UNIVERSITY OF NAIROBI

SCHOOL OF COMPUTING AND INFORMATICS

OPTIMISING RATIONAL DECISION MAKING WHEN REASONING ABOUT ENHANCING PKI SECURITY FOR eGOVERNMENT: *A Quantitative Decision Support Approach*

by

GEOFFREY CHEMWA P80/P/83761/2012

SUPERVISOR: PROFESSOR WILLIAM OKELO-ODONGO

This Research Thesis is Submitted in Fulfillment
of the
Requirements for the Degree of Doctor of Philosophy in Information Systems
School of Computing and Informatics
University of Nairobi

DECLARATION

This research is my original work and has not been presented for a degree award or any other awards in any other university

-- -------------------------------

Geoffrey Wekesa Chemwa Date

Reg. No: P80/P/83761/2012

This thesis has been submitted with my approval as the University Supervisor;

-- -------------------------------

Professor William Okelo – Odongo Date

DEDICATION

To God in heaven for making a way where there was none. To parents Hezron and Grace for believing in me no matter what! To my family, Virginia, Cecilia, Seth and Bradley, however busy I am, you know you come first.

ABSTRACT

The security of eGovernments is a frontline issue in any meaningful discussion about trustworthy electronic transactions and service delivery. At the center of electronic service delivery is identity management. Trust can only be achieved through secure electronic identity and access management. Most eGovernments across the globe e.g. Britain, Australia, Estonia, Kenya etc. prefer implementing Public Key Infrastructures in their identity and access management systems as a means of achieving strong authentication mechanisms for its users. This is because eGovernments face massive threats from a knowledge society that has easy access to hacking knowledge and tools, and also well-funded hacker groups. These threats can easily compromise any system whose security is not properly enhanced. We are cognizant of the fact that in most governments, the planners, implementers and assessors of PKI rely on quality management systems like ISO to qualitatively measure compliance to best practices through relevant audits. Such strategies are paperwork intensive and try to ensure process adherence but lack the capacity to quantitatively measure non-functional quality properties like security, interoperability, availability, privacy, reliability, performance among others. We propose a quantitative approach when reasoning about PKI security attributes. Optimisation of decisions needed to ensure cyber secure PKI solutions for e-Government requires a good decision support system informed by quantitative measures of key security quality attributes. Although PKI is a universal concept, its design and implementation in different contexts means that each context offers emergent challenges that requires unique solutions. This thesis proposes a decision optimisation tool for PKI security derived from existing models. The research demonstrates how security can be modeled using variables that influence its optimisation in PKI solutions. The research uses regression analysis and specifically partial least squares to perform relevant inference on PKI security influencing factors and present the various statistical measures to security managers in an easy to visualize manner. The Structure Case, Culture, People, Process and Technology (CPPT) and Partial Least Squares Structural Equation Modeling (PLS-SEM) frameworks are all used in the study. The output is a generic quantitative PKI security rational decision optimisation tool.

ACKNOWLEDGEMENTS

Thank you Professor Okelo-Odongo for your dedicated guidance throughout this research study. I would not have made it without you. Thank you Dr. Wanjiku Ng'ang'a for your guidance during the early days of my study. Special thanks to the staff at the Communications Authority of Kenya, and Kenya ICT Authority, especially Evans and Esther. Thank you the staff at Kenya Revenue Authority, Department of Immigration and Registration of Persons and Huduma Centers. Without you where would I have got the data? Thank you Dr. Stephen Mburu for introducing SmartPLS to me when I really needed it. Thank you the team behind SmartPLS for giving me a student license. It really helped me in data analysis. Thank you my colleagues, friends and critics, for lending an ear and providing relevant guidance as my study progressed.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

BPEL4WS	Business Process Execution Language For Web Services
CC/PP	Composite Capabilities/Preference Profiles
CDSSO	Cross Domain Single Sign On
CoT	Circle of Trust
EAC	Extended Access Control
eIAMS	Electronic Identity and Access Management
EIS	European Information Society
FIDIS	Future of Identity in the Information Society
GoK	Government of Kenya
G2C	Government to Citizen
G2B	Government to Business
GSM	Global Systems for Mobile Communications
ICT4D	Information Communication and Technology for Development
IAM	Identity and Access Management
IIMS	Interoperable Identity Management System
IMS	Identity Management System
LDAP	Light Weight Directory Access Protocol
OASIS	Advancing Open Standards for the Information Society
P3P	Platform for Privacy Preferences Project
RFID	Radio Frequency Identification
SAML	Security Assertion Markup Language
SAT	SIM Application Toolkit

SIM	Subscriber Identity Module
SOAP	Simple Object Access Protocol
SP	Service Provider
SSO	Single Sign On
UDDI	Universal Description Discovery and Integration
WS	Web Service
XML	eXtensible Markup Language
W3C	World Wide Web Consortium

1. INTRODUCTION

the electronic person has no physical
body but still has the full rights as a
citizen

1.1 Background

E-Government (henceforth referred to as eGovernment in this thesis) is the short form for electronic government. It can be defined as the application of information and communication technologies to deliver government services and facilitate electronic interactions between the government and citizens, organisations and other countries (Satyanarayana, 2004; Waema and Adera, 2011). The security of eGovernment depends on its identity and access management strategy. Most eGovernments in the world like Australia (Australian Government, 2009a), United Arab Emirates (Al-Khouri, 2011a), Britain and Estonia (Choudhury et al., 2002) etc. prefer implementing Public Key Infrastructure (PKI) solutions to enforce strong authentication on transacting entities and create a cyber-secure environment (Al-Khouri, 2011a). PKI uses the public key cryptography scheme.

Public Key Cryptography is a cryptographic strategy in which communicating parties use a pair of keys each, one public and another private, to secure their communications over networks by encrypting and decrypting sent messages using the key pairs (RSA, 1999). Public Key Infrastructure (PKI) can be defined as a set of legally recognized information technologies and essential services which manage the issuance and use of the public-private encryption key pairs and their related digital certificates to identify and authenticate transacting entities on computer networks as a means of creating a cyber-secure and trustworthy environment (ibid). PKI can be adopted by both private and public organisations. Therefore, enhancing the security of Public Key Infrastructure (PKI) solutions for eGovernment is very important today in a world in which the digital economy is demanding protection of sensitive data from increasingly knowledgeable citizens, well-funded hacking groups and the proliferation of hacking knowledge and tools.

The term enhancing here refers to the modification of the quality attributes of an information system in order to make it more efficient or meet certain stringent conditions required to deliver maximum business value (Al-Khouri, 2011b). In this research we believe that enhanced security of PKI solutions for eGovernment can be achieved through optimising the rational decisions of PKI security managers when reasoning about the best course of action between competing alternatives.

In decision theory, "an optimal decision" is one which when chosen gives the best outcome as compared to all other available alternatives (Johnson and Ekstedt, 2007). In this study, we define PKI security rational decision optimisation as a quantitative data driven decision making approach which directly affects PKI security quality attributes in such a manner as to positively improve them or move them towards a predetermined optimal state. The study believes that if all the attributes are improved, then the overall security of the eGovernment PKI will be enhanced.

Identity Management (IM) and Access Management (AM) in PKI solutions work together to give a three sixty degrees (360^0) solution in identifying users of a system and controlling their access to ICT resources respectively. Identity and Access Management (IAM) is the set of formal procedures that collectively identify entities and manage access to information resources, enabling only authorized entities to access online resources at the correct time for the right reasons (Ernst & Young, 2013; Wagner, 2010). Electronic Identity and Access Management (eIAM) forms the foundation of delivering public electronic services like e-taxation, e-procurement, e-import/export management, e-land administration, e-social security management, e- health management, e-passports and driving license among others. A public-private partnership may lead to e-banking, e-insurance, e-payments etc. and other private services linking to the national identity infrastructure to authenticate citizens during critical transactions as a means of creating trust between the transacting entities.

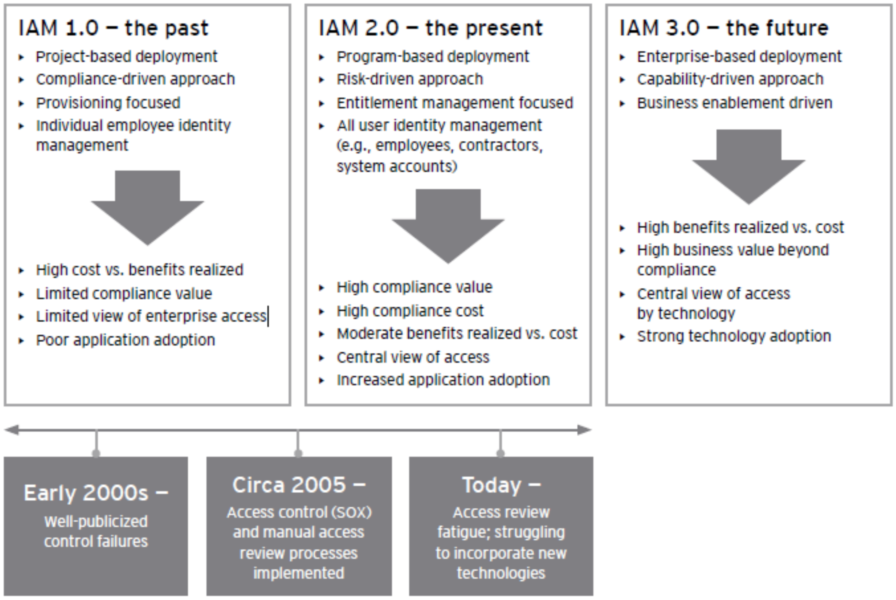
The main goal of this study was to develop a quantitative rational decision optimisation tool that would help managers to reason about how best to enhance the security of eGovernment PKI solutions by supporting their rational decision making process. In so doing, the research proposes that the security of the PKI solution would be moved towards optimisation. One of the main methods of optimising information security decisions relies on checklists which provide Yes/No answers to recognized standards compliance queries (ISC, 2003). The checklists are used to test security controls in information systems (ISACA, 2015) and are widely applied in procedures such as automated software tools, documentation reviews, walk-throughs, inquiry and observation (ibid). Although the Yes/No values indicate compliance deficiencies, they lack ability to represent these deficiencies using quantitative or statistical measures that can depict the level of attainment of various security quality attributes in the solutions. This research addresses this gap.

This study proceeded through three phases. It distilled key PKI security quality requirements and their attributes from literature. Using these attributes, a conceptual framework was formulated. The research then used partial least squares structural equation modeling to capture the causal relations in the conceptual framework. Data collection tools captured the key data for each attribute modeled. The models effectiveness was evaluated through data collection and analysis. Data was collected using questionnaire led face to face interviews with experts in the field.

Gaps and opportunities at each stage not only helped in the initial determination of various prior quantitative measures of the quality factors but also informed on what needed to be done by stakeholders. For example, knowing that a system administrator was highly trained in ICT security gave a “high” rating when thinking about his/her capability in designing security solutions and a “low” if the person did not have relevant training.

IAM in Government needs to create an enabling environment where both physical and electronic identities interoperate creating a kind of “identity ecosystem” since not all identity attributes of a person can be digitised (Smedinghoff, 2010). A cameo preview of the evolution of identity and access management shows a general trend – a move from compliance driven to risk reduction driven approaches. Futuristic systems are envisioned to be more capability driven to meet mission critical business needs (Ernst & Young, 2013). The past is full of siloed systems i.e. implemented in separate Government service points, usually controlling a few applications and focused on provisioning technology (Lips and Pang, 2008). Government agencies not only need huge efforts to understand and meet stringent compliance requirements but after attaining it they need to remain compliance savvy. Targets include reducing risks posed by new technologies and their costs. The drive towards centrally controlled systems that are standardized and automated is still hindered by siloed administrative functions, and management of time consuming processes such as manual approval, provisioning and access control (Ernst & Young, 2013). Siloed administrative functions in matters identity means that different government agencies require different identity credentials from citizens, hence creating identity islands within the same government or organisation. Mobile and cloud computing is revolutionizing IAM yet again, with future focus shifting from compliance-based programs to personalized applications and business capabilities enablement. Figure 1.1 summarizes key features in the IAM evolution path.

Figure 1.1: Evolution of IAM Systems


IAM 1.0 – the past
Project-based deployment
Compliance-driven approach
Provisioning focused
Individual employee identity management
High cost vs. benefits realized
Limited compliance value
Limited view of enterprise access
Poor application adoption
IAM 2.0 – the present
Program-based deployment
Risk-driven approach
Entitlement management focused
All user identity management (e.g., employees, contractors, system accounts)
High compliance value
High compliance cost
Moderate benefits realized vs. cost
Central view of access
Increased application adoption
IAM 3.0 – the future
Enterprise-based deployment
Capability-driven approach
Business enablement driven
High benefits realized vs. cost
High business value beyond compliance
Central view of access by technology
Strong technology adoption
Early 2000s –
Well-publicized control failures
Circa 2005 –
Access control (SOX) and manual access review processes implemented
Today –
Access review fatigue; struggling to incorporate new technologies

Source: Ernst & Young, 2013

In IAM 1.0 various IAM efforts are seen as independent projects usually with limited scope targeting single or few agencies and are compliance driven. The cost of compliance versus benefits is usually high. IAM 2.0 moves from the project view to the program view where similar projects are grouped together under one program and there are concerted efforts across board to move beyond compliance by adopting a risk-based approach hence enabling realization of moderate benefits. Lastly but not least, IAM 3.0 focuses on enterprise based deployment in which a global goals of the organisation are taken into account in order to harness all the available capabilities for high business value beyond just compliance. These efforts lead to a realization of a centralised view of IAM that would drive all profitable identity and access management reliant services.

Kenya is still at IAM 1.0 although it is quietly but surely setting the stage for IAM 2.0. Various identity projects like the public key infrastructure, biometric citizen digital database, electronic passports, electronic identity cards, electronic driving license among others, are either already live or are being developed in the background. Kenya has strongly continued adoption of ICT across the economy. For example, the eCitizen portal has so far registered 1.7 million Kenyans and has supported over 2.4 million electronic transactions which have generated Ksh. 4.2 Billion for the government. The Kenya National Single Window System (Tradenet System) meant to facilitate international trade through electronic transactions has seventeen modules (17) operational with the remaining three (3) projected to come live by December, 2016 (GoK Ministry of Finance, 2016). However, all these are not PKI enabled and citizens still have a mix of electronic and paper identities e.g. some have electronic passports, others registered mobile SIM cards, paper/plastic identity cards, Driving Lisences, PINs, birth certificates etc.

The Government is trying to break down siloed administration through creation of centralised service centers (Huduma Service eCenters) where key services can be accessed under one roof. The term siloed administration here means each government agency or department authenticates entities separately i.e. requiring its own set of identity documents from citizens before delivering services to them. At the same time, implementation of an electronic citizen database and public key infrastructure is at its last stages. It is planned that in the near future, a biometrics driven database which will act as a focal point for rolling out the new smartcard based identity cards with digital certificates among others for all Kenyans will be created. This would move the country towards IAM 2.0. Table 1.1 below highlights some of these efforts:

Table 1.1: eGovernment Initiatives in Kenya

Initiative/Target Year of Completion	**Agency(ies)**	**Status**	**Challenges**	**Main Aim(s)**
Public Key Infrastructure (PKI) / October 2013	Ministry of ICT, Communications Authority of Kenya, Kenya ICT Authority.	Implemented. Waiting for electronic citizen database and National Payment Gateway.	High cost required to start certificate authorities. Devolved government and agencies.	Enforce cyber security and non-repudiation in electronic transactions.
Electronic Citizen Database / March 2015	Ministry of Interior and Coordination of National Government, Kenya Citizens and Foreign Nationals Management Service	Final development/ implementation/ testing stages.	High cost of doing physical authentication of those Kenyans without ID's, Birth Certificates or who acquired them fraudulently	Tackle national security challenges
Electronic Identity Cards.	Ministry of Interior and Coordination of National Government, Kenya Citizens and Foreign Nationals Management Service	Final development/ implementation/ testing stages	High cost of creating electronic ID's.	Tackle national security challenges

Source: Kenya Integrated Population and Registration System, 2014

1.2 Problem Statement

Most eGovernments are adopting Public Key Infrastructure (PKI) solutions within their identity and access management systems as a means of enforcing strong authentication of entities wishing to access electronic services or transact online (Al-Khouri, 2011a). The security of PKI solutions in eGovernment is therefore increasingly coming under serious threat due to the rise of the knowledge society which has ready access to hacking knowledge and tools. Of greater concern are profit driven hackers and well-funded foreign government hacker groups perpetrating cyber warfare (Aaviksoo, 2010). Enhancing PKI security at all times is therefore paramount in maintaining a cyber-secure environment where eGovernment can thrive.

One of the main methods of ensuring PKI security enhancement relies on providing decision support information to security managers collected through system audits which are process centered (ISO or other standards driven), and whose main tools are checklists (ISC, 2003; Goulet, 2009). These audits can be done quarterly, yearly or after every three (3) years (Federal Public Key Infrastructure Authority, 2015), depending on policy guidelines adopted and relevant budget cycles. Although checklists are good, they present their information to the decision maker mainly in a Yes/No format indicating compliance or non-compliance to standards and procedures. This lacks quantitative or statistical inferencing on the studied security attributes and their level of attainment in the solutions.

This research proposes a new way of visualizing the level of compliance to PKI security standards by converting the data collected using checklists, walkthroughs, interviews etc. into quantitative measures i.e. the research demonstrates that during data collection, the data can be represented based on an interval Likart scale. The data can then be processed using regression analysis and best fit methods to generate relevant statistical measures on the attributes and their causal relationships. Rational decision makers can then use the generated measures to quickly make best decisions that enhance the PKI security hence move it towards the required optimum state.

1.3 Purpose of the Study

The purpose of the study is to develop a decision support tool that would optimise rational decisions made on PKI security for e-Government.

Although Kenya is ranked one hundred and nineteen (119) in the annual United Nations E-Government Survey (UNPAN, 2014), many electronic identity and access management initiatives are still at the infancy stage e.g. a discussion with eGovernment experts in Kenya revealed that up to the date of publication of this thesis, eGovernment initiatives like iTax and eCitizen do not share a global, centralised, and interoperable identity repository for citizens. The only credentials that have been digitized in Kenya are the Passport (in line with International Civil Aviation (ICAO) standards) and the voter’s card. The process of applying for the passport, title deed and driving license in Kenya is also semi-automated, through the eCitizen portal. The portal is still on the open internet and not behind a PKI enabled gateway since users just require a username and password to log in and don’t need a digital certificate to transact any business on the platform.

There are many challenges that face implementation of PKI solutions in Government, most of them are change management related. In order for forward looking change to take place, certain tenets need to be in place including the right vision, enough resources, capable workforce and other factors as captured in the Table 1.2. The table can be interpreted as follows:

Row 1: If there exists the right vision, enough resources, a capable workforce, capable processes, a good organizational culture, the right incentives and a good action plan, then the required change in identity and access management for eGovernment can be realised.

Row 2: If all the tenets in #1 above exist but there lacks the right vision, this would breed confusion.

Row 3: If all the tenets in #1 above exist but there lacks enough resources, this would breed anxiety and frustration within the project team.

Row 4: If all the tenets in #1 above exist but there lacks a capable workforce, this would breed slow or little progress.

Row 5: If all the tenets in #1 above exist but there lacks a capable processes, this would lead to adoption of solutions that already exist or reinvention of the wheel.

Row 6: If all the tenets in #1 above exist but there lacks a good organizational culture that supports positive change, this would create many barriers for the proposed change.

Row 7: If all the tenets in #1 above exist but there lacks the right global incentives within the organisation, this would cause sporadic change patterns.

Row 8: If all the tenets in #1 above exist but there lacks the right action plan within the organisation, this would cause many false starts.

This table emphasizes the fact that change management is a complex issue that must be carefully taken into consideration. In eGovernment, change management is a critical success factor since eGovernment services have a wide implementation span carried out by both public and public-private partnered teams in various sectors (Nograšek, 2011). Hence care must be taken to provide all relevant requirements for effective realization of required change.

Table 1.2: Managing Change during Planning, Design and implementation of PKI

#	Vision	Resources	Capable Workforce	Capable Processes	Organisational Culture	Incentives	Action Plan	Outcomes
1	√	√	√	√	√	√	√	Required change
2	X	√	√	√	√	√	√	Confusion
3	√	X	√	√	√	√	√	Anxiety & frustration
4	√	√	X	√	√	√	√	Slow or little progress
5	√	√	√	X	√	√	√	Reinventing the wheel
6	√	√	√	√	X	√	√	Barriers to change
7	√	√	√	√	√	X	√	Sporadic change
8	√	√	√	√	√	√	X	False starts

Source: Managing Complex Change; AmbroseDelarose, 1987

1.4 Research Objectives

Main Objective:

To develop a PKI security rational decision optimisation tool based on quantitative assessments of security quality attributes.

Supporting objectives:

a) To identify key PKI security quality attributes that need to be measured and optimised in order to have a secure public key infrastructure solution for eGovernment;

b) To utilize the attributes identified in (a) in proposing a PKI security rational decision optimisation conceptual framework;

c) Develop a PKI security rational decision optimisation tool from the CF in (b) and

d) Evaluate the tools' effectiveness.

1.5 Research Questions

Main Question:

What PKI security managerial decision optimisation tool can be developed based on quantitative assessment of security quality attributes?

Supporting Questions:

i. What security quality attributes define a good public key infrastructure solution for eGovernment and how can they be quantitatively assessed?

ii. What conceptual framework best captures and proposes a quantitative decision optimisation model based on the attributes identified in (a) above.

iii. How can the model in (ii) be converted into a PKI security rational decision optimisation tool and how can its effectiveness be evaluated?

iv. What is the effectiveness of the tool?

1.6 Research Map

A broad view of identity management is covered in Chapter 1. Chapter 2 is the literature review which culminates in a conceptual framework. Chapter 3 looks at the methodology. Finally, Chapter 4 and 5 present the results and conclusions respectively.

The main objective of the research was achieved by the entire study since all the efforts and activities were aimed at delivering the PKI security rational decision optimisation tool. Supporting objectives (a) was achieved in Chapter 2 of the research specifically in Table 2.6 which identifies PKI security quality properties and their influencing variables. Supporting objective (b) was achieved in Chapter 2 and specifically as captured in the conceptual framework Figures 2.23. The main research question and supporting questions (i) and (ii) guide this process of identifying PKI security quality factors from literature and modeling them in a conceptual framework (CF). Supporting objective (c) is achieved in Chapter 3 in which the the conceptual framework is translated into a PLS-SEM model then tool to be used for decision support. Its effectiveness as posed by question (iii) and (iv) is evaluated by collecting data and populating the model, then analysing to generate various statistical measures in Chapter 4. Lastly, Chapter 5 gives the conclusions and recommendations.

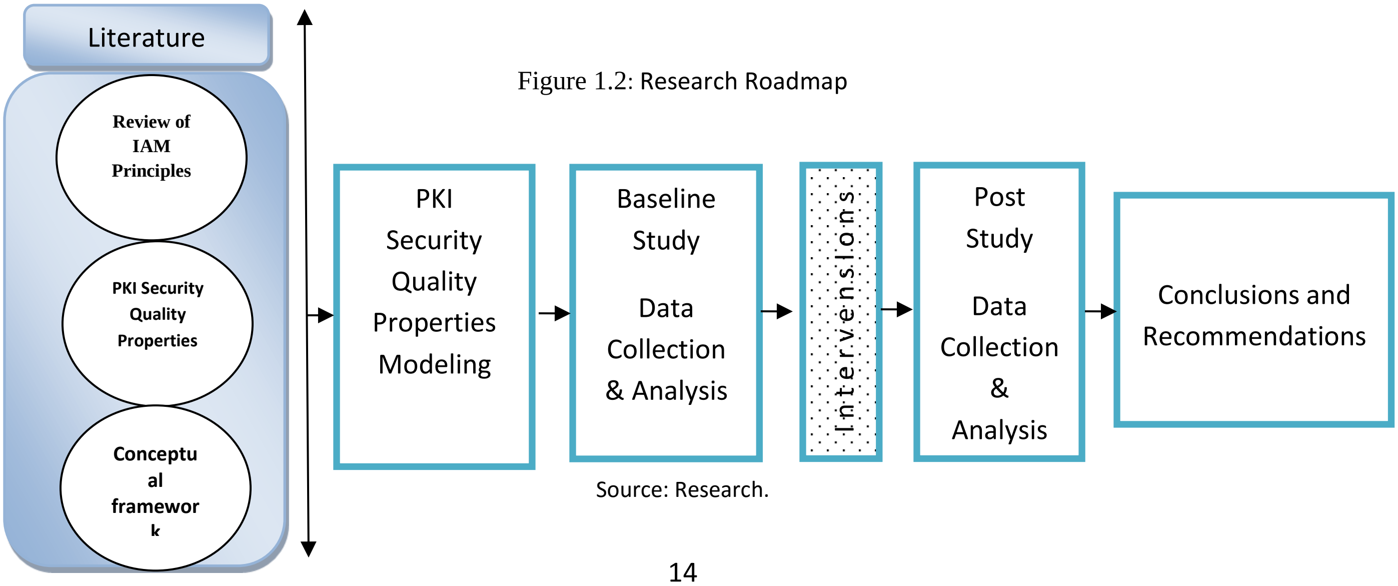


Figure 1.2: Research Roadmap

Literature
Review of IAM Principles
PKI Security Quality Properties
Conceptu al framewor k
PKI Security Quality Properties Modeling
Baseline Study
Data Collection & Analysis
Intervensions
Post Study
Data Collection & Analysis
Conclusions and Recommendations

Source: Research.

1.7 Significance of the Study

More often than not information security managers find themselves with the difficult task of making the best decisions regarding how best to enhance the security of information systems. To solve this problem, organisations either engage external decision support consultants or implement decision support systems (DSS) (Turpin and Marais, 2004). In both options, the underlying fact is that data and or information about the target area has to be collected, analysed and presented to the decision maker in a form and or format that is not only meaningful but would quickly help in the cognitive process of making the optimum decisions (Ghani et al., 2009). Optimisation of PKI security rational decisions takes a similar approach. Data is usually collected mainly using checklists and presented to decision makers in a Yes/No format indicating responses to compliance queries(ISC, 2003 ; Goulet, 2009). A spot check by the research team during such an audit at Kenya's first Certificate Authority (Kenya ICT Authority) during a security audit revealed that a single checklist could generate upto 50 responses covering one area e.g. firewall configuration. We contend that such large volume of decision support data is cumbersome to work with if not presented in quantitative or pictorial format which can quickly help a decision maker to grasp difficult concepts, follow plotted trends etc.

This research proposes a new way of visualising such data for PKI security rational decision makers when reasoning about the security of PKI solutions in eGovernment. The decision support tool that the research proposes makes the following significant contributions:

1. Instead of presenting the decision support data as a long list of Yes/No responses, the research demonstrates how to convert such responses into quantitative data presented

to the decision maker as statistical measures e.g. t-values, p-values, composite reliability etc. (Wong, 2013).

2. The statiscal data can be presented using pictorial or graphical format which would help decisiom makers to see the analytics presented visually, hence enabling them to grasp difficult concepts a bit more easily, see new patterns and or track past trends. This in essence is one goal of a good decision support system (United Nations, 2009).

The above two significant contributions gives the research its motivations. It is also worthwhile to note that most avatar countries like Australia (Australian Government, 2009b), United States of America (Smedinghoff, 2010) and the European Union (FIDIS WP3, 2005) which have succeeded to a good measure in eGovernment, fall outside the African continent. Developing countries like Kenya though at an advantage of adopting such technologies that have already succeeded out there cannot do so directly because of context specific issues. For example, PKI solutions are tightly coupled to the contexts within which they are implementes e.g. the legal, regulatory and institutional frameworks of a country greatly impact PKI solutions (Dempsey, 2004). This research is therefore significant in that it proposes one way of enhancing the security of PKI solutions through PKI security rational decision optimisation and is grounded in a developing economy scenario, i.e. Kenya.

2. LITERATURE REVIEW

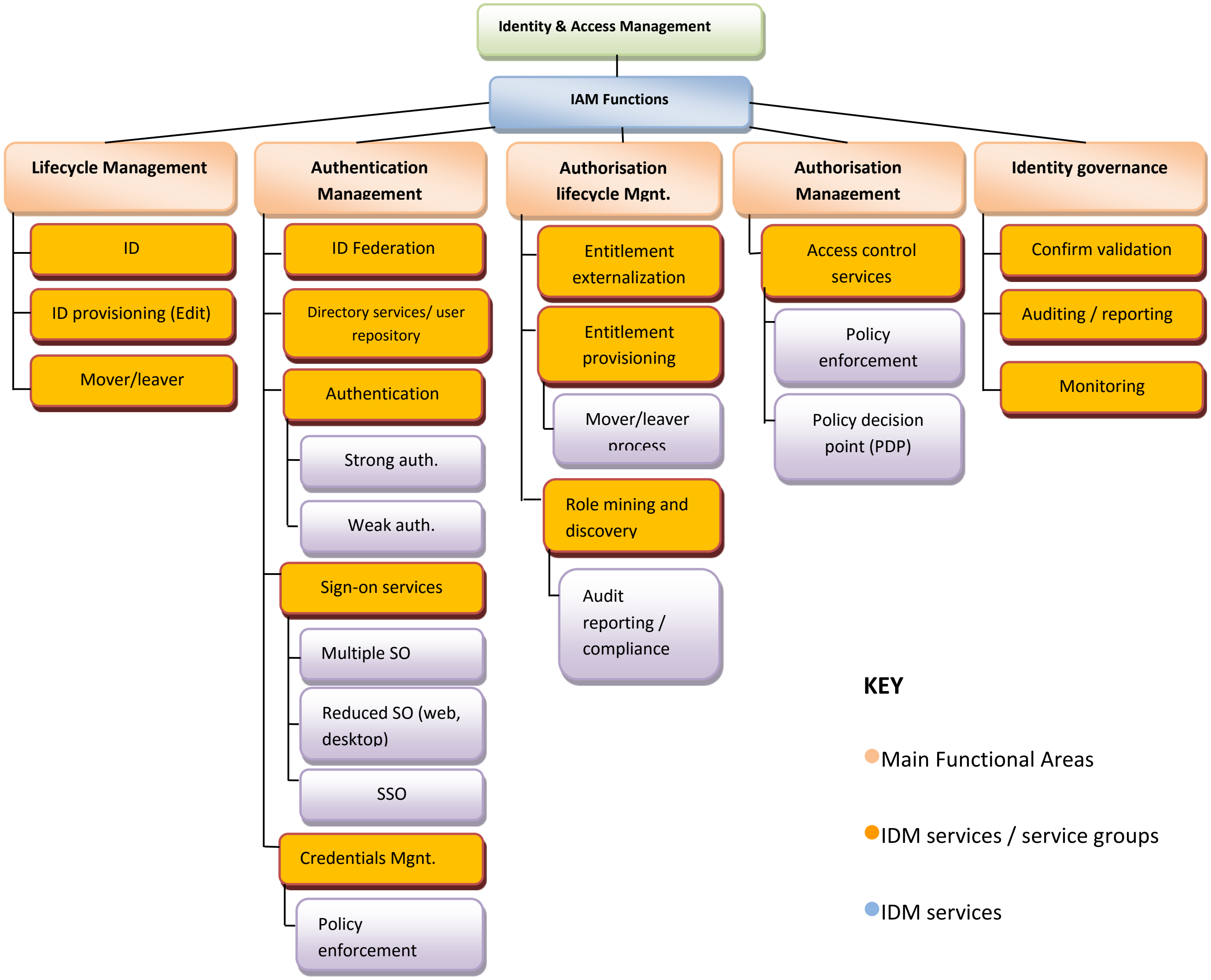
The simpler the identity
the easier the interaction

~Unknown

2.1 eGovernment Identity and Access Management Framework

A good Identity and Access Management (IAM) program for e-Government that is PKI enabled requires an overarching strategy and governance model (SMEDINGHOFF, 2010; Ernst & Young, 2013) as is depicted in Figure 2.1.

Figure 2.1: Sample eGovernment IAM Framework



Source: Open Data Center Alliance, 2012

IAM can be divided into lifecycle, authentication and authorization management:

1. Lifecycle management: deals with creating identities, providing credentials, editing changes, replacing lost ones and deleting those that leave the system.
2. Authentication management: determines the true identity of an entity.
3. Authorization management: binds rights/privileges to an identity; define access levels.
4. Identity Governance: manages risks through administration, monitoring accesses to resources, auditing and reporting.

Although not directly inferable from the framework above, it is the position of this study that in order to enhance PKI security in eGovernment, there must be put in place proper security controls governing the entire lifecycle of a digital identity from the time it is created to the time it is retired or deleted. To achieve this, optimised decisions must be taken when reasoning about the PKI security throughout the lifecycle and management of electronic identities in order to achieve enhanced security.

2.2 PKI and eGovernment Identity and Access Management

The Oxford University Dictionary defines identity as the "fact of being who or what a person is or thing is"(Angus, 2010). This implies that it is always desirable to verify or know what a person or thing is or what they claim to be at all times before engaging or transacting with them. The Collins Dictionary defines identity as "the state of having unique identifying characteristics held by no other person or thing" and also "the individual characteristics by which a person or thing is recognized"(Collins, 2014). In eGovernment, entities transact with one another using digital identities in cyberspace i.e. virtual or electronic representations their owners. This underlines the great need for proper identification and authentication of the digital identities in order to create a trustworthy transacting environment devoid of impersonation and which supports non-repudiation. To achieve these, most eGovernments e.g. Australia, USA, United Arab Emirates and Kenya(Australian Government, 2009 ; Smedinghoff, 2010 ; Al-Khouri, 2011a) prefer implementing the Public Key Infrastructure solutions as a means of enforcing cyber security. PKI achieves this by registering key identity details of each citizen or organisation and then issuing them with unique digital certificates and keys based on public asymmetric encryption schemes (usually based on Rivest-Shamir-Adleman (RSA) or Elliptic curve schemes) (RSA, 1999). By setting up trusted certification authorities which securely provide end users with both public and private keys, and having secured certificate status assurance and lookup mechanisms for third parties, PKI enables cyber security in complex electronic transaction environments (Al-Khouri, 2011a). From this point view, it becomes evident that the security of the PKI system itself comes into sharp focus. If the PKI system is compromised, then the entire security of eGovernment will be compromised. Therefore there is great need to make sure that PKI security is properly enhanced at all the times to protect against known and emergent threats.

2.2.1 Public Key Infrastructure (PKI)

PKI can be defined as the a framework of legally recognized services that consists of hardware, software, policies, and procedures for managing entity identity keys and certificates for trustworthy identification and authentication in cyber-space (Choudhury et al., 2002). It is the desire of many a government to have a twenty four hour (24 hr) self service e-Government model based on key quality factors in a secure electronic communication environment like that offered by PKI (Ijaz, 2012; Al-Khouri, 2011b). The key drivers for adoption of PKI include:

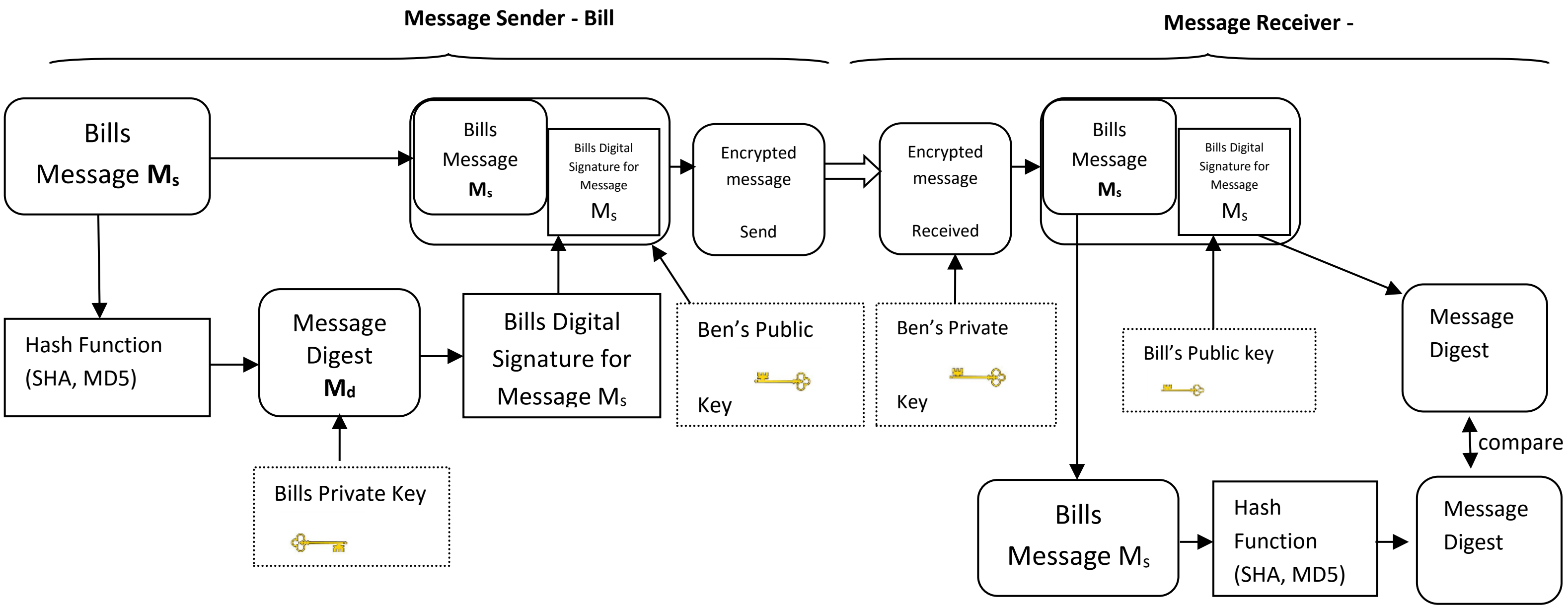
- Non-repudiation: a transacting party cannot at a later date disown information it sent. PKI accomplishes this through *digital signatures* which peg a message to its owner.

- Integrity: data should not be modified or altered in transit or while in storage. This is accomplished through message *hashing* – a process that produces a message digest based on the entire message, which is usually a value from a single value function. If the message is changed, the digest changes hence modification can easily be detected.
- Confidentiality: messages should be protected against unauthorised access during transmission. PKI achieves this through *encryption.*
- Authentication: This is a means of identifying genuine users. PKI achieves this through *digital certificates* issued by a certification authority.
- Access Control: only people with the required security priviledges have access to information. PKI achieves this through digital certificates at the authentication point.

PKI is not just the information technology infrastructure; it also consists of policies, laws, software, people, processes and standards that govern secure electronic communications based on public key cryptography scheme (Mjølsnes et al., 2008). This study adopts Smedinghoffs definition which sees PKI as the wide-ranging information technology infrastructures together with people, laws, policies, procedures and standards that are deployed to offer high levels of information security during online transactions, non-repudiation and protection of important communications (Smedinghoff, 2010).

PKI is grounded in cryptography but solves identified problems inherent in symmetric key and asymmetric key (public key) cryptography. Although symmetric key cryptography is fast performance wise because it uses only one key (Thakur and Kumar, 2011), there exists a key passing problem between communicating parties i.e. the shared private key has to be securely passed between the two communicating entities (ibid). On the other hand, although public key cryptography solves the key passing problem by having two keys (private and public), it is slow and there is the difficulty of ascertaining whether a public key published in the directory actually belongs to the claimant (Choudhury et al., 2002), which can lead to impersonation attacks. A digital certificate given by PKI CA's solves the impersonation problem by introducing a trusted third party to issue and verify keys between transacting entities (Chen et al., 2014). Figure 2.2 depicts how digital signatures are used to encrypt messages, create trust and enable checking of message integrity between communicating entities.

Figure 2.2: Generating a Digital Signature



Source: Public Key Infrastructure Implementation and Design; Choudhury et al., 2002

There are many types of digital signatures but they can all be classified as either direct or arbitrated (Choudhury et al., 2002). In the direct approach, a person encrypts the message and the digest with his/her private key and sends it to the recipient. The problem with this one is that non-repudiation is not assured since the sender can later claim to have lost his/her private key. In the arbitrated approach, a signed message by the sender is first sent to a trusted arbitrator who checks that the signature actually belongs to the sender before forwarding the message to the receiver. This later approach ensures non-repudiation since the arbitrator is usually the certification authority who knows which certificate is valid, compromised or expired (Ibid). A good national eGovernment PKI should be built around the arbitrator in which case the service providers communicate with the CA's to verify the credentials of all those seeking services.

The main components of PKI are summarized in Table 2.1 below:

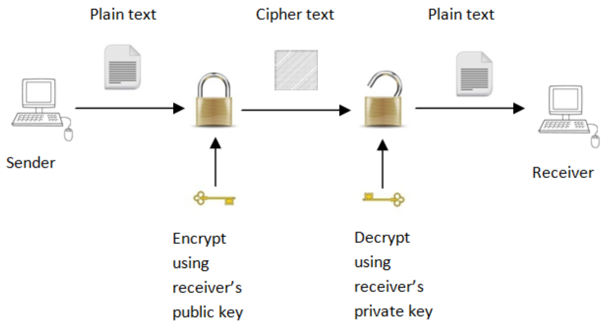
Table 2.1: Components of PKI

Component	**Description**
Digital Certificates and Signatures	• A digital certificate is a signed electronic identity card issued by a certification authority having trusted identity credentials (public and private cryptographic keys) of its holder used to authenticate him/her to third parties. • A digital signature is a one way protocol or algorithm that mathematically computes a single value (digest) from an entire message used by the sender to sign the message and by the recipient to test authenticity of the message received.
Certification Authorities (CA)	• Issues and revokes digital certificates
Registration Authorities (RA)	• Usually hived off from CA to validate certificate requests from end entities.
Certificate Repositories	• Store certificates for authentication purposes. • Provide Certificate Revocation Lists (CLR's)
Archives	• A long time historical information storage repository
Certification Policies and Practice Specifications	• Specifications that outline how the CA and its certificates are to be utilized; level of trust; indemnity or liability issues in case of broken trust.
End Entities	• These are end users who get issued with certificates.

Source: Public Key Infrastructure Implementation and Design; Choudhury et al., 2002

In a public key communication system, every user generates a public cryptographic key and a private one. When sending a message, the user encrypts it with the recipient's public key which is known to everybody. However that message can only be decrypted by the recipient's private key as depicted in Figure 2.3 below. Depending on the level of trust required by the PKI, a user either presents his or her identity proof online to a RA or appears there in person with relevant identification credentials. Once positively identified, the CA issues Digital Certificates which contain a user's public key and identity and continues to manage them henceforth. The user can then generate a private key known only to him or her. The digital certificates are used to verify the digital signatures between transacting parties hence ensuring data integrity and proper authentication. In another arrangement, the two keys can also be generated by a smartcard or software on a user's computer. Security is assured by the fact that the private key is only known by its owner. A well planned PKI also ensures confidentiality through encryption and gives assurance that a particular digital signature belongs to a given person (non-repudiation).

Figure 2.3: Public Key Cryptography


Plain text
Cipher text
Plain text
Sender
Receiver
Encrypt using receiver's public key
Decrypt using receiver's private key

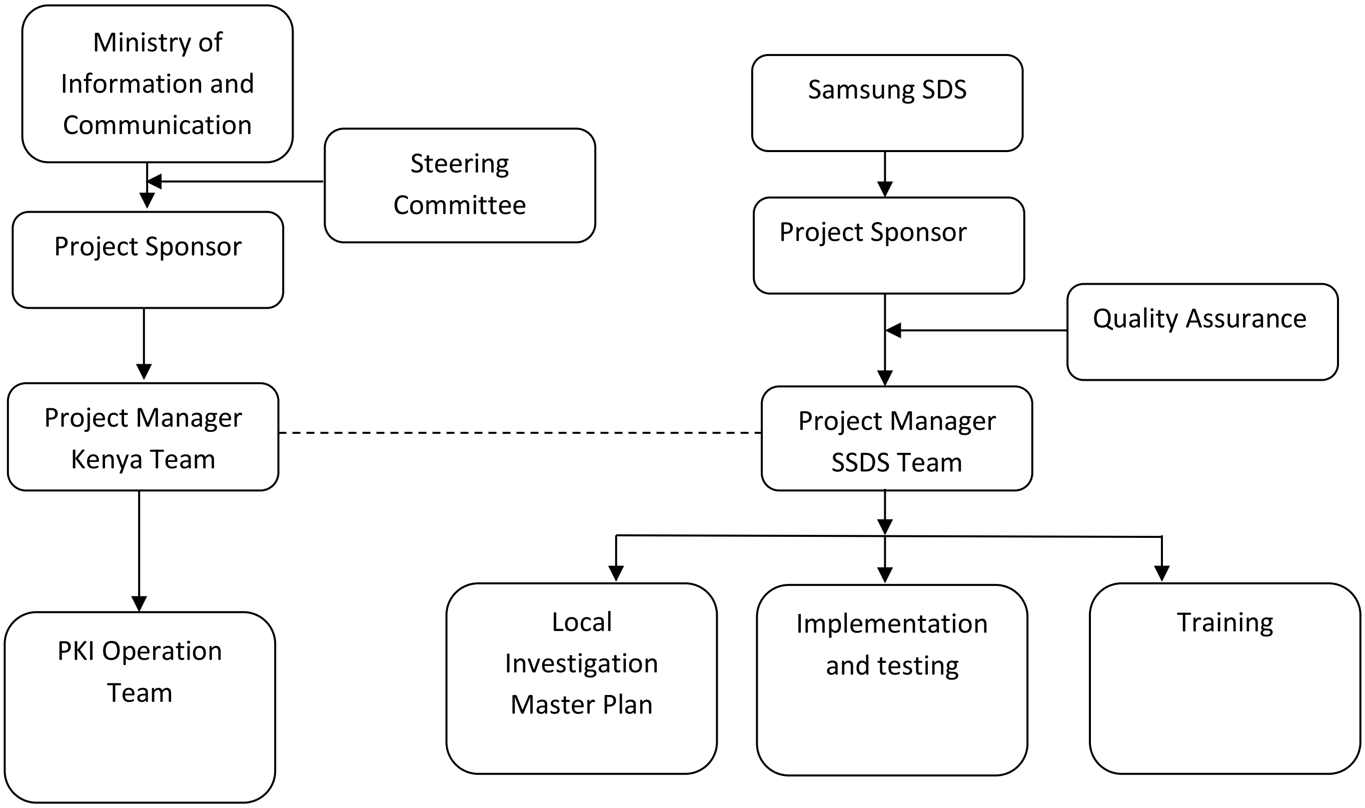
Source: Research; Public Key Cryptography.

2.2.2 PKI in Kenya

The introduction of the public key infrastructure (PKI) scheme in Kenya championed by the Communications Authority of Kenya is a good step to enable full electronic transactions between the state, citizens and other legal entities. The company that won the tender to roll out the

National PKI solution is the same that did it for South Korea (Samsung SDS). The project falls under the Kenya Transparency & Communications Infrastructure Project (KTCIP) sponsored by the World Bank and is aligned to achieve Kenya's Vision 2030 ICT pillar. The project teams were organised as shown in Figure 2.4. Adopting PKI for e-Government creates both opportunities and challenges.

Figure 2.4: National PKI Project Team

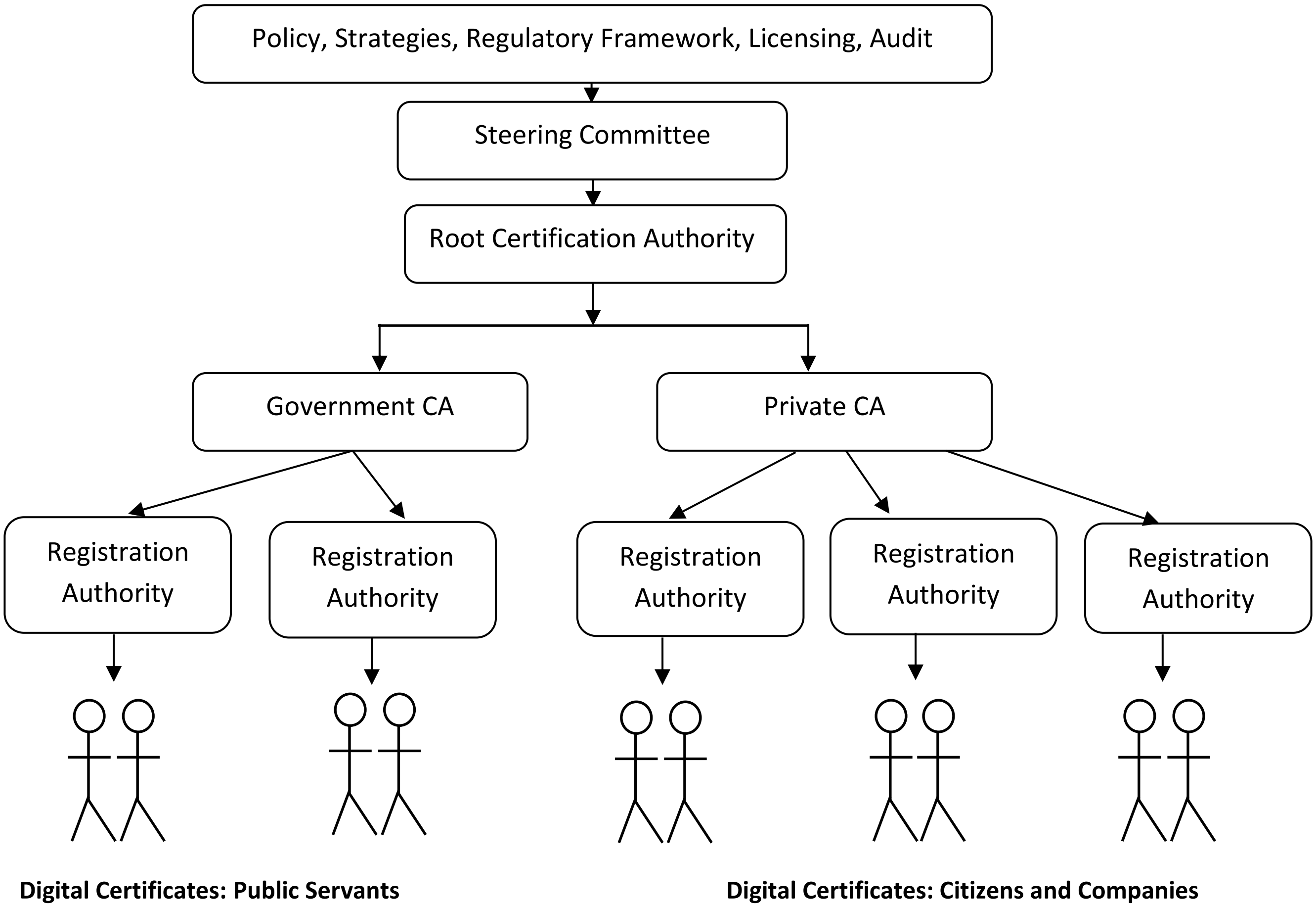

Ministry of Information and Communication
Steering Committee
Project Sponsor
Project Manager Kenya Team
PKI Operation Team
Samsung SDS
Project Sponsor
Quality Assurance
Project Manager SSDS Team
Local Investigation Master Plan
Implementation and testing
Training

Source: Research; Ministry of Information and Communication, Kenya; 2014

The opportunities include more efficient remote service delivery to the citizens, more transparent and accountable processes in public administration, secure transactions, and nonrepudiation. On the other hand, lack of appropriate human capacity, high cost of technology, low literacy levels among others are some of the challenges.

The current situation in Kenya is that since the digital citizen database and PKI are not live yet, many public and private agencies have their own identity management systems. Agencies are either at the presence level (mostly with static websites) with some elements of interaction level e.g. simple database search ability at Kenya National Examinations Council (KNEC). It is envisaged that once the government develops the fundamental infrastructure to capture and manage the digital identities of citizens, these organisations will rely on e-Government PKI enabled authentication to roll out real time secure transactions. This study envisions a PKI scheme integrated together with smartcard based identity credentials. The structure of the National PKI is as shown in Figure 2.5.

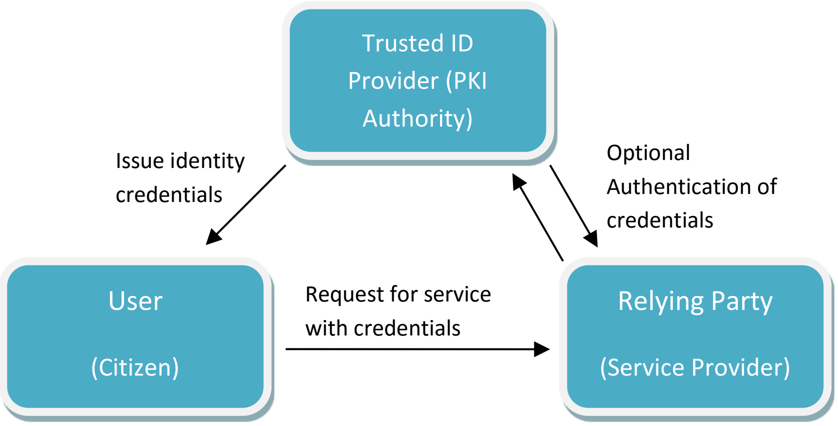
Figure 2.5: National PKI Structure


Policy, Strategies, Regulatory Framework, Licensing, Audit
Steering Committee
Root Certification Authority
Government CA
Private CA
Registration Authority
Registration Authority
Registration Authority
Registration Authority
Registration Authority
Digital Certificates: Public Servants
Digital Certificates: Citizens and Companies

Source: Research; Ministry of Information and Communication, Kenya; 2014

User (U) requests a service from a relying party (RP). ID provider verifies user identity.

Figure 2.6: Building Trust in Identity Claims


Trusted ID Provider (PKI Authority)
Issue identity credentials
Optional Authentication of credentials
User
(Citizen)
Request for service with credentials
Relying Party
(Service Provider)

Source: Public Key Infrastructure Implementation and Design; Choudhury et al., 2002

2.3 PKI Security Decision Optimisation Theories and Frameworks

Designing optimised systems that have good multi-quality properties and tradeoffs is not a small task (Koziolek and Reussner, 2011). The objective of any optimisation task is to maximise the quality of solutions under given constraints. In mathematics and science, measuring the length of a physical phenomena, or the mass of a tangible object can be classified among trivial problems. However, measuring other properties like the percentile composition of elements in a sample requires considerable effort. Information systems like PKIs also have those properties that can be directly measured and those that are fairly difficult to measure or approximated.

Since it is impossible to measure the quality properties of information systems like security, interoperability etc. directly due to their non-functional nature (Johnson et al., 2014), this research proposes identification of relevant quality attributes for each quality property under study, then quantitatively assessing the attributes to estimate the level of their attainment in the solution under investigation. The measures can then inform rational decisions on how best to allocate resources or act in order to affect their improvement. In this study we concentrate on eGovernments that have adoppted PKI technology as a means of enhancing their security through strong authentication mechanisms (Al-Khouri, 2011a).

The security threats to the information assests of eGovernments and especially their PKI enabled identity and access management systems are real. Many eGovernments like Israel, Estonia, USA etc. (Vaidya, 2015) have reported massive threats and exploits on their information infrastructures either by individuals, organised hacking groups or by foregn governments. A spot check on the global state of hacking activity on a website such as www.norse.com reveals that hadly does a moment go by without hacking activities taking place worldwide. This underlines the need for enhancing PKI security at all times in order to prevent exploits directed against them from succeeding.

One way of trying to understand and improve the quality of information systems is to create models of their quality attributes and their relationships (Narman et al., 2007). If the quality property is say security, the model made up of security quality attributes would enable analysis and measurement of such non-functional quality property hence providing important information to planners, implementers, managers etc. with the aim of optimising their decision making (Johnson et al., 2007; Johnson and Ekstedt, 2007; Koziolek and Reussner, 2011). In this section, we not only look at PKI security attributes but also study several decision optimisation theories that can be used to optimise rational decisions that affect their improvement then justify why we finally settled on Partial Lease Squares Structural Equation Modeling (PLS-SEM).

Partinently, holistic enhancement of a quality property like PKI security involves representing it using its partinent measurable quality attributes whose evaluation consists the objective function(s) / fitness function (Aleti et al., 2013).

2.3.1 Some PKI Security attributes

A careful review of ISC (2003), Australian Government (2009a), Federal Public Key Infrastructure Authority (2015) and discussions with PKI security experts revealed a consistent set of PKI security attributes whose controls need to be tested and reported during relevant audits. A summary of some of the attributes and their controls as identified are captured in Table 2.2.

Table 2.2: PKI security attribute controls

No	Main Attribute	Control Statement
1	Certificate Policy (CP) and Certificate Practice Statement (CPS)	CPS must conform to / derived from its corresponding CP
2	Initial Identification of Entity before issuance of Digital Certificate	Certificate Authority (CA) / Registration Authority (RA) must positively identify a person or entity based on legal requirements before issuance of digital certificate
3	Certificate Revocation	There must be a secure procedure for revoking certificates after key compromise i.e. the process must be triggered by reception of an authenticated request.
4	Cryptographic Modules	All Certificate Authority equipment including cryptographic modules shall be protected from unauthorized access at all times.
5.	Physical Access Control	Pertaining to Certificate Authorities: ensure no unauthorised physical access to hardware is permitted. Also all removable media and paper containing

		sensitive information must be stored in secured containers.
6.	Backups	Full system backups sufficient to recover from system failure shall be made on a periodic schedule. Backups are to be performed and stored off-site not less than once per week.
7	Role Segragation	Two or more trusted persons must participate during sensitive PKI operations e.g. CA key generation, CA signing key activation, CA private keys backup etc.
8	Background Investigations	CA personnel must pass the following minimum background checks: Education, Employment, Residence, Good Conduct and References.
9	Training	All CA personnel must receive continuous training in all operational areas, and critical areas like disaster recovery and continuity planning.
10	Contract Staff	They must meet the personnel requirements set in the certificate policy.
11	Audit Logs	There shall be audit logs generated for all auditable events touching on the security of the CA. Automation of the collection mechanism is preferable.
12	CA Key Lifecycle Management	Secure lifecycle management of CA and client keys
13	Professional Ethics	There should be in place a professional code of ethics for all personell derived from the CP
14	Legal/Regulatory	There must be in place a lagal, regulatory and institutional framework that properly adminiters issues of risks, indemnity etc.
15	Disaster Recovery Planning	There should be in place proper continuity planning – plans, budgets and readiness.
16	How are Entity Private Keys Generated	If CA generates entity private keys then there must be a secure process of delivering them to the entity. If the entity generates own keys, then the question is how entity has to deliver the public key to the CA for issuance of relevant certificate.

17	Key Size	Must follow NIST SP 800 131A guidelines.
18	Certificates Version	X.509 certificates must be used for current PKI
20	Type of Cryptographic Algorithm	Is it RSA, Digital Signature Algorithm, Elliptic Curve Digital Signature etc. Elliptic curve gives better security.
21	Audit Log Protection	Audit log should be protected against modification. Strict personell controls on who can access or backup the logs must be in place.
22	Backup Policy	Data should be backed up securely in off-site locations. Backed data should be encrypted.
23	Personnel Controls	There must be defined trusted roles for personnel handling sensitive information PKI assets. Sensitive tasks must have dual or multile personnel controls
24	Private Key Recovery	There should exist a secure procedure for recovering escrowed private keys e.g. trusted chain of custody, split knowledge control, dual control etc. The procedures for selecting an escrow agent need to be secure and tested.
25	Time Stamping Services	There needs to exist a secure time stamping server to service PKI enabled transactions e.g. the time stamp provides strong evidence that a particular digital record or transaction existed at a particular time before expirtion of the signing certificate etc.
26	Online Revocation/Status Checking	There should exist real-time information about the status of a certificate in the certificate revocation list (CRL). The best is to have On-line Certificate Status Protocol with stapling for best optimisation.
27	Certificate Revocation Procedures	The CA should have a secure procedure for gathering certificate revocation requests e.g. upon receipt of a properly signed email request.
28	Routine Certificate Renewal	The CP or the CPS must have secure rules and procedures on how to renew or replace a certificate that is expiring.

Source: (ISC, 2003); (Australian Government, 2009a); (Federal Public Key Infrastructure Authority, 2015)

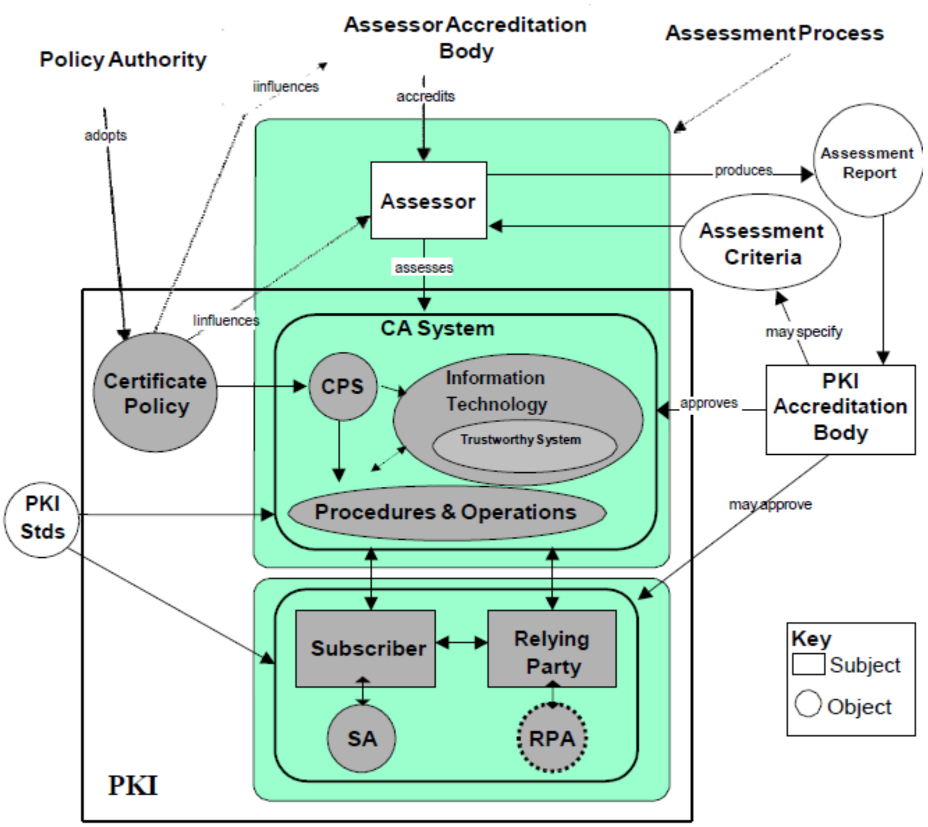
These security attributes in Table 2.2 plus many others which are not listed here are very important when reasoning about the security of a PKI system. If they are measured and managers helped to make the right decisions about their improvement, the security of the PKI will be enhanced.

2.3.2 Decision Optimisation Theories

eGovernment PKI security rational decision optimisation usually relies on actionable information provided to security managers generated from independent audits to find out whether the PKI was implemented and operated according to the requirements in its certificate authorities certificate policy (CP) and Certificate Practice Statement (CPS) (Federal Public Key Infrastructure Authority, 2015). Due to the high cost of such audits and the need to plan and budget for them, it is not strange to find that most eGovernment agencies prefer a two tier kind of audit cycle: one, a full compliance audit at the beginning before commencement of operations then repeated every three years; and two, yearly audits with each proceeding year testing the non-compliance aspects of the previous year to see whether they were corrected (Ibid).

PKI assessment audit can either be carried out by internal staff (internal audit) or by independent assessors (independent audit). Though the internal audit is good for self assessment and improvement, it is the independent audit which carries a lot of weight since it is assumed the external assessors did not participate in the setting up of the security controls hence are likely to be more objective (NIST, 2014). The PKI security independent assessment process involves several stakeholders including the certificate authority (CA) under investigation, professional and independent assessors who are properly accredited by a reputable accreditation body, policy authority and lastly the PKI accreditation body (ISC, 2003). Once selected, the assessors spearhead the process of using the assessment criteria provided by the PKI accreditation body or creating one by involving all stakeholders if it does not exist. This assessment criteria can then be used to assess and report on the state of the PKIs security controls. The data collected is organised and presented to management for decision making. The report can also be presented to the PKI accreditation authority for them to know whether the certificate authority is complying to their certificate policy, certificate practice statement, legal and regulatory requirements as set out in the policy. This PKI security assessment and enhancement model can be represented diagramatically as shown in Figure 2.7.

Figure 2.7: PKI security assessment model



Source: ISC (2003)

A good security audit usually takes a risk based approach to guide the assessor in performing either compliance testing or substantive testing of the security controls. Compliance testing finds out whether an organisation's compliance with security control procedures is okey. Substantive testing usually gathers relevant evidence to evaluate the integrity of data and transactions on the information system (ISACA, 2015). This study concentrated more on compliance testing.

A number of other information systems improvement or enhancement methodologies exist. Examples include the Information Technology Balanced Score Card (IT BSC) (Kaplan, 2012), Six Sigma (Anand et al., 2012), and Architecture Optimisation approaches like the Zachman Model(Zachman, 2008) and its related derivative like Enterprise Architecture Analysis (Johnson et al., 2007).

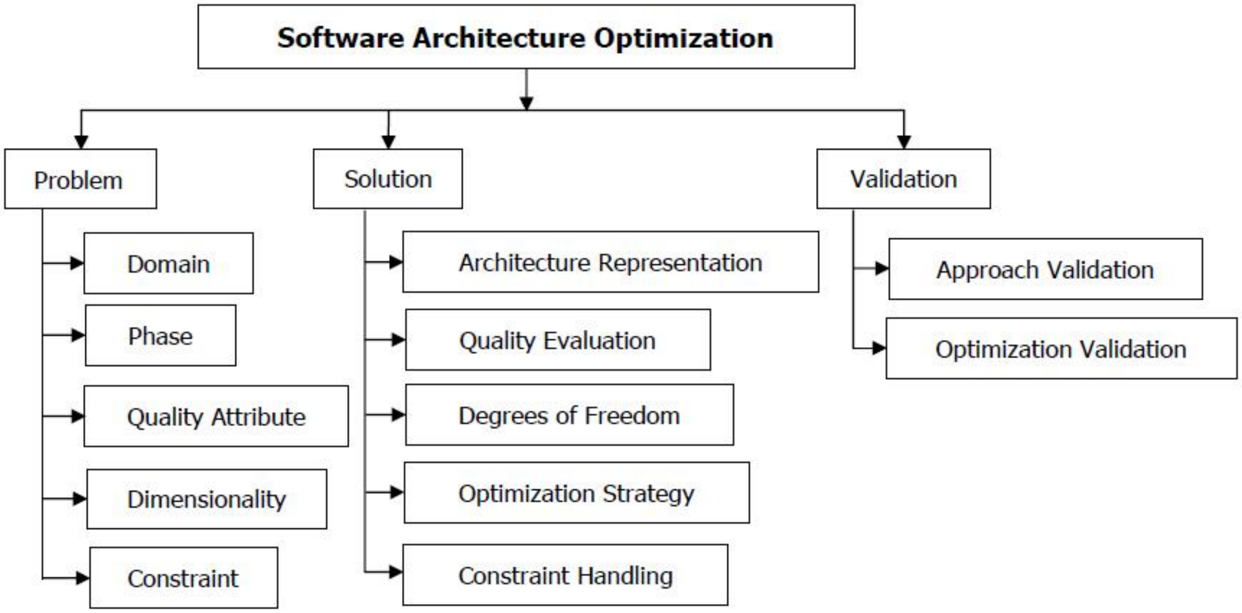
The balanced score card method is a process centered evaluation technique which does not only looks at the financial aspects but also evaluates things like customer satisfaction and the innovative abilityof the entities under investigation. Its output is usually metrics in the form of key performance indicators (KPIs) which can be used to optimise business related information technology decisions (ISACA, 2015 ; Kaplan, 2012). A key performance indicator (KPI) can be defined as a measure of how well a particular process is being executed as a means of attaining the set goal i.e. by using KPIs it is possible to know whether a particular goal will be attained or not through assessment of performance drivers e.g. skills, business capabilities, resource availability etc. It is the view of this research that this method therefore cannot be used to measure the level of attainment of key quality attributes in an information system but rather to inform the strategic managers on whether the set goals are achievable or not based on the performance drivers identified.

The six sigma methodology is a process driven approach which seeks to improve quality through the progressive identification and elimination of defects and minimisation of variation in process output (Anand et al., 2012 ; MIT, 2012). Using the two sigma tools i.e. DMAIC (Define, Measure, Analyse, Improve, Control) and the DMADV (Design, Measure, Analyse, Design, Verify) and exisitng process can be improved or a new one introduced respectively. Sigma six uses statistical measures to compute defects per million opportunities i.e. how many times does a process fail measured in terms of process output failing to meet customer specifications given one million instances. Process improvement is measured as reducing defects per million opportunities (DPMO) (Ibid). Although sigma six is good, many industry players refuse to adopt it and where it has been adopted failure rates are high, attributed to lack of an iplementation model that can effectively guide rollout of the program in organisations (Fursule et al., 2012).

In the last few decades, there has been great interest in adopting the engineering approach to information systems architecture design, with great emphasis being placed on the need to treat the architecture as a blueprint that embodies the true qualities of final solutions (Johnson and Ekstedt, 2007). Such architectures should be able to provide a foundation on which quality properties can be assessed, measured and predicted as accurately as possible to the real world solution, hence helping to minimise post implementation costs (Koziolek and Reussner, 2011). One reason why architecture models are appealing is because they can be used to structure complex software systems and provide a blueprint on which future references or improvements can be made (Aleti et al., 2013). The decisions taken during architecture design have future economic and quality implications; things like selection of hardware and software, their mapping to each other, system topology, quality measurements among others (ibid). When looking at eGovernment solutions, it is not enough to look at hardware and software only because citizen service oriented solutions need to be holistic. Success of solutions in the real world is not only pegged on excellent information systems but also other human aspects like culture, literacy levels, process efficiencies among others. For example a culture of corruption is likely to spawn high cost procurement procedures that deliver low quality solutions. Therefore, when selecting tools to model the quality of information systems, it is important to have a framework like the Culture, People, Process and Technology (CPPT) forming a criteria on which a modeling tool can be selected so as to capture all relevant factors that can influence achievement of quality in the final solutions.

Aleti et al(2013) provides a taxonomy for classifying optimisation techniques for software architectures after researching 188 papers. She proposes that an optimisation technique can be categorised according to the type of problem it can solve (how the problem is formulated), the techniques used to solve the problem and how the technique can be validated for a particular optimisation problem, see Figure 2.8. She also notes that before embarking on any optimisation task, it is good to know which domain the technique was created for i.e. embedded systems, information systems or general. She concludes that exact optimisation techniques are only possible where the search space is small and exhaustive search for the optimum solution can be done. However, most problems have a large search space hence approximate (metaheuristic) optimisation techniques are the most feasible. Harman et al. (2012) pitches for optimisation problems to be generaly called Search Based Software Engineering (SBSE) problems. He mentions optimisation techniques such as local search, simulated annealing, and genetic algorithms. Although the paper is comprehensive and tries to classify various optimisation strategies for various needs like requirements, architecture, design, development, testing among others it does not give specific guidance on how to use specific models or tools in a situation but just mentions them.

Figure 2.8: A Taxonomy of Software Architecture Optimisation Techniques

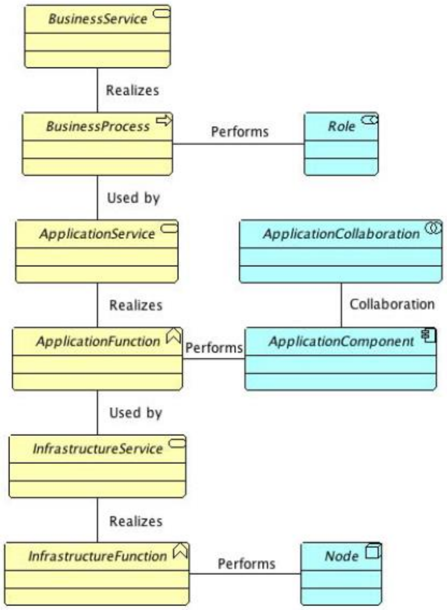


PROBLEM: Domain: Embedded, IS or General; **Phase:** Place in development process (design time/run-time); **Quality Attribute:** System attributes to be measured; **Dimensionality:** single/multiple objective optimisation. **SOLUTION: Architecture Representation:** process input describing architecture to optimise; **Quality Evaluation:** Used evaluation procedures; **Degrees of freedom:** how many alternative architecture configurations are available in the search space; **Optimisation strategy:** exact/approximate solutions? **VALIDATION: Approach validation:** practicality and accuracy of approach; **Optimisation validation:** how well a solution approximates global optimum / relative performance

Source: Software Architecture Optimisation Methods: A Systematic Literature review; Aleti et al., 2007

Enterprise Architecture Analysis (EAA) is yet another initiative that can be used to optimise information systems architectures through rational decision support. Examples of EAA initiatives include The Zachman Framework (Zachman, 2008), Open Group Architecture Framework (TOGAF), Enterprise Architecture Planning among others. (Johnson and Ekstedt, 2007). Some tools that implement EAA modeling include but are not limited to Archimate which was ratified by The Open Group in 2012, Metis, System Architect, Qualiware, and EAA Tool (EAAT) (ibid). These tools use methods such as tradeoff analysis, Monte Carlo techniques among others for optimising architectures. EAA is a holistic approach to modeling enterprise information systems architecture and it uses UML notation to represent business services, business processes, application services, infrastructure services, components and their collaborations, roles among others, and quality their attributes in a causal model (Figure 2.9 shows the metamodel).

Figure 2.9: A EAA Metamodel - ArchiMate



Source: IT Management with Enterprise Architecture; Johnson et al., 2014

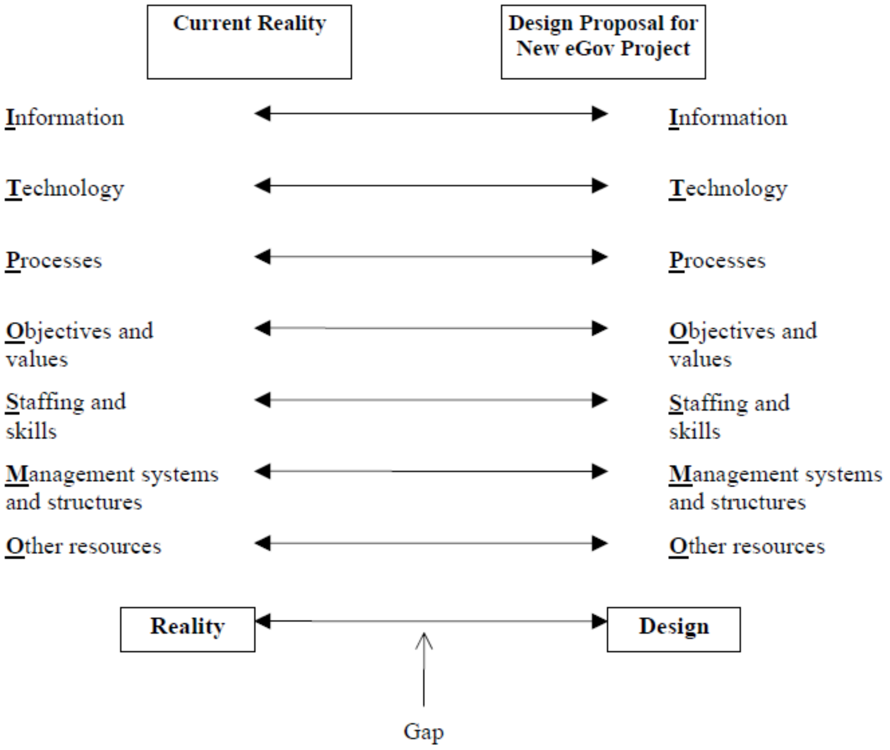
Constraints on the model attributes can then be specified using an architecture languages like Object Constraints Language (OCL) before evaluating the quality factors. Although EAA is a good methodology, architecture programming in languages like OCL is not a simple task even to a seasoned programmer because of the declarative nature of the languages used to describe meta-objects in the models.

The other method, Sturctural Equation Modeling (SEM), is a multivariate data analysis method that can be used to assess the influence of certain indicators on key quality attributes of an information system. Though this approach was originally developed and theorised by Ringle, Wande and Will in 2005 for use in marketing research, it has found application in many other areas including information systems (Wong, 2013). Its main features include ability to test linear and additive causal models, has good user interface and advanced reporting features. Since the technique can accept non-functional quality attributes and their causal relationships, it can be applied to a wide range of research modeling and data analysis needs (ibid). The technique uses regression analysis to compute the influence of independent variables on the dependent ones.

Heeks (2003) in his working paper “Most eGovernment-for-Development Projects Fail: How Can Risks be Reduced?” presents an optimisation technique based on gap analysis as a means of optimising eGovernment initiatives. He points out that there exists ‘design reality gaps’ and oversize gaps between project design and on-the-ground reality often lead to partial or total failure of projects. If the gaps are identified and situations in which failures are likely to occur are properly addressed then success of projects is realised. When assessing eGovernment projects, the dimensions of Information, Technology, Processes, Objectives and values, Staffing and skills, Management systems and structures and Other resources like time and money (ITPOSMO) need to be considered (ibid). Heeks model therefore seeks to manage the reality gaps between the current state and proposed design state in order to spur success in project implementation. Heeks model is depicted in Figure 2.10.Heeks approach requires the user to have a numerical scale of 0-10 which can then be used to estimate the size of the design reality gap as follows:

- **0** – no change between design proposal and current reality hence no gap.
- **5** – some degree of change between design proposal and current reality.
- **10 –** complete or radical change between design proposal and current rating.

Figure 2.10: Heeks Model – eGovernment Initiatives Gap Analysis



Source: iGovernment Working Series Paper No. 14 - Most eGovernment-for-Development Projects Fail: How Can Risks be Reduced? Heeks, 2003

After estimating values for all the seven areas, a simple addition can be made and interpretation of results can be done as follows:

- 57-70: eGovernment initiative will fail.
- 43-56: eGovernment initiative may fail if action is not taken to close the gap.

- 29-42: Initiative may fail totally or partially if action is not taken to close the gap.
- 15-28: Initiative will partially fail unless action is taken to close the gap.
- 0-14: Initiative may succeed without many hitches.

Since the problem at hand is to model a decision making support tool based on analysis of key quality factors, the research borrows the ideas expressed by Heeks model to come up with a seven level interval likart scale ranging from 1 – 7 {1 – not optimised (strongly disagree) -----> 7 – optimised (strongly agree)} for various security indicators. Responses from interview / questionnaires are used to populate a Structural Equation Model (SEM) created in SmartPLS. Heeks approach provides a foundation for designing relevant interval based assessment questions while PLS-SEM is used to analyse responses and provide viasual decision alternatives. This approach is preferred for this research because apart from wanting to gauge whether the best alternative has been selected when reasoning about the quality of individual attributes in PKI solutions, we also want to perform ANOVA on the data set collected to discover important statistical indicators. Based on this reasoning, a mean as close as possible to one (7) with as little variance as possible for example would indicate that a certain quality attribute is optimised while one that is close to seven (1) means the attribute is not optimised. One limitation of SEM however is that when the data set is large, it can generate erroneous results(Wong, 2013). This research has a small data set of less than one hundred respondents hence is not too large for the innaccuracies associated with SEM technique to kick in.

2.3.3 Regression Analysis and PLS - SEM

Partial Least Squares Structural Equation Modeling is an extension of the multiple linear regression analysis technique (StatSoft, 2000). A linear regression model helps a researcher to study the causal relationship that one variable (called the independent variable say X) has on a dependent variable say Y. Suppose for example we wish to observe the relationship between education E and salaries of information security experts S based on the two variables only and ignoring all the others that could have an effect on S. Let S be the earnings and E the independent variable influencing S based on number of years spent at school. Assuming that data about the salaries and education levels of the experts were collected and plotted in a chart as shown in Equation 1 then it would indeed appear that the more the number of years in education a person has the higher the income. This hypothesized relationship can be captured as follows in a simple regression model shown in Equation 1:

$$S = C_0 + \beta E + \varepsilon \text{ ……………………………………………………………………………….. (1)}$$

Where: S = salary of the expert (called the dependent or endogenous variable);

C_0= is baseline/constant earning with zero education;

β= is the positive effect on earnings for every year spent in school (called the regression coefficient) and

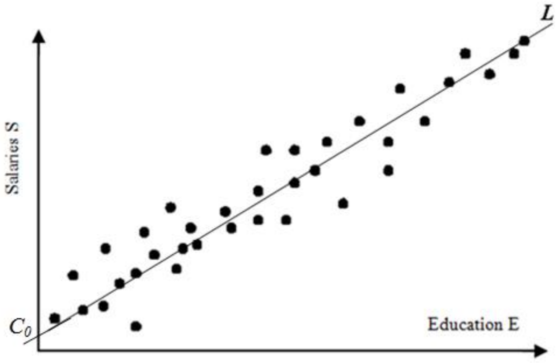
E = is the independent/exogenous/explanatory variable.

However, a careful study of the scatter chart may lead the researcher to conclude that it is not education alone that may influence earnings since there is no strict linearity displayed. Other unaccounted for factored like experience, productivity among others could have a significant impact. The researcher therefore includes an error term ε which represents all those variables that have a causal relationship on the income but are not directly observable (at times referred to as noise) (Skyles, 1992). If we set ε = 0 as in most cases, then the regression equation becomes the equation of a straight line in a 2-dimensinal plane with C_0 becoming the y-intercept and (E, S) being arbitrary points (x, y) that lie on the line and β the slope of the line as shown in (2).

$$S = C_0 + \beta E \qquad (2)$$

Now this means that somewhere on the scatter chart we can find a line which satisfies (2) and this can be found by estimating (predicting) the values of C_0 and β a task which requires considerable effort because many lines fit the bill. Hence the task is to find the best fit – a line L which best generalizes the data as shown in Figure 2.11. One way of achieving this is selecting the line that has the minimum sum of square errors.

Figure 2.11: Regression Analysis – Selecting Best Fit Solution



Source: An Introduction to Regression Analysis; Skyles 1992.

We now move on to PLS-SEM.Structural Equation Models, also called simultaneous equation models are multivariate or multiple linear regression analysis models (Fox and Weisberg, 2010). Unlike equation 1 where we only have a single influencing variable, we can model more variables say we add experience X to the model (1) resulting in (3). γ is modeled to be positive.

$$S = C_0 + \beta E + \gamma X + \varepsilon \qquad (3)$$

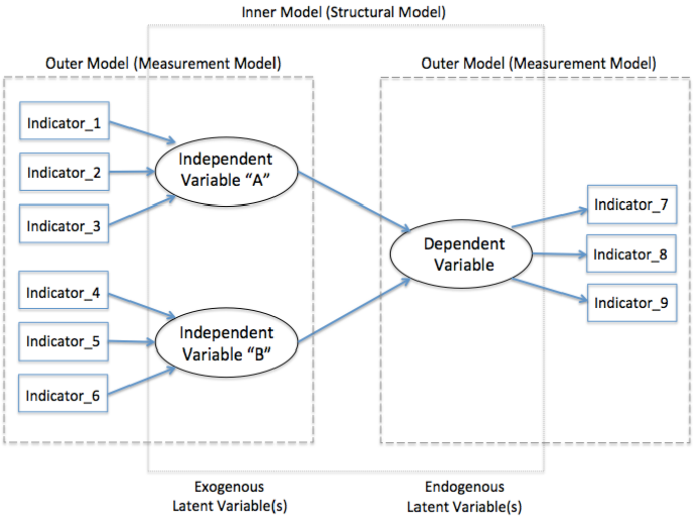
Equation 3 now has become a multi-regression and multivariate in nature. It now has two regression coefficients. It means that S is influenced by E and X and the task of estimating (predicting) values of C0, β and γ is no longer within 2-D space but 3-D, and on a plane rather than a simple straight line and relies purely on observable variables S, E and X. Unlike humans who find it challenging to reason in more than 3-D, the computer can perform analysis of many variables in n-D space(Schumacker and Lomax, 2012). Each factor enters the analysis independently and its causal impact can also be assessed independently e.g. possibility of answering questions like "Holding education constant, how does experience influence earnings?"

Partial Least Square (PLS) is an extension of multiple linear regression analysis equations(Wong, 2013). The **O** observations described by **D** dependent variables are stored in an **O×D** matrix denoted by **I**. The values of **P** predictors on the observations are stored in an **O×P** matrix **F**. PLS does not aim to find hyper planes of minimum variance between responses and independent variables, but to predict **I** from **F** by finding a linear regression model through creation of new spaces where observed and predicted variables can be plotted (University of North Carolina, 2007).Structural Equation Modeling is a technique for depicting relationships between variables with the aim of quantitatively testing the theory hypothesized by the researcher e.g. whether an independent variable influences the dependent one or not. In our case we use PLS-SEM tool that helps a person to model and do Analysis of Variance (ANOVA). A PLS-SEM model would have:

- Exogenous variables: independent variables. All causal arrows point away from it.
- Endogenous variables: dependent variables. Arrows point to it - causal effects.
- Indicators: observed measures or variablesused to infervalue of the latent variable.

Diagrammatically, a model take the form of Figure 2.12 (Wong, 2013) although in our case the model is reflective hence all indicator arrows point away from the variables.

Figure 2.12: PLS Structural Model



Source: Partial Least Squares Structureal Equation Modeling (PLS-SEM) Techniques Using SmartPLS; Wong, 2013

2.4 Decision Theories

In this study we believe that optimising decisions when reasoning about the security of PKI solutions eventually optimises PKI information security. Table 2.3 shows some decison theories:

Table 2.3: Decision Optimisation Theories

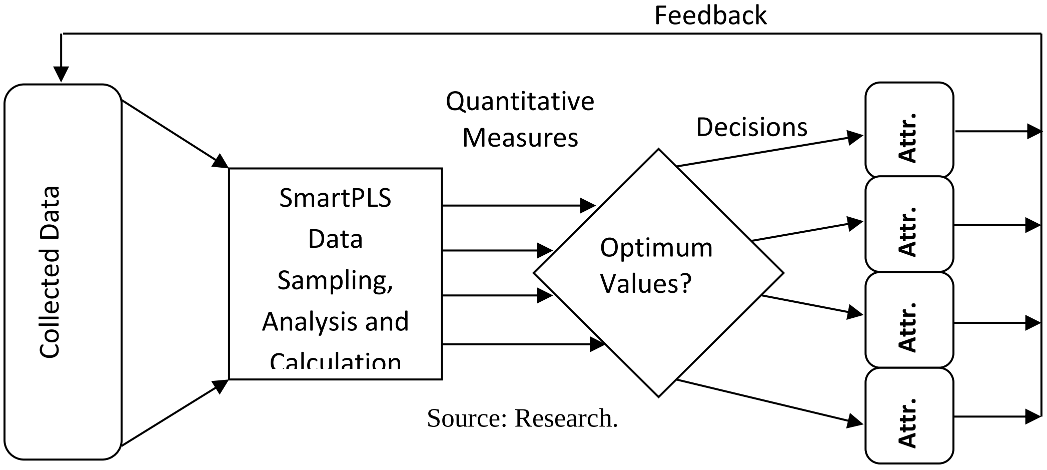
DECISION THEORY	CHARACTERISTICS
Rational Comprehensive Model	• Informed "economic-man" reasoning. • Process clearly defined: Intelligence, Design, Choice and review. • Clear problem definition, goals, objectives. • A cost/benefit analysis – rank alternatives. • Choose best alternative.
	Critique: Problems are hadly clearly defined; It needs a lot of information to make predictions and make decisions.
Incremental Theory	• Increment actions and keep the strategy open to adjustment • Work from status quo (current policy) only decisions that improve existing policy increamentally are selected. • Concensus among policy makers is key.
	Critique: Slow; Old policy still influences current decisions; Takes effort/difficult to bring all policy makers on board.

Bounded Rationality	• Based on Simon (1979) – the rational decision maker does not always have complete information and optional choices not always required. • Sequential searching and satisficing approach used; take first fittting solution.
	Critique: Decision taken may suffer from local optima problem.
Organisational Procedures View	• Decisions a result of standard procedures evoked by organisation sub-units (pre-programmed in existing procedures).
	Critique: inflexible; maintains status quo at expense of innovation.
Political View	• Decision making influenced by individual interests. • Organised influence and power plays a big role (coalitions).
	Critique: power games may lead to loss of objectivity; majority not always right.
Garbage Can View	• Power groups interact without much organisation to match problems to available solutions in a group or "can". Once the problem is solved the can is disbanded.
	Critique: solutions subjective to the composition of the can.

Source: Turpin and Marais, 2004: Decision Making: Theory and Practice

This research utilised the rational comprehensive model since we believed that the quantitative measures produced by this research would enable the decision maker have a clear perspective of the various quality attributes that influence security and the levels of their achievement in current solutions. A decision can be taken whether to give attention or allocate resources to a weak area.

Figure 2.13: Decision Optimisation Process



Source: Research.

The decision optimisation function is based on various statistical measures that were generated using regression analysis in SmartPLS software. Generally, the optimised decisions were a result of the process represented in Figure 2.13.

2.5 Approaches to IAM Systems

IAM approaches can be categorized into four:

1. **Silo/ Isolated Identity Management:** Each agency or organisation in government acts as an independent identity manager or provider. The agency relies on a government given token to verify a person's identity or issues its own identity credentials to citizens (OECD, 2011). Although this strategy is the best in terms of simplicity and privacy, it causes credential overload to citizens and password fatigue. It is also very expensive in the long run since duplication of infrastructure across various identity silos in government means spiraling costs.
2. **Centralised Identity Management:** In this approach, a single trusted entity acts as the identity provider, manager and verifier for all other interested parties. It eliminates some of the problems of the silo approach i.e. due to centralisation, Single Sign On (SSO) is achieved hence eliminating credential overload / password fatigue. Privacy of the users can be easily compromised because it creates a focal point of attack.
3. **Federated Identity Management:** In this approach, instead of having one trusted entity, several trusted entities form a circle of trust. A user registered by one entity is trusted across board hence removing the weaknesses of the centralised approach (Birrell and Schneider, 2013). The challenge here is

administrative – the need for inter-communication between various trusted entities to verify identity assertions (FIDIS WP2, 2005).

4. **User Centric Identity Management:** In this approach, the user is in charge determining what identity credentials are used in which environment. Systems have to request for the users permission before forwarding identity credentials to any requesting party(FIDIS WP2, 2005).

Let us look at the above approaches from the industries perspective based on real solutions available on the market. Microsoft .NET Passport, SAML Tokens, Liberty Alliance Project, OpenID, Web Services Federation (WS-Federation), Microsoft's InfoCard among others are all examples of various solutions out there on the market, some obsolete while others have larger clientele. We shall also look at the principles and technological motivations that have impacted this field.Current research on electronic and identity management from consortiums and companies such as Microsoft, FIDIS, Open Group, OASIS, Liberty, Shibboleth among others informed this thesis immensely because they bring together many experts in identity management.

Identity management has generally evolved from service based to centralized, then from centralized to federated. Service based required users to register with a service and it was the responsibility of the service to maintain a database of user identities. This model created fatigue to users who had to maintain long lists of credentials. It was because of this that Microsoft in 1999 came up with a centralised Single Sign-On (SSO) solution known as Microsoft .NET Passport. The main aim is to allow a user to have only one .NET passport issued credential and use it to access other sites that are members of the .NETPassport group (Goldtack, 2006).

The centralized nature of Passport and its dependence on cookies opened a leeway for attackers to commit exploits by not only having a focal point of attack but also acquiring session cookies and passport tickets that have authentication primitives hence enabling them to impersonate genuine users. Users were also wary of one provider being in custody of all their identities (Ibid). In 2002, the Organisation for the Advancement of Structured Information Standards (OASIS) group developed the Security Assertion Markup Language (SAML). This is an XML-based open standard for exchanging authentication, entitlement and attribute information across domains.

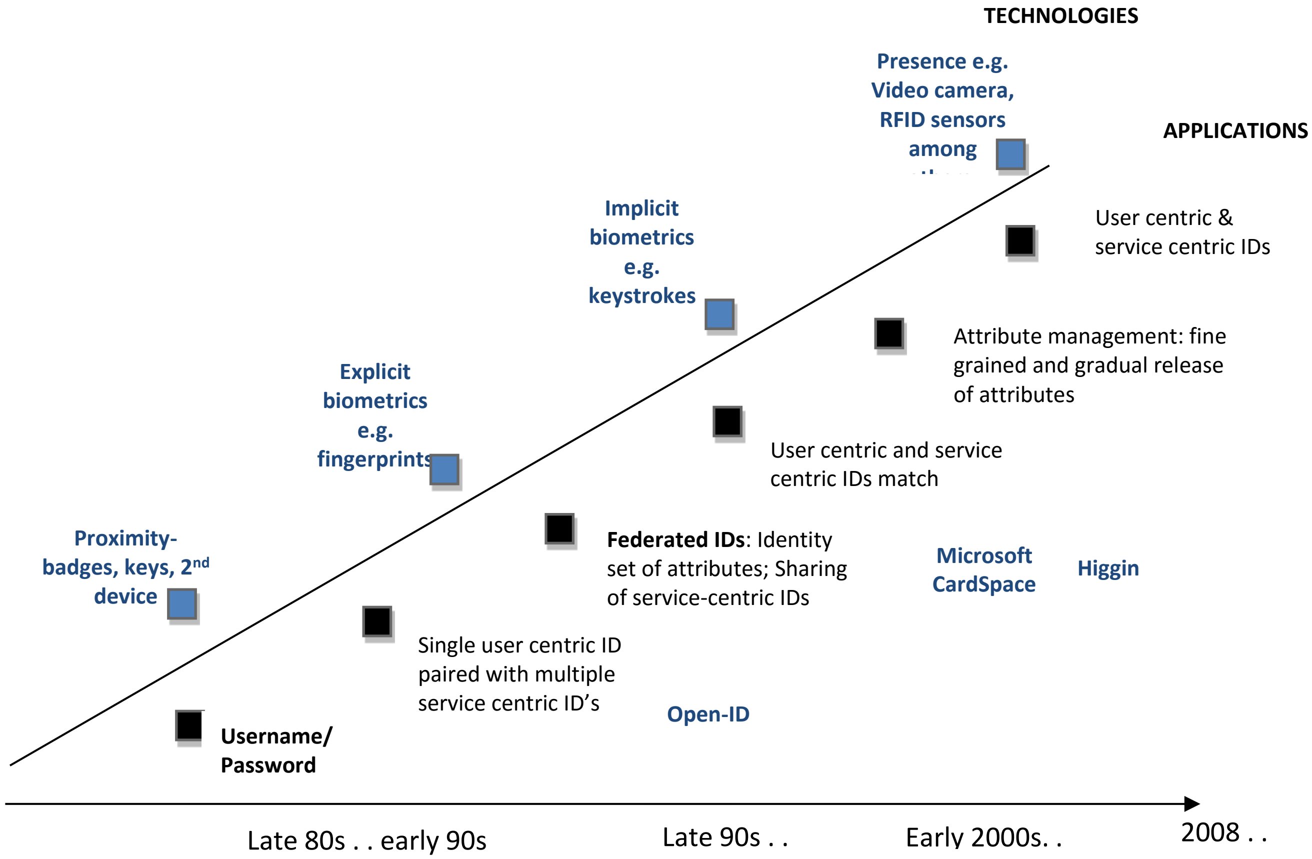
The Liberty Alliance, OASIS web services and Internet2 Shibboleth project adopted SAML as their communication standard. SAML requires a Prover to register with a local verifier. The verifier is more closely trusted by the service provider. The verifier provides authentication services and assures other service providers that the Prover is whom they claim to be, and this creates a trust system (Wisniewski et al., 2005). Federation is the key principle in today's identity management scenarios. SAML is one of the most widely accepted standards that has been and can be used to create this identity management layer. This layer defines mechanisms and formats for communication of identity information between various domains. OASIS web services (WS-Security) committee, WS-* group, OASIS' XACML and the Web2 Shibboleth project all use SAML. SAML has two major components:

- SAML *assertions*: they describe security tokens that represent users. These assertions in XML format transfer user identity (and attributes) from identity providers to service providers in a neutral way.

- SAML *bindings:* they represent bindings and profiles for a single sign on protocol.

Figure 2.14 below shows a summary of technologies and applications as they evolved.

Figure 2.14: Evolution of Identity Management Applications


TECHNOLOGIES
APPLICATIONS
Presence e.g. Video camera, RFID sensors among
Implicit biometrics e.g. keystrokes
Explicit biometrics e.g. fingerprints
Proximity- badges, keys, 2nd device
User centric & service centric IDs
Attribute management: fine grained and gradual release of attributes
User centric and service centric IDs match
Federated IDs: Identity set of attributes; Sharing of service-centric IDs
Microsoft CardSpace
Higgin
Single user centric ID paired with multiple service centric ID's
Open-ID
Username/ Password
Late 80s . . early 90s
Late 90s . .
Early 2000s. .
2008 . .

Source: ITU-T; Adopted from ITTU-T IdMFG Framework for Workgroup; ITU-T; 2003

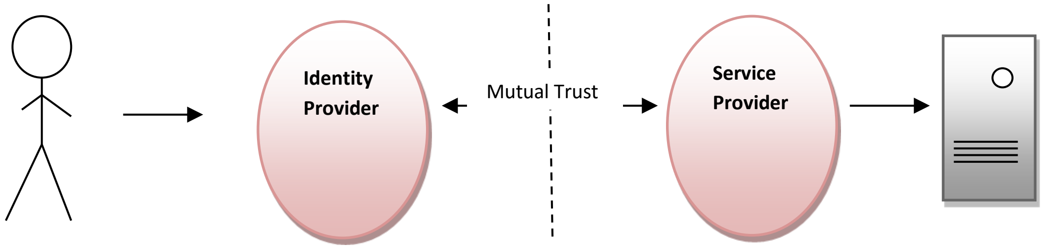
SAML 1.x (1.1 and 1.2) were adopted for use in the Liberty Alliance ID-FF 1.2 to provide SSO functionality while SAML 2.0 is more advanced and addresses security concerns that faced 1.x versions.

Microsoft did not give up after the failure of .NET Passport. Kim Cameron who was Microsoft's Chief Architect of Access management came up with the famous Seven Laws of Identity on which Windows CardSpace (also named InfoCard) is based.

CardSpace selected different user identity data by generating objects from an identity class and displaying them on the screen as virtual information cards or credentials that the user can select for use in a particular context. It is built on the Web Services Protocol stack but does not interoperate with SAML. This lack of interoperability means that if one government agency implements solutions based on CardSpace and another on SAML then it would be expensive and difficult to integrate the two platforms

The Web Services Federation (WS-Federation) defines ways in which different security frameworks can federate and interoperate. This is done through brokerage of identity, authentication, attributes, and authorization assertions between different identity management frameworks and enforcement of privacy of federated claims (OASIS, 2009). WS-Federation specification builds on WS-Trust, WS-Security, WS-Policy and WS-* standards. It uses Extensible Markup Language (XML), Simple Object Access Protocol (SOAP) and Web Services Description Language (WSDL) extensibility models to create a building block that can interoperate with other web services, transport and application specific protocols to implement secure solutions. The important roles in identity federation are can be summarized into two: the identity provider and the service provider. A mutual trust relationship exists between the two organisations as shown in Figure 2.15 below.

Figure 2.15: Mutual Trust between Actors



Source: Research.

The Liberty alliance project was set up to create open, technical specifications that enable single sign on (SSO) mechanism, authorization, identity mapping, account linking and directory services in federated networks. Another goal was to enable a permission based attribute-sharing network where users have a say on the use and sharing of their personal information. It uses SAML to implement its identity management layer (Cole et al., 2003). Since managing different user profiles separately on different sites is cumbersome and costly to the user and service providers, the Liberty Alliance Project aimed at putting in place mechanisms that would make it possible to share identity information across domains. The idea was to enable users connect to multiple sets of identity information across e-commerce sites in order to create one simple-to-manage federated identity. This allowed single-sign-on and cheaper management of user profiles across a federation (Alsaleh and Adams, 2006). The Liberty group gave birth to the Identity Federation framework (ID-FF), Identity Web Services Framework (ID-WSF) and Identity Service Interface Specifications (ID-SIS).

The ID-FF defines a scheme for federating identities and mechanisms for SSO using this identity. It does this by enabling users having different accounts at different liberty enabled sites to link them and use them for SSO. On the other hand, ID-WSF defines a framework targeting web services which makes sure that service providers can share user identities in a carefully crafted permission based manner. Features of ID-WSF include permission based attribute sharing, identity service discovery (find out identity and attribute providers) and interaction service which seeks to get user permissions. Finally, ID-SIS defines interfaces for interoperable exchange of identity attributes between different providers, including interfaces for sharing registration, contact book, geo-location data among others.

The main critiques of the Liberty Project point out that the single-sign-on (SSO) mechanism means that an identity thief can perform great exploits within the Circle of Trust (CoT) if he/she manages to compromise log in credentials of one Liberty Enabled (LE) site.

Last but not least, openID which is a Single Sign-on (SSO) centralized web identity management system based on API was developed. It provides a username and password to a registered user. Once signed in, OpenID tells all other subscribed websites that the user is who they are without showing them the actual credentials, hence avoiding possible compromise of the credentials through insecure sites (Patel and Oza, 2013). Now, with all the good intentions of providing open, decentralized, free framework for single identity management, OpenID has been criticized as follows; Loskot (2008) summarizes the problems as vulnerability to phishing attacks, browser exploits based on XSS and CSRF, and a Trojan or key logger attack can steal the single username and password hence compromise all other accounts under the scheme.

2.5.1 Identity and Access Management Maturity Levels

Many eGovernment maturity models have been proposed by various researchers, research groups and experts in eGovernment. Though they do not directly address IAM maturity, they shall be reviewed as a means of extracting important aspects that do. We give more emphasis to research groups although a few individual researchers were considered. Table 2.4 gives a summary of various models. X means stage not available.

Table 2.4: Comparative Maturity Models

MODEL	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6
Gartner, 2000	Website presence: static websites, informational	Interaction: email, web forms, chat, download pdf/doc files	Transaction: online transactions e.g. cash payments.	Transformation: integrated & personalized services	X	X

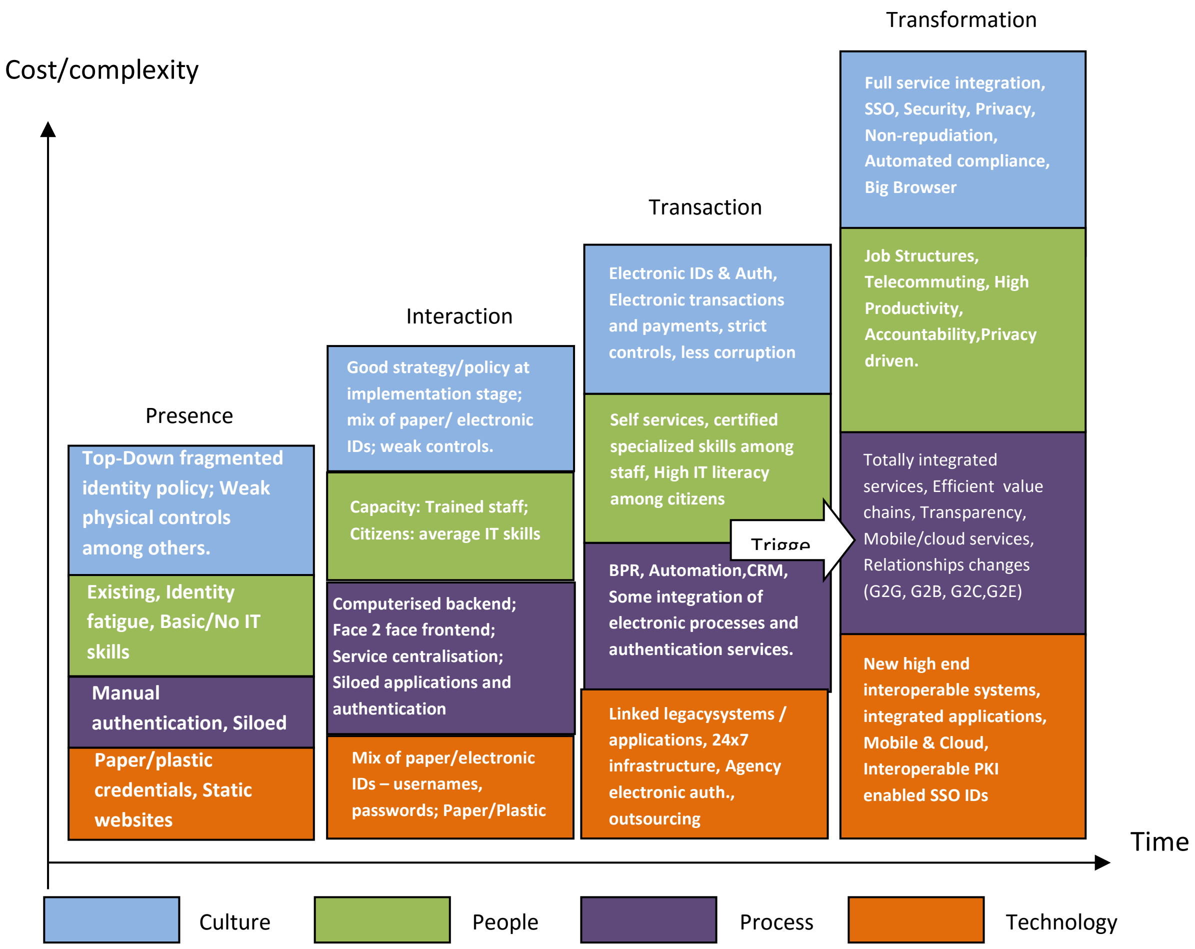
United Nations, 2012	Emerging information: websites with static information.	Enhanced information services: one way/simple two way communication	Transactional services: two way interaction.	Connected services: citizen/customer centric services; proactive services with feedback loop.	X	X
World Bank, 2003	Publish: information published on websites e.g. rules, regulations, documents, forms	Interact: citizens give feedback/comments on legislative/policy proposals	Transact: perform secure transactions online.	X	X	X
Ernst &Young,	Initial: Manual, Informal, non-standardised processes	Repeatable: Similar manual processes across board	Defined: Standardised and documented processes, SLA’s	Managed: automation, fragmented processes, SLA compliance monitoring	Optimised: Integrated automated processes, Exceed SLA’s	X
Accenture	Presence: information published online	Basic capability:	Service availability:	Mature delivery:	Service transformation:	X
KPMG	Immature: manual, ad-hoc per application; no common security policy	Aware: manual, ad-hoc per application groups; per application authentication	Capable: Common services; no common IAM modules; automated where appropriate	Mature: automated user management per classes of application	Industry leading: integrated IAM across applications	X
UK National Audit, 2002	Basic sites: few sites, basic information.	Electronic Publishing: websites with many pages	ePublishing: Emails with prompt responses; customized services	Transactional:		
Deloitte and Touche, 2000	Information publishing	Two way transactions	Multipurpose portals	Portal personalisations	Clustering of common purposes	Full integration and online transactions

Kuppinger Cole, 2007	Basic IM: basic provisioning; basic web access	Advanced IM: decentralised access management integration	Service Oriented IM: centralized federated services	Business Driven IM: integrated ID driven business systems;	X	X
Al-Khaouri, 2011	Presence: public approval; website markup; people hardly use technology.	Interaction: Email services; support skills; searchable databases;	Transaction: Full online transactions; eAuthentication; high skills.	Transformation: Integrated services; performance driven	X	X

Source: Research

e-Government scales the phases of growth depicted in Figure 2.16 over time:

Figure 2.16: Phases of eGovernment IAM Maturity


Cost/complexity
Time
Presence
Top-Down fragmented identity policy; Weak physical controls among others.
Existing, Identity fatigue, Basic/No IT skills
Manual authentication, Siloed
Paper/plastic credentials, Static websites
Interaction
Good strategy/policy at implementation stage; mix of paper/ electronic IDs; weak controls.
Capacity: Trained staff; Citizens: average IT skills
Computerised backend; Face 2 face frontend; Service centralisation; Siloed applications and authentication
Mix of paper/electronic IDs – usernames, passwords; Paper/Plastic
Transaction
Electronic IDs & Auth, Electronic transactions and payments, strict controls, less corruption
Self services, certified specialized skills among staff, High IT literacy among citizens
Trigge
BPR, Automation,CRM, Some integration of electronic processes and authentication services.
Linked legacysystems / applications, 24x7 infrastructure, Agency electronic auth., outsourcing
Transformation
Full service integration, SSO, Security, Privacy, Non-repudiation, Automated compliance, Big Browser
Job Structures, Telecommuting, High Productivity, Accountability,Privacy driven.
Totally integrated services, Efficient value chains, Transparency, Mobile/cloud services, Relationships changes (G2G, G2B, G2C,G2E)
New high end interoperable systems, integrated applications, Mobile & Cloud, Interoperable PKI enabled SSO IDs
Culture
People
Process
Technology

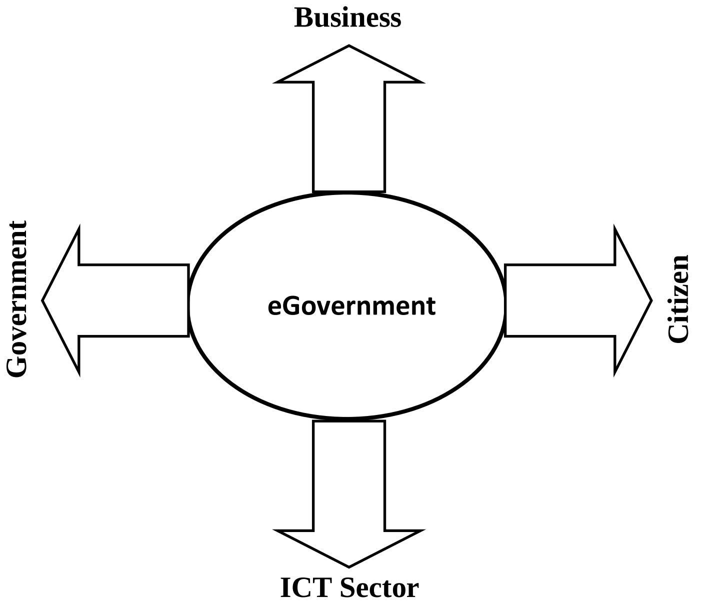
Source: Modified from People Process Technology e-Government Maturity Model; Al-Khouri, 2011

From Figure 2.16, we can conclude that as e-Government matures, so is the need for better electronic identity and access management. At the presence and interaction stages of growth, front office/telephone and simplistic user name/password kind of authentications are sufficient. However, as e-Government grows and starts supporting mission critical services at the transaction and transformation level, secure and trustworthy electronic identity authentications supported by a legal framework are required to mitigate emergent risks.

2.6 Sample PKI Enabled e-Government Initiatives

In the words of Satyanarayana (2004), eGovernment is not about the "e" but about government, neither is it about hardware and software but about services to the citizens. The main stakeholders in any eGovernment undertaking are summarised in Figure 2.17.

Figure 2.17: eGovernment Stakeholders



Source: e-Government: The Science of the Possible; Satyanarayana, 2004.

Although establishing e-Government is a complex process, it can be generalised to follow a few simple steps as listed below:

1. Secure political will/support for the e-Government initiative at all levels of Government. The process should be initiated from the highest office with a coordinating agency reporting directly to the leader of government/state. Champions at each level to be appointed.
2. Develop the right legal and regulatory framework for e-Government. The legal framework should clearly define the rights, responsibilities and obligations of all stakeholders.
3. Provide enough budgetary support for e-Government, to cover all necessary areas like acquisition of hardware, software, networks, consultancy, human capital, training, incentives to the private sector among others.
4. Establish Government agencies in all key sectors that would enable strategic planning and operationalisation of initiatives, to champion the e-Government agenda. These agencies report to the main cordinating agency.
5. Establishment of a secure national high speed data communications backbone linking various political regions and government agencies.
6. Establishment of secure local area networks, intranets and extranets linking various government agencies through the backbone.
7. Establish an electronic trading, taxation and payment platform.
8. Establish an electronic database for citizens holding all relevant biodata.

9. Establish service oriented and secure Government distributed information systems, servers, portals, databases and Public Key Infrastructure (PKI) to support full electronic service delivery to citizens and companies in relevant sectors.
10. Reengineer processes and train civil servants at all levels in relevant ICT skills - to man Government ICT resources and utilise them to deliver electronic services.
11. Be inclusive by establishing multilingual ICT enabled call centers, public access points, automated electronic help lines and centralised service delivery centers to cut red tape.
12. Train the public in information technology skills in order to increase uptake of electronic Government services.
13. Have top class maintenance and continuos improvement strategies for Government ICT resources.

PKI plays a central role in e-Government service delivery mechanism because it is usually integrated into the eGovernment identity and access management system to enable stronger authentication of online (Al-Khouri, 2011a). This is because certain Government services like those requiring exchange of sensitive information or involving payments need not only to be secure but support nonrepudiation. Therefore, an identity and access management agency must be put in place to spearhead all matters identity in e-Government. There are several models or architectures of deploying eGovernment. We looked at how PKI enabled identity and access management defines these architectures by studying a few examples across the world then come up with a model that is easy to set up and that is modular in nature.

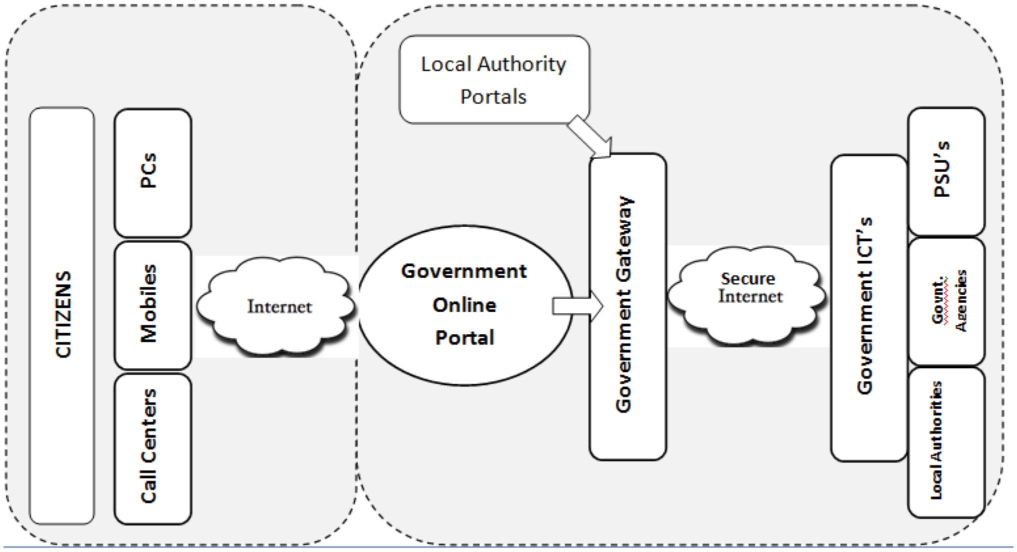
The United Kingdom Model (UK) was chosen for study because according to the United Nations 2014 eGovernment ranking, UK stands at position 8 worldwide, but third in Europe after France and Netherlands (UNPAN, 2014). Therefore, UK is in a good position to serve as a good country to learn from in matters eGovernment. Estonia's eGovernment has faced massive cyber attacks e.g. in 2007, from allegedly Russian hackers since inception(Ashmore, 2009; Aaviksoo, 2010). These cyber attacks against Estonia led to collapse of nearly all ICT supported services in the country e.g. education, banking, national security etc. (Ashmore, 2009). These attacks raised awareness of the vulnerabilities that eGovernment and any other IT structures connected to the internet face. With the help of the United Nations, Europe and the North Atlantic Trade Organisation (NATO), Estonia has managed to overcome most of the threats to set up one of the best eGovernments in the world (UNPAN, 2014). Studying the estonian model is therefore a good initiative for anyone who wishes to learn how a small country overcame adversity to set up one of the best eGovernments in the world.

2.6.1 The United Kingdom Model

The United Kingdom (UK) is one of the worlds leading countries when it comes to adoption of technology in Government. In Europe, other countries that lead in this area include Denmark, German, Sweden, Netherlands among others.

In the UK model, all users of e-Government services are authenticated through a common Government Gateway (GG). The GG performs the following services: (1) It is the central registration platform (2) It is the main identification agency (3) It ensures security of Government to Citizen (G2C) and Government to Business (G2B) transactions over the internet through proper authentication. After successful authentication, the user is allowed to access various government services via Government Secure Internet (Gsi) which links various government agencies as shown in Figure 2.18. Users are authenticated using either a digital certificate where relevant or based on a username/password issued by the GG for transactions that do not require certificates. One advantage of this model is the centralisation of authentication services for all government agency online services. The government therefore can invest all resources and effort to make sure that the gateway enforces the required standards of security. The disadvantage of course is creation of a single point of failure and creating a focal point of attack.

Figure 2.18: United Kingdom eGovernment Architecture

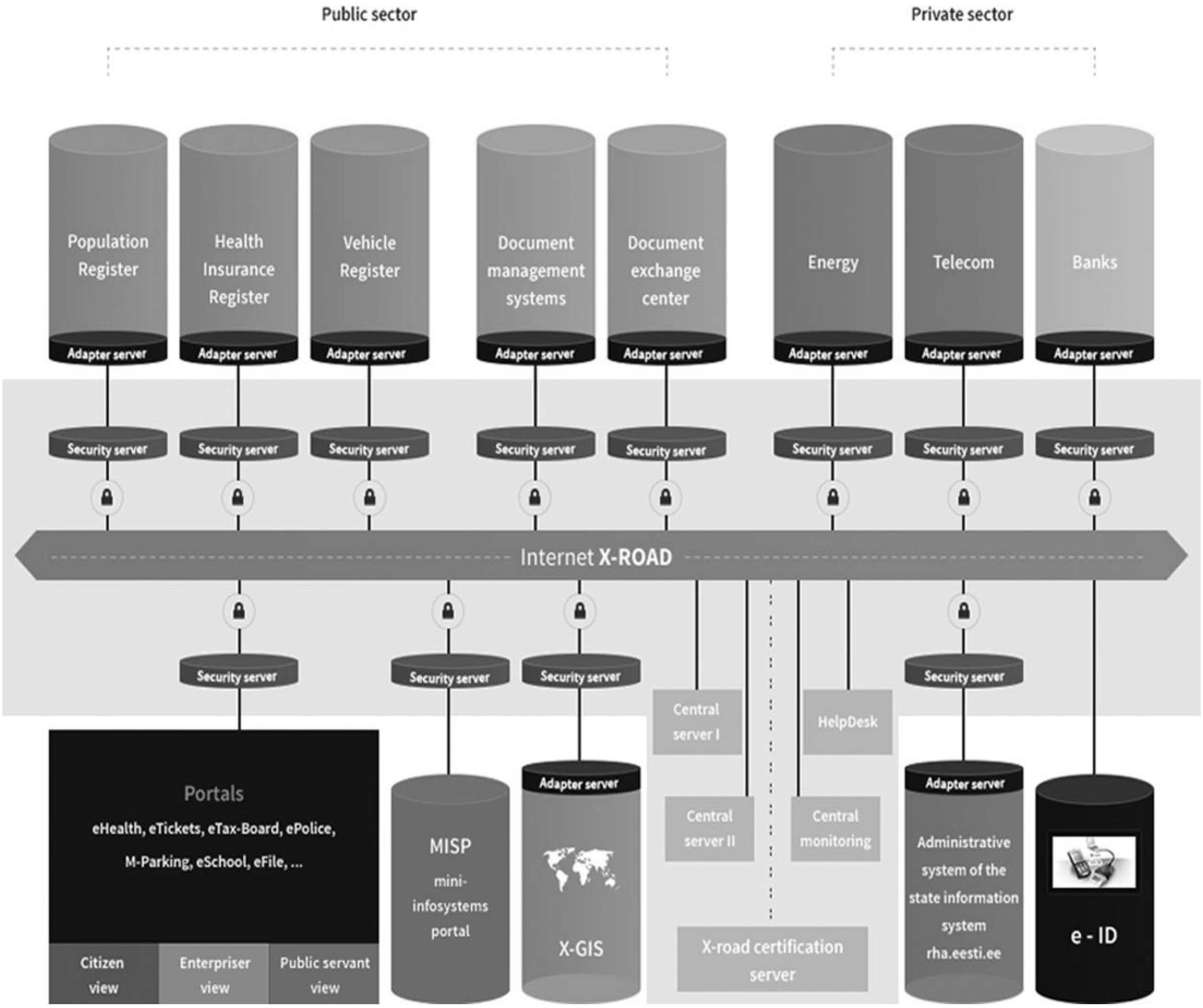


Source: Adopted from e-Government: The Science of the Possible; Satyanarayana, 2004.

2.6.2 The Estonian Model

The Estonian model differs from the United Kingdoms' one in that instead of a single government gateway, each agency connects to a high speed internet backbone through its own secured gateway as shown in Figure 2.19. This model distributes risk across various agencies and gives each one of them the power to authenticate those who wish to access its online services. Although the single point of failure and attack is eliminated, the administrative challenge is scaled and the need to make sure that all departments adhere to the same standards. It also follows that each agency has the freedom to provide their own digital certificate used to authenticate citizens hence a user with an electronic ID can be authenticated at each server.

Figure 2.19: Estonian eGovernment Architecture



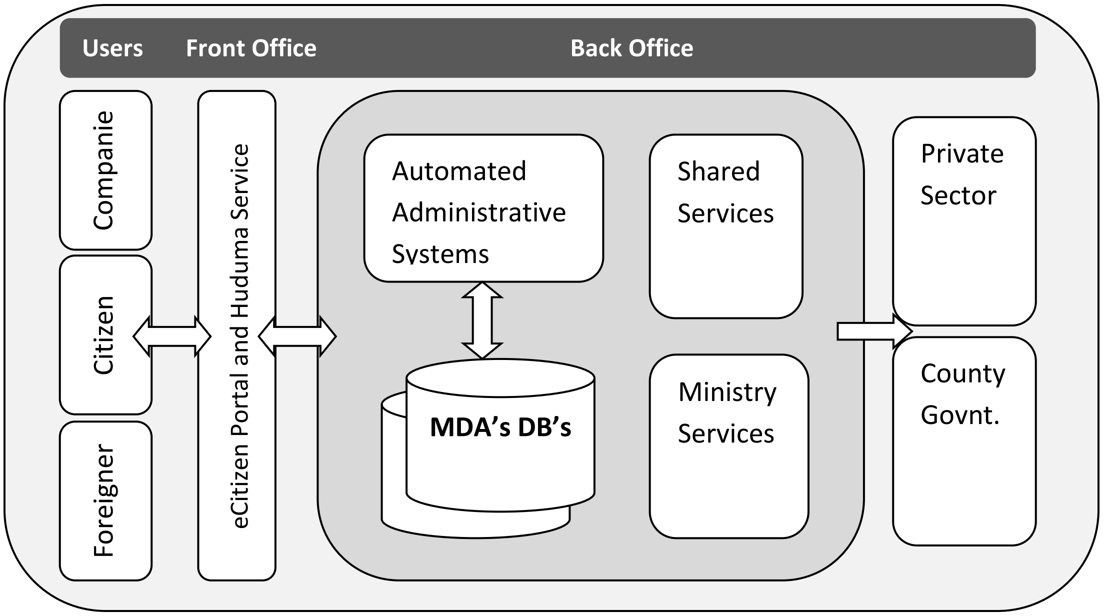
Source: Estonian Example of Integration eGovernment Services; Kalja, 2005.

A review of the USA (Whitehouse,2012), Egypt (Eid, 2009), among others shows that the best practice is to craft eGovernment models that envision secure access to key government services.

2.6.3 The Kenyan Model

The Kenyan Government has not developed a unique model but is currently relying on a generic model as depicted in Figure 2.20. From this model, one that combines eGovernment and front office contact with citizens through Huduma Centers was derived. In both models, there is lack of a clear full automated online security and authentication layer. The Huduma model relies on a hybrid where there is a mix of paper credentials and online data queried by a government official. A scan across government shows that individual agencies are at different stages of adopting technology. There is need to move to full automation. Secondly there is need to develop an eGovernment model, complete with a national security and authentication strategy covering services offered by both the central and county governments. There should be a clear separation between the external internet common to everybody, from the government PKI gateway secured internet where government systems run. Privare sector installations should not be part of the secure government infrastructure as is in the generic model.

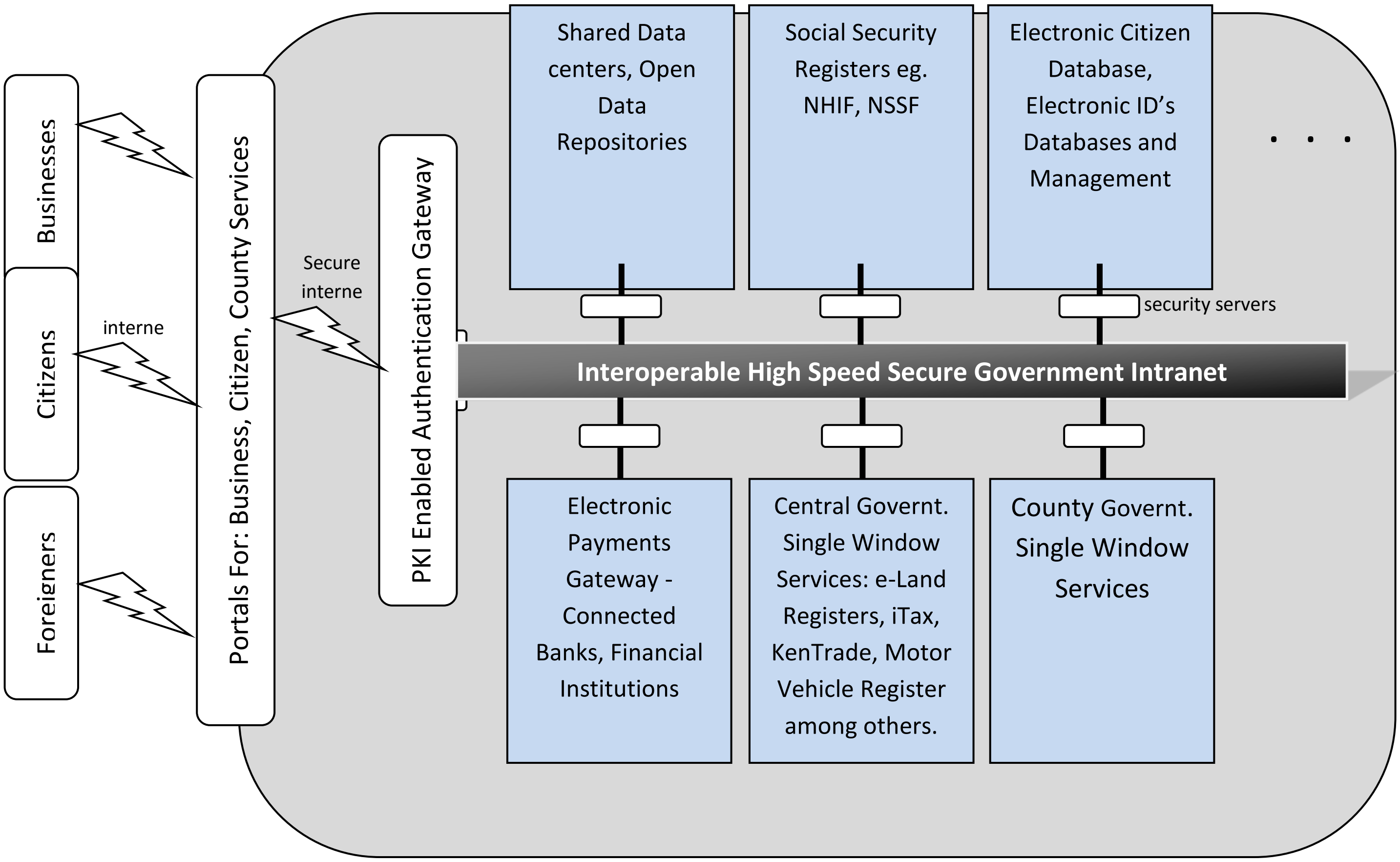
Figure 2.20: Generic eGovernment Model Adopted by Kenya



Source: GoK; Modified from The Kenya ICT Masterplan 2013/14-17/18; GoK, 2014.

The model in Figure 2.21 proposes integration of county and central government services in order to standardise service delivery across board. This is specifically important for Kenya because the devolved units have just set out planning, procuring and deploying their ICT infrastructure. There is great need to have their planning, procurement and deployment under the national ICT and eGovernment strategy to avoid pulling in different directions or setting up many silos all over that do not interoperate in harmony.The proposed model must have multilevel security designed within its architecture with one main security gateway and each agency having a security server too. Though expensive, multiple security servers are very important in enhancing inter-agency security even behind the PKI gateway e.g. the military may be keen to protect its assets from other government agencies or users on the secure internet.

Figure 2.21: Proposed e-Government Architecture Modular Arrangement


Businesses
Citizens
Foreigners
interne
Portals For: Business, Citizen, County Services
Secure interne
PKI Enabled Authentication Gateway
Shared Data centers, Open Data Repositories
Social Security Registers eg. NHIF, NSSF
Electronic Citizen Database, Electronic ID's Databases and Management
. . .
security servers
Interoperable High Speed Secure Government Intranet
Electronic Payments Gateway - Connected Banks, Financial Institutions
Central Govern t. Single Window Services: e-Land Registers, iTax, KenTrade, Motor Vehicle Register among others.
County Govern t. Single Window Services

Source: Research;

Knowing and listing various initiatives is one thing but measuring and predicting the qualityfactors in the final solutions in order to optimise decision making is a worthy challenge. When planning and implementing PKI solutions, managers are often faced by many potential solutions in a search space and have to make one choice over the other. At one end of the decision-making process is the goal while on the other end are the decision alternatives (Johnson and Ekstedt, 2007). The quality of final solutions lies in a well informed and optimised decision making process. This thesis argues that predicting the influence of various attributes on quality properties enables mangers make the right rational decisions not only on the critical success factors but how to invest resources for optimum quality achievenent. Solutions that have already been implemented can be improved based on such modeling. Achievement of the goal/success criteria depends on the selection of available alternativese.g. one quality factor relevant to setting up service portals could be security. For example, we know that firewalls, encryption, intrusion detection systems among others when utilised contribute towards secure systems. Similarly different firewals for example have varing levels of security assuarance in final solutions. If these alternatives are modelled and quantitative comparisons of their possible future impacts are generated, managers can use the results to make rational decisions on what would give optimum security when designing, implementing or improving existing enterprise information systems.

2.7 Current State of Electronic Identity Management in Kenya

Kenya is quickly adopting technology in Government as a means of delivering services to citizens and businesses. The merger of the Kenya ICT Board, the Directorate of e-Government (DeG) and the Government Information Technology Services (GITS) into one entity The Kenya ICT Authority (KICTA) was a good strategic move as a means of consolidating all information technology functions. KICTA now falls under the Ministry of ICT. Also, the Kenya Information and Communication Amendment Act 2013 has been signed into law and establishes the Communications Authority of Kenya (CAK) to replace the Communication Commission of Kenya. The Directorate of e-Government mission statement reads as follows:

> “ to provide quality information and services to and enable online interactions with the public, businesses and other Government units in a convenient and secure manner through the innovative use of ICT”

There is concerted effort everywhere in Government to transit from paper based processes to electronic under the guidance of KICTA, CAK among others. Many ministries and their agencies have succeeded at the very least to set up static websites that help to disseminate information to the public. Although most services still remain manual, citizens and companies are quickly acquiring electronic identities in many ways and some Government departments like Kenya Revenue Authority (KRA) have started offering some electronic services to individuals and businesses. In fact, KRA have been granted a Certification Authority (CA) licence in order to effectively roll out their iTax platform which has services such as PIN application, tax returns and import/export duty payments done online. Table 2.5 shows some examples of the most common identity tokens currently available in Kenya. In the last column to the right, a (-) means the token is not digitised while an (X) means it is digitised.

Table 2.5: Identity Tokens Available in Kenya

Identity Token Name	**Issuing Agency**	**Attributes on Token**	**Digital**
National Identity Card (ID)	National Registration Bureau – Ministry of State for Immigration & Registration of Persons	ID number, Photo, Fingerprint, First name, Middle name, Surname, DOB, District, Location, Sub-location, Tribe.	-

Birth Certificate	Department of Civil Registration – Ministry of State for Immigration & Registration of Persons	Name of father, Name of mother, Place of Birth, Name of child, Date of Birth, Gender	-
Passport	Department of Immigration – Ministry of State for Immigration & Registration of Persons	Passport number, First name, middle name, surname, citizenship, Normal and biometric photos	X
PIN	Kenya Revenue Authority	PIN, First name, Second name, surname.	X
Voters Card	Independent Boundaries and Electoral Commission	Names, ID number, Fingerprint, voting station	X
Educational Certificates	KNEC, Colleges and Universities	Names, Examining body, Grades, Year of examination, Certificate number	-
Bank Cards	Banks	Names, Bank, Branch, Account number, Validity	X
Mobile SIM cards	Mobile Service Operators	IMEI, SIM Number, Geo-location, Owner names, Owner national ID number.	X
Driving License	Kenya Revenue Authority/Kenya Police Traffic Department	Names, DOB, Vehicle class, Drivers photo, National ID number, DL number	-
Title Deeds	Ministry of Lands	Names, Deed number, Land Number, Location, Size	-
Motor Vehicle Log Book (s)	Kenya Revenue Authority	Names of owner, Vehicle Type, YOM, Ownership history, Vehicle color, Chassis number, Number plate	-
Central Depository System Account (CDS)	Capital Markets Authority (CMA)	Names of account holder, NationalID number, PIN	X
User Name and Password	www.ecitizen.go.ke	ID Number, First Name on National ID	X
Digital Signature/Certificate based on PKI	Communications Authority of Kenya (CCK) and Approved CAs like KICT Authority	Digital signatures, private and public encryption keys, digital certificates	X

Source: Research

Currently, the management of the seven main types of Government sponsored identities in Kenya is performed by the following departments which now fall under the Kenya Citizens and Foreign Nationals Management Service under the Ministry of Interior and Co-ordination of National Government:

1. National Identity Card (NID): The National Registration Bureau (NRB) - deals with issuance/replacing/changing details on identity cards.

2. Birth/Death certificates: The Department of Civil Registration (DCR)

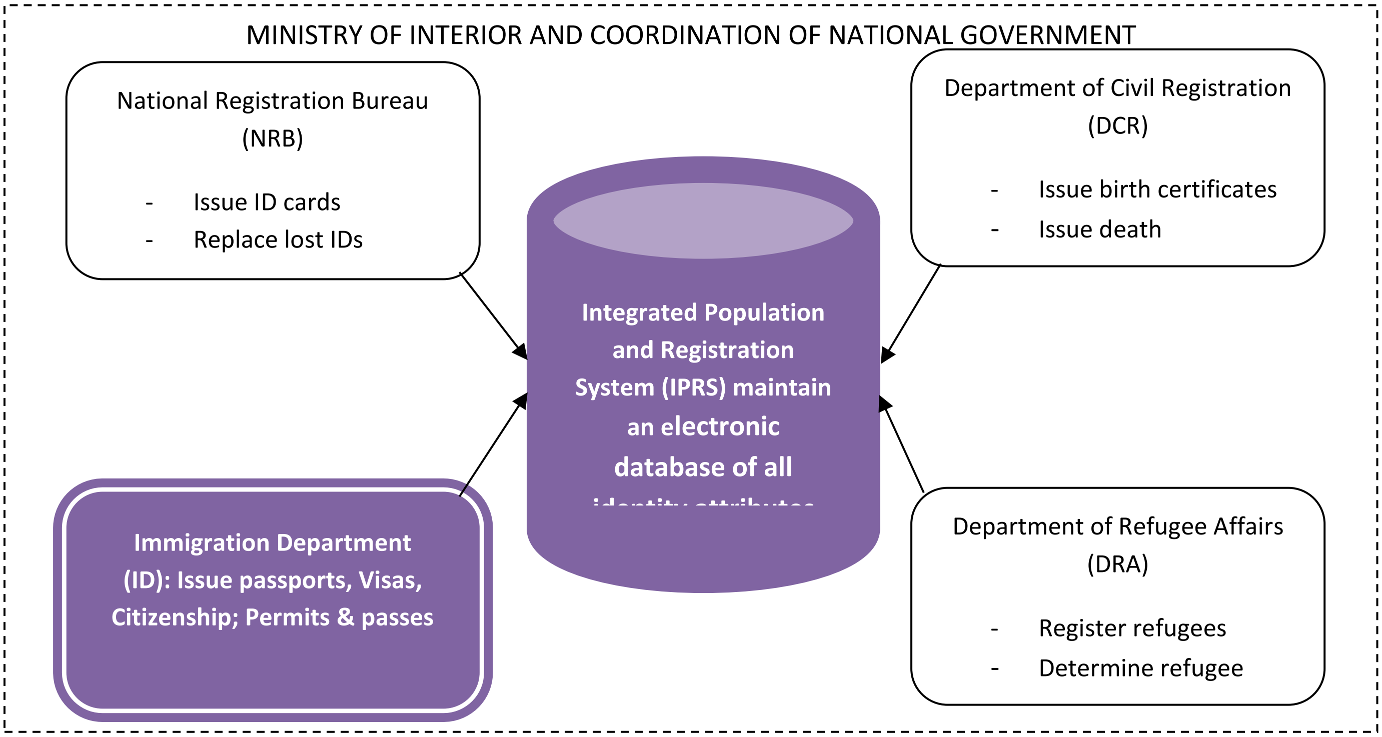
3. Passports/Visas/Application for Citizenship: The Immigration Department (ID)

4. Refugee Registration: Department of Refugee Affairs (DRA) - registers refugees/ issues movement pass / determines status.

5. The Integrated Population and Registration System (IPRS): stores the information from the first four departments above in a computer based database.

Figure 2.22 below shows a top level diagram depicting how the above agencies relate with the IPRS. The Government was to use this data to authenticate and develop a new electronic database for citizens which had been planned to start by March, 2015.

Figure 2.22: Integrating Identity Attributes


MINISTRY OF INTERIOR AND COORDINATION OF NATIONAL GOVERNMENT
National Registration Bureau (NRB)
- Issue ID cards
- Replace lost IDs
Department of Civil Registration (DCR)
- Issue birth certificates
- Issue death
Integrated Population and Registration System (IPRS) maintain an electronic database of all
Immigration Department (ID): Issue passports, Visas, Citizenship; Permits & passes
Department of Refugee Affairs (DRA)
- Register refugees
- Determine refugee

Source: Research; Ministry of Interior and Coordination of National Government, 2014

Other significant citizen identity providers in Kenya are:

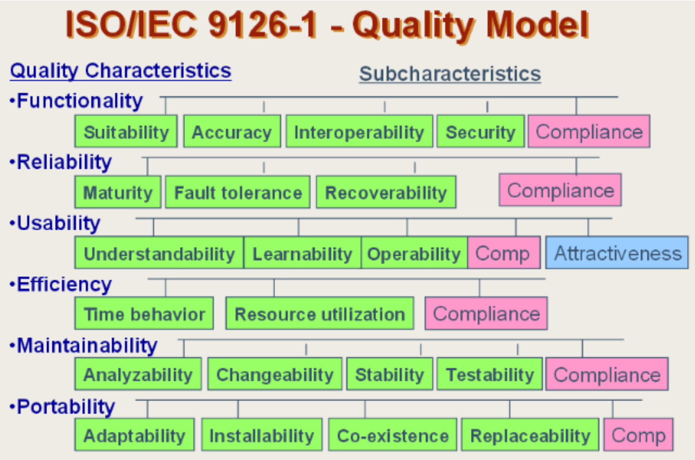
1. Biometric Voter Register: Independent Electoral and Boundaries Commission. (IEBC).
2. The driving license and PIN: Kenya Revenue Authority (KRA)
3. The Digital Certificates and Signature: Communications Authority of Kenya (CAK)

2.8 Research Gap Identification

As has been previosly discussed, eGovernment initiatives like electronic identity cards (eIDs), electronic taxation (eTax), electronic procurement (eProcurement), electronic commerce (eCommerce), electronic land register, electronic social security management, electronic health (eHealth), electronic passport (ePassport), electronic banking (eBanking), electronic driving licence, electronic birth certificates among others are all identity driven. It therefore follows that the quality of identity and access managent systems like PKI in eGovernment impacts greatly on all sectors of electronic service delivery hence underlining the need to have internal assessment frameworks that can quantitatively measure partinent quality attributes.

Our contribution here is developing a quantitative decision optimisation tool when assessing, measuring and predicting the achievement of PKI quality attributes. Managers are usually faced by great challenegs when reasoning about the quality of information systems due to a large space full of many alternatives, different context factors, lack of standards, challenging implementation requirements, many attributes to be measured, evolving requirements among others. (Hays et al., 2005; ISC, 2003;Choudhury et al., 2002). Most governments and their agencies rely on the winning contactor or supplier when doing quality assessments, mostly assessing only whether an implemented system works or not (Heeks, 2003). Better still, a few have put in place benchmarking and quality management systems like International Organisation for Standards (ISO) and insist on all actors being compliant. Now benchmarking is limited in the scope of qualities it can be used to assess (mostly measurable) while ISO is document trail intensive. Another quality management system, the balanced score card is good but relies mostly on financial metrics as a measure of organisational performance (Kaplan and School, 2010). The ISO 9126 – (1 – 4) and later ISO 25030 (which is part of the Software Quality and Requirements Evaluation (SQuaRE) the ISO 25000 series), set the non-functional quality characteristics/properties that need to be assessed when evaluating software as *functionality*, *reliability*,*usability*, *efficiency*, *maintainability* and *portability* as shown in Figure 2.23 below (Narman et al., 2007; Zubrow, 2004). Notice that each characteristic has functional sub-characteristics which refine the model and help us to measure the main characteristic. Asessment criteria and and relevant properties are sourced from literature where research has already been done. In most cases, a mix of already researched and empirical study has to be adopted in order to come up with the right assessment criteria and to identify relevant properties.

Figure 2.23: ISO 9126-1 Quality Model



Adopted from Zubrow,2004.

The American Information Security Committee (ISC, 2003) came up with the PKI Assessment Guide (PAG) which details various ways of assessing the quality of PKI systems. The PAG divides assessment into two: those done before starting operation and those done periodically to ensure compliance with the Certificate Policies (CP) and the Certificate Practise Statement (CPS).

2.9 Quality Properties and Attributes Derivation

The PKI security decision optimisation model proposed by this thesis was derived from a careful study of various papers written and published in journals, by research groups or industry practitioners. Although an exhaustive review of all literature on identity management is possible, it is not necessary since a representative sample of the works from leading authorities in the field quickly identifies trends and those qualities or attributes they agree on. To build on the ISO and SEI publications, standards which identify general quality attributes when assessing software quality, twenty four extra publications specifically written to address identity and access management have been reviewed and by coincidence twenty four partinent quality properties identified. To measure the variables, relevant measureable attributes have to be used either sourced from literature, from current industry practise or empirically where they do not exist. All the quality properties are important, whether their frequency is low or high and therefore they should be represented in the generalised model derived from Table 2.6.

Assessing each endogenic variable is not trivial and hence after coming up with a generalised model for all, the variable with the highest frequency was chosen for more specialised modeling and measurement as a demonstration of how all the others can also be assessed. In this case, there is an agreement even with the experts in the field that security is the most important.

Table 2.6: PKI Quality Properties from Literature

Sources Property	Aleti et al., 2013	Johnson et al., 2007	I T U - T	Vanamali,2004	Al - Khouri, 2011b	Smedingoff,2010	Aicholzer, 2010	K P M G, 2008	K upinger Cole	Ramskrishnan, 2011	Baltzer, 1990	Kwinter et al., 2008	F I D I S WP4, 2006	Oostdijk, 2009	Stephan, 2010	Stefanova, 2006	I B M, 2007	Entrust, 2009	Information ACT Kenya	Gartner, 2010	PAG, 2003	Accenture, 2010	Ijaz, 2012	Hong Kong Gov. 2008	Frequency
Security	X		X	X	X	X	X	X			X			X	X	X	X	X	X	X	X	X	X	X	19
Compliance	X		X	X	X	X		X	X		X		X	X	X	X	X	X	X	X	X			X	18
Cost	X		X		X	X	X	X			X			X	X	X	X		X	X	X	X		X	16
Privacy			X	X	X	X	X	X						X	X	X	X		X	X	X	X		X	15
Efficiency			X	X	X	X	X				X		X		X		X	X	X	X	X	X	X		14
Interoperability		X	X		X	X		X	X						X	X	X		X		X	X	X		13
Risk				X	X	X		X			X		X	X	X	X	X			X	X	X			13
Availability	X	X	X										X		X	X	X	X	X	X	X		X		12
Human Capacity					X	X		X		X	X		X				X		X	X	X		X	X	12
Performance	X		X		X	X					X		X				X	X		X	X	X	X		12
Trust				X	X	X					X			X	X		X	X	X		X	X		X	12
Legal / Policies	X							X	X				X	X		X		X	X		X	X		X	11
Usability		X	X				X	X	X			X	X		X			X			X				10
Integration					X			X	X						X	X	X	X		X	X	X			10
Accounting			X		X	X									X		X	X	X		X	X		X	10
Reliability	X	X	X		X	X							X		X				X		X				09
Strategy				X	X	X		X	X							X	X				X	X			09
Integrity				X	X	X								X		X	X		X		X		X		09
Culture				X		X			X	X	X						X		X		X				08
Automation			X	X	X			X	X												X				06
Accuracy		X	X			X										X	X				X				06
Disaster Recovery			X					X										X		X	X				05
Confidentiality																X			X		X		X		04
Modifiability	X	X																							02

Source: Research Literature Review.

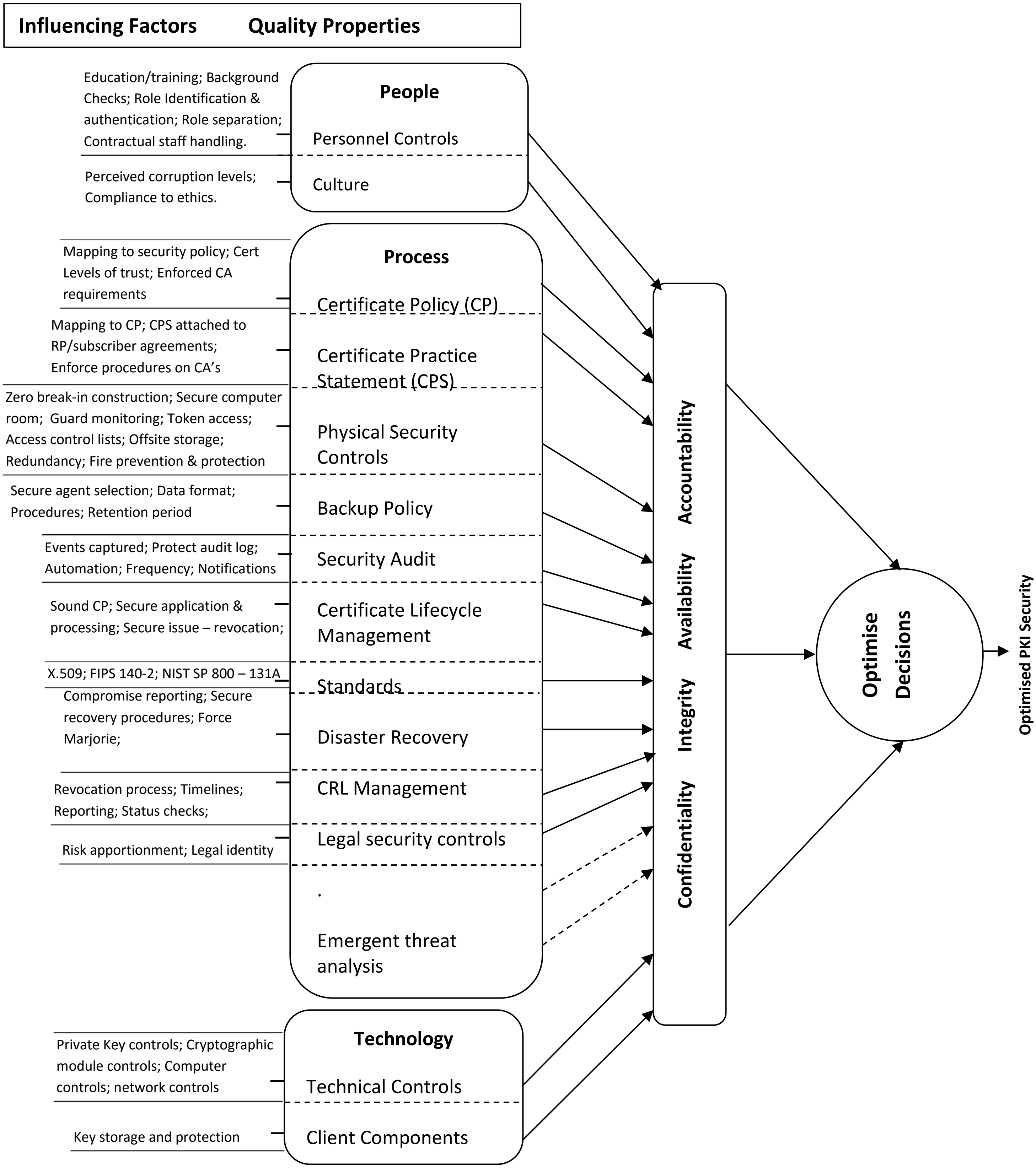
Based on the ISO 9126 model, security is classified as a variable under functionality. In this thesis it stands out alone as one of the main quality properties to be assessed. The variables that were used to assess PKI security are captured in Table 2.6 arranged based on CPPT framework. The variables in Table 2.7 are not directly derived from Table 2.6 but from the literature.

Table 2.7: Proposed PKI Quality Assessment Properties

QUALITY PROPERTIES	**INFLUENCING VARIABLES**
PEOPLE	
Personnel Controls	Education, Identification, Role separation, Contract staff.
Culture	Code of ethics, Perceived corruption.
PROCESS	
Certificate Policy	Mapping to security policy (SP); Certificate levels of trust; Interoperability.
Certificate Practice Statement	Mapping to CPs; Completeness - RP/Subscriber Agreements.
Physical Security Controls	For CAs, RAs & Subscribers
Backup Policy	Data types; Protection of backups; Retention period; Backup procedures
Security Audit	Types of events captured for audit; Protection of audit log; Frequency of audits; Audit collection system; Notifications.
Certificate Lifecycle Management	Sound CP; Secure application & processing; Secure issue – revocation
Standards	X.509; FIPS 140-2; NIST SP 800 – 131A
Disaster Recovery	Redundancy; Secure facility; Revoked public key; Compromised private keys; Operation after force majeure; Backup policy; Reporting.
CRL Management	Common CRL? Online Certificate Status Protocol (OCSP)? Certificate List; Extensions; Version numbers; Distribution Points (CLDP)?
Legal/Business Risk Management	Legal responsibilities; Accountability; Risk apportionment.
TECHNOLOGY	
Technical Security Controls	Network security controls; Computer security controls; Cryptographic module controls; Algorithm selection; Key size; Key pair generation; Private key delivery;
Client Side Components	Smartcards; Components in OS/Applications

Source: Research

Figure 2.24: Conceptual Framework


Influencing Factors
Quality Properties
People
Education/training; Background Checks; Role Identification & authentication; Role separation; Contractual staff handling.
Personnel Controls
Perceived corruption levels; Compliance to ethics.
Culture
Process
Mapping to security policy; Cert Levels of trust; Enforced CA requirements
Certificate Policy (CP)
Mapping to CP; CPS attached to RP/subscriber agreements; Enforce procedures on CA's
Certificate Practice Statement (CPS)
Zero break-in construction; Secure computer room; Guard monitoring; Token access; Access control lists; Offsite storage; Redundancy; Fire prevention & protection
Physical Security Controls
Secure agent selection; Data format; Procedures; Retention period
Backup Policy
Events captured; Protect audit log; Automation; Frequency; Notifications
Security Audit
Sound CP; Secure application & processing; Secure issue – revocation;
Certificate Lifecycle Management
X.509; FIPS 140-2; NIST SP 800 – 131A
Standards
Compromise reporting; Secure recovery procedures; Force Marjorie;
Disaster Recovery
Revocation process; Timelines; Reporting; Status checks;
CRL Management
Risk apportionment; Legal identity
Legal security controls
.
Emergent threat analysis
Technology
Private Key controls; Cryptographic module controls; Computer controls; network controls
Technical Controls
Key storage and protection
Client Components
Accountability
Availability
Integrity
Confidentiality
Optimise Decisions
Optimised PKI Security

Source: Research;

2.10 PKI Security Variables Explanations

The main variables that have been selected to characterize the security of PKI solutions for eGovernment are summarized in the conceptual framework in Figure 2.27. Each variable has various attributes that influence it. These variables and attributes were selected from a pool of many more. The research made an effort to choose those variables that have the greatest impact if compromised but does not in any way suggest that those which were left out are not important. The conceptual framework is extensible and more can be added as need arises.

2.10.1 Personnel Controls

Personnel control deal with those measures that ensure that the human resource employed within the PKI industry are well educated/trained, qualified, competent, of high integrity and are in a position to offer or run highly available, efficient and trustworthy operations. The influencing attributes under this variable include but are not limited to:

i. Education and training requirements: staff must have the best ICT and security qualifications and competencies.

ii. Background checks: before employing staff, thorough background checks need to be performed e.g. validation of educational certificates, references, and security background checks among others.

iii. Role identification and authentication: each staff must be properly identified and authenticated on the PKI system with strict privilege administration.

iv. Role separation: separation of roles is very important. Also, sensitive tasks should have more than one staff authentication requirement.

v. Contractual staff: sourcing and management of contract staff is important. Proper background checks and binding contracts on non-exposure of sensitive data and secrets must be signed.

2.10.2 Culture

The culture of a country or organisation plays a key role in PKI security. A culture of corruption and nepotism results in low quality solutions that are procured expensively. It also leads to easy compromise of staff which may lead to internally instigated systems attacks. The two attributes selected here are:

i. Corruption perception: it is important to gauge the level of corruption perception among employees.

ii. Professional ethics: there should be well documented and enforced ethical standards.

2.10.3 Certificate Policy (CP)

A certificate Policy (CP) is a very important document for any PKI. It is usually derived directly from the national information security or cyber security strategy. A CP is a named set of rules that specifies how a particular certificate is applicable to particular applications or communities that have shared security requirements. It outlines the requirements that stakeholders need to meet in order to be part of a particular PKI. It is the certificate policy which structures digital certificates into levels of trust e.g. basic, medium and high security. Also, the regulating authority uses the CP to form, assess and grant operating rights to Certificate Authorities. The attributes here include:

i. Mapping to security policy: there should be a strong mapping between the national cyber security policy and the certificate policy. This mapping makes sure that there is a disciplined PKI environment across board as every CA must be evaluated based on the CP. Also, CAs must have a head start when coming up with their Certificate Practice Statements (CPS).

ii. Certificates levels of trust: whether the CP organizes certificates into levels of trust.

iii. Enforce CA requirements: whether the CP is used to enforced strict requirements to be met by aspiring or operating CAs.

2.10.4 Certificate Practice Statement (CPS)

This document is usually derived from the CP by the Certificate Authority. A CPS is a complete set of rules that a CA uses when issuing digital certificates. As such, the CPS is an extremely important document. The attributes that were used to assess this include:

i. The mapping of the CPS to CP.

ii. Whether the CPS is usually attached to agreements with clients or relying parties.

iii. Whether the CPS is used to enforce procedures on the CA.

2.10.5 Physical Security Controls

Physical security controls are very important. The site hosting the PKI needs to be constructed based on the total security zone principles. Logical controls alone are not capable of providing total protection to the cryptographic keys, hence the need for strict physical controls. Break ins, fires, natural disasters among others are great threats. The attributes identified her include:

i. Access control lists: it is important that every visitor should be authorized and they must sign in and out. In some place, only those on a special list should be allowed entry.

ii. Offsite storage: critical storage equipment, cryptographic keys among others should be stored safely away from the site to reduce the risk of theft, destruction by disasters among others.

iii. Redundancy: servers, storage disks among others should have redundant sites.

iv. Fire prevention and protection: availability of fire extinguishers and other fire prevention measures like smoke detectors.

2.10.6 Backup Policy

The backup policy usually deals with the types of data backed up, retention periods for backups, protection of backups and backup procedures. We shall assess this variable using the following attributes:

i. Secure agent selection: when selecting a third party backup agent, the process of identifying the agent must be credible and ensure security of the data.

ii. Data format: was the backed up data encrypted or not? Encryption gives better security.

iii. Backup procedures: all procedures must ensure creation and protection of the backups, their integrity, and ensure unauthorized access or modification.

iv. Retention period: the period should be sufficient and meet the requirements of the law and contractual obligations.

2.10.7 Security Audit

The security audit tries to ascertain whether standard procedures are being followed to the letter in all implementations and operations of the PKI. In this case, the main standard utilised is the X.509 standard. We use the following attributes to assess this variable:

i. Events captures: whether there was a clear policy on which events were captured

ii. Protection of the audit log: the audit log must be protected from modification even by those who handle it.

iii. Automation: whether all audit events are automatically logged.

iv. Frequency: how frequently is the audit done?

v. Notifications: is there a mechanism for informing audit offenders? How fast?

2.10.8 Certificate Lifecycle Management

This variable assesses the full lifecycle management of a digital identity. All the procedures must be secure. It looks at secure measures from registration to issuance, revocation, renewal or replacement, disposal of digital identities. The attributes here include:

i. A sound certificate policy: the certificate policy sets the standards for the registration and regulation of CA. It also enforces key requirements on creation and revocation of certificates.

ii. Secure application and processing: the application process must be transparent and secure to avoid capture of wrong identity data in the first place. All the processing of digital certificates need to be secure.

iii. Secure issue and revocation: after the identity has been created, how is it passed to the owner? The issue process can be compromised easily if enough care is not taken. It is better to have physical delivery to owner after proof of identity. For eGovernment, a phone call or code followed by transfer of certificates over an encrypted link may be adopted. Also, how is a certificate revoked? Is there a requirement for signed email? etc.

2.10.9 Standards

Check which standards the PKI implementation adheres to:

i. X.509: this is the default standard.

ii. FIPS 140-2: Federal Information Processing Standard (**FIPS**) Publication **140-2**is used to accredit cryptographic modules. It covers eleven (11) areas dealing with design and implementation of cryptographic module. Developed by Cryptographic Module Testing

(CMT) laboratories and accredited by National Voluntary Laboratory Accreditation Program (NVLAP) of USA that does testing of cryptographic modules.

iii. SP800-131A: A standard that requires compliance to stronger cryptographic keys and robust algorithms i.e. Keysize = 2048 for strict adherence.

2.10.9.1 Disaster Recovery

This refers to ability to continue operations in the face of disasters. Proper plans and mitigation measures must be put in place. The following attributes were selected:

i. Compromise reporting: do we have secure channels of reporting compromise of keys? How many?

ii. Secure recovery procedures: in case of a disaster, do we have well laid procedures for recovery?

iii. Force Marjorie: in case of floods, fires, earthquakes, solar flares among others do we have mitigation measures?

2.10.9.2 CRL Management

CRL management is a key feature of any secure PKI. Revoked certificates can be a source of insecurity i.e. replay attacks if the lists are not well managed. The attributes we chose here include:

i. Revocation process: needs to be secure – a procedure to trigger revocation needs to meet certain set standard e.g. secure email signed with keys among others.

ii. Timeliness: strict timelines on compromise reporting and lifecycle processing to come up with new certificate.

iii. Reporting: how many channels of reporting are there? The more the better.

iv. Status checks: does the PKI use online certificate status protocol (OCSP) with stapling to increase efficiency? Stapling requires that an end user queries the CRL and gets an all clear report which is then attached to a relying party request. The RP therefore does not have to query the CRL and this improves efficiency.

2.10.9.3 Legal Security Controls

Every PKI should be rooted in a robust legal framework. The main issue here is whether electronic signatures are recognized as the proper identity of a person and if there is a good scheme to deal with matters of or risk and indemnity. The attributes here are:

i. Risk apportionment: clear risk apportionment in case a need arises.

ii. Indemnity issues: who was to take the blame and meet the cost in case breaches or other unplanned for occurrences take place?

2.10.9.4 Technical Controls

This are the technical aspects of the PKI and how they are handled. The key attributes here are:

i. Private Key controls: storage of the private key. Is it escrowed?

ii. Computer controls: is the computing base secure?

iii. Network controls: do we have a secure firewall?

2.10.9.5 Client Components

This is concerned with the packaging of the digital certificates at the clients end. Smartcards that require two factor authentications would be more secure for example than if the certificates were in a flash disk. The attributes here include:

i. Key storage: where are the keys stored at the client end? In a flash disk, computer or smartcard?

ii. What protective measures are end users obligated to observe in order to protect private keys?

2.11 Variable Operationalisation

Table 2.8 shows how various security quality properties and their influencing variables were mapped from the conceptual framework to the PLS-SEM tool and how data about each one of them was captured. This table can be interpreted as follows:

Column 1: This are the main PKI security quality attributes identified in the conceptual framework. Each quality attribute has its influencing variables in Column 3

Column 2: It shows which question on the questionnaire in Appendix 1 was designed to collect data about each influencing variable in Column 3 of the table.

Column 3: This are the security quality variables for each of the quality attributes in column 1. Data shall be collected for each of this variables.

Column 4: It indicates how each of the variables shall be measured i.e. what type of data shall be collected about each variable e.g. Yes/No or values of a Likart scale.

Column 5: It shows the label for each of the variables in Column 3 as an indicator in the PLS-SEM model as depicted in Figure 3.4.

Table2.8: Variable Operationalisation

Main Quality	Questionnaire Question	Variables	Measurements	PLS-SEM Model Indicator
Personnel Controls	1	Relevant Education/Training	Likart scale	**Education**
	2	Background checks	Likart scale	**Backchecks**
	3	Role identification &authentication	Likart scale	**RoleIDAuth**
	4	Multi-staff authentication	Likart scale	**MultiStaffAuth**
	5	Contractual staff / consultants policy (bonding, agreements, indemnity, vetting,	Yes/No	**ContractStaffVet**
Culture	6	Compliance to code of conduct	Yes/No	**CULTUREEthics**
	7	Corruption index/perception	Likart scale	**CULTURECorrupt**
Certificate Policy (CP)	8	Degree of one on one mapping to security policy	Yes/No	**CPtoSPMapp**
	9	Structure certificates into levels of trust?	Yes/No	**CertTrustLevels**
	10	CP enforces auditable requirements & controls on all CAs	Yes/No	**CPnCAControls**
Certificate Practice Statement (CPS)	11	Does it map to CP directly?	Yes/No	**CPSnMaptoCP**
	12	Whole CPS attached to RP/Subscriber agreements/contracts?	Yes/No	**CPSonAgreements**
	13	Enforces auditable procedures on CAs based on CP?	Yes/No	**CPSenforceCAProcedure**
Physical Security Controls	14	Zero break in physical access controls {High security zone construction, CA in	Yes/No	**SecureBase**
	15	Media storage: offsite/separate location, redundancy.	Yes/No	**OffsiteMediaStorage**
	16	Fire prevention and protection	Yes/No	**FirePrevention**
Backup policy	17	Secure procedure selecting backup/archival agent?	Yes/No	**BackupSecureAgent**
	18	Format of backed up data (plaintext / encrypted / split key)	Likart	**BackupDataFormat**
	19	Data backup procedures?	Yes/No	**BackupSecureProcedure**
	20	Backup retention periods policy/sufficiency	Yes/No	**BackupRetention**
Security audit	21	Policy on types of events captured	Yes/No	**AuditEvents**
	22	Protection of audit log (can be changed, cannot be changed)	Yes/No	**AuditLogProtection**

Main Quality	Questionnaire Question	Variables	Measurements	PLS-SEM Model Indicator
	23	Level of automation - audit collection systems (manual, semi-automated, fully	Yes/No	AuditAutomation
	24	Frequency of audits (never, not often, often, very often)	Likart scale	AuditFrequency
	25	Notification process to offenders/action	Yes/No	AuditNotification
Certificate /	26	Secure sound certificate policy	Yes/No	LifecycleCP
	27	Secure application process & processing	Yes/No	LifecycleRekey
	28	Secure Rekeying/modification/renewal/recovery	Yes if ALL; No if Not ALL.	LifecycleSecure
Standards	29	X.509, FIPS 140-2, NIST SP 800 – 131A	Yes if ALL; No if Not ALL.	Standards
Compromise & Disaster	30	Multiple/quick compromise reporting mechanisms	Yes if > 1; No if only 1	CDRMultiReporting
	31	Secure recovery procedures: resource corruption / key comp	Yes/No	CDRRecovery
	32	Force Marjorie – strategies to ensure continuity in aftermath	Likart	CDRForceMajorie
CRL Management	33	Revocation: Secure / Clear guidelines – DS signed message?	Yes/No	CRLRevoProcedure
	34	Strict timelines – Reporting & CA certificate reprocessing	Yes/No	CRLTimelines
	35	Is CRL repository OCSP enabled / mandatory status checks	Yes/No	CRLOSSP
Technical Security				
Private Key Controls	36	Who generates private/public key? CA, RA or Subscriber?	Yes/No	TECHPKGen
	37	Key passing SSL/signed email/snail mail/token/porter?	Yes if ALL; No for less than 2	TECKKeyPass
	38	Restriction mechanism on key usage – x.509 v. 3 certificates?	Yes/No	TECHx509Certs
Cryptographic module	39	Is the module hardware/software/firmware/hybrid?	Likart	CRYPTModuleType
	40	Key lengths – comply to NIST SP 800 – 131A	Likart	CRYPTKeyLength
	41	Private key under n of m (n=m=2)	Yes/No	CRYPTPKNofM
	42	Is private key escrowed? If so agent well vetted?	Yes/No	CRPTEscrowed
	43	Secure activation/deactivation/destruction private key	Yes/No	PKActivation
Computer	44	Use of trusted computing base	Yes/No	COMPSecureBase
Network security	45	High Security firewalls	Yes/No	Firewalls
	46	Trusted time source for time stamping	Yes/No	TimeSource
Client Side Security	47	Key storage computer/smartcard/removable media	Likart scale	KeyStorage
Legal Security Controls	48	Risk apportionment in case of compromise/losses	Yes/No	RiskApportioning

3. RESEARCH METHODOLOGY

The simpler the identity
the easier the interaction

~Unknown

3.1 Overview

Research is based on some philosophical persuasions or paradigms and methodologies that are deemed suitable or relevant for the pursuit of knowledge development in a particular field. In order to successfully carry out a good research, it is important to know these persuasions, their assumptions, strengths and weaknesses so as to make informed choices based on the subject area under study (Carroll and Swatman, 2000). Knowledge of the foregoing alone is not enough because more often than not, the researcher also has to contend with other factors such as resource constraints, imperfect reasoning processes and even human failings (Simon, 2011). All these call for the researcher to honestly assess their weaknesses, and the limitations of the choices they make so as to map the best way forward during the study (Ibid).

This research idea crystallised after the researcher was exposed to the difficulty in measuring achievement of key software quality factors in software while developing mobile applications, and how to express the same to third parties. This exposure spurred the researcher towards studying how such quality factors can be measured. The choice of doing so by studying eGovernment identity and access management was arrived at based on the facts that it is an emergent area of study and it offers the right mix of opportunities and challenges for research. The research critically looks at the initiatives needed to achieve this and how to optimise security in one of them, the public key infrastructure. The research therefore has three main parts:

- the general part which looks at relevant identity and access management initiatives, and the desirable qualities of the public key infrastructure – these are mainly captured from literature survey;

- identification of security quality variables that influence security of the PKI solutions for eGovernment leading to development of the research Conceptual Framework (CF). The proposed CF is derived from existing models.
- development of the decision support model based on the conceptual framework using SmartPLS and evaluating the model through data collection, analysis and focus group discussions.

3.2 Limitations and Study Assumptions

The area of identity and access management is very wide and therefore this study cannot cover everything under it. The research concentrates on identity management in eGovernment and specifically developing a decision support model which would quantitatively assesses the security of public key infrastructure solutions. Some limitations of the study include:

1. It is difficult to access sensitive government records but the researcher was able to access as much as possible after acquiring permission from relevant ministries.
2. The difficulty of getting access to all the expert staff since most of them have very busy schedules and are often out of the office attending to important assignments.

Apart from the limitations, the study also made some assumptions:

1. That the model developed for the study captures the most important variables that influence the security of PKI solutions. There are many other variables which may not have been included in order to reduce the complexity of the model but the model is extensible as indicated.

2. Although variables have interrelationships between themselves this was not captured in the model. Instead, the model assumes that each variable affects core security attributes of accountability, availability, integrity and confidentiality individually or in tandem with others.

3.3 The Research Path

Generally, the research process is as detailed below:

3.3.1 Problem Specification

In September 2011, the researcher was involved in a mobile applications development course. As the course progressed, it became evident that there are certain software quality factors like security, availability, interoperability, modifiability, accuracy among others which are difficult to measure directly. This left a lingering impact on the researcher which was to be fully activated upon attending a conference hosted by the then Communications Authority of Kenya to discuss Kenya's Public Key Infrastructure initiative. The presenters expressed a host of identity and access management challenges and the researcher learned of a host of Government initiatives aimed at rolling out eGovernment services. One key problem that was highlighted was how to measure the quality of ICT systems in government like PKI solutions. That is how this research was born.

3.3.2 Literature Review

Pertinent literature was reviewed in order to build a conceptual framework or grounding for the research. A precise definition of terms such as attributes, credentials, authentication and identity was done. After this, an overview of identity management approaches in eGovernment and the public key initiative was identified as key to securing electronic transactions. Qualities that determine a good eGovernment public key infrastructure were isolated and the research chose to concentrate on security and how to optimise it in the same. A security optimisation conceptual framework was developed then modeled using Partial Least Squares Structural Equation Modeling (PLS-SEM) with the aim of presenting a new way of optimizing decision making when reasoning about the security of public key infrastructure solutions. The frameworks and strategies studied enabled construction of a model which guided the development of interview questionnaires, data collection, and interpretation.

3.3.3 Methods Used in the Research Process

Four methods were identified to guide the researcher in different aspects of the research:

1. The Structure Case approach
2. Culture, People, Process and Technology.
3. Survey of Kenya's PKI implementation.
4. PLS - SEM.

This implies that this research combined both the qualitative and quantitative approaches to come up with a model for optimizing rational decisions on PKI security as a key element of the identity and access management initiative for e-Government. In a nutshell, supporting objective (a), (b) and (c) were studied using purely qualitative methods while (d) quantitative ones. Research questions (i – iv) helped to lead the investigation for achieving each objective. Using the Structure Case method, the research iterates through literature until it comes up with a Conceptual Framework (CF) that captures the key attributes that influence PKI security. Although many variables are identified, the research makes an effort to select those which are most important to form an extensible CF. The importance of each variable was gauged through discussions with the experts on the ground where the research was grounded i.e. the cyber-security experts at the Kenya ICT Authority which is the premier PKI Certificate Authority in Kenya.

After the CF, interviews and discussions are carried out with relevant PKI experts to try and establish a baseline showing the current state of implementation of the PKI initiative in Kenya. The research targeted experts at the Communications Authority, ICT Authority and KRA who worked to plan and implement Kenya's PKI solution in their relevant agencies. This enabled the research to formulate relevant questionnaire questions.

After the baseline data was collected, the CF was directly mapped into a Partial Least Square Structural Equation Model (PLS-SEM) in SmartPLS as shown in Figure 3.5. PLS-SEM was chosen because the research noticed that the tool has capability to directly map the CF and its causal relationships into the model. It also has various algorithms for statistical analysis of data. The data collected in the baseline survey was used to populate the PLS-SEM model and to generate various measures of the current state of PKI implementation. To this extent, we can say that the research was seeking to find out whether the model can effectively capture the relevant data and compute the various statistical measures as is the aim of objective (c).

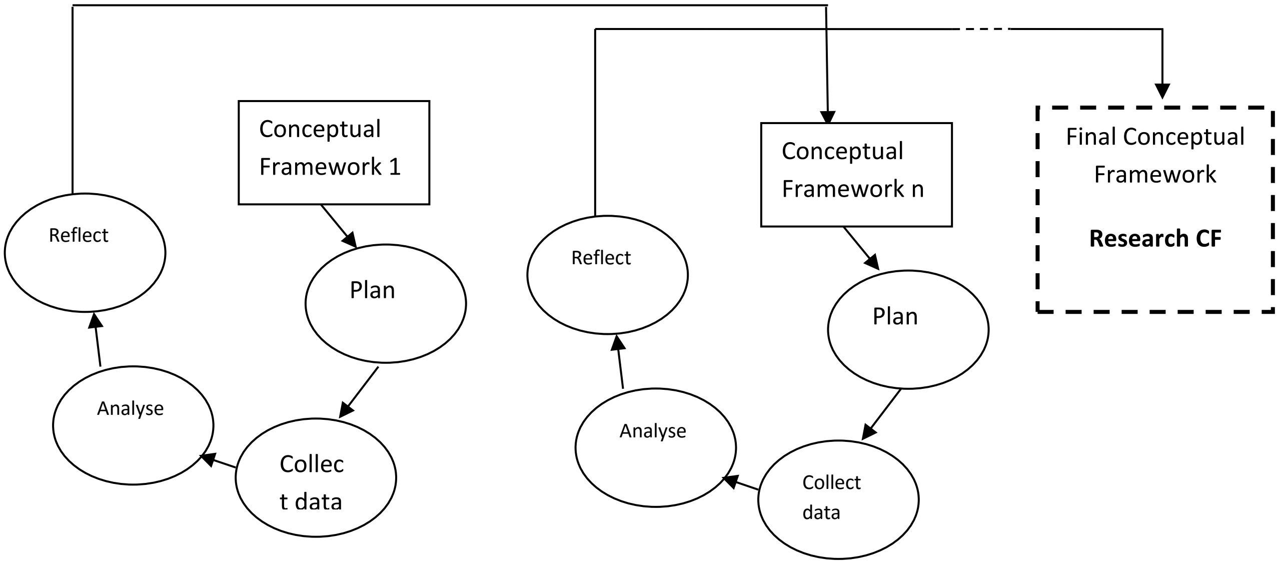
The statistical measures generated by the baseline study were shared with the relevant PKI staffs who were participating in the study. The aim was to have them note areas that seemed to fall below the expected standards, identifying the reasons and adjusting the system or processes appropriately. After three months, more data was collected using interviews guided by a questionnaire method and used to populate the model. The new measures were compared with the baseline measures to find out whether improvements in indicators had been achieved or not.

3.3.3.1 Structure Case Framework and its Link to Others

Starting with an original rough concept as captured in the conceptual framework, the study iterates through the planning, data collection, analysis and reflection stages until a clear solution or framework is developed (Carroll and Swatman, 2000). The plan phase includes the research paradigm, concepts and their relationships in the conceptual framework, determination of research design, methods of data analysis determined. The collect data phase involved collection and recording of data. In this case data was collected from literature and in-depth interviews with experts. The analyse phase qualitatively analysed the data collected. Finally, the reflect phase involved introspection and reflection to confirm concepts or come up with new ones.

The initial Conceptual Framework 1 was the researcher's initial understanding and is used to lay down the research scope and guides the first research cycle. There is a continuous iteration through new knowledge until a final CF is reached.

Figure 3.1: Research methodology using Structure Case framework and PLS-SEM



Source: Carrol and Swatman, 2000

Data collection mainly involved searching online databases for relevant work e.g.

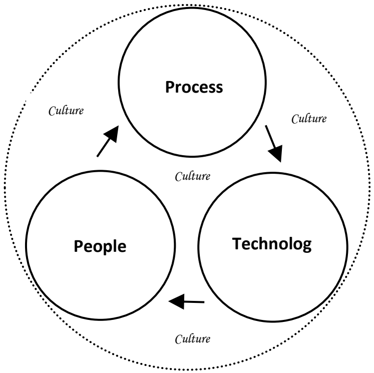
- Springer Link (http://springerlink.com)
- Google Scholar (http://scholar.google.com)
- IEEE Explore(http://ieeexplore.ieee.org)
- ACM Digital Library (http://dl.acm.org)

The databases were selected because they are known sources of high quality ICT publications. However, publications of experiences from avatar countries were found using generic search.

3.3.3.2 Culture, People, Process and Technology Strategy (CPPTS)

The importance of People, Process and Technology (PPT) approach, when seeking to optimise initiatives that would lead to the improvement of people's capacity, process efficiency and technological effectiveness in any organisation cannot be overstated (Valdez et al, 2008). A holistic approach to process improvement as espoused in Aristotle's philosophy which states that the whole is more important than the sum of its parts, rather than a granular one which analyses elements such as economic, process, technology, and people among others dependently, is better. Culture is the context within which interactions occur as shown in Figure 3.2.

Figure 3.2: CPPT Model



Source: Research

3.3.3.3 Partial Least Squares Structural Equation Modeling (PLS - SEM)

PLS modeling is a multivariate structural equation modeling analysis methodology initially developed by Wold (1972,1982), Lohmoller (1989) and Ringle et al. (2005) as cited in Ringle et al., (2012). It initially found popularity in marketing research due to its ability to model linear and additive causal models but has since become popular among researchers from all fields as long as they are working with latent variables which have causal relationships, some independent and others dependent (Wong, 2013; Monecke and Leisch, 2012). The SmartPLS tool is particularly liked for its friendly user interface and ease of use. This tool was chosen for this research due to its ability to identify quality factors that are difficult to measure directly (latent variables) like interoperability, security, performance, reliability among others of the target systems as dependent variables, modeling the causal relationships with their attributes and subjecting them to analysis based on data collected resulting in quantitative measurements of the attributes. This makes sure that diagrammatic descriptions of systems and their environments are formally modeled and quantitatively assessed (ibid).

SEM achieves this through the following steps:

1. Based on the conceptual model, an inner model made of latent variables is constructed in the PLS-SEM tool similar to Figure 3.4.

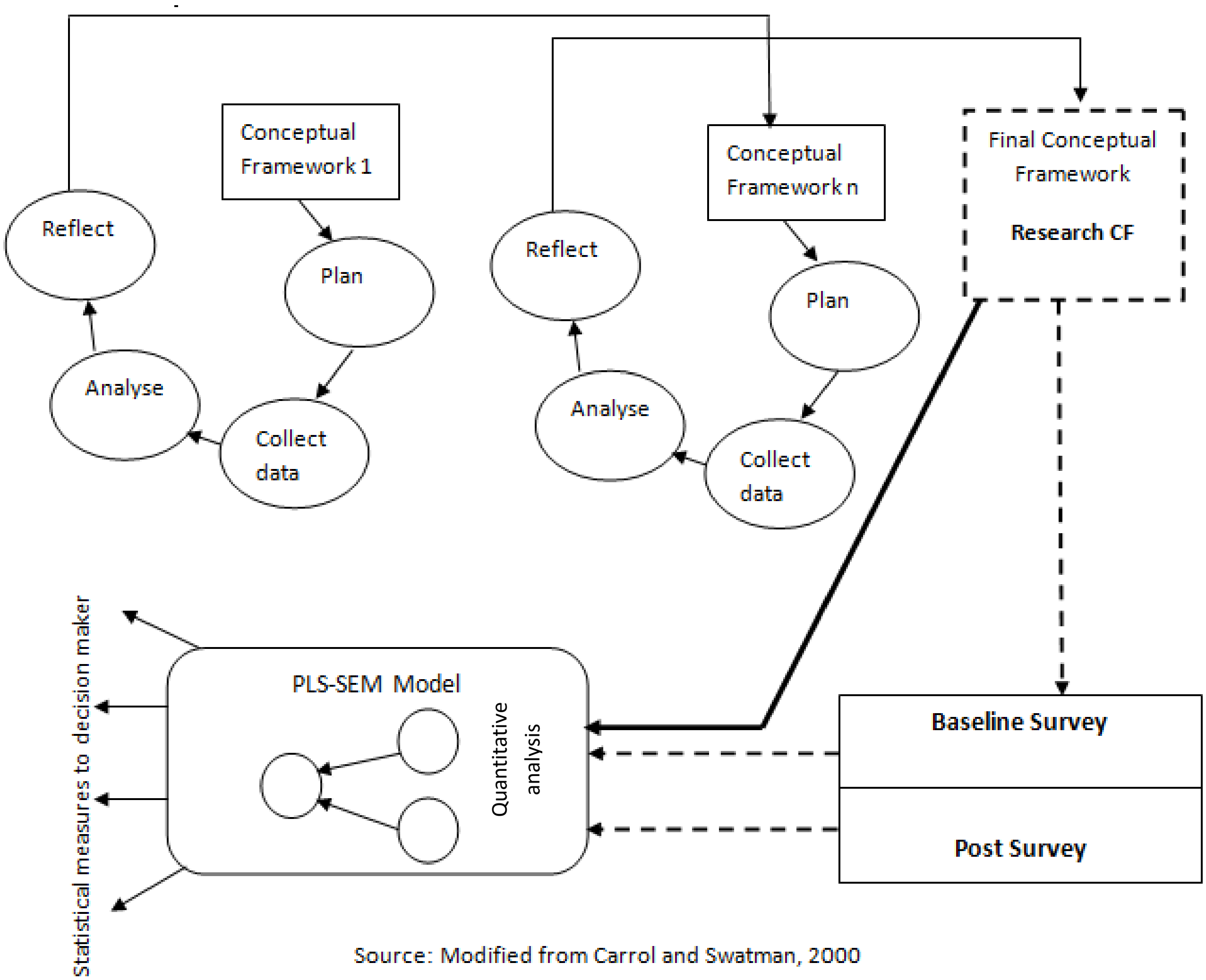
2. Evidence collection: Data (evidence) is collected to initialize the outer model made of variable indicators that directly influence the independent latent variables.
3. Analysis: In this step, quantitative values of the quality attributes are calculated. Quantitative values are visualized on the model links but other advanced reports are generated too.

3.4 Extended Research Methodology Model

The final CF is directly translated into a PLS-SEM model in SmartPLS which is able to capture directly all the causal relationships. A baseline survey is done to get the initial data which is fed into the model. The statistical measures obtained formed the basis of discussion with various stakeholders and experts. After three months, a post survey was done. The data collected populated the model again, and after analysis the initial measures and latest ones were compared. The statistical measures at each stage are shown to decision makers to influence relevant decisions as captured in the decision optimisation model.

In the extended model, analysis happens at two levels. The first happens during the iterative phase of the research, during literature review. The aim of this analysis is to discover the PKI security quality attributes from literature. In the final box labelled PLS-SEM Model, quantitative analysis of the collected data is done in order to generate the statistical measures of the PKI security quality attributes. Figure 3.3 shows the extended model.

Figure 3.3: Extended Model



Source: Modified from Carrol and Swatman, 2000

3.5 Data Collection Strategies

A baseline study was be conducted through discussions with experts. The data collected during this study was used to initialize the SmartPLS model and give initial quantitative measurements of the level of achievement of various security attributes in Kenya’s PKI solution. After this, it was expected that managers used the measures generated to improve areas that seemed to lag behind. Data was collected once every month for three months to assess whether the model impacted positively the decision making process or not. This was evident based on whether areas that seemed to lag behind were continuously corrected over the time that the model was in use.

Table 3.1: Data Collection Strategies

Research method	Objective	Data collection method
1. Structure case **2. Culture, People, Process, Technology** **3. PLS-SEM in SmartPLS**	To develop a decision optimisation model for optimizing public key infrastructure security based on quantitative assessments of the security quality attributes.	Literature review PLS-SEM
3. Structure case **4. Culture, People, Process, Technology**	a. To identify key PKI security quality attributes that need to be measured and optimised in order to have a secure public key infrastructure solution for eGovernment. b. To utilize the attributes identified in (a) in proposing a PKI security rational decision optimisation conceptual framework. c. Develop a PKI security rational decision optimisation tool from the CF in (b) and	Literature review and one on one informal discussion with industry experts.
5.PLS-SEM in SmartPLS	d. Evaluate the level of effectiveness of the proposed tool.	Surveys, Results from SmartPLS model, Focus group.

Source: Research;

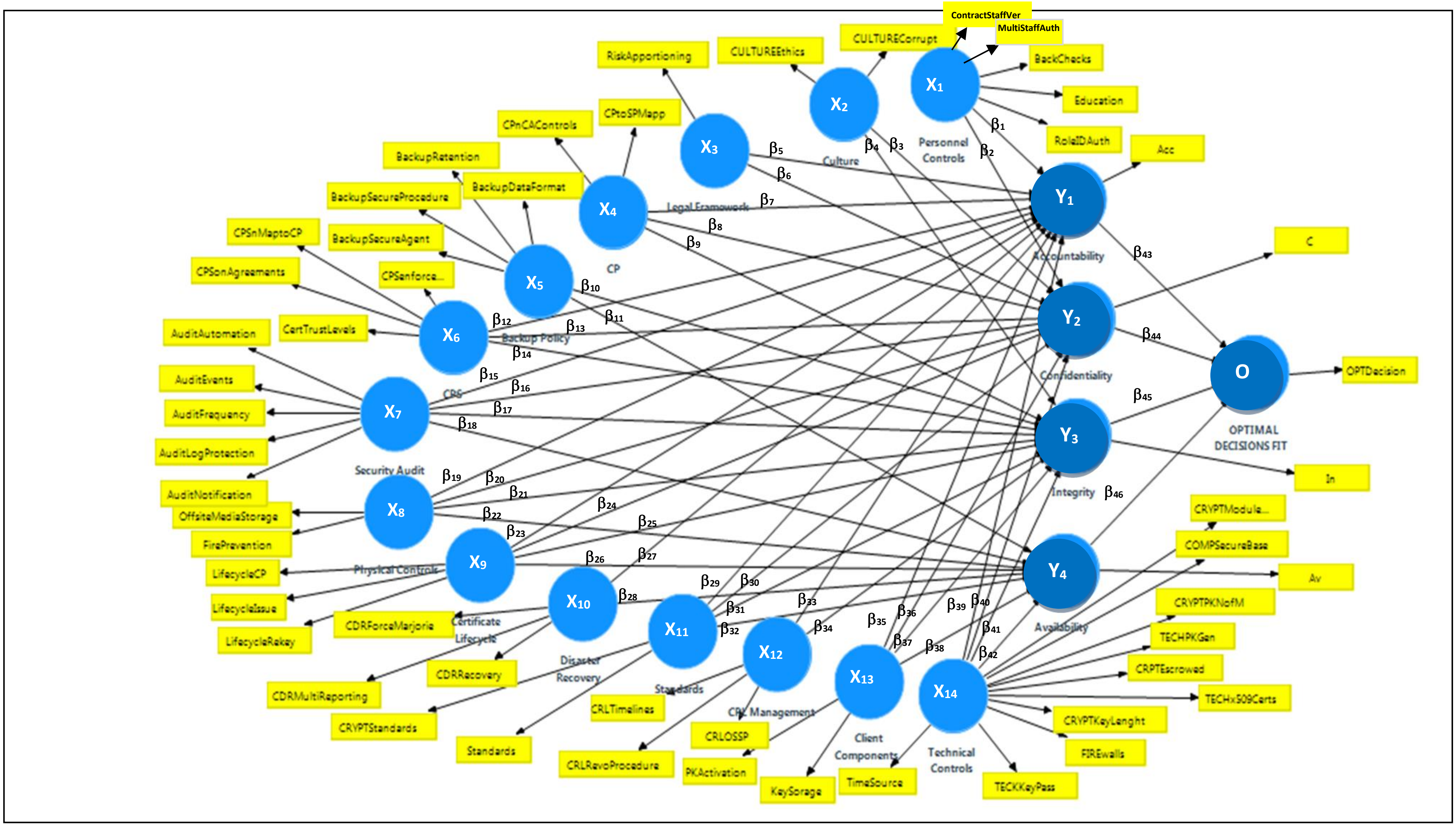
3.6 Modeling

The conceptual framework Figure 2.23 was converted into a PKI security decision optimisation tool in SmartPLS as shown in Figure 3.4. The tool has fourteen exogenous variables labelled X_1 to X_{14} mapping directly to the quality properties identified in the conceptual framework. Each of the exogenous variables maps directly to the influencing factors identified in the conceptual framework and given a label as shown in Table 2.9 e.g. the quality property "Personnel Controls" is labelled X_1 in the SmartPLS tool. This property has an influencing variable "Education/Training" captured as "Relevant Education/Training" in Table 2.9. This variable is labelled "Education" as an indicator on the "Personnel Controls" exogenous variable in the

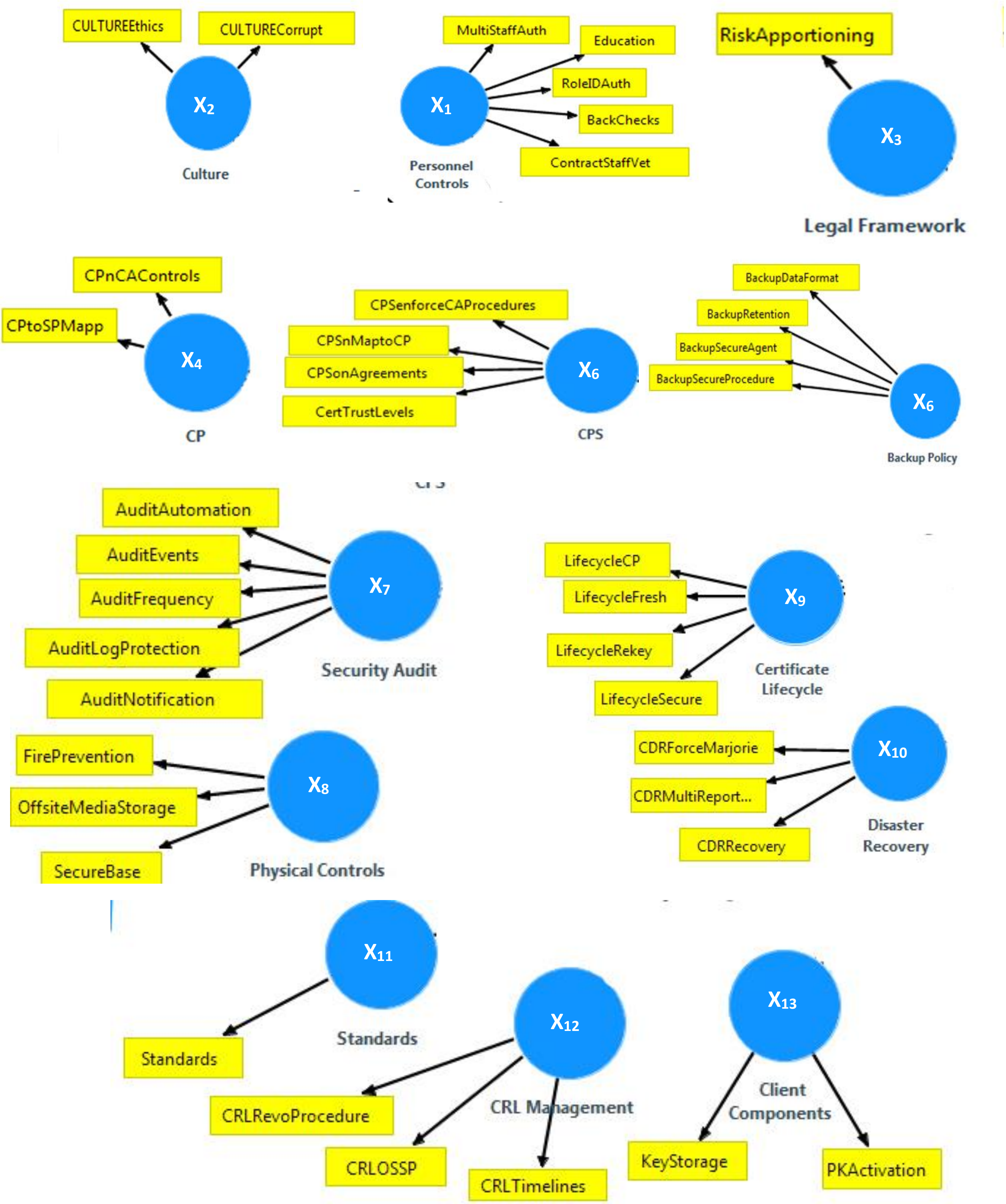
SmartPLS tool. All the other exogenous variables and indicators can be traced back to the conceptual framework and Table 2.9 in this manner.

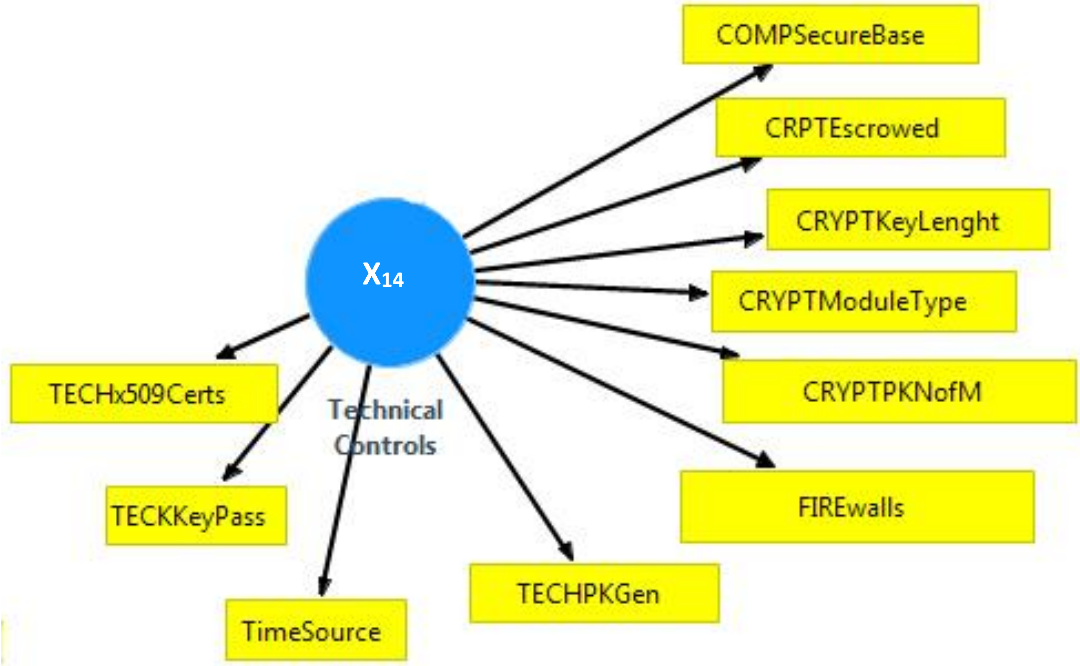
3.7 The model in SmartPLS

Figure 3.4: SmartPLS Structuctural Equation Model



A close up view of the latent variables and their indicators follows complete with their labels and indicators:

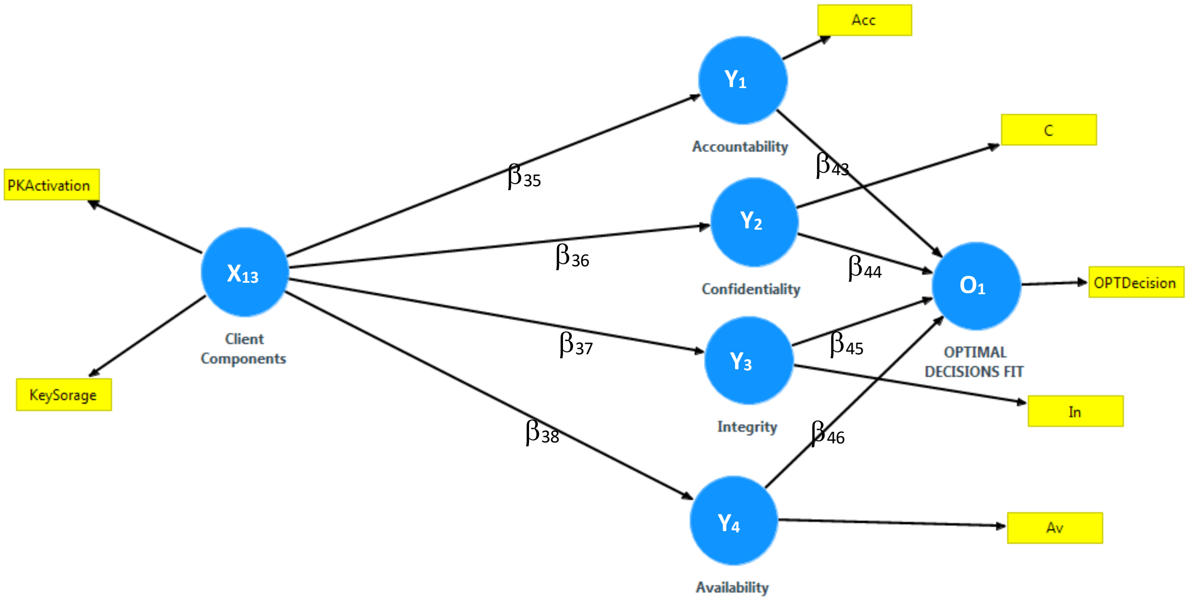




3.8 The Measurement and Structural Model

The measurement model (outer model) consists of all the indicators pertinent to each exogenous variable. Each variable has its own block of indicators. A variable must have at least one indicator. For demonstration purposes, notice that an exogenous variable like Client Components (which refers to controls affecting digital keys at the users end) has two indicators: *KeyStorage* (assessing whether digital keys are stored on a smartcard, in a browser or on the computer) and *PKActivation* (which assesses the level of security attached to the process of activating a private key after being issued). The structural (inner) model consists of the exogenous variables and their relationships with the endogenous ones. It describes the causes and consequences of the variables on the integrity, accountability, confidentiality and availability of the PKI system. These then have effects on the final decision fit. Figure 3.5 below shows an extract from the SmartPLS model:

Figure 3.5: Model Extract



Source: research

From this extracted diagram we can create a 6x6 adjacency matrix **D** as shown in Table 3.2. In essence, a table for the whole model can be constructed using the same rules. In our case such a table for Figure 3.2 would be a 19 x 19 matrix.

Table 3.2: Adjacency Matrix for Model Extract of Figure 3.4

	X_{13}	Y_1	Y_2	Y_3	Y_4	O_1
X_{13}	1	1	1	1	1	0
Y_1	1	1	0	0	0	1
Y_2	1	0	1	0	0	1
Y_3	1	0	0	1	0	1
Y_4	1	0	0	0	1	1
O_1	0	1	1	1	1	1

If d_{ij} is equal to 1 then exogenous variable i is a predecessor of endogenous variable j

Source: Research

From our earlier equations on regression, it is not difficult to represent the relationships between the various variables in the inner model using matrices as follows:

$$Y_i = \beta_0 + \beta_i X_i + \varepsilon \qquad (4)$$

Where:

Y_i denotes the matrix for the various variables endogenous variables
β_0 is the baseline
β_i is coefficient matrix whose values are set to 0 where those of adjacency matrix X are 0
X_i denotes the matrix for various exogenous variables
ε are error terms which are centered i.e. set to 0 with maximun variance of 1.

Therefore the endogenous variables in our model of Figure 3.4 have the following equations depicting their relationships with the exogenous variables:

$$Y_1 = B_{ACC} + \beta_1 X_1 + \beta_5 X_3 + \beta_7 X_4 + \beta_{12} X_6 + \beta_{15} X_7 + \beta_{19} X_8 + \beta_{23} X_9 + \beta_{27} X_{10} + \beta_{29} X_{11} + \beta_{33} X_{12} + \beta_{35} X_{13} + \beta_{39} X_{14} + \varepsilon_{15} \ldots\ldots\ldots (5)$$

$$Y_2 = B_C + \beta_1 X_1 + \beta_3 X_2 + \beta_6 X_3 + \beta_8 X_4 + \beta_{13} X_6 + \beta_{16} X_7 + \beta_{20} X_8 + \beta_{24} X_9 + \beta_{30} X_{11} + \beta_{36} X_{13} + \beta_{40} X_{14} + \varepsilon_{16} \ldots\ldots\ldots (6)$$

$$Y_3 = B_I + \beta_4 X_2 + \beta_9 X_4 + \beta_{10} X_5 + \beta_{14} X_6 + \beta_{17} X_7 + \beta_{21} X_8 + \beta_{25} X_9 + \beta_{31} X_{11} + \beta_{34} X_{12} + \beta_{37} X_{13} + \beta_{41} X_{14} + \varepsilon_{17} \ldots\ldots\ldots (7)$$

$$Y_4 = B_{AV} + \beta_{11} X_5 + \beta_{18} X_7 + \beta_{22} X_8 + \beta_{26} X_9 + \beta_{28} X_{10} + \beta_{32} X_{11} + \beta_{38} X_{13} + + \beta_{42} X_{14} + \varepsilon_{18} \ldots\ldots\ldots (8)$$

$$O_1 = B_{ODF} + \beta_{43} X_{15} + \beta_{44} X_{16} + \beta_{45} X_{17} + \beta_{46} X_{18} + \varepsilon_{19} \ldots\ldots\ldots (9)$$

Equations (5) to (9) describe the object functions of the model. The statistical measures generated by PLS algorithms on the data based on these functions are compared to expected values as indicated in Table 3.3 that indicate optimum states hence decision makers can easily know if security in the PKI solution is optimized or not.

3.9 Data Analysis Strategies

The data collected ranged in value from 1 – 7 on interval likart scale. In cases where the options are Yes/No then only two values, one (1) or (7) were collected. In cases where the answer was “Don’t know” in a Yes/No situation, the value 1 was entered. A one (1) means the worst case scenario in which the reality gap is big, best practice was not followed and decisions are not optimised. A seven (7) means best practice and outcome are experienced, that reality gaps are small due to well-fitting design solutions after best decisions were made.

Intermediary values like 2, 3, 4, 5 and 6 when using a likart scale mean possibilities of partial failures or successes unless relevant interventions are made. For each likart scale, the possible responses are arranged from worst case to best case hence providing an interval scale (Wong, 2013). To measure whether a particular variable is optimized or not depends on the Analysis of Variance (ANOVA) on the data received for each attribute. A mean of 1 means that that attribute is not optimized at all while a mean of 7 with little variance means the attribute is optimised. Although the metrics are not the same, we seek to quantitatively show the level of achievement of each variable in the PKI solution. Those that are optimized have low variance to indicate consistency and their means as close to 7 as possible. Table 3.3 gives a summary of expected optimal values for various measurements.

Table 3.3: Optimal Values for Various Statistical Measures

MEASURE NAME	SYMBOL	IMPORTANCE	ALGORITHM/CONDITIONS	OPTIMAL VALUES	Source(s)
T- Test Inner model	t-Value	Tests significance of relationships. It assesses the statistical significance of path coefficients and indicator loadings. Answers two questions: Does the relation exist? If so, how strong is the relationship?	**Bootstrapping.** 1300 subsamples are drawn randomly from the observed sample size with replacement. This is possible because the bootstrapping algorithm performs sampling with replacement i.e. although desirable the sample sets don't have to be made up of different elements each time. The ideal is to have 5000 subsamples but it is infeasible to get such a large number of unique samples in our case because we have a small sample size (34 observations for each variable). We then compare the t-Values to critical values from the standard normal distribution. This determines whether they are significantly different from zero.	The larger the better. Indicator loadings should be positive.	(Wong, 2013) (Hair et al., 2013) (Ringle et al, 2014)
P-Test Inner model	p-Value	In combination with the t-Value they indicate which relationship is most significant.	**Bootstrapping**. p-Value here used to test the significance predicted by the t-Values.	IF • t>1.96 p<=0.05 • t > 2.576 p<=0.01 • t>3.29 p <= 0.001 among others.	(Wong, 2013) (Hair et al., 2013)

Indicator reliability Outer model	Outer Loadings $\varnothing$	Indicator reliability. Square each of the outer loadings.	**PLS Algorithm** Find $\varnothing^2$	Values => 0.7	Hulland (1999) as cited in (Wong (2013)
Effect Size	f^2	It measures change in R^2 when a variable is omitted	**PLS Algorithm** Find f^2	0.02 => small effect 0.15 => medium 0.35 => large	(Hair et al., 2013)
Internal Consistency Reliability		Measures the degree to which the data is consistent. High variance indicates less consistency. An internal consistency test should be done before any PLS analysis of data is done.	**PLS Algorithm**	Values => 0.7	(Wong, 2013) (Hair et al., 2013)
Average Variance Extracted	AVE	Measures the validity of the latent variables in the model	**PLS Algorithm**	AVE > 0.5	(Wong, 2013) (Hair et al., 2013)
Coefficient of Determination	R^2	Explains the % of variance that exogenous/intermediary variables have on their final target(s).	**PLS Algorithm**	Scale → 0.0 – 1.0; the higher the better. R^2 => 0.7	(Wong, 2013) (Hair et al., 2013)
Path Coefficients	P_c	The path coefficients indicate to us the correlations between the exogenous and the endogenous variables	**PLS Algorithm**	Scale **1**>= **P_c**>= **-1** Also report t>1.96 and p<0.05 for bootstrapping algorithm.	(Hair et al., 2013)

Sources: As stated in the Table.

4. RESULTS

4.1 Data Analysis

The main objective of the study is to quantitatively measure the key PKI security quality attributes in order to help optimise the decisions of managers when reasoning about PKI security. The data analysis of the study was done at two levels:

i. Analysis of the results of the structural model: This looked at the results of the relationships between the latent variables. This covered looking at the path coefficients and coefficient of determination (R^2).

ii. Analysis of the measurement model. This was done by studying the indicators, their validity and reliability.

The analysis was done using bootstrapping and partial least squares structural equation modeling in SmartPLS. The model has 19 latent variables; fourteen (14) exogenous and five (5) endogenous ones. The final target one variable named Optimal_Decision_Fit represents the state of optimisation of security decisions based on the four core PKI security variables of accountability, confidentiality, integrity and availability.

The outer model has forty nine (49) indicators which were all linked to the questionnaire questions and were populated with data from both the baseline and post survey to form the basis of statistical analysis.

4.2 Baseline and Post Study

The study covered the months of January – July 2015. The study targeted cyber security experts, public key infrastructure implementers and identity management experts in the Ministry of Information, State Department of Interior, and Ministry of Devolution and planning. A total of 30 respondents participated in the pre-study and 34 in the post study. A pre-study was carried out first to set the baseline. A questionnaire was developed to collect data in the post study.

Due to the small number of technical experts in the cyber-security departments that participated in the research (one example is that by the time of the research Kenya ICT Authority cyber-security center had only five (5) employees), there was no need to do sampling or respondents but rather all of them were engaged by the researcher. The same questionnaire was used to guide the interviews in all places because the respondents were all technical staff who had hands on experience implementing the PKI system in their respective places.

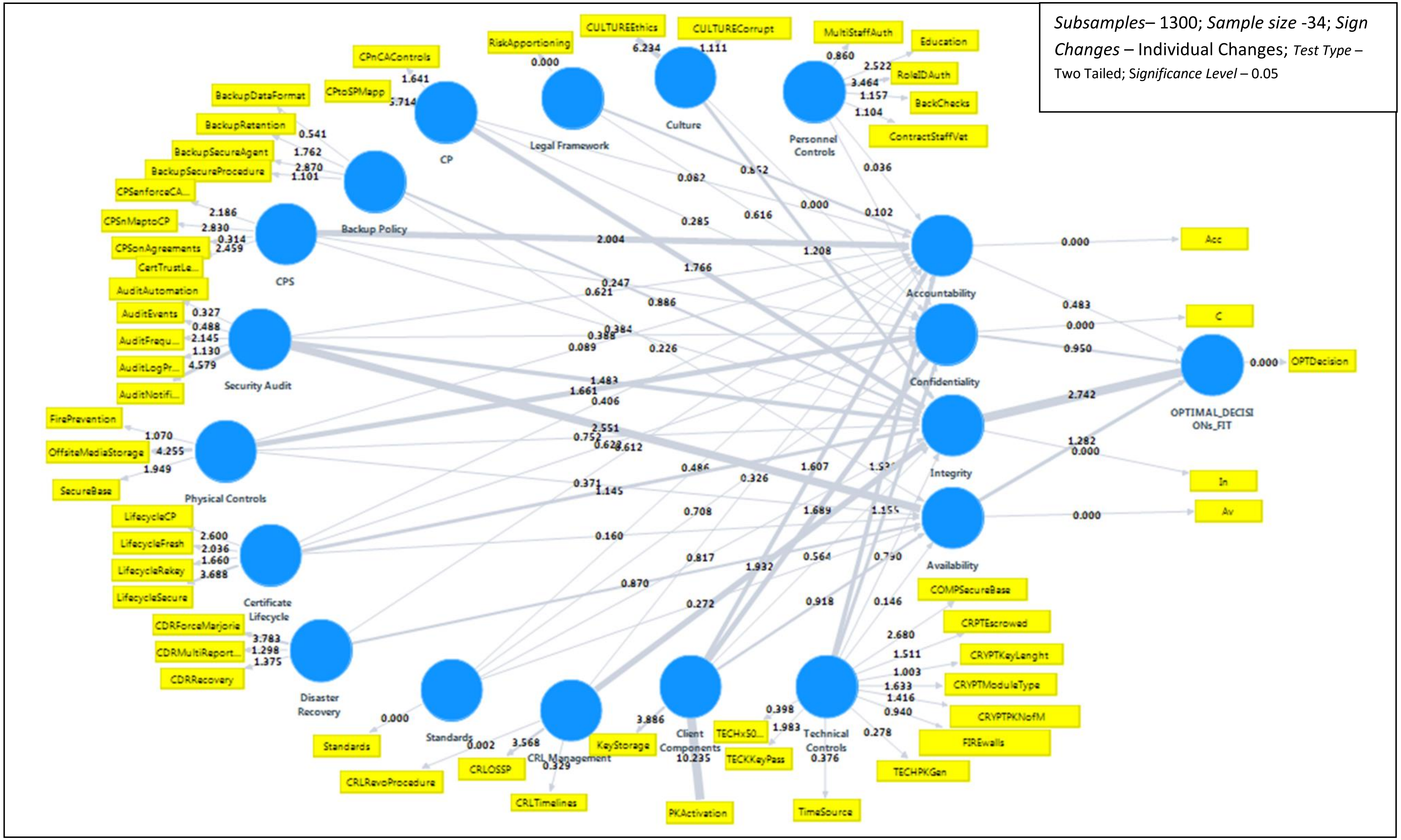
4.3 Bootstrapping Algorithm

4.3.1 Indicator Outer Weights – Bootstrapping with Replacement

Figures 4.1 (A) and (B) show the relative significance of each of the exogenous variables on their respective endogenous variables as analysed for the pre-study and post study respectively. For example, notice that Integrity has the greatest significance on OPTIMAL_DECISION_FIT. Similarly, Certificate Lifecycle and Disaster Recovery seem to be quite significant when reasoning about availability in Figure 4.1 (B). Table 4.1 tabulates indicator outer weights in terms of their t-Values and respective p-Values.

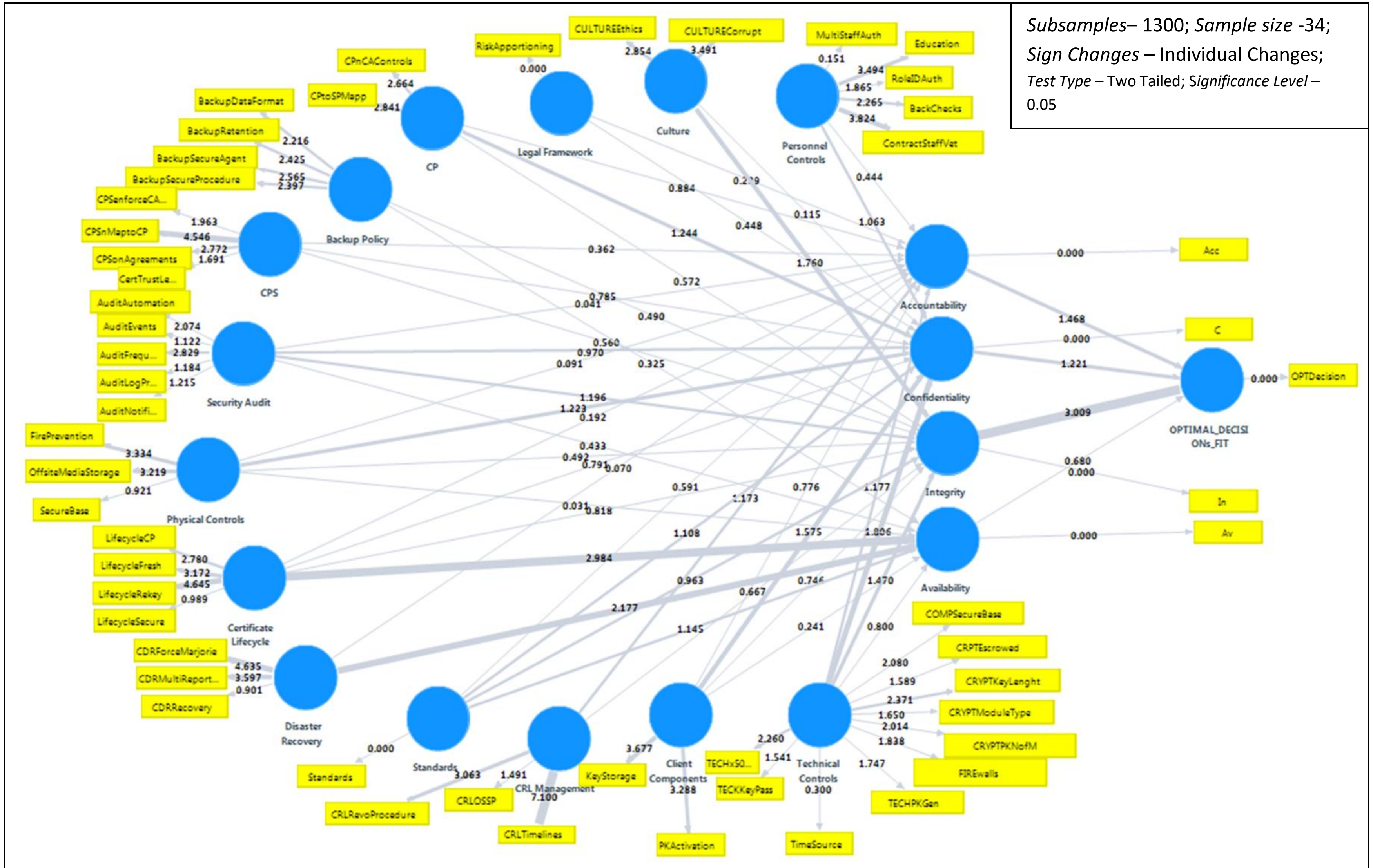
The indicator (manifest variables) outer weights are important because they show their significance to the exogenous variable that they are attached to in the outer model. It is important to have as high a t-value as possible and as low a p-value as possible for each indicator. However, p-values depend on the magnitude of the t-value as detailed in Table 3.2. Every time the Bootstrap algorithm runs, it produces slightly different t-values due to the random sampling error. However, this is acceptable because it does not significantly affect the final results.

Figure 4.1:(A) BASELINE STUDY: Path Coefficients Bootstrapping Algorithm



Subsamples– 1300; *Sample size* -34; *Sign Changes* – Individual Changes; *Test Type* – Two Tailed; *Significance Level* – 0.05

Figure 4.1:(B) POST STUDY: Path Coefficients Bootstrapping Algorithm



Subsamples– 1300; *Sample size* -34;
Sign Changes – Individual Changes;
Test Type – Two Tailed; *Significance Level* – 0.05

Source: Research model in SmartPLS

Table 4.1: BASELINE and POST TEST Outer Weights t-Values

Outer Weights Values								COMMENTS
	BASELINE				POST-TEST			
Indicator	**t-Value**	**p-Value**			**t-**	**p-**		
AuditAutomation<-Security Audit	0.060	0.952	O		2.151	0.032	S	Improved. Good. Adheres to global
AuditEvents <-Security Audit	1.520	0.129	O		0.720	0.472	O	Some events (DoS) not logged*
AuditFrequency<-Security Audit	2.150	0.032	S		2.776	0.006	S	Improved. Good. Adheres to global
AuditLogProtection<-Security Audit	1.396	0.163	O		0.445	0.656	O	Always a big challenge here*
Auditnotification<-Security Audit	4.735	0.000	S		1.024	0.306	O	Not notified unreliable baseline data*
BackChecks<-Personel Controls	0.660	0.509	O		1.398	0.162	O	Not enough checks*
BackupDataFormat<-Backup Policy	0.154	0.878	O		2.184	0.029	S	Improved. Baseln data was incomplete.
BackupRetention<-Backip Policy	1.766	0.078	O		2.335	0.020	S	Sufficient periods.
BackupSecureAgent<-Backup Policy	3.350	0.001	S		2.031	0.042	S	Good.
BackupSecureProcedure<-Backup Policy	1.487	0.137	O		2.445	0.015	S	Baseline data scanty
CDRForceMarjorie<-Disaster Recovery	3.564	0.000	S		3.166	0.002	S	Good
CDRMultiReporting<-Disaster Recovery	1.070	0.285	O		2.637	0.008	S	Noted need to increase acceptable
CDRRecovery<-Disaster Recovery	1.183	0.237	O		0.335	0.738	O	More effort needed here*
COMPSecureBase<-Technical Controls	3.262	0.001	S		2.038	0.042	S	Good
CPSenforceCAProcedures<-CPS	2.316	0.021	S		1.727	0.084	O	None. A Certificate Policy needed.
CPSnMapptoCP<-CP	3.116	0.002	S		3.931	0.000	S	Good
CPSonAgreements<-CPS	0.846	0.398	O		2.010	0.045	S	Good.
CPnCAControls<-CP	0.946	0.344	O		1.883	0.060	O	Needs to be enforced.
CPtoSPMapp<-CP	4.618	0.000	S		2.756	0.006	S	Good
CRLOSSP<-CRL Management	3.868	0.000	S		0.840	0.401	O	OSSP found with no stapling*
CRLRevoProcedure<-CRL Management	0.313	0.754	O		2.832	0.005	S	Good. Initial data not available
CRLTimeLines<-CRL Management	1.134	0.257	O		3.941	0.000	S	Good. Initially data not available.
CRPTEscrowed<-Technical controls	1.396	0.163	O		2.000	0.005	S	Good. Initially data not available.
CRYPTKeylength<-Technical Controls	2.605	0.009	S		2.570	0.010	S	Good but length needs improvement
CRPTModuleType<-Technical Controls	2.302	0.021	S		2.402	0.020	S	Good.
CRPTPKNofM<-Technical Controls	2.170	0.030	S		2.394	0.017	S	Good.
CULTURECorrupt<-Culture	0.625	0.532	O		2.652	0.008	S	Good. Initially data not available.
CULTUREEthics<-Culture	4.858	0.000	S		2.081	0.038	S	Good.
CertTrustLevels<-CPS	2.405	0.016	S		1.136	0.256	O	Not yet implemented but in plans
ContractStaffVet<-Personnel Controls	1.277	0.202	O		2.961	0.003	S	Good.
Education<-Personnel Controls	2.147	0.032	S		2.220	0.027	S	Good.
FIREwalls<-Technical Controls	2.292	0.022	S		1.169	0.243	O	DoS bypassed firewall - undetected*
FirePrevention<-Physical Controls	1.087	0.277	O		2.732	0.006	S	Good.
KeyStorage<-Client Components	2.559	0.011	S		3.054	0.002	S	Good
LifecycleCP<-Certificate Lifecycle	1.552	0.121	O		2.346	0.019	S	Good
Lifecyclefresh<-Certificate Lifecycle	0.839	0.402	O		2.342	0.019	S	Good
LifecycleRekey<-Certificate Lifecycle	1.205	0.229	O		4.046	0.000	S	Good
LifecycleSecure<-Certificate Lifecycle	3.103	0.002	S		1.393	0.164	O	Initial data was not reliable
MultiStaffAuth<-Personnel Controls	0.105	0.917	O		0.387	0.699	O	Bad. Few personnel. Role overlaps
OffsiteMediaStorage<-Physical Controls	4.495	0.000	S		3.097	0.002	S	Good
PKActivation<-Client Components	5.464	0.000	S		2.395	0.017	S	Good
RoleIDAuth<-Personnel Controls	3.181	0.002	S		0.740	0.459	O	Bad. Few personnel. Role overlaps
SecureBase<-Physical Controls	1.909	0.056	O		0.781	0.435	O	Site not initially securely built
TECHPKGen<-Technical Controls	0.945	0.345	O		0.921	0.357	O	256bit key length not secure enough*
TECHx509Certs<-Technical Controls	0.910	0.363	O		2.369	0.018	S	Good. Initial data incomplete.
TECHKeyPass<-Technical Controls	3.500	0.000	S		1.434	0.152	O	Need to improve
TimeSource<-Technical controls	1.255	0.210	O		0.232	0.816	O	Unavailable. Need to be created.

Source: Research

The results in this section show clearly that the proposed model was able to capture relevant data about the state of each of the attributes. In the baseline study, some of the information received was incomplete leading to some variables getting a low score. However, after the information was provided in the post study, it also reflects. The model discovered some deficiencies as indicated by red or blue rows in the table. We discuss some of them below:

1. AuditEvents <-Security Audit: During one of the site visits, a non-logged denial of service attack was recorded by the research. This means that the ability of the system to log some events needed to be improved.
2. AuditLogProtection<-Security Audit: Protection of the audit log was noted as a key issue especially from insiders. The question of protection of the same in case of powerful external influences needs to be carefullt considered.
3. Auditnotification<-Security Audit: the notification of offenders on the system should be immediate. Again the low score here reflects the DoS attack and the inability to trace the source.
4. BackChecks<-Personel Controls: Based on the questionnaire and responses, we believe one reason for this outcome is because we targeted employees instad of employers to answer this question. Most acknowledged their certificates and Ids were vetted but references and background security checks were not selected. Therefore we conclude that the data collected for this attribute was unreliable.
5. CDRRecovery<-Disaster Recovery: The low score here emanates from the fact that the only Force Marjeure threat well prepared for was fire outbreak. Others like earthquakes, flood, solar flares among others scored no hit.
6. CPSenforceCAProcedures<-CPS: The low score here emanates from the general lack of a Certificate Policy (CP) document which normally forms the basis for CPS documents.

7. CPnCAControls<-CP: There is no certificate policy. This document should be created by the regulatory agency (Communications Authority) to enable Certificate Authorities develop standard Certificate Practice Statements across board that foster certificate security, interoperability among others.
8. CRLOSSP<-CRL Management: Lack of stapling leads to inefficient PKI as relying parties have to keep on asking for certificate verification from CAs before a transaction.
9. TECHPKGen<-Technical Controls: RSA 256 bit key is not secure.
10. TimeSource<-Technical controls: Lack of a secure time stamp server for a PKI spells disaster. It is impossible to enforce non-repudiation where transaction time can be manipulated.

Due to the sensitivity of the security of PKI systems, the researcher was not allowed to interact with the system directly but through proxy i.e. a technical person who is knowledgeable and has hands on access to the system. For example, after discussing a particular shortcoming visualised using the decision optimisation tool during the baseline study e.g. looking at Table 4.1 the indicator CDRMultiReporting had a very low score with a p-value way above 0.05 i.e 0.285. This had been caused by there being only two methods of reporting a key or certificate compromise to the certification authority i.e. telephone and signed email. However, after discussion about the importance of having many more ways of allowing users to report such cases, the contact person initited a process where users could also report through social media and this improved the score in the post study to a p-value of 0.008 hence the change from red to green in the table.

4.3.2 R Square Values – Bootstrapping with Replacement

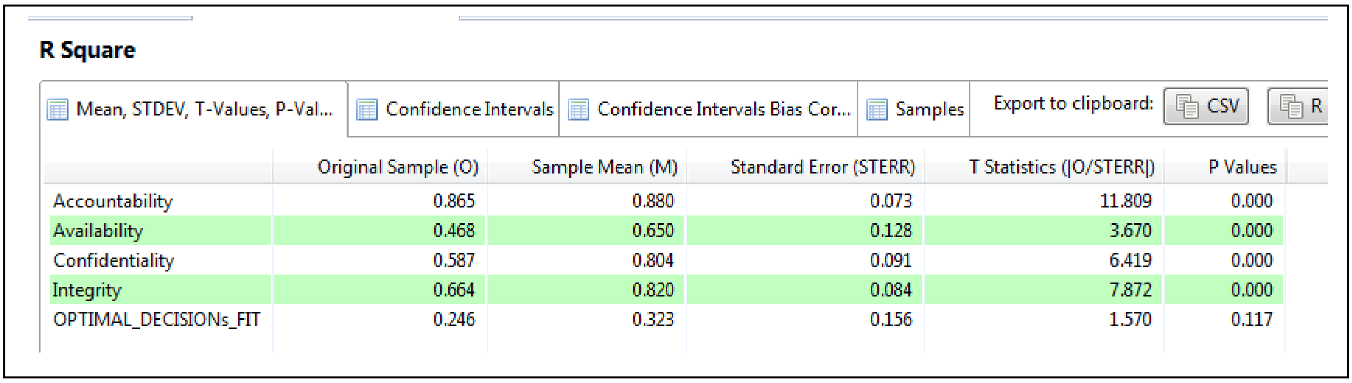
Table 4.2 (A) and (B) shows R square values for endogenous variables. There is a marked improvement between the time when the baseline was taken and three months later when a second set of data was collected and analysed. For example:

- In Table 4.2 (A) all the T values except for Accountability are lower than in Table 4.2 (B)
- The p-values for Availability (0.047) and Confidentiality (0.001) in Table 4.2 (A) all improve to become 0.000 in Table 4.2 (B).
- The value for OPTIMAL_DECISION_FIT in Table 4.2 (A) (0.047) improves to become 0.044 in Table 4.2 (B).

This indicates that the proposed method of assessing and optimizing various security quality attributes is viable and can be applied as suggested in this thesis. The improvement in the t-values and p-values after the baseline can be explained to mean some of the weaknesses identified were improved upon hence yielding better results in the successive assessments.

Table 4.2 (A) and (B): Bootstrapping Results – R^2 Values for Endogenous Variables and t-Values

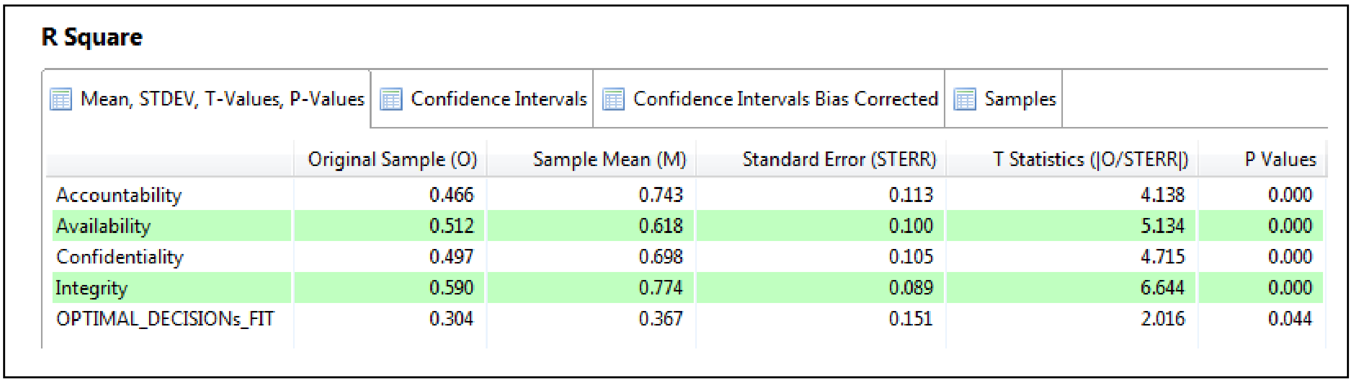
Table 4.2 (A) Baseline R Square Values - Bootsrapping


Mean, STDEV, T-Values, P-Val...
Confidence Intervals
Confidence Intervals Bias Cor...
Samples
Export to clipboard:
CSV
R

R Square

	Original Sample (O)	Sample Mean (M)	Standard Error (STERR)	T Statistics (|O/STERR|)	P Values
Accountability	0.865	0.880	0.073	11.809	0.000
Availability	0.468	0.650	0.128	3.670	0.000
Confidentiality	0.587	0.804	0.091	6.419	0.000
Integrity	0.664	0.820	0.084	7.872	0.000
OPTIMAL_DECISIONs_FIT	0.246	0.323	0.156	1.570	0.117

Table 4.2 (B): POST TEST R Square Values - Bootstrapping


Mean, STDEV, T-Values, P-Values
Confidence Intervals
Confidence Intervals Bias Corrected
Samples

R Square

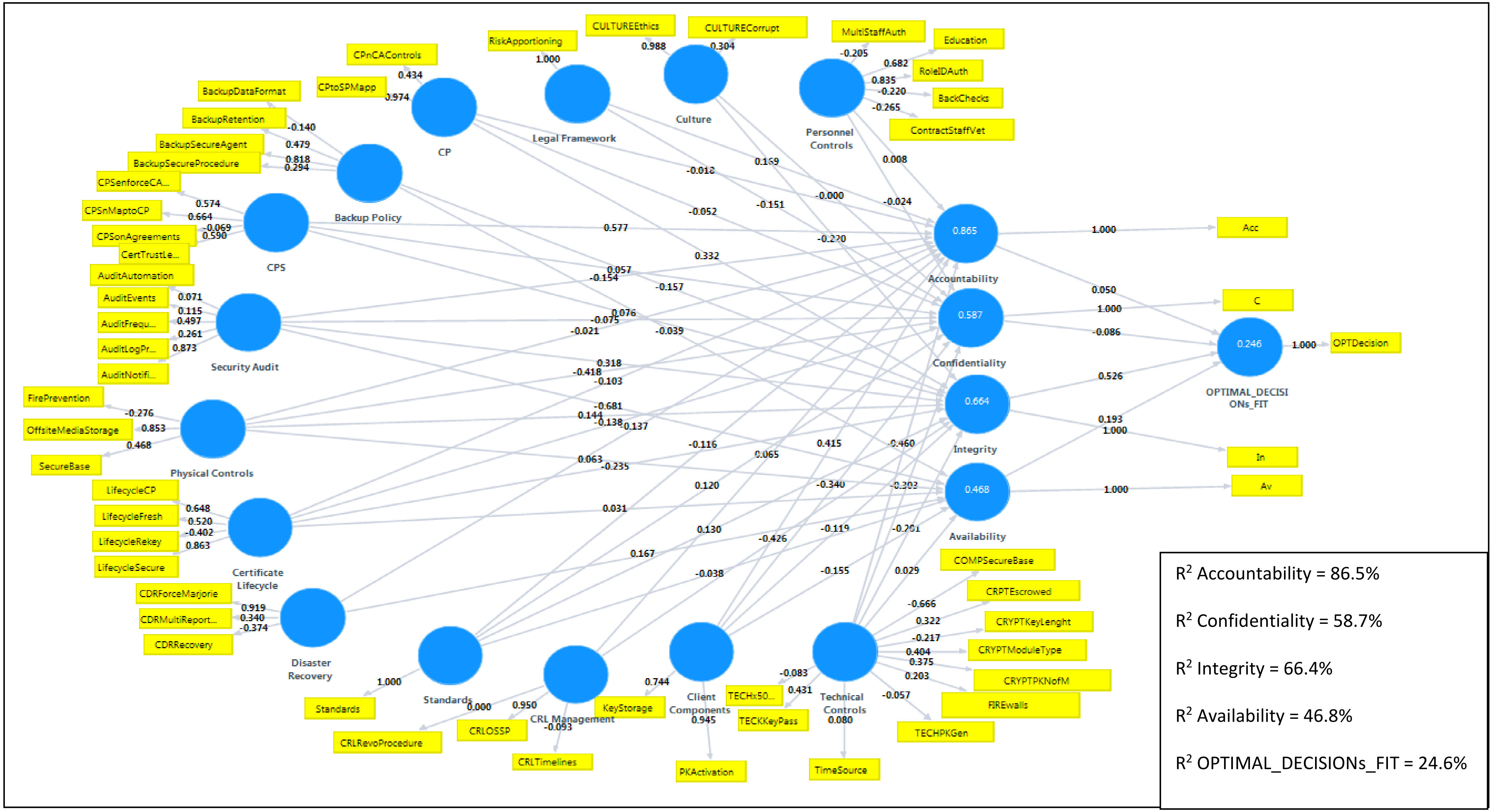
	Original Sample (O)	Sample Mean (M)	Standard Error (STERR)	T Statistics (|O/STERR|)	P Values
Accountability	0.466	0.743	0.113	4.138	0.000
Availability	0.512	0.618	0.100	5.134	0.000
Confidentiality	0.497	0.698	0.105	4.715	0.000
Integrity	0.590	0.774	0.089	6.644	0.000
OPTIMAL_DECISIONs_FIT	0.304	0.367	0.151	2.016	0.044

Source: Research Model in SmartPLS 3

The t and p-values of the OPTIMAL_DECISION_FIT show improvement.

4.3.3 PLS Algorithm

Figure 4.2: (A): PRE TEST Specifications: Weighting Scheme - path; Maximum Iterations - 300; Stop Criterion – 10-7; Initial Weights – 1.0.



R^2 Accountability = 86.5%

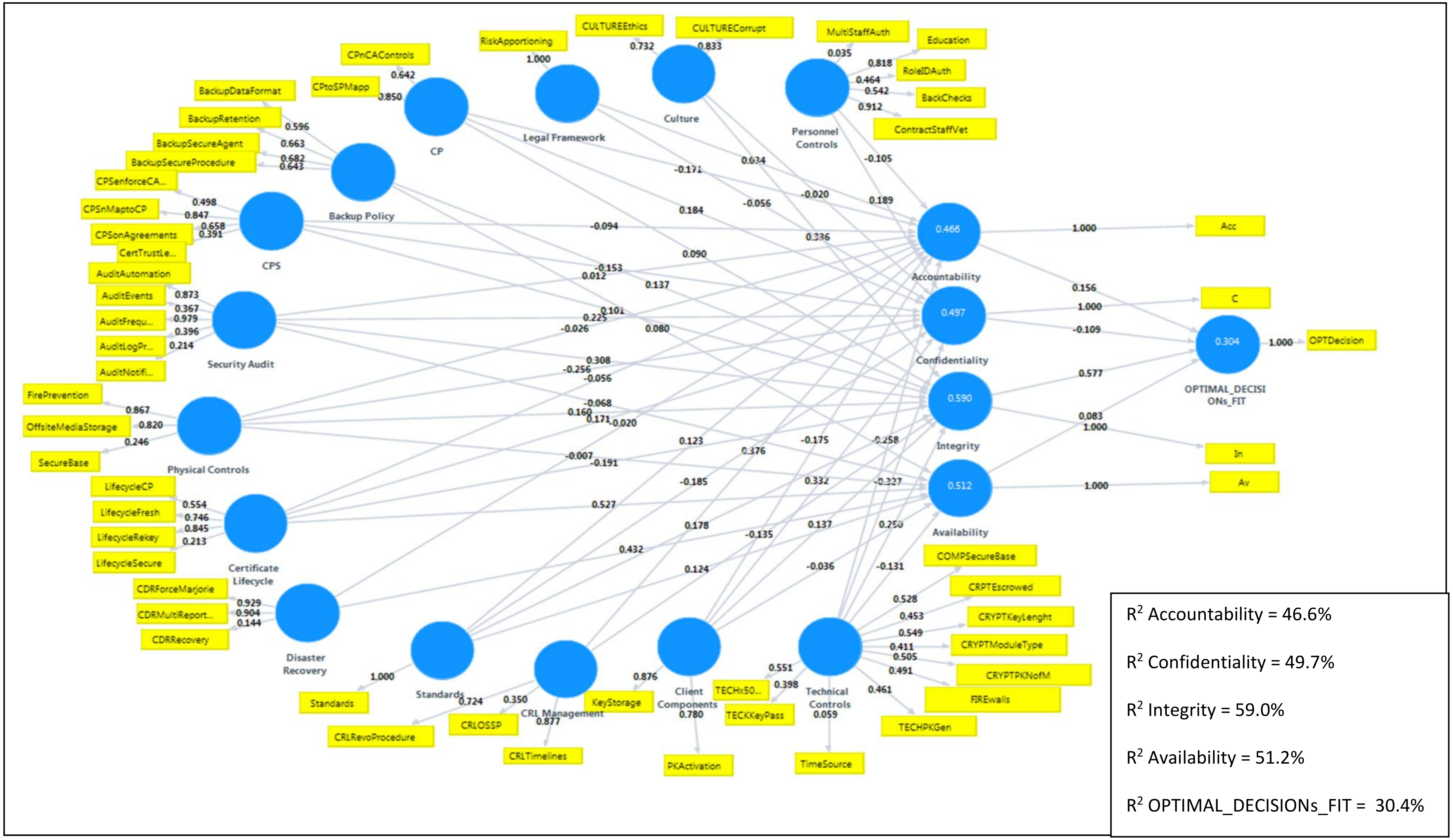
R^2 Confidentiality = 58.7%

R^2 Integrity = 66.4%

R^2 Availability = 46.8%

R^2 OPTIMAL_DECISIONs_FIT = 24.6%

Figure 4.2:(B) POST TEST Specifications: Weighting Scheme - path; Maximum Iterations - 300; Stop Criterion – 10-7; Initial Weights – 1.0



R^2 Accountability = 46.6%

R^2 Confidentiality = 49.7%

R^2 Integrity = 59.0%

R^2 Availability = 51.2%

R^2 OPTIMAL_DECISIONs_FIT = 30.4%

4.3.4 Path Coefficients – PLS Algorithm

The path coefficients indicate to us the correlations between the exogenous and the endogenous variables. It is an estimate of how much a change in the standard deviation of the exogenous variable would change that of the endogenous variable. We therefore look at the standard deviations of path coefficients. The relationship is defined using the following general equation:

$\sigma_Y = \sigma_x * P_c$ where

σ_Y is the standard deviation of the endogenous variable

σ_x is the standard deviation of the exogenous variable

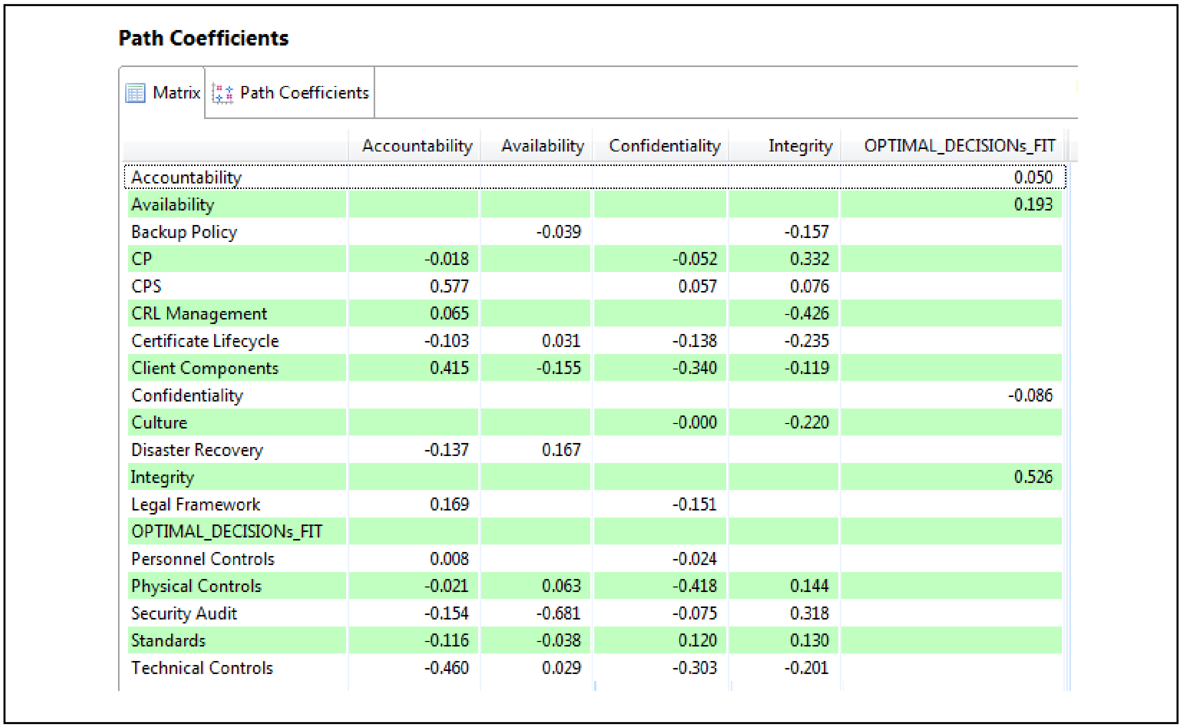
P_c is the path coefficient

It means if the standard deviation of the exogenous variable σ_x changes by 1, that of the endogenous variable σ_Y will change by P_c if P_c is positive and will reduce by P_c if P_c is negative.

Chin (1998) points out that this value should be greater than 0.200 (above 0.300 is better). Looking at Table 4.3 (A) and (B) stronger relationships are shown in (B) since most of the values in the standard error (SEM) field are greater. To get the standard deviations, we multiply (SEM x Square Root of N).

Based on the fact that P_c should have values between 1 and -1 as stated in Table 4.3 we can conclude that the path coefficients in both tables (A) and (B) are significant. However, in Table 4.3(B) there is a weakening of correlations between 20 pairs of variables (marked by red asterisk) and a strengthening of 25 correlations (marked by purple asterisk). Although a bit baffling, we suspect it is caused by the suppressor effect (Falk and Miller, 1992) when the sign of the path coefficients is different from that of the correlation coefficient.

TABLE 4.3 (A): BASELINE PLS Algorithm Path Coefficients



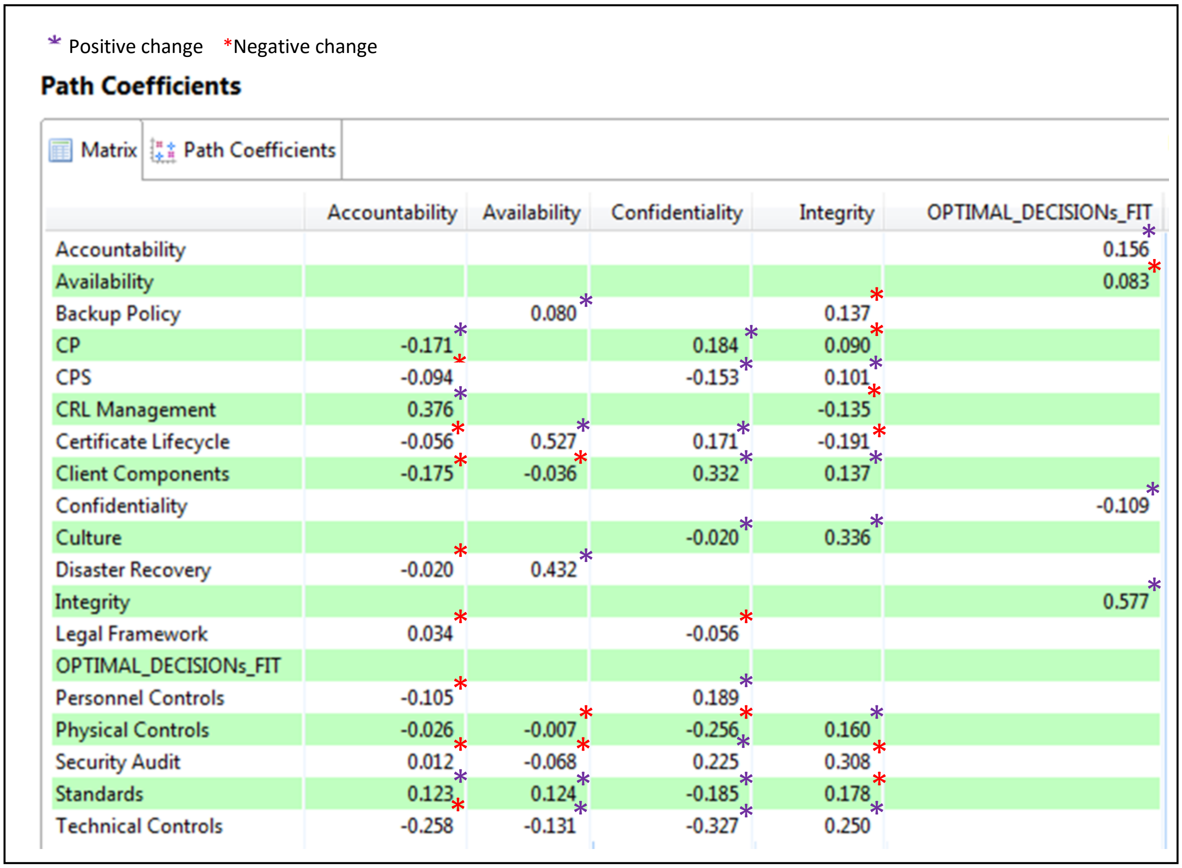
Path Coefficients

	Accountability	Availability	Confidentiality	Integrity	OPTIMAL_DECISIONs_FIT
Accountability					0.050
Availability					0.193
Backup Policy		-0.039		-0.157	
CP	-0.018		-0.052	0.332	
CPS	0.577		0.057	0.076	
CRL Management	0.065			-0.426	
Certificate Lifecycle	-0.103	0.031	-0.138	-0.235	
Client Components	0.415	-0.155	-0.340	-0.119	
Confidentiality					-0.086
Culture			-0.000	-0.220	
Disaster Recovery	-0.137	0.167			
Integrity					0.526
Legal Framework	0.169		-0.151		
OPTIMAL_DECISIONs_FIT					
Personnel Controls	0.008		-0.024		
Physical Controls	-0.021	0.063	-0.418	0.144	
Security Audit	-0.154	-0.681	-0.075	0.318	
Standards	-0.116	-0.038	0.120	0.130	
Technical Controls	-0.460	0.029	-0.303	-0.201	

Looking at Table 4.3 (A) all the path coefficients that are below 0.1 either positive or negative means the relationships are very weak hence not significant (Wong, 2013). For example, the correlation between the Certificate Policy (CP) and Confidentiality is -0.052 reflecting to be very weak. This may be explained by the fact that on the ground, there is lack of a proper certificate policy hence it may impact directly on confidentiality. This needs to be improved.

We note that all the values fall between 1 and -1 as they should but only a few attain the > 0.3 e.g. Security Audit -> Integrity at 0.308 and Certificate Lifecycle -> Availability at 0.527etc. A strong + towards +1 means there exists a strong relationship between the exogenous and endogenous variable while a strong –ve towards -1 indicates weak correlation as mentioned earlier.

Table 4.3 (B): POST TEST PLS Algorithm: Path Coefficients



* Positive change *Negative change

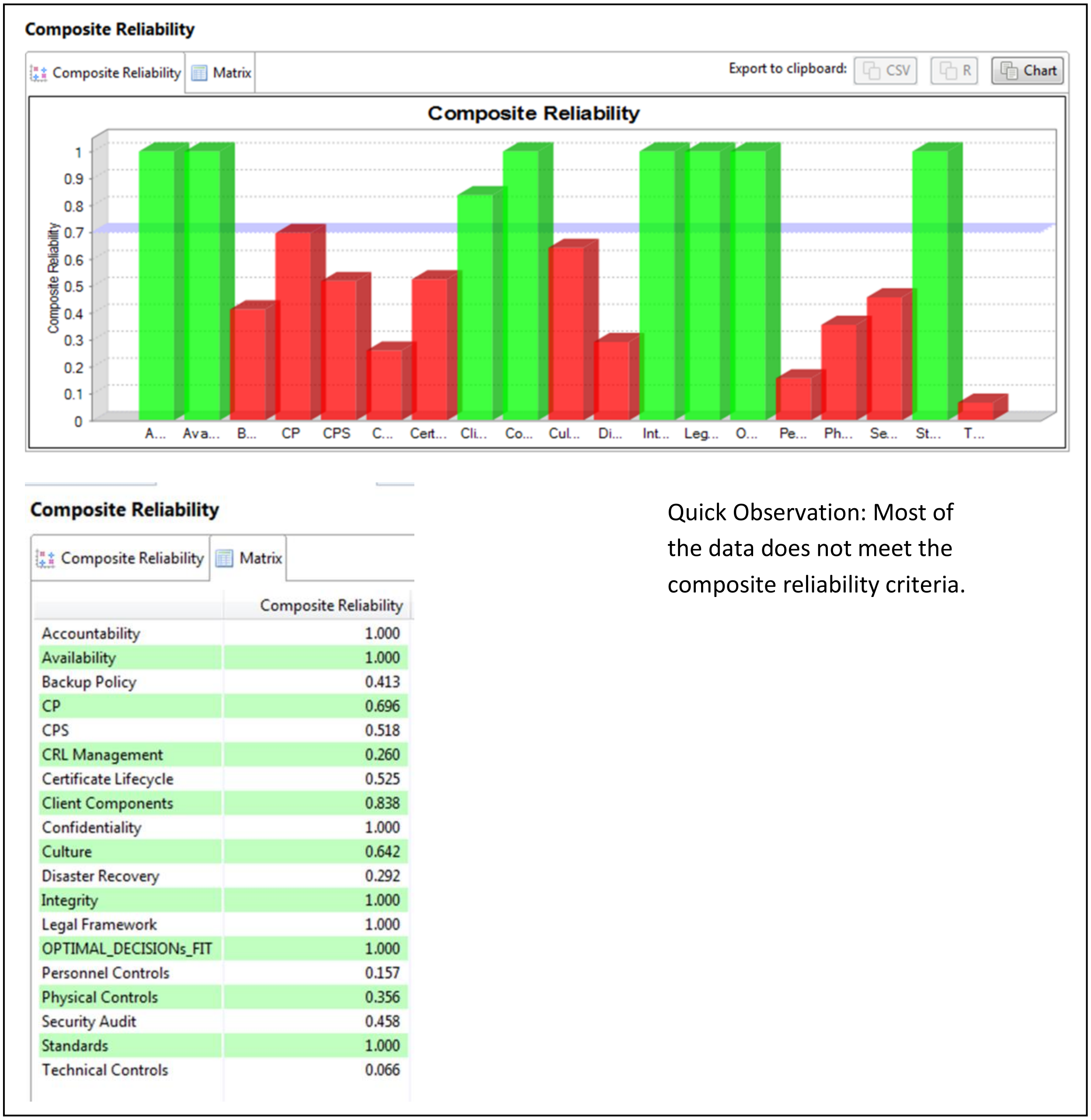
Path Coefficients

	Accountability	Availability	Confidentiality	Integrity	OPTIMAL_DECISIONs_FIT
Accountability					0.156*
Availability					0.083*
Backup Policy		0.080*		0.137*	
CP	-0.171*		0.184*	0.090*	
CPS	-0.094*		-0.153*	0.101*	
CRL Management	0.376*			-0.135*	
Certificate Lifecycle	-0.056*	0.527*	0.171*	-0.191*	
Client Components	-0.175*	-0.036*	0.332*	0.137*	
Confidentiality					-0.109*
Culture			-0.020*	0.336*	
Disaster Recovery	-0.020*	0.432*			
Integrity					0.577*
Legal Framework	0.034*		-0.056*		
OPTIMAL_DECISIONs_FIT					
Personnel Controls	-0.105*		0.189*		
Physical Controls	-0.026*	-0.007*	-0.256*	0.160*	
Security Audit	0.012*	-0.068*	0.225*	0.308*	
Standards	0.123*	0.124*	-0.185*	0.178*	
Technical Controls	-0.258*	-0.131*	-0.327*	0.250*	

4.3.5 Composite Reliability – PLS Algorithm

Before performing any other analysis on the data, it is important to find out whether the data is reliable. Composite reliability for each construct should be equal to or more than 0.7. This was achieved as displayed in Figure 4.3.

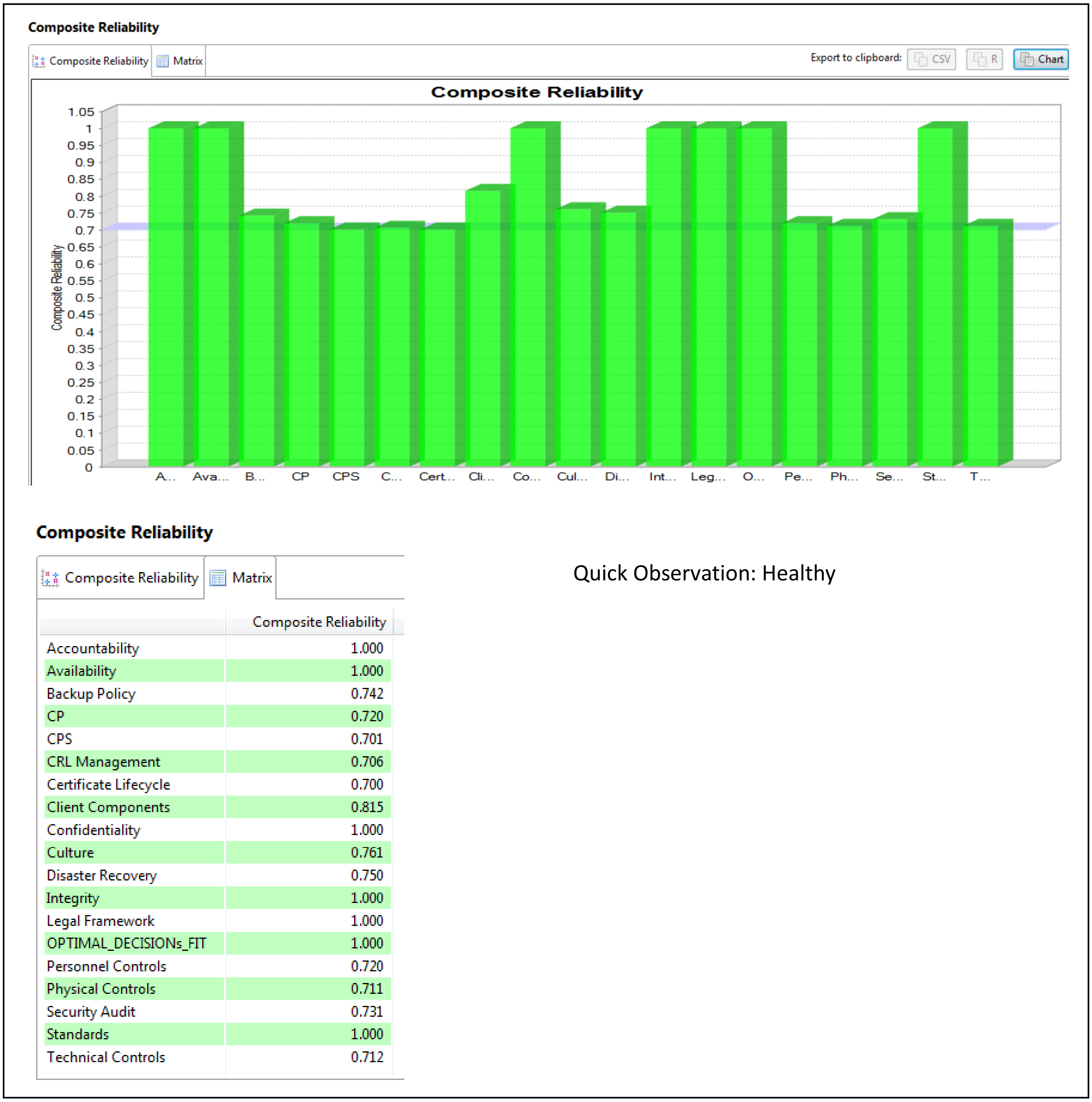
Figure 4.3: (A): PRE-TEST Composite Reliability – PLS Algorithm



	Composite Reliability
Accountability	1.000
Availability	1.000
Backup Policy	0.413
CP	0.696
CPS	0.518
CRL Management	0.260
Certificate Lifecycle	0.525
Client Components	0.838
Confidentiality	1.000
Culture	0.642
Disaster Recovery	0.292
Integrity	1.000
Legal Framework	1.000
OPTIMAL_DECISIONs_FIT	1.000
Personnel Controls	0.157
Physical Controls	0.356
Security Audit	0.458
Standards	1.000
Technical Controls	0.066

Quick Observation: Most of the data does not meet the composite reliability criteria.

Figure 4.3 (B): POST TEST Composite Reliability After Three Months – PLS Algorithm



	Composite Reliability
Accountability	1.000
Availability	1.000
Backup Policy	0.742
CP	0.720
CPS	0.701
CRL Management	0.706
Certificate Lifecycle	0.700
Client Components	0.815
Confidentiality	1.000
Culture	0.761
Disaster Recovery	0.750
Integrity	1.000
Legal Framework	1.000
OPTIMAL_DECISIONs_FIT	1.000
Personnel Controls	0.720
Physical Controls	0.711
Security Audit	0.731
Standards	1.000
Technical Controls	0.712

Quick Observation: Healthy

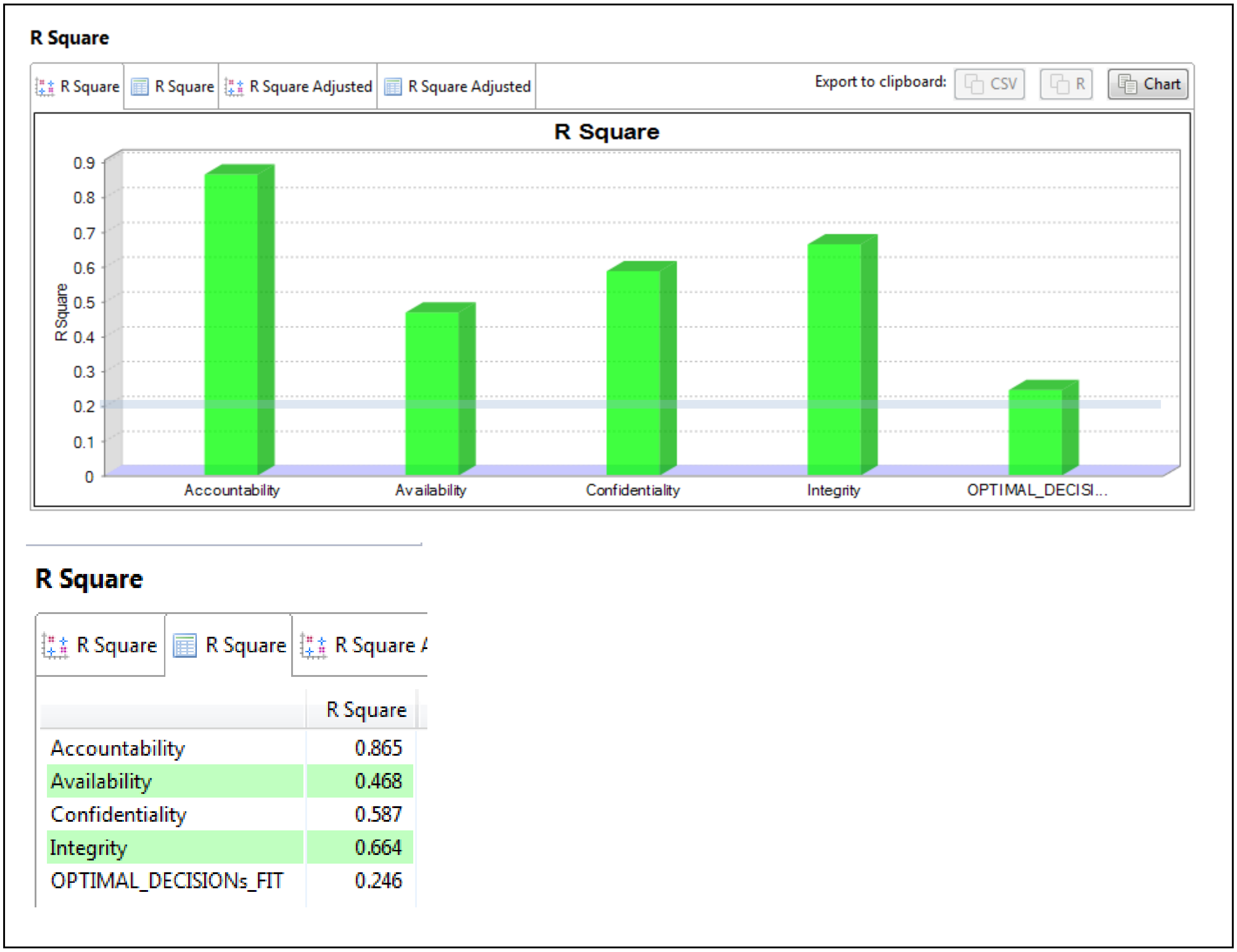
Source: Research Model in SmartPLS 3

4.3.6 R Square – PLS Algorithm

The R square (R^2) value shows how closely the data fits the regression line. In PLS, it is also called the coefficient of multiple regressions. We can say a model fits the data well if the differences between the observations and predicted values are small and unbiased. R^2 therefore indicates the percentage of the target variable variance explained by the linear model.

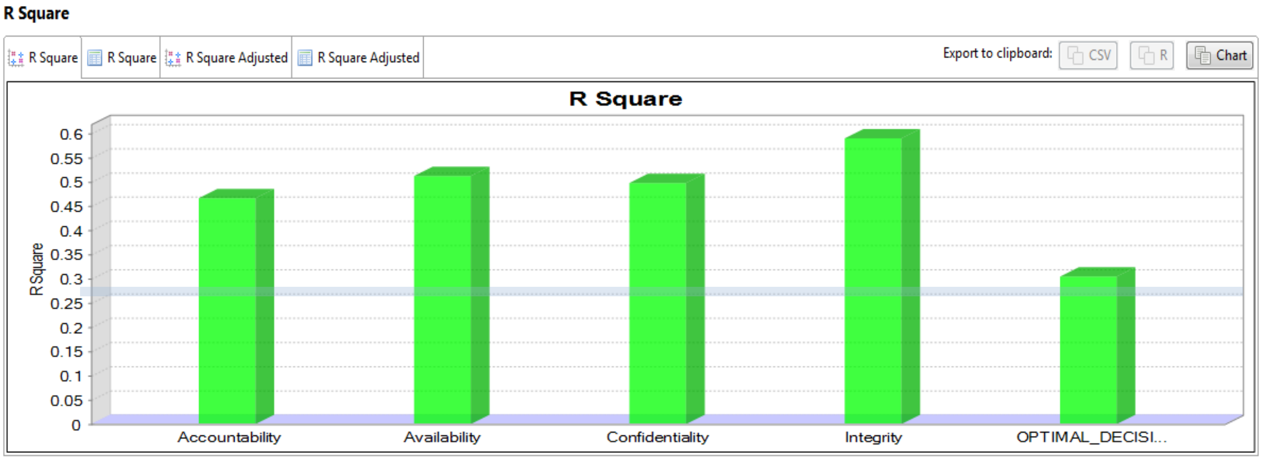
$$R^2 = \text{Explained Variation / Total Variation.}$$

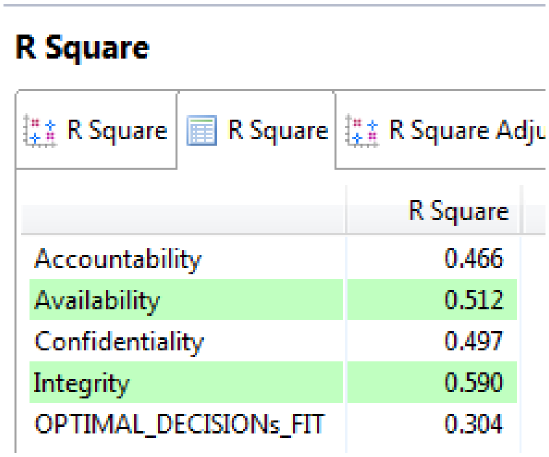
Figure 4.4(A): R^2 for Endogenous Variables Baseline Study – PLS Algorithm



	R Square
Accountability	0.865
Availability	0.468
Confidentiality	0.587
Integrity	0.664
OPTIMAL_DECISIONs_FIT	0.246

Figure 4.4(B): R^2 for Endogenous Variables After Three Months – PLS Algorithm





	R Square
Accountability	0.466
Availability	0.512
Confidentiality	0.497
Integrity	0.590
OPTIMAL_DECISIONs_FIT	0.304

The higher the R^2 value the better. In the baseline study, OPTIMAL_DECISION_FIT had a value of 24.6%. After three months, this value rose to 30.4% which shows a clear improvement and movement towards the desired results. R^2 of upto 50% is expected for this type of hard to measure problems.

Source: Research Model in SmartPLS 3

4.4 Interpretations

1. **Outer Weights Bootstrapping:** Table 4.1 shows that the indicator variables t-values and p-values. In PLS-SEM data is not distributed normally (Hair et al., 2013) hence parametric significance tests used in regression analysis cannot be used to test whether coefficients such as outer weights, outer loadings and path coefficients are significant. Instead, the non-parametric bootstrap procedure is used to calculate parameter estimates. Subsamples are then used to calculate standard errors and eventually the t-values. Therefore, each t-value is gauged as set out in Table 3.2. Based on this standard we can conclude as follows by looking at Table 4.1:
    a. PRE-TEST:
        i. 22 indicators meet the threshold of $p<0.05$ significance level marked by letter **S** and the row is green in color. The others which fall below the threshold have letter **O** and the rows are color filled.
        ii. The reasons why some indicators fall below the expected level are varied, from insufficient baseline data to an outright gap or poor performance.
    b. POST TEST:
        i. 29 indicators meet the threshold of $p < 0.05$ significance level.
        ii. Some indicators which were healthy in the pre-test lost significance in the post test either because the initial information received during the baseline study about them was unreliable or incomplete; or the data about them remained static as that of the others in the same block improved.

Generally there is an improvement. Even where the threshold is not reached in the post test results, most of the values show a positive increase except a few, pointing to the fact that there was a positive impact overall.

2. **R square (R^2) Bootsrapping**: Tables 4.2 (B) shows an improvement in the t-value of the OPTIMAL_DECISIONs_FIT from 1.570 in the pre-test to 2.016 in the post test. The p value in Table 4.2(B) for this value indicates 95% significance level as predicted. This means that the model has a positive impact on optimisation of the decision making process.
3. **Path Coefficients PLS:** Table 4.3 (A) and (B) indicate the relative significance of the exogenous variables on each of the endogenous variables. From Table 4.3 (B), we see that we have twenty five (25) correlations strengthening (marked by purple asterisk) and twenty (20) weakening. However, all the coefficients fall between the healthy range of 1 >= P_c>= -1. We can therefore say the model captures the correlations very well.

4. **PLS-SEM Composite Reliability:** Comparison of Figure 4.3(A) and (B) shows a great improvement composite reliability of data during the post test. In (A) eleven constructs fall below the 0.7 required threshold. In (B) all the latent variables attain a composite reliability of 0.7 and above. This indicates a positive impact on the ground.
5. **PLS-SEM R Square:** As shown in Figures 4.4 (A) and (B) the R^2 values are above 0.25 or 25%. This is the lowest acceptable value. However, the higher the R^2 value the better. Again Integrity seems to explain the biggest percentage of variance in the OPTIMAL_DECISIONs_FIT. The R^2value of the OPTIMAL_DECISIONs_FIT in the post test is higher i.e

PRE-TEST: 24.6%

POST TEST: 30.4%

In essence, the R^2 value should be as high as possible. From the figure above, we can say that after the three months period, there is a positive change and the model had a better fit to the data that was collected during the post-test than the pre-test.

6. **PLS-SEM f^2 Test:** f^2 measures the impact on R^2 of an endogenous variable when one of the exogenous variables is excluded. Although nearly all the constructs fall below the 0.35 minimum value, at least they are all positive.

4.5 Critical Observations and Gaps Identified

Based on Table 4.1, we carefully look at each quality attribute, its baseline values, interventions if any and finally its post-test value. We use colors in the comments column to indicate the state of various attributes. Red color indicates that the attribute is below the expected level while green shows that the attribute is healthy. Blue color indicates that there exists a gap in the current system which needs to be plugged i.e. a gap here means a missing technology or capability. Although we explain briefly the likely cause of each in the comments column, let us point out the most critical areas here.

1. **BackChecks<-Personnel Controls:** academic qualifications (as presented in certificates), national IDs, birth certificates and refereed resumes is good. However, PKI staff require far much more stringent background checks carried out by state investigative agencies to avoid employing disguised foreign agents especially in a country like Kenya which has porous borders.

2. **CPSenforceCAProcedures<-CPS:** The fact that the research discovered that there is no Certificate Policy (CP) document which normally forms the basis of coming up with a certification authority's procedures needs to be looked into and corrected.
3. **CPnCAControls<-CP:** Lack of a certificate policy. This anomaly should be corrected. A certificate policy (CP) lays down the requirements that PKI participants must have in order to operate within a PKI. A CP identifies different uses for certificates and the various end users that can participate in the PKI. A good CP ensures that even when we have more than one certificate authorities, there shall be interoperability between them.
4. **CRLOSSP<-CRLManagement:** Although there is Online Certificate Status Protocol in use, stapling is not supported. It is important to enforce stapling for a more efficient PKI implementation especially when eGovernment proliferates to serve the masses.
5. **TECHPKGen**<-**TechnicalControls:**The research failed to establish whether the cryptography method used is elliptic key cryptography (ECC) or RSA. However, it established that the key length is 256 bits. For ECC 256 bit key lengths is sufficient but if it is RSA or other methods, we need 1024 and above to reduce the threat of cryptanalysis and brute force attacks as specified in NIST-SP 800 – 131A standard.
6. **TimeSource<-TechnicalControls:** The research noted the lack of an in depended secure time stamp server for the PKI.
7. The physical location of the Root Certification Authority was not initially constructed as a high security zone.
8. There is need for confidence among staff that the computing architecture of the certification authority was constructed on secure computing base standards.
9. Need to come up with relevant strategies to prevent zero day attacks.

10. The number of employees implementing and maintaining the public key infrastructure at the ICT Authority needs to be increased. Currently, the cyber security section has only five employees working under one head.

11. Kenya's eGovernment initiative needs to be modeled around a secure framework like the one proposed in Figure 2.21. The one currently deployed as eCitizen.go.ke is on the open insecure internet. Similarly, each important agency needs to have its own server to form a distributed system or a cloud instead of centralizing everything at the treasury.

5. CONCLUSIONS

5.1 What We Set Out to Do

We set out to develop a quantitative decision support tool for optimizing PKI security rational decisions. After developing it, we aimed to evaluate its effectiveness by quantitatively measuring the PKI security quality attributes in order to serve as a foundation for improving or optimizing the decision making process affecting the security attributes. Looking back at the research objectives and questions, we conclude as follows:

1. On the main research question 1 and main objective 1: we can conclude that the research developed a quantitative model in PLS –SEM after deriving it from other models due to its suitability to our conceptual framework and research design.
2. On the supporting objective (a) and research question (i) we conclude that key quality attributes that influence the security of PKI were identified and captured in a conceptual framework. We also identified the need to statistically measure their achievement in PKI solutions using PLS – SEM analysis.
3. On supporting objective (b) and research question (ii) we conclude that a fitting conceptual framework was developed by deriving it from other existing frameworks.
4. On supporting objective (c) and research question (iii) we conclude that a fitting PLS-SEM model was developed mapped directly from the conceptual framework hence fulfilling this objective and research question.
5. On supporting objective (d) and research question (iv), we point out that the fact that the tool developed in PLS-SEM could be populated with data and that it computed statistical measures for various security quality attributes during the baseline and post-study; we conclude that the model is effective and it can be used to quantitatively assess PKI security attributes. The rational decision optimisation tool developed would assess

whether each measure meets the set level or not hence help a decision make to take the relevant action.

5.2 Contributions

This research has made contributions in the following areas:

1. **Best Practices:** Proposed a new statistical methodology of measuring the level of achievement of key information security quality factors in a public key infrastructure solution. The statistical measures can then be used by decision makers to optimise the security of PKI solutions.
2. **New Conceptual Models:** The study came up with a new conceptual models derived from existing models:
    a. The decision optimisation process captured in the model of Figure 2.13
    b. The Conceptual Model as captured in Figure 2.24
3. **Proposed eGovernment Model:** A proposed new PKI enabled eGovernment Model as detailed in Figure 2.21. The new model overcomes the weaknesses identified in three common models as depicted by those of the United Kingdom, Estonia and Kenya.

The research has disseminated its finding through seminars, local and international conferences which has led to some of its works being accepted for publishing.

5.3 Recommendations

More work needs to be done in this area in order to identify even more metrics that can be used to assess the level of achievement of security quality factors IN PKI solutions. One area that is troubling security experts to date is zero day attacks. Such attacks are difficult to predict or prepare for and therefore conventional methods fall short when it comes to assessing or modeling them hence the need for further research.

5.4 Conclusions

The research model developed, operationalised and evaluated in this research has stood scrutiny in various security metrics forums both locally and internationally. Based on its ability to measure the attributes identified in the study, we conclude that it is effective. Although we could not carry out the predictive relevance of the model (Q^2) due to the small sample size, we can conclude that the model was able to capture key indicator measures including gaps that required to be plugged. It is therefore the humble submission of this research that the model can be adopted to be used as a way of quantitatively measuring security metrics of public key infrastructure and other software systems.

REFERENCES

Aaviksoo, J., 2010. Cyberattacks Against Estonia Raised Awareness of Cyberthreats. Defense Against Terrorism Review 3, 13–22.

Aleti, A., Buhnova, B., Grunske, L., Koziolek, A., Meedeniya, I., 2013. Software Architecture Optimization Methods: A Systematic Literature Review. IEEE Transactions on Software Engineering 39, 658–683. doi:10.1109/TSE.2012.64

Al-Khouri, A.M., 2011a. PKI in Government Identity Management Systems (arXiv e-print No. 1105.6357).

Al-Khouri, A.M., 2011b. Optimizing Identity and Access Management (IAM) Frameworks. International Journal of Engineering Research and Applications 1, 461–477.

Alsaleh, M., Adams, C., 2006. Enhancing consumer privacy in the liberty alliance identity federation and web services frameworks, in: Proceedings of the 6th International Conference on Privacy Enhancing Technologies, PET'06. Springer-Verlag, Berlin, Heidelberg, pp. 59–77. doi:10.1007/11957454_4

Anand, V., Saniie, J., Oruklu, E., 2012. Security Policy Management Process within Six Sigma Framework. Journal of Information Security 3, 49–58.

Angus, S., 2010. Oxford Dictionary of English, 3rd ed. Oxford University Press, Oxford,UK.

Ashmore, W.C., 2009. Impact of Alleged Russian Cyber Attacks. Baltic Security & Defence Review 11, 4–40.

Australian Government, 2009a. GATEKEEPER PKI FRAMEWORK THREAT AND RISK ASSESSMENT TEMPLATE.

Australian Government, 2009b. National e-Authentication Framework: Better Practice Guides Vol. 1 Identity e-Authentication.

Birrell, E., Schneider, F.B., 2013. Federated Identity Management Systems: A Privacy-Based Characterization. IEEE Security & Privacy 11, 36–48.

Carroll, J.M., Swatman, P.A., 2000. Structured-case: a methodological framework for building theory in information systems research. European Journal of Information Systems 9, 235–242. doi:10.1057/palgrave.ejis.3000374

Chen, D., Cremers, C., Hyun-Jin Kim, T., Perrig, A., Sasse, R., Szalachoeski, P., 2014. ARPKI: Attack Resilient Public-Key Infrastructure, in: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security. Presented at the 2014 ACM SIGSAC Conference on Computer and Communications Security, ACM, New York, pp. 382–393. doi:10.1145/2660267.2660298

Choudhury, S., Bhatnagar, K., Haque, W., NIIT, 2002. Public Key Infrastructure: Implementation and Design. Hungry Minds M&T Books, New York, NY 10022.

Cole, P., Duserick, W., Lesser, J., Podorowsky, G., Sibieta, P., Thornby, C., 2003. Privacy and Security Best Practices (Report No. Version 2.0). Liberty Alliance.

Collins, D., 2014. Collins English Dictionary. Collins English Dictionary.

Dempsey, J.X., 2004. Creating the Legal Framework for Information and Communications Technology Development: The Example of E-Signature Legislation in Emerging Market Economies. MIT Information Technologies and International Development 1, 39–52.

Ernst & Young, 2013. Identity and Access Management: Beyond Compliance.

Federal Public Key Infrastructure Authority, 2015. Federal Public Key Infrastructure (F PKI) Compliance Audit Requirements.

FIDIS WP2, 2005. D2.3: Models: Future of IDentity in the Information Society (deliverable No. D2.3), Future of Identity in the Information Society. FIDIS.

FIDIS WP3, 2005. D3.1:Structured Overview on Prototypes and Concepts of Identity Management Systems" (Deliverable No. D3.1). European Information Society FIDIS, United Kingdom.

Fox, J., Weisberg, H.S., 2010. An R Companion to Applied Regression, Second Edition edition. ed. SAGE Publications, Inc, Thousand Oaks, Calif.

Fursule, N.V., Bansod, S.V., Fursule, S.N., 2012. Understanding the Benefits and Limitations of Six S igma Methodology. International Journal of Scientific and Research Pu blications 2.

Ghani, E., Laswad, F., Tooley, S., Jusoff, K., 2009. The Role of Presentation Format on Decision-makers' Behaviour in Accounting. International Business Research 2, 183–195.

GoK Ministry of Finance, 2016. Budget Statement for the fiscal Year 2016/2017.

Goldtack, M., 2006. Lesson learned: From MS Passport to CardSpace.

Goulet, W., 2009. Analyzing Enterprise PKI Deployments.

Hair, J.F., Ringle, C.M., Sarstedt, M., 2013. Editorial - Partial Least Squares Structural Equation Modeling: Rigorous Applications, Better Results and Higher Acceptance (SSRN Scholarly Paper No. ID 2233795). Social Science Research Network, Rochester, NY.

Harman, M., Mansouri, S.A., Zhang, Y., 2012. Search-based Software Engineering: Trends, Techniques and Applications. ACM Comput. Surv. 45, 11:1–11:61. doi:10.1145/2379776.2379787

Hays, R.T., Stout, R.J., Ryan-Jones, D.L., 2005. Quality Evaluation Tool For Computer and Web-Delivered Instruction (U).

Heeks, R., 2003. iGovernment Working Paper No. 14 - Most eGovernment-for-Development Projects Fail: How Can Risks be Reduced?

Ijaz, I., 2012. Design and Implementation of PKI For Multi Domain Environment. International Journal of Computer Theory and Engineering 4, 505–509.

ISACA, 2015. CISA Review Manual, 26th ed. ISACA, Rolling Meadows, USA.

ISC, 2003. PKI Assessment Guidelines.

Johnson, P., Ekstedt, M., 2007. Enterprise Architecture Models and Analyses for Information Systems Decision Making. Studentlitteratur, Stockholm.

Johnson, P., Johansson, E., Sommestad, T., Ullberg, J., 2007. A Tool for Enterprise Architecture Analysis, in: Enterprise Distributed Object Computing Conference, 2007. EDOC 2007. 11th IEEE International. Presented at the Enterprise Distributed Object Computing Conference, 2007. EDOC 2007. 11th IEEE International, pp. 142–142. doi:10.1109/EDOC.2007.25

Johnson, P., Lagerstrom, R., Ekstedt, M., Osterlind, M., 2014. IT Management with Enterprise Architecture. Royal Institute of Technology, Stockholm, Sweden.

Kaplan, R.S., 2012. Conceptual Foundations of the balanced ScoreCard.

Kaplan, R.S., School, H.B., 2010. Conceptual Foundations of the Balanced Scorecard. Harvard Business School.

Koziolek, A., Reussner, R., 2011. Towards a Generic Quality Optimisation Framework for Component-based System Models, in: Proceedings of the 14th International ACM Sigsoft Symposium on Component Based Software Engineering, CBSE '11. ACM, New York, NY, USA, pp. 103–108. doi:10.1145/2000229.2000244

Lips, M., Pang, C., 2008. IDENTITY MANAGEMENT IN INFORMATION AGE GOVERNMENT EXPLORING CONCEPTS, DEFINITIONS, APPROACHES AND SOLUTIONS.

Loskot, M., 2008. Learning the OpenID problems [WWW Document]. URL http://mateusz.loskot.net/2008/05/14/learning-the-openid-problems/ (accessed 7.23.12).

MIT, 2012. Six Sigma Basics.

Mjølsnes, S.F., Mauw, S., Katsikas, S., 2008. Public Key Infrastructure: 5th European PKI Workshop: Theory and Practice, EuroPKI 2008 Trondheim, Norway, June 16-17, 2008, Proceedings. Springer Science & Business Media.

Monecke, A., Leisch, F., 2012. semPLS: Structural Equation Modeling Using Partial Least Squares. JSS 48.

Narman, P., Johnson, P., Nordstrom, L., 2007. Enterprise Architecture: A Framework Supporting System Quality Analysis, in: Enterprise Distributed Object Computing Conference, 2007. EDOC 2007.

11th IEEE International. Presented at the Enterprise Distributed Object Computing Conference, 2007. EDOC 2007. 11th IEEE International, pp. 130–130. doi:10.1109/EDOC.2007.39

NIST, 2014. Assessing Security and Privacy Controls in Federal Information Systems and Organizations Building Effective Assessment Plans.

Nograšek, J., 2011. Change Management as a Critical Success Factor in e-Government Implementation. Business Systems Research 2, 1–56.

OECD, 2011. DIGITAL IDENTITY MANAGEMENT Enabling Innovation and Trust in the Internet Economy.

Patel, R.R., Oza, B., 2013. Enhanced OpenID Protocol in Identity Management. International Journal of Application or Innovation in Engineering and Management 2.

Ringle, C.M., Sarstedt, M., Straub, D.W., 2012. A Critical Look at the Use of PLS-SEM in MIS Quarterly. MIS Quarterly 36, iii–xiv.

RSA, 1999. Understanding Public Key Infrastructure: An RSA Data Security White Paper.

Satyanarayana, J., 2004. E Government: The Science of the Possible. PHI Learning Pvt. Ltd.

Schumacker, R.E., Lomax, R.G., 2012. A Beginner's Guide to Structural Equation Modeling: Third Edition. Routledge.

Simon, M.K., 2011. Dissertation and Scholarly Research: Recipes for Success.

Skyles, A.O., 1992. An Introduction to Regression Analysis.

Smedinghoff, T., 2010. Building an Online Identity Legal Framework: The Proposed National Strategy (Report No. 800-372–1033), Privacy and Security Law. The Bureau of National Affairs, USA.

StatSoft, 2000. Partial Least Squares (PLS) [WWW Document]. Partial Least Square. URL http://www.uta.edu/faculty/sawasthi/Statistics/stpls.html (accessed 2.18.15).

Thakur, J., Kumar, N., 2011. DES, AES and Blowfish: Symmetric Key Cryptography Algorithms Simulation Based Performance Analysis. International Journal of Emerging Technology and Advanced Engineering 1.

Turpin, S.M., Marais, M.A., 2004. Decision-making: Theory and Practice.

United Nations, 2009. Making Data Meaningful.

University of North Carolina, 2007. Variance and Design of Experiments [WWW Document]. Variance. URL http://www.unc.edu/courses/2007spring/psyc/530/001/variance.html (accessed 2.27.15).

UNPAN, 2014. 2014 Global E-Government Survey [WWW Document]. 2014 Global E-Government Survey. URL http://www.unpan.org/egovkb/global_reports/08report.htm

Vaidya, T., 2015. 2001-2013: Survey and Analysis of Major Cyberattacks.

Waema, T., Adera, E., 2011. Local Governance and ICT's in Africa: Case Studies and Guidelines for Implementation and Evaluation. Pambazuka Press, United Kingdom.

Wagner, R., 2010. Identity and Access Management:Key Initiative Overview.

Wisniewski, T.W., Nadalin, T., Cantor, S., Hodges, J., Mishra, P., 2005. SAML V2.0 Executive Overview (No. sstc-saml-exec-overview-2.0-cd-01). OASIS, 25 Corporate Drive Suite 103 Burlington, MA 01803-4238 USA.

Wong, K.K.-K., 2013. Partial Least Squares Structural Equation Modeling (PLS-SEM) Techniques Using SmartPLS. Marketing Bulletin Technical Note.

Zachman, J.A., 2008. About The Zachman Framework^{TM} [WWW Document]. John Zachmans Concise Definition of the Zachman Architecture. URL https://www.zachman.com/about-the-zachman-framework (accessed 9.19.14).

Zubrow, D., 2004. Software Quality Requirements and Evaluation, the ISO 25000 Series.

APPENDIX I: RESEARCH QUESTIONNAIRE

University of Nairobi
School of Computing and Informatics
RESEARCHER: Geoffrey Chemwa

This data is collected as part of a PHD research titled "Optimising Security in Public Key Infrastructure Solutions for eGovernment". The responses will be fed into a rational decision optimisation tool. All the data collected will be held in utmost confidence. All permissions from relevant authorities have been granted.

Please take a few minutes to fill out the questionnaire form. We greatly appreciate time and effort helping us towards completion of the study.

1. I would welcome further education and training to enhance my capacity in my current job/role as a PKI expert

    ○ No, I have all the knowledge and skills that I need to meet any challenge
    ○ No, when I meet challenges I have capacity to research, self-learn and apply
    ○ Yes, some of the challenges require just a bit of training
    ○ Yes, I require moderate training
    ○ Yes, I require frequent training
    ○ Yes, I require very frequent training
    ○ Yes, I require to be retrained fully

2. Before I got employed here I was thoroughly vetted and a background check done on the following:

    1. Birth certificate 2. National ID 3. Academic certificates 4. References

    ○ I was not vetted
    ○ Academic certificates and References I provided
    ○ Birth certificate, National ID and the References
    ○ Birth certificate, National ID and Academic certificates
    ○ National ID, Academic certificates and References
    ○ All the above
    ○ All the above and more: State __________________

3. All PKI staff are identified and authenticated on the PKI system using the following:

    ○ Code generators together with Username and Password
    ○ Biometrics together with Username and password
    ○ Biometrics together with Code generators

○ Digital keys together with Username and password

○ Digital keys together with Biometrics

○ Hardware keys together with Usernames and passwords

○ Hardware keys together with Digital keys

4. Multi-staff identification and authentication is required to unlock/perform sensitive PKI modules/tasks

○ None of the tasks

○ Very few of the tasks

○ Few of the tasks

○ A few of the tasks

○ Some of the tasks

○ Most of the tasks

○ All of the sensitive tasks

5. All contract staff go through rigorous vetting, are closely monitored before selection and they sign binding confidentiality/bonding agreements.

○ No ○ Yes ○ Don’t know

6. We have a strong professional code of conduct which is actively enforced

○ No ○ Yes ○ Don’t know

7. As far as I know, corruption/nepotism in this organisation can be rated at:

○ Rampant

○ Very common

○ A few cases

○ Few cases

○ A bit common

○ Very few cases

○ Zero

8. Our certificate policy was derived from Kenya's cyber security policy and maps to it strongly

○ No ○ Yes ○ Don’t know

9. The Certificate Policy structures digital certificates into levels of trust general purpose / commercial among others.

○ No ○ Yes ○ Don’t know

10. The Certificate Policy enforces strict auditable requirements on the CA, RP and subscribers

○ No ○ Yes ○ Don't know

11. Our Certificate Practice Statement (CPS) was derived directly from the Certificate Policy

○ No ● Yes ○ Don't know

12. The CPS is a critical element of any contractual agreements with relying parties and subscribers

○ No ○ Yes ○ Don't know

13. The Certificate Practice Statement enforces strict auditable security procedures on CAs, RPs and subscribers

○ No ○ Yes ○ Don't know

14. All facilities that host the Root Certification Authority were built as a high security zone with all the zero physical break in controls stated below:
1. Highly secured computer room 2. Access control lists 3. Token access 4. Armed guards

○ No ● Yes ● Don't know

15. Data storage media is stored in offsite/separate locations and ready state redundancy is ensured.

○ No ● Yes ○ Don't know

16. There is a good fire prevention policy / training / equipment.

○ No ○ Yes ○ Don't know

17. There exist secure procedures for vetting and selecting a data and private keys backup agent

○ No ○ Yes ○ Don't know

18. Data is backed up in the following formats:

○ Plain text

○ Some plain text and some with unsplit key encrypted

○ Some plain text and some with split key encrypted

○ All with unsplit key encrypted

○ Some with unsplit and other with split key encrypted

○ All with split key encrypted

○ Better than split key encrypted: Other ____________________

19. There are secure data backup procedures in place which are strictly enforced

○ No ○ Yes ○ Don't know

20. We have sufficient backup retention period guidelines

○ No ○ Yes ○ Don't know

21. There is a sound audit policy which identifies what auditable events should be captured

○ No ○ Yes ○ Don't know

22. The security audit log is properly protected against modification / deletion

○ No ○ Yes ○ Don't know

23. The system collects all the data required for audit and generates all audit reports automatically

○ No ○ Yes ○ Don't know

24. The frequency of the system security audit can be said to be:

○ Once a year ○ Once every half year ○ Once every three months

○ Once every month ○ Once every two weeks ○ Once a week

○ Once a day

25. Offenders who cause audit queries are promptly informed based on a set of procedures

○ No ○ Yes ○ Don't know

26. There is a sound secure certification policy that manages the certificate/key lifecycle

○ No ○ Yes ○ Don't know

27. All digital certificate applicants will be registered afresh after authentication of their existing identity tokens

○ No ○ Yes ○ Don't know

28. There are secure procedures that govern the following processes
Check against that which you agree with

□ Private key re-keying □ Certificate modification □ Certificate renewal □ Certificate recovery

29. The following standards are intrinsic to all our procedures
Check against that which you agree with

□ X.509 □ FIPS 140-2 for cryptographic module □ NIST SP 800 - 131A

30. Certificate compromise is reported through the following mechanisms:

□ Telephone □ Signed email □ Social media e.g. Skype □ Other _______________

31. After resource corruption or certificate/key compromise we have secure recovery procedures

○ No ○ Yes ○ Don’t know

32. We have a sound continuity plan that covers advanced human exploits and force marjorie occurrences:
Select those you feel you are well prepared to deal with in case they happen:

□ Solar flares/radiations

□ Earthquakes

□ Floods

□ Power outages

□ Fires

□ Internal compromise

□ Zero day attacks

33. There are secure procedures on how to initiate certificate revocation requests
A digitally signed message is required

○ No ○ Yes ○ Don’t know

34. Strict reporting timelines are observed on certificate compromise and processing progress

○ No ○ Yes ○ Don’t know

35. The following Online Certificate Status Protocol (OCSP) is used:

○ OCSP with stapling ○ OCSP without stapling ○ Don’t know

36. There are secure procedures on who is allowed to generate private/public keys

○ No ○ Yes ○ Don’t know

37. After generation, keys are passed securely to their relevant owners through:
Check as appropriate:

□ Post office unregistered mail

□ Post office registered mail

□ Unsigned email

□ Non-SSL protected session

□ SSL protected session

□ Signed email

□ Personal collection after identity authentication

38. All certificates issued are X.509 Version 3 certificates

○ No ○ Yes ○ Don’t know

39. The cryptographic module is of the following type

□ Software □ Firmware □ Hybrid □ Hardware

40. We use the following key lengths to generate private and public keys:

○ 32 bits ○ 64 bits ○ 128 bits ○ 256 bits ○ 512 bits ○ 1024 bits ○ 2048 or more

41. The CA's private key is under very secure split security management i.e. n of m principle ?

○ No ○ Yes ○ Don’t know

42. The CA private keys are escrowed. If so there was a secure procedure for engaging escrow agent

○ No ○ Yes ○ Don’t know

43. There is a secure procedure for activation/deactivation/destruction of CA, RA and subscriber private keys

○ No ○ Yes ○ Don’t know

44. The CA's computing architecture is secure and adheres to the trusted computing base standards

○ No ○ Yes ○ Don’t know

45. Network resources are secured using high security firewall

○ No ○ Yes ○ Don’t know

46. There is a trusted time source for time stamping data

○ No ○ Yes ○ Don’t know

47. On the client side, private keys are stored in:

□ Computer always connected to internet without firewall

□ Computer always connected to internet with firewall

□ Stand-alone computer connected once-in-a-while to internet without firewall

□ Stand-alone computer connected once-in-a-while to internet with firewall

□ Removable media like hard disk or flash disk

□ Magnetic cards

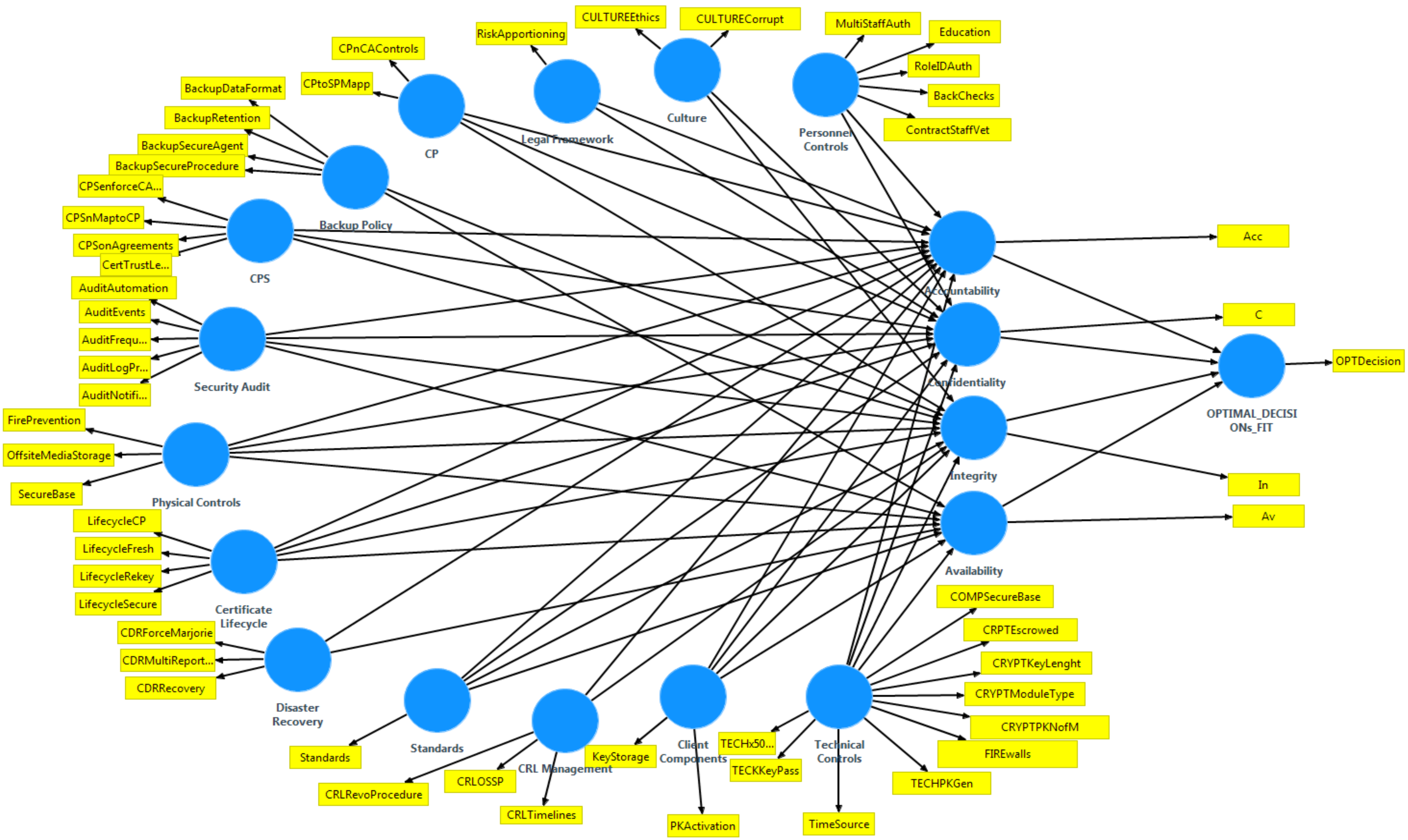
□ Smartcard

48. There is a sound legal and regulatory framework that apportions risk and indemnity of all PKI operations

○ No ○ Yes ○ Don’t know

APPENDIX II: PLS SEM MODEL

[



PLS-SEM Model