disk/storage) within the environment	-	.754**	.715**	.715**	.708**
In3. My organisation leverage encryption to protect virtual machine images during transport between hypervisor instances		-	.627**	.573**	.675**
In4. My organisation supports data stewardship			-	.889**	.884**
In5. My organisation support secure deletion (e.g., degaussing/cryptographic wiping) of archived data				-	.884**
In6. Data protection for sensitive data My organisation supports additional protection for user to store sensitive data					-

**. Correlation is significant at the 0.01 level (2-tailed).

6.6.2.1.5 Correlations for Availability (Av)

The availability factor shows the item logically segments each user data to have significant correlation to backup/redundancy mechanism, r (30) = 0.82, recovery of data, r (30) = 0.59, and verifying the restored data, r (30) = 0.54, (all $p<0.01$).

Table 6.9 Correlations for Availability Factor

	Av2	Av3	Av4
Av1. My organisation segments data logically segment for each user	.822**	.587**	.536**
Av2. My organisation has backup or redundancy mechanisms	-	.507**	.495**
Av3. My organisation has data recovery mechanisms		-	.777**
Av4. My organisation verifies data authenticity after restore process			-

**. Correlation is significant at the 0.01 level (2-tailed).

6.6.2.1.6 Correlations for Non-Repudiation (Nr)

The correlation for non-repudiation factor has shown that the key management policies binding key to identifiable owners have significant correlation to synchronised time-service protocol, r (30) = 0.81, geographical location as authentication factor, r (30) = 0.65, and restriction of user data to specific countries or geographical location, r (30) = 0.54, (all $p<0.01$). There is also a significant correlation between using synchronised time-service protocol and restriction of user data to specific countries or geographic locations, r (30) = 0.36, p<0.05.

Table 6.10 Correlations for Non-Repudiation Factor

	Nr2	Nr3	Nr4
Nr1. My organisation has key management policies binding keys to identifiable owners	.807**	.651**	.543**
Nr2. My organisation uses synchronised time-service protocol (e.g., NTP etc.)	-	.896**	.385*
Nr3. My organisation uses geographical location as an authentication		-	.204
Nr4. My organisation has restriction of user data to specific countries or geographic locations			-

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

6.6.2.1.7 Correlations for Authenticity (At)

The results from Table 6.11 show the correlations from the authenticity factor. The cryptographic protection is significantly correlated to ensuring origin authentication, r(30) = 0.82, verification for session authenticity, r(30) =0.76, anti-counterfeiting policy, r(30) = 0.55, (all p<0.01), and session-level protection, r(30) = 0.46, p<0.05.

Table 6.11 Correlations for Authenticity Factor

	At2	At3	At4	At5
At1. My organisation has cryptographic protection mechanisms	.817**	.758**	.549**	.460*
At2. My organisation ensures the origin authentication	-	.729**	.424*	.350
At3. My organisation has verification assurances to ensure session authenticity		-	.688**	.604**
At4. My organisation has anti-counterfeiting policy			-	.536**
At5. My organisation provide session-level protection where needed				-

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

6.6.2.1.8 Correlations for Reliability (Re)

The reliability factor shows multi-failure disaster recovery is significantly correlated to system maintenance, r(30) = 0.69, malicious code protection mechanism, r(30) = 0.57, (all p<0.01), and continuous monitoring, r(30) = 0.37, p<0.05. Patch management has a significant correlation to continuous monitoring, r (30) = 0.73, and malicious code protection, r (30) = 0.70, (all p<0.01).

Table 6.12 Correlations for Reliability Factor

	Re2	Re3	Re4	Re5
Re1. My organisation has multi-failure disaster recovery	.690**	.359	.365*	.567**
Re2. My organisation has system maintenance process/policy	-	.357	.570**	.612**
Re3. My organisation has patch management policy/process		-	.730**	.696**
Re4. My organisation has continuous monitoring process/solutions			-	.785**
Re5. My organisation has malicious code protection mechanisms at entry and exit points				-

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

6.6.2.1.9 Correlations for Accountability (Ac)

The correlation for accountability factor has shown that conformance with external standards have significant correlation to mechanisms to put internal security policies in effect, r (30) = 0.75, supports transparency and participation to conformance process with internal standards/policy, r (30) = 0.77, ensures clarity of Service Level Agreement/Guarantee (SLAs/SLGs), r (30) = 0.78, and conducts penetration tests of cloud service infrastructure regularly as prescribed by industry best practices/guidance, r (30)=0.88, (all p<0.01).

Table 6.13 Correlations for Accountability Factor

	Ac2	Ac3	Ac4	Ac5
Ac1. My organisation has process to conformance with external standards	.750**	.768**	.776**	.883**
Ac2. My organisation has mechanisms to put internal security policies in effect	-	.854**	.584**	.592**
Ac3. My organisation supports transparency and participation to conformance process with internal standards/policy		-	.637**	.656**
Ac4. My organisation ensures clarity of Service Level Agreement/Guarantee (SLAs/SLGs)			-	.871**
Ac5. My organisation conducts penetration tests of cloud service infrastructure regularly as prescribed by industry best practices/guidance				-

**. Correlation is significant at the 0.01 level (2-tailed).

6.6.2.1.10 Correlations for Auditability (Au)

The results from Table 6.14 show the correlations from the auditability factor. The audit assertions using a structured, industry accepted format (e.g., Cloud Audit/Cloud Trust, ISACA's Cloud Computing Management Audit, etc.) is significantly correlated to reviews audit logs on a regular basis, r (30) = 0.58, on-demand audit review, r (30) =0.58, generates audit report, r (30) = 0.55, and ensures the audit log is in original content or time, r (30) = 0.71, all p<0.01.

Table 6.14 Correlations for Auditability Factor

	Au2	Au3	Au4	Au5
Au1. My organisation produces audit assertions using a structured, industry accepted format (e.g., Cloud Audit/Cloud Trust, ISACA's Cloud Computing Management Audit, etc.)	.579**	.579**	.681**	.709**
Au2. My organisation reviews audit logs on a regular basis	-	.739**	.528**	.533**
Au3. My organisation has on-demand audit review		-	.812**	.774**
Au4. My organisation generates audit report			-	.813**
Au5. My organisation ensures the audit log is in original content or time				-

**. Correlation is significant at the 0.01 level (2-tailed).

6.6.3 Reliability of the SecRaS Instrument

There are two reliability test methods which are widely used: internal consistency and test-retest reliability (Pallant 2013). Internal consistency is the extent to which the items are interrelated and internally consistent to a specific construct, whereas test-retest reliability refers to conducting the same test with the same group on different occasions; the correlation between the two results indicates the degree of reliability (Pallant 2013).

The study used an internal consistency reliability test which is the Cronbach's Alpha (α) test. The Cronbach Alpha (α) test is a statistical method calculated through SPSS. The results provide the average correlation of all items in the same construct (Pallant 2013). The reliability scores obtained using Cronbach alpha range between 0 and 1; a result closer to 1 indicates higher reliability. However, the reliability scores rely on the size of the questions' scales: if the scales were ten or less, the minimum score of reliability accepted is 0.7 (Hair et al. 2014). A reliability value of 0.5 is accepted for item-to-total correlation (Sekaran 2003, Hair et al. 2014). Table 6.15 shows the reliability score range and the level of acceptance of the study, based on the literature review.

Table 6.15 Cronbach's Alpha Reliability Scores (Pallant 2013, Hair et al. 2014)

Cronbach alpha	Level of Internal Consistency
α ≥ 0.9	Excellent

Cronbach alpha	Level of Internal Consistency
0.9 > α ≥ 0.8	Good
0.8 > α ≥ 0.5	Acceptable
α < 0.5	Poor

6.6.3.1 Internal consistency for security factors

The internal consistency reliability test was undertaken for the nine factors in a stepwise process; the factor's reliability is checked and if the Cronbach's alpha value for the factor is low, suitable items in the factor that could raise the alpha value will be removed. Items with low item-to-item and item-to scale correlations, which would raise alpha value if deleted, would be considered for elimination. The factors have shown a good alpha value therefore, no items from any factors were eliminated. The overall Cronbach's Alpha as shown in Table 6.16 for the instrument is 0.954 which indicates an excellent level of internal consistency.

Table 6.16 Total Reliability Statistics for SecRaS

Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
.954	.952	43

The findings of the reliability test has demonstrated an acceptable level of reliability for each the constructs as shown in Table 6.17. The Cronbach's alpha value is between 0.5 and 0.9 indicating an acceptable level of internal consistency.

Table 6.17 Reliability Test

Factor	Number of Items	Cronbach's α
Cloud Storage Security	3	0.933
Confidentiality	6	0.749
Integrity	6	0.936
Availability	4	0.874
Non-repudiation	4	0.841
Authenticity	5	0.879
Reliability	5	0.867
Accountability	5	0.933
Auditability	5	0.518

6.6.3.2 Internal consistency for Cloud Storage Security (CS) factor

The Cronbach's α = .933 shown in Table 6.18 is considered an excellent value for indicating the consistency of three items in CS factor. Therefore, all items were remained.

Table 6.18 Reliability Statistics for Cloud Storage Security (CS) factor

Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
.933	.933	3

Table 6.19 Item-Total Statistics for Cloud Storage Security (CS) factor

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
CS1	5.87	5.706	.867	.758	.898
CS2	6.07	5.926	.879	.774	.888
CS3	6.27	6.271	.841	.708	.919

6.6.3.3 Internal consistency for Confidentiality (Co) factor

The Cronbach's α = .749 shown in Table 6.20 is considered an acceptable value for indicating the consistency of six items in Co factor. Although if deleting item Co6, the alpha value will increase but the total alpha value is sufficient. Therefore, all items were remained.

Table 6.20 Reliability Statistics for Confidentiality (Co) factor

Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
.749	.666	6

Table 6.21 Item-Total Statistics for Confidentiality (Co) factor

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Co1	15.07	16.823	.649	.589	.667
Co2	14.97	16.171	.683	.654	.654
Co3	15.30	14.976	.687	.497	.647
Co4	15.80	15.821	.709	.658	.645
Co5	15.87	16.740	.584	.543	.684
Co6	17.17	28.144	-.464	.243	.864

6.6.3.4 Internal consistency for Integrity (In) factor

The Cronbach's α = .936 shown in Table 6.22 is considered an excellent value for indicating the consistency of six items in In factor. Therefore, all items were remained.

Table 6.22 Reliability Statistics Integrity (In) factor

Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
.936	.940	6

Table 6.23 Item-Total Statistics Integrity (In) factor

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
In1	14.73	25.513	.783	.805	.932
In2	14.73	25.857	.873	.864	.916
In3	15.00	29.034	.716	.646	.936
In4	14.77	27.771	.857	.846	.919
In5	14.77	27.909	.844	.853	.921
In6	14.83	28.833	.846	.859	.922

6.6.3.5 Internal consistency for Availability (Av) factor

The Cronbach’s α = .874 shown in Table 6.24 is considered a good value for indicating the consistency of four items in Av factor. Therefore, all items were remained.

Table 6.24 Reliability Statistics for Availability (Av) factor

Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
.874	.875	4

Table 6.25 Item-Total Statistics for Availability (Av) factor

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Av1	11.57	8.944	.792	.869	.815
Av2	11.63	9.344	.733	.853	.838
Av3	12.00	9.310	.719	.654	.844
Av4	12.30	9.114	.683	.618	.860

6.6.3.6 Internal consistency for Non-repudiation (Nr) factor

The Cronbach’s α = .841 shown in Table 6.26 is considered a good value for indicating the consistency of four items in Nr factor. Therefore, all items were remained.

Table 6.26 Reliability Statistics Non-repudiation (Nr) factor

Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
.841	.847	4

Table 6.27 Item-Total Statistics Non-repudiation (Nr) factor

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Nr1	8.45	10.470	.821	.722	.733
Nr2	8.41	10.037	.848	.893	.717
Nr3	8.41	11.966	.679	.830	.798
Nr4	8.66	12.734	.412	.350	.916

6.6.3.7 Internal consistency for Authenticity (At) factor

The Cronbach's α = .879 shown in Table 6.28 is considered a good value for indicating the consistency of five items in At factor. Therefore, all items were remained.

Table 6.28 Reliability Statistics for Authenticity (At) factor

Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
.879	.879	5

Table 6.29 Item-Total Statistics for Authenticity (At) factor

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
At1	12.10	18.852	.795	.736	.833
At2	12.13	20.120	.701	.720	.856
At3	11.73	18.823	.867	.757	.816
At4	12.10	21.472	.648	.518	.868
At5	11.53	21.982	.561	.410	.888

6.6.3.8 Internal consistency for Reliability (Re) factor

The Cronbach's α = .867 shown in Table 6.30 is considered a good value for indicating the consistency of five items in Re factor. Therefore, all items were remained.

Table 6.30 Reliability Statistics for Reliability (Re) factor

Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
.867	.870	5

Table 6.31 Item-Total Statistics for Reliability (Re) factor

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Re1	15.17	15.316	.583	.567	.868
Re2	14.93	15.168	.679	.624	.842
Re3	14.73	15.651	.635	.604	.852
Re4	14.80	14.510	.743	.738	.825
Re5	14.90	14.714	.834	.728	.806

6.6.3.9 Internal consistency for Accountability (Ac) factor

The Cronbach’s α = .933 shown in Table 6.32 is considered an excellent value for indicating the consistency of five items in Ac factor. Therefore, all items were remained.

Table 6.32 Reliability Statistics for Accountability (Ac) factor

Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
.933	.933	5

Table 6.33 Item-Total Statistics for Accountability (Ac) factor

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Ac1	13.23	20.185	.901	.863	.902
Ac2	13.47	20.947	.771	.768	.927
Ac3	13.57	20.185	.814	.771	.919
Ac4	13.07	21.099	.791	.769	.923
Ac5	12.93	20.961	.837	.877	.915

6.6.3.10 Internal consistency for Auditability (Au) factor

The Cronbach’s α = .518 shown in Table 6.34 is considered an acceptable value for indicating the consistency of five items in Au factor. Therefore, all items were remained.

Table 6.34 Reliability Statistics for Auditability (Au) factor

Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
.518	.515	5

Table 6.35 Item-Total Statistics for Auditability (Au) factor

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Au1	12.79	18.241	.717	.588	.406
Au2	11.86	6.909	.546	.634	.265
Au3	11.93	5.709	.874	.806	.205
Au4	11.83	6.933	.753	.870	.163
Au5	11.72	6.635	.700	.850	.164

6.6.4 Discussion of the Validation Study

A careful consideration has been made to validate the instrument in two experiments; pre-test and validation study. The outcome of the experiments has shown that SecRaS instrument has provided an effective measure of the developed constructs.

1. Pre-test

The content validity result as presented in Table 6.4 shows the content validity ratio (CVR) for each factors and items in the instrument. On the scale of 0 to 1, the CVR value must be more than .50. This is to indicate that more than 50% of the experts agree with the suitability of the items and its importance to the instrument (Ayre and Scally 2014). Thus the instrument consist of nine factors and 52 items was evaluated. The results shown that from a pool of 52 items, only 43 are accepted; obtained an acceptable value of CVR.

2. Validation Study

Correlation analyses was conducted to analyse the strength of the relationship between factor to factor, item to item and the instrument as a whole. The result has shown that there is a significant relationship among the factors and among the items in each factors. As a whole the correlation result suggest a moderate and strong relationship. Therefore, it is considered to measure the underlying concept in Cloud Storage Security Framework (CSSF).

The internal reliability test was conducted to examine the internal consistency of the instrument as a whole, the factors, and the items for each factors. In previous research internal reliability was conducted as an indicator for stable and valid construct, especially newly developed instrument. The results have shown the factors to have a good internal consistency. Therefore, this result indicates that the items are measuring the same related concept. Thus it is concluded that the nine factors and 43 items have good internal reliability.

The items for the factors has shown a statistically significant measure for the measured content. Therefore, the final version of SecRaS instrument is concluded to have a reliable nine factors and 43 items.

6.7 Chapter Summary

In this chapter, a detail explanation was described on the development of a measuring instrument, Security Rating Score (SecRaS) using Cloud Storage Security Framework (CSSF) as a reference guide. The SecRaS instrument was developed following the Goal-Question-Metric (GQM) approach. The approach emphasises that all measurement should be goal-oriented and each measurement collected is stated in terms of the major goals. Questions are then derived from the goals to refine and determine if the goals can be achieved. The metrics measurements that are collected are then used to answer the questions in a quantifiable manner.

A study was later designed to validate the SecRaS instrument. Based on the CSSF, nine factors were selected and 52 items were generated for further consideration. Next, a pre-test that uses content validity ratio (CVR) was conducted with five experts to refine the instrument after which only 43 items remained in the revised instrument. Following the pre-test, a validation study with 30 security practitioners was conducted using the revised SecRaS. Data analysis was conducted using correlation analysis to examine the relationship between each item in a factor and the relationship between the factors and the instrument as a whole. Results suggest that the SecRaS has statistically significant correlations between items and factors and towards the instrument as a whole. Reliability analysis showed that SecRaS has good internal consistency reliability and may be used in a research scenario. Information gathered using SecRaS can be useful and insightful to inform the security practitioners on the security factors in cloud storage implementation. Next, Chapter 7 will demonstrate a study to establish the relationship(s) among security factors that affect the cloud security implementation in cloud storage. This study makes use of the validated SecRaS in a research scenario.

Chapter 7: Using Structural Equation Modelling to Establish the Relationship(s) Among Security Factors that Affects Security in Cloud Storage Implementation

This chapter will explain about structural equation modelling (SEM) that was carried out on the data obtained from the Security Rating Score (SecRaS) instrument. In answering the third research question, "What are the relationship(s) among security factors identified from Security Rating Score (SecRaS)?", this chapter aims to establish the relationship(s) among the security factors and identify which relationship(s) will affect security implementation in cloud storage using factor analysis and SEM.

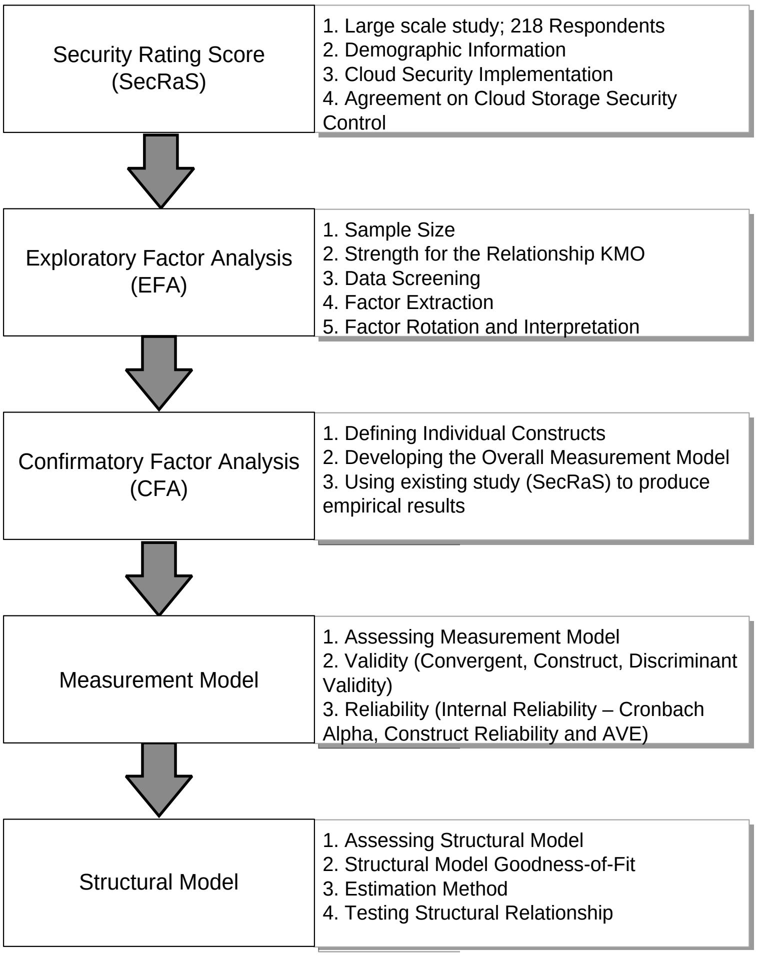

Security Rating Score (SecRaS)
1. Large scale study; 218 Respondents
2. Demographic Information
3. Cloud Security Implementation
4. Agreement on Cloud Storage Security Control
Exploratory Factor Analysis (EFA)
1. Sample Size
2. Strength for the Relationship KMO
3. Data Screening
4. Factor Extraction
5. Factor Rotation and Interpretation
Confirmatory Factor Analysis (CFA)
1. Defining Individual Constructs
2. Developing the Overall Measurement Model
3. Using existing study (SecRaS) to produce empirical results
Measurement Model
1. Assessing Measurement Model
2. Validity (Convergent, Construct, Discriminant Validity)
3. Reliability (Internal Reliability – Cronbach Alpha, Construct Reliability and AVE)
Structural Model
1. Assessing Structural Model
2. Structural Model Goodness-of-Fit
3. Estimation Method
4. Testing Structural Relationship

Figure 7.1 Steps to Establish the Relationship that Affects Security Implementation in Cloud Storage

The following sections are presented as follows, an introduction of SecRaS instrument that was developed and validated in Chapter 6, applying SecRaS in a large-scale study, analysis of the results; demographic and summarising the data using exploratory factor analysis (EFA). Later a comprehensive analysis using SEM; how Confirmatory Factor Analysis and SEM was used to measure how well the item indicators represent a construct. In the next section, the measurement model is tested, analysed and results are presented. The final section explains on testing the structural model, the analysis and results of the structural model. The diagram presented in Figure 7.1 shows the process involved in building a model using data obtained from SecRas in this chapter.

7.1 Security Rating Score (SecRaS)

The validated instrument in Chapter 6 has shown an acceptable value for each factors and therefore, SeCRaS was carried out in a larger sample. The population of the study only targets security experts and practitioners in ICT and preferably with cloud storage experience as the context of the study is from the view of stakeholders. The sampling is non-probability sampling, in which the participant responses are based on their willingness and availability (Gravetter and Forsano 2012). A large number of researcher have claimed that there is no fixed number for sample size but an adequate sample size is required in order to ensure the reliability of the study and allowing possibility of generalising the results from the data collection (Saunders et al. 2009, Hair et al. 2014). Selecting the sample size is also based on the preliminary test considerations (Saunders et al. 2009); in this case the results retrieved will be tested using Structural Equation Modelling (SEM). According to Kline (2005), SEM requires at least 200 sample in order to run adequately. Furthermore, the number of respondents of the study was established based on the observation that most published articles that have applied SEM as the analysis technique have at least 200 cases. The number of respondents in this study is 218.

The security variables in the study are inter-related, 43 questions regarding security policies and controls were given to security practitioners in Malaysia. The survey (or questionnaire) was posted online using iSurvey. iSurvey is survey generation and research tool for distributing online questionnaires provided by the University of Southampton. The survey link was distributed in cloud security groups in Malaysia (Linkedin and Facebook). The groups also include, Persatuan Juruanalisa Sistem Sektor Awam (PERJASA), which is the Information Technology (IT) officers group for the Government of Malaysia. Their experience and expertise in security will help the study to identify the significant aspect to protect data in cloud storage. Moreover, the 43 variables were only answered by security practitioners that have at least two years' experience in cloud security. Each variable for this analysis has a five point Likert-type scale; from strongly disagree (which is equal to one) to strongly agree (which is equal to five).

7.2 Analysis and Results of Security Rating Score (SecRaS)

This section presents the analysis and results of SecRaS instrument. First, missing data from the collected data is discussed and then the internal reliability of instrument are also shown in details. Next, the data is analysed for demographic information and results are presented. The IBM SPSS 22 software is used to perform the analysis.

7.2.1 Missing Data

Missing data is a crucial issue in data analysis stage when a questionnaire is the form of data collection method. Thus, before conducting data analysis, missing data must be resolved. A variety of methods are used in resolving missing data, but when considering the application of SEM, it is recommended to use the Listwise Deletion (LD), Pairwise Deletion (PD), Multiple Imputation (MI), and Full Information Maximum Likelihood (FIML) (Graham et al. 2007, Graham 2009, Hair et al. 2014). MI is a statistical technique that works by the process of replacing missing values with estimated values (Graham et al. 2007), whereas in the FIML technique, the process works by estimating parameters directly from the raw data for each individual. However, generating values for missing values by applying the MI method may lead to bias, and therefore invalid outcomes. The FIML methods provide more accurate results but due to their computational complexity and sensitivity (Enders and Bandalos 2001), this process is not applied in the present study. LD is a method in which any case that contains single or multiple missing data from the analysis is eliminated. This method may affect the sample if there are many missing items in relation to the data size which can result in reducing the statistical power (Hair et al. 2014). On the other hand, PD takes into consideration all the non missing data. Another consideration in selecting an approach is the specification of sample size; if the missing data are random, less than 10% of the observations, and factor loadings are high (0.7 or greater), then any approach are appropriate (Enders and Bandalos 2001). In this study, we have used the pairwise deletion (PD) to handle missing data issues as it provides fewer problems with convergence, factor loading estimates relatively free of bias and easy to implement using the AMOS program (Hair et al. 2014).

7.2.2 Reliability

The study applied a measure of construct reliability based on the Cronbach Alpha test. The overall Cronbach Alpha value is 0.939 and in Table 7.1, the Cronbach Alpha values for most constructs are shown between 0.8 and 0.9, which indicates very good internal consistency of items, whereas the Cronbach Alpha reliability of the constructs; integrity, accountability and auditability is above 0.9 which indicates excellent internal consistency (Pallant 2013).

Table 7.1 Reliability Analysis using Cronbach Alpha

Concept measured	Item used	Cronbach Alpha	Reliability Results
Cloud Storage Security (CS)	3	0.937	Excellent
Confidentiality (Co)	6	0.855	Very good
Integrity (In)	6	0.917	Excellent
Availability (Av)	4	0.893	Very good
Nonrepudiation (Nr)	4	0.872	Very good
Authenticity (At)	5	0.855	Very good
Reliability (Re)	5	0.893	Very good
Accountability (Ac)	5	0.938	Excellent
Auditability (Au)	5	0.901	Excellent

7.2.3 Demographic Data

Demographic background of the respondents is described in this section. Table 7.2 shows the demographic data of the participant responses. The demographic section includes four questions about the respondents' domain/sector, job/position, experience in ICT and the percentage of involvement in a cloud storage environment. The objective of this section is to focus on ICT security practitioners involved with a cloud storage environment.

Table 7.2 Demographic data

Questions	**Answer Options**	**Frequencies**	**Percentage**
Domain/Sector	Academic/Education	28	12.8
	Government	127	58.3
	Industry	45	20.6
	Others	18	8.3
Job/Position	IT Officer/Assistant IT Officer	136	62.4

Questions	Answer Options	Frequencies	Percentage
	IT Manager	18	8.3
	IT Technical	16	7.3
	Consultant	9	4.1
	Analyst/Expert	10	4.6
	Policy Maker	3	1.4
	Researcher	4	1.8
	Others	22	10.1
Experience in ICT security	2-5yrs	64	29.4
	6-10yrs	107	49.0
	more than 10 years	47	21.6
Percentage of involvement in cloud storage environment	0-25%	99	45.4
	26-50%	54	24.8
	51-75%	40	18.3
	76-100%	25	11.5

The majority respondents are from the government sector, 58.3 percent and the largest group of participant was IT officers/Assistant IT officers. The questionnaire also asked about the participant's experience in ICT security; the result shows that 29.4 percent of the respondent had at least two years of experience, 49 percent had over six years of experience and over 20 percent had over ten years of experience in ICT security. From this data, it can be concluded that the majority of respondents have good and well established knowledge in ICT security. To obtain further information about the respondents' involvement in a specific domain, respondents are asked of their involvement in cloud storage and the approximate percentage of involvement in a cloud storage environment. The results showed 99 respondents have none or less than 25 percent involvement in a cloud storage environment. Nevertheless, the majority number of respondents have been involved in a cloud storage environment (119 respondents in total).

7.2.4 Cloud Storage Security Policy, Procedures and Processes/Practice Implementation

The security policy, procedure & practice/process information is described in this section. Figure 7.2 shows the implementation of cloud security policy(s), procedures, and processes for cloud storage in their respective organisations. This section was intended to give the researcher a general knowledge about the participants' organisation's existing security policies in their existing ICT security policies targeting the cloud security policies and whether the security procedures and practices have been implemented.

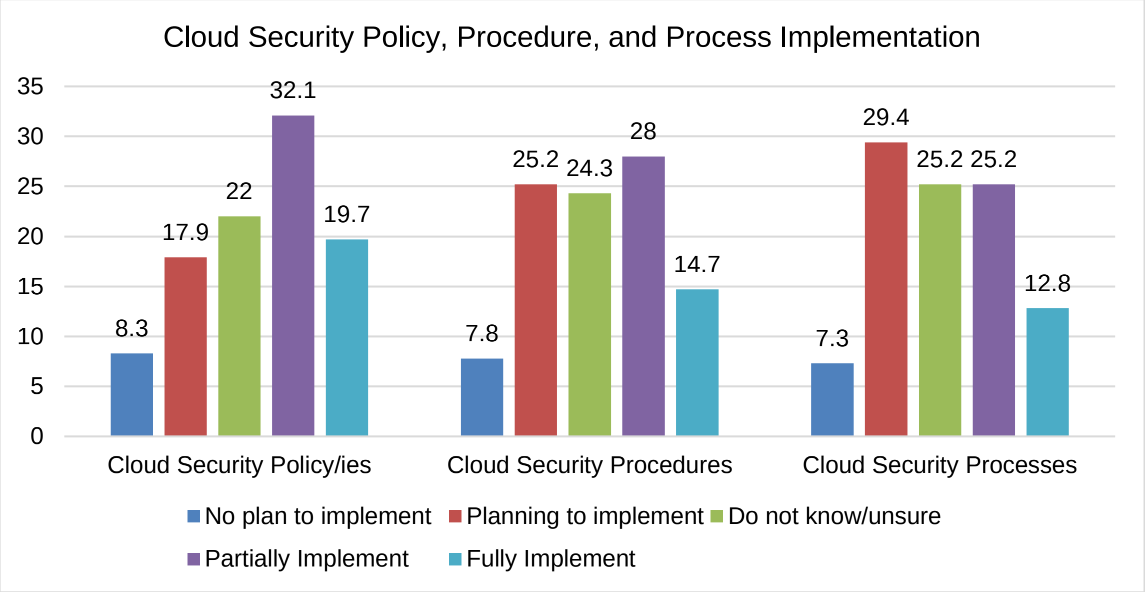


Figure 7.2 Cloud Security Implementation in Organisations

The results show that more than 50% or the respondents have cloud security in place in their organisations. Although most of the procedures may not be implemented, a high number of organisations have partially implemented the cloud security procedures and have cloud security process/practices planned for implementation in their organisations.

7.2.5 Security Measures in Cloud Storage

The final section of the questionnaire describes on the security in cloud storage, this part was designed to obtain knowledge about the security of cloud storage in their organisations. The questions were used to measure how much does an organisation follows the Cloud Storage Security Framework (CSSF). Questions were asked to what extent they agree with the statements associated with the security factors. Each factors have several items indicators. This section was concerned with the empirical measurements for the suggested factors and their item indicators. Overall, the results shows that there was a strong agreement on the measuring items for each of the security factors.

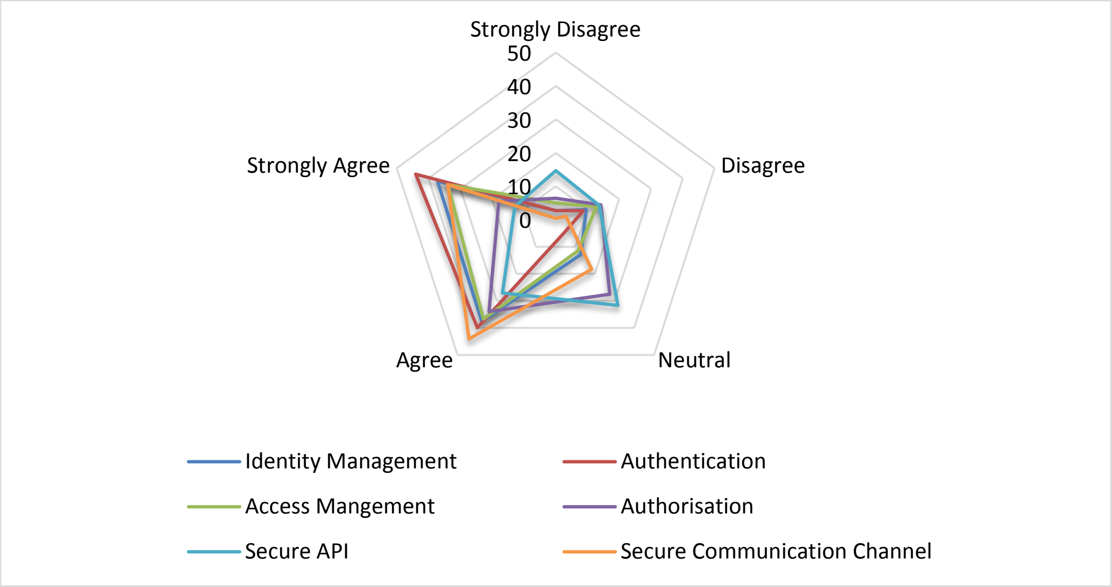


Figure 7.3 Confidentiality Factor

The practitioners agreed on the required security measures, it is predicted to perceive that most of the respondents strongly agree on the importance of security control implementations to protect data accessed in cloud storage. The controls includes; Authentication, Identity Management, Secure Communication Channel, and Access Management as shown in Figure 7.3 for the confidentiality factors. In addition, results shown in Figure 7.4 have shown how practitioners perceived in having controls to protect data stored in cloud storage. In short, the practitioners agrees on having key management, encryption, and protection for sensitive data.

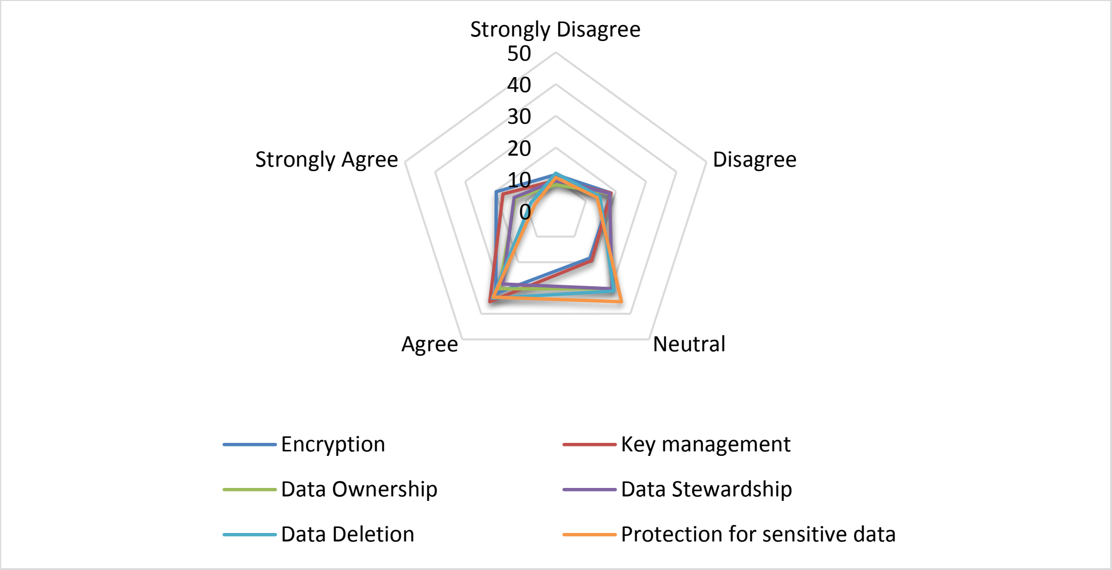


Figure 7.4 Integrity Factor

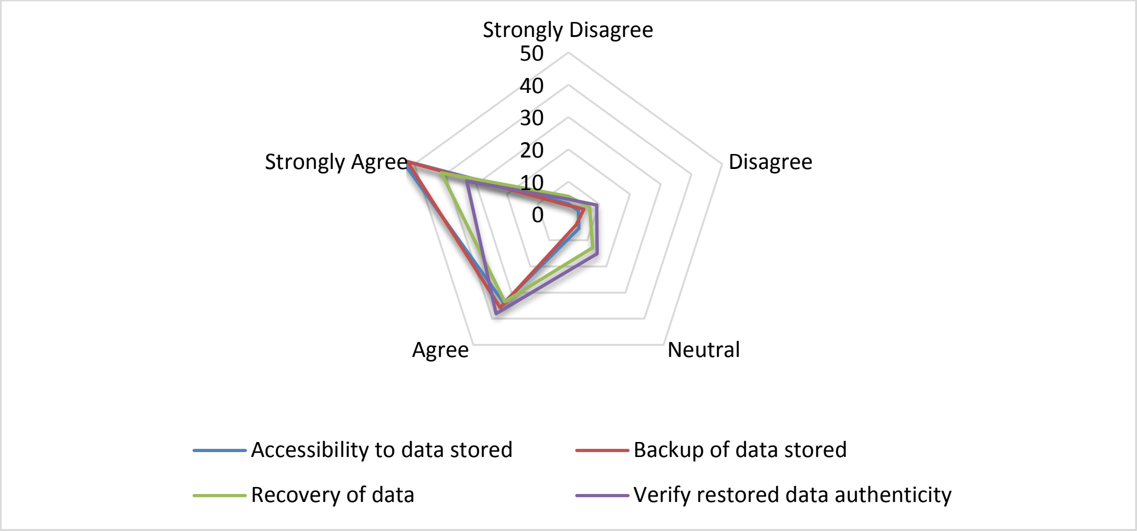


Figure 7.5 Availability Factor

In ensuring the availability of data, the majority of practitioners strongly agrees on accessibility of data stored and the backup of data stored in cloud storage; more than 50% supports both controls (Figure 7.5). The agreement on the security measures in cloud storage also includes the non-repudiation factors showing a neutral agreement on most of the items indicators; bind and validate identities, geographical location as authentication, time stamp and restriction of storage according to location. Although 20% of the respondents strongly disagrees that storage access should be restricted based on the geographical location, most of the other controls are being agreed. This shows an interesting response from the practitioners as having the restriction may affect the accessibility to the data in general (Figure 7.6). Restriction of access to data based on geographical location was discussed as efforts to track the provenance i.e. the chronology of the ownership, custody, or location of the data in cloud storage.

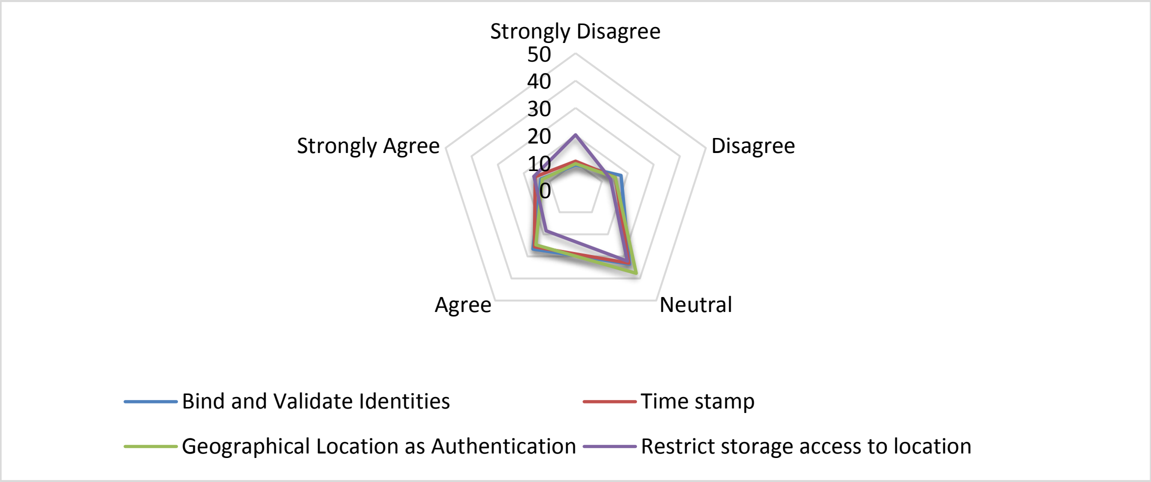


Figure 7.6 Non-repudiation Factor

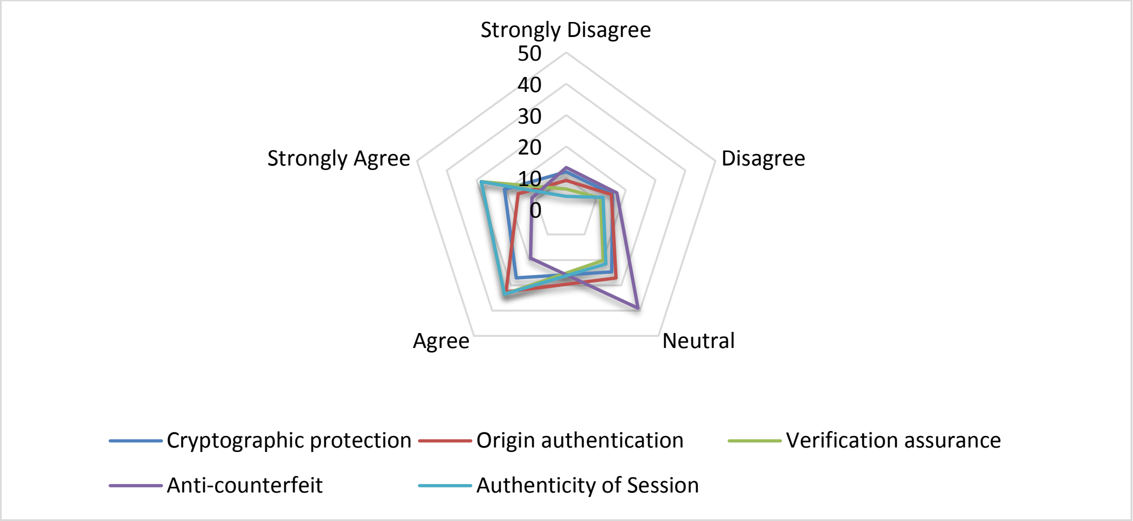


Figure 7.7 Authenticity Factor

As for the authenticity factor, most of the respondents agrees on the authenticity of data and verification assurance of data stored and accessed in cloud storage as shown in Figure 7.7.

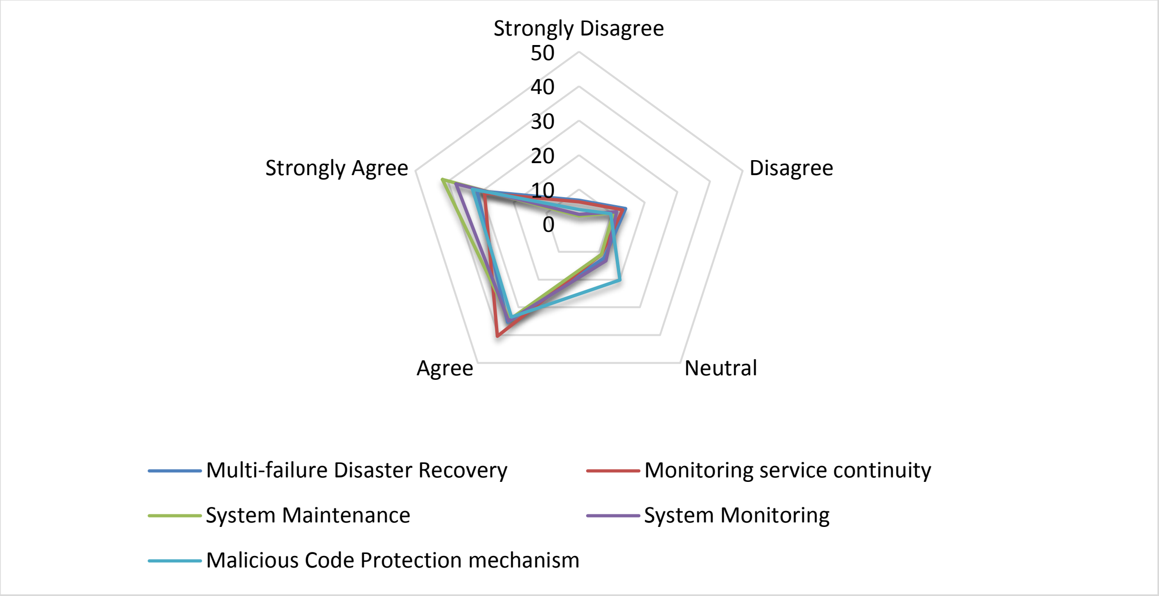


Figure 7.8 Reliability Factor

In terms of reliability of cloud storage, the respondents strongly agrees that system maintenance and system monitoring are among the important security measures to ensure service reliability. The service continuity has more than 40% of agreement although a small number of respondents disagrees Figure 7.8. In Accountability factors, there is a fair distribution of respondents between strongly agrees and agrees on conformance to external standards, security functionality, and security assurance. The majority of respondents have neutral ratings on the accountability measuring items as shown in Figure 7.9.

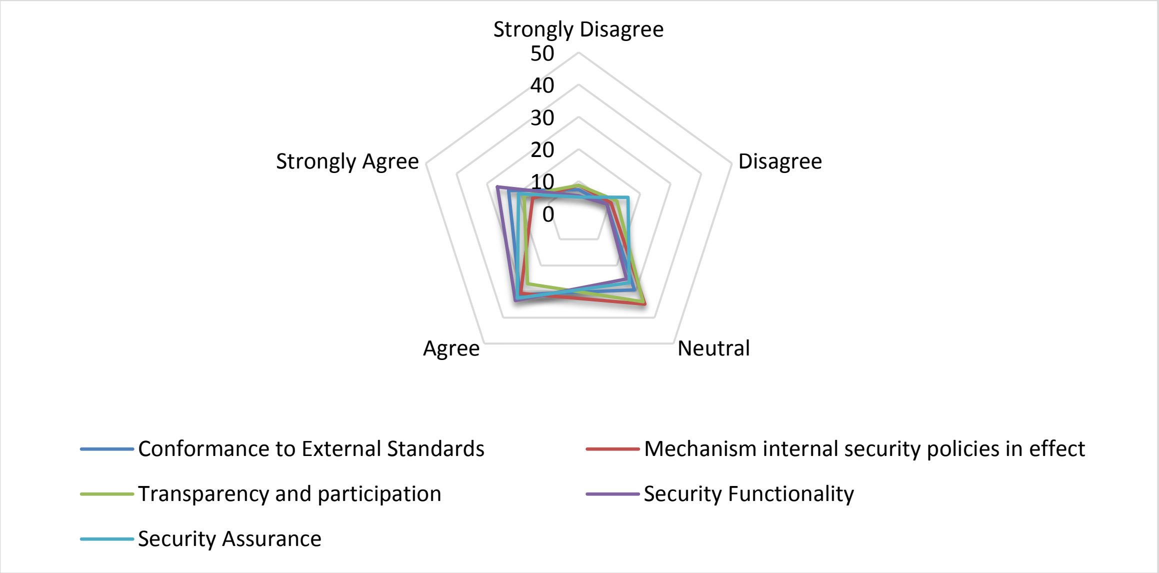


Figure 7.9 Accountability Factor

On the other hand, the Auditability factor shows strong agreement on Audit Policies, Audit Logs, and Report Generation with agreements of more than 40% for each items, presented in Figure 7.10.

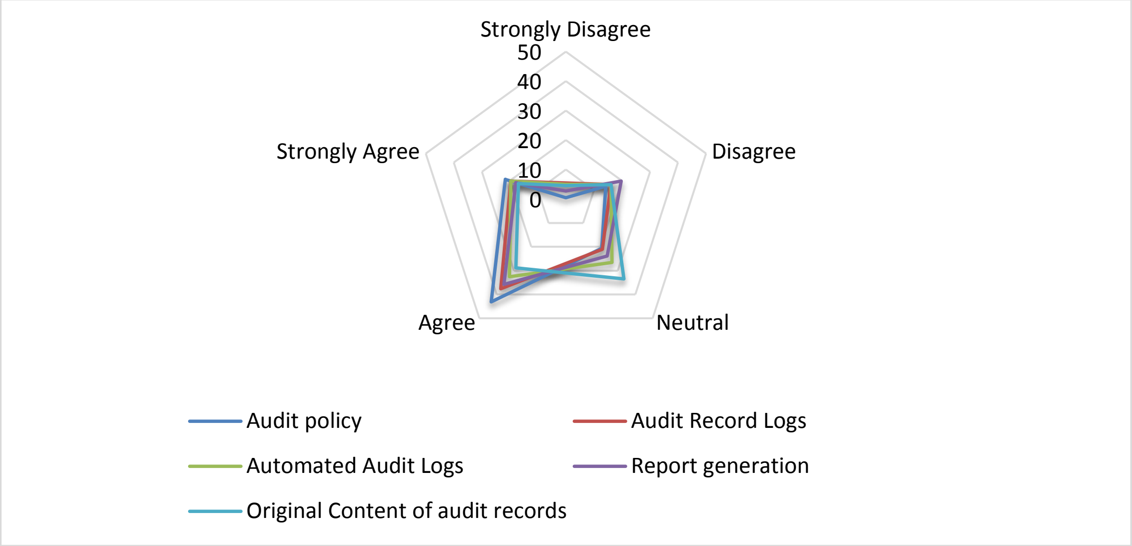


Figure 7.10 Auditability Factor

7.3 Factor Analysis

Factor analysis is carried out to explain the underlying structure among variables (Hair et al. 2014). In other words, factor analysis is used to regroup variables into a limited set of clusters based on shared variance. Hence, it helps to isolate constructs and concepts. The numerous observed variables have correlated patterns of responses since they are all related with a latent (i.e. not specifically measured) variable. The relationship with an underlying latent variable; the factor, which cannot be directly measured is assumed to be identified with various quantifiable variables. The general purpose of factor analysis is to summarise data so that relationships and patterns can be easily interpreted and understood (Gie Yong and Pearce 2013). There are several assumptions when applying factor analysis; linear relationship, no multicollinearity, includes relevant variables, and correlation between variables and factors. By applying factor analysis in this study, the security variables (or indicator items) are analysed to construct the security factors using factor analysis.

7.4 Exploratory Factor Analysis (EFA)

Exploratory Factor Analysis (EFA) is utilised to explore the underlining hypothetical structure and distinguish the structure of relationship between the variables (Cramer and Howitt 2004). EFA is also conducted to understand the measurements and significance of the variables from the survey (or questionnaire). In addition, EFA can help to provide a summary for data inter-relationship and place those variables into their groups accordingly. EFA tries to uncover patterns by exploring the dataset; in this study EFA will summarise the relationships between data retrieved from SecRaS and groups these variables accordingly in an unconstrained manner i.e. the items are not pre-specified to belong to any cluster/group prior the analysis. The IBM SPSS Statistics 22 software is utilised to perform EFA.

7.5 Using EFA to Understand the Underlying Relationship between Item Indicators and Constructs in SecRaS

The suitability and appropriateness of conducting EFA on the data need to be checked. There are several issues to consider when determining the suitability of the data; the sample size, the strength of the relationships among the variables (Kaiser-Meyer-Olkin (KMO) measure), data screening, factor extraction, factor rotation and interpretation, correlation, and the analysis of the factors. These are discussed below.

7.5.1 Sample Size

The reliability of factor analysis depends on the sample size. The common rule to apply to sample size is that a study has at least 10 to 15 participants per variable (Field 2013). It was also recommended to have between 5 and 10 participants per variable which add up to a total of 300 participants (Kass and Tinsley, 1979). A smaller sample size was also suggested to be sufficient if solutions have several variables with higher loading (above 0.80) (Tabachnick and Fidell 2007). In this analysis, the sample size is 218; most of the variables have loadings above 0.80 (see Table 7.5).

7.5.2 Strength for the Relationship using Kaiser-Meyer-Olkin (KMO) Measure

Another test to ensure the data is suitable for factor analysis is by observing the strength of inter-correlations among the variables. One of the statistical measures used to identify this is called Kaiser-Meyer-Olkin (KMO); a measure of sampling adequacy which ranges from 0 to 1. If the value yields more than 0.7, then the correlation on the whole are sufficient to perform factor analysis. Values between 0.5 and 0.7 are mediocre, values between 0.7 and 0.8 are good, values between 0.8 and 0.9 are great and lastly values above 0.9 are superb (Kaiser 1974). A KMO with 0.6 is suggested as the minimum value for a good factor analysis (Tabachnick and Fidell 2007). As measured from the sample, a KMO value of 0.856 was acquired from the data (Table 7.3). Therefore, it is justified that the factor analysis is suitable for these data sets.

Table 7.3 KMO and Bartlett’s Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.856
Bartlett's Test of Sphericity	Approx. Chi-Square	7818.674
	df	903
	Sig.	<0.001

7.5.3 Data screening

Before running the analysis, data was screened to remove any variables that should be excluded before the analysis is run. Some of the test includes detecting for outliers. Reasons to check for outliers include to eliminate the possibility of incorrect data entry, failure to specify missing values in the computer syntax so missing values are read as real data, identify if respondent is not the population intended to sample or if the outlier is representative of the population intended to sample but population has more extreme scores than a normal distribution (Tabachnick and Fidell 2007). Many statistical methods are sensitive to outliers so it is important to identify outliers and make decisions about what to do with them. Factor analysis can be sensitive to outliers, so as part of the preliminary data screening process, outliers are detected through extreme values; multivariate outlier is an extreme score on one or more variables (Pallant 2013). Another data screening involves observing the correlation matrix with all variables. The matrix will indicate which variables that do not correlate with any other variables or correlate very highly with other variables (r < 0.9) (Field 2013). None of the variables in this study fits the description therefore all the variables are included in the analysis.

7.5.4 Factor Extraction

Factor extraction is performed as one of the steps in factor analysis. It involves finding the minimum number of factors (or factors) that can be identified to best represent the interrelations among the set of variables. The most commonly used approach is principal components analysis (PCA) and common factor analysis. The selection of approach depends on the objective of factor analysis. In this analysis, the objective of factor analysis is for data summarisation and to define structure, therefore PCA is more suitable to summarise most of the original information (variance) in a minimum number of factors prediction (Hair et al. 2014).

In PCA, the techniques that can be used to assist in the decision concerning the number of factors to retain are Kaiser's criterion and Scree test. In order to determine how many numbers of factors are extracted, eigenvalues (or Kaiser criterion) and scree plot (Field 2013, Pallant 2013) are two sets of information that can be referred to.

7.5.4.1 Kaiser Criterion

The first method, eigenvalues or Kaiser's criterion will extract and retain the factors that have eigenvalues greater than 1 for further investigations. The eigenvalue of a factor represents the amount of the total variance explained by that factor. Table 7.4 summarises the factors that have eigenvalues greater than one (factor 1 to 9).

Table 7.4 Total Variance Explained

Factors	Eigenvalues (Total)	Eigenvalues (% of Variance)	Eigenvalues (Cumulative %)
1	12.485	29.034	29.034
2	3.690	8.582	37.616
3	2.917	6.783	44.399

Factors	Eigenvalues (Total)	Eigenvalues (% of Variance)	Eigenvalues (Cumulative %)
4	2.603	6.052	50.451
5	2.289	5.324	55.775
6	2.125	4.942	60.717
7	1.940	4.512	65.230
8	1.768	4.113	69.342
9	1.626	3.782	73.124
10	0.961	2.234	75.358
.	.	.	.
..	..	..	..
43	0.048	0.111	100.000

7.5.4.2 Scree Plot

On the other hand, using the scree plot, the point at which the curve changes direction and becomes horizontal suggest the number of factors. This involves plotting each of the eigenvalues of the factors and inspecting the plot to find a point at which the shape of the curve changes direction and becomes horizontal. It is recommended to retain all factors above the elbow, or break in the plot, as these factors contribute the most to the explanation of the variance in the data set. As described above, the scree plot suggests retaining only factors above this point (Figure 7.11).

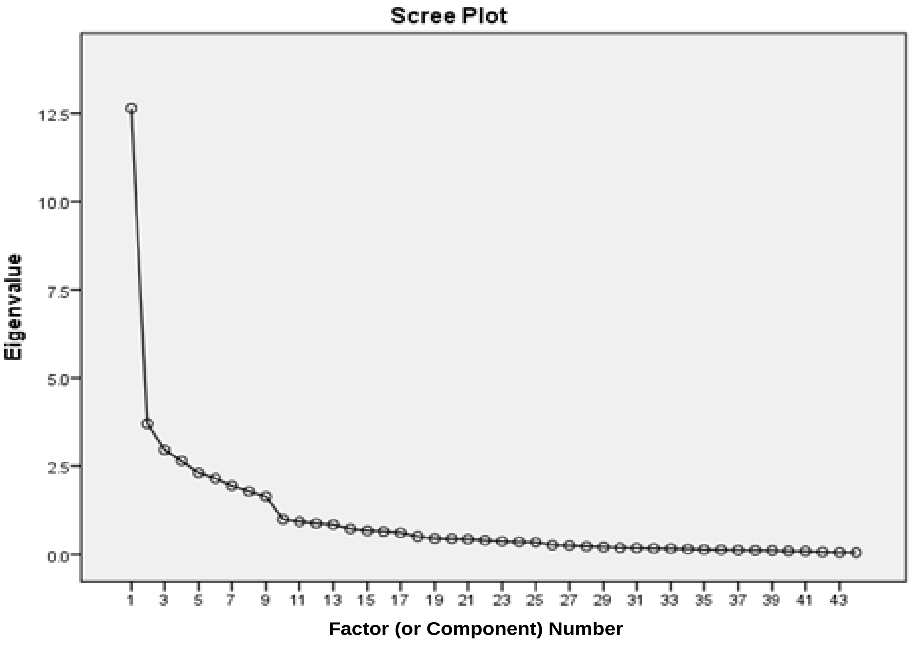


Figure 7.11 Factor Analysis Scree Plot

7.5.5 Factor Rotation and Interpretation

After the number of factors has been identified, the next step is to interpret the set of grouped variables. Factor rotation is useful to assist in this process. The factors are presented in the pattern of loadings in a manner that is easier to interpret. There are two techniques in rotating factors; orthogonal (varimax) and oblique (oblimin). In order to see which rotation technique is appropriate for our data, we tried both orthogonal and oblique techniques (Pallant 2013). In oblique rotation, the pattern matrix contains the factor loadings after the rotation while the structure matrix describes the relationship between the factors. The interpretation is mainly completed from the pattern matrix; however the structure matrix is useful for the purpose of double checking (Field 2013). Table 7.5 provides summary for the nine factors and their related indicators.

Table 7.5 Factor Loadings using orthogonal rotation

Factors (or Components)	Item Variables	Loadings
1	My organisation has cloud storage security policy in place	0.861
	My organisation has implemented security procedures to comply to industry standards	0.930
	My organisation has enforced security processes to support cloud security storage policy/ies	0.902
2	My organisation has identity management policy	0.685
	My organisation has user-based authentication process	0.798
	My organisation has access management policy	0.794
	My organisation has process to specify rights and restrictions for user access to data	0.793
	My organisation follows industry standards to build in security for systems	0.684
	My organisation has a secure communication channel policy	0.617
3	My organisation has documentation of encryption management practices/guidelines	0.789
	My organisation encrypt user data at rest (on disk/storage) within the environment	0.784
	My organisation leverage encryption to protect virtual machine images during transport between hypervisor instances	0.786
	My organisation supports data stewardship	0.763
	My organisation support secure deletion (e.g., degaussing/cryptographic wiping) of archived data	0.772
	My organisation supports additional protection for user to store sensitive data	0.818
4	My organisation segments data logically for each user	0.802
	My organisation has backup or redundancy mechanisms	0.800
	My organisation has data recovery mechanisms	0.830

Factors (or Components)	Item Variables	Loadings
	My organisation verifies data authenticity after restore process	0.823
5	My organisation has key management policies binding keys to identifiable owners	0.866
	My organisation uses synchronised time-service protocol (e.g., NTP etc.)	0.886
	My organisation uses geographical location as an authentication	0.860
	My organisation has restriction of user data to specific countries or geographic locations	*0.465
6	My organisation has cryptographic protection mechanisms	0.841
	My organisation ensures the origin authentication	0.821
	My organisation has verification assurances to ensure session authenticity	0.765
	My organisation has anti-counterfeiting policy	0.636
	My organisation provide session-level protection where needed	0.587
7	My organisation has multi-failure disaster recovery	0.714
	My organisation has system maintenance process/policy	0.741
	My organisation has patch management policy/process	0.830
	My organisation has continuous monitoring process/solutions	0.799
	My organisation has malicious code protection mechanisms at entry and exit points	0.728
8	My organisation has process to conformance with external standards	0.822
	My organisation has mechanisms to put internal security policies in effect	0.842
	My organisation supports transparency and participation to conformance process with internal standards/policy	0.846
	My organisation ensures clarity of Service Level Agreement/Guarantee (SLAs/SLGs)	0.840
	My organisation conducts penetration tests of cloud service infrastructure regularly as prescribed by industry best practices/guidance	0.840
9	My organisation produces audit assertions using a structured, industry accepted format (e.g., Cloud Audit/Cloud Trust, ISACA's Cloud Computing Management Audit, etc.)	0.601
	My organisation reviews audit logs on a regular basis	0.729
	My organisation has on-demand audit review	0.816
	My organisation generates audit report	0.865
	My organisation ensures the audit log is in original content or time	0.853

*loading below 0.5

Factor loadings with 0.5 and below is an indicator of which variables to drop if the KMO measure value is low (Woolford 2015). The factor loading below 0.5 is included in this study as the KMO measure value is more than 0.6. For a sample size of 200 a loading should be greater than 0.36 (Tabachnick and Fidell 2007, Field 2013, Gie Yong and Pearce 2013).

7.5.6 Analysis of the Factors

This section presents the analysis of factors obtained from the Security Rating Score (SecRaS) instrument. Figure 7.12 shows the nine construct retrieved from the 43 item indicators in SecRaS instrument.

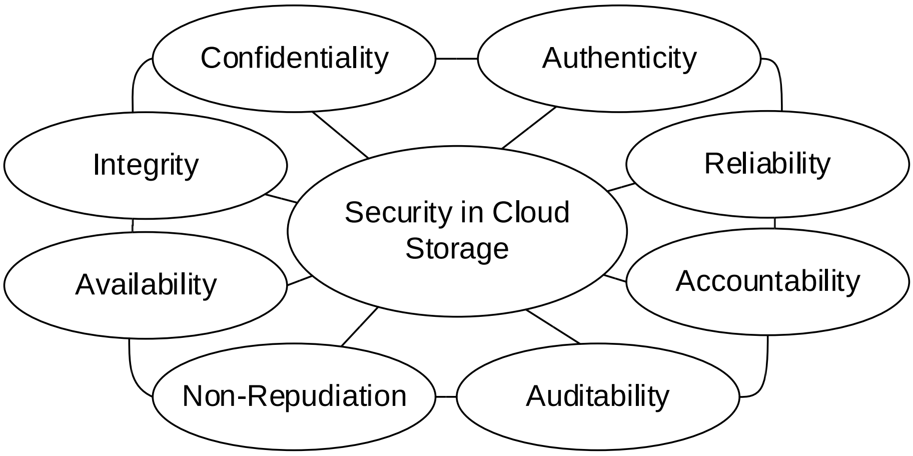


Figure 7.12 Nine Construct in Security Rating Score (SecRaS) Instrument

7.5.6.1 Factor 1: The Security Implementation for Protecting Data in Cloud Storage

This factor can be interpreted as the cloud security implementation in general. The indicator item with the highest loadings is security procedures implementations in an organisation. This factor looks into the policy implementation, procedure/process and security measures/practices in an organisation. With these loadings, this factor can be interpreted as the ‘**Security Implementation for Cloud Storage Aspects**’.

7.5.6.2 Factor 2: The Confidentiality of Data accessed in Cloud Storage

This factor shows the importance of security measures implementation for data accessed in cloud storage. The indicator items with highest loading are access management, authorisation and authentication. In general, the controls associated with this cluster describes securing data access to the data stored in cloud storage. Having these loadings, factor 2 is interpreted as '**Confidentiality Aspects**'.

7.5.6.3 Factor 3: The Integrity of Data stored in Cloud Storage

The second factor is loaded by six indicators. The highest loadings explain the important of data stewardship in a cloud storage service. Data stewardship involves the management of data. This is followed by the encryption mechanisms items that indicates the importance of having data encrypted at rest. From the loadings, it can be seen that practitioners are concern with the key management. It is considered reasonable to name these three loadings as '**Integrity Aspects'**.

7.5.6.4 Factor 4: The Availability of Data stored in Cloud Storage

The fourth factor resulting from factor analysis explained on the accessibility to the data aspect. A variable, '*My organisation has data recovery mechanisms are in place in case of a security event'* shows that the practitioners strongly agree that recovery of data ensures the availability of data stored in cloud storage. The rest of the loadings in this factor are clearly showing the importance of the availability of access to the data stored in cloud storage. This factor is interpreted as '**Availability Aspects**'.

7.5.6.5 Factor 5: The Non-repudiation of Data stored in Cloud Storage

The factor contains loadings that provide meanings about non-repudiation of data stored in cloud storage. The top items are time stamp, bind and validations between identities and geographical location as authenticating factor. Time stamp involves using a synchronised time-service/protocol (e.g., Network Time Protocol (NTP) etc.). The factor also explains that the practitioners agreed that binding keys to identifiable owners and support integration of location as an authentication factor (e.g. generation location-based cryptographic keys etc.). However, capability to restrict the storage of user data to specific countries or geographic locations has a loading of less than 0.5. Therefore, these loadings are interpreted as '**Non-repudiation Aspects'**.

7.5.6.6 Factor 6: The Authenticity of Data stored and accessed by authorised user in Cloud Storage

This factor has five loadings representing the importance of authenticity of data stored and accessed in cloud storage. The highest indicator is cryptographic protection, '*My organisation has cryptographic protection mechanisms (e.g. digital signatures, signed* *hashes using asymmetric cryptography, key to generate the hash, public key to verify the hash information etc.)'.* The loadings for the factor can be described as '**Authenticity Aspects**'.

7.5.6.7 Factor 7: The Reliability of Service provided by Cloud Storage

Factor seven is showing loadings about the reliability and consistency of cloud storage services. The loadings describe the importance of service continuity in cloud storage. This involves the disaster recovery, system maintenance and patch management, system monitoring and malicious protection. In this analysis, there is one loading in this factor ('*My organisation has Patch Management or System Maintenance policy'*) which is the highest indicator loading in the survey. The loadings demonstrate on consistency of services in cloud storage and can be described as the '**Reliability Aspects'**.

7.5.6.8 Factor 8: The Accountability of Service provided by Cloud Storage

This factor has five loadings in total. Three of the loadings which are describing the conformance with external and internal and the transparency of the responsibilities etc. The key loading- '*My organisation allows for transparency and external participation'* indicates the importance of clarity, or in other words – '*In my organisation, the clarity of Service Level Agreement/Guarantee (SLAs/SLGs) is emphasised'*. This reflects the importance of conformance with external, internal etc. responsibilities are vital. Besides that, the loading also indicates the importance of security functionality and security assurance. All of these have supported the interpretation of factor 8 '**Accountability Aspects**'.

7.5.6.9 Factor 9: The Auditability of Data stored and accessed in Cloud Storage

In the last factor, all the five loadings are describing mainly on the needs of having a well audit policy, audit log review and report generation in cases of security events. The indicator, *'My organisation does no alteration from the original content or time ordering of audit records'* has the highest agreement with 0.896 showing emphasis of the agreement that audit logs are valid and cannot be changed in terms of content and time of the security incident. Other loading are also high (greater than 0.8); automated audit logs and report generation. Therefore, these loadings are best to be described as '**Auditability Aspects**'.

The analysis of factors have summarised that when conducting the analysis in an unconstrained manner through EFA, a structure for data obtained from SecRaS was identified; 43 items indicators are clustered into nine constructs. The next section will test the structure identified in a constrained manner using CFA. CFA will validate how much the 43 item indicators explains the nine constructs through a measurement model.

7.6 Confirmatory Factor Analysis (CFA)

Confirmatory factor analysis (CFA) is a type of structural equation modelling (SEM) that deals specifically with measurement models, that is, the relationships between observed measures or indicators (e.g., test items, test scores, behavioural observation ratings) and latent variables or factors (Brown 2006). CFA approach will then validate the relationship of item indicators in the constructs. The result will test how well the measured variables represent the number of constructs in SecRaS. The IBM SPSS AMOS 22 software is used to perform CFA. AMOS is an additional module for SPSS that has the functionality to undertake CFA and Structural Equation Modelling (SEM).

7.7 Structural Equation Modelling (SEM)

Structural Equation Modelling (SEM) is a statistical analysis technique developed to analyse the inter-relationships among multiple variables in a conceptual model (Zainuddin 2012). It is a multivariate technique which combines both factor analysis and multiple regression analysis to simultaneously examine a series of interrelated dependence relationships among the latent and observed variables or between the latent variables themselves (Hair et al. 2014). Latent variables, which are known as factors, constructs or unobserved variables, are the hypothetical constructs of interest in a study which cannot be measured directly; these include attitude or knowledge whereas observed variables, which are known as latent indicators, manifesting variables or measuring items, are variables that can be measured directly using developed instruments or tests. Observed variables are used to define or infer latent variables.

SEM could be seen as a statistical methodology which takes a confirmatory approach rather than an exploratory to analyse a structural theory involving certain phenomenon (Byrne 2010). In the confirmatory approach, a researcher hypothesises a specific conceptual model, gathers data, and tests whether the data fit the model. In this model, the conceptual model is either confirmed of disconfirmed, based on a chi-square statistical test of significance and/or meeting acceptable model-fit criteria (Schumacker and Lomax 2010). SEM was used in the study in the measurement model and structural model (Hair et al. 2014). The measurement model analysis allows the researcher to evaluate how well observed variables logically and systematically represent hypothesised constructs (Hair et al. 2014). The measurement model is the primary step in SEM and without applying it the analysis will be inadequate (Kline 2005). Through measurement analysis, the researcher needs to verify the factor structure of a set of indicators and this allows the researcher to define the relationship between a set of measured variables and a set of latent variables (Suhr 2006). Moreover, verification of construct validation and construct reliability is completed through the measurement model (Hair et al. 2014). Besides testing a measurement model, SEM can also be used to test a structural model. The structural model is a conceptual representation of structural relationship(s) between constructs; expressed with a set of structural equations and typically illustrated with a diagram (Hair et al. 2014). In this research, the structural relationship between the construct (security factors) is represented empirically by structural estimates known as path estimate. The path estimates show how the structural model i.e. the causal model assumes that the relationships meet the conditions necessary for causation.

7.8 The Measurement Model

To determine whether the security rating score instrument (SecRaS) is a good model for measuring the security of cloud storage, it was tested using Structural Equation Modelling (SEM). CFA through SEM was used to validate the relationship between item indicators for constructs in SecRaS.

The process begins by defining individual constructs; listing constructs that will comprise the measurement model. Using the results obtained in SecRaS, an EFA analysis was carried out (described in the previous section) to define the underlying structure of the data. The pattern structure has identified that the 43 item indicators can be grouped into nine constructs. The defined constructs are cloud storage security, confidentiality, integrity, availability, non-repudiation, authenticity, reliability, accountability, and auditability.

The measurement model of the study was performed to nine construct or latent variables (unobserved variables) that were measured by 43 item indicators or measured variables (observed variables) in the proposed model. The latent variables and their indicators is shown in Table 7.6. Measured variables were concepts acquired from the initial study on Cloud Storage Security Framework (CSSF); previous studies on goal-driven security factors in cloud storage domain in chapter 4 and chapter 5, which were then used as a reference to develop the SecRaS instrument in chapter 6.

Table 7.6 Latent constructs and indicator variables

	Latent Variable	Items' Code	Items' used
1	Cloud Storage Security	CS	CS1, CS2, CS3
2	Confidentiality	Co	Co1, Co2, Co3, Co4, Co5, Co6
3	Integrity	In	In1, In2, In3, In4, In5, In6
4	Availability	Av	Av1, Av2, Av3, Av4
5	Non-repudiation	Nr	Nr1, Nr2, Nr3, Nr4
6	Authenticity	At	At1, At2, At3, At4, At5
7	Reliability	Re	Re1, Re2, Re3, Re4, Re5
8	Accountability	Ac	Ac1, Ac2, Ac3, Ac4, Ac5
9	Auditability	Au	Au1, Au2, Au3, Au4, Au5

Table 7.7 AMOS output for the hypothesised model summary statistics

Computation of degrees of freedom:	
Number of distinct sample moments	630
Number of distinct parameters to be estimated	106
Degrees of freedom (630-106)	524
Result:	
Chi-Square	872.313
Degree of freedom	524
Probability Level	<0.001

7.8.1 The Measurement Model Validity

Before assessing fit of the measurement model, it would be beneficial to look at the details related to SEM; validity, and reliability test used in this study.

7.8.1.1 Validity

Validity of a measurement model is the ability to measure what is supposed to measure for a construct in a measurement model (Zainuddin 2012). It includes convergent validity, construct validity and discriminant validity. The Table 7.8 presents an overview of the validity requirement for a measurement model. A detail description of each validity requirement will be explained in the next sub-section.

Table 7.8 Requirement for the validity of the measurement model

Type of validity	Requirement

Convergent validity	Average Variance Extracted (AVE) ≥ 0.5

Construct validity	Fitness index is achieved as:
	GFI ≥ 0.90; CFI ≥ 0.90;
	RMSEA ≤ 0.08;
	Chisq/*df* < 5.0

Discriminant validity	Free from redundant items Correlation coefficient between each pair of latent construct ≤ 0.85

7.8.1.1.1 Convergent Validity

Convergent validity assesses the degree to which two measures of the same concept are correlated (Hair et al. 2014). In other words, all the items that indicate a specific construct should share a high proportion of variance in common. It could be verified using factor loading value as a high value indicates that the item converge on the same latent factor. To get a high convergent validity, factor loadings should be statistically significant and the value should be 0.5 or higher. The best is more than 0.7 as the square of standardised factor loading represents how much variation in an item is explained by the latent factor. Or, it could also be verified through average variance extracted (AVE), and to achieve adequate convergence the value of AVE should be 0.5 or higher (Zainuddin 2012). This high correlation value indicated that the scale is measuring the intended concept. With CFA, the average variance explained (AVE) is calculated as the mean variance extracted for the items loading on a construct and is the summary indicator of convergence. This value can be calculated using standardised loadings (see equation 7.1):

$$AVE = \frac{\sum_{i=1}^{n} L_i^2}{n} \qquad 7.1$$

The $\boldsymbol{L_i}$ represents the standardised factor loading and $\boldsymbol{i}$ is the number of items. So for $\boldsymbol{n}$ items, AVE is computed as the total of all squared standardised factor loadings (squared multiple correlations) divided by the number of items. In other words, it is the squared completely standardised factor loadings or communality. Using the same logic, an AVE of 0.5 or higher is a good rule of thumb suggesting good convergence. An AVE of less than 0.5 indicates that on average more error remains in the items than variance explained by the latent factor structure imposed in the measure. An AVE measure should be computed for each latent construct in a measurement model.

7.8.1.1.2 Construct Validity

Construct validity is the extent to which a set of items actually reflect the theoretical latent construct those items are designed to measure (Hair et al. 2014). The validity could be verified through the fitness indices value (Zainuddin 2012). Construct validity defines how well a test or experiment measures up to its claims. It refers to whether the operational definition of a variable actually reflect the true theoretical meaning of a concept. In this study, construct validity refers to whether the data obtained from SecRaS instrument measures the construct adequately. An example is a measurement of the security of cloud storage measured through confidentiality, integrity, availability, non-repudiation, authenticity, reliability, accountability, and auditability.

7.8.1.1.3 Discriminant Validity

Discriminant validity means that individual measured items should represent only one latent construct (Hair et al. 2014, p. 778). It could be achieved when the measurement model is free from redundant items. This is checked from the modification indices value and then followed by the deletion of item or by setting the correlated pair of items as 'free parameter estimates'. And, to look for distinctiveness between constructs, the correlation between constructs should be less than 0.85 (Kline 2005). Discriminant validity is "extent to which a construct is truly distinct from other constructs both in term of how much it correlate with other constructs and how distinctly measured variables represent only this single construct" (Hair et al. 2014, p.662). Discriminant validity is measured by comparing the Average Variance Extracted value (AVE) for a construct with the square correlation estimate between the construct and another construct, in other words, comparing the square root of AVE with the correlation estimate between these constructs (Hair et al. 2014). To pass the discriminant validity test, the value of AVE for each construct is higher than the square correlation estimate between constructs (Hair et al. 2014). In Table 7.10, the AVE of construct factors: all AVE were higher than the square correlation estimate between the constructs. Therefore, this is sufficient evidence of discriminant validity of the constructs.

7.8.1.2 Reliability

Reliability of a measurement model is a measure of the degree to which a set of indicators of a latent construct is internally consistent in their measurements (Hair et al. 2014). A reliable model does not guarantee that the model is valid. However, reliability could be an indicator of convergent validity (Hair et al. 2014). In this study, three types of reliability were employed. They are internal reliability, construct reliability (CR) and average variance extracted (AVE).

Internal reliability refers to the degree to which all the items are measuring the same underlying construct (Pallant 2013). Internal reliability is measured by Cronbach's alpha coefficient and it is quite sensitive to the number of items measured. The study has calculated the reliability by using Cronbach alpha (α), in fact α can be used for estimating reliability only when the number of indicators are equally loaded on a constructs' variable or for the model underlying a single construct (Novick and Lewis 1966).

Construct reliability (CR) is intended to determine the consistency of the items representing a latent construct (Zainuddin 2012). To achieve CR, a value of square of total standardised loading divided by the sum of a square of total standardised loading and measurement error (Said et al., 2011) and the value should be 0.6 or more to represent the existence of internal consistency (Hair et al., 2006). Construct reliability or composite reliability (CR) or is often used in conjunction with SEM to examine the reliability of the construct and "*measure of reliability and internal consistency of the measured variables representing a latent construct*" (Hair et al. 2014, p.601). The next formula (see equation 7.2) was used to calculate CR (Hair et al. 2014):

$$CR = \frac{(\sum_{i=1}^{n} L_i)^2}{\left(\sum_{i=1}^{n} L_i\right)^2 + \left(\sum_{i=1}^{n} e_i\right)} \qquad 7.2$$

The equation above is based on standardised factor loading $\boldsymbol{L_i}$, where $\boldsymbol{n}$ is the number of items and there are $\boldsymbol{e_i}$ error variance terms for a construct. A reliability of between 0.6 and 0.7 is acceptable, but a good reliability is higher than 0.7, and internal consistency is increased with high reliability (Hair et al. 2014).

Average variance extracted (AVE) refers to the average percentage of variation explained by the items in a construct (Zainuddin 2012). A value of 0.5 or higher indicates the items share a high proportion of variance in common whereas a value less than 0.5 indicates that on average, more error remains in the items than variance explained by the latent factor structure imposed on the measure (Hair et al. 2014). The requirements of all the reliability are shown in Table 7.9.

Table 7.9 Requirement for the reliability of the measurement model

Type of reliability	Requirement
Internal reliability	Cronbach Alpha ≥ 0.7
Construct reliability (CR)	CR ≥ 0.6
Average Variance Extracted (AVE)	AVE ≥ 0.5

7.8.2 Assessing the CFA Measurement Model

In reviewing the validity and reliability of the measurement model, the incorporation of the error covariance or in other words, by allowing six measurement errors (e8-e9, e8-e10, e14-e15, e27-e28, e35-e36, and e38-e39), to be correlated had made a substantially large improvement to model fit.

Table 7.10 The CFA results for the measurement model

Construct	Item	Standardised Factor Loading (>0.5)	CR Alpha (>0.7)	CR (>0.6)	AVE (>0.5)
Cloud Storage Security (CS)	CS1	0.83	0.937	0.941	0.918
	CS2	0.97			
	CS3	0.94			
Confidentiality (Co)	Co1	0.70	0.855	0.860	0.714
	Co2	0.84			
	Co3	0.77			
	Co4	0.74			
	Co5	0.59			
	Co6	0.60			
Integrity (In)	In1	0.66	0.917	0.899	0.775
	In2	0.67			
	In3	0.66			
	In4	0.84			
	In5	0.83			
	In6	0.94			
Availability (Av)	Av1	0.68	0.893	0.878	0.805
	Av2	0.68			
	Av3	0.92			
	Av4	0.90			
Non-repudiation (Nr)	Nr1	0.90	0.872	0.902	0.844
	Nr2	0.99			
	Nr3	0.93			
	Nr4	0.44			
Authenticity (At)	At1	0.90	0.855	0.858	0.746
	At2	0.86			
	At3	0.78			
	At4	0.56			
	At5	0.57			
Reliability (Re)	Re1	0.60	0.893	0.880	0.775
	Re2	0.66			
	Re3	0.89			

Construct	Item	Standardised Factor Loading (>0.5)	CR Alpha (>0.7)	CR (>0.6)	AVE (>0.5)
	Re4	0.91			
	Re5	0.76			
Accountability (Ac)	Ac1	0.87	0.938	0.930	0.853
	Ac2	0.91			
	Ac3	0.88			
	Ac4	0.81			
	Ac5	0.78			
Auditability (Au)	Au1	0.60	0.901	0.898	0.802
	Au2	0.69			
	Au3	0.81			
	Au4	0.97			
	Au5	0.88			

Note: CR Alpha = Cronbach Alpha, CR = Construct Reliability, AVE = Average Variance Extracted

7.9 The Structural Model

Previous sections have presented verification of construct validity and composite reliability; hence, the structural model stage will assess the hypotheses that proposed the relationships among the construct variables that are represented as a causal path. Table 7.11 displays the hypotheses represented by the path’s estimation.

Table 7.11 Hypotheses assessed in structural model

Construct	Hypotheses	Hypothesised relationships
Confidentiality of Data Accessed (Co)	H1	Co → CS
Integrity of Data Stored (In)	H2	In → CS
Availability of Data Stored (Av)	H3	Av → CS
Non-repudiation of Data stored (Nr)	H4	Nr → CS
Authenticity of Data Accessed and Stored (At)	H5	At → CS
Reliability of Cloud Storage Services (Re)	H6	Re → CS
Accountability of Cloud Storage Services (Ac)	H7	Ac → CS
Auditability of Data Accessed and Stored (Au)	H8	Au → CS

Note: CS = Cloud Storage Security

7.9.1 Structural Model Goodness of Fit (GoF)

Goodness-of-fit indices are fit indices of SEM. There are many measures developed but only some will be utilised in this study, as shown in Table 7.12 (Kline 2005, Hooper et al. 2008, Hair et al. 2014). In general, there are three groups of measures: practical fit measures, absolute fit measures and incremental fit measures. It is suggested that the study should report at least three fit indexes with at least one from each category

(Hair et al. 2014). Chi-square statistics (X^2), also known as CMIN (minimum discrepancy), is a value representing the discrepancy between the unrestricted sample covariance matrix and the restricted covariance matrix (Byrne 2010). But, it has limitations. With large sample sizes, it is common to get a large X^2 value relative to degrees of freedom. So, large samples mean poor fit. However, with small samples, the Chi-square statistics lack power and yet could not discriminate between good and poor fitting models (Hooper et al. 2008). A fitness index developed to minimise the sensitivity of sample size on the Model Chi-Square is normed chi-square (X^2/df) and it is grouped under practical, subjective or ad hoc indices of fit. Normed chi-square is a value representing value of X^2 divided by the degrees of freedom resulting in a lower value (Kline 2005). And, the recommended value for a reasonable fit model is between 2.0 and 5.0. Next is an absolute fit index which measure of how well the hypothesised model fits in comparison with no other models (Byrne 2010). They estimate the proportion of variability in the sample covariance matrix explained by the predicted covariance matrix in the model (Kline 2005). Included in this category are GFI (Goodness-of-Fit), AGFI (Adjusted GFI) and RMSEA (Root Mean Square Error of Approximation) (Hooper et al. 2008). GFI is an index measuring the amount of covariance between the latent variables in the model (Kline 2005). For data with low factor loading and sample size, the cut-off is suggested to be 0.95 instead of 0.9 (Hooper et al. 2008). Furthermore, with a large number of degree of freedom compared to the number of sample size, its value tends to have a downward bias. Recently, this index is becoming less popular due to its sensitivity. It is even recommended not to be used. AGFI also adjusts the GFI based upon degrees of freedom in the specified model. There will be greater reduction of values of GFI when it comes to more complex models (Kline 2005). However, AGFI is becoming less popular nowadays as it is not well-performed in some computer simulation study.

RMSEA is an index measuring the discrepancy between the observed and estimated covariance matrices per degree of freedom (Hoe 2008). It is not sensitive to sample size as the discrepancy measured is in terms of the population and not the sample. However, it is quite sensitive to the number of estimated parameters in the model (Byrne, 2010) and it favours models with the lesser number of parameters (Hooper et al. 2008). Next are incremental or comparative fit indices, NFI (Normed Fit Index), CFI (Comparative Fit Index) and TLI (Tucker Lewis Index). These fit indices compare the chi square value of the hypothesised model against the chi square value of some standard, such as the null model. The null model assumes zero population co-variances among the observed variables (Kline 2005). In other words, the null model assumes that all measured variables are uncorrelated. Since incremental indices measure the increase in fit relative to a null model, higher value of the indices indicate larger improvement in fit (Lei and Wu 2007). CFI is actually a revised form of NFI. When choosing the index, CFI should be given priority compared to NFI, as suggested by Bentler in Byrne (2010) because NFI tends to underestimate the fit of a small sample sized model (Kline 2005) whereas CFI is the least effected by sample size (Hooper et al. 2008). With NFI, the fit of a sample size less than 200 will be underestimated. CFI is also the most widely used due to its relative insensitivity to model complexity (Hair et al. 2014). Furthermore, CFI provides a measure of complete co-variation in the data (Byrne 2010). With regards to which indices should be reported, Hooper et al. (2008) believe that CFI and RMSEA are preferred over other indices for they are the most insensitive to sample size, model misspecification and parameter estimates.

Table 7.12 Goodness of Fit Indices

Goodness-of-fit Index	Acceptable Value	Comments
Chi-square (X^2)	ρ > 0.05 (non-significant)	Indicates exact fit of the model. Value is sensitive to large sample size
Practical indices of fit:		
Normed chi-square (X^2/df)	[2.00, 5.00]	This is to reduce the sensitivity of X^2 to sample size $X^2/df < 3.0$: good fit
Absolute fit index:		
The Goodness-of-Fit Index(GFI)	[0.00, 1.00]	GFI = 1.00: perfect fit GFI > 0.9: good fit
The Adjusted GFI (AGFI)	[0.00, 1.00]	Values close to 1.00: good fit GFI > 0.9: good fit
Root Mean Square Error of Approximation (RMSEA)	RMSEA ≤ 0.08	RMSEA < 0.05: good fit RMSEA 0.05 - 0.08: adequate fit Values up to 0.10: poor fit
Incremental fit indices:		
Normed Fit Index (NFI)	NFI ≥ 0.90	NFI = 1.00: perfect fit Values close to 0.00: poor fit
Comparative Fit Index (CFI)	CFI ≥ 0.90	0.00 > CFI > 1.00 for acceptance
Tucker-Lewis Index (TLI)	TLI > 0.90	0.0 > TLI > 1.00 for acceptance

7.9.1.1 Estimation Method

In this study, maximum likelihood (ML) estimation method is used. ML is an estimation method used in generating parameter estimates of structural equation models. It is an estimation method that works simultaneously where estimates of model parameters for all variables are calculated all at once (Kline 2005). Being the most widely used, it is an iterative procedure that seeks to minimise a discrepancy between the model and the sample covariance. The Maximum Likelihood (ML) estimation technique was used to calculate the GoF indices using AMOS (version 22). The GoF statistics for the structural model are displayed in Table 7.13, and it is clear that the indices confirm that the model has a good fit with the observed data.

The magnitude of the factor loadings was substantially significant. For this model, it showed that X^2/df =1.650; GFI=0.794; AGFI=0.761; NFI=0.846; CFI=0.932; TLI=0.925 and RMSEA=0.055 indicating a good fit. The fit indices indicated that the nine-factor model schematically portrayed in Figure 7.13 was the most optimal model representing security of cloud storage implementation. In other words, the hypothesised nine-factor model fits the sample data. In addition, (Byrne 2010) asserted that assessment of model adequacy should be based on theoretical, statistical and practical considerations.

Table 7.13 Goodness of Fit indices for the structural model

Chi-square X^2 = 1163.372, ρ < 0.001	The Proposed Model fit indices	Model fit indices for sample size < 250 (Hair et al. 2014)
N	218	
Normed chi-square X^2/df	1.650	<3.00
RMSEA	0.055	<.08
CFI	0.932	≥ .900
RMR	0.092	< 0.1
Standardised RMR	0.0703	<.09

Note: *N = number of sample, df* = degree of freedom, X^2/df = normed chi-square or ratio of likelihood (X^2) to degrees of freedom, RMSEA = Root mean square error of approximation, CFI = Comparative fit index, RMR = Root Mean Square Residual

7.9.2 Assessing the Construct Relationships

Although it is confirmed above that there is a good fit between the proposed model and the observed data, a good fit alone is insufficient evidence to support the proposed structural model. Thus, the hypothetical relations among the construct variables will be assessed by examining the following variables: ρ-value, regression coefficients (standardised path coefficient β), Z- value and squared multiple correlations (R^2) (Hair et al. 2014).

ρ-value is used to evaluate how statistically significant the relationship is between measured variables and latent variables at the level 0.05. The standardised path coefficient for each variable indicates the size of its effect on the model: standardised path coefficients with values less than 0.1 indicate a small effect, while values larger than 0 indicate a large effect (Suhr 2008). Critical Ratio (CR, or t-value) refers to standard normal distribution; the t-value is computed through dividing the unstandardised regression coefficient by the standard error (SE). A coefficient value is considered significant at the .05 level (1.96 or higher, -1.96 or lower) (Hair et al. 2014). The squared multiple correlations (R^2) represent "the proportion of variance that is explained by the predictors of the variable in question" (Byrne 2010), so through the R^2 value the strength of the structural relation will be defined: for a stronger relationship between two variables, it is close to 1, whereas a value close to 0 indicates to a weak relationship.

Table 7.14 shows the standardised path coefficient and t-values for all hypotheses. The paths estimated for hypotheses H1, H2, and H6 were positive and statistically significant and exogenous variables have strong relationships with endogenous variables. Furthermore, ρ-values of hypotheses H1, H2 and H6 were above the critical value 0.05 with critical ratio values of 0.637, 0.655 and 0.772 respectively. The path estimated for hypothesis H3, H4, H5, H7, and H8 was below the critical t-value of Type I error, 0.05; also ρ-value was greater than 0.05, indicating a not statistically significant relationship with security implementation (CS) to protect data in cloud storage. Therefore, availability, non-repudiation, authenticity, accountability, and auditability have no direct relationship with security implementations (CS). By including the effects of the interacting variables, a larger proportion of the respective variances in security implementation ($R^2 = 0.56$) are accounted for according to Figure 7.13.

Table 7.14 Hypotheses analysis

Hypothesised Path	β	Critical Ratio (CR)	ρ
H1: Co → CS	0.32	6.37	< 0.001***
H2: In → CS	0.36	6.55	< 0.001***
H3: Av → CS	0.01	-3.58	0.732
H4: Nr → CS	0.02	-1.72	0.086
H5: At → CS	0.07	0.34	0.735
H6: Re → CS	0.82	7.72	< 0.001***
H7: Ac → CS	0.14	0.47	0.086
H8: Au → CS	0.02	0.23	0.735

Note: β = Standardised estimates, CR/t-value = Critical ratio, *** probability (ρ) < .001, CS = Cloud Storage Security, Co = Confidentiality, In = Integrity, Av = Availability, Nr = Non-

Hypothesised Path	β	Critical Ratio (CR)	ρ
repudiation, At = Authenticity, Re = Reliability, Ac = Accountability, Au = Auditability			

7.9.3 Assessment of Hypotheses

Path analysis was used in the study in examining the hypothesised relationship of the proposed model, through using the standardised path coefficients, as shown in the previous sections. This section will discuss in detail the proposed hypothesised relationships that have been tested and supported by the data.

H1: Confidentiality (Co) will positively affect the Security Implementation of Cloud Storage (CS) to protect data accessed in cloud storage. Confidentiality (Co) was found to have a significant direct influence and positive effect on security implementations to protect data in cloud storage: the standardised regression weight of CS (β), 0.32, with t-value of 6.37, suggests that the path between Co and CS is statistically significant at the $\rho<0.001$ level. Therefore, this result indicates strong support for the hypothesis (H1), as proposed in the conceptual model. From the result it is clear that a one standard deviation increase in confidentiality scores, is associated with increasing security implementation to protect data stored in a cloud storage, by 0.32 points (based on the standardised Beta coefficient value).

H2: Integrity (In) will positively affect the Security Implementation of Cloud Storage (CS) to protect data accessed in cloud storage. Integrity (In) was found to have a significant direct influence and positive effect on security implementations to protect data in cloud storage: the standardised regression weight of CS (β), 0.36, with t-value of 6.55, suggests that the path between In and CS is statistically significant at the $\rho<0.001$ level. Therefore, this result indicates strong support for the hypothesis (H2), as proposed in the conceptual model. From the result it is clear that a one standard deviation increase in integrity scores, is associated with increasing security implementation to protect data stored in a cloud storage, by 0.36 points (based on the standardised Beta coefficient value).

H3: Availability (Av) was found to have no direct effect on the security implementation of cloud storage, at a level of p-value = $0.732 > 0.05$. Thus hypothesis H3 is not supported.

H4: Non-repudiation (Nr) was found to have no direct effect on the security implementation of cloud storage, at a level of p-value = $0.086 > 0.05$. Thus hypothesis H4 is not supported.

H5: Authenticity (At) was found to have no direct effect on the security implementation of cloud storage, at a level of p-value = 0.735 > 0.05. Thus hypothesis H5 is not supported.

H6: Reliability (Re) will positively affect the Security Implementation of Cloud Storage (CS) to protect data accessed in cloud storage. Reliability (Re) was found to have a significant direct influence and positive effect on security implementations to protect data in cloud storage: the standardised regression weight of CS (β), 0.82, with t-value of 7.72, suggests that the path between Re and CS is statistically significant at the $\rho<0.001$ level. Therefore, this result indicates strong support for the hypothesis (H6), as proposed in the conceptual model. From the result it is clear that a one standard deviation increase in reliability scores, is associated with increasing security implementation to protect data stored in a cloud storage, by 0.82 points (based on the standardised Beta coefficient value).

H7: Accountability (Ac) was found to have no direct effect on the security implementation of cloud storage, at a level of p-value = 0.086 > 0.05. Thus hypothesis H7 is not supported.

H8: Auditability (Au) was found to have no direct effect on the security implementation of cloud storage, at a level of p-value = 0.735 > 0.05. Thus hypothesis H8 is not supported.

The structural model in Figure 7.13 consists of significant paths indicated by straight arrows connecting one factor to another. The figure also shows that 56% variance of security implementation in cloud storage is described by eight factors. The symbol of * indicates a significant path shown in the structural model. E1 – E9 represents the error/disturbance from each factor. Table 7.14 and Figure 7.13 can be represented by the following equations:

$$F1 = 0.32\ \text{Confidentiality}^{*} + E2 \qquad (1)$$

$$F2 = 0.36\ \text{Integrity}^{*} + E3 \qquad (2)$$

$$F3 = 0.01\ \text{Availability} + E4 \qquad (3)$$

$$F4 = 0.02\ \text{Non-repudiation} + E5 \qquad (4)$$

$$F5 = 0.07\ \text{Authenticity} + E6 \qquad (5)$$

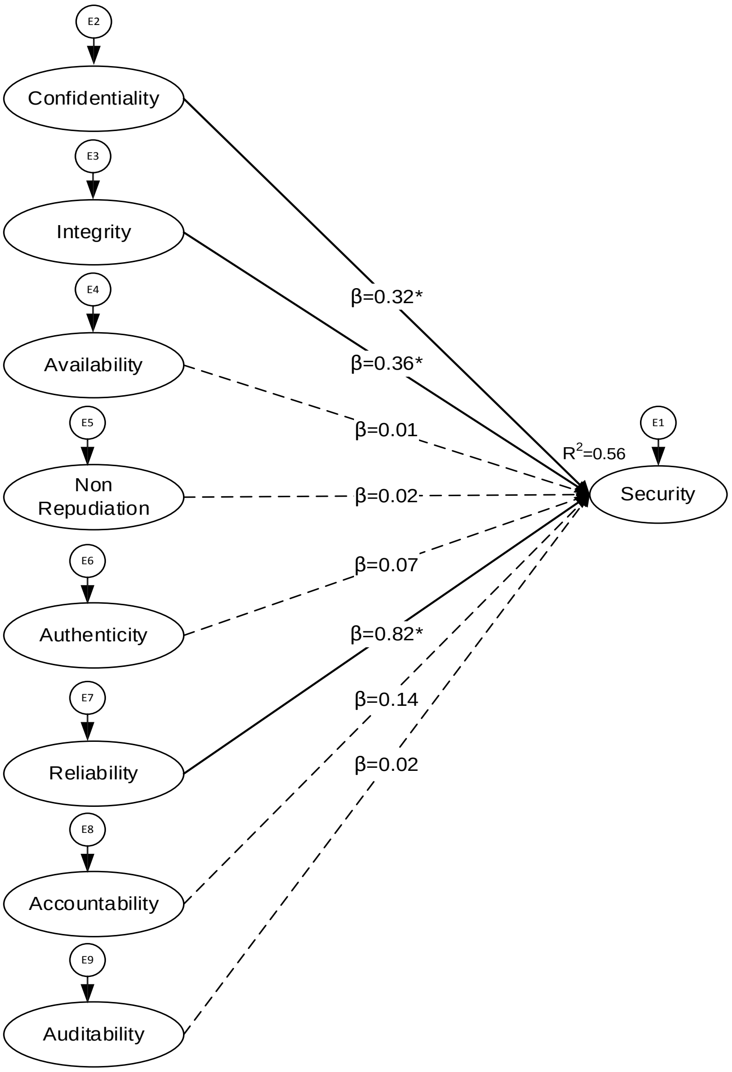
$$F6 = 0.82\ \text{Reliability}^{*} + E7 \qquad (6)$$

$$F7 = 0.14\ \text{Accountability} + E8 \qquad (7)$$

$$F8 = 0.02\ \text{Auditability} + E9 \qquad (8)$$

From equation F1 until F8, an estimated value for security implementation in cloud storage can be calculated using the following equation:

$$\text{Security} = F1^* + F2^* + F3 + F4 + F5 + F6^* + F7 + F8 + E1 \qquad (9)$$



Note:

⬭ represent a latent variables

Path coefficient
⟶ represents direct/significant effects
- - -> represents indirect/insignificant effects

β = standardised coefficients
R^2 = squared correlations
* ρ < 0.001
E = error/disturbance

For simplicity, the observed variables and their corresponding paths have been removed from the figure

Figure 7.13 Path Diagram of the Structural Model

The relationships for the factors are shown in the above equation. There are three significant paths in the model for relationships and hypotheses of:
H1: Confidentiality factor is positively related to the Security Implementation in Cloud Storage.
H2: Integrity factor is positively related to the Security Implementation in Cloud Storage.
H6: Reliability factor is positively related to the Security Implementation in Cloud Storage.

The rest of the relationships are shown to be insignificant.

7.10 Chapter Summary

In this chapter, the Security Rating Score (SecRaS) instrument was distributed to security practitioners in Malaysia and a number of 218 usable responses were obtained. The data was analysed using IBM SPSS 22 for demographic information. The study applied a measure of construct internal reliability based on the Cronbach’s Alpha test. The study has an overall Cronbach Alpha value of 0.939 and the Cronbach’s Alpha values for most constructs are between 0.8 and 0.9, which indicates very good internal consistency of items rating score. The study shows that the majority respondents are from IT officers in government sector. The result also shows more that 60 percent of the respondents have over six years of experience in ICT security and 50 percent are involved in a cloud storage environment. From this data, it can be concluded that the majority of respondents have a good and well-established knowledge in ICT security. More than 50 percent of respondents rated that their organisation have cloud security policy in place in their organisation although not all procedures and practices have been implemented. Nevertheless, most of the respondents rated that their organisation had planned for an implementation of cloud security practices and processes. The study also reveals that there is a strong agreement for cloud storage security implementations in eight factors (confidentiality, integrity, availability, non-repudiation, authenticity, accountability, and auditability).

Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) was carried out to test the data obtained from security rating score instrument (SecRaS). EFA was performed in an unconstrained manner whereas CFA was performed in a constrained manner. When EFA was performed, nine factors were extracted and retained for further investigation following the eigenvalue rules. Each cluster of factors was found to have high loadings (greater than 0.8). A Kaiser-Meyer-Olkin (KMO) statistical measure was carried out and resulted a value of more than 0.8 indicating a good value for factor analysis. Therefore, the sample of data has undergone initial considerations of suitability for performing factor analysis. When applying correlation between the extracted factors, the correlation matrix has shown a moderate relationship among most of the factors. After the rotation is performed, the indicators that were loaded into those nine factors are interpreted and they are defined as: (I) Factor 1: the security implementation to protect data in cloud storage, (II) Factor 2: the Confidentiality of Data accessed in Cloud Storage, (III) Factor 3: the Integrity of Data stored in Cloud Storage, (IV) Factor 4: the Availability of Data stored in Cloud Storage, (V) Factor 5: the Non-repudiation of Data stored in Cloud Storage, (VI) Factor 6: the Authenticity of Data stored and accessed by authorised user in Cloud Storage, (VII) Factor 7: the Reliability of Service provided by Cloud Storage, (VIII) Factor 8: the Accountability of Service provided by Cloud Storage and lastly for (IX) Factor 9: the Auditability of Data stored and accessed in Cloud Storage.

The next stage was performing Confirmatory Factor Analysis (CFA) to the data; a hypothesised measurement model has pre-specified 43 item indicators and nine constructs. Four main steps are undertaken in CFA; defining individual constructs, developing the measurement model, using existing study (SecRaS) to produce empirical result and assessing model validity. In assessing the model validity, several validity and reliability tests were taken into considerations such as convergent validity, construct reliability, discriminant validity and the reliability of the measurement model. After several error terms are co-varied; better indices values are obtained. The model has revealed fit indices value; X^2/df =1.650; GFI=0.794; AGFI=0.761; NFI=0.846; CFI=0.932; TLI=0.925 and RMSEA=0.055. The result has indicated that the model is a good fit with the dataset. In structural model analysis, the result has shown that there were direct effects among construct variables. Through standardised path coefficients it was found that there were positive influences towards security implementation in cloud storage between confidentiality, integrity, and reliability. In contrast, insignificant relationships were found between availability, non-repudiation, authenticity, accountability, and auditability towards security implementation in cloud storage. The main goal of conducting SEM is to provide predictive factors for organisation to consider a secure cloud storage implementation.

Chapter 8: Conclusion and Future Work

This chapter provides an overview of the study, presents conclusions drawn from results, and outlines the future research.

8.1 Conclusion

The cloud is an environment where users share the resources and to store their data online. To secure data in the cloud, many different security frameworks have been proposed. The existing security frameworks, industry accepted standards, and best practices have focused on securing the generic cloud implementation. A review of existing studies revealed a lack of insights and guidelines to secure data specifically in cloud storage facilities. There has been no study conducted on synthesising security factors specifically for cloud storage. Cloud storage has not been well explored. Many researchers have considered cloud storage within a generic cloud computing model. However, since cloud storage by necessity stores data outside the control of the data owner, this has raised specific security concerns. In order to understand security challenges for securing data in cloud storage, the security concerns are abstracted from the threat landscape at a higher level. As the result, security measures can be addressed more effectively by specifying the important security concerns. What is particularly missing in the existing literature is a research to identify the security factors, security measures affecting these factors, and the underlying relationships between the factors.

In order to understand more on the security factors and measures, a framework was developed by synthesising existing research, industry-accepted standards and best practices on cloud security and cloud storage. The framework, Cloud Storage Security Framework (CSSF) was further discussed with experts and practitioners before it was confirmed. Following the confirmed framework, an instrument was developed to measure how much organisation follows the framework. The instrument, Security Rating Score (SecRaS) can be applied to help stakeholders assess how much their organisations are following the framework. SecRaS was developed using goal-question-metrics approach and validated using a content validity ratio (CVR), the internal reliability and correlation analyses. The validated instrument was then distributed to the security practitioners in Malaysia whereby it discovered a three causal relationship model that shows a significant effect towards the implementation of security in cloud storage.

8.2 Cloud Storage Security Framework (CSSF): A summary

In this section, a summary of the development, confirmation and the applicability of Cloud Storage Security Framework (CSSF) is briefly explained.

8.2.1 Development of the framework

Initially, a synthesis process of assessing, combining, and comparing data discovered from a detailed literature review of security frameworks, industry-accepted standards, and technical white papers, was conducted in order to identify the security factors as shown in Table 2.2 and Table 2.3. Similar patterns were obtained from the synthesis and these were used as an initial set of factors presented in Table 4.4. The purpose of this investigation was to identify the factors needed to construct the Cloud Storage Security Framework (CSSF).

8.2.2 Confirmation of CSSF

The next stage involved a confirmation study that employed both qualitative and quantitative methods using the triangulation approach. Semi-structured interviews were conducted with 20 security experts to assess and explore if there were additional factors. The findings revealed that all the proposed factors are important. The experts suggested two additional factors (Accountability and Auditability). The final stage involved an online survey that was distributed to 34 security practitioners. The aim of the questionnaire was to confirm the recommended factors. The results showed that all the factors are statistically significant. The factors are confidentiality, integrity, availability, non-repudiation, authenticity, reliability, accountability, and auditability. As a result, a security framework for cloud storage was constructed based on the literature synthesis, recommendations from experts and supported by the practitioner's survey. The confirmed framework has nine security factors as shown in Table 5.7.

8.2.3 Application of CSSF

CSSF was used as a reference to build a measuring instrument, Security Rating Score (SecRaS). SecRaS is used to measure how much an organisation follows CSSF. The Goal-Question-Metrics (GQM) approach was used following the development of the instrument. Each factor in CSSF was constructed as a goal, and each goal has questions describing the metrics. Based on CSSF, nine factors were selected and 52 items were generated for further consideration. Next, a pre-test was conducted with five experts to refine the instrument after which only 43 items remained in the revised instrument. Following the pre-test, a validation study with 30 security practitioners was conducted using the revised SecRaS. Data analysis was conducted using correlation analysis to examine the relationship between each item in a factor and the relationship between the factors and the instrument as a whole. Results suggest that the SecRaS has statistically significant correlations between items and factors and towards the instrument as a whole. Reliability analysis showed that SecRaS has good internal consistency reliability. The refined SecRaS consists of nine factors with 43 items. These results suggest that SecRaS has the required level of reliability and may be used in a research scenario.

Information gathered using SecRaS can be useful and insightful to inform the security practitioners on the security factors in cloud storage implementation. Information gathered gives understanding for the causality of the security factors.

The study has suggested that for successfully implementing a secure cloud storage in an organisation, security measures that look into protecting (i) the confidentiality of data, (ii) the integrity of data, and (iii) the reliability of cloud storage provider should be considered which subsequently lead to secure cloud storage implementation.

Significant paths from the Structural Equation Modelling (SEM) can serve as guidelines to the managers to understand how each of the factors has an influence on each other in a secure cloud storage implementation.

8.3 Contributions of the research

Three main contributions were made by this research; the development of the framework that could address the security of data in cloud storage, the application of the framework by developing a measuring instrument and after validating the instrument, finally the application of the instrument is carried out in a research scenario. Results obtained from the large scale survey proposes a three way relationship model instead of the eight relationship model tested on the data.

8.3.1 A Framework: Cloud Storage Security Framework (CSSF)

The framework, CSSF is the main contribution of this research. The aspects that made this possible were:

- Incorporation of appropriate cloud computing and computer security literature in order to improve knowledge in cloud storage security that acts as the guideline
- Analysis of cloud storage threats
- A list of cloud storage security factors
- Introducing a potentially measurable security factors in cloud storage based on the security threats

8.3.2 An Instrument: Security Rating Score (SecRaS)

The instrument was developed to measure how much does organisation follows CSSF. A set of metrics was developed using a goal oriented approach. Each security factors have goals and metrics and a set of questions were developed to measure the metrics in a form of a questionnaire. The questionnaire was then validated with experts and practitioners. During the development and validation process, related contributions were:

- Analysis of suitable approach to instrument development
- Application of exploratory approach while following widely accepted methodologies for instrument development
- Integration of established methodologies for instrument validation
    - Content validity was undertaken whether the instrument measures the security concepts
    - Correlation analysis was used to look into the relationship of the questions items, and
    - Internal reliability looks into the consistency of the instrument.

8.3.3 A Security Model: Establishing the Relationship(s) among the Security Factors

Finally, SecRaS has provided some insights into its first generalisation in demonstrating applicability. Besides developing measuring instruments with different objectives and usages, more importantly this research demonstrated that SecRaS has the potential to be generalised in developing security tools across related domains. In addition to applying the instrument in a research scenario, this study further constructs a model consisting of factors that affects the security implementation for cloud storage. The model proposes three type of causal relationships in terms of how the security implementation in cloud storage could be affected by the security factors. The proposed relationship based on the structural model may be tested as hypotheses that can be explore in the future. In addition, the model also points out that the security factors have different strengths. Lastly this study represents one of the first efforts that attempt to theorise how stakeholders can make cloud security implementation successful in some depth. It contributes to the accumulation of knowledge in the area of organisation-wide IT innovation by systematically identifying patterns and constructs with details using quantitative data.

8.4 Implications

This research has made an effort to produce significant contributions to protecting data in cloud storage systems. Furthermore, the obtained research data and findings will give input to policy makers, practitioners and researchers. The implications of the findings and results with regards to methodological and practical are presented in this section.

8.4.1 For Security Managers

For security managers, CSSF opens a new paradigm on how stakeholders can make cloud security implementation successful in some depth i.e. the measurement to manage security improvements and to outline the appropriate way to design and produce a comprehensive security metrics. Insights discovered from CSSF have shown patterns that can be applied in their organisations. The SEM models from this study can help provide security managers with some checklists and predictions for accomplishing a secure cloud storage system implementation in their organisation.

8.4.2 For Security Practitioners

For security practitioners, the CSSF enables deconstruction of the concept of security in cloud storage into smaller, conceptually distinct and manageable factors to guide the design of security in cloud storage.

8.4.3 For Security Researchers

For researchers, the CSSF provides a common framework to conceptualise their research and make it easier to see how the security factors fit into the larger picture.

8.5 Limitations of the study

The proposed framework is not free of lacks. The framework makes the first attempt to integrate the goal driven security measures that take into accounts the value of confidentiality, integrity, availability, non-repudiation, authenticity, reliability accountability and auditability. Because the subjective factors of security and new findings outside the studied security environment, the specification could often experience changes. Therefore, security is not static.

The subjective factor has been reduced as much as possible but as in all human analysis, it is hard to remove it completely. Because security is related to the way it is perceived, there is no objective security. Thus most probably the same scenario analysed by other person leads to a slightly different specification.

8.6 Future Work

This research opens different directions and works in the field of cloud security management. While the framework provides new insights and benefits to the security implementation in cloud storage, this section presents a research and development plan designed to improve the understanding of security in cloud storage. The underlying goal is to help make sure that cloud storage security is developed systematically with scientific validation principles. The result of this plan will be a set of tools to improve approaches to securing security in cloud storage. Continuing research is aimed at developing an enhanced SecRaS. The plan is divided into five years and ten years milestone with specific research tasks to be achieved. Based upon the typical maturation span, the aim for a short term plan is after doctoral research i.e. the 5 year span. A long term goal is roughly 10 years. The plan of future work is described below:

- Short term plan (5 years)
    - Further validation of the measuring instrument
    - Initial prototype of an automated instrument tool
    - Refine measuring tool
    - Empirical results can show how the instrument supports multiple domains

- Long term plan (10 years)
    - Automated and fully adaptable measuring instrument tool

- Guidelines for designing security in cloud storage
- Systemise policies, measures, and standards through an automated software tool allowing a more accurate control and measurement.

8.6.1 An Automated Measuring Tool to Enable Evaluation of Cloud Storage Security

The first step in the measurement development process is to specify the constructs i.e. security factors to be measured. This requires analysis of the content domain into its constituent lists of items. The long-term goal is to develop an automated measuring tool, available to a wider community of academics and practitioners, and, in addition, for the measuring tool to be adaptable to different domains. Currently, work planned for the first three years has been completed and the Security Rating Score (SecRaS) was produced. A series of experiments was conducted to validate and demonstrate the first use of SecRaS in a research scenario. Further work is needed to prototype the measuring tool based on the SecRaS, then conduct experiments to refine and validate the measuring tool across different domains.

8.6.2 Designing Security Guidelines to Enable Integration of Cloud Storage Security

A focus of this research task is to develop guidelines, processes, methodologies, and tools to enable the integration of security in cloud storage. A key goal is to simplify the steps in designing security implementation in cloud storage. Designing tools are needed to simplify the security design processes with few intermediate steps and the ability to hide the underlying complexity of the security design work. Existing and emerging security models need to be re-examined, while experiments that demonstrate the use of these security tools are needed. At present, synthesis of the relevant literature in cloud storage and cloud security applications has been completed. Further work in this area is needed to produce a set of requirements so that an automated tool prototype can be developed. Then a series of experiments needs to be conducted to validate, refine and demonstrate the use of the automated tool. The long-term results of this research would produce security guidelines of strategies for cloud storage that can be applied to different domains.

8.7 Final Remarks

The researcher has found that although empirical research is challenging, it is very interesting and most importantly when applied correctly, valuable results are obtained. Finally, it was found that applying statistical analyses can be highly useful in empirical cloud security research; however finding the suitable statistical analysis for the collected data requires good understanding of the data itself, the goals of the research as well as the statistical test available to fulfill the specified goal.

List of References

Abu-Libdeh, H., Princehouse, L., and Weatherspoon, H., 2010. RACS: A Case for Cloud Storage Diversity. *Proc. of the 1st ACM Symposium on Cloud Computing*, 229–240.

AlZain, M. A., Soh, B., and Pardede, E., 2011. MCDB: Using Multi-Clouds to Ensure Security in Cloud Computing. *Proceedings - IEEE 9th International Conference on Dependable, Autonomic and Secure Computing, DASC 2011*, 784–791.

Amazon S3, 2016. *Amazon Simple Storage Service (S3) — Cloud Storage — AWS* [online]. Available from: https://aws.amazon.com/s3/ [Accessed 5 May 2016].

ASD, 2012. *Top Four Mitigation Strategies to Protect your ICT System* [online]. Australian Signals Directorate (ASD). Available from: http://www.asd.gov.au/publications/protect/Top_4_Mitigations.pdf [Accessed 22 Aug 2014].

ASD, 2014a. *Strategies to Mitigate Targeted Cyber Intrusions - Mitigation Details* [online]. Australian Signals Directorate (ASD). Available from: www.asd.gov.au/publications/Mitigation_Strategies_2011.pdf [Accessed 16 Aug 2014].

ASD, 2014b. *Australian Government Information Security Manual Controls* [online]. Australian Signals Directorate (ASD). Available from: http://www.asd.gov.au/publications/Information_Security_Manual_2014_Principles.pdf [Accessed 12 Oct 2014].

Ateniese, G., Burns, R., Curtmola, R., Herring, J., Kissner, L., Peterson, Z., and Song, D., 2007. Provable data possession at untrusted stores. *Proceedings of the 14th ACM conference on Computer and communications security - CCS '07*, (1), 598.

Ayre, C. and Scally, A. J., 2014. Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement & Evaluation in Counseling & Development (Sage Publications Inc.)* [online], 47 (1), 79–86. Available from: 10.1177/0748175613513808%5Cnhttp://offcampus.lib.washington.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=92969452&site=ehost-live.

Basili, V. R., 1992. Software Modeling and Measurement: The Goal/Question/Metric Paradigm. *Technical Report CS-TR- 2956, Department of Computer Science, University of Maryland, College Park, MD 20742*.

Basili, V. R., 1993. Applying the Goal/Question/Metric Paradigm in the Experience Factory. *Software Quality Assurance and Measurement A Worldwide Perspective*.

Basili, V. R., Caldiera, G., and Rombach, H. D., 1994. Goal Question Metric Paradigm. *Encyclopedia of Software Engineering*.

Bessani, A., Correia, M., Quaresma, B., Andre, F., and Sousa, P., 2011. DEPSKY : Dependable and Secure Storage in a Cloud-of-Clouds. *In*: *EuroSys'11 - Architecture*. 31–45.

Bhattachejee, A., 2012. *Social Science Research: Principles, Methods, and Practices*. Textbooks Collection Book 3, Global Text Project.

Bolderston, A., 2012. Conducting a Research Interview. *Journal of Medical Imaging and Radiation Sciences*, 43, 66–76.

Borgmann, M., Hahn, T., Herfert, M., Kunz, T., Richter, M., Viebeg, U., and Vowe, S., 2012. *On the Security of Cloud Storage Services. SIT Technical Reports*.

Bowers, K. D., Juels, A., and Oprea, A., 2009a. Proofs of Retrievability: Theory and Implementation. *CCSW '09 Proceedings of the 2009 ACM Workshop on Cloud Computing Security* [online], 43–54. Available from: http://doi.acm.org/10.1145/1655008.1655015%5Cnhttps://eprint.iacr.org/2008/175.

Bowers, K. D., Juels, A., and Oprea, A., 2009b. HAIL : A High-Availability and Integrity Layer for

Cloud Storage. *In*: *CCS*. 187–198.

Boxcryptor, 2016. *Encryption software to secure cloud files | Boxcryptor* [online]. Available from: https://www.boxcryptor.com/en/ [Accessed 5 May 2016].

Braun, V. and Clarke, V., 2006. Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, 3, 77–101.

Brock, M. and Goscinski, A., 2010. Toward a Framework for Cloud Security. *In*: *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 254–263.

Brown, T. A., 2006. *Confirmatory Factor Analysis for Applied Research*. The Guilford Press.

Brumley, D., 1999. Invisible Intruders: Rootkits in Practice. *Intrusion Detection Special Issue*, 9.

Buglione, L. and Abran, a, 2000. Balanced Scorecards and GQM: What are the Differences? *3rd European Software Measurement Conference, FESMA-AEMES 2000*, 18–20.

Byrne, B. M., 2010. *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. 2nd ed. Taylor and Francis Group.

Cachin, C. and Haas, R., 2010. Dependable Storage in the Intercloud. *IBM Research Report*, 1–6.

Caronni, G. and Waldvogel, M., 2003. Establishing Trust in Distributed Storage Providers. *Proceedings - 3rd International Conference on Peer-to-Peer Computing, P2P 2003*, 128–133.

Catteddu, D. and Hogben, G., 2009. *Cloud Computing: Benefits, Risks and Recommendations for Information Security White Paper*. European Network and Information Security Agency (ENISA).

Cohen, J., 1988. *Statistical Power Analysis for Behavioral Sciences (revised ed.)*. Second. Lawrence Erlbaum.

Cotten, S. R., Tashakkori, A., and Teddlie, C., 1999. Mixed Methodology: Combining Qualitative and Quantitative Approaches. *Contemporary Sociology*, 28, 752.

Covic, V., 2015. *Wuala Shutting Down, P2P Storage Service Recommends Tresorit* [online]. Available from: https://www.cloudwards.net/news/wuala-shutting-down-p2p-storage-service-recommends-tresorit-9405/ [Accessed 5 May 2016].

CPNI, 2014a. *The Critical Security Controls for Effective Cyber Defense V5.0 Report* [online]. Centre for the Protection of National Infrastructure (CPNI). Available from: http://www.cpni.gov.uk/documents/publications/2014/2014-04-11-critical-security-controls.pdf?epslanguage=en-gb [Accessed 22 Aug 2014].

CPNI, 2014b. *Reducing the Cyber Risk in 10 Critical Areas White Paper* [online]. Centre for the Protection of National Infrastructure (CPNI). Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/395716/10_steps_ten_critical_areas.pdf [Accessed 22 Aug 2014].

Cramer, D., 1998. *Fundamental Statistics for Social Research: Step-by-step Calculations and Computer Techniques Using SPSS for Windows*. Psychology Press.

Cramer, D. and Howitt, D., 2004. The SAGE Dictionary of statistics. *Statistics*, 188.

Creswell, J. W., 2003. *Research design - Qualitative, Quantitative and Mixed Methods Approaches*. Second Edi. SAGE Publication.

Creswell, J. W. and Clark, V. L., 2011. *Designing and Conducting Mixed Methods Research*. SAGE Publications.

Cronbach, L. J. and Shavelson, R. J., 2004. My Current Thoughts on Coefficient Alpha and Successor Procedures. *Educational and Psychological Measurement*, 64 (3), 391–418.

CSA, 2009. *Security Guidance for Critical Areas of Focus in Cloud Computing V2.1 White Paper* [online]. Cloud Security Alliance (CSA). Available from:

https://cloudsecurityalliance.org/csaguide.pdf [Accessed 22 Aug 2014].

CSA, 2010. *Top Threats to Cloud Computing V1.0 Report* [online]. Cloud Security Alliance (CSA). Available from: https://cloudsecurityalliance.org/topthreats/csathreats.v1.0.pdf [Accessed 22 Aug 2014].

CSA, 2013a. *Cloud Computing Vulnerability Incidents : A Statistical Overview Report* [online]. Cloud Security Alliance (CSA). Available from: https://cloudsecurityalliance.org/download/cloud-computing-vulnerability-incidents-a-statistical-overview/ [Accessed 22 Aug 2014].

CSA, 2013b. *The Notorious Nine: Cloud Computing Top Threats in 2013 Report* [online]. Cloud Security Alliance (CSA). Available from: https://downloads.cloudsecurityalliance.org/initiatives/top_threats/The_Notorious_Nine_Cloud_Computing_Top_Threats_in_2013.pdf [Accessed 22 Aug 2014].

CSA, 2013c. *The Cloud Control Matrix V3.0.1 White Paper* [online]. Cloud Security Alliance (CSA). Available from: https://cloudsecurityalliance.org/download/cloud-controls-matrix-v3-0-1 [Accessed 22 Aug 2014].

CSA, 2014. *Consensus Assessments Initiative (CAIQ)* [online]. [online]. Available from: https://cloudsecurityalliance.org/research/cai/.

Cyra, L. and Górski, J., 2008. Extending GQM by Arguement Structures. *Cee-Set 2007*, 44 (5), 26–39.

Deluca, D., Gallivan, M. J., and Kock, N., 2008. Furthering Information Systems Action Research : A Post-Positivist Synthesis of Four Dialectics. *Journal of the Association for Information Systems*, 9 (2), 48–72.

Dewan, H. and Hansdah, R. C., 2011. A Survey of Cloud Storage Facilities. *2011 IEEE World Congress on Services*, 224–231.

Doane, D. P. and Seward, L. E., 2011. Measuring Skewness : A Forgotten Statistic? *Journal of Statistics Education*, 19 (2), 1–18.

Dropbox, 2016. *Dropbox* [online]. Available from: https://www.dropbox.com/ [Accessed 5 May 2016].

duplicity, 2016. *duplicity: Main* [online]. Available from: http://duplicity.nongnu.org/ [Accessed 5 May 2016].

Dwivedi, Y. K., Choudrie, J., and Brinkman, W.-P., 2006. Development of a Survey Instrument to Examine Consumer Adoption of Broadband. *Industrial Management & Data Systems*, 106 (5).

El-Booz, S. A., Attiya, G., and El-Fishawy, N., 2016. A Secure Cloud Storage System Combining Time-based One Time Password and Automatic Blocker Protocol. *2015 11th International Computer Engineering Conference: Today Information Society What's Next?, ICENCO 2015* [online], 188–194. Available from: http://dx.doi.org/10.1186/s13635-016-0037-0.

EncFS, 2016. *Valient Gough | EncFS* [online]. Available from: http://www.arg0.net/encfs [Accessed 5 May 2016].

Enders, C. K. and Bandalos, D. L., 2001. The Relative Performance of Full Information Maximum Likelihood Estimation for Missing Data in Structural Equation Models Equation Models.

ENISA, 2009. *Glossary — ENISA* [online]. Available from: http://www.enisa.europa.eu/activities/risk-management/current-risk/risk-management-inventory/glossary.

Ertaul, L., Singhal, S., and Saldamli, G., 2010. Security Challenges in Cloud Computing. *Security & Management*, 36–42.

Field, A., 2013. *Discovering Statistics using IBM SPSS Statistics*. 4th ed. SAGE Publications.

Fink, A., 2003. *The Survey Handbook*. 2nd ed. SAGE Publications.

Firesmith, D., 2004. Specifying Reusable Security Requirements. *Journal of Object Technology*, 3 (1), 61–75.

Gartner, 2012. Newsroom: Consumers Will Store More Than a Third of Their Digital Content in the Cloud by 2016. *Press Release*.

Garver, M. and Mentzer, J., 1999. Logistics Research Methods: Employing Structural Equation Modeling to Test for Construct Validity. *Journal of Business Logistics*, 20 (1).

Gie Yong, A. and Pearce, S., 2013. A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutorials in Quantitative Methods for Psychology*, 9 (2), 79–94.

Golafshani, N., 2003. Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report*, 8, 597–606.

Google, 2016a. *Google Drive - Cloud Storage; File Backup for Photos, Docs & More* [online]. Available from: https://www.google.com/drive/ [Accessed 5 May 2016].

Google, 2016b. *Cloud Storage - Online Data Storage | Google Cloud Platform* [online]. Available from: https://cloud.google.com/storage/ [Accessed 5 May 2016].

Graf, S., 2014. Flexible Secure Cloud Storage.

Graf, S., Lang, P., Hohenadel, S. A., and Waldvogel, M., 2012. Versatile Key Management for Secure Cloud Storage. *Proceedings of the 31st IEEE Symposium on Reliable Distributed Systems (SRDS '12)* [online], 469–474. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6424897.

Graham, J. W., 2009. Missing Data Analysis: Making It Work in the Real World. *Annual Review Psychology* [online], 60, 549–576. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18652544.

Graham, J. W., Olchowski, A. E., and Gilreath, T. D., 2007. How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory. *Prevention Science*, 8 (3), 206–213.

Granneman, J., 2013. *IT Security Frameworks and Standards: Choosing the Right One* [online]. Available from: http://searchsecurity.techtarget.com/tip/IT-security-frameworks-and-standards-Choosing-the-right-one [Accessed 12 Aug 2015].

Gravetter, F. and Forsano, L., 2012. *Research Methods for the Bahavioral Sciences*. 6th ed. Belmont, CA: Wadsworth Cengage Learning.

Gruschka, N. and Jensen, M., 2010. Attack Surfaces: A Taxonomy for Attacks on Cloud Services. *2010 IEEE 3rd International Conference on Cloud Computing*, 276–279.

GTISC and GTRI, 2013. *Emerging Cyber Threats Report 2014* [online]. Georgia Tech Information Security Center (GTISC) and Georgia Tech Research Institute (GTRI), Georgia Tech Cyber Security Summit 2013. Available from: https://www.gtisc.gatech.edu/pdf/Threats_Report_2014.pdf [Accessed 22 Aug 2014].

Guest, G., Bunce, A., and Johnson, L., 2006. How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. *Field Methods*, 18 (1), 59–82.

Guion, L. a, Diehl, D. C., and Mcdonald, D., 2011. Triangulation : Establishing the Validity of Qualitative. *University of Florida/IFAS*, 2–4.

Habiba, U., Masood, R., Shibli, M., and Niazi, M., 2014. Cloud Identity Management Security Issues & Solutions: A Taxonomy. *Complex Adaptive Systems Modeling* [online], 2 (1), 1–37. Available from: http://dx.doi.org/10.1186/s40294-014-0005-9.

Haeberlen, A., 2010. A Case for the Accountable Cloud. *ACM SIGOPS Operating Systems*

Review, 44 (2), 52–57.

Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E., 2014. *Multivariate Data Analysis: A Global Perspective*. Seventh. Pearson New International Edition. Pearson New International Edition.

Hoe, S. L., 2008. Issues and Procedures in Adopting Structural Equation Modeling Technique. *Journal of Applied Quantitative Methods*, 3 (1), 76–83.

Honan, M., 2012. Kill the Password: Why a String of Characters Can’t Protect Us Anymore. *WIRED*, 9–16.

Honer, P., 2013. Cloud computing Security Requirements and Solutions: A Systematic Literature Review. *19th Twente Student Conference on IT, Enshede, The Netherlands*.

Hooper, D., Coughlan, J., and Mullen, M., 2008. Structural Equation Modelling: Guidelines for Determining Model Fit. *The Electronic Journal of Business Research Methods* [online], 6 (1), 53–60. Available from: www.ejbrm.com.

Huang, H. M., Rauch, U., and Liaw, S. S., 2010. Investigating Learners’ Attitudes toward Virtual Reality Learning Environments: Based on a Constructivist Approach. *Computers and Education* [online], 55 (3), 1171–1182. Available from: http://dx.doi.org/10.1016/j.compedu.2010.05.014.

ISECOM, 2001. *ISECOM - Open Source Security Testing Methodology Manual (OSSTMM)* [online]. Institute for Security and Open Methodologies (ISECOM). Available from: http://www.isecom.org/mirror/OSSTMM.3.pdf [Accessed 22 Aug 2014].

ISECT, 2014a. ISO / IEC 27017 — Information technology — Security techniques — Code of practice for information security controls based on ISO / IEC 27002 for cloud services (DRAFT), 3–5.

ISECT, 2014b. ISO / IEC 27018 : 2014 — Information technology — Security techniques — Code of practice for protection of Personally Identifiable Information (PII) in public clouds acting as PII processors, 27001–27003.

Islam, S. and Falcarin, P., 2011. Measuring Security Requirements for Software Security. *IEEE 10th International Conference in Cybernetic Intelligent System (CIS)* [online], 70–75. Available from: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6169137.

ISO/IEC, 2016. *ISO/IEC 27000:2016(en), Information technology — Security techniques — Information security management systems — Overview and vocabulary* [online]. Available from: https://www.iso.org/obp/ui/#iso:std:iso-iec:27000:ed-4:v1:en [Accessed 1 Jan 2017].

ISO/IEC JTC 1/SC 27, 2004. ISO/IEC 13335-1:2004 — Information technology — Security techniques— Management of information and communications technology security — Part 1: Concepts and models for information and communications technology security management. [online]. Available from: https://www.iso.org/standard/39066.html [Accessed 8 May 2015].

Jones, J. A., 2005. An Introduction to Factor Analysis of Information Risk (FAIR). *Risk Management Insight*, 1 (614).

Ju, J., Wu, J., Fu, J., and Lin, Z., 2011. A Survey on Cloud Storage. *Journal of Computers*, 6 (8).

Juels, A. and Jr, B. S. K., 2007. PORs : Proofs of Retrievability for Large Files. *CCS ’07, Alexandra, Virginia, USA*, 584–597.

Kaiser, H. F., 1974. An Index of Factorial Simplicity. *Psychometrika* [online], 39 (1), 31–36. Available from: http://dx.doi.org/10.1007/BF02291575.

Kamara, S. and Lauter, K., 2010. Cryptographic Cloud Storage. *In*: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 136–149.

Kamara, S., Papamanthou, C., and Roeder, T., 2011. CS2 : A Searchable Cryptographic Cloud Storage System. *Microsoft Research*, 1–25.

Kassou, M. and Kjiri, L., 2012. A Goal Question Metric Approach for Evaluating Security in a Service Oriented Architecture Context. *International Journal of Computer Science Issues*, Vol 9.

Kline, R. B., 2005. *Principles and Practice of Structural Equation Modeling*. 2nd ed.

Ko, R. K. L., Jagadpramana, P., Mowbray, M., Pearson, S., Kirchberg, M., Liang, Q., and Lee, B. S., 2011a. TrustCloud: A framework for accountability and trust in cloud computing. *In*: *Proceedings - 2011 IEEE World Congress on Services, SERVICES 2011*. 584–588.

Ko, R. K. L., Jagadpramana, P., Mowbray, M., Pearson, S., Kirchberg, M., Liang, Q., and Lee, B. S., 2011b. TrustCloud: A Framework for Accountability and Trust in Cloud Computing. *In*: *Proceedings - 2011 IEEE World Congress on Services, SERVICES 2011*. 584–588.

Lawshe, C. H., 1975. A Quantitative Approach to Content Validity. *Principles of personnel testing (2nd ed.). New York: McGraw-Hill.*, 28 (4), 563–75.

Lei, P.-W. and Wu, Q., 2007. Introduction to Structural Equation Modeling: Issues and Practical Considerations. *Educational Measurement: issues and practice*, 26 (3), 33–43.

Likert, R., 1932. A Technique for the Measurement of Attitudes. *Archives of Psychology*, 22 (140), 1–55.

Lynn, M. R., 1986. Determination and Quantification of Content Validity. *Nursing Research*, 35 (6), 382–385.

Ma, Q., Johnston, A. C., and Pearson, J. M., 2008. Information Security Management Objectives and Practices: A Parsimonious Framework. *Information Management & Computer Security*, 16 (3), 251–270.

Mahajan, P., Setty, S., Lee, S., Clement, A., Alvisi, L., Dahlin, M., and Walfish, M., 2011. Depot : Cloud Storage with Minimal Trust. *In*: *Proceedings of the 9th Symposium on Operating Systems Design and Implementation, Vancouver, Canada*.

Mapp, G., Aiash, M., Ondiege, B., and Clarke, M., 2014. Exploring a New Security Framework for Cloud Storage Using Capabilities. *In Proceedings - IEEE 8th International Symposium on Service Oriented System Engineering (SOSE)*, 484–489.

Marlinspike, M., 2009. More Tricks For Defeating SSL In Practice. *Black Hat USA*.

Martini, B. and Choo, K. K. R., 2013. Cloud Storage Forensics: OwnCloud as a Case Study. *Digital Investigation*, 10, 287–299.

Mather, T., Kumaraswamy, S., and Latif, S., 2009. *Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance*. International Journal Of Policy And Administration. O'Reilly Media, Inc.

Mell, P. and Grance, T., 2009. The NIST Definition of Cloud Computing. *National Institute of Standards and Technology*, 53 (6).

Merkle, R. C., 1988. A Digital Signature Based on a Conventional Encryption Function. *Crypto*, 10.

Mertens, D. M., 2010. Publishing Mixed Methods Research. *Journal of Mixed Methods Research*, 5, 3–6.

Microsoft, 2015. *Security Threats* [online]. Microsoft Developer Network (MSDN). Available from: https://msdn.microsoft.com/en-us/library/cc723507.aspx [Accessed 22 Apr 2015].

Microsoft, 2016a. *Microsoft OneDrive* [online]. Available from: https://onedrive.live.com/about/en-us/ [Accessed 5 May 2016].

Microsoft, 2016b. *Microsoft Azure: Cloud Computing Platform & Services* [online]. Available from: https://azure.microsoft.com/en-us/ [Accessed 5 May 2016].

Milenkoski, A., Iosup, A., Kounev, S., Sachs, K., Ding, J., and Rosenberg, F., 2013. Cloud Usage Patterns : A Formalism for Description of Cloud Usage Scenarios. *Tech Report SPEC-RG-2013-001 v1.0.1, SPEC Research Group*, 12–13.

Miller, R., 2013. How Dropbox Stores Stuff for 200 Million Users. *Data Center Knowledge*, 2013–2016.

Mouratidis, H., Islam, S., Kalloniatis, C., and Gritzalis, S., 2013. A Framework to Support Selection of Cloud Providers Based on Security and Privacy Requirements. *Journal of Systems and Software*, 86 (9), 2276–2293.

Mu, S., Chen, K., Gao, P., Ye, F., Wu, Y., and Zheng, W., 2012. μlibCloud: Providing High Available and Uniform Accessing to Multiple Cloud Storages. *Proceedings - IEEE/ACM International Workshop on Grid Computing*, 201–208.

Myagmar, S., Lee, A. J., and Yurcik, W., 2005. Threat Modeling as a Basis for Security Requirements. *In Proceedings of the 2005 ACM Workshop on Storage Security and Survivability (StorageSS '05)*, 94–102.

Na, S., Kim, K., and Huh, E., 2013. A Methodology for Evaluating Cloud Computing Security Service-Level Agreements. *International Journal of Advancements in Computing Technology*, 5 (13), 235–242.

Nepal, S., Chen, S., Yao, J., and Thilakanathan, D., 2011. DIaaS: Data Integrity as a Service in the Cloud. *Proceedings - 2011 IEEE 4th International Conference on Cloud Computing, CLOUD 2011*, 308–315.

NIST, 2004. *Standards for Security Categorization of Federal Information and Information Systems*. National Institute of Standards and Technology (NIST) Special Publication FIPS 199.

NIST, 2013a. Security and Privacy Controls for Federal Information Systems and Organizations. *National Institute of Standards and Technology (NIST) Special Publication 800-53 Revision 4*.

NIST, 2013b. NIST Cloud Computing Security Reference Architecture. *National Institute of Standards and Technology (NIST) Special Publication 500-299*, 1–204.

Novick, M. R. and Lewis, C., 1966. *Coefficient Alpha and the Reliability of Composite Measurements*. Technical Report Number One Office of Naval Research Contract Nonr-4866(Oo) Project Designation NR 042-249.

OwnCloud, 2016. *ownCloud.org* [online]. Available from: https://owncloud.org/ [Accessed 5 May 2016].

Pallant, J., 2013. *SPSS Survival Manual: A Step by Step Guide to Data Analysis using IBM SPSS*. Third Edit. Allen & Unwin.

Pearson, S., Tountopoulos, V., Catteddu, D., Sudholt, M., Molva, R., Reich, C., Fischer-Hubner, S., Millard, C., Lotz, V., Jaatun, M. G., Leenes, R., Rong, C., and Lopez, J., 2012. Accountability for Cloud and Other Future Internet Services. *4th IEEE International Conference on Cloud Computing Technology and Science Proceedings*, 629–632.

Pohlman, M., 2010. Using the CSA Control Matrix and ISO 27017 Controls to Facilitate Regulatory Compliance in the Cloud. *Cloud Security Alliance* [online]. Available from: http://docbox.etsi.org/workshop/2012/201201_SECURITYWORKSHOP/3_INTERNATIONAL_STANDARDIZATION/EMC_CSA_POHLMANN.pdf.

Polit, D. F. and Beck, C. T., 2006. The Content Validity Index: Are You Sure You Know What's Being Reported? Critique and Recomendations. *Research in nursing & health*, 29, 489–497.

Popa, R., Lorch, J., and Molnar, D., 2011. Enabling Security in Cloud Storage SLAs with CloudProof. *Proc. USENIX* [online], 355–368. Available from: http://dl.acm.org/citation.cfm?id=2002181.2002212%5Cnhttp://www.usenix.org/events/atc11/tech/final_files/Popa.pdf%5Cnhttp://static.usenix.org/event/atc11/tech/final_files/Popa.p df.

Razali, N. M. and Wah, Y. B., 2011. Power Comparisons of Shapiro-Wilk , Kolmogorov-Smirnov , Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, 2 (1), 21–33.

Rouse, M., 2005. *What is storage?* [online]. TechTarget. Available from: http://searchstorage.techtarget.com/definition/storage [Accessed 12 Mar 2015].

Rubin, H. and Rubin, I., 2005. *Qualitative Interviewing : The Art of Hearing Data*. 2nd ed. Thousand Oaks, CA: SAGE.

Rubio, D., Berg-Weger, M., Tebb, S. S., Lee, E. S., and Rauch, S., 2003. Objectifying Content Validity: Conducting a Content Validity Study in Social Work Research. *Social Work Research*, 27 (June 2015), 94–104.

Ryan, M. D., 2013. Cloud Computing Security: The Scientific Challenge, and a Survey of Solutions. *Journal of Systems and Software*, 86 (9), 2263–2268.

Sabahi, F., 2011. Cloud Computing Security Threats and Responses. *2011 IEEE 3rd International Conference on Communication Software and Networks*, 245–249.

Satran, J., Meth, K., Sapuntzakis, C., and Chadalapaka, M., 2004. Internet Small Computer Systems Interface (iSCSI). [online]. Available from: http://www.ietf.org/rfc/rfc3720.txt [Accessed 5 May 2016].

Saunders, M., Lewis, P., and Thornhill, A., 2009. *Research Methods for Business Students*. Business. Pearson Education Limited.

Schumacker, R. E. and Lomax, R. G., 2010. *A Beginner's Guide to Structural Equation Modeling*. 3rd ed. Taylor and Francis Group.

Sekaran, U., 2000. *Research Methods for Business*. Research methods for business. John Wiley & Sons, Inc.

Sekaran, U., 2003. *Research Methods for Business A Skill-Building Approach*. Fourth. John Wiley & Sons.

Shaikh, F. B. and Haider, S., 2011. Security Threats in Cloud Computing. *6th International Conference on Internet Technology and Secured Transactions, Abu Dhabi, UAE*, (December), 11–14.

Shapiro, S. S. and Wilk, M. B., 1965. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52 (3/4), 591–611.

Shraer, A., Cachin, C., and Cidon, A., 2010. Venus: Verification for Untrusted Cloud Storage. *Workshop on Cloud* [online], 19–29. Available from: http://dl.acm.org/citation.cfm?id=1866841.

Singh, A. and Chatterjee, K., 2015. Identity Management in Cloud Computing through Claim-Based Solution. *2015 Fifth International Conference on Advanced Computing & Communication Technologies* [online], 524–529. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7079139.

Singh, R., Kumar, S., and Agrahari, S. K., 2012. Ensuring Data Storage Security in Cloud Computing. *IOSR Journal of Engineering*, 2 (12), 17–21.

SparkleShare, 2016. *SparkleShare - Self hosted, instant, secure file sync* [online]. Available from: https://www.sparkleshare.org/ [Accessed 5 May 2016].

SpiderOak, 2016. *SpiderOak* [online]. Available from: www.spideroak.com [Accessed 5 May 2016].

Srinivasan, M. K. and Rodrigues, P., 2012. State-of-the-art Cloud Computing Security Taxonomies - A Classification of Security Challenges in the Present Cloud. *In*: *ICACCI '12*. 470–476.

Stefanov, E. and Dijk, M. Van, 2012. Iris: A Scalable Cloud File System with Efficient Integrity Checks. *ACSAC '12 Proceedings of the 28th Annual Computer Security Applications Conference* [online], 229–238. Available from: http://stst.elia.pub.ro/news/so/os_base/Storage/DISK/ECC/585.pdf.

Straub, D., Boudreau, M.-C., and Gefen, D., 2004. Validation Guidelines for Is Positivist. *Communications of the Association for Information Systems*, 13 (24), 380–427.

Subashini, S. and Kavitha, V., 2011. A Survey on Security Issues in Service Delivery Models of Cloud Computing. *Journal of Network and Computer Applications*, 34 (1), 1–11.

Suhr, D., 2006. Exploratory or Confirmatory Factor Analysis? *Statistics and Data analysis* [online], 1–17. Available from: http://140.112.142.232/~PurpleWoo/Literature/!DataAnalysis/FactorAnalysis_SAS.com_200-31.pdf.

Suhr, D., 2008. Step Your Way through Path Analysis. *Western Users of SAS Software Conference* [online], 1–10. Available from: http://wuss.org/proceedings08/08WUSS Proceedings/papers/pos/pos04.pdf.

Suntharam, V. S., Reddy, K. V, and Puspalatha, N., 2013. Data Storage Security in Cloud Computing and Verification of Metadata by Encryption. *International Journal of Computer Science and Electronics Engineering*, 2 (3).

Swiderski, F. and Snyder, W., 2004. *Threat Modeling*. Microsoft Press.

Tabachnick, B. G. and Fidell, L. S., 2007. Multivariate Analysis of Variance and Covariance. *Using Multivariate Statistics*, 3, 402–407.

Takabi, H., Joshi, J. B. D., and Ahn, G. J., 2010. SecureCloud: Towards a Comprehensive Security Framework for Cloud Computing Environments. *In*: *Proceedings - International Computer Software and Applications Conference*. 393–398.

Tawalbeh, L., Darwazeh, N. S., Al-Qassas, R. S., and AlDosari, F., 2015. A Secure Cloud Computing Model Based on Data Classification. *Procedia Computer Science* [online], 52 (1), 1153–1158. Available from: http://dx.doi.org/10.1016/j.procs.2015.05.150.

TeamDrive, 2016. *TeamDrive.com | Sync your data fast and securely* [online]. Available from: https://www.teamdrive.com/en/ [Accessed 5 May 2016].

Teddlie, C. and Tashakkori, A., 2010. Overview of Contemporary Issues in Mixed Methods Research. *In*: *Sage handbook of mixed methods in social & behavioral research*. 1–44.

The Open Group, 2009. *Risk Taxonomy* [online]. Available from: http://pubs.opengroup.org/onlinepubs/9699919899/toc.pdf.

Thurmond, V. A., 2001. The Point of Triangulation. *Journal of nursing scholarship : an official publication of Sigma Theta Tau International Honor Society of Nursing / Sigma Theta Tau*, 33, 253–258.

TrueCrypt, 2016. *TrueCrypt* [online]. Available from: http://truecrypt.sourceforge.net/ [Accessed 5 May 2016].

Vaquero, L. M., Rodero-Merino, L., Caceras, J., and Lindner, M., 2009. A Break in the Clouds: Towards a Cloud Definition. *ACM SIGCOMM Computer Communication Review*, 39 (1), 50–55.

Vrable, M., Savage, S., and Voelker, G. M. G., 2012. Bluesky: A Cloud-Backed File System for the Enterprise. *Fast '12* [online], 19. Available from: http://cseweb.ucsd.edu/~voelker/pubs/bluesky-fast12.pdf%5Cnhttp://static.usenix.org/event/fast12/tech/full_papers/Vrable.pdf%5Cnhttp://dl.acm.org/citation.cfm?id=2208461.2208480.

Vu, Q. H., Colombo, M., Asal, R., Sajjad, A., El-Moussa, F. A., and Dimitrakos, T., 2015. Secure Cloud Storage: A Framework for Data Protection as a Service in the Multi-Cloud Environment. *2015 IEEE Conference on Communications and NetworkSecurity, CNS*

2015, 638–642.

Wang, C., Chow, S. S. M., Wang, Q., Ren, K., and Lou, W., 2013. Privacy-Preserving Public Auditing for Secure Cloud Storage. *IEEE Transactions on Computers*, 62 (2), 362–375.

Wang, C., Ren, K., Lou, W., and Li, J., 2010. Toward Publicly Auditable Secure Cloud Data Storage Services. *IEEE Network*, 24 (4), 19–24.

Wang, C., Wang, Q., Ren, K., Cao, N., and Lou, W., 2012. Toward Secure and Dependable Storage Services in Cloud Computing. *IEEE Transactions on Services Computing*, 5, 220–232.

Wang, Q., Wang, C., Ren, K., Lou, W., and Li, J., 2011. Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing. *IEEE Transactions on Parallel and Distributed Systems*, 22, 847–859.

Wei, L., Zhu, H., Cao, Z., Jia, W., and Vasilakos, A. V., 2010. SecCloud: Bridging Secure Storage and Computation in Cloud. *In*: *Proceedings - International Conference on Distributed Computing Systems*. 52–61.

Weinhardt, C., Anandasivam, A., Blau, B., Borissov, N., Meinl, T., Michalk, W., and Stosser, J., 2009. Cloud Computing - A Classification, Business Models, and Research Directions. *Business & Information Systems Engineering*, 391.

Weiss, A., 2007. Computing in the Clouds. *netWorker Magazine - Cloud computing: PC functions move onto the web*, (Volume II, Issue 4), 16–25.

Woolford, S., 2015. (Factor) Analyze This: PCA or EFA.

Wu, J., Ping, L., Ge, X., Ya, W., and Fu, J., 2010. Cloud storage as the Infrastructure of Cloud Computing. *In*: *Proceedings - 2010 International Conference on Intelligent Computing and Cognitive Informatics, ICICCI 2010*. 380–383.

Yao, C., Xu, L., and Huang, X., 2013. A Secure Cloud Storage System from Threshold Encryption. *Proceedings - 5th International Conference on Intelligent Networking and Collaborative Systems, INCoS 2013*, 541–545.

Yao, J., Chen, S., Nepal, S., Levy, D., and Zic, J., 2010. TrustStore: Making Amazon S3 Trustworthy with Services Composition. *CCGrid 2010 - 10th IEEE/ACM International Conference on Cluster, Cloud, and Grid Computing*, 600–605.

Zainuddin, A., 2012. *A Handbook on SEM Structural Equation Modelling*. 6th ed. Kampus Kota Bharu: UiTM Kelantan.

Zarandioon, S., Yao, D., and Ganapathy, V., 2012. K2C: Cryptographic Cloud Storage with Lazy Revocation and Anonymous Access. *In*: *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*. 59–76.

Zhang, R. and Chen, P., 2012. A Dynamic Cryptographic Access Control Scheme in Cloud Storage Services. *In*: *Proceedings - 2012 8th International Conference on Computing and Networking Technology (INC, ICCIS and ICMIC), ICCNT 2012*. 50–55.

Zhao, R. and Yue, C., 2014. Toward a Secure and Usable Cloud-Based Password Manager for Web Browsers. *Computers & Security*, 46, 32–47.

Zhou, L., Varadharajan, V., and Hitchens, M., 2013. Achieving Secure Role-Based Access Control on Encrypted Data in Cloud Storage. *Information Forensics and Security, IEEE Transactions on*, 8 (12), 1947–1960.

Zikmund, W., Babin, B., Carr, J., and Griffin, M., 2012. *Business Research Methods*. 9th ed. Cengage Learning.

Zissis, D. and Lekkas, D., 2012. Addressing Cloud Computing Security Issues. *Future Generation Computer Systems*, 28 (3), 583–592.

Appendix A: Confirming Cloud Storage Security Framework (Initial Study)

A.1 Interview Questions

Interview Questions

Part I General

Background questions

1. **What is your organisation domain?**

( ) Industry ( ) Education/Academic ( ) Government
( ) Others, please specify: ____

2. **Which of these roles fits your job description?**

( ) IT/Technical (Application Security, Digital Forensics Investigation, Threat Intelligence etc.)
( ) Consultant/Advisory (Consultant, Industry Research/Analyst)
( ) Security Researcher (Academic, PhD Researcher etc.)
( ) Security Policy Maker (CIO, CTO, CSO, Chief Security Information Officer etc.)
( ) Others, please specify: ____

3. **How long have you been working in computer and information technology security i.e Cyber Security?**

( ) 5 years ( ) 6-10 years ( ) More than 10 years

4. **What is the estimated percentage of data stored in your information or cloud systems (example: cloud storage)?**

( ) 0% - 25% ( ) 25%- 50% ( ) 50%- 75% ( ) 75%- 100%

5. **Do you use a cloud storage?**

( ) Yes ( ) No

6. **What type of Cloud Storage do you frequently use?**

( ) Commercial
( ) Public Cloud Storage provided by your organisation
( ) Private Cloud Storage provided by your organisation
( ) Others, please specify: ____

Thank you for your answers. In Part II, we will discuss some of the important security factors objectives in a cloud system.

Part II Security Factors

7. How important are these security factors in the cloud?
   a) Security Implementation in Cloud Storage (policy, procedures etc.)
   b) Confidentiality
   c) Integrity
   d) Availability
   e) Non-Repudiation
   f) Authenticity
   g) Reliability
8. In your opinion, are there any other factors or issues that you think would matter besides the seven factors above?

Part III Security Concerns

A technical report by Cloud Security Alliance (CSA) and Georgia Technology University etc. has identified these concerns as below.

   a) Data Breach
   b) Data Loss & Leakage
   c) Insecure APIs
   d) Account/Service Hijacking
   e) Denial of Service
   f) Malicious Insiders
   g) Abuse of Cloud Service
   h) Inadequate Cloud Planning/Design
   i) Cloud-Related Malware
   j) Closure of cloud service
   k) Natural Disaster
   l) Hardware Failure
   m) Shared Technology Vulnerabilities
   n) Insufficient due to diligence

9. Let's discuss about this. What do you think of the concerns?
10. In your experience, what are the important and high risk concerns?

A.2 Participant Information

Participant Information

Ethics reference number: **ERGO/FPSE/14962**	Version: 1	Date: 2015-05-05
Study Title: A Security Framework to Protect Data in Cloud Storage		
Investigator: Farashazillah B Yahya		

Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form. Your participation is completely voluntary.

What is the research about? This is a research project that aims to investigate the appropriate security framework to protect data in cloud storage. The study is sponsored by the University of Southampton at the end of the study, you may request how your data is used as the study findings.

Why have I been chosen? You have been approached because of your expertise in the field of Computer and Information Technology Security.

What will happen to me if I take part? You will be informed of the study purpose and procedures. Your participation is voluntary, and it's your right to unconditionally withdraw at any time and for any reason. You will first sign a consent and then will be asked some questions with regard to the study. It will take about 50 minutes in total give and a 10 minutes break.

Are there any benefits in my taking part? The study will contribute to current knowledge about managing threats in cloud systems and protecting data in cloud storage.

Are there any risks involved? There are no particular risks associated with your participation.

Will my participation be confidential? All data collected is anonymous and your data will be kept confidential on a password-protected computer and used only for the purposes of this study. It will be linked to your consent form by a code. The investigator will destroy it once the study is done. If you would like to access your data after your participation, change it, or withdraw it, please contact the investigator (e-mail: fara.yahya@soton.ac.uk) or project supervisor (rjw1@soton.ac.uk) who will arrange this.

What happens if I change my mind? You may withdraw at any time and for any reason. You may access, change, or withdraw your data at any time and for any reason prior to its destruction. You may keep any benefits you receive.

What happens if something goes wrong? Should you have any concern or complaint, contact me if possible (investigator e-mail: fara.yahya@soton.ac.uk), otherwise please contact the FPSE Office (e-mail: lg11@soton.ac.uk) or any other authoritative body such as Dr Martina Prude, Head of Research Governance (02380 595058, mad4@soton.ac.uk).

A.3 Ethics Consent Form (Interview)

Consent Form for the Interview

Ethics reference number: **ERGO/FPSE/14962**	Version: 1	Date: 2015-05-05
Study Title: A Security Framework to Protect Data in Cloud Storage		
Investigator: Farashazillah B Yahya		

Please initial the box(es) if you agree with the statement(s):

I have read and understood the Participant Information (version 1 dated 2015-05-05) and have had the opportunity to ask questions about the study. □

I agree to take part in this study and agree for my data to be used for the purpose of this study. □

I understand my participation is voluntary and I may withdraw at any time and for any reason. □

Data Protection

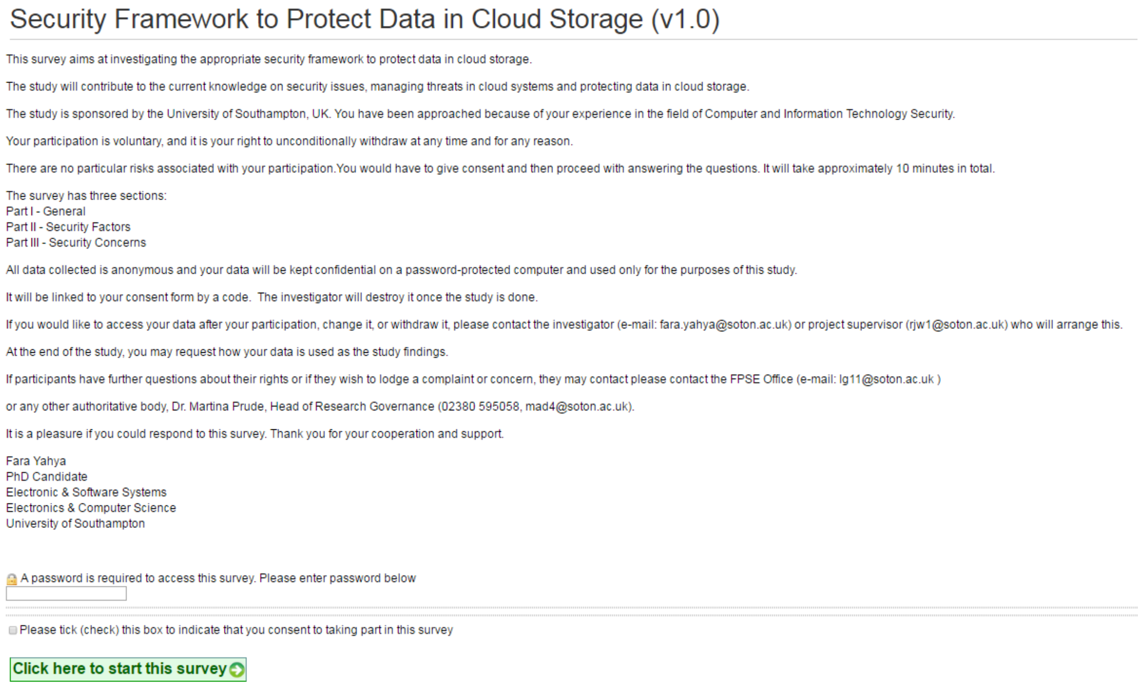
I understand that information collected and recorded during my participation in this study is completely anonymous and will be stored on a password protected computer and that this information will only be used for the purpose of this study.

Name of participant (print name)……………………………………………………

Signature of participant………………………………………………………………….

Date……………………………………………………………………………………..

A.4 Survey Questions



Security Framework to Protect Data in Cloud Storage (v1.0)

This survey aims at investigating the appropriate security framework to protect data in cloud storage.

The study will contribute to the current knowledge on security issues, managing threats in cloud systems and protecting data in cloud storage.

The study is sponsored by the University of Southampton, UK. You have been approached because of your experience in the field of Computer and Information Technology Security.

Your participation is voluntary, and it is your right to unconditionally withdraw at any time and for any reason.

There are no particular risks associated with your participation.You would have to give consent and then proceed with answering the questions. It will take approximately 10 minutes in total.

The survey has three sections:
Part I - General
Part II - Security Factors
Part III - Security Concerns

All data collected is anonymous and your data will be kept confidential on a password-protected computer and used only for the purposes of this study.

It will be linked to your consent form by a code. The investigator will destroy it once the study is done.

If you would like to access your data after your participation, change it, or withdraw it, please contact the investigator (e-mail: fara.yahya@soton.ac.uk) or project supervisor (rjw1@soton.ac.uk) who will arrange this.

At the end of the study, you may request how your data is used as the study findings.

If participants have further questions about their rights or if they wish to lodge a complaint or concern, they may contact please contact the FPSE Office (e-mail: lg11@soton.ac.uk)

or any other authoritative body, Dr. Martina Prude, Head of Research Governance (02380 595058, mad4@soton.ac.uk).

It is a pleasure if you could respond to this survey. Thank you for your cooperation and support.

Fara Yahya
PhD Candidate
Electronic & Software Systems
Electronics & Computer Science
University of Southampton

A password is required to access this survey. Please enter password below

□ Please tick (check) this box to indicate that you consent to taking part in this survey

Practitioners Survey

Part I General

1. What is your organisation domain?
   □ Industry
   □ Education/Academic
   □ Government
   □ Others, please specify: ………………………………………………………………

2. Which of these roles fits your fits your job description?
   □ IT/Technical (Application Security, Digital Forensics Investigation, Concern Intelligence etc.)
   □ Consultant/Advisory (Consultant, Industry Research/Analyst)
   □ Security Researcher (Academic, PhD Researcher etc.)
   □ Security Policy Maker (CIO, CTO, CSO, Chief Security Information Officer etc.)
   □ Others, please specify: ………………………………………………………………

3. How long have you been working in Cyber Security?
   □ 2-5 years □ 6-10 years□ More than 10 years
4. What is the estimated percentage of data stored in your cloud storage? This includes files stored in private cloud storage (your organisation) and public cloud storage including Dropbox, Google Drives etc.
   □ 0% - 25% □ 26%- 50% □ 51%- 75% □ 76%- 100%

5. Do you use a cloud storage?

   □ Yes □ No

6. What type of Cloud Storage do you frequently use?

   □ Commercial
   □ Public Cloud Storage provided by your organisation
   □ Private Cloud Storage provided by your organisation
   □ Others, please specify: ………………………………………………………………

7. In which country are you working at the moment?

□ Malaysia
□ United Kingdom
□ Others, please specify: ……………………………………………………………………

Part II Security Factors

8. Please state whether you find the security objectives important.

Factors	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Cloud storage security policy must be implemented in an organisation	○	○	○	○	○
Procedure must be implemented to fulfill the implementation of security policy(s)	○	○	○	○	○
Sensitive data must not reach the wrong person	○	○	○	○	○
Access must be restricted to those authorised to view the data	○	○	○	○	○
Data must not be changed or altered by unauthorised people	○	○	○	○	○
Data is hidden from those that are not supposed to see it	○	○	○	○	○
A functioning system environment must be correctly maintained	○	○	○	○	○
Keeping up with the latest necessary system upgrades	○	○	○	○	○
Data correctly reflects the object	○	○	○	○	○
Individual owner of an account must not allow other user to use his/her account	○	○	○	○	○
A user whose authentication request is approved becomes authorised to access the accounts of that account holder	○	○	○	○	○
Data across the system should be in synch with each other	○	○	○	○	○
Critical components or functions of a system are duplicated to increase reliability of the system	○	○	○	○	○
A proof of the integrity and origin of data must be provided	○	○	○	○	○
The data source is trustworthy	○	○	○	○	○
Accountable against data loss or interruptions	○	○	○	○	○
A source must be able to provide proof of identity	○	○	○	○	○
Data are protected with policies by accredited bodies	○	○	○	○	○

Part III Security Concerns

9. The literature identified the following issues as the most frequent concerns, could you list the highest risk of cloud storage concerns in your organisation?

Concerns	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Data breached intentionally to an untrusted environment	○	○	○	○	○
Unintentionally data disclosure to an untrusted environment	○	○	○	○	○
Accidental data deletion by the cloud provider	○	○	○	○	○
Data leaked unintentionally to a trusted/untrusted environment	○	○	○	○	○
Current or former employee, contractor, or others who has or had intentionally exceeded authorised access to organisation's resources	○	○	○	○	○
Sabotage of the organisation's systems or data.	○	○	○	○	○
Interfaces are not designed to protect against both accidental and malicious	○	○	○	○	○
Bypassing authentication defenses and data validation via third party APIs	○	○	○	○	○
System slowdown due to inordinate amounts of finite system resources consumption	○	○	○	○	○
Service outages as system isnt responding	○	○	○	○	○
Credentials are stolen	○	○	○	○	○
Phishing, fraud, and exploitation of software vulnerabilities	○	○	○	○	○
Abuse of legitimate cloud infrastructure	○	○	○	○	○
Computing resources are used without authorisations	○	○	○	○	○
Vulnerability or misconfiguration leading to a compromise across an entire cloud.	○	○	○	○	○
Compromise of an integral piece of shared technology	○	○	○	○	○
Malfunction within the cloud service	○	○	○	○	○
Services are unavailable	○	○	○	○	○
Hostile or intrusive software, including computer viruses, worms, trojans, ransomware, spyware, adware, scareware, and other malicious programs	○	○	○	○	○
Data compromised by software used to disrupt computer operation, gather sensitive information, or gain access to private computer systems	○	○	○	○	○
Poor and inadequate cloud planning and design	○	○	○	○	○
Unrealistic expectations and lack of proper training for cloud users	○	○	○	○	○

A.5 Interview Analysis - Thematic Analysis

Codes	Themes	Expert	Flag
Cloud Storage Security Implementation	All agreed that security implementation in cloud storage must be initiated by having a well-defined security policy in place.	L, P, E, H, A, J, Q, T, B, S, F, G, M, N, R, I, K, O, C, D, F	C
	It is important to have a defined security policy and process	S	C
	A security policy will inform users, staff, and managers, specify mechanisms for security and provide a baseline	A	C
	Some organisation when going for certification and compliance to standards that policy implementation is compulsory. Policy is the best compliance tool – legislatively	P	C
Confidentiality	All agreed that confidentiality is an important security factor.	L, P, E, H, A, J, Q, T, B, S, F, G, M, N, R, I, K, O, C, D, F	C
	Cloud storage should only show documents to authorised users	J	C
	Systems require passwords as part of access control.	R	C
	Experts mentioned that confidentiality, integrity and availability are most basic and common security factors. A system is secured through these three factors.	A, E, L, O, P, and S	C
	Experts mentioned that identification, authentication, authorisation, access control and encryption are practices and measures to prove confidentiality and integrity.	F, I, J, R	C
Integrity	The majority eighteen over twenty agreed that integrity is an important security factor	A, G, K, E, L, O, P, N, Q, S, T, B, C, D, F, H, I, J	C
	Experts mentioned that confidentiality, integrity and availability are most basic and common security factors. A system is secured through these three factors.	A, E, L, O, P, S	C
	Data integrity ensures data are not leaked or lost in the cloud.	N	C
	Without integrity, data that is received or sent cannot be trusted.	Q	C
Availability	All agreed that availability is an important security factor.	L, P, E, H, A, J, Q, T, B, S, F, G, M, N, R, I, K, O, C, D, F	C
	Cloud is becoming the only real choice when faced with the pressure of finance, availability, security, ease-of-use, and scalability	T	C
	Experts mentioned that confidentiality, integrity and availability are most basic and common security factors. A system is secured through these three factors.	A, E, L, O, P, and S	C
	Expert mentioned that although availability is more like a functional factor of a system, it is important as it ensures the data can always be accessed.	J	C
	Expert mentioned that availability is the most important factor, as the user will always expect the system to be accessible, regardless of any concerns.	P	C
Non-repudiation	The majority fourteen over twenty agreed that non-repudiation is important.	P, H, I, A, J, K, L, B, M, N, O,	C

Codes	Themes	Expert	Flag
		C, D, E,	
	Non-repudiation can reduce fraud and promote the legal enforceability of electronic agreements and transactions	P	C
	Expert explained that non-repudiation is an assurance that any party cannot deny sending or receiving the data. This includes obligations for contracts, standards, etc.	H	C
	A contract or scope of work is an assurance that security protection is met - assurance	I	O
Authenticity	The majority fifteen over twenty agreed that authenticity is an important factor.	J, K, F, L C, D, E, O R, M, N, P H, B, C	C
	Solid authentication defends a system against the security risk of impersonation, in which a sender or receiver uses a false identity to access a system. Digital certificates can provide a more secure method of authentication while offering other security benefits as well	R	C
	Expert revealed that currently IT systems are also reviewed, based on their quality of delivering authentic data. This involves verifying the source as genuine.	M	C
	Verification is part of authenticity	N	O
Reliability	All experts mentioned that reliability is an important factor.	L, P, E, H, A, J, Q, T, B, S, F, G, M, N, R, I, K, O, C, D, F	C
	Measuring whether the system is consistent	G	C
	Is the system reliable and valid?	T	C
Accountability	Expert mentioned that accountable services are provided on trust basis by having a contract or SLAs with a clear and concise definition of security policy.	A	A
	Expert emphasises that accountability and trust is an important factor that must be fulfilled by providers (internal or external). They have made it compulsory for Security Level Agreement in many organisations as an extension to Service Level Agreements (SLAs).	K	A
	Accountability involves the processes, policies, and controls necessary to trace actions to their source. These develop trust among systems -Accountability is overlapping with trust	Q	A
Auditability	Expert specified that organisations are providing security policies to ensure work task follows the guidelines and best practices.	C	A
	Expert stated that IT security standards are being implemented to gain confidence. The system must enable assessment, examination and audit to be completed smoothly.	E	A
	A cloud storage should provide full access logs, allowing organisations to see how data is being accessed, shared and used in real time. This audit data is available through APIs to real time systems allowing organisations to respond to data governance issues and provide full audit logs where required	T	A

C: Confirmed, I: Irrelevant, A: Additional and O: Overlapped

A.6 Ethics Consent Form (Survey)

Consent Form for the Questionnaire

Ethics reference number: **ERGO/FPSE/14962**	Version: 1	Date: 2015-05-05
Study Title: A Security Framework to Protect Data in Cloud Storage		
Investigator: Farashazillah B Yahya		

Please press 'I agree' if you agree with the following statement(s):

I have read and understood the Participant Information sheet provided on the previous page (version 1 dated 2015-05-05) □

I agree to take part in this research project and agree for my data to be used for the purpose of this study. □

I understand my participation is voluntary and I may withdraw at any time without consequence and my data will be deleted if I withdraw at any time. □

Data Protection

I understand that that this study is anonymous and no personal data will be stored.

When 'I agree' is pressed on the online the participant will be moved to the next page to start the questionnaire.

A.7 Survey Analysis - Statistical Analysis

One sample t-test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
CloudStorageSecurity1	34	1.76	1.046	0.179
CloudStorageSecurity2	34	1.62	0.954	0.164
Confidentiality1	34	1.65	0.774	0.133
Confidentiality2	34	1.59	0.743	0.127
Integrity1	34	1.79	0.914	0.157
Integrity2	34	1.76	1.046	0.179
Availability1	34	1.62	0.954	0.164
Availability2	34	1.76	1.017	0.174
Nonrepudiation1	34	1.94	0.919	0.158
Nonrepudiation2	34	1.82	0.869	0.149
Authenticity1	34	1.79	0.914	0.157
Authenticity2	34	1.68	0.843	0.145
Reliability1	34	1.88	0.844	0.145
Reliability2	34	1.74	0.828	0.142
Accountability1	34	1.88	0.880	0.151
Accountability2	34	1.85	0.857	0.147
Auditability1	34	1.82	0.758	0.130
Auditability2	34	1.88	0.769	0.132

One-Sample Test

	Test Value = 2.5					
				Mean Differenc e	95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)		Lower	Upper
CloudStorageSecurity1	-4.098	33	0.001	-0.735	-1.10	-0.37
CloudStorageSecurity2	-5.393	33	0.001	-0.882	-1.22	-0.55
Confidentiality1	-6.426	33	0.001	-0.853	-1.12	-0.58
Confidentiality2	-7.152	33	0.001	-0.912	-1.17	-0.65
Integrity1	-4.504	33	0.001	-0.706	-1.02	-0.39
Integrity2	-4.098	33	0.001	-0.735	-1.10	-0.37
Availability1	-5.393	33	0.001	-0.882	-1.22	-0.55
Availability2	-4.217	33	0.001	-0.735	-1.09	-0.38
Nonrepudiation1	-3.545	33	0.002	-0.559	-0.88	-0.24
Nonrepudiation2	-4.537	33	0.001	-0.676	-0.98	-0.37
Authenticity1	-4.504	33	0.001	-0.706	-1.02	-0.39
Authenticity2	-5.698	33	0.001	-0.824	-1.12	-0.53
Reliability1	-4.265	33	0.001	-0.618	-0.91	-0.32
Reliability2	-5.386	33	0.001	-0.765	-1.05	-0.48
Accountability1	-4.095	33	0.001	-0.618	-0.92	-0.31
Accountability2	-4.400	33	0.001	-0.647	-0.95	-0.35
Auditability1	-5.206	33	0.001	-0.676	-0.94	-0.41
Auditability2	-4.682	33	0.001	-0.618	-0.89	-0.35

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Databreach1	34	2.59	1.328	0.228
Dataleak2	34	2.06	.919	0.158
Dataleak1	34	2.38	1.303	0.223
Databreach2	34	2.62	1.206	0.207
Insecureapi1	34	1.94	0.952	0.163
Insecureapi2	34	2.09	1.083	0.186
Insider1	34	2.68	1.249	0.214
Insider2	34	2.47	1.237	0.212
Dos1	34	2.85	1.105	0.189
Dos2	34	2.59	1.104	0.189
Acchijacking1	34	2.50	1.354	0.232
Acchijacking1	34	2.32	1.007	0.173
Abuse1	34	2.71	1.244	0.213
Abuse2	34	2.56	0.824	0.141
Sharedtech1	34	2.26	1.189	0.204
Sharedtech2	34	2.29	1.001	0.172
Malware1	34	1.97	0.834	0.143
Malware2	34	2.09	1.026	0.176
Planning1	34	2.56	1.021	0.175
Planning2	34	2.74	0.828	0.142

One-Sample Test

	Test Value = 2.5					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Databreach1	-1.807	33	0.080	-0.412	-0.88	0.05
Dataleak2	-5.970	33	0.000	-0.941	-1.26	-0.62
Dataleak1	-2.764	33	0.009	-0.618	-1.07	-0.16
Databreach2	-1.848	33	0.074	-0.382	-0.80	0.04
Insecureapi1	-6.488	33	0.000	-1.059	-1.39	-0.73
Insecureapi2	-4.907	33	0.000	-0.912	-1.29	-0.53
Insider1	-1.511	33	0.140	-0.324	-0.76	0.11
Insider2	-2.496	33	0.018	-0.529	-0.96	-0.10
Dos1	-0.776	33	0.443	-0.147	-0.53	0.24
Dos2	-2.174	33	0.037	-0.412	-0.80	-0.03
Acchijacking1	-2.153	33	0.039	-0.500	-0.97	-0.03
Acchijacking1	-3.918	33	0.000	-0.676	-1.03	-0.33
Abuse1	-1.379	33	0.177	-0.294	-0.73	0.14
Abuse2	-3.124	33	0.004	-0.441	-0.73	-0.15
Sharedtech1	-3.607	33	0.001	-0.735	-1.15	-0.32
Sharedtech2	-4.112	33	0.000	-0.706	-1.06	-0.36
Malware1	-7.194	33	0.000	-1.029	-1.32	-0.74
Malware2	-5.182	33	0.000	-0.912	-1.27	-0.55
Planning1	-2.520	33	0.017	-0.441	-0.80	-0.09
Planning2	-1.864	33	0.071	-0.265	-0.55	0.02

Appendix B Security Rating Score (SecRaS)

B.1 Participant Information

Participant Information

Ethics reference number: **ERGO/FPSE/18945**	Version: 1	Date: 2015-02-01
Study Title: A Security Rating Score (SecRaS) Instrument to Evaluate Security in Cloud Storage		
Investigator: Farashazillah B Yahya		

Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form. Your participation is completely voluntary.

What is the research about? This is a research project that aims to develop and validate the security rating score (SecRaS) instrument based on the security framework to protect data in cloud storage. The instrument is constructed from a set of questions that describes the component in the security framework. The study is sponsored by the University of Southampton at the end of the study, you may request how your data is used as the study findings.

Why have I been chosen? You have been approached because of your expertise in the field of Computer and Information Technology Security.

What will happen to me if I take part? You will be informed of the study purpose and procedures. Your participation is voluntary, and it's your right to unconditionally withdraw at any time and for any reason. You will first sign a consent and then will be asked some questions with regard to the study. It will take about 60 minutes in total give and a 10 minutes break.

Are there any benefits in my taking part? The study will contribute to current knowledge about security in cloud systems and protecting data in cloud storage.

Are there any risks involved? There are no particular risks associated with your participation.

Will my participation be confidential? All data collected is anonymous and your data will be kept confidential on a password-protected computer and used only for the purposes of this study. It will be linked to your consent form by a code. Once submitted, you will no longer be able to interact with, modify or remove your input, since it can no longer be identified with you. The investigator will destroy the data once the study is done.

What happens if I change my mind? You may withdraw your participation at any time and for any reason. If you would like to withdraw your participation, please contact the investigator (e-mail: fara.yahya@soton.ac.uk) or project supervisor (rjw1@soton.ac.uk) who will arrange this.

What happens if something goes wrong? Should you have any concern or complaint, contact me if possible (investigator e-mail: fara.yahya@soton.ac.uk), otherwise please contact the FPSE Office (e-mail: lg11@soton.ac.uk) or any other authoritative body such as the Head of the Research Governance Office (02380 595058, rgoinfo@southampton.ac.uk).

B.2 Consent Form

Consent Form for the expert review and validation study (questionnaire)

Ethics reference number: **ERGO/FPSE/18945**	Version: 1	Date: 2015-02-01
Study Title: A Security Rating Score (SecRaS) Instrument to Evaluate Security in Cloud Storage		
Investigator: Farashazillah B Yahya		

Please press 'I agree' if you agree with the following statement(s):

I have read and understood the Participant Information sheet provided on the previous page (version 1 dated 2016-02-01)

I agree to take part in this research project and agree for my data to be used for the purpose of this study.

I understand my participation is voluntary and I may withdraw at any time without consequence and my data will be deleted if I withdraw at any time.

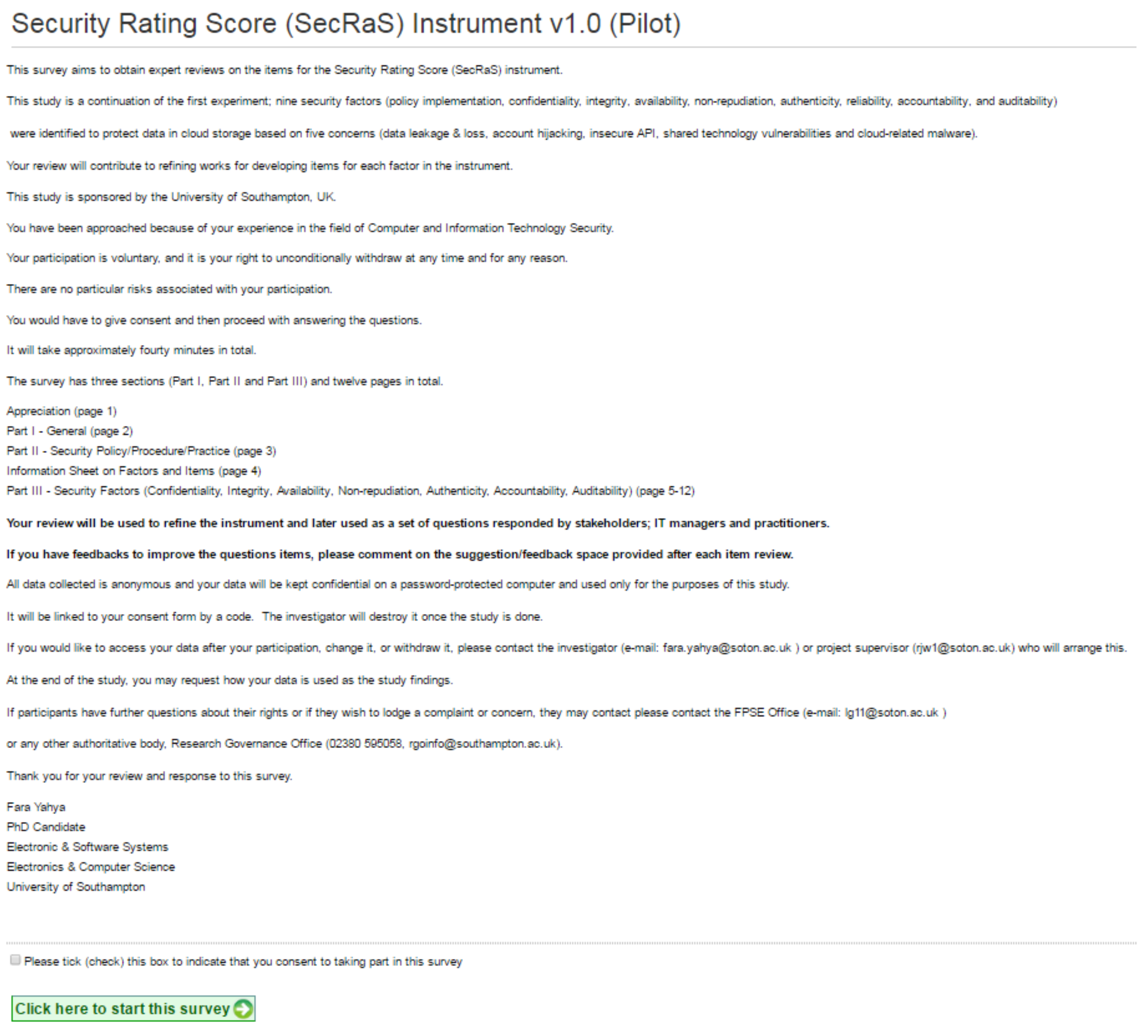
Data Protection

I understand that information collected during my participation in this study is completely anonymous / will be stored on a password protected computer/secure University server and that this information will only be used in accordance with the Data Protection Act (1998). The DPA (1998) requires data to be processed fairly and lawfully in accordance with the rights of participants and protected by appropriate security.

□ ***I agree***

When 'I agree' is pressed on the online the participant will be moved to the next page to start the questionnaire.

B.3 Initial Instrument



Security Rating Score (SecRaS) Instrument v1.0 (Pilot)

This survey aims to obtain expert reviews on the items for the Security Rating Score (SecRaS) instrument.

This study is a continuation of the first experiment; nine security factors (policy implementation, confidentiality, integrity, availability, non-repudiation, authenticity, reliability, accountability, and auditability)

were identified to protect data in cloud storage based on five concerns (data leakage & loss, account hijacking, insecure API, shared technology vulnerabilities and cloud-related malware).

Your review will contribute to refining works for developing items for each factor in the instrument.

This study is sponsored by the University of Southampton, UK.

You have been approached because of your experience in the field of Computer and Information Technology Security.

Your participation is voluntary, and it is your right to unconditionally withdraw at any time and for any reason.

There are no particular risks associated with your participation.

You would have to give consent and then proceed with answering the questions.

It will take approximately fourty minutes in total.

The survey has three sections (Part I, Part II and Part III) and twelve pages in total.

Appreciation (page 1)
Part I - General (page 2)
Part II - Security Policy/Procedure/Practice (page 3)
Information Sheet on Factors and Items (page 4)
Part III - Security Factors (Confidentiality, Integrity, Availability, Non-repudiation, Authenticity, Accountability, Auditability) (page 5-12)

Your review will be used to refine the instrument and later used as a set of questions responded by stakeholders; IT managers and practitioners.

If you have feedbacks to improve the questions items, please comment on the suggestion/feedback space provided after each item review.

All data collected is anonymous and your data will be kept confidential on a password-protected computer and used only for the purposes of this study.

It will be linked to your consent form by a code. The investigator will destroy it once the study is done.

If you would like to access your data after your participation, change it, or withdraw it, please contact the investigator (e-mail: fara.yahya@soton.ac.uk) or project supervisor (rjw1@soton.ac.uk) who will arrange this.

At the end of the study, you may request how your data is used as the study findings.

If participants have further questions about their rights or if they wish to lodge a complaint or concern, they may contact please contact the FPSE Office (e-mail: lg11@soton.ac.uk)

or any other authoritative body, Research Governance Office (02380 595058, rgoinfo@southampton.ac.uk).

Thank you for your review and response to this survey.

Fara Yahya
PhD Candidate
Electronic & Software Systems
Electronics & Computer Science
University of Southampton

□ Please tick (check) this box to indicate that you consent to taking part in this survey

Practitioners Survey

Part I General

1. What is your organisation domain?
   □ Academic
   □ Government
   □ Industry
   □ Others, please specify: ……………………………………………………………………

2. Which of these roles fits your fits your job description?
   □ IT/Technical (Application Security, Digital Forensics Investigation, Concern Intelligence etc.)
   □ Consultant/Advisory (Consultant, Industry Research/Analyst)
   □ Security Analyst/Expert (Information Security)
   □ Security Policy Maker (CIO, CTO, CSO, Chief Security Information Officer etc.)
   □ Security Researcher (Academic, PhD Researcher etc.)
   □ Others, please specify: ……………………………………………………………………

3. How long have you been working in computer & IT security (i.e. cyber security?
   □ 5 years □ 6-10 years □ More than 10 years

4. What is the estimated percentage of data stored in your/your organisations' cloud storage?.
   □ 0% - 25%   □ 26%- 50%   □ 51%- 75%   □ 76%- 100%

Part II Security Policy/Procedure/Practice

The Security Rating Score (SecRaS) has been developed with the idea that IT security policy, procedure and practice are in place. Please refer to the rating score table shown below.

5. These are a set of questions that describes whether policy, procedure and practice are implemented in their organisation. Please indicate your agreement.

Items	Rating Score				
	No Plan to Implement	Planning to Implement	Do Not Know	Partially Implement	Fully Implement
My organisation has cloud storage security policy in place	○	○	○	○	○
Please provide details: _______________________ Example: The organisation is ISO27001 compliance					
My organisation has implemented security procedures to comply to industry standards	○	○	○	○	○
Please provide details: _______________________					
My organisation has enforced security processes to support cloud storage security policy/ies	○	○	○	○	○
Please provide details: _______________________					
My organisation has imposed security controls to support cloud storage security policy/ies	○	○	○	○	○
Please provide details: ______________________					

Part III Security Factors

6. This is an overview of the Security Rating Score (SecRaS). SecRaS consists of eight factors (Confidentiality, Integrity, Availability, Non-repudiation, Authenticity, Accountability, and Auditability) and 48 items. Please indicate your agreement.

Items	Rating Score				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Confidentiality					
My organisation has identity management policy	○	○	○	○	○
My organisation has user-based authentication process	○	○	○	○	○
My organisation has access management policy	○	○	○	○	○
My organisation has process to specify rights and restrictions for user access to data	○	○	○	○	○
My organisation follows industry standards to build in security for systems	○	○	○	○	○
My organisation has a secure communication channel policy	○	○	○	○	○
Integrity					
My organisation has documentation of encryption management practices/guidelines	○	○	○	○	○

Items	Rating Score				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
My organisation encrypt user data at rest (on disk/storage) within the environment	○	○	○	○	○
My organisation leverage encryption to protect virtual machine images during transport between hypervisor instances	○	○	○	○	○
My Organisation supports data stewardship	○	○	○	○	○
My organisation support secure deletion (e.g., degaussing/cryptographic wiping) of archived data	○	○	○	○	○
My organisation supports additional protection for user to store sensitive data	○	○	○	○	○
Availability					
My organisation segments data logically for each user	○	○	○	○	○
My organisation has backup or redundancy mechanisms	○	○	○	○	○
My organisation has data recovery mechanisms	○	○	○	○	○
My organisation verifies data authenticity after restore process	○	○	○	○	○
My organisation documents the restore or redundancy mechanisms	○	○	○	○	○
My organisation has define restore procedure for responding to requests for user data from governments or third parties	○	○	○	○	○
Non-Repudiation					
My organisation has key management policies binding keys to identifiable owners	○	○	○	○	○
My organisation uses synchronised time-service protocol (e.g., NTP etc.)	○	○	○	○	○
My organisation uses geographical location as an authentication	○	○	○	○	○
My organisation has restriction of user data to specific countries or geographic locations	○	○	○	○	○
My organisation provides the means for authorised individuals to determine the identity of the data producer	○	○	○	○	○
My organisation support integration of location as an authentication factor	○	○	○	○	○
Authenticity					
My organisation has cryptographic protection mechanisms	○	○	○	○	○
My organisation ensures the origin authentication	○	○	○	○	○
My organisation has verification assurances to ensure session	○	○	○	○	○

Items	Rating Score				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
authenticity					
My organisation has anti-counterfeiting policy	○	○	○	○	○
My organisation provide session-level protection where needed	○	○	○	○	○
My organisation provides mechanisms to protect the authenticity of communications sessions	○	○	○	○	○
Reliability					
My organisation has multi-failure disaster recovery	○	○	○	○	○
My organisation has system maintenance process/policy	○	○	○	○	○
My organisation has patch management policy/process	○	○	○	○	○
My organisation has continuous monitoring process/solutions	○	○	○	○	○
My organisation has malicious code protection mechanisms at entry and exit points	○	○	○	○	○
My organisation has conducted failover test	○	○	○	○	○
Accountability					
My organisation has process to conformance with external standards	○	○	○	○	○
My organisation has mechanisms to put internal security policies in effect	○	○	○	○	○
My organisation supports transparency and participation to conformance process with internal standards/policy	○	○	○	○	○
My organisation ensures clarity of Service Level Agreement/Guarantee (SLAs/SLGs)	○	○	○	○	○
My organisation conducts penetration tests of cloud service infrastructure regularly as prescribed by industry best practices/guidance	○	○	○	○	○
My organisation has means of remediation for internal enforcement	○	○	○	○	○
Auditability					
My organisation produces audit assertions using a structured, industry accepted format (e.g., Cloud Audit/Cloud Trust, ISACA's Cloud Computing Management Audit, etc.)	○	○	○	○	○
My organisation reviews audit logs on a regular basis	○	○	○	○	○
My organisation has on-demand audit review	○	○	○	○	○
My organisation generates audit report	○	○	○	○	○
My organisation ensures the audit log is in original content or	○	○	○	○	○

Items	Rating Score				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
time					
My organisation has a process to audit records for events of interest	○	○	○	○	○

B.4 Description of Security Rating Score (SecRaS) Instrument

Goal: Understanding the cloud storage security implementation

Factor	Cloud Storage Security Implementation (CS)	
Description	Realising Security Policies Implementation in Cloud Storage Achieving security is a challenge not solvable by applying technical solutions only. As described in the definition of security goals (confidentiality, integrity, availability, non-repudiation, authenticity, reliability), providing throughout security exceeds all technical possibilities. Cloud services have such an impact, that covering other areas such as policy implementation becomes crucial. Recent developments show the need for techniques as well as for establishing policies and procedures to protect user data. Regulations and technical measures use different terminologies resulting in different interpretations questioning if the same problems have to be solved. Therefore, consistent terminology in terms of security in cloud storage must be well-defined in cloud storage policies and procedures.	
Goal	Understanding the cloud storage security implementation	
Question	What are the security implementation initiated for securing a cloud storage?	
Metric	Security implementation in cloud storage	
Items	CS1	My organisation has cloud storage security policy in place
	CS2	My organisation has implemented security procedures to comply to industry standards
	CS3	My organisation has enforced security processes to support cloud storage security policy/ies
	CS4	My organisation has imposed security controls to support cloud storage security policy/ies

Goal: Assessing the confidentiality of data accessed in cloud storage from the stakeholder's viewpoint

Factor	Confidentiality (Co)	
Description	Confidential data handling in cloud storage is the protection of data by allowing only the intended recipient to access the data. Data should be handled correctly to prevent unauthorised exposure (Firesmith 2004). Brock and Goscinski (2010) has characterise security concerns of clouds by proposing a Cloud Security Framework (CSF) that takes into consideration cloud infrastructure protection to ensure confidentiality. Data can be protected by applying access controls, authentication and authorisation while handling data effectively (Vrable et al. 2012, Mapp et al. 2014, El-Booz et al. 2016). Ensuring confidentiality is performed straight-forward: Before uploaded in the cloud, user access to the data needed to be decided (El-Booz et al. 2016). Security in any system including cloud storage involves primarily ensuring that the right user gets access to only the authorised data in the authorised format at an authorised time and from an authorised location. Identity and access management is of prime importance in this regard (Habiba et al. 2014, Singh and Chatterjee 2015).	
Goal	Assessing the confidentiality of data	
Question	What are the security measures for securing the confidentiality of data cloud storage?	
Metric	Security measure affecting confidentiality of data in cloud storage	
Items	Co1	My organisation has identity management policy
	Co2	My organisation has user-based authentication process
	Co3	My organisation has access management policy
	Co4	My organisation has process to specify rights and restrictions for user access to data
	Co5	My organisation follows industry standards to build in security for systems
	Co6	My organisation has a secure communication channel policy

Goal: Assessing the integrity of data stored in cloud storage from the stakeholder's viewpoint

Factor	Integrity (In)	
Description	Integrity is the ability of a provider to detect changes or modifications to an original status of remote data stored in cloud storage. Some techniques implement integrity across a packet header and/or data field by creating a hash across the contents of the packet (Firesmith 2004). Most approaches ensuring confidentiality care about integrity of the data as well: Inconsistent access to the data automatically harms the modification to its correct status. Therefore, having encrypting the data before uploaded to the cloud storage safeguards the remote data. Data is therefore equipped with checksums and probes. The status of the data must be checked continuously. The remote location makes incessantly checks hard to perform. The Proof of Retrievability (POR) (Bowers et al. 2009a) tackles this problem. A Message Authentication Code (MAC) combined with an Error Correction Code (ECC) is applied on the buckets. The MAC detects large errors and is relying on units in the buckets. The ECC protects the bucket against small errors. The number of units for the MAC and the size of the ECC gives an assumption about a possible successful retrieval. This assumption is provided as probability to successfully access data. This technique can be combined with a cloud-of-clouds approach (Bowers et al. 2009b).The proof of data possession (PDP) (Ateniese et al. 2007) represent a similar approach offering a probability of possession using sampling. Current approaches working with single clouds are used mainly for synchronisation (Mahajan et al. 2011). Checksums are combined with encryption to guard data (Yao et al. 2010). These approaches use the versioned data by generating a chain of hashes. Other approaches (Wang et al. 2012) combine sampling with erasure codes similar to the POR but are working on single clouds only. The usage of multiple clouds (Abu-Libdeh et al. 2010, Cachin and Haas 2010, Mu et al. 2012) needs sophisticated integrity checks guarding the data against single, faulty clouds. All data receives a version number ensuring reliability additionally to the availability.	
Goal	Assessing the integrity of data	
Question	What are the security measures for securing the integrity of data cloud storage?	
Metric	Security measure affecting integrity of data in cloud storage	
Items	In1	My organisation has documentation of encryption management practices/guidelines
	In2	My organisation encrypt user data at rest (on disk/storage) within the environment
	In3	My organisation leverage encryption to protect virtual machine images during transport between hypervisor instances
	In4	My organisation supports data stewardship
	In5	My organisation support secure deletion (e.g., degaussing/cryptographic wiping) of archived data
	In6	My organisation supports additional protection for user to store sensitive data

Goal: Assessing the availability of data stored in cloud storage from the stakeholder's viewpoint

Factor	Availability (Av)
Description	Availability guarantees the access to the data. Availability on the server-side is hard to be assured from a users' perspective. The status of the cloud as well as the connectivity stays out of focus of a user. Measures to increase availability include mirroring the data in multiple clouds as well as local caching. Availability is the idea that the data is accessible to all authorised users at all times. Its unavailability may occur in a physical way, as the failure of critical network components, power disruptions, and physical plant

	disruptions, either malicious or natural (Firesmith 2004, Takabi et al. 2010). Availability can also be impacted in a logical way, in the form of improper addressing or routing, and through the use of Denial-Of-Service attacks, which are the deliberate insertion of unwanted data into the network (Brock and Goscinski 2010). This is often associated with address spoofing, which associates the introduction of unwanted data with a trusted end node. Zissis and Lekkas (2012) recommended a cloud system designed and maintained with important aspects, which include contingency planning for power failures and disaster recovery, is also part of a system availability (Firesmith 2004, Mapp et al. 2014). The availability of the data is nevertheless dictated by cloud storage providers. Ways to overcome this dependability are local caching of the data and/or the usage of multiple clouds (Cachin and Haas 2010). Similar to the idea of RAID, several approaches distribute the data in disjoint clouds (Abu-Libdeh et al. 2010, Cachin and Haas 2010). The “Proof of Retrievability” (POR) (Juels and Jr 2007, Bowers et al. 2009a) or “Proof of Data Possession” (PDP)(Bowers et al. 2009a) generates knowledge about the integrity and thereby indirectly about the availability. These approaches focus on cloud storage only. Similar techniques also exist in the area of P2P storage (Caronni and Waldvogel 2003).	
Goal	Assessing the availability of data	
Question	What are the security measures for ensuring availability of data cloud storage?	
Metric	Security measure affecting availability of data in cloud storage	
Items	Av1	My organisation segments data logically for each user
	Av2	My organisation has backup or redundancy mechanisms
	Av3	My organisation has data recovery mechanisms
	Av4	My organisation verifies data authenticity after restore process
	Av5	My organisation documents the restore or redundancy mechanisms
	Av6	My organisation has define restore procedure for responding to requests for user data from governments or third parties

Goal: Assessing the non-repudiation of data stored in cloud storage from the stakeholder’s viewpoint

Factor	Non-repudiation (Nr)	
Description	Non-repudiation is to assign attribution, i.e., provenance, to a data that a third party could verify and be confident that it cannot be disputed. It can also prevent a recipient from denying data was received. Firesmith (2004) highlights that non-repudiation attempts to provide a comprehensive interaction (e.g., transaction and transmission of data) is prevented from successfully repudiating (i.e., denying) any aspect of the interaction. Non-repudiation thus assumes data integrity so that a party cannot argue that the data was modified. Focusing on concurrent access, some approaches (Shraer et al. 2010) use optimistic, time-stamped writes. Combinations of integrity checks with probabilistic tests on remote data result in higher-level architectures (Kamara and Lauter 2010, Wei et al. 2010). Examples for these architectures are cloud-based file systems (Kamara et al. 2011, Stefanov and Dijk 2012, Vrable et al. 2012) or database systems (AlZain et al. 2011). These approaches use the idea of Merkle-Trees (Merkle 1988). The folder structure leverages from the tree structure in combination with remote integrity checks. Optionally, the task of integrity checks can be delegated to untrusted cloud components (Nepal et al. 2011). In this scenario, encrypting and the computation of checksums are performed by different cloud services.	
Goal	Assessing the non-repudiation of data	
Question	What are the security measures for maintaining non-repudiation of data cloud storage?	
Metric	Security measure affecting non-repudiation of data in cloud storage	
Items	Nr1	My organisation has key management policies binding keys to

		identifiable owners
	Nr2	My organisation uses synchronised time-service protocol (e.g., NTP etc.)
	Nr3	My organisation uses geographical location as an authentication
	Nr4	My organisation has restriction of user data to specific countries or geographic locations
	Nr5	My organisation provides the means for authorised individuals to determine the identity of the data producer
	Nr6	My organisation support integration of location as an authentication factor

Goal: Assessing the authenticity of data accessed and stored in cloud storage from the stakeholder's viewpoint

Factor	Authenticity (At)	
Description	Authenticity of data refers to its original conception by its owner or author. Maintaining this relationship of data and network communications is performed with the use of public key encryption and a process called digital signing (Brock and Goscinski 2010, Zissis and Lekkas 2012). To create a digital signature, a hash is created across the data. A hash ensures the data is coming from an authentic source (Mapp et al. 2014). When ownership of a digital signature secret key is bound to a specific user, it demonstrates that the data was sent by a valid user. Thus, authenticating the source of data. Focusing on collaborative use cases, the challenge is to provide a suitable key management. The key management should make use of the scalability and availability of the cloud (Cachin and Haas 2010). Challenges are especially the key distribution and flexible access to versatile groups (Popa et al. 2011). Consequently, in preserving authenticity these approaches should also satisfy integrity and availability. Authenticity must include adaptable access rights mapped to different versions.	
Goal	Assessing the authenticity of data	
Question	What are the security measures for guaranteeing authenticity of data cloud storage?	
Metric	Security measure affecting authenticity of data in cloud storage	
Items	At1	My organisation has cryptographic protection mechanisms
	At2	My organisation ensures the origin authentication
	At3	My organisation has verification assurances to ensure session authenticity
	At4	My organisation has anti-counterfeiting policy
	At5	My organisation provide session-level protection where needed
	At6	My organisation provides mechanisms to protect the authenticity of communications sessions

Goal: Assessing the reliability of service provided by cloud storage from the stakeholder's viewpoint

Factor	Reliability (Re)
Description	Reliability refers to the ability of a provider to a consistent intended service (Brock and Goscinski 2010, Zissis and Lekkas 2012). Operational reliability and flexibility is needed in cloud environments using capabilities (Mapp et al. 2014). The proposed capabilities are functions developed into mechanisms using a capability-based approach. Several different models guarantee a reliable cloud service. Examples include logging and monitoring (Ko et al. 2011a), establishing procedural approaches (Pearson et al. 2012), combinations of sampling, replaying modifications and time-stamping or establishing an entire life cycle using all of these measures (Ko et al. 2011a). Focusing on cloud storage only, modifications must be traceable by a user. Some approaches guard integrity by versioning hashes. These approaches care about reliability as well (Mahajan et al. 2011). Other approaches put the data directly under version control including adjacent metadata (Shraer et al.

	2010, Bessani et al. 2011, Stefanov and Dijk 2012).	
Goal	Assessing the reliability of data	
Question	What are the security measures for sustaining reliability of data cloud storage?	
Metric	Security measure affecting authenticity of data in cloud storage	
Items	Re1	My organisation has multi-failure disaster recovery
	Re2	My organisation has system maintenance process/policy
	Re3	My organisation has patch management policy/process
	Re4	My organisation has continuous monitoring process/solutions
	Re5	My organisation has malicious code protection mechanisms at entry and exit points
	Re6	My organisation has conducted failover test

Goal: Assessing the accountability of services provided by cloud storage from the stakeholder's viewpoint

Factor	Accountability (Ac)	
Description	Cloud storage be made accountable to their users, both the users and the cloud storage providers stand to benefit –because they can check whether their computations are being performed correctly, and the latter because they can more easily handle complaints and resolve disputes. Accountability is an opportunity for the cloud industry: it can mitigate risks for both the user and the provider, and it can enable an entirely new range of cloud-based applications (Haeberlen 2010). Accountability are provided on trust basis by having a contract or SLAs with a clear and concise definition of security policy. Accountability emphasises that trust is an important factor that must be fulfilled by providers (internal or external). They have made it compulsory for Security Level Agreement in many organisations as an extension to Service Level Agreements (SLAs). Accountability involves the processes, policies, and controls necessary to trace actions to their source. These develop trust among systems.	
Goal	Assessing the accountability of data	
Question	What are the security measures for continuing accountability of data cloud storage?	
Metric	Security measure affecting authenticity of data in cloud storage	
Items	Ac1	My organisation has process to conformance with external standards
	Ac2	My organisation has mechanisms to put internal security policies in effect
	Ac3	My organisation supports transparency and participation to conformance process with internal standards/policy
	Ac4	My organisation ensures clarity of Service Level Agreement/Guarantee (SLAs/SLGs)
	Ac5	My organisation conducts penetration tests of cloud service infrastructure regularly as prescribed by industry best practices/guidance
	Ac6	My organisation has means of remediation for internal enforcement

Goal: Assessing the auditability of data accessed and stored in cloud storage from the stakeholder’s viewpoint

Factor	Auditability (Au)	
Description	Auditability ensures that a specified organisation(s) are providing security policies to ensure work tasks follow guidelines and best practices. IT security standards are being implemented to gain confidence. The system must enable assessment, examination and audit to be completed smoothly. A cloud storage should provide full access logs, allowing organisations to see how data is being accessed, shared and used in real time. This audit data is available through APIs to real time systems allowing organisations to respond to data governance issues and provide full audit logs where required.	
Goal	Assessing the auditability of data	
Question	What are the security measures for ensuring auditability of data cloud storage?	
Metric	Security measure affecting authenticity of data in cloud storage	
Items	Au1	My organisation produces audit assertions using a structured, industry accepted format (e.g., Cloud Audit/Cloud Trust, ISACA's Cloud Computing Management Audit, etc.)
	Au2	My organisation reviews audit logs on a regular basis
	Au3	My organisation has on-demand audit review
	Au4	My organisation generates audit report
	Au5	My organisation ensures the audit log is in original content or time
	Au6	My organisation has a process to audit records for events of interest

B.5 Expert Evaluation Feedback Form

Expert Evaluation Feedback Form

Preamble: Note to Experts for Evaluation Purposes

Dear Sir/Madam

This document is referred to the Expert Evaluation Feedback Form.

You have also been provided with a document referred to as the 'SecRaS Instrument'. You have been provided with these documents because you are an expert in the field. Please evaluate the SecRaS Instrument using the evaluation criteria.

The evaluation items have criteria for evaluation as below:

Evaluation Criteria of the Items/Scale	Definition
Important	The question is essential to describe SecRaS in cloud storage systems. It must be included and its absence would affect SecRaS negatively.
Neither important nor unimportant	The question may be useful but NOT essential to describe SecRaS in cloud storage systems.
Unimportant	The question is NOT necessary to describe SecRaS in a cloud storage systems. Its absence would not affect SecRas in cloud storage systems.

This evaluation is to get the understanding whether the questions are important or not and whether the scales appropriately measuring the question or not.

Thank you for your cooperation. Your help is greatly appreciated.

Yours sincerely
Fara

This table consists of SecRas factors, the item number for each element and the evaluation criteria. The item number for each factor referred to the same number and the same factor in the SecRaS instrument.

Please evaluate the item by placing a tick (✓) in the appropriate part of the evaluation criteria.

		Evaluation Criteria of the Items		
Security Factor	Item Number	Important	Neither Important nor unimportant	Unimportant
Cloud Storage Security (CS)	CS1	○	○	○
	CS2	○	○	○
	CS3	○	○	○
	CS4	○	○	○
Confidentiality (Co)	Co1	○	○	○
	Co2	○	○	○
	Co3	○	○	○
	Co4	○	○	○
	Co5	○	○	○
	Co6	○	○	○
Integrity (In)	In1	○	○	○
	In2	○	○	○
	In3	○	○	○
	In4	○	○	○
	In5	○	○	○
	In6	○	○	○
Availability (Av)	Av1	○	○	○
	Av2	○	○	○
	Av3	○	○	○
	Av4	○	○	○
	Av5	○	○	○
	Av6	○	○	○
Non-repudiation (Nr)	Nr1	○	○	○
	Nr2	○	○	○
	Nr3	○	○	○
	Nr4	○	○	○

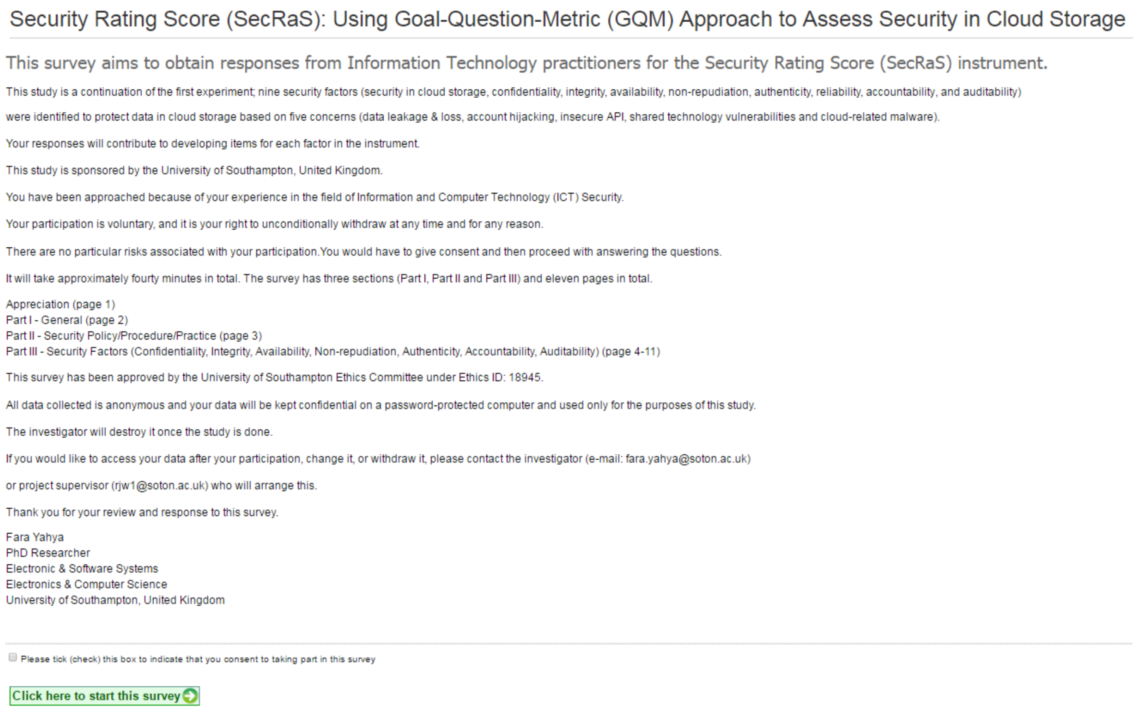
		Evaluation Criteria of the Items		
Security Factor	Item Number	Important	Neither Important nor unimportant	Unimportant
	Nr5	○	○	○
	Nr6	○	○	○
Authenticity (At)	At1	○	○	○
	At2	○	○	○
	At3	○	○	○
	At4	○	○	○
	At5	○	○	○
	At6	○	○	○
Reliability (Re)	Re1	○	○	○
	Re2	○	○	○
	Re3	○	○	○
	Re4	○	○	○
	Re5	○	○	○
	Re6	○	○	○
Accountability (AC)	Ac1	○	○	○
	Ac2	○	○	○
	Ac3	○	○	○
	Ac4	○	○	○
	Ac5	○	○	○
	Ac6	○	○	○
Auditability (Au)	Au1	○	○	○
	Au2	○	○	○
	Au3	○	○	○
	Au4	○	○	○
	Au5	○	○	○
	Au6	○	○	○
Total Item	52			

B.6 Content Validity Ratio Analysis

CVR for 52 potential items

Factor	Total of Items	Significant Items	CVR item 1	CVR item 2	CVR item 3	CVR item 4	CVR item 5	CVR item 6	Average CVR
CS	4	3	1	1	1	0.2	-	-	0.80
Co	6	6	0.6	0.6	1	1	0.6	1	0.80
In	6	6	0.6	1	0.6	0.6	1	1	0.80
Av	6	4	1	1	1	1	0.2	0.2	0.73
Nr	6	4	1	0.6	1	1	0.2	0.2	0.67
At	6	5	0.6	1	1	0.6	0.6	0.2	0.67
Re	6	5	1	0.6	1	0.6	0.6	0.2	0.67
Ac	6	5	1	0.6	1	0.6	0.6	0.0	0.63
Au	6	5	1	0.6	0.6	0.6	0.6	0.2	0.60
Total	52	43							

B.7 Improved Instrument



Security Rating Score (SecRaS): Using Goal-Question-Metric (GQM) Approach to Assess Security in Cloud Storage

This survey aims to obtain responses from Information Technology practitioners for the Security Rating Score (SecRaS) instrument.

This study is a continuation of the first experiment; nine security factors (security in cloud storage, confidentiality, integrity, availability, non-repudiation, authenticity, reliability, accountability, and auditability)

were identified to protect data in cloud storage based on five concerns (data leakage & loss, account hijacking, insecure API, shared technology vulnerabilities and cloud-related malware).

Your responses will contribute to developing items for each factor in the instrument.

This study is sponsored by the University of Southampton, United Kingdom.

You have been approached because of your experience in the field of Information and Computer Technology (ICT) Security.

Your participation is voluntary, and it is your right to unconditionally withdraw at any time and for any reason.

There are no particular risks associated with your participation.You would have to give consent and then proceed with answering the questions.

It will take approximately fourty minutes in total. The survey has three sections (Part I, Part II and Part III) and eleven pages in total.

Appreciation (page 1)
Part I - General (page 2)
Part II - Security Policy/Procedure/Practice (page 3)
Part III - Security Factors (Confidentiality, Integrity, Availability, Non-repudiation, Authenticity, Accountability, Auditability) (page 4-11)

This survey has been approved by the University of Southampton Ethics Committee under Ethics ID: 18945.

All data collected is anonymous and your data will be kept confidential on a password-protected computer and used only for the purposes of this study.

The investigator will destroy it once the study is done.

If you would like to access your data after your participation, change it, or withdraw it, please contact the investigator (e-mail: fara.yahya@soton.ac.uk)

or project supervisor (rjw1@soton.ac.uk) who will arrange this.

Thank you for your review and response to this survey.

Fara Yahya
PhD Researcher
Electronic & Software Systems
Electronics & Computer Science
University of Southampton, United Kingdom

□ Please tick (check) this box to indicate that you consent to taking part in this survey

Practitioners Survey

Part I General

1. What is your organisation domain?
   □ Academic
   □ Government
   □ Industry
   □ Others, please specify: ……………………………………………………………………

2. Which of these roles fits your fits your job description?
   □ IT/Technical (Application Security, Digital Forensics Investigation, Concern Intelligence etc.)
   □ Consultant/Advisory (Consultant, Industry Research/Analyst)
   □ Security Analyst/Expert (Information Security)
   □ Security Policy Maker (CIO, CTO, CSO, Chief Security Information Officer etc.)
   □ Security Researcher (Academic, PhD Researcher etc.)
   □ Others, please specify: ……………………………………………………………………

3. How long have you been working in computer & IT security (i.e,. cyber security?
   □ 5 years □ 6-10 years□ More than 10 years

4. What is the estimated percentage of data stored in your/your organisations' cloud storage?.
   □ 0% - 25% □ 26%- 50% □ 51%- 75% □ 76%- 100%

Part II Security Policy/Procedure/Practice

The Security Rating Score (SecRaS) has been developed with the idea that IT security policy, procedure and practice are in place. Please refer to the rating score table shown below.

5. These are a set of questions that describes whether policy, procedure and practice are implemented in their organisation. Please indicate your agreement.

Items	Rating Score				
	No Plan to Implement	Planning to Implement	Do Not Know	Partially Implement	Fully Implement
My organisation has cloud storage security policy in place	○	○	○	○	○
Please provide details: ________________________ Example: The organisation is ISO27001 compliance					
My organisation has implemented security procedures to comply to industry standards	○	○	○	○	○
Please provide details: ________________________					
My organisation has imposed security controls to support cloud security policy/ies	○	○	○	○	○
Please provide details: _______________________					

Part III Security Factors

6. This is an overview of the Security Rating Score (SecRaS). SecRaS consists of eight factors (Confidentiality, Integrity, Availability, Non-repudiation, Authenticity, Accountability, and Auditability) and 40 items. Please indicate your agreement.

Items	Rating Score				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Confidentiality					
My organisation has identity management policy	○	○	○	○	○
My organisation has user-based authentication process	○	○	○	○	○
My organisation has access management policy	○	○	○	○	○
My organisation has process to specify rights and restrictions for user access to data	○	○	○	○	○
My organisation follows industry standards to build in security for systems	○	○	○	○	○
My organisation has a secure communication channel policy	○	○	○	○	○
Integrity					
My organisation has documentation of encryption management practices/guidelines	○	○	○	○	○
My organisation encrypt user data at rest (on disk/storage) within the environment	○	○	○	○	○
My organisation leverage encryption to protect virtual machine images during transport between hypervisor instances	○	○	○	○	○
My Organisation supports data stewardship	○	○	○	○	○
My organisation support secure deletion (e.g., degaussing/cryptographic wiping) of archived data	○	○	○	○	○
My organisation supports additional protection for user to	○	○	○	○	○

Items	Rating Score				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
store sensitive data					
Availability					
My organisation segments data logically for each user	○	○	○	○	○
My organisation has backup or redundancy mechanisms	○	○	○	○	○
My organisation has data recovery mechanisms	○	○	○	○	○
My organisation verifies data authenticity after restore process	○	○	○	○	○
Non-Repudiation					
My organisation has key management policies binding keys to identifiable owners	○	○	○	○	○
My organisation uses synchronised time-service protocol (e.g., NTP etc.)	○	○	○	○	○
My organisation uses geographical location as an authentication	○	○	○	○	○
My organisation has restriction of user data to specific countries or geographic locations	○	○	○	○	○
Authenticity					
My organisation has cryptographic protection mechanisms	○	○	○	○	○
My organisation ensures the origin authentication	○	○	○	○	○
My organisation has verification assurances to ensure session authenticity	○	○	○	○	○
My organisation has anti-counterfeiting policy	○	○	○	○	○
My organisation provide session-level protection where needed	○	○	○	○	○
Reliability					
My organisation has multi-failure disaster recovery	○	○	○	○	○
My organisation has system maintenance process/policy	○	○	○	○	○
My organisation has patch management policy/process	○	○	○	○	○
My organisation has continuous monitoring process/solutions	○	○	○	○	○
My organisation has malicious code protection mechanisms at entry and exit points	○	○	○	○	○
Accountability					
My organisation has process to conformance with external standards	○	○	○	○	○
My organisation has mechanisms to put internal security policies in effect	○	○	○	○	○
My organisation supports transparency and participation to conformance process with internal standards/policy	○	○	○	○	○
My organisation ensures clarity	○	○	○	○	○

Items	Rating Score				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
of Service Level Agreement/Guarantee (SLAs/SLGs)					
My organisation conducts penetration tests of cloud service infrastructure regularly as prescribed by industry best practices/guidance	○	○	○	○	○
Auditability					
My organisation produces audit assertions using a structured, industry accepted format (e.g., Cloud Audit/Cloud Trust, ISACA's Cloud Computing Management Audit, etc.)	○	○	○	○	○
My organisation reviews audit logs on a regular basis	○	○	○	○	○
My organisation has on-demand audit review	○	○	○	○	○
My organisation generates audit report	○	○	○	○	○
My organisation ensures the audit log is in original content or time	○	○	○	○	○

B.8 Practitioner’s Validation Study

B.8.1 Correlation Analysis

Correlations

		Co_1.1	Co_1.2	Co_1.3	Co_1.4	Co_1.5	Co_n1.6
Co_1.1	Pearson Correlation	1	.707**	.537**	.482**	.484**	-.363*
	Sig. (2-tailed)		.000	.002	.007	.007	.049
	N	30	30	30	30	30	30
Co_1.2	Pearson Correlation	.707**	1	.576**	.625**	.363*	-.314
	Sig. (2-tailed)	.000		.001	.000	.048	.092
	N	30	30	30	30	30	30
Co_1.3	Pearson Correlation	.537**	.576**	1	.628**	.546**	-.344
	Sig. (2-tailed)	.002	.001		.000	.002	.062
	N	30	30	30	30	30	30
Co_1.4	Pearson Correlation	.482**	.625**	.628**	1	.661**	-.452*
	Sig. (2-tailed)	.007	.000	.000		.000	.012
	N	30	30	30	30	30	30
Co_1.5	Pearson Correlation	.484**	.363*	.546**	.661**	1	-.400*
	Sig. (2-tailed)	.007	.048	.002	.000		.029
	N	30	30	30	30	30	30
Co_n1.6	Pearson Correlation	-.363*	-.314	-.344	-.452*	-.400*	1
	Sig. (2-tailed)	.049	.092	.062	.012	.029	
	N	30	30	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Correlations

		In_1.3	In_1.4	In_1.5	In_2.1	In_2.2	In_2.3
In_1.3	Pearson Correlation	1	.878**	.592**	.683**	.683**	.611**
	Sig. (2-tailed)		.000	.001	.000	.000	.000
	N	30	30	30	30	30	30
In_1.4	Pearson Correlation	.878**	1	.754**	.715**	.715**	.708**
	Sig. (2-tailed)	.000		.000	.000	.000	.000
	N	30	30	30	30	30	30
In_1.5	Pearson Correlation	.592**	.754**	1	.627**	.573**	.675**

	Sig. (2-tailed)	.001	.000		.000	.001	.000
	N	30	30	30	30	30	30
In_2.1	Pearson Correlation	.683**	.715**	.627**	1	.889**	.884**
	Sig. (2-tailed)	.000	.000	.000		.000	.000
	N	30	30	30	30	30	30
In_2.2	Pearson Correlation	.683**	.715**	.573**	.889**	1	.884**
	Sig. (2-tailed)	.000	.000	.001	.000		.000
	N	30	30	30	30	30	30
In_2.3	Pearson Correlation	.611**	.708**	.675**	.884**	.884**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	
	N	30	30	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations

		Av_2.2	Av_2.3	Av_3.2	Av_3.4
Av_2.2	Pearson Correlation	1	.922**	.587**	.536**
	Sig. (2-tailed)		.000	.001	.002
	N	30	30	30	30
Av_2.3	Pearson Correlation	.922**	1	.507**	.495**
	Sig. (2-tailed)	.000		.004	.005
	N	30	30	30	30
Av_3.2	Pearson Correlation	.587**	.507**	1	.777**
	Sig. (2-tailed)	.001	.004		.000
	N	30	30	30	30
Av_3.4	Pearson Correlation	.536**	.495**	.777**	1
	Sig. (2-tailed)	.002	.005	.000	
	N	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations

		Nr_1.1	Nr_1.2	Nr_1.3	Nr_3.1
Nr_1.1	Pearson Correlation	1	.807**	.651**	.543**
	Sig. (2-tailed)		.000	.000	.002
	N	29	29	29	29
Nr_1.2	Pearson Correlation	.807**	1	.896**	.385*
	Sig. (2-tailed)	.000		.000	.039

	N	29	29	29	29
Nr_1.3	Pearson Correlation	.651**	.896**	1	.204
	Sig. (2-tailed)	.000	.000		.288
	N	29	29	29	29
Nr_3.1	Pearson Correlation	.543**	.385*	.204	1
	Sig. (2-tailed)	.002	.039	.288	
	N	29	29	29	29

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Correlations

		At_1.1	At_1.2	At_1.3	At_2.1	At_3.1
At_1.1	Pearson Correlation	1	.817**	.758**	.549**	.460*
	Sig. (2-tailed)		.000	.000	.002	.010
	N	30	30	30	30	30
At_1.2	Pearson Correlation	.817**	1	.729**	.424*	.350
	Sig. (2-tailed)	.000		.000	.019	.058
	N	30	30	30	30	30
At_1.3	Pearson Correlation	.758**	.729**	1	.688**	.604**
	Sig. (2-tailed)	.000	.000		.000	.000
	N	30	30	30	30	30
At_2.1	Pearson Correlation	.549**	.424*	.688**	1	.536**
	Sig. (2-tailed)	.002	.019	.000		.002
	N	30	30	30	30	30
At_3.1	Pearson Correlation	.460*	.350	.604**	.536**	1
	Sig. (2-tailed)	.010	.058	.000	.002	
	N	30	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Correlations

		Re_1.1	Re_1.3	Re_2.2	Re_2.3	Re_3.2
Re_1.1	Pearson Correlation	1	.690**	.359	.365*	.567**
	Sig. (2-tailed)		.000	.051	.047	.001
	N	30	30	30	30	30
Re_1.3	Pearson Correlation	.690**	1	.357	.570**	.612**
	Sig. (2-tailed)	.000		.053	.001	.000

	N	30	30	30	30	30
Re_2.2	Pearson Correlation	.359	.357	1	.730**	.696**
	Sig. (2-tailed)	.051	.053		.000	.000
	N	30	30	30	30	30
Re_2.3	Pearson Correlation	.365*	.570**	.730**	1	.785**
	Sig. (2-tailed)	.047	.001	.000		.000
	N	30	30	30	30	30
Re_3.2	Pearson Correlation	.567**	.612**	.696**	.785**	1
	Sig. (2-tailed)	.001	.000	.000	.000	
	N	30	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Correlations

		Ac_1.1	Ac_1.2	Ac_1.4	Ac_2.2	Ac_2.4
Ac_1.1	Pearson Correlation	1	.750**	.768**	.776**	.883**
	Sig. (2-tailed)		.000	.000	.000	.000
	N	30	30	30	30	30
Ac_1.2	Pearson Correlation	.750**	1	.854**	.584**	.592**
	Sig. (2-tailed)	.000		.000	.001	.001
	N	30	30	30	30	30
Ac_1.4	Pearson Correlation	.768**	.854**	1	.637**	.656**
	Sig. (2-tailed)	.000	.000		.000	.000
	N	30	30	30	30	30
Ac_2.2	Pearson Correlation	.776**	.584**	.637**	1	.871**
	Sig. (2-tailed)	.000	.001	.000		.000
	N	30	30	30	30	30
Ac_2.4	Pearson Correlation	.883**	.592**	.656**	.871**	1
	Sig. (2-tailed)	.000	.001	.000	.000	
	N	30	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations

		Au_n1.6	Au_2.1	Au_2.2	Au_3.1	Au_3.3
Au_n1.6	Pearson Correlation	1	-.579**	-.579**	-.681**	-.709**
	Sig. (2-tailed)		.001	.001	.000	.000
	N	29	29	29	29	29

Au_2.1	Pearson Correlation	-.579**	1	.739**	.528**	.533**
	Sig. (2-tailed)	.001		.000	.003	.003
	N	29	29	29	29	29
Au_2.2	Pearson Correlation	-.579**	.739**	1	.812**	.774**
	Sig. (2-tailed)	.001	.000		.000	.000
	N	29	29	29	29	29
Au_3.1	Pearson Correlation	-.681**	.528**	.812**	1	.913**
	Sig. (2-tailed)	.000	.003	.000		.000
	N	29	29	29	29	29
Au_3.3	Pearson Correlation	-.709**	.533**	.774**	.913**	1
	Sig. (2-tailed)	.000	.003	.000	.000	
	N	29	29	29	29	29

**. Correlation is significant at the 0.01 level (2-tailed).

B.8.2 Internal Reliability

Reliability Statistics

Cronbach's Alpha	N of Items
.954	43

Reliability Statistics

Cronbach's Alpha	N of Items
.933	3

Reliability Statistics

Cronbach's Alpha	N of Items
.749	6

Reliability Statistics

Cronbach's Alpha	N of Items
.936	6

Reliability Statistics

Cronbach's Alpha	N of Items
.874	4

Reliability Statistics

Cronbach's Alpha	N of Items
.841	4

Reliability Statistics

Cronbach's Alpha	N of Items
.879	5

Reliability Statistics

Cronbach's Alpha	N of Items
.867	5

Reliability Statistics

Cronbach's Alpha	N of Items
.933	5

Reliability Statistics

Cronbach's Alpha	N of Items
.518	5

Appendix C Exploratory Factor Analysis

C.1 KMO and Bartlett's Test

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.856
Bartlett's Test of Sphericity	Approx. Chi-Square	7818.674
	df	903
	Sig.	.000

C.2 Total Variance Explained

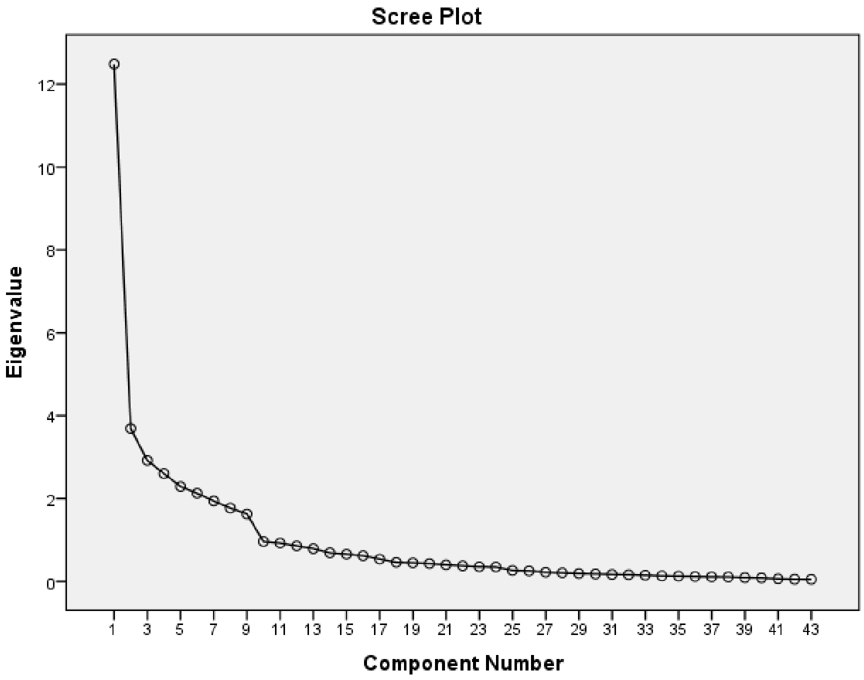
Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	12.485	29.034	29.034	12.485	29.034	29.034	4.429	10.300	10.300
2	3.690	8.582	37.616	3.690	8.582	37.616	4.064	9.451	19.751
3	2.917	6.783	44.399	2.917	6.783	44.399	3.704	8.613	28.364
4	2.603	6.052	50.451	2.603	6.052	50.451	3.669	8.532	36.896
5	2.289	5.324	55.775	2.289	5.324	55.775	3.621	8.421	45.317
6	2.125	4.942	60.717	2.125	4.942	60.717	3.268	7.600	52.917
7	1.940	4.512	65.230	1.940	4.512	65.230	3.062	7.122	60.039
8	1.768	4.113	69.342	1.768	4.113	69.342	2.952	6.865	66.904
9	1.626	3.782	73.124	1.626	3.782	73.124	2.675	6.220	73.124
10	.961	2.234	75.358						
11	.925	2.150	77.508						
12	.857	1.992	79.501						
13	.790	1.838	81.339						
14	.688	1.600	82.938						
15	.655	1.524	84.462						
16	.620	1.441	85.903						
17	.537	1.249	87.153						
18	.460	1.069	88.222						
19	.447	1.039	89.262						
20	.430	1.000	90.262						
21	.399	.928	91.190						
22	.376	.875	92.065						

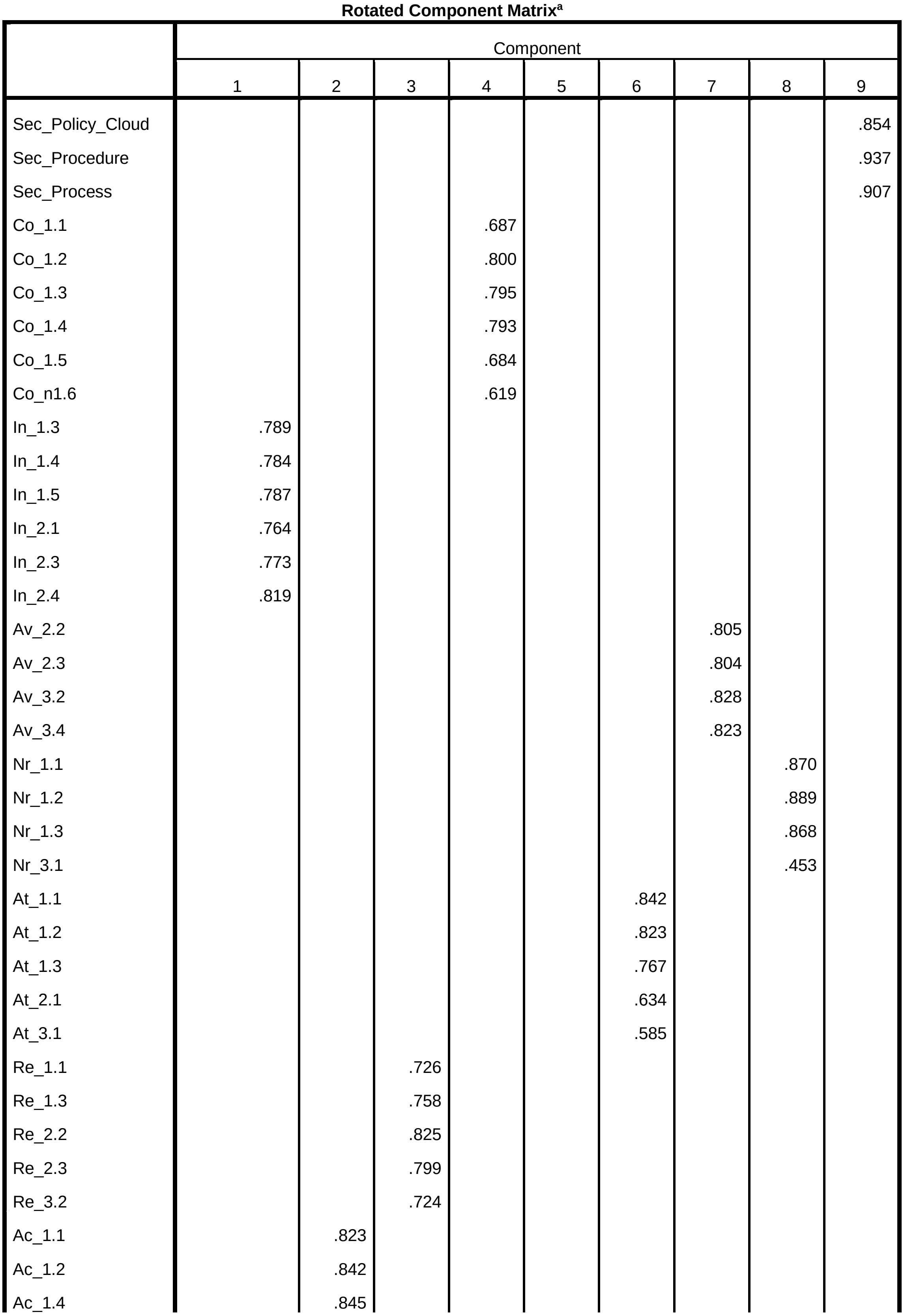
23	.351	.817	92.881						
24	.347	.807	93.689						
25	.264	.614	94.303						
26	.251	.585	94.888						
27	.221	.514	95.402						
28	.207	.482	95.884						
29	.189	.440	96.323						
30	.179	.417	96.740						
31	.166	.387	97.127						
32	.159	.370	97.496						
33	.148	.344	97.840						
34	.133	.309	98.149						
35	.127	.296	98.445						
36	.117	.273	98.717						
37	.109	.254	98.971						
38	.106	.247	99.219						
39	.090	.209	99.428						
40	.083	.194	99.622						
41	.063	.147	99.769						
42	.052	.120	99.889						
43	.048	.111	100.000						

Extraction Method: Principal Component Analysis.

C.3 Scree plot



C.4 Rotated Component Matrix



Rotated Component Matrix[a]

	Component								
	1	2	3	4	5	6	7	8	9
Sec_Policy_Cloud									.854
Sec_Procedure									.937
Sec_Process									.907
Co_1.1				.687					
Co_1.2				.800					
Co_1.3				.795					
Co_1.4				.793					
Co_1.5				.684					
Co_n1.6				.619					
In_1.3	.789								
In_1.4	.784								
In_1.5	.787								
In_2.1	.764								
In_2.3	.773								
In_2.4	.819								
Av_2.2							.805		
Av_2.3							.804		
Av_3.2							.828		
Av_3.4							.823		
Nr_1.1								.870	
Nr_1.2								.889	
Nr_1.3								.868	
Nr_3.1								.453	
At_1.1						.842			
At_1.2						.823			
At_1.3						.767			
At_2.1						.634			
At_3.1						.585			
Re_1.1			.726						
Re_1.3			.758						
Re_2.2			.825						
Re_2.3			.799						
Re_3.2			.724						
Ac_1.1		.823							
Ac_1.2		.842							
Ac_1.4		.845							

Ac_2.2		.841							
Ac_2.4		.841							
Au_n1.6					.597				
Au_2.1					.731				
Au_2.2					.820				
Au_3.1					.867				
Au_3.3					.853				

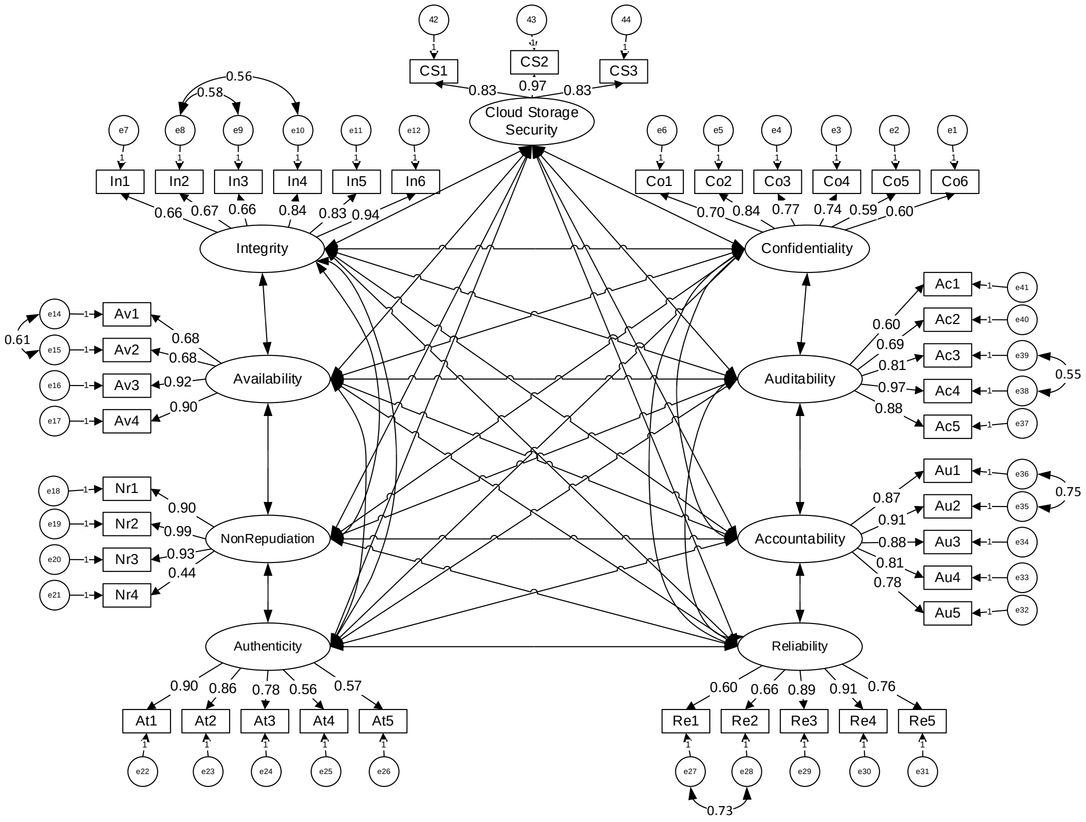
Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

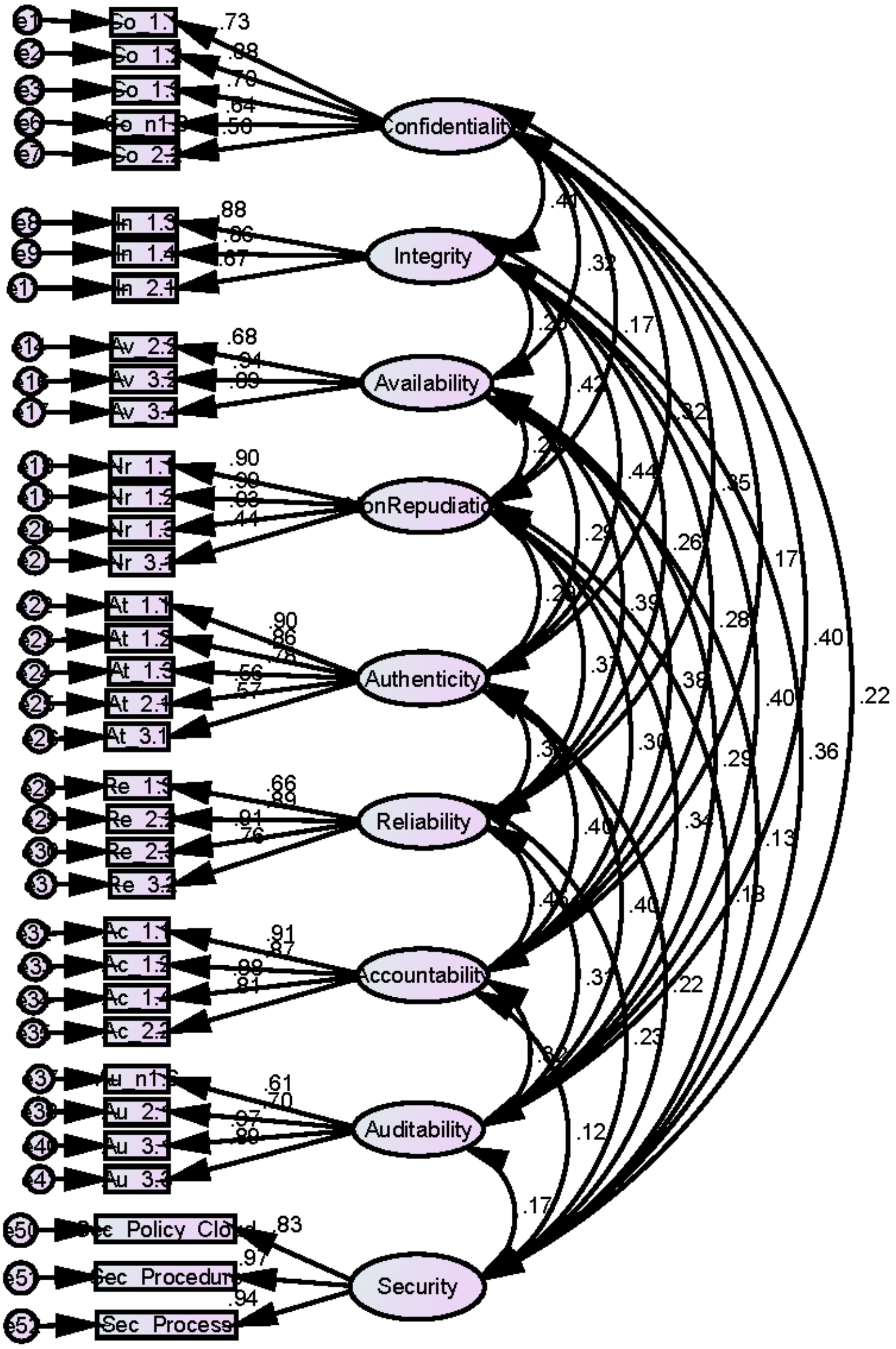
a. Rotation converged in 7 iterations.

Appendix D Structural Equation Modelling

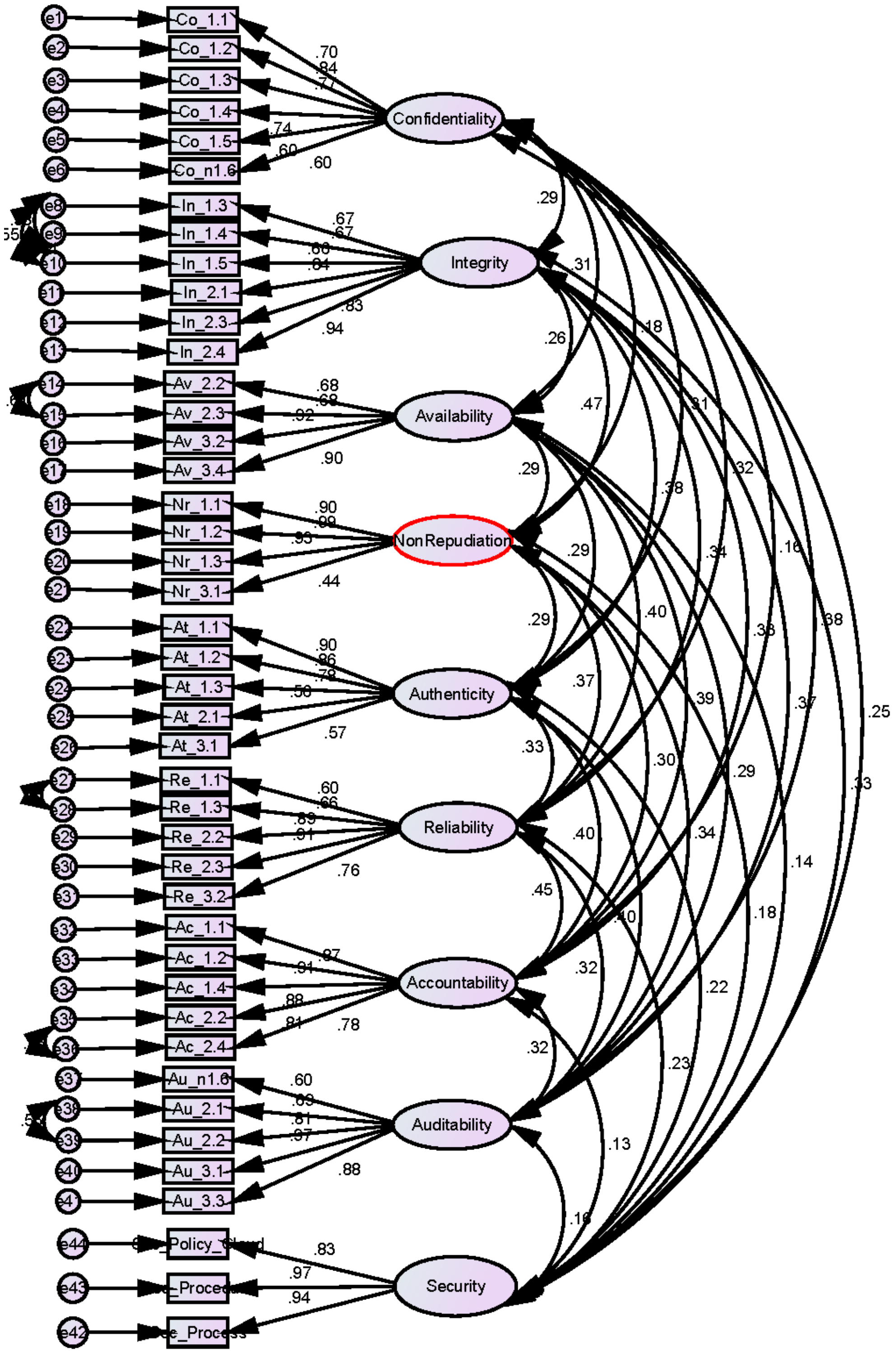
D.1 Confirmatory Factor Analysis



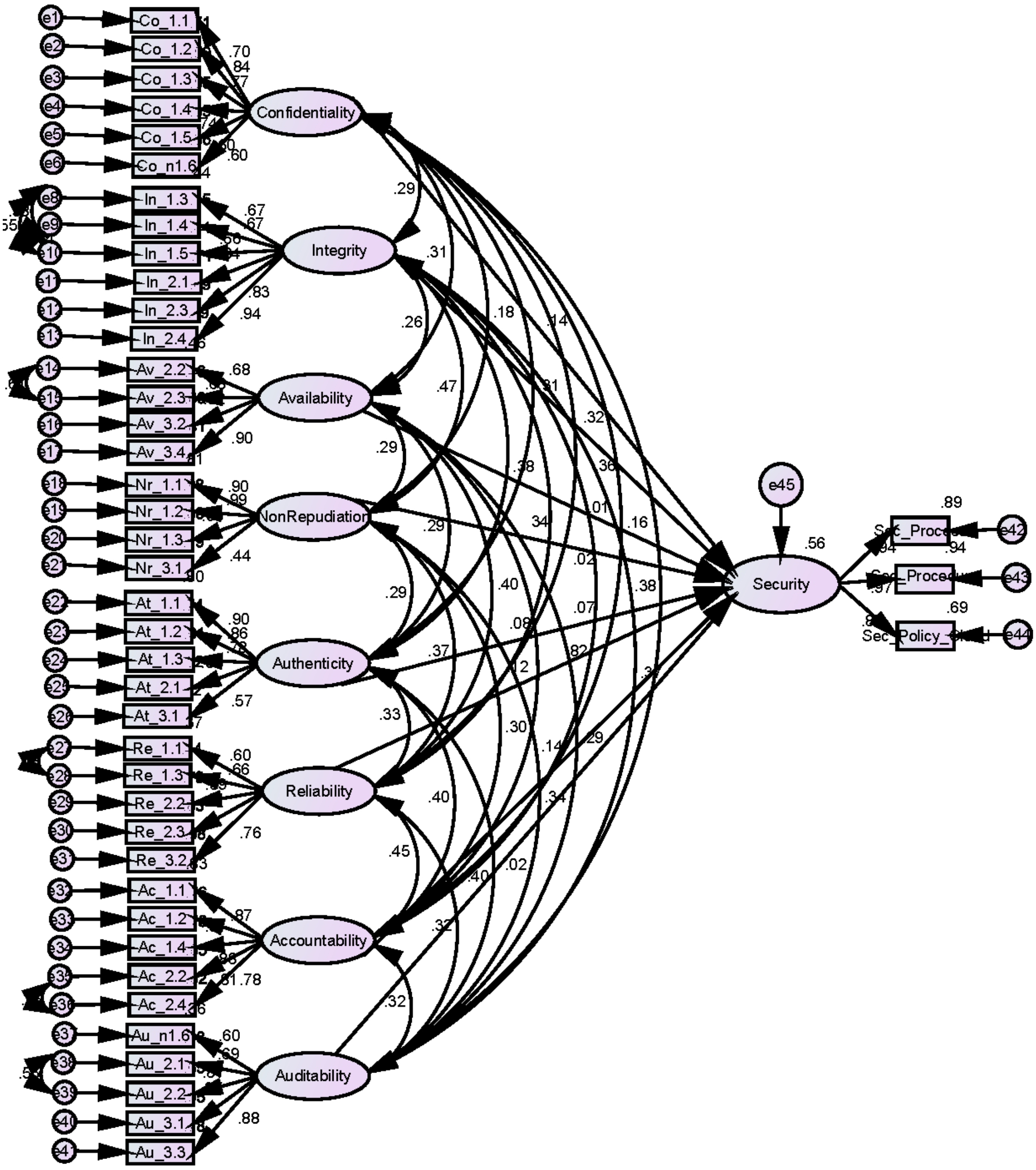
D.2 Confirmatory measurement with AMOS



D.3 Confirmatory measurement with modification indices



D.4 Structural model with AMOS



CMIN/DF = 1.650
CFI = 0.932
RMSEA = 0.055
PCLOSE = 0.083
SRMR = 0.0703

D.5 Discriminant validity analysis

	CR	AVE	MSV	MaxR (H)	Co	In	Av	Nr	At	Re	Ac	Au	CS
Co	0.860	0.510	0.142	0.995	0.714								
In	0.910	0.630	0.219	0.961	0.289	0.775							
Av	0.878	0.648	0.160	0.923	0.311	0.264	0.805						
Nr	0.902	0.712	0.219	0.989	0.183	0.469	0.294	0.844					
At	0.855	0.551	0.156	0.990	0.311	0.379	0.295	0.295	0.746				
Re	0.880	0.601	0.200	0.991	0.320	0.338	0.399	0.371	0.334	0.775			
Ac	0.930	0.727	0.200	0.992	0.159	0.333	0.387	0.302	0.403	0.446	0.853		
Au	0.898	0.643	0.154	0.993	0.377	0.367	0.295	0.338	0.398	0.315	0.323	0.802	
CS	0.941	0.843	0.119	0.994	0.249	0.326	0.139	0.183	0.222	0.234	0.126	0.163	0.918

D.6 Computation of degrees of freedom (Default model)

Number of distinct sample moments:	820
Number of distinct parameters to be estimated:	115
Degrees of freedom (820 - 115):	705

D.7 Result (Default model)

Minimum was achieved
Chi-square = 1163.372
Degrees of freedom = 705
Probability level = .000

D.8 Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	115	1163.372	705	.000	1.650
Saturated model	820	.000	0		
Independence model	40	7565.640	780	.000	9.700

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.092	.794	.761	.683
Saturated model	.000	1.000		
Independence model	.394	.207	.166	.197

Baseline Comparisons

Model	NFI	RFI	IFI	TLI	CFI
	Delta1	rho1	Delta2	rho2	
Default model	.846	.830	.933	.925	.932
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.904	.765	.843
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	458.372	368.610	556.022
Saturated model	.000	.000	.000
Independence model	6785.640	6510.236	7067.552

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	5.361	2.112	1.699	2.562
Saturated model	.000	.000	.000	.000
Independence model	34.865	31.270	30.001	32.569

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.055	.049	.060	.083
Independence model	.200	.196	.204	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	1393.372	1446.951	1782.589	1897.589
Saturated model	1640.000	2022.045	4415.286	5235.286
Independence model	7645.640	7664.276	7781.020	7821.020

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	6.421	6.007	6.871	6.668
Saturated model	7.558	7.558	7.558	9.318
Independence model	35.233	33.964	36.532	35.319

HOELTER

Model	HOELTER	HOELTER
	.05	.01
Default model	144	149
Independence model	25	26