A Model for Optimising Security in Public Key Infrastructure Solutions for eGovernment

A case study of Kenya

Geoffrey Wekesa Chemwa

Department of Computing, Institute of Computer Science and Information Technology (ICSIT), Jomo Kenyatta University of Agriculture and Technology, P. O. Box 62000 00200, NAIROBI, KENYA.

Tel: +254722681057, Fax: +254 312 123 4567, E-mail: chemwex@icsit.jkuat.ac.ke

Abstract-Public Key Infrastructure (PKI) quality attributes like security, availability, integrity, interoperability etc. are latent in nature meaning they cannot be measured or observed directly. This presents a problem on how they can be optimized since as Drucker's maxim goes, if you can't measure it, you can't manage it. 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 quarterly audits. Such strategies are paperwork intensive and try to ensure process adherence but lack the capacity to quantitatively measure non-functional quality properties. eGovernments and their cyber security strategies, face massive threats from a knowledge society that has easy access to hacking tools, and also well-funded hacker groups, some sponsored by foreign governments.In this work,we derive a conceptual framework from existing frameworks then model a quantitative decision support tool using path analysis techniques, specifically Partial Least Square Structural Equation Modeling.The data used to initialize the model is real data collected from an ongoing PKI implementation. We opine that if key decisions are optimized during planning, implementation and auditing, then the security of the a PKI solution will also be optimized. We also provide an eGovernment arrangement that relies on PKI security for identification, authentication and authorization. It is worthwhile to note that although PKI is a universal concept, its design and implementation in different contexts means that each context offers emergent challenges that require unique security solutions.

Keywords-Public Key Infrastructure; Digital Certificate; eGovernment; Cyber Security; Structural Equation Modelling.

1. Introduction

Governments are adopting new ways of doing business through the digital platform and are embracing online and mobile applications not only to improve internal efficiencies but offer their citizens delightful service.The security of eGovernment relies on secure identification and authentication of all stakeholders during transactions to make sure that only authorized parties get access to the relevant resources at the right time [1,2]. One reliable method of ensuring this in such complex environments is the use of public key infrastructure solutions to register all players, issue them with digital certificates and ensure that all communications are signed with digital signatures [3]. However, there is no best fit formula for optimizing all PKI solutions globally since each PKI operates in different contexts and each context offers emergent challenges that require to be addressed in a unique way [4].

In this paper, we shall contribute to knowledge by developing a quantitative model for rational decision optimization when reasoning about PKI security in developing economies. We are cognizant of the fact that in most governments, PKI regulators, planners, implementers and assessors rely on quality management systems like ISO and standards such as x.50x in their PKI quarterly audits or reviews. In fact ISO 9126-(1-4) and later ISO 25030 (which is part of the Software Quality and Requirements Evaluation (SQuaRE) the ISO 25000 series) forms the basis of

identifying security as a worthwhile topic worth researching. Audits based on the standards above are paperwork intensive and try toensure process and requirements compliance mainly through checklists but lack the capacity to measure latent quality properties like security, interoperability, availability, privacy, reliability, performance etc. which are not explicit hence cannot be observed directly. We demonstrate how security can be modeled using multivariate assessment of factors that have causal relationships using partial least squares structural equation models. After collecting data using questionnaires and interview methods, we use regression analysis in the form of Partial Least Squares Structural Equation Modeling to model and perform measurements in SmartPLS Version 3 [5]. The output is a generic but extensible quantitative PKI security rational decision optimization model. The model shall display variable relationships and their quantitative weights in such a manner that decision makers can use them to prioritize resources and or take corrective actions where needed during audits or when predicting scenarios [6].

2. Materials Theories and Methods

2.1. Software Quality Optimisation

In this section we briefly review other software quality optimization approaches presented in other works before justifying why we chose to utilize Partial Least Squares, Structural Equation Modeling (PLS-SEM). The term optimization is not new when talking about software systems. Reference [7] presents software cost optimization using linear COCOMO equations. As is well known, COCOMO concentrates on effort and cost and ignores other important software quality properties like security. Reference [8] and [9] propose Enterprise Architecture Analysis (EAA) techniques to optimize non-functional quality attributes like security, availability, interoperability, integrity etc. This is good and in line with this paper. However EAA tools are derived from Unified Modeling Language (UML) and modeling follows Open Group's ArchiMate. Enforcing quality attribute constraints using Object Constraint Language (OCL) requires considerable programming effort that many researchers would find difficult to learn. Reference [10] also presents the Architecture Tradeoff Analysis methodology (ATAM) initially developed by the Software Engineering Institute as a software architecture quality optimization framework. ATAM performs architecture analysis and design tradeoff decisions in order to achieve desired quality attributes such as security, performance, availability etc. in the final solutions. ATAM is good but its results depend on the quality of the architecture. It concentrates more on tradeoffs but lacks an inference or predictive capability based on quantitative assessments of the latent quality variables. Besides, PKI security is so critical that such pareto optimal [11] techniques may compromise the entire system (tradeoffs may introduce loopholes which can be used to commit exploits). Reference [12] also presents a comparative study on software quality optimization either using case-based or parametric methods. However, they view optimization from the point of view of the discovery and removal of defects only and ignore other important quality attributes. Reference [13] discusses how to optimize the quality of e-learning systems components using multi-criteria evaluation, and specifically mention security as one of the key criteria that must be optimized. However, their model is too broad and does not give the security aspect the in-depth treatment it deserves.

Other works like [14] suggest search based software engineering (SBSE) techniques as a means of searching for optimal solutions when faced by a large search space of potential solutions. SBSE strategies include automated tools that utilize simulated annealing and genetic algorithms to optimize activities such as requirements engineering, costing, project management, maintenance, quality assessment etc. (ibid). However, SBSE techniques use meta-heuristic algorithms to search large solution spaces to arrive at optimal solutions. This is computationally intensive and requires significant execution time that may render such techniques infeasible[15]. Lastly but not least, [6] presents the partial least square structural equation modeling (PLS-SEM) technique which is a multivariate data analysis method that can test theoretically supported linear and additive causal models. In our case, we adopt PLS-SEM to model software quality properties like security, performance etc. and the multi-variables that influence them in a user friendly and easy to understand environment. Other factors that influenced our choice for this framework are the ability to represent causal relationships in path models and perform

predictive quantitative assessments on them even with small sample sizes.

2.2. EGovernment Security and PKI

Developing economies are both at an advantage and disadvantage when it comes to technology adoption in Government. They are advantaged because they adopt technologies that have already been tested live in the first world and hence most bugs and teething problems would have been removed or understood. However, they are disadvantaged because developing economies have their own unique socio-economic, socio-cultural and socio-political contexts which require solutions that are customized for them. When reasoning about PKI for instance, each country solution requires a unique technical, policy , legal and regulatory framework developed and customized in the country of implementation [3,16]. The poor ICT infrastructure, low incomes, low literacy on e-business, low trust levels, insecure transaction services, high costs of connectivity etc. present enormous challenges [17]. One big context challenge in developing economies for instance is the entrenched culture of corruption as detailed in the Transparency International (TI) report 2014 showing Nigeria, Kenya etc. ranking very poorly at positions 136 and 145 respectively [18]. Fig. 1 presents an extensible model for eGovernment demonstrating

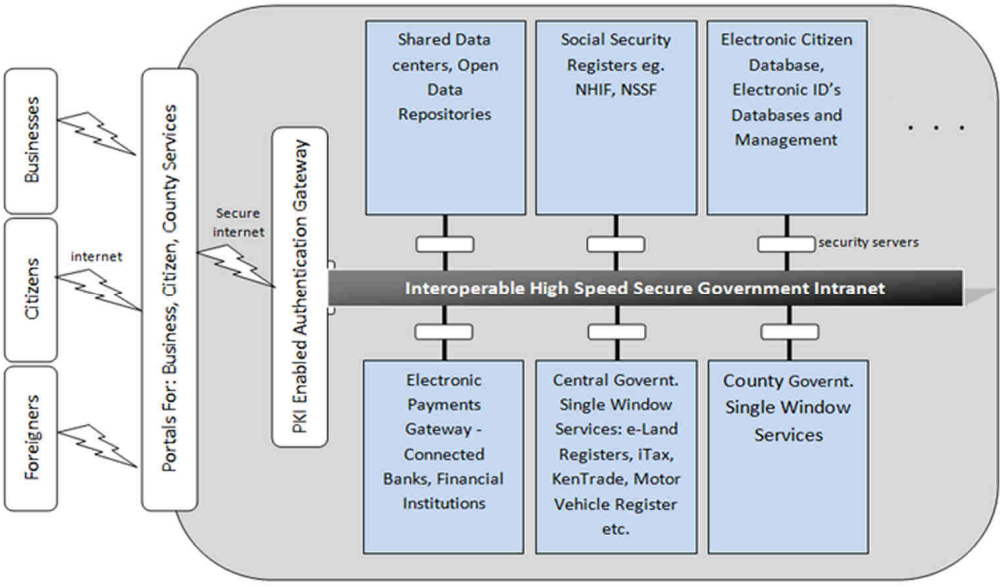

Fig. 1. PKI Enabled e-Government Model.

Estonia [20], Australia [21], Kenya [22] etc. we conclude that the PKI solution should act as the secure gate keeper that identifies and authenticates all parties transacting online by providing an environment that is secure, trustworthy and supports non-repudiation [23]. A PKI enabled gateway makes sure that access to the secure government intranet is only allowed for parties that successfully authenticate using digital certificates. In so doing all government information resources across board are secured. The UK model also has a similar gateway but some of its information resources like for local authorities lie external to the secure intranet [19]. The Estonian model does not have a secure gateway but each agency connects to the common internet called the X-ROAD via a security server. Now that means it is possible to insecurely access the X-ROAD but be kept at bay by individual agency security servers. We propose that an amalgamation of the two though expensive would provide several layers of security that would be difficult to break. This would also enhance user privacy since different agencies require different identity information and a context sensitive smartcard based identity management system running on the intranet and agency servers would enforce it. More recent developments point to the fact that eGovernment is quickly moving towards the cloud [24].

In Kenya, eGovernment and hence cyber security initiatives like PKI rest on a host of development, legal and regulatory frameworks namely:

1. The Kenyan Constitution 2010 which recognizes electronic transactions.
2. The Kenya Vision 2030 which identifies ICT as one of the important foundations for economic development bearing the theme "strengthening the foundation for a knowledge economy" and therefore achieving transformation in the government to make it responsive to the citizen [22].

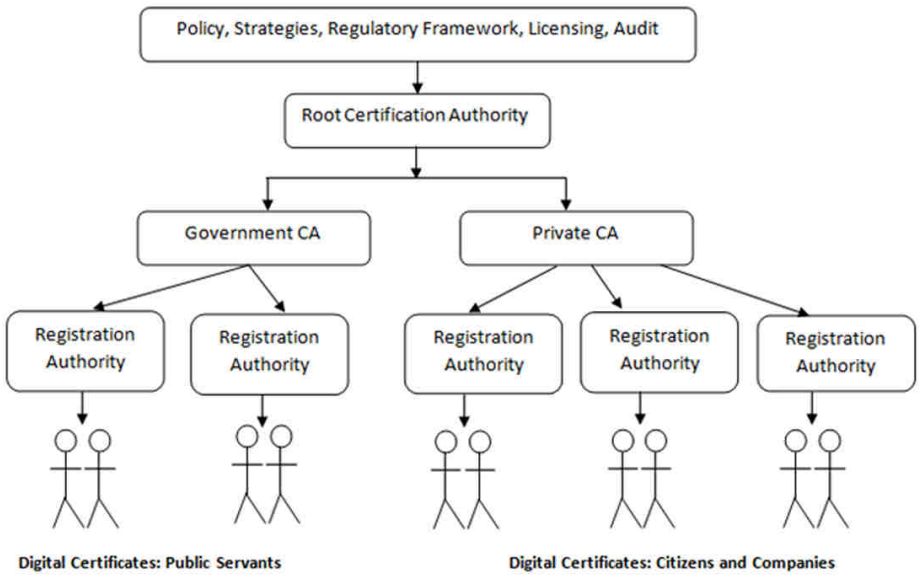

Fig. 2. National PKI Infrastrure, Kenya; Source Communications Authority of Kenya, 2014.

Under this, the Kenya Transparency & Communications Infrastructure Project (KTCIP) sponsored by the World Bank helped set up Kenya's PKI.

3. The National Cyber Security Management Framework which incorporated the
4. Information and Communications Act 1998 and specifically Cap 411A which sets up a legal framework for eCommerce in Kenya. It also amalgamated the Electronic Certification and Domain Name Administration Regulations 2010 which sets relevant conditions that must be met for electronic communications to be authentic and provides for a national PKI.
5. The National Computer Incident Report (NCIR) team.

The Kenyan PKI model closely shadows that of South Korea since the company that won the tender to implement it (Samsung SDS) is from there. The project architecture is as captured in Fig. 2.

2.3. Assessing Security in PKI Solutions

Some generic security threats in PKI as identified by the Australian Government include but are not limited to inappropriate evidence of identity, accidental/deliberate submission of wrong identity documents during initial registration, failure of necessary checks during registration, staff collusion, corrupt CA staff, poor record keeping, data entry mistakes, interception, database corruptions, social engineering attacks on certificate/registration authority (C/RA) staff/help desk, revocation failures, RA spoofing, compromised private key, private key media failure, Relying Party (RP) fails to check revocation status /certificate path, poor infrastructure security [25] etc. The document also gives mitigation measures for the identified threats andvulnerabilities.

Reference [4] presents an assessment model when assessing PKI solutions to ensure interoperable and trustworthy systems as shown in Fig. 3. The model envisions a highly interdependent environment in which the policy body, assessor, assessors accreditation body and PKI accreditation body work in tandem to make sure that the CP, CPS, Standards etc. are applied to the CA's Information Technology (IT) infrastructure and its procedures and operations to

ensure qualitative systems. This paper draws most of its assessment criteria from [4] only that we provide a quantitative way of assessing the key attributes identified other than relying purely on checklists. The ISO 9126 standards [26] and SEI

do not exist [28]. Table 1 identifies the exogenous variables that influence security in the model and some of their indicators mainly drawn from the

PKI x.509 standards [29] and the PKI assessment guidelines [4] and other literature. This forms the

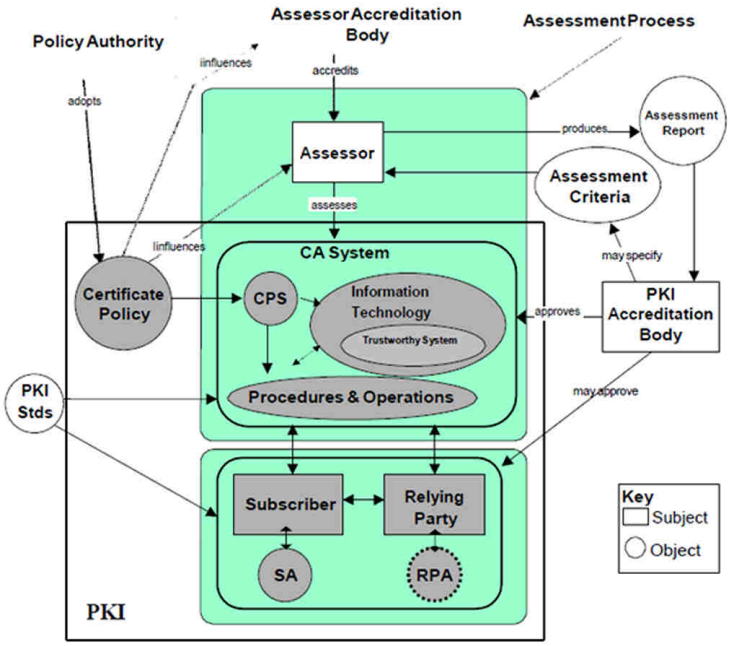

Fig. 3.PKI Assessment Model: Adopted from [4].

standards[26] identify general attributes when assessing software quality.

PKI security is intrinsic to the quality of the Certificate Policy (CP) and the resultant Certificate Practice Statement (CPS) and how strictly they are enforced during planning, implementation and daily management [4]. We now briefly look at each of the variables identified in the CF before moving on to methodology. We recognize the fact that the variables and indicators selected for this model may not be the only ones available, hence the CF is extensible as indicated. Some had to be left out in order to make the study manageable. The major works from which these are drawn include [4,27,25].

Each quality property is a latent variable that has measurement indicators/ attributes. However, when assessing a quality property say security, it becomes necessary to studywhat variables influence it and how they can be measured. To measure the latent variable, indicators are used and they have to be sourced from literature, from current industry practice or empirically where they

basis for coming up with the conceptual framework shown in Fig. 4 from which the PLS-SEM model was drawn.

Just like in [4] we divide our security assessment based on seven key areas:

- Policy, legal and regulatory assessments
- Initial registration controls
- Certificate lifecycle controls
- Management, operation and physical controls
- Technical security controls
- Certificate, CRL and OCSP profiles
- Specification administration.

However, when modeling, we do not specifically structure the model as such because we are more interested in the relevant security variables regardless of from which segment they come from and how they affect the four cornerstones of PKI security, namely

confidentiality, integrity, availability and accountability [30]. The identified variables and theirindicators are tabulated in Table 1. Notice however that in the table, we try to capture where a particular variable falls in the People, Process and Technology model (PPT) since this model has been widely applied in eGovernment studies. However we pitch for the CPPT model with C standing for Culture [31]. The culture variable is very important in developing economies because we argue that however well all the other variables are met, a culture of corruption for example can totally wipe out any gains and totally compromise security of PKI systems. The table translates to Fig. 6, the Conceptual Framework (CS).

A. Personnel Controls

We investigate whether key personnel have the right qualifications to handle their jobs in a trustworthy manner. Other controls include role separation for sensitive tasks and the trustworthinessof the process of engaging contract staff.

B. Culture

The research would investigate if there is a professional code of ethics and whether it is strictly enforced. We shall also find out the perception in terms corruption/nepotism on how tenders, contracts are issued and how staff are employed.

C. Certificate Policy (CP)

A CP is the cornerstone of a PKI. It is usually owned by the root certification authority and is usually drawn from the eGovernment security policy.

A CP is a set of rules which govern the requirements that any PKI participant must meet in order to operate within the PKI and it lays the ground for various CA interoperability.

Table 1. Exogenous variables and their indicators

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

D. Certificate Practice Statement (CPS)

Derived from a CP, a CPS states the practices and procedures that a single CA would use in all its operations.

E. Physical Security Controls

These measures for CAs/RAs/Subscribers try to minimize the risk of key compromise through break-ins, theft, force marjorie, power failures etc.

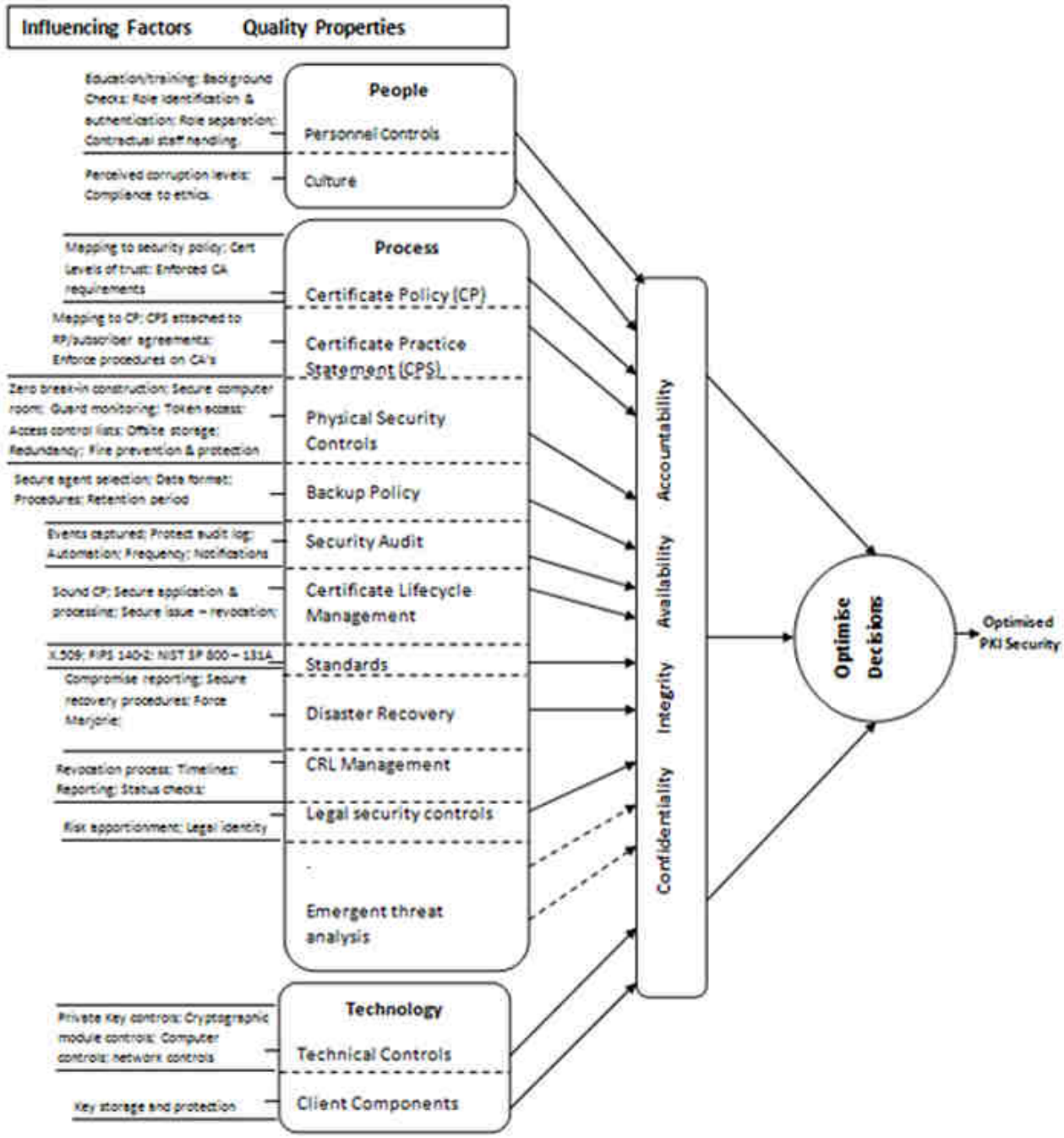

Fig. 4.PKI Security Decision Optimisation Conceptual Framework.

F. Backup Policy

The backup policy is a very sensitive area. If a backup agent is engaged, it is important to have a secure selection process. Also the format in which the data (and especially the keys) are stored is important i.e. plaintext or encrypted. Also, data retention periods, redundancy requirements etc. have to be met.

G. Security Audit

One important consideration is whether the assessors will be in-house or contracted. External auditors are likely to do an unbiased job. Other measurements include protection of the audit log against alteration or destruction since it may have important evidence.

H. Certificate Lifecycle Management

All the processes starting with initial registration, processing, issue, activation, deactivation, revocation etc. of certificates should be done in a secure and trustworthy manner.

İ. Standards

These are very important to any PKI solution. A checklist of the relevant standards will be used to assess adherence to best practice.

J. Disaster Recovery

Readiness to deal with disastrous events that can bring the PKI to its knees is very important. Good provisions forcompromise reporting tested and tried recovery procedures after disasters etc. need to be assessed.

K. CRL Management

Certificate Revocation Lists (CRL) management is very important and can prove very costly if not handled properly. The reporting process when revocation is required should be secure e.g. who requests for a revocation and does the revocation messages have to be digitally signed? Also, strict reporting timelines are important.

L. Legal Security Controls

The legal, policy and regulatory framework should be sufficient in order to deal with difficult scenarios like risk apportionment, potential liability management, indemnity, legal responsibilities etc. for all players like CA, RA, RP, Subscriber, repositories etc.

M. Technical Controls

These are a raft of assessments that would touch on a wide range of technical concerns like the logical security of the private key, security of the cryptographic module, computer and network controls.

N. Client Components

The term client here mainly refers to relying parties and citizens. The concern here is mainly on how the private key is stored, is it on the client computer or in a smartcard? How is the private key generated and or passed to the client after generation? Is the process secure? This is important sincemost cases of key loss and or compromise may emanate from this end.

2.4. Partial Least Squares Structural Equation Modeling (PLS-SEM)

Partial Least Squares Structural Equation Modeling (PLS-SEM) is an extension of the multiple linear regression analysis technique [32]. 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 Fig. 2 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 (1):

$$S= C_0+\beta E+\varepsilon \quad (1)$$

Where: *S* = salary of the expert (called the dependent or endogenous variable); C_0 is baseline/constant earning with zero education; $\beta$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 influenceearnings since there is no strict linearity displayed. Other unaccounted for factored like experience, productivity etc. 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 [33]. If we set $\varepsilon = 0$ as in most cases, then the regression equation becomes the equation of a straight line in a 2-dimensinal plane with C0 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).

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

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 βa 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 Fig. 5.

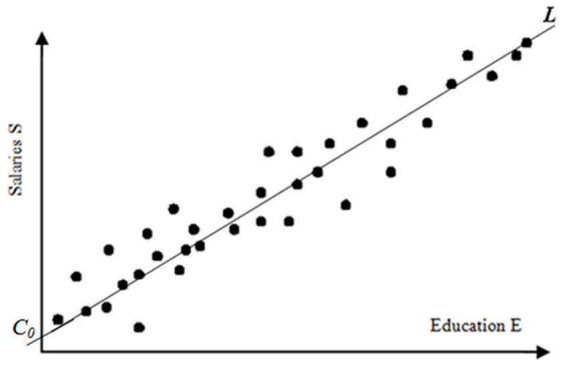

Fig. 5. Scatter chart of S/E

One way of achieving this is selecting the line that has the minimum sum of square errors. We now move on to PLS-SEM.

Structural Equation Models, also called simultaneous equation models are multivariate or multiple linear regression analysis models [34]. 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). $\gamma$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 C_0, β and $\gamma$is nolonger 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 [35]. 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 [6]. The O observations described by D dependent variables are stored in an O×Dmatrix 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 [36].

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 relationship arrows point away from it.

Endogenous variables: dependent variables. Path arrows point to it showing causal effects.

Indicators: observed measures or variables used to infer the value of the latent variable.

Diagrammatically, a model takes the form of Fig. 6 [6].

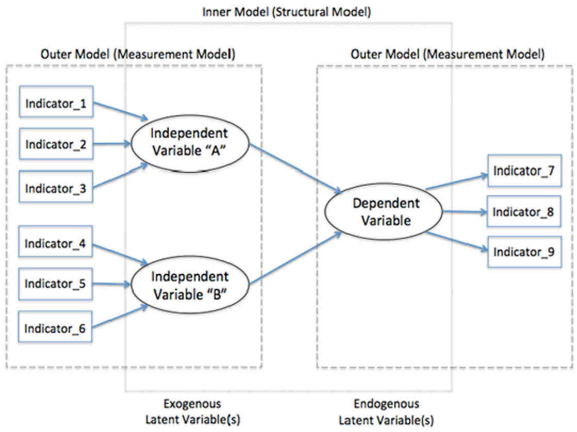

Fig. 6. Structural Equation Model

2.5. **Methodology**

Literature review was done using the structure case strategy [37]. This helped the research to distill key PKI security quality attributes and develop the CF. The CF was then directly modeled in SmartPLS as shown in Fig. 7. Each quality property in the CF becomes a latent variable in SmartPLS. The influencing factors become indicators in a reflective relationship. Intermediate variables of confidentiality, integrity, availability and accountability become endogenous variables to which all arrows from exogenous variables point, representing various influences. The last endogenous variable in the chain OPTIMAL_DECISION_FIT represents the collective state of PKI security after assessment.

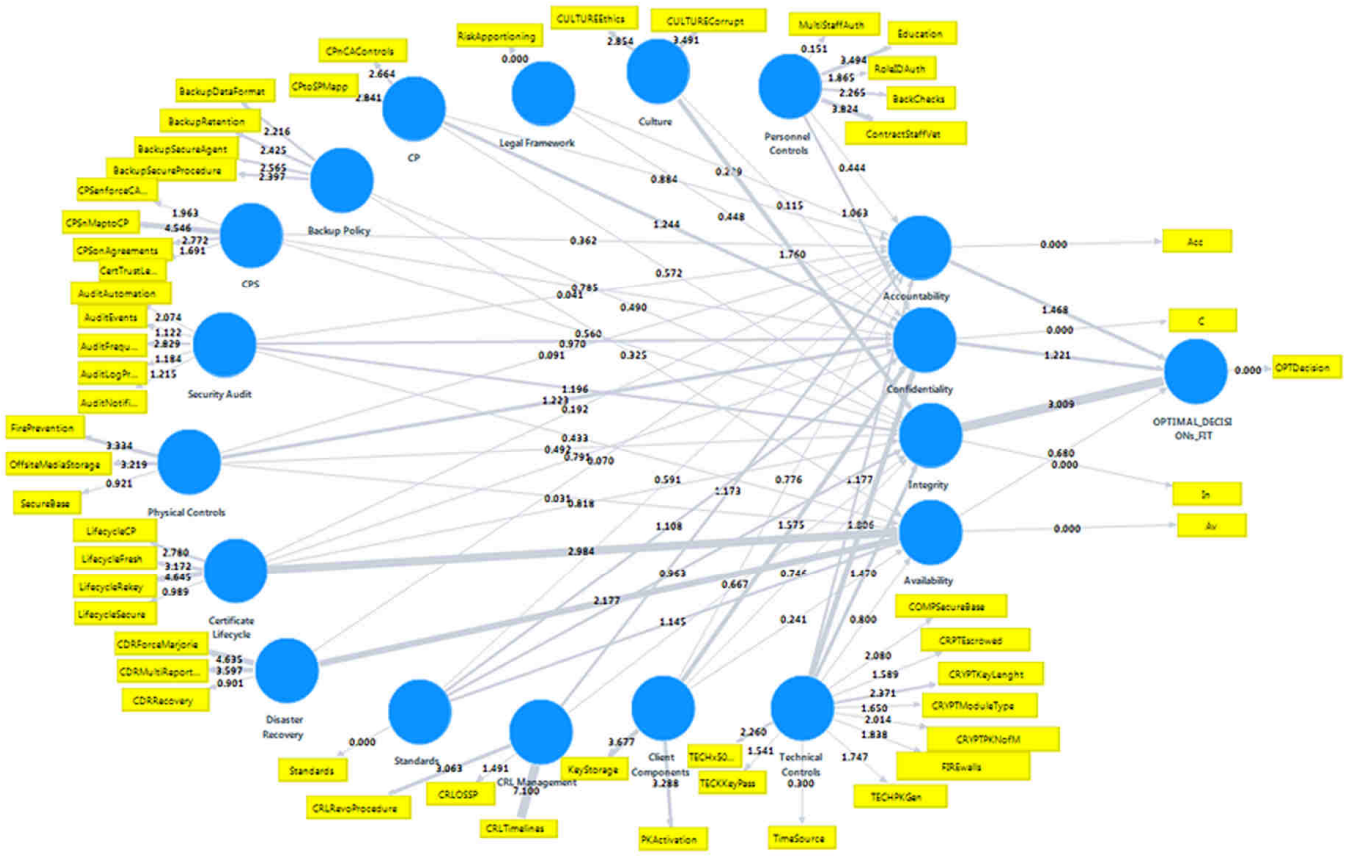

Fig. 7. Model in SmartPLS: results of post study bootstrapping after 300 iterations

All indicators are ideally translated into a questionnaire or interview question with interval measures on the responses. We use a seven level Likart scale system with the worst case scenario scoring the least (1) while the best the most (7). Where we have used No/Yes field, the No is scored as 1 and the Yes as a 7. For example, under technical controls on the CF, to measure the attribute Cryptographic Module Controls, one of the indicators is the key length (CRPTKeyLenght) as modelled in SmartPLS. The questionnaire question was presented as follows:-

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

☐ 64 bits ☐ 128 bits ☐ 512 bits ☐ 256 bits

☐ 1024 bits ☐ 2048 bits ◉ >2048 bits

In this case we assume 64 bits to be the worst case scenario while >2048 bits the best. The responses were coded into interval values 1 – 7 and captured in an Excel file which was used to populate the model. Initially, we carry out a baseline survey to establish the level of attainment of various attributes. We initially use a sample size of 30 to collect data about the various attributes. Then after three months, we collect data again in a post study to see whether the weaknesses identified earlier have been improved by comparing various statistical measures generated. We use a reflective measurement scale because we assume the indicators have correlations among themselves.

3. Calculations

A. Bootstrapping Algorithm

Figure 7 shows the results of the bootstrapping algorithm after 300 iterations and the significance level set at 0.05. The values on the causal lines between variables in the inner model represent the t-values. Looking at the model, we can say for example that integrity, accountability and confidentialityare very significant factors when coming up with optimized decisions due to their high t-values. However, Availability-> Optimal Decisions Fit has the lowest t-value among the three (0.680). Table 2a and 2b below shows a comparison between the p and t-values of some of the indicators during the baseline and post study. We could not fit the entire table because of the limited space. In Table 2a, there was an improvement in the indicator values while in Table 2b there was a decline.

Table 2a. Examples of indicators that showed improvement.

	BASELINE			POST-TEST	
Indicator	**t-value**	**p-value**		**t-value**	**p-value**
AuditAutomation	0.060	0.952		2.151	0.032
BackChecks	0.660	0.509		1.398	0.162
BackupSecurePro cedure	1.487	0.137		2.445	0.015

Table 2b. Examples of indicators that showed decline.

	BASELINE			POST-TEST	
Indicator	**t-value**	**p-value**		**t-value**	**p-value**
Auditnotification	4.735	0.000		1.024	0.306
CRLOCSP	3.868	0.000		0.840	0.401
FIREwalls	2.292	0.022		1.169	0.243

Explanations for improvement of the **t** values of indicators can be found in the fact that in some cases e.g. AuditAutomation (the level of automation in collection of system audit data); the baseline data was scanty and incomplete. Although the improvement may be argued to be a false impact, it is worthwhile to note that at least the model was able to capture and measure any anomalies and represent the true position when complete data was entered. In Table 2b, the CRLOCSP (whether checking CRLs uses the online certificate status protocol) declined because although the baseline established that OCSP is implemented in the PKI, the post study detected that it is OCSP without stapling. The study therefore provided a recommendation for the managers to consider implementing stapling for a more efficient PKI.

B. Composite Reliability

After running the PLS algorithm, Figure 8a shows the composite reliability of the data collected during the baseline survey while 8b shows that of the post study data. Notice that it is easy to notice the improvement in the data's reliability indicating more consistency in the

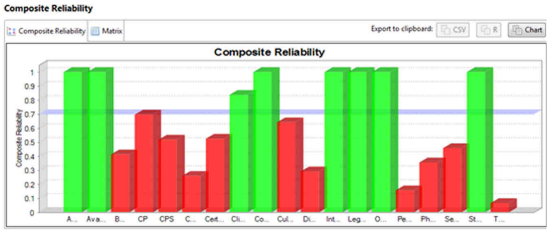

Fig. 8a. Composite reliability pre-study

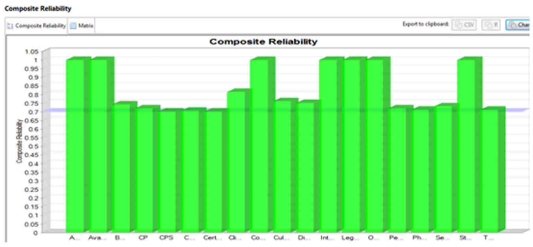

Fig. 8b. Composite reliability post study

positive responses of the respondents regarding the state of implementation of the various security attributes. In Figure 8b, all the attributes have attained the target value 0.7 and above hence we can conclude that the PKI is in a healthy state.

C. R^2 Value

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 = Explained Variation / Total Variation(4)

The R^2 value of the OPTIMAL_DECISION_FIT post study (30.4%) is higher than that of the baseline study (24.6%) as shown in Figure 9a and 9b respectively. This

indicates some improvement in the total security of the PKI solution since the exogenous variables have increased their total effects on the optimal decisions.

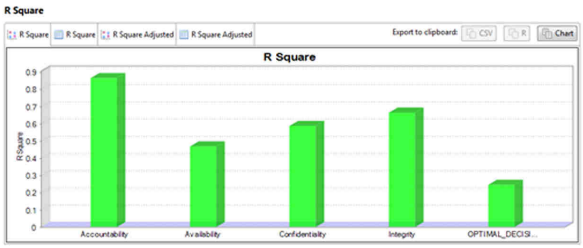

Fig. 9a. R^2 of OPTIMAL_DECISION_FIT baseline

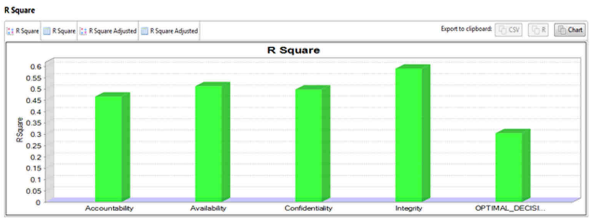

Fig. 9b. R^2 of OPTIMAL_DECISION_FIT post study

4. Conclusions

PLS-SEM is a flexible method for modeling variables, their relationships and performing predictions. When used in assessing and optimising PKI security, it can be used to capture all relevant latent variables together with their indicators to come up with a structural model which can be used to optimise rational decision making. Ideally, the analysis of variance leads the assessor to answer important questions that help to enhance variables that seem to fall below expected values.

PLS-SEM is a flexible method for modeling variables, their relationships and performing predictions. When used in assessing and optimising PKI security, it can be used to capture all relevant latent variables together with their indicators to come up with a structural model which can be used to optimise rational decision making. Ideally, the analysis of variance leads the assessor to answer important questions that help to enhance variables that seem to fall below expected values.

5. Acknowledgements

The author thanks Professor Okelo-Odongo, Esther and Evans at the ICT Authority, Kenya, Jomo Kenyatta University of Agriculture and Technology, The National Council of Science and Technology, Kenya for their support.

6. References

[1] Ernst & Young. Identity and Access Management: Beyond Compliance. Technical report. Ernst & Young, http://www.ey.com/Publication/vwLUAssets/EY_-_Evolving_identity_and_access_management/$FILE/EY-Evolving-identity-and-access-management.pdf, 2013.

[2] R. Wagner, “Identity and Access Management: Key Initiative Overview.” Gartner Inc., 2010.

[3] T. Smedinghoff, “Building an Online Identity Legal Framework: The Proposed National Strategy,” The Bureau of National Affairs, USA, Report 800-372-1033, 2010.

[4] ISC, “PKI Assessment Guidelines.” Information Security Committee, American Bar Association, 2003.

[5] C. M. Ringle, S. Wende, and J.-M. Becker, “SmartPLS 3,” A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), 2014. [Online]. Available: http://www.smartpls.com. [Accessed: 03-Jan-2015].

[6] K. K.-K. Wong, “Partial Least Squares Structural Equation Modeling (PLS-SEM) Techniques Using SmartPLS,” Marketing Bulletin, vol. Technical Note, no. 1, 2013.

[7] N. Merlo-Schett, M. Glinz, and A. Mukhija, “COCOMO,” presented at the Seminar on Software Cost Estimation, Zurich, Switzerland, 2002.

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

[9] G. Chemwa, “Electronic Identity and Access Management for E-Government: Optimising Public Key Infrastructure Initiatives Through Probabilistic Assessment of Quality Attributes,” in Proceedings of the 2014 JKUAT Scientific, Technological and Industrialisation Conference, Nairobi, Kenya, 2014, pp. 542–551.

[10] R. Kazman, M. Barbacci, M. Klein, S. J. Carrière, and S. G. Woods, “Experience with Performing Architecture Tradeoff Analysis,” in Proceedings of the 21st International Conference on Software Engineering, New York, NY, USA, 1999, pp. 54–63.

[11] V. Veerappa and E. Letier, “Understanding Clusters of Optimal Solutions in Multi-objective Decision Problems,” in Proceedings of the 2011 IEEE 19th International Requirements Engineering Conference, Washington, DC, USA, 2011, pp. 89–98.

[12] A. Brady and T. Menzies, “Case-Based Reasoning vs Parametric Models for Software Quality Optimisation.” ACM, 2010.

[13] E. Kurilovas and V. Dagiene, "Multiple Criteria Evaluation of Quality and Optimisation of e-Learning System Components," Electron. J. E-Learn., vol. 8, no. 2, pp. 141–151, 2010.
[14] M. Harman, U. Ph, and B. F. Jones, "Search-Based Software Engineering," Inf. Softw. Technol., vol. 43, pp. 833–839, 2001.
[15] S. Yoo, M. Harman, and S. Ur, "Highly Scalable Multi Objective Test Suite Minimisation Using Graphics Cards," in Search Based Software Engineering, M. B. Cohen and M. Ó. Cinnéide, Eds. Springer Berlin Heidelberg, 2011, pp. 219–236.
[16] T. Waema and E. Adera. Local Governance and ICT's in Africa: Case Studies and Guidelines for Implementation and Evaluation. Pambazuka Press UK, 2011(Book).
[17] A. Ntoko, "e-Business: A Technology Strategy for Developing Countries," e-Business:A Technological Strategy for Developing Countries, 2010. [Online]. Available: http://www.itu.int/ITU-D/cyb/publications/archive/wmrcjune00/ntoko.html. [Accessed: 17-Feb-2015].
[18] Transparency International, "Corruption Perception Index 2014," Transparency International, Berlin, Germany, Report, 2014.
[19] J. Satyanarayana, E Government: The Science of the Possible. PHI Learning Pvt. Ltd., 2004.
[20] A. Kalja, N. Reitsakas, and N. Saard, "eGovernment in Estonia: Best Practices," Technol. Manag. Unifying Discip. Melting Boundaries, pp. 500–506, 2005.
[21] Australian Government, "National e-Authentication Framework." Australian Government Infromation Management Office, 2009.
[22] GoK, "The Kenya National ICT Master Plan 2013/14 - 2017/18." ICT Authority, 2014.
[23] GoK, "Cybersecurity Strategy." Ministry of Information and Communication, 2014.
[24] R. K. Das, S. Patnaik, and A. K. Misro, "Adoption of Cloud Computing in e-Governance," in Advanced Computing, N. Meghanathan, B. K. Kaushik, and D. Nagamalai, Eds. Springer Berlin Heidelberg, 2011, pp. 161–172.
[25] Australian Government, "Gatekeeper PKI Framework Threat and Risk Assessment Template," Australian Government Information Management Office, 2009.
[26] D. Zubrow, "Software Quality Requirements and Evaluation, the ISO 25000 Series." Software Engineering Institute, Carnegie Mellon, 2004.
[27] Accenture, "Citizen Identity Authentication Programs: Challenges and Benefits." 2008.
[28] P. Johnson and M. Ekstedt, Enterprise Architecture Models and Analyses for Information Systems Decision Making. Stockholm: Studentlitteratur, 2007.
[29] Cooper et al., "RFC 5280: Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile." The Internet Society, 2008.
[30] E. Wheeler, Security Risk Management: Building an Information Security Risk Management Program from the Ground Up. Elsevier, 2011.
[31] P. Buttles-Valdez, A. Svolou, and F. Valdez, "A Holistic Approach to Process Improvement Using the People CMM and the CMMI-DEV: Technology, Process, People & Culture, The Holistic Quadripartite," presented at the Software Engineering Institute Tutorial, CarnegieMellon, 2005.
[32] StatSoft, "Partial Least Squares (PLS)," Partial Least Square, 2000. [Online]. Available: http://www.uta.edu/faculty/sawasthi/Statistics/stpls.html. [Accessed: 18-Feb-2015].
[33] A. O. Skyles, "An Introduction to Regression Analysis." The University of Chicago Law School, 1992.
[34] J. Fox and H. S. Weisberg, An R Companion to Applied Regression, Second Edition edition. Thousand Oaks, Calif: SAGE Publications, Inc, 2010.
[35] R. E. Schumacker and R. G. Lomax, A Beginner's Guide to Structural Equation Modeling: Third Edition. Routledge, 2012.
[36] University of North Carolina, "Variance and Design of Experiments," Variance, 2007. [Online]. Available: http://www.unc.edu/courses/2007spring/psyc/530/001/variance.html. [Accessed: 27-Feb-2015].
[37] J. M. Carroll and P. A. Swatman, "Structured-case: a methodological framework for building theory in information systems research," Eur. J. Inf. Syst., vol. 9, no. 4, pp. 235–242, Dec. 2000.
[38] J. Hulland, "Use of partial least squares (PLS) in strategic management research: a review of four recent studies," Strateg. Manag. J., vol. 20, no. 2, pp. 195–204, Feb. 1999.
[39] R. P. Bagozzi and Y. Yi, "On the evaluation of structural equation models," J. Acad. Mark. Sci., vol. 16, no. 1, pp. 74–94, Mar. 1988.