

**INTEGRATED AND AUTOMATED MOBILE DATA
INFORMATION SYSTEM IN STATISTICS**

BY

CHENELA NKHOWANI

A Dissertation submitted to the University of Zambia in partial
fulfilment of the requirement of Degree of Master of Engineering in
Information Communications Technology

THE UNIVERSITY OF ZAMBIA

LUSAKA

2021

COPYRIGHT DECLARATION

All rights reserved. No part of this dissertation may be reproduced or stored in any form or by any means without permission in writing from the author or the University of Zambia.

DECLARATION

I Chenela Nkhowani, the undersigned hereby declare that this dissertation represents my own work and has never been submitted for any other qualification at this or any other University or Institution. All sources that I have used or quoted have been acknowledged by complete references.

Author: Chenela Nkhowani

Signed:

Date:

I, **Dr Charles Lubobya**, having being the supervisor hereby declare that I supervised this dissertation and is the true work for the undersigned student herein called Chenela Nkhowani and has not been submitted for any assessment or any other purposes in any other institution. All sources that have been used or quoted have been acknowledged by complete references.

SUPERVISOR: Dr Charles Lubobya

Signed:

Date:

I, **Abraham Ngoyi**, having being the co-supervisor hereby declare that I supervised this dissertation and is the true work for the undersigned student herein called Chenela Nkhowani and has not been submitted for any assessment or any other purposes in any other institution. All sources that have been used or quoted have been acknowledged by complete references.

CO-SUPERVISOR: Abraham Ngoyi

Signed:

Date:

CERTIFICATE OF APPROVAL

This dissertation of Chenela Nkhowani is approved as fulfilling the partial requirements for the award of the Degree of Master of Engineering in ICT by the University of Zambia (UNZA).

Examiner 1:Signature:Date:

Examiner 2:Signature:Date:

Examiner 3:Signature:Date:

Chairperson :Signature:Date:

ABSTRACT

The swift expansion of Mobile Phone Technology in the developing world has created rapid development and deployment of a wide array of Potential Applications, including Mobile Banking, Data Collection, Patient Tracking, and Behavioural Interventions which has allowed many Organizations to Monitor Performance, Making Decisions about our Programs, and report on our progress cheaper, faster, and better.

The basis of the study is to create an Electronic Mobile Data Application to be used in collecting data electronically via Cspro Application. The system will be designed to allow Enumerators to log into the System using the User ids to collect data and send it to the server.

The system consists of the User Menu, which is made up of actions that one needs to do namely Data Collector, Send Data/Update System, and Logout and Exit the program.

During the data collection the Enumerator first needs to Log into the System by entering the User Name and the password. The system checks if the userid is in the system to grant accordingly. The user menu will open where they will be asked what they want to do. The samples are already entered in the system.

Depending on what cluster number has been pick in a particular province, all the Identification particulars such as Province, District, Constituency, Ward, Region, Census Supervisory Area (csa) and Enumeration area (ea) would be prefilled.

Using a detailed literature review, the results of the study confirmed that the time it takes to analyse and write reports is reduced in Electronic Data Collection. The study identified errors and inconsistencies which, was resolved at interview time with the help of people who had knowledge of the data they were collecting, The data was more accurate in the questionnaire due to automated skip patterns and mandatory items which are programmed in the system for instance, if someone is aged 11 and the marital status asks people who are 12 years then this person was skipped because this was not applicable to the question. It was discovered also that Mobile Data Application also improved field management and Real Time Monitoring of Enumeration Activities as data is sent to the Central location.

Based on the findings Organisations need to adopt Mobile Data Application, which as shown reduces errors during data capture when transcribing data from paper to mobile device and Boost field data capture productivity.

Keywords: Mobile Application, Cspro Application, Real Time Monitoring, Monitor Performance.

DEDICATION

To my husband Costain, am so grateful for your love and support. To my children, Samson, Jackson and Mutinta, you gave me the inspiration to work hard and achieve this not so easy task. To my family, I thank you for your patience and understanding. To the Lord God Almighty for everything.

ACKNOWLEDGMENTS

I would like to acknowledge God, for keeping me in good health throughout the research study, and for keeping my mind focused during the long hours days and nights, spent on searching and researching. My sincere gratitude goes to Mr. Abraham Ngoyi and Dr Charles Lubobya and for the guidance and encouragement given to me during the process of working on this research. The work seemed to be impossible in my eyes, but through their valuable support it became possible.

Finally, I would like to express my very profound gratitude to my husband, children, parents, brothers and sisters for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

ACRONYMS

NSO	National statistical office
CAPI	Computer assisted person interview
MDP	Mobile digital platform
MCC	Mobile cloud computing
GPS	Global positioning system
CSPRO	Census and survey processing system
IBM	International Business Machines
IOS	Operating system
MDC	Mobile data collection
EMIT	Enzyme multiplied immunoassay technique
CSO	Central statistical office
CSA	Census supervisory area
EA	Enumeration area
ODK	Open Data Kit
EDC	Electronic data collection
PDA	Personal digital assistant
ERB	Energy regulation board
PAPI	Paper and pencil interview
AE's	adverse events
SPSS	Statistical Package for Social Sciences
CSENTRY	Cspro data entry
ZAMSTATS	Zambia Statistics Agency

TABLE OF CONTENTS

COPYRIGHT DECLARATION	i
DECLARATION	ii
CERTIFICATE OF APPROVAL	iii
ABSTRACT.....	iv
DEDICATION	v
ACKNOWLEDGMENTS	vi
ACRONYMS.....	vii
LIST OF FIGURES	x
LIST OF TABLES	x
CHAPTER ONE	1
INTRODUCTION	1
1.1 Introduction.....	1
1.2 Background to the Study	1
1.3 Statement of the Problem	2
1.4 Aim of the Study	3
1.5 Objectives of the study	3
1.6 Research Questions	4
1.7 Significance of the study	4
1.8 Scope of the Study.....	4
1.9 Organisation of the Dissertation.....	5

1.10	Summary.....	5
CHAPTER TWO		6
LITERATURE REVIEW		6
2.1	Introduction.....	6
2.2	History of Electronic Mobile Application.....	6
2.3	Types of Mobile Applications.....	7
2.3.1	Native application	7
2.3.2	Hybrid Applications.....	8
2.3.3	Web applications.....	9
2.4	Integrated Computerised Case Management.....	10
2.4.1	What is data collection	10
2.5	Mobile Data Collection	11
2.6	Rate of Adoption of Mobile Data Collection in Developing Countries.....	12
2.7	Cloud Computing	13
2.8	Data Integrity.....	15
2.9	Authorization.....	16
2.10	Related Works	16
2.11	Chapter Summary	24
CHAPTER THREE		25
METHODOLOGY		25
3.1	Introduction.....	25

3.2	Research Design.....	25
3.3	Area of Study	25
3.4	Study Population	25
3.5	Sample size and sampling technique.....	26
3.6	Data Collection Methods.....	26
3.7	Data Processing, Analysis and Presentation	27
3.7.1	Primary Data Analysis	27
3.8	System Automation.....	27
3.9	Ethical Consideration	28
3.10	Limitations of the Study	28
3.11	Summary.....	29
CHAPTER FOUR.....		30
ANALYSIS OF STUDY FINDINGS AND DISCUSSION.		30
4.1	Introduction.....	30
4.2	Socio-Demographic Data	30
4.3	Challenges of Paper and Paper Interviews (PAPI).....	31
4.3.1	Delay in results	31
4.3.2	Dissemination Period	32
4.3.3	Skip Patterns	32
4.4	Benefits of the CAPI	33
4.4.1	Data Accuracy.....	33

4.4.2	Productivity Gain	33
4.4.3	Cost-Effectiveness	34
4.5	Implementation of a New automated mobile data collection application system.....	34
4.5.1	Current System.....	34
4.6	Results from Installation on the CAPI	37
4.7	Research Instrument.....	42
4.7.1	CAPI Menu Code.....	42
4.8	Saving of Cases	44
4.9	Synchronizing and Sending of Data.....	45
4.10	Retrieving data from the Server.....	48
4.11	Tabulation of Data and Report Writing.....	49
4.12	Testing	51
4.13	Summary.....	52
CHAPTER FIVE		53
CONCLUSIONS AND RECOMMENDATIONS		53
5.1	Introduction.....	53
5.2	Conclusions.....	53
5.3	Recommendation.....	54
REFERENCES		55
APPENDIX A: Data collection permission letter.....		61
APPENDIX B: System prototype code.....		65

APPENDIX C: International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Certification	86
---	----

LIST OF FIGURES

Figure 1.1-1: National Statistical System	3
Figure 2.1-1: Example of a native application (Wei, Lei, 2017)	7
Figure 2.1-2:Hybrid application (Abusaimeh, Hesham, 2009).....	8
Figure 2.1-3:web application (Cho, Ying-Chiang, 2017).....	9
Figure 2.1-4: Cloud Computing Architecture[42]	14
Figure 2.1-5: cloud Service Model (Chandrashekhar , 2017).....	15
Figure 4.1-1: Distribution of Drawbacks of Using Paper Questionnaires	31
Figure 4.1-2: Distribution of Period of Dissemination using PAPI.....	32
Figure 4.1-3: Skip patterns are difficult with PAPI	33
Figure 4.1-4: Skip patterns are difficult with PAPI	33
Figure 4.1-5:Paper Questionnaire	34
Figure 4.1-6: Paper Questionnaire Flow Chart.....	35
Figure 4.1-7: Integrated System.....	36
Figure 4.1-8b: Diagram of mobile data system	36
Figure 4.1-9: Workflow of an integrated System	37
Figure 4.1-10: CSEntry Menu.....	38
Figure 4.1-11: CSEntry Settings	39
Figure 4.1-12: Installation of the Application.....	40
Figure 4.1-13: Installed CSEntry ApplicatioSsource: Research Data	41
Figure 4.1-14:Terms of Service for Cspro	42
Figure 4.1-15: CAPI Menu Application	44
Figure 4.1-16: Data Entry into the System	44
Figure 4.1-17: Data and update application.....	46

Figure 4.1-18: Connection to the Server.....	46
Figure 4.1-19: Data Being put to the Server	47
Figure 4.1-20: Synchronisation of Data.....	47
Figure 4.1-21: Updating of Application.....	48
Figure 4.1-22: Successfully Updated.....	48
Figure 4.1-23: Data Retrieval from the Server	49
Figure 4.1-24: Data Ready for Tabulation and Analysis	50

LIST OF TABLES

Table 2.1.1: Popular platforms for mobile data collection	11
Table 4.1.1: Socio-Demographic Data.....	30
Table 4.1.2: Implementation of Testing Evaluation of the Prototype.....	51

CHAPTER ONE

INTRODUCTION

1.1 Introduction

This chapter discusses the background of the study, the problem statement and the aim of the research is also covered. Moreover, It outlines the research objectives and research questions followed by the significance, scope. Finally, the organization of the thesis and a summary of the chapter are presented.

1.2 Background to the Study

National Statistical Offices (NSOs) the world over are charged with the pivotal responsibility of compiling, processing and disseminating statistics on a wide array of subjects. Of primary concern is the quality of the data and the need for the NSO to provide plausible information and knowledge. Statistical undertakings have over the years been conducted through manual pen and paper interviewing, but in the recent past the advent of electronic Computer Assisted Personal Interviewing (CAPI) revolutionised the way in which statistical data is collected (De leeuw 2008). This is but one of the improvements most NSOs have made to modernise their business process and advance their business strategy. The use of mobile technology seems to be effective as an enabling technology for resource limited settings such as in the developing world. Kenya, an East African nation, has set standards of mobile phone use on the continent (Shao D 2012). Being the fastest growing group of mobile technology users on the continent, 92% of Kenyans use their mobile phone to go online (kelvin onoka 2017). The swift expansion of mobile phone technology in the developing world has created rapid development and deployment of a wide array of potential applications, including mobile banking, data collection, patient tracking, and behavioural interventions. The Zambia Information Communication Telecommunication Agency (ZICTA) estimates that the mobile market in the country has maintained an upward trend with an increase in the number of mobile subscriptions of 7.7 million in 2019 representing a 47.08 percent penetration rate. Technology has allowed many organizations to monitor our performance, make decisions about our programs, and report on our progress cheaper, faster, and better.

Mobile devices such as smart phones and tablets offer great flexibility and improved productivity for mobile users. The use of mobile devices has increased greatly in recent years

and this use is predicted to continue to rise (Harrison et al 2013). Using mobile devices for data collection provides a snapshot of individuals' everyday perceptions, experiences, and interactions with their environments. The use of mobile phones for the assessment of individuals' daily lives is not a new research method (Fahrenberg et al., 1996). However, because smartphones have now become so widespread throughout the population, are low in cost, and are equipped with sensor technology and ready for data collection through apps (Miller, 2012; Cartwright, 2016; Harari et al., 2016; Beierle et al., 2018a).

Having a clear understanding of the challenges associated with pen and paper personal interviews (PAPI) is therefore key to managing the automated mobile data collection system that would deliver accurate data, which in turn supports the needs of the business. However, Zambia Statistical Agency is still using the PAPI in some of its programs. This research therefore intended to explore ways in which PAPI could be improved through the development of an integrated and automated mobile data collection application system. This is expected to help overcome these challenges, which are high cost of printing and dispatch, large amount of work involved in transcribing the data and low response rate and deliver the data with such accuracy and on time.

1.3 Statement of the Problem

Getting data of good quality on time has always been a concern. The quality of research data determines readiness for decision-making (Chen H, Hailey D, Wang N, Yu P 2014). The advent of electronic data capture has revealed that paper-based data collection can actually be costly and time consuming (Asare-aboagye Y 1995).

While organizations and projects are using mobile phone technology to collect information such as agriculture and crop information, employment statistics and market prices, natural disaster information, census, opinion polls from the field, Mobile phone technology makes it possible to access information and interact with the databases (Lehr, 2007). The development of this technology is based on a technology which uses widely available mobile technology gadget. Despite an upward trend in demand for Internet services and reduced cost of Internet enabled mobile devices (GSMA report 2019), Zambia Statistics Agency still has challenges in the collection of quality data in good time. Therefore, this is a gap that needs to be explored because there is no much published research or existing research on a similar subject for Zambia. Therefore, this study attempted to develop an integrated and automated mobile data collection application system that would improve on the quality data in good time.

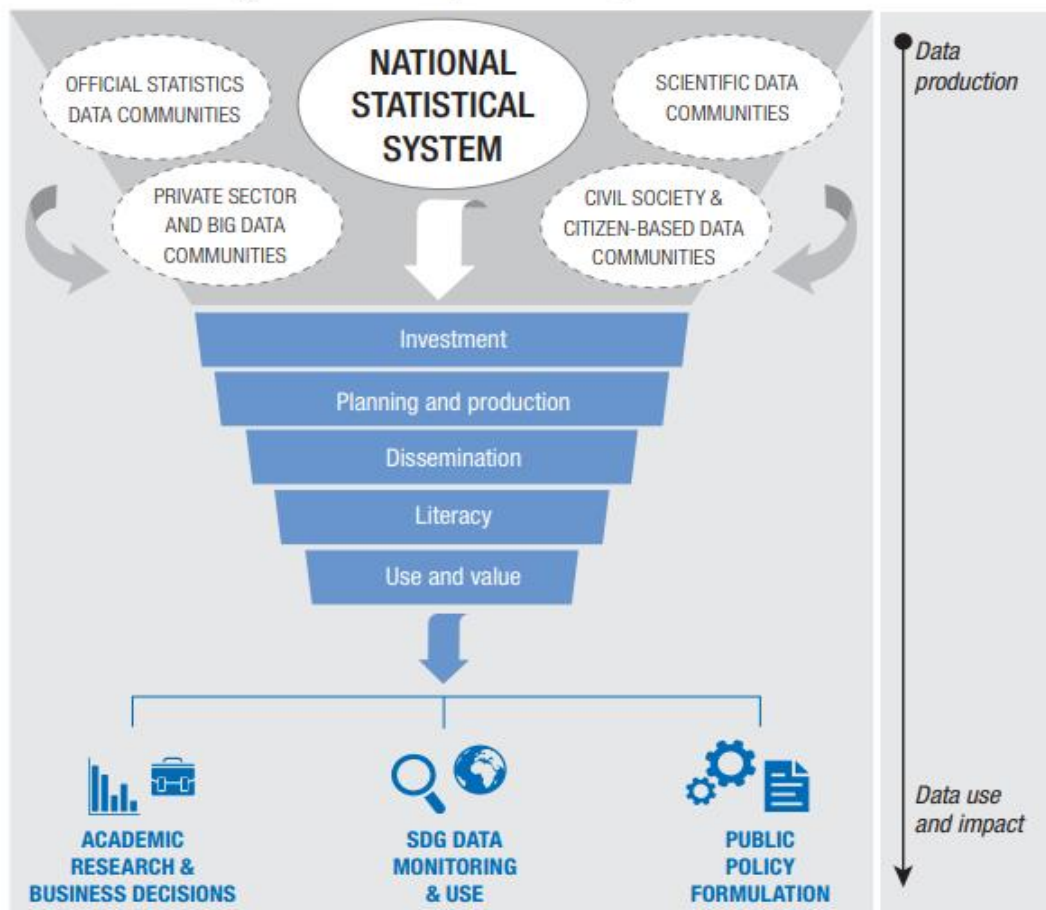


Figure 1.1-1: National Statistical System

1.4 Aim of the Study

The aim of this study was to design and a develop an integrated and automated mobile application and implement the designs into a functional program to improve the data collection process at ZamStats.

1.5 Objectives of the study

The objectives of the study were as follows:

- (i) To investigate the challenges faced by Zambia Statistics Agency in the process of data collection using PAPI;
- (ii) To establish the benefits of using an integrated and automated mobile data collection system; and
- (iii) To develop and design an integrated and automated mobile data information system that could be used to improve the data collection process.

1.6 Research Questions

The research questions were as follows:

- (i) What challenges does Zambia Statistics Agency face when collecting data using PAPI?
- (ii) What are the benefits of using an integrated and automated mobile data collection process?
- (iii) Can an integrated and automated mobile data collection system be used to improve the data collection process?

1.7 Significance of the study

The Zambia Statistics Agency (ZamStats) will benefit greatly from this study as it will provide ways in which data can be collected with minimum cost and in real time. There is little research done on the development of an integrated and automated mobile applications on data collection process in Zambia. The available literature is too general and does not address the specific areas of the subject matter. Henceforth this study will be beneficial to academicians and researchers in that it will contribute to the body of knowledge on the subject matter.

1.8 Scope of the Study

The research investigated how mobile electronic application effectively addressed factors that have made the existing data collection system ineffective in the context of Accuracy, timeliness, data integrity and productivity gain. The research limited its scope to designing and developing an integrated and automated mobile application system.

1.9 Organisation of the Dissertation

This report has six chapters; in the first chapter is introduction of the study, the statement of the problem, the aim and objective as well as the research questions, significance and scope of study.

In chapter two, there is a review of the relevant literature and in chapter three there is a discussion on the study methodology used in order to address the research problem and a justification for adopting the used methodology.

In chapter four, the study highlights the research findings.

Chapter five discusses the findings in answering to the research questions.

The last chapter which is chapter six, draws conclusions and makes recommendations.

1.10 Summary

This section discussed the introduction of the study covering the study background, statement of the problem. The chapter further looked at the aim and objectives as well as the research questions that guided the study. The significance and scope of the study are discussed in this section. The next chapter looks at the review of literature by other scholars in order to broaden the understanding of the subject matter.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter discusses the relevant literature reviewed in the study with respect to history of electronic mobile application and how they became popular because they simplified life at the time. The chapter gives an overview of three types of applications namely Native Applications, Hybrid Applications and Web Applications. The genesis of electronic mobile application and its developments over time, gives an insight into compatibility and the different types of Operating Systems. Lastly, this chapter looks at related works to automated electronic mobile data collection besides other methods of data collection such as focus group discussions (FGD), participatory rural appraisal /assessment (PRA), Observation Method, Case Study Method, Principal Component Analysis and so on and so forth.

2.2 History of Electronic Mobile Application

The first smartphone was Simon launched by IBM in 1994. It had more than 10 built-in applications. Some of them were, calendar and calculator. Mobile applications instinctively became popular because they simplified life at that time. (Barzilay, Micky 2019) in his journal explains that 2500 years ago the Romans decided that they wanted to collect taxes from each citizen in the Roman Empire .in order to do this they understood that they first needed a census to count the citizens of the empire. The roman empire was vast so they quickly realised that it would be impossible to get everyone to go through the gate of Rome to record each person. The Romans solved the issue by sending officials from Rome to all over the empire to collect data for the census from a specific area and then sent back the data to Rome (Lk 2:3-5). As indicated the data collection procedure was complex and tough and used many resources and much time was spent to collect the data. Today there are tremendous changes that have happened thanks to mobile technology we are able to gather millions of data records, data bytes from a large region in seconds.

With this growth in the technological industry, the expectations of mobile users also increased and gave birth to new innovations. This is how the mobile application industry came into action. While developing a mobile application, the developer creates an application that can run on any operating system (Z. Zhou, C. Sun, J. Lu and F. Lv 2018).

2.3 Types of Mobile Applications

There are three different mobile application and these are:

2.3.1 Native application

This application has full support for application stores and markets. It is easier to locate the application in the catalogue of the applications store. Native applications are not only easy to use but in addition to the built-in functions, they have security, support, efficiency and convenience to provide maximum use of the system hardware for the user.

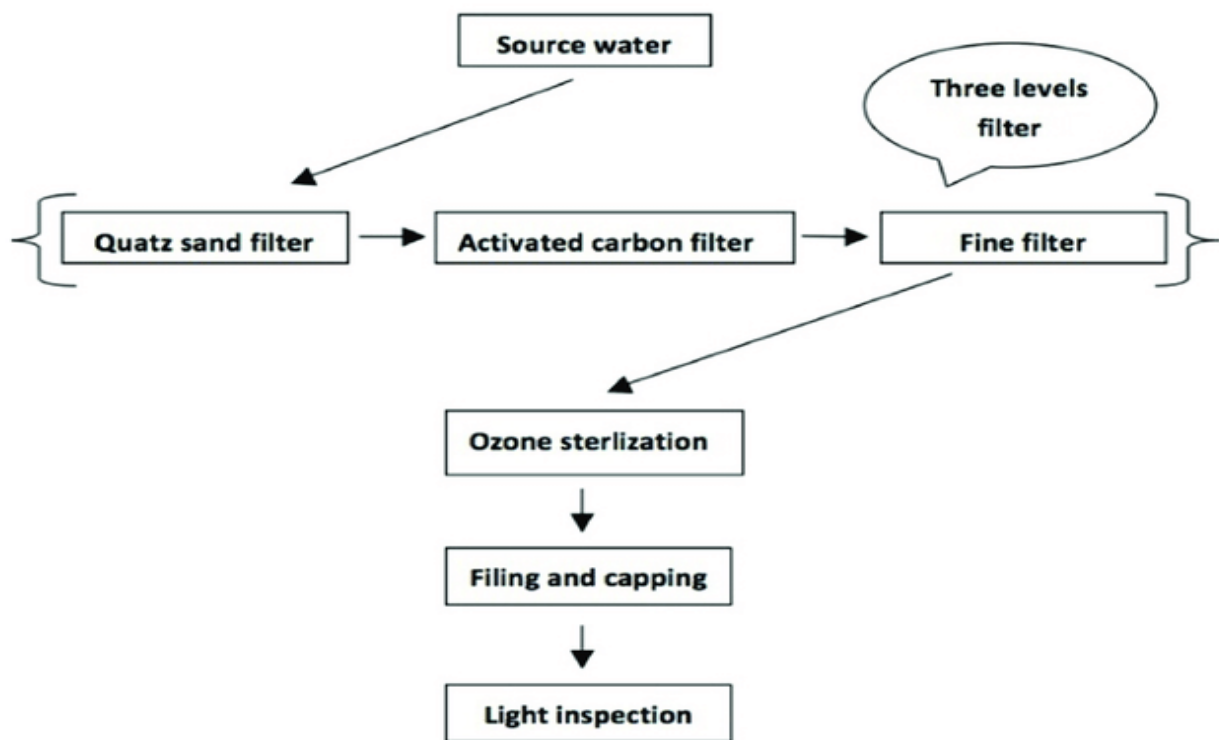


Figure 2.1-1: Example of a native application (Wei, Lei, 2017)

2.3.2 Hybrid Applications

A Hybrid Application is a platform where a mobile application developer, designs a code base that is compatible for both iOS and Android platforms. Some features or designs are not supported on both devices, which can result in continuous changes.

If impeccable and fast performance is essential for the core functionality of your application, a hybrid application should not be on your list. Debugging is another important disadvantage of hybrid applications. The lack of experience of some developers in the specific platform, which is to discover the exact cause of a problem, can be complicated.

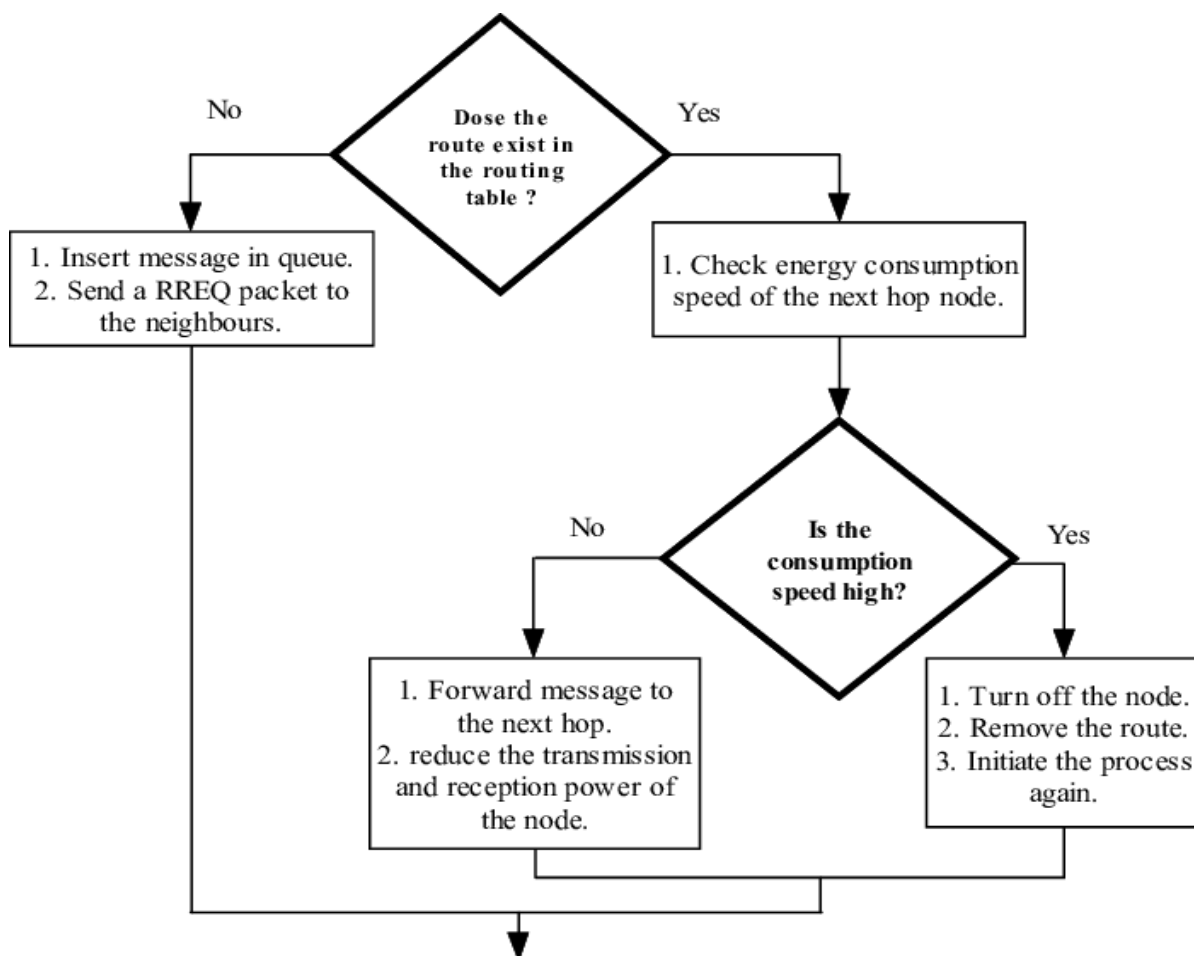


Figure 2.1-2:Hybrid application (Abusaimeh, Hesham, 2009)

2.3.3 Web applications

This Mobile web applications supports multiple mobile browsers, which increases development and maintenance costs. And it has a limited scope when it comes to accessing the functions of a device. This application could be important if you are running a small business and having an Internet connection to run it.

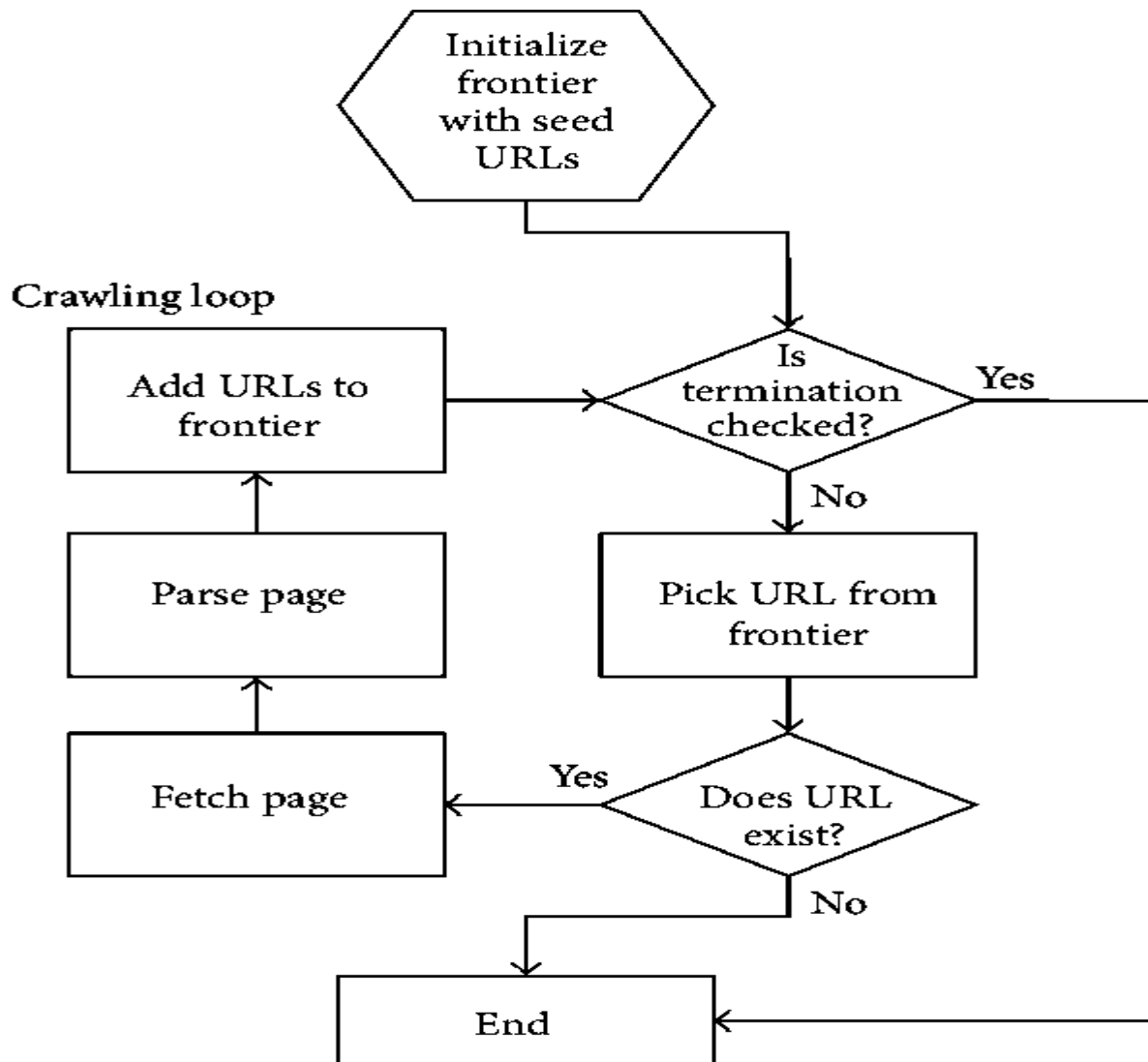


Figure 2.1-3:web application (Cho, Ying-Chiang, 2017)

2.4 Integrated Computerised Case Management

A computerized Case Management System should be able to provide relevant individual client information as well as standardized and flexibly aggregated data to a large variety of end-users. The design and implementation of such a system is often a laborious and highly lengthy process. The design of such a system is affected by the concept the initiators originally have in mind. The computerized case management system in the Mental Health Department of the Israel Defence Forces was designed as a means to provide clinicians with relevant data and thus improve service delivery. The system developed through four consecutive attempts from initiation to completion. The comparative analysis of these iterations produced a set of guidelines for the development and implementation of similar systems. (monnickendam manachem, 1989).

2.4.1 What is data collection

Data collection is a process of obtaining quality and useful information which enables government, organisations, business and academic institutions to analyse information and based on it to answer research questions to help make better decisions, improve efficiency, establish facts and solve problems.

The growing importance of data has increased in a feasible way to approach impact measurement and data collection. According to (Onoka, 2017) online service providers along with companies and investment markets are all gathering information and feedback about the services they provide. They analyse the data allowing them to learn and service the customers better.

2.5 Mobile Data Collection

With the rapid development of mobile communication technology, mobile applications have been widely promoted (G. Jing, H. Wei, W. Ziyang, Z. Chao and J. Haitao 2019).

Mobile data collection is an improved way for data capturing and storage. It involves devices such as PDAs, cellular phones, smartphones, tablets, netbook and notebook (Shao D. A 2012.) Currently, on the market, there are many types of hand-held devices running different operating systems. The most used and well known are Android, iOS for iPhone and iPad, Windows mobile, and Blackberry (Jasmeet, Singh P.2014). The following table shows the popular platforms for mobile data collection. Most of them use WiFi connectivity and they all use 3G.

Table 2.1.1: Popular platforms for mobile data collection

Tool	License type	Mobile compatibility	Data type	Network protocol	Storage
Atlas.ti mobile	Proprietary	Android tablets and iPad	Text,images,audio And video	GPRS,3G,Wifi	Hosted
CSEntry CS pro data entry	Open source	Android phone and tablet	Text	GPRS,3G,Wifi, Bluetooth	Local
EMIT	Proprietary	Java phones	Text	GPRS	Hosted
EpiCollect	Open source	Android and IOS phones and tablets	Text,video ,Audio,images ,GPS ,Barcodes	GPRS,3G,Wifi	Hosted
FrontlineSMS	Open source	Basic phones	Text(SMS)	GSM	Local
GIS mobile data collection	Open source	Android and IOS phones and tablets	Text,video ,Audio,images ,GPS ,Barcodes	GPRS,3G,Wifi	Hosted

Tool	License type	Mobile compatibility	Data type	Network protocol	Storage
Kobocollect	Open source	Android and IOS phones and tablets	Text,video ,Audio,images ,GPS ,Barcodes	GPRS,3G,Wifi	Hosted or local
Magpi	Proprietary	Android and IOS phones and tablets	Text, video ,Audio,images ,GPS ,Barcodes	GPRS,3G,Wifi	Hosted
Mobenzi researcher	Proprietary	Java, Blackberry and Android phones and tablets	Text, video ,Audio,images ,GPS ,Barcodes	GPRS,3G,Wifi	Hosted
ODK Collect	Open source	Android smart phones and tablets	Text, video ,Audio,images ,GPS ,Barcodes	GPRS,3G,Wifi	Local or hosted
Pendragon	Proprietary	IOS and Android and Windows phones and tablets	Text, video ,Audio,images ,GPS ,Barcodes	GPRS,3G,Wifi	Local or hosted
RapidSMS	Open-source	Basic phones	Text(SMS)	GSM	Local
SurveyCTO	Proprietary	Android phones and tablets	Text, video ,Audio,images ,GPS ,Barcodes	GPRS,3G,Wifi	Hosted

2.6 Rate of Adoption of Mobile Data Collection in Developing Countries

The rate of penetration of mobile devices has increased, they are more cell phones than individuals in the world. Meanwhile, the gap between developed and developing countries remains wide

(Furukawa MF, 2014). In 2013, the rate of penetration of mobile phones was 128% in developed countries, 89% in developing countries, and 63% in Africa (ITU. ICT facts and figures. 2013). this high cell phones penetration rate is considered as a factor that arouses the use of hand-held devices for data gathering in Africa (Anokwa Y, 2009). However, the use of mobile devices for health research or public health data collection seems low in developing countries, and studies on this topic are scarce. According to Tomlinson et al, in developing 10 countries, “the use of mobile technology as a research instrument is in its infancy.

2.7 Cloud Computing

Cloud computing has been envisioned as the next generation paradigm in computation. In cloud computing environment, both applications and resources are delivered on-demand over the Internet as services. Cloud is an environment of the hardware and software resources in the data centres that provide diverse services over the network or the Internet to satisfy user's requirements (Sun, Yunchuan ,2014). Cloud computing has emerged as a technology primarily focusing on large scale resource sharing and low cost for big data storage technology. Cloud computing has proved to be an emerging technology capable of providing business models for organizations to utilize deferent computing services at a minimum cost. Cloud is defined as a metaphor describing a web as a space where computing has been preinstalled and exist as a pool of services such as information, infrastructure, applications, storage and processing power on the web ready to be shared (K. Kavitha,2014) . The cloud computing architecture consists of the front end and back end components as shown in figure 2.1-4.

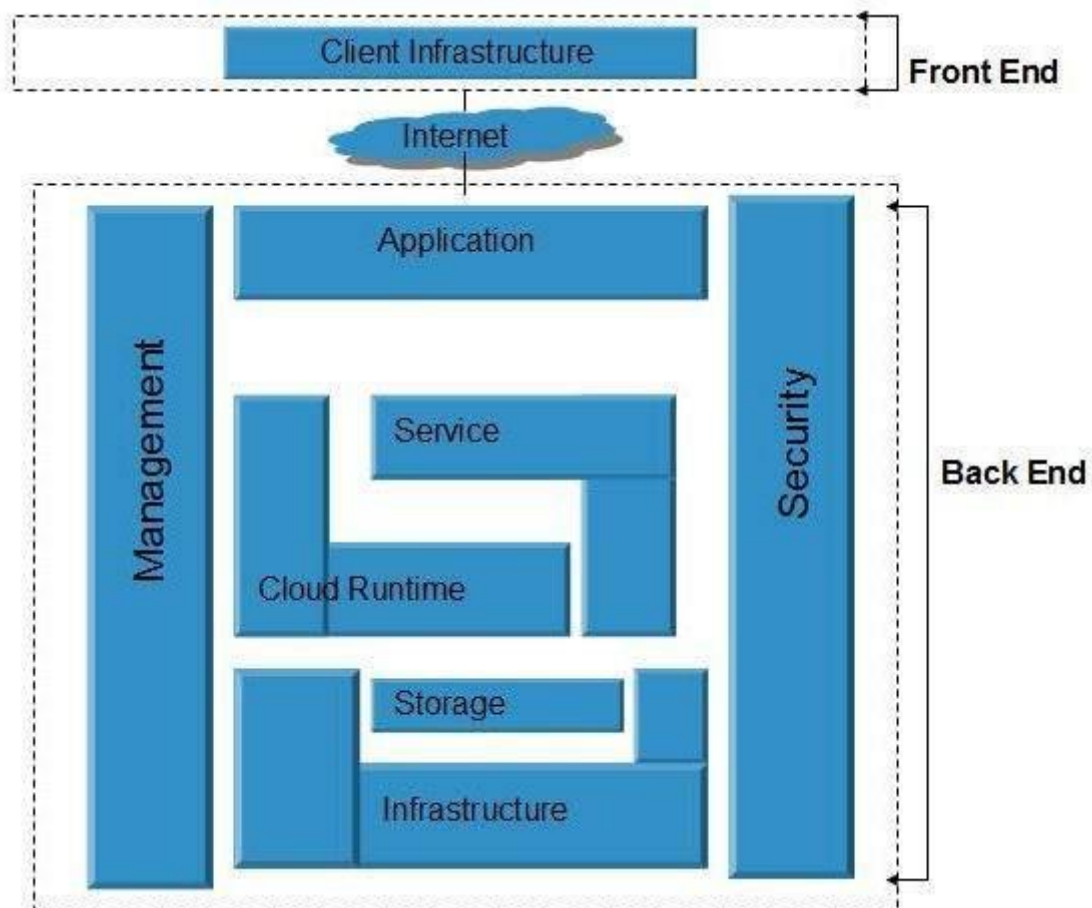


Figure 2.1-4: Cloud Computing Architecture[42]

Cloud computing can be considered as a new computing archetype that can provide services on demand at a minimal cost. The three well-known and commonly used service models in the cloud paradigm are software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS). In SaaS, software with the related data is deployed by a cloud service provider, and users can use it through web browsers. In PaaS, a service provider facilitates services to users with a set of software programs that can solve specific tasks. In IaaS, the cloud service provider facilitates services to the users with virtual machines and storage to improve their business capabilities as shown in figure 2.1-5.

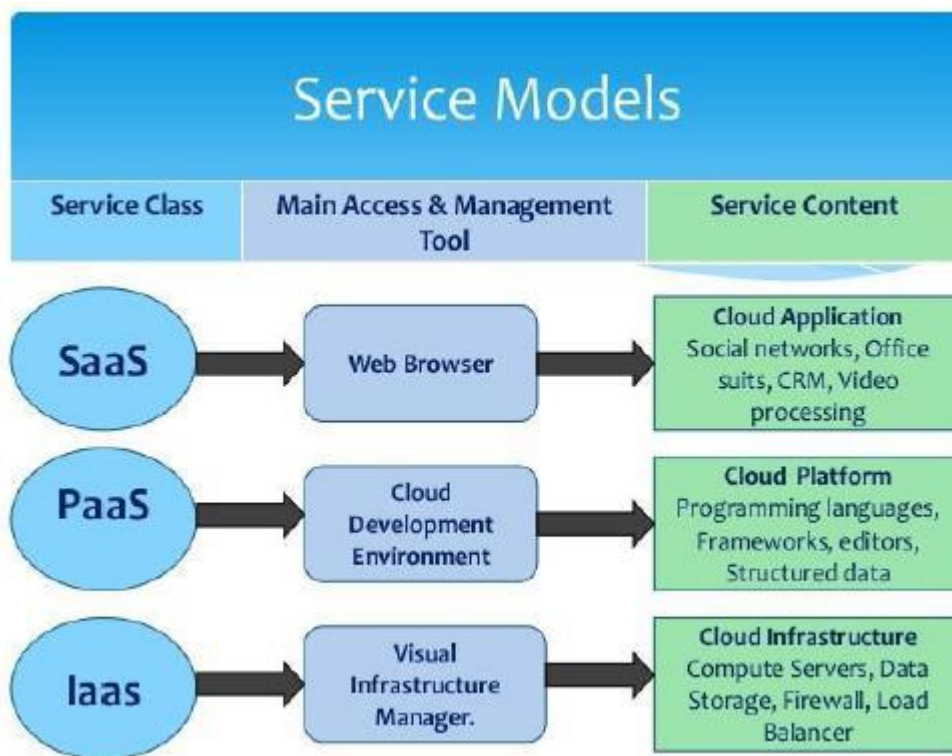


Figure 2.1-5: cloud Service Model (Chandrashekhar , 2017)

2.8 Data Integrity

Data integrity is one of the most critical elements in any information system. Generally, data integrity means protecting data from unauthorized deletion, modification, or fabrication. Managing entity's admittance and rights to specific enterprise resources ensures that valuable data and services are not abused, misappropriated, or stolen.

Data integrity is easily achieved in a standalone system with a single database. Data integrity in the standalone system is maintained via database constraints and transactions, which is usually finished by a database management system (DBMS). Transactions should follow ACID (atomicity, consistency, isolation, and durability) properties to ensure data integrity. Most databases support ACID transactions and can preserve data integrity.

2.9 Authorization

Authorization is used to control the access to data. It is the mechanism by which a system determines what level of access a particular authenticated user should have to secure resources controlled by the system. (Trnka, Michal et al ,2018). It also defines users' permissions in terms of access to digital resources and the extent of its usage. Authorization is granted to the successfully authenticate users according to his/her rights information available in the Access Management System (AMS) (Lynch, 2009). Authorization also addresses the issue of responsibilities assigned to different personnel involved in the development of a digital repository/library and their respective authorities in terms of addition, deletion, editing and uploading of records into a digital collection. Authorization is more challenging than authentication, especially for widely distributed digital content providers. Conventional access control architecture denotes an access control policy as a subject (user) is authorized to exercise some permission on an object. This usual model implicitly assumes that the user population is known more or less. But in a digital content environment, the user population is vast, dynamic and impossible to predict all the users. Thus conventional authorization or access control mechanisms that rely on knowing the user and associating permissions with them fail significantly in digital repositories. So, this digital environment demands some further challenge for access control (Bertino, 2002).

2.10 Related Works

Most statistical offices and agencies in developed countries like the USA and Netherlands already migrated from tradition paper and pencil interviews to computer-assisted personal interview (Danielsson and Maarstad, 1982). The first national household survey that used CAPI for all of its data collection was the Netherlands Labour Force Survey in 1987 (vanBastelaer et al. 1988). In the same year the first U.S. national household survey, the Nationwide Food Consumption Survey, was conducted by national analysts, using CAPI for at least part of the data collection (Rothschild and Wilson, 1988). Central statistical office in most developing countries in Africa are not equipped with enough tablets and software's for the CAPI data collection system. African countries have either a medium or low IDI (ITU, 2011) with those in North Africa (Algeria, Morocco and Tunisia) and Southern Africa (South Africa, Mauritius and Botswana) having relatively higher scores. According to Abdoukarim and Rugege (2013), the highest-ranked African country is ranked 70 globally (out

of 155), and more efforts must be made to ensure that Africa is a more active participant of the information society. These efforts can be driven by more ICT for development (ICT4D) projects on the African continent. The importance of ICT4D projects within Africa can be seen in the following quotation of the Secretary-General of the ITU, Dr Hamadoun I. Touré (2013).

Data collection using PDAs was more accurate and complete than paper-based data collection. The rate of omitted information among 32 variables collected in the hospital was 6% (342/5760; 95%CI: 5.4 to 6.6) for the paper based-data collection compared to none in the electronic data collection ($p < 0.05$). Mistakes such as typographical errors, decimal point faults or illogical values for the variables glucose, haemoglobin, blood pressure, heart rate and culture bottle weight before and after incubation of blood were reduced from 7% (65/900; 95%CI: 5.7 to 9.1) in the paper-based questionnaires to 1% (95/11,045; 95%CI: 0.7 to 1.0) in the electronic data collection ($p < 0.05$) respectively. For a 19-month study period, a total expenditure of USD 23,500 was calculated for the paper-based data collection and entry, compared with USD 17,710 for the PDA-based system (David G. Dillona, 2014).

De Leeuw and Nicholls (1996) point out that the benefits of moving from PAPI to CAPI are cost savings and a reduction in the time elapsed between fieldwork and the availability of the data for analysis. For academic studies like BHPS and SOEP, the potential improvement in data quality is the most important benefit. The quality improvement need not only be due to CAPI itself but can partly be the result of a self-selection process among interviewers: if the professional interviewers want to work with CAPI, the quality of surveys conducted using PAPI decline, due to negative self-selection into the group of remaining interviewers.

Bet Caeyers (2010) points out that checks are believed to lead to more accurate data capture, in Capi because they were run *during* the interview, at a time when they could still be resolved with the respondent. The check procedure does not run automatically but is activated by the interviewer by manually clicking check-buttons. They are run at various Stages during the interview, typically after completing all the questions on one screen. A final, global check can be run at the end of the interview and this improved the data quality. In addition, he also states in his paper that errors leading to missing variables in PAPI are virtually eliminated in CAPI.

Dylan Knowles (2014) in his paper Design and Use of a Smartphone Data Collection Tool says Smartphones reduce many of the day-to-day problems encountered by traditional pen and paper

tools and custom hardware. Smartphones are frequently woven into their owners' lives and are generally carried by their owners. Smartphone-based data collection, therefore, has the potential to capture much information about participants' lives. Smartphone apps (henceforth, apps) can issue reminders, can launch traditional tools exactly when required, can provide sensed contexts such as location to survey and dairying applications, and are weightless and space less beyond the hardware platform on which they run. Paper surveys and diaries, in contrast, are bulky, silent, and have no knowledge of their previous use nor current spatial context. Apps have upload capabilities that can transmit data on regular intervals, which reduces the likelihood that survey entries will be lost. Smartphones are built to withstand the abuse of careless human users and environmental risks (e.g., static discharge) reducing the likelihood that collected data will be lost due to damage.

Kamala Thriemer^{1*}, (2012) says that the main concern with using the paper forms was the time interval between data collection and data checking prior to computerization. This frequently resulted in the detection of erroneous data at a point when it was too late to make corrections (e.g., after the patient was discharged). The main errors detected were omissions and illogical data. The inadequate adherence to inclusion and exclusion criteria during the screening process posed an additional challenge. When using paper-based CRFs study enrolment criteria were inconsistently applied, resulting in missing eligible patients and difficulties tracking the proportion of patients missed for the study but were missed, creating a potential bias when reporting the results. These problems led to the introduction of PDAs for electronic data collection.

Adama Baguiya(2016) points out in his paper an offline mobile data capture module for health and demographic surveillance system introducing Information and Communication Technologies (ICT) and mobile technologies for data collection may significantly improve data quality and ease data processing and data management. So far, the level of use of mobile devices in public health research data collection and its determinants are not clearly known.

Edgar Ferrer-Moreno et al (2013) say Electronic Medical Records (EMRs) are promising to improve quality and efficiency in healthcare services. Several models for adopting EMRs in healthcare centres have been developed, these are complex models that follow a clinician driven workflow.

Teresa Arroyo-Gallego, (2016) states that using smartphones as tools for longitudinally tracking symptoms could enable personalization of drug regimens and improve patient monitoring.

Parkinson's disease (PD) is an ideal candidate for these tools. At present, evaluation of PD signs requires trained experts to quantify motor impairment in the clinic, limiting the frequency and quality of the information available for understanding the status and progression of the disease. Mobile technology can help clinical decision making by completing the information on motor status between hospital visits.

Schick-Makaroff and Molzahn (2015) in their paper say that the benefits of tablet computers for the collection of electronic patients reported outcomes such as:

- Immediate access to retrievable data.
- Potential to link ePRO data with electronic health records.
- Capacity to trigger automated alerts when patients identify problems.
- Immediate scoring and presentation of data.
- More economical in terms of resources and time Greater likelihood of reporting sensitive Information.
- Prevention of errors in data-entry.
- Observation of survey completion in real-time.
- Improved data quality.
- Less missing information With minor modifications from paper to electronic Formats, validation is not likely necessary.

Thriemer et al (2012) in their paper replacing paper data collection forms with electronic data entry in the field: findings from a study of community-acquired bloodstream infections in Pemba, Zanzibar say that Electronic data collection offers direct data entry at the initial point of contact. It has numerous advantages and has the potential to replace paper-based data collection in the field. The availability of information and communication technologies for direct data transfer has the potential to improve the conduct of public health research in resource-poor settings.

Carina King et al (2014) points out that Fieldworkers preferred using EDC to paper-based systems, although some struggled with the technology at first. Highlighted features include in-built skip patterns for all systems and specifically the 'case' function that CommCare offers. MIVA as a standalone app required considerably more time and expertise than the other systems to create and

could not be customised for our specific research needs; however, it facilitates standardised routine data collection. CommCare and ODK Collect both have user-friendly web-interfaces for form development and good technical support. CommCare requires the Internet to build an application and download it to a device, whereas all steps can be done offline with ODK Collect, a desirable feature in low connectivity settings. Pendragon required more complex programming of logic, using a Microsoft Access application, and generally had less technical support. Start-up costs varied between systems, and all were considered more expensive than setting up a paper-based system; however, running costs were generally low and therefore thought to be cost-effective over the course of our projects.

Multiple studies confirm productivity gains of between 15 and 25% when the mobile workforce is equipped with mobile computers and wireless access (Olst, n.d. 2004). These productivity savings come about in three ways: time savings, flexibility, and quality of work. In his paper on the impact of mobile application technology on today's workforce, Rochford outlines how this technology enables significant structural change to how organisations perform their tasks and accomplish business goals. This can now be achieved by moving existing business processes, automated or manual, beyond the organisation's office to wherever and whenever those tasks can be carried out efficiently. One of the key enablers to this move is through the use of GIS. Today there are several organisations with field workforces that continue to perform daily operations using the traditional paper-based, clipboard system. Paper-based processes have long been the *de facto* method for data collection and information management. Research shows that some 70% of all errors in corporate databases can be attributed to the manual entry of paper-based data. The development of mobile computers has brought a completely different perspective and approach on how humans communicate with computers. Interaction with mobile computers is vastly different compared to desktop computers. Desktop computing takes place in a static environment where the user is presented with a large, high-resolution user interface and information is stored either locally on the computer, or may be accessed via a data network. Cobb states that "usability" is important for any mobile computing application or device, but it is even more important for capability-constrained devices with limited display or input, Thus usability analysis should encompass any wireless / PDA application project from start to completion.

M. Reichert et al (2017) says that a large amount of data needs to be collected in a rather short time. In this context, smart mobile devices can be a feasible instrument to foster data collection scenarios. To enable domain experts to create and maintain mobile data collection applications

themselves, the QuestionSys framework relies on a model-driven approach to digitize paper-based questionnaires. This digital transformation is based on manual as well as automated tasks. The manual tasks applied by the domain experts can be eased by the use of change patterns. They describe features to easily add or delete the elements of a questionnaire. This work summarizes crucial change patterns and shows how they can be applied in practice.

Lauraitis, et al R (2019) in their paper say that new technology enables a constant boost to the powers of mobile devices, which in the previous years have transformed from simple mobile phones to smart phones. Computational powers of these electronics enable actions that previously were possible only for computers. By the use of special applications, we may benefit from sensors and multimedia capabilities of operating systems. Therefore, a new era for devoted implementations opens, in which a smart application can take the role of computing system to estimate the symptoms of diseases by evaluating signals coming from a human body. Methods: We propose a model of an application implemented for mobile android systems, which can be used for examination of central nervous system motor disorders occurring in patients suffering from Huntington (HD), Alzheimer, or Parkinson diseases.

K. Biron, et al (2019) in Crime Tracking System (CTS) say that the process of automating data entry, storage, and retrieval using a mobile interface and a database is efficient in terms of storage of data. Crime investigation plays a significant role in keeping the modern cities safe. The emergence of the field of data science has opened new doors that have high potential in facilitating the process of the criminal investigation. Data science enables criminal investigators to create tangible and quantifiable information on criminal social network profiles then analyzes crime concurrence based on location proximity. Aiming at availing the relevance of integrating data science analysis techniques in the criminal investigation, their study examined the suitability of cluster graphs and geolocation proximity (through heat maps) as two of the data science analysis techniques to facilitate the criminal investigation process. On one hand, cluster graph analyses the social network of a suspect and examines its centrality, parallel betweenness and community affiliation. On the other hand, geolocation proximity provides an investigator with previous crime occurrences through the use of the heat map visualizing technique. These concepts are tested by embedding them in a mobile application developed to facilitate the investigation process. The findings of this study demonstrate that the integration of data science and investigation activities is

feasible and can extend the investigation process beyond the scope of the police profile structured database based on manual interrogation activities.

A.Dankar and P. P. Kundapur (2019) say one of the most common academic processes that institutions/universities follow is that of maintaining student/staff attendance. However, it has been observed that the conventional method of taking students attendance on registers to confirm their physical presence is still prevalent. This method is time-consuming, inefficient and prone to human errors. In order to address the attendance issue, this paper proposes a simple user-friendly mobile application “Automated Mobile Attendance System” (AMAS). AMAS is interfaced with a website in the backend for data entry and report generation. The application is able to track students using GPS and Bluetooth beacons to confirm and verify their presence in classrooms. The application maintains a record of the absentees that is synced with the tables in a remote database server regularly. This application reduces the time required to take attendance, prevents the loss of data as well as provisions to edit incorrect responses. AMAS is developed using Android Studio 3.0.1 and is compatible with 4.4 (KitKat version). The mobile and web application integration is via WAMP server. The database used is MySQL.

Konstantinos Rammos (2018) says the Internet of things (IoT) becomes important and variously used in many applications. In the case of smart health, patients and healthcare practitioners manage their appointment easier. He further said that with an automated patient appointment patients can receive appointment notification before the appointment date, be able to postpone or cancel the appointment easily when they cannot come to see the doctor or when the patient gets better. In addition, the doctors and patients can observe their appointment via mobile application.

J. Schobel (2016) says that the widespread dissemination of smart mobile devices offers promising perspectives for a variety of healthcare data collection scenarios. Usually, the implementation of mobile healthcare applications for collecting patient data is cumbersome and time-consuming due to scenario-specific requirements as well as continuous adaptations to already existing mobile applications. Emerging approaches, therefore, aim to empower domain experts to create mobile data collection applications themselves. This paper discusses flexibility issues considered by a generic and sophisticated framework for realizing mobile data collection applications. Thereby, flexibility is discussed along with different phases of data collection scenarios. Altogether, the realized

flexibility significantly increases the practical benefit of smart mobile devices in healthcare data collection scenarios.

M. Katarahweire (2017) say Mobile data collection systems (MDCS) in the health sector are of great benefit to health care providers and community workers especially in low-resource settings. MDCS enable the extension and provision of health services closer to the community by enabling data collection and diagnosis without the patient being in a hospital setting.

H. Zhong *et al* (2010) say that traditional data collection methods for geological investigation, profile measurement, and other field survey in the research community of geosciences are generally based on the manual measuring and recording by investigators. The most popular “instruments” used in the traditional survey are pencils as well as printed hard copy charts. The methods are obviously complicated and inefficient, at the same time the collected data are always inaccurate and not compatible with the digital process in the computer.

Y. Anokwa (2009) points out that Organizations in low-income regions need tools for collecting and reporting on data. Existing paper-based approaches are often slow and incomplete when compared to data collection tools on mobile devices.

J. Schobel, R. Pryss and M. Reichert (2015) in their paper state that In future, more and more clinical trials will rely on smart mobile devices for collecting structured data from subjects during trial execution. Although there have been many projects demonstrating the benefits of mobile digital questionnaires, the scenarios considered in literature have been rather limited so far. In particular, the number of subjects is rather low in respective studies and a well controllable infrastructure is usually presumed, which not always applies in practice. This paper gives insights into the lessons learned in a clinical psychology trial when using tablets for mobile data collection. In particular, more than 1.700 subjects have participated so far, providing us with valuable feedback on collecting trial data with smart mobile devices on a large scale. Furthermore, issues related to an insufficient infrastructure (e.g., unstable Internet connections) have been addressed as well.

However, this study investigated challenges facing the Zambia statistics agency in the processing of data collection using PAPI, and provides valuable insights which in turn an integrated and automated mobile data collection system was developed to improve the data collection process.

2.11 Chapter Summary

This chapter discussed the relevant literature reviewed in the study with respect to history of electronic mobile applications and how they became popular because they simplified life at the time. The chapter gave an overview of the types of applications, the genesis of electronic mobile application and their developments over time, an insight into compatibility and the different types of operating systems. Lastly, this chapter closed by looking at related works to automated electronic mobile data collection application system.

One of the most striking traits of automated electronic mobile data collection application system is that they offer productivity in terms of time, flexibility, and quality of work.

CHAPTER THREE

METHODOLOGY

3.1 Introduction

This chapter discusses the research design that was employed especially with respect to the choice of the design. It also looks at the area of study, the population of the study, sample design under which the sampling frame, technique and size was determined. Thereafter it discussed data collection methods as well as data analysis and data presentation methods that were employed in the study. Lastly, it dealt with the ethical considerations and limitations of the study.

3.2 Research Design

This study adopted an exploratory research design. Exploratory research is carried out to find out about a research problem when there are no or very few earlier studies to refer to (Ayakonya 2013). This design was chosen because there is little studies done on the subject matter in Zambia. The study aimed at collecting information to describe the existing situation and attempted to find solutions to the challenges and developed an integrated and automated mobile application system. The researcher also used primary data by administering the questionnaires to test the acceptability of the system to the users.

3.3 Area of Study

The study was conducted in Lusaka at the Zambia Statistics Agency (ZSA) head office and Lusaka provincial office. The Zambia Statistics Agency headquarters are located along Nationalist Avenue and the Provincial office located at Indeco house along Cairo road.

3.4 Study Population

The study population included all members of staff who were working in the research and development unit and data processing branch in the Information Technology (IT) Department of the Central Statistical Office at the headquarters and the Provincial Office in Lusaka province.

3.5 Sample size and sampling technique

The sample was 20 which were purposively sampled. The purposive sampling technique that was used was the homogenous purposive sampling. Homogenous purposive sampling was chosen because the researcher's focus was on particular characteristics of the population that are of interest to enable the researcher answer the research questions which are related to data collection processes.

3.6 Data Collection Methods

Data collection was carried out over a period of Five weeks starting from the third week of December 2019 to the last week of January 2020. During data collection, the researcher administered the self-administered questionnaires to the respondents at all the target locations in Lusaka and gave the respondents enough time, which was spread out into a few days to answer the questionnaires. The researcher collected all the answered questionnaires in the second and third weeks of the data collection period. In the Sixth week, the researcher started designing the template for entering data that was collected to be analysed. This research draws knowledge from already existing technological systems, and as such, the development of the new system mostly depended on the existing secondary data on the subject of database migration combined with the theoretical knowledge of the author. Secondary sources in form of published literature from journals, books and internet sources were used for this research. These were used to review what other people have done, or related work on the subject. However, the research also made use of primary data by interviewing experts or people that have used these database migration systems for a long time.

The research draws knowledge from already existing technological systems, and as such, the development of the new automated mobile information system mostly depended on the existing

secondary data on the subject of mobile data collection systems combined with the theoretical knowledge of the author. Secondary sources in form of published literature from journals, books and internet sources were used for this research. These were used to review what other people have done, or related work on the subject. However, the research also made use of primary data by interviewing experts or people that have used these systems for a long time.

3.7 Data Processing, Analysis and Presentation

The method of analysis for this research was divided into two categories, and these categories were mainly guided by the three objectives of the research.

3.7.1 Primary Data Analysis

The first method of analysis was the analysis of primary data collected from the individual IT personnel. This involved getting information from individual IT personnel on the existing data collection methods and identifying the challenges associated with it. To achieve this, IT experts were purposefully selected from the Information and Technology Department. The questionnaires were carefully structured with both closed and open-ended questions. The closed ended questions were meant to capture the quantitative aspects of the responses, while the open-ended questions were designed to capture the qualitative responses of different expert opinions. After the collection of the data, it was entered using a template in the Statistical Package for Social Sciences (SPSS v23) where it was edited and cleaned. Analysis was done using descriptive analysis, which sort to establish the sample distribution by displaying sample characteristics of the respondents as well as their individual views and opinions in relation to the subject under study. Data was presented in form of tables and figures such as bar charts for better understanding.

3.8 System Automation

The second method of analysis involved designing the modules that made up the system. This was based on the third objective of the research, which was to develop an integrated and automated mobile data collection system and implement the designs into a functional

program, whose aim was to address the challenges associated with the current data collection system. In so doing, the following were expected from the new system:

- (i) Fully automate the mobile data collection process to reduce the cost of printing and transporting bulky questionnaires to different parts of the country where we have the statistical offices.
- (ii) Improve data quality by providing the skip patterns in the automated mobile application system to make it easy for enumerators to follow through.
- (iii) Finally, to reduce the dissemination period as data is instantly digitised and analysis can start almost immediately.

3.9 Ethical Consideration

The study ensured the following:

- (i) Consent: The subjects in this study were reminded that their participation, in this case, was purely voluntary and that if they wished they could decline to respond or taking part in the study.
- (ii) Right of Respondents: This was done to enable the researcher to respect the autonomy and privacy of the respondents. The respondents therefore were assured of ultimate confidentiality and anonymity. To that effect, respondents were informed that they needed not to indicate their names on the questionnaires.
- (iii) Confidentiality: The information that the participants brought would be treated as confidential and no part of it would be leaked out for any purposes than was intended. To ensure this, the researcher used no name or any identifying marks to recognise the respondents.

3.10 Limitations of the Study

The study faced a lot of lack of interest from the respondents. It was not easy to collect enough questionnaires back on time causing delays in data processing and analysis. Moreover, financial constraints were another limiting factor.

3.11 Summary

This chapter outlined the stages in ensuring that the objectives stated in chapter one were followed. This included the research design, data collection method, sampling and analysis of the data collected. This next chapter outlines an analysis of the findings of the study.

CHAPTER FOUR

ANALYSIS OF STUDY FINDINGS AND DISCUSSION.

4.1 Introduction

This chapter is dedicated at analysis and discussion of the findings of the research in relation to the challenges associated with the old system of collecting data, as well as the implementation of the new system, aimed at addressing the challenges of the old system. As such, the chapter consists of two components, i.e. the descriptive analysis of the primary data that was collected from participants of the research, as well as the practical implementation of the new system described in the methodology chapter of the research.

4.2 Socio-Demographic Data

Participant's socio-demographic data was important in order to identify other variables that may not have been part of the central objectives of the study. The socio-demographic characteristics include age, sex and level of education. The study found out that 50% of the respondents were aged between 30-39 years of age. Most of the respondents were males representing a percentage share of 60%, 45% of the respondents possess a bachelors' degree. See Table 2.

Table 4.1.1: Socio-Demographic Data

Variable	Frequency	Percentage
AGE RANGE		
20-29	2	10%
30-39	10	50%
40-49	5	25%
50+	3	15%
Total	20	100%
SEX		
Male	12	60%
Female	8	40%
Total	20	100%
LEVEL OF EDUCATION		
Diploma	6	30%
Bachelors degree	9	45%
Masters Degree	0	25%
Total	20	100%

4.3 Challenges of Paper and Paper Interviews (PAPI)

The study was interested in establishing the challenges that come with Paper and Pencil Interviews (PAPI) as a data collection system. Participants were asked what challenges they face when collecting data using PAPI. The question sought to find out the extent to which PAPI is affecting the data collection process.

4.3.1 Delay in results

The study revealed that one of the challenges of the current manual system is the delay in results due to data entry and data cleaning. When the respondents were asked point out the drawbacks of paper questionnaires, the majority of them representing a percentage share of 39% said that that delay in results a serious drawback. This is shown in figure 4.1-1.

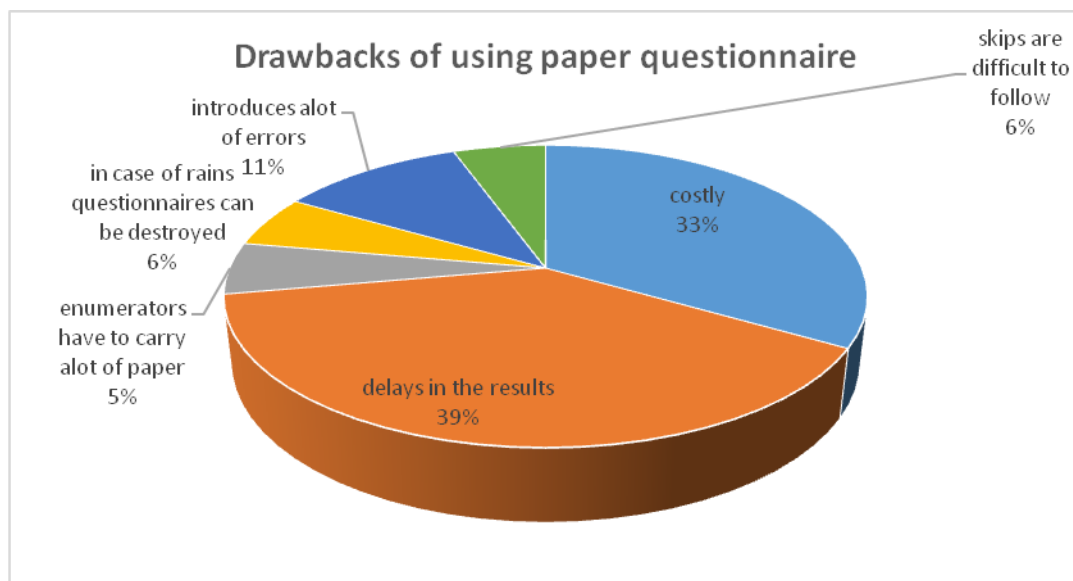


Figure 4.1-1: Distribution of Drawbacks of Using Paper Questionnaires

4.3.2 Dissemination Period

The study revealed that the dissemination period was longer by using the paper questionnaire. A Total of 94.4% of the respondents strongly agreed that the dissemination period is longer with PAPI than with CAPI compared to 5.6% who disagreed as shown by figure 4.1-2.

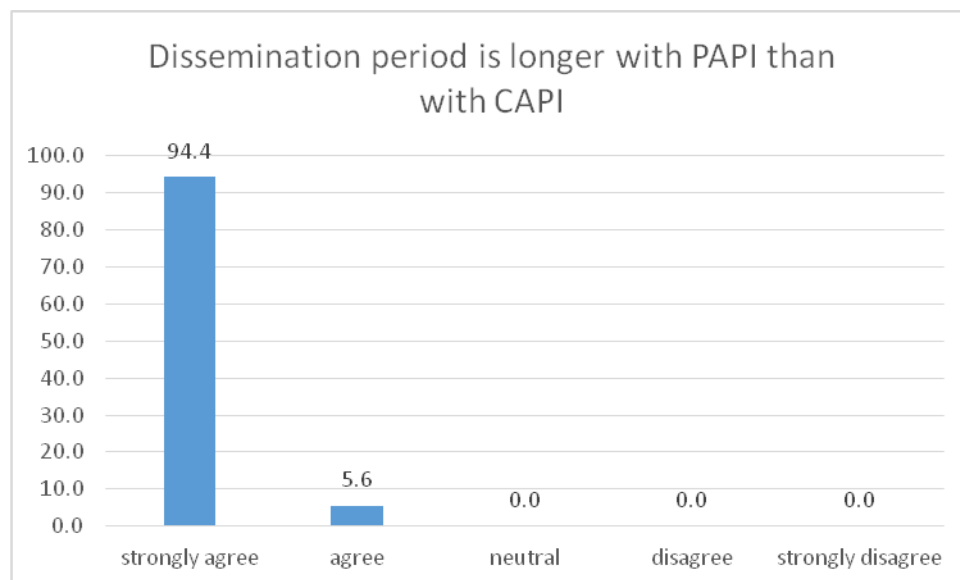


Figure 4.1-2: Distribution of Period of Dissemination using PAPI

4.3.3 Skip Patterns

Another challenge with PAPI is with skip patterns. The study revealed that, the major weakness of the paper questionnaire is that skip patterns can be difficult for enumerators to follow and was seen by the 61.1 % that strongly agreed and only 38.9% disagreed. This is shown in the figure below

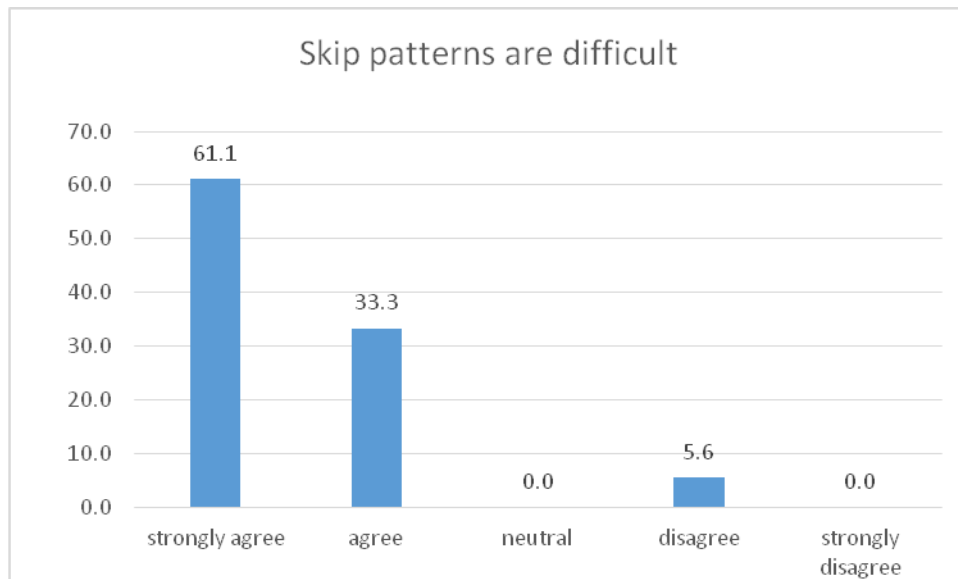


Figure 4.1-3: Skip patterns are difficult with PAPI

4.4 Benefits of the CAPI

The study was interested to establish the benefits of using an integrated and automated mobile data collection system. This was in answering to research question two which sought to know the benefits of the proposed system of data collection. The study found out the following benefit.

4.4.1 Data Accuracy

It was found that data collection using PDAs was more accurate and complete than paper-based data collection. The study found out that the rate of omitted information among variables collected is more for the paper based-data collection compared to the electronic data collection which is said to be almost none. 90% of the respondents echoed that mistakes such as typographical errors, decimal point faults or illogical values for the variables reduced in the electronic data collection.

4.4.2 Productivity Gain

The study found out that there are productivity gains when the mobile workforce is equipped with mobile computers and wireless access. These productivity savings come about in three

ways: time savings, flexibility, and quality of work. 80% of the respondents agreed that technology enables significant structural change to how organisations perform their tasks and accomplish business goals.

4.4.3 Cost-Effectiveness

The study revealed that using mobile devices for data collection is not significantly more expensive than using a paper-based system, especially for a small scale study. The initial cost of devices is what usually makes electronic device solutions slightly more expensive than manual data collection. However, it is proved that for large scale studies, tablet-based data collection definitely becomes cheaper than the traditional method. Moreover, irrespective of the duration of the study, with mobile data collection, the total cost of ownership is lower as mobile devices and software can be reused for a long period of time over several studies.

4.5 Implementation of a New automated mobile data collection application system

The design of the integrated and automated mobile data collection application system was effectively implemented designed using CSPRO and Android Mobile Application Technologies. The system consists of the User Menu which helps the user to Collect Data, Send/Receive Data, Update the System, Logout and Exit the Program.

4.5.1 Current System



Figure 4.1-5:Paper Questionnaire

Paper and Pencil Workflow

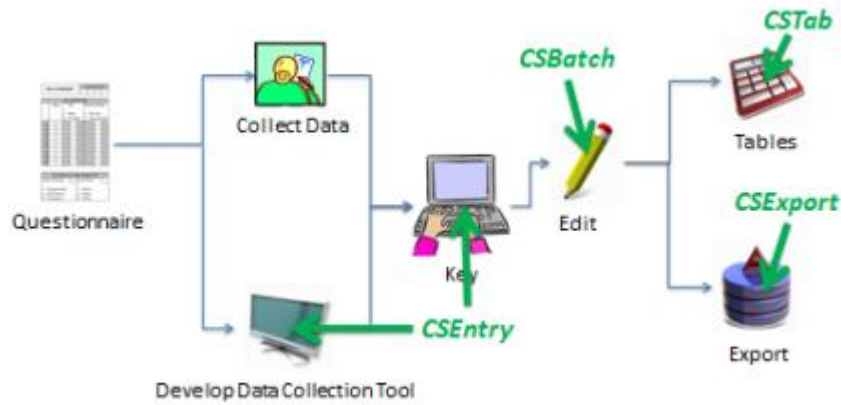


Figure 4.1-6: Paper Questionnaire Flow Chart

The main concern with using the paper forms was the time interval between data collection and data checking prior to computerization. This frequently resulted in the detection of erroneous data at a point when it was too late to make corrections, for example after the interviewer had come back from the field. The main errors detected were omissions and illogical data. When using paper-based question to collect data study criteria were inconsistently applied, resulting in missing eligible respondents and difficulties tracking the proportion of the respondents missed for the study missed, creating a potential bias when reporting the results. These problems led to the introduction of electronic mobile data collection application system.

4.5.2 Integrated System

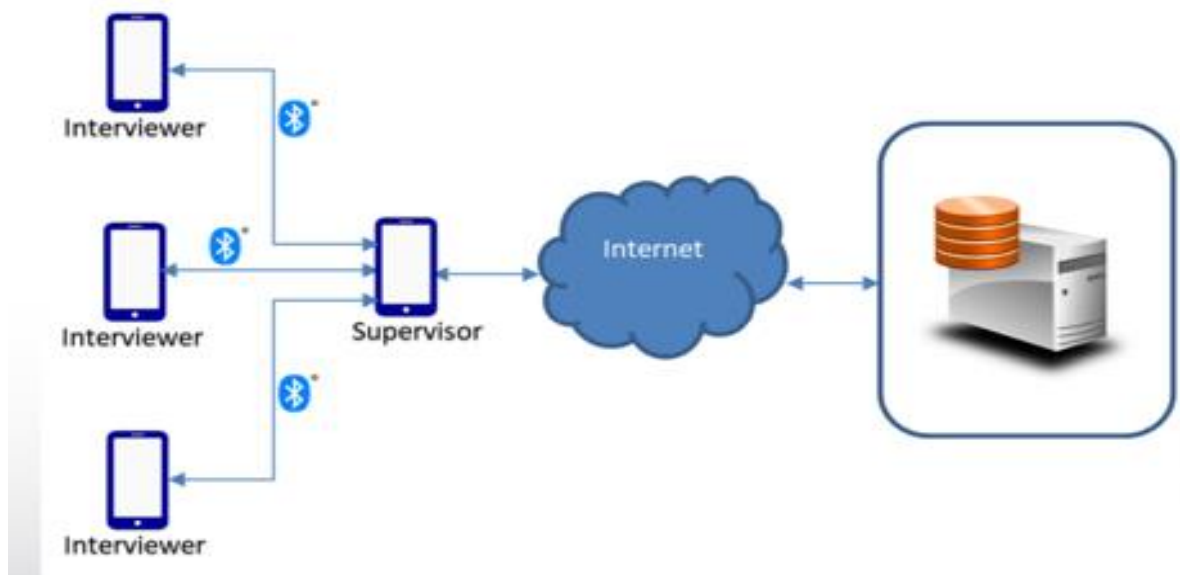


Figure 4.1-7: Integrated System

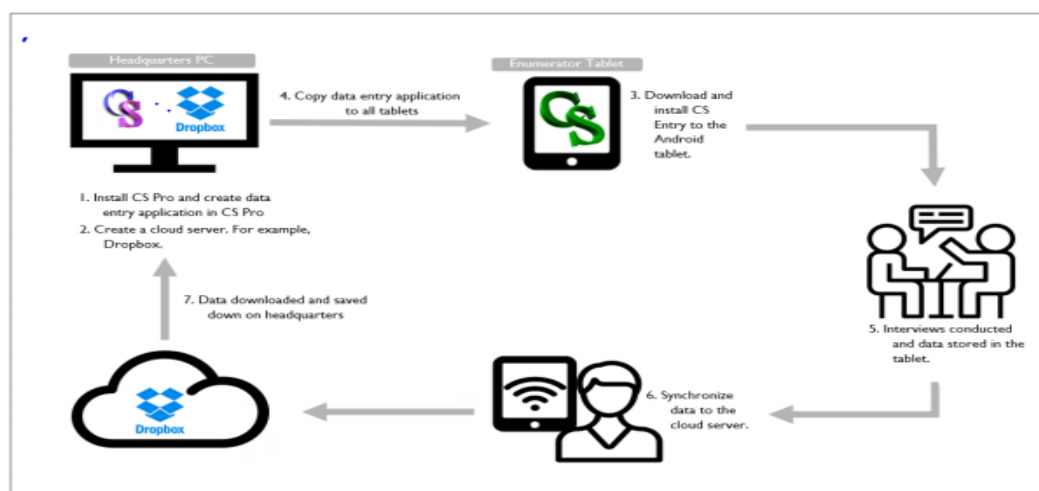


Figure 4.1-8b: Diagram of mobile data system

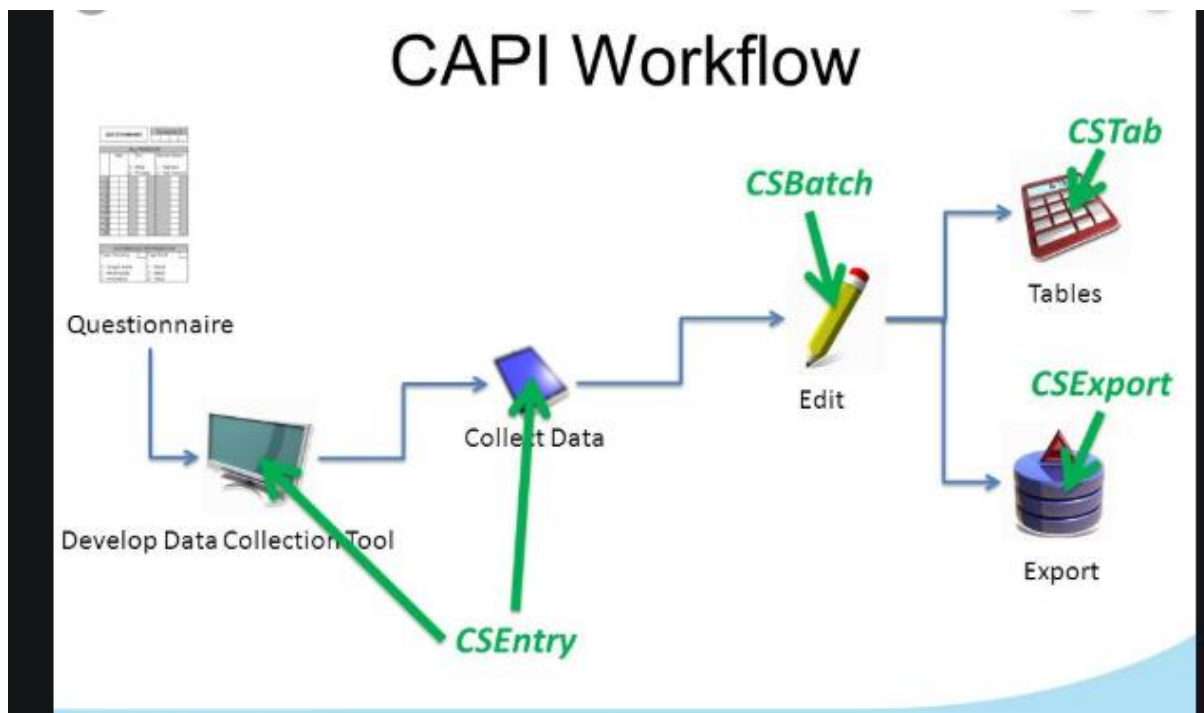


Figure 4.1-9: Workflow of an integrated System

4.6 Results from Installation on the CAPI

The final result of the system was an integrated and automated mobile data application system. The installation of the CSEntry and the results of the survey that was conducted using the electronic mobile data collection application system, which was designed and developed. To install the application, CSPro installation program has to run on each computer and the tablets in readiness of the data collection. The installation setup can be run from removable media such as a CD or USB (flash) drive, or it can be placed on a networked drive. The CSPro installer has the file name **cspro73.apk**. Following are the steps to install CSPro:

- (i) Double-click on the file.
- (ii) Read and accept the U.S. Census Bureau's license agreement
- (iii) Select the components that you want to install. You will generally want to install all of the components.
- (iv) Select the folder where you want to install CSPro. You will generally want to install CSPro in the suggested directory.

After CSPro has been installed on the computer or tablet a CSPro 7.3 icon will be instated on the desktop, which after double-clicking it options to run the program and/or view the release notes will be provided for the user.

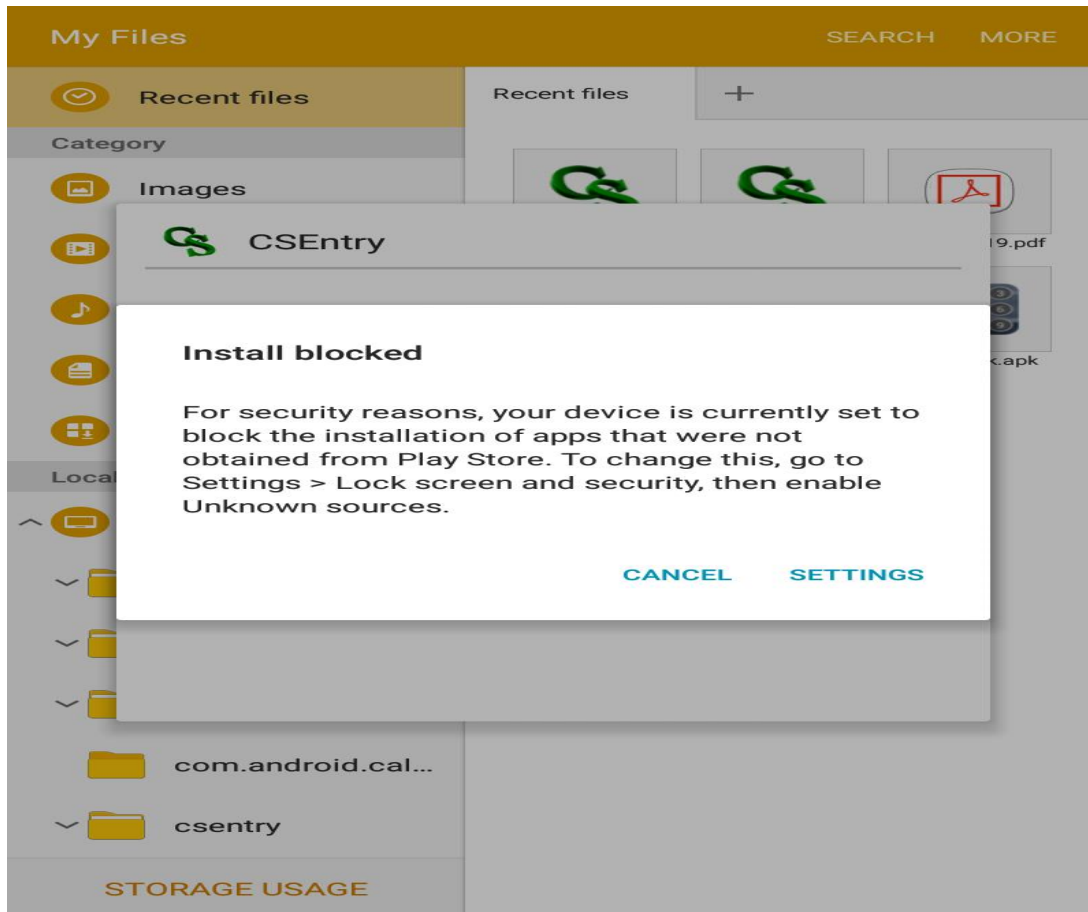


Figure 4.1-10: CSEntry Menu

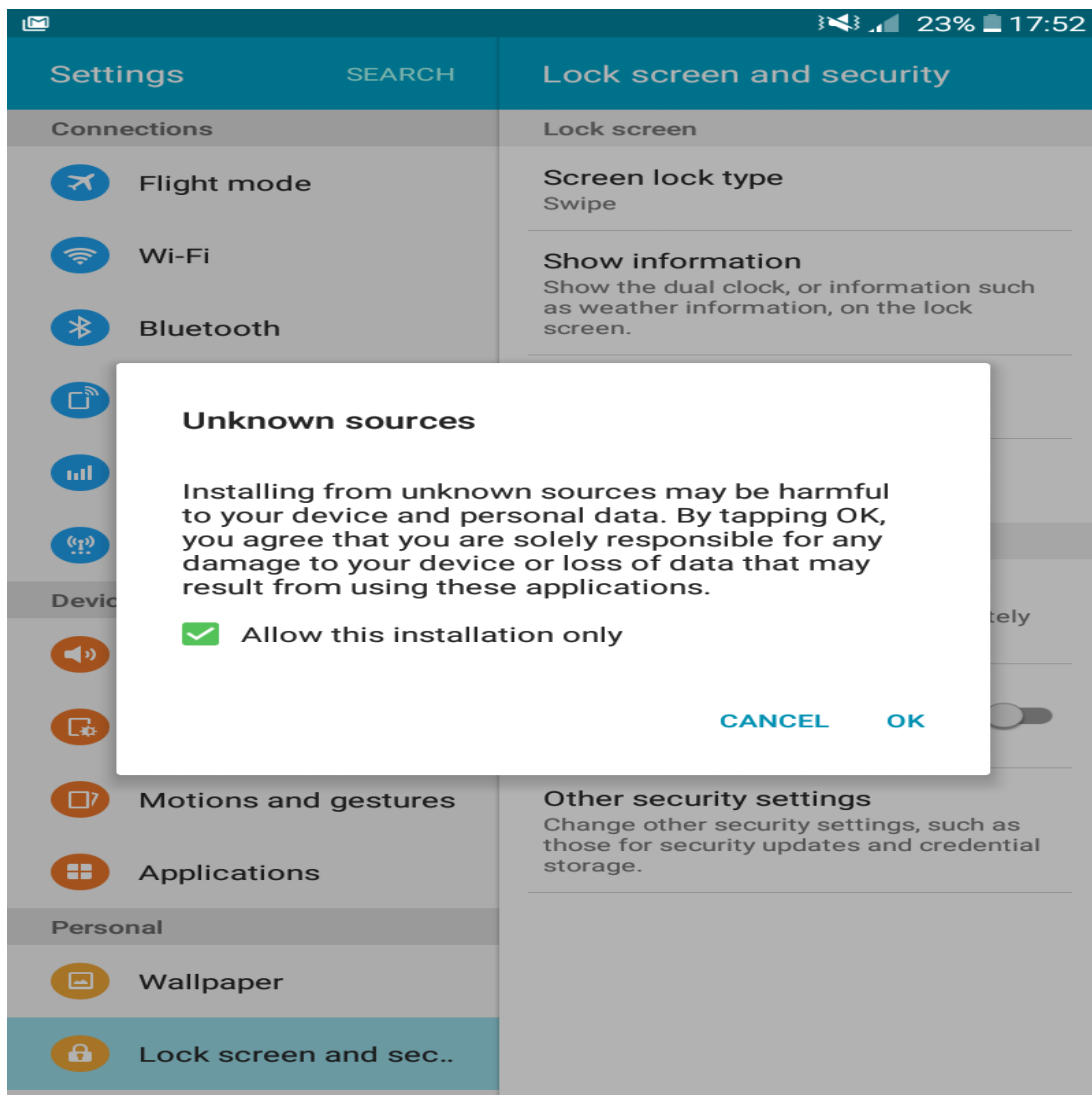


Figure 4.1-11: CSEntry Settings

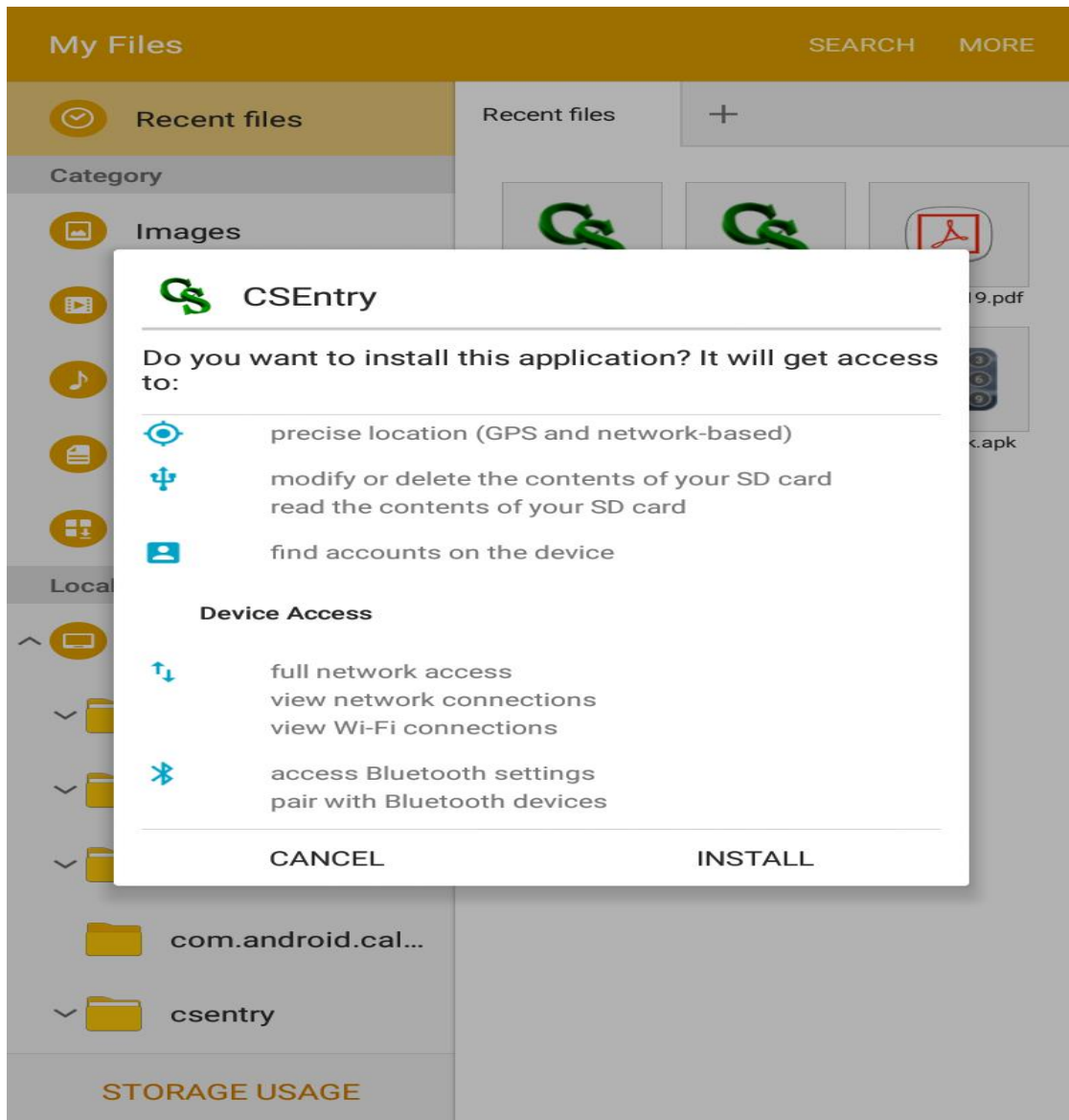


Figure 4.1-12: Installation of the Application

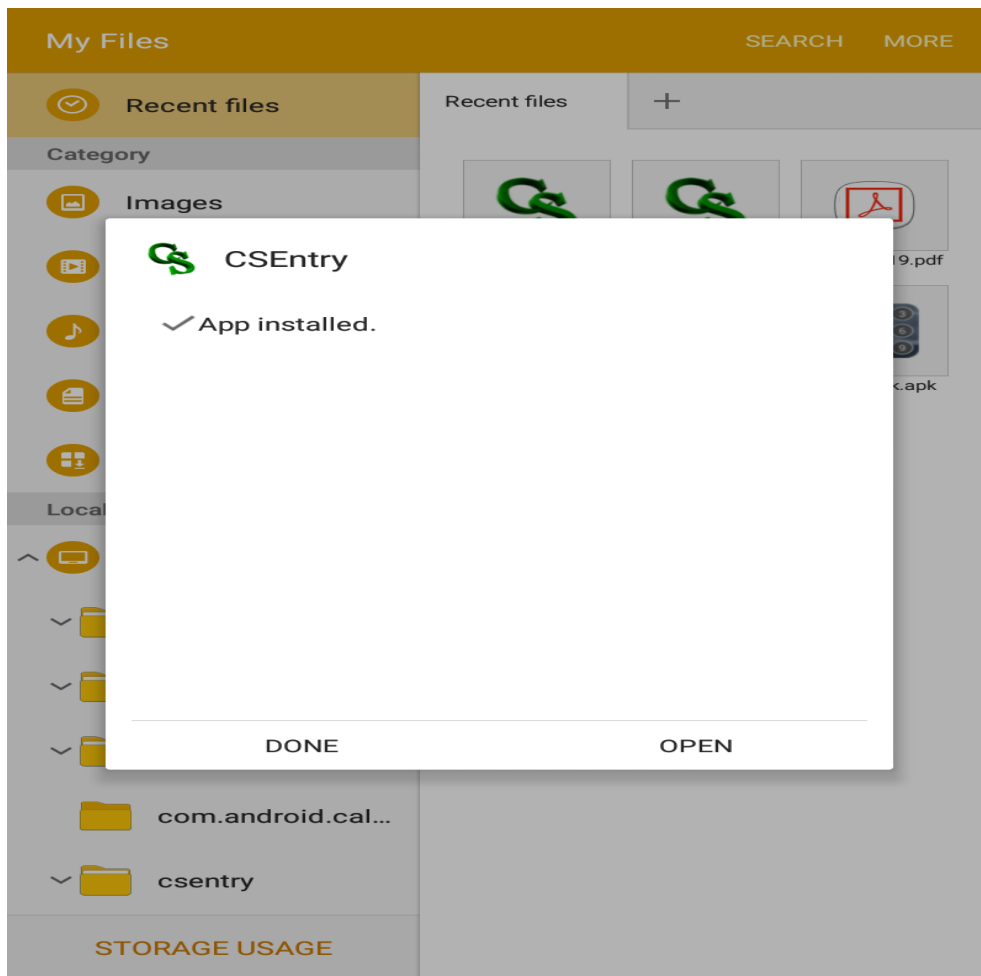


Figure 4.1-13: Installed CSEntry ApplicationSource: Research Data

CSEntry

United States Census Bureau

Terms of Service
Please Read Before Accepting

Terms of Service for CSPro

"CSPro" is an application made available by the U.S. Census Bureau (Census) that provides individuals and entities a tool for data processing.

Census does not require individuals or entities to register or provide any personally identifiable information (PII) to Census as a condition of downloading or using CSPro. Census will not have access to any information regarding individuals or individuals' devices using the CSPro application.

The Terms of Service for use of the above application are as follows:

A. CSPro is provided to You, as the user and those who may take by, through or under it, as is, without any warranty (whether express or implied) or representation whatsoever, including but not limited to any warranty of merchantability. CSPro is taken hereunder without any right to support or to any improvements, extensions, or modifications, except as may be agreed to separately, in writing, by Census.

B. You hereby and forever waive, release and discharge the United States/Department of Commerce and all its instrumentalities, including Census (together, Commerce), from any and all liabilities and obligations in connection with the use, application, sale or conveyance of CSPro. You agree that this term should be given the broadest interpretation possible in favor of the release of Commerce under applicable law.

C. You will not use the logos, emblems, or trademarks of Commerce on any product or service that is directly or indirectly related to these terms without the prior written approval of Commerce. You will not in any way imply that Commerce endorses your use of CSPro.

D. You agree to make a good faith effort to use CSPro in a way that does not cause damage, harm, or embarrassment to Commerce. Commerce expressly reserves all rights and remedies.

CANCEL **ACCEPT**

Figure 4.1-14: Terms of Service for Cspro

4.7 Research Instrument

4.7.1 CAPI Menu Code

Conduct the interview function

The conduct interview function opens the main application.

Send data /update application

After collecting data in the field, one needs to get the data from the interviewer's devices back to headquarters to create a combined data file. While this can be done manually by connecting each device to a laptop, copying the individual data files and combining them using the Concatenate Data tool, this is rather inconvenient for large surveys. Instead, it is possible to send the data from each tablet over the internet to a central server that will combine the data into a single data file to be used for further processing. In CSPro this is called synchronization. Using the send data/update application, which is a two-way function

as you can send data as well as update the program. In case there are any changes to the application, the update will be done when you select the send data/update application.

Logout

The logout function allows one to logout of the user account.

Exit the application

The exit application function helps to exit.

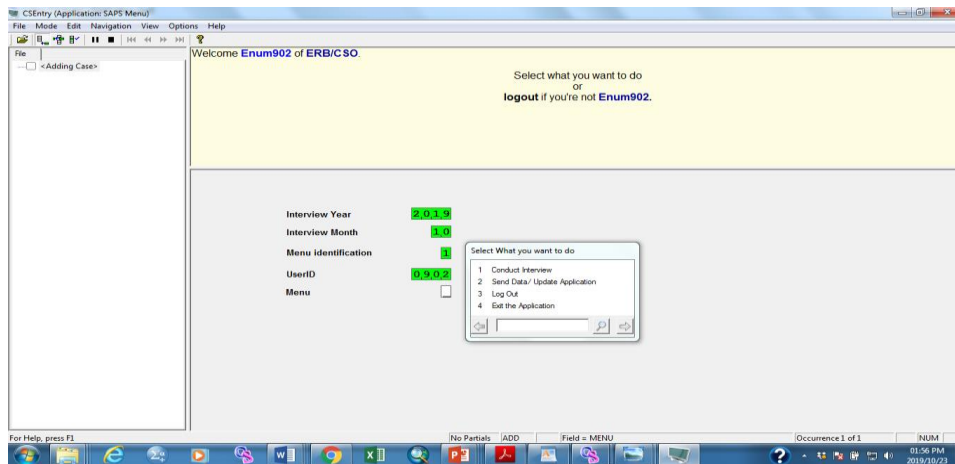


Figure 4.1-15: CAPI Menu Application

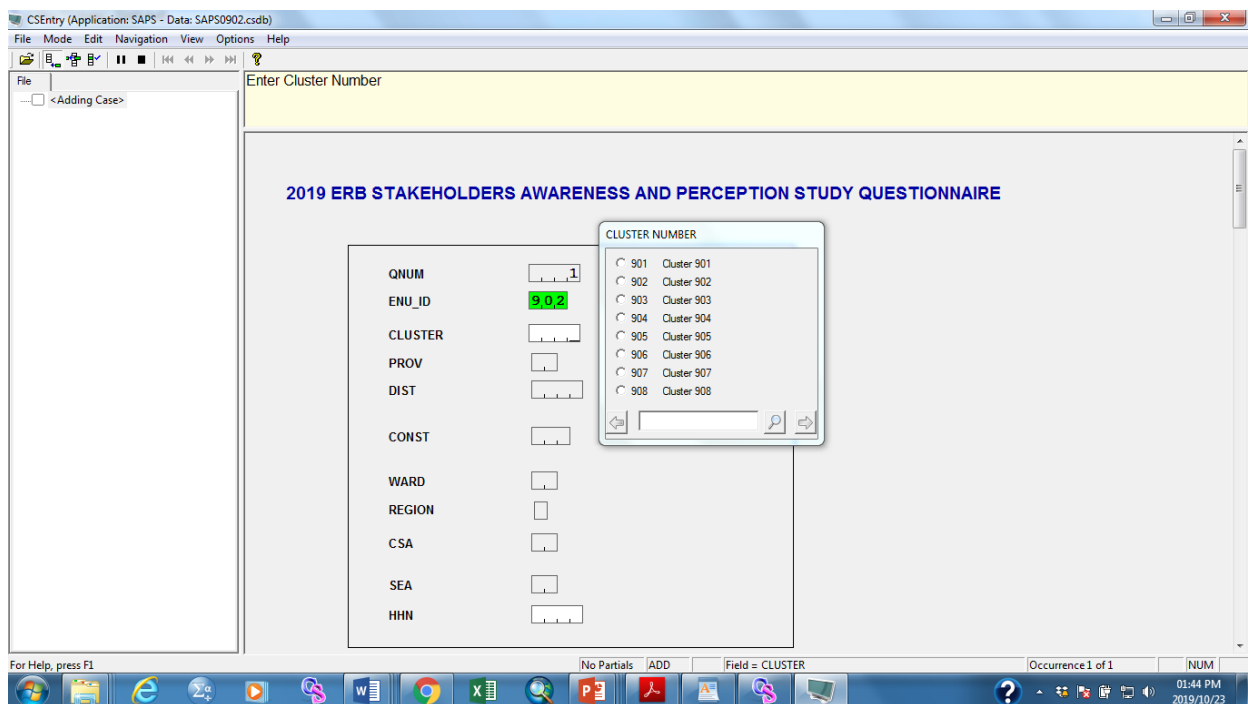


Figure 4.1-16: Data Entry into the System

4.8 Saving of Cases

Saving a case writes the contents of the current case to the data file, including any changes you would have made. This includes all records, of all types, into which data had been entered. If a record type had been declared **Required** (in the data dictionary), and no data was entered for this record type, a blank record of the required type would be written as part of the questionnaire, and a message informing you of this action will be issued. There is no way to suppress this message. Records are written to the output file in the order in which the record types were defined in the data dictionary.

If the application is defined as **operator-controlled**, you may use the **F12** or **Ctrl+F12** keys to save a case. If the application is more than one level, use the **Ctrl+F12** key.

If the application is defined as a **system-controlled** type, you may use the **F10** key (Advance to End) to save a case. This key would execute all the procedures from the current field through the end of the case. The application may contain underlying logic which stops you at a certain point and displays a message. The system will only save the case if there are no such messages.

The following are true for both operator and system-controlled mode:

- (i) If one is in **add** mode, saving the case will keep the user in add mode and allowing the operator to continue adding cases.
- (ii) If the user is in **modify** mode, saving the case will terminate modify mode. The operator has to select a new case and to get back into modify mode.

4.9 Synchronizing and Sending of Data

A data entry application may be able to synchronize data with a remote server. If that is the case, the **Synchronize** menu option (off the **File** menu) would be enabled when the user is not currently working in any case. When one selects this option, CSEntry will attempt to connect to the remote server and then synchronize data, as well as possibly the data entry application itself, using parameters specified by the application programmer. You may have to enter a username and/or password before the synchronization process begins.

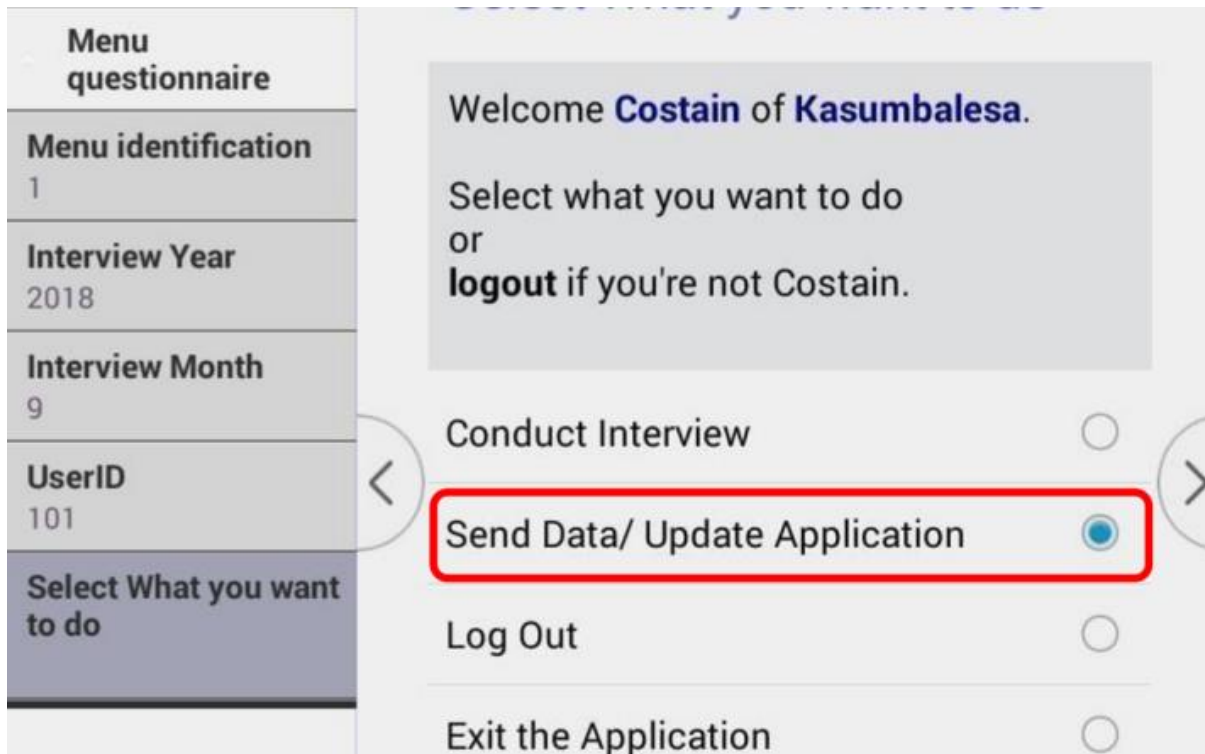


Figure 4.1-17: Data and update application

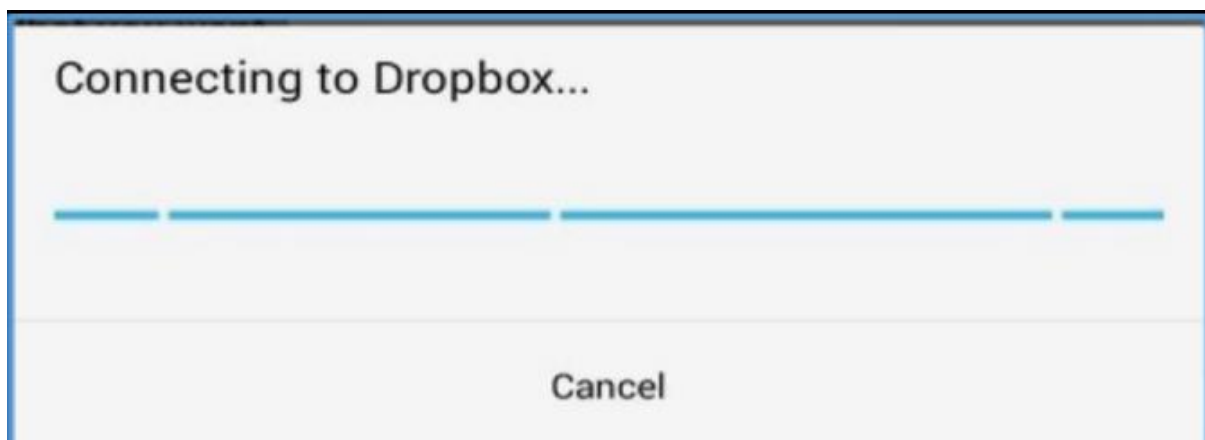


Figure 4.1-18: Connection to the Server

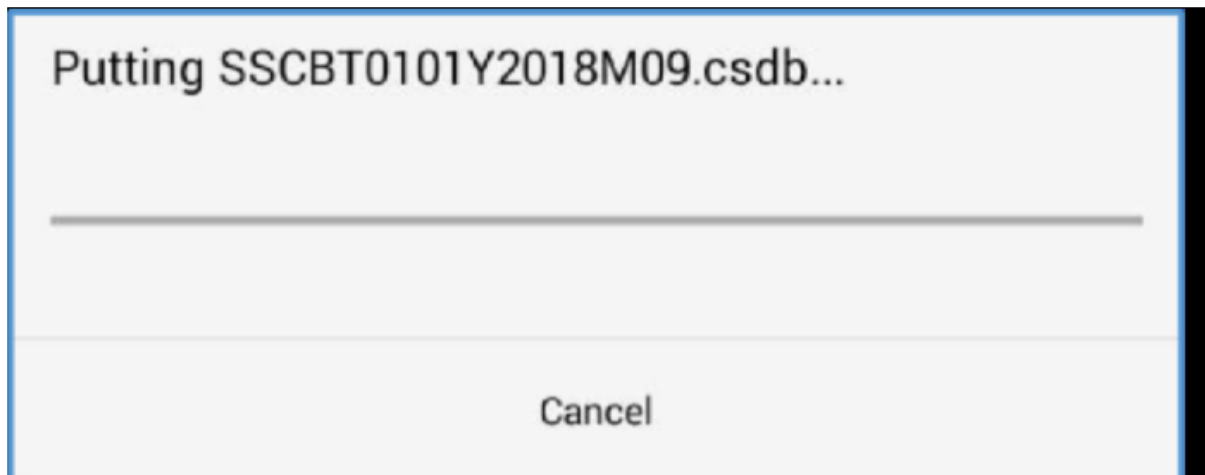


Figure 4.1-19: Data Being put to the Server

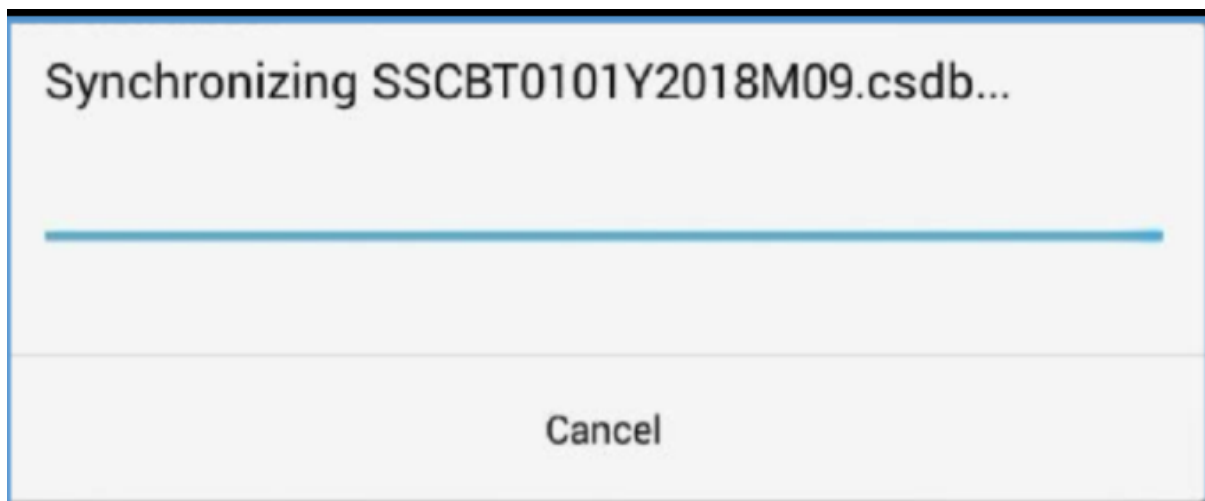


Figure 4.1-20: Synchronisation of Data

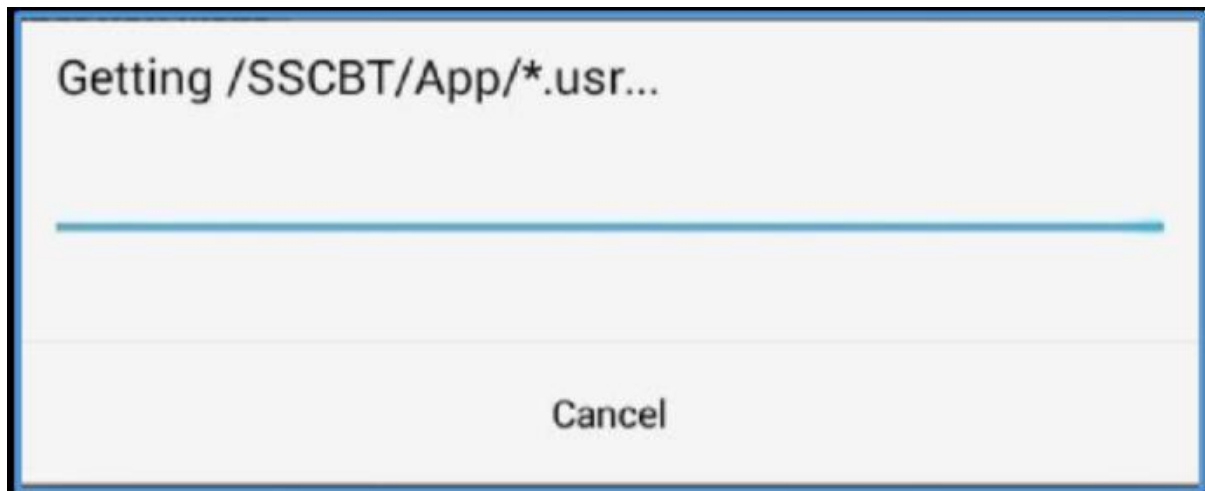


Figure 4.1-21: Updating of Application

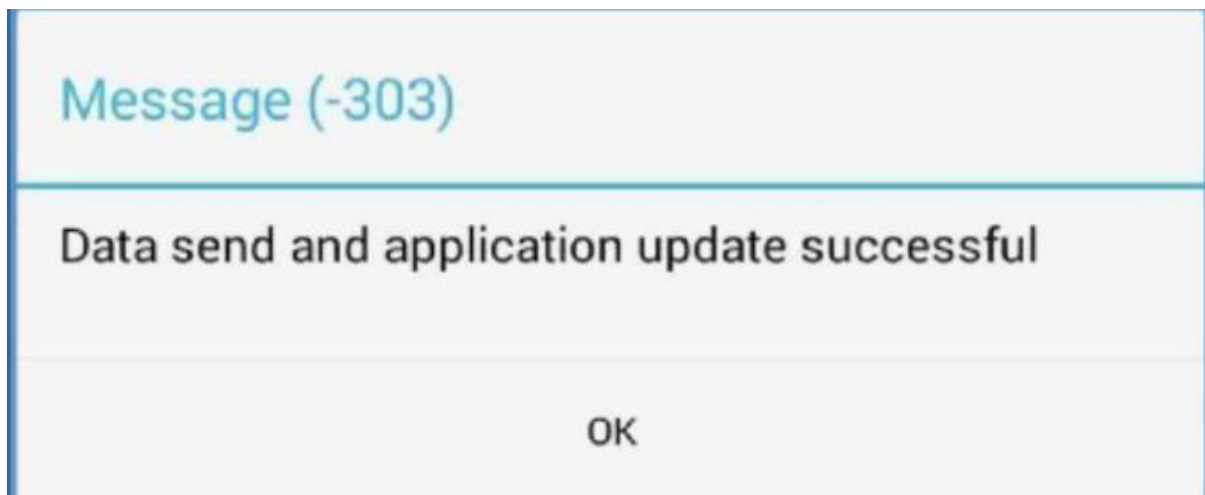


Figure 4.1-22: Successfully Updated

4.10 Retrieving data from the Server

CSPRO's Dropbox sync does not store a single data file in Dropbox. Instead, it stores a file in /CSPRO/DataSync/ for every synchronization. Because of this, the user will not see the data as a single file. In order to obtain the combined data file, use the Data Viewer to download the data from Dropbox. Each data file has an enumerator id and to have one data file these different data files which are in cspro database format have to be downloaded and later concatenated into one file. The figure below shows how the data comes in from each tablet.

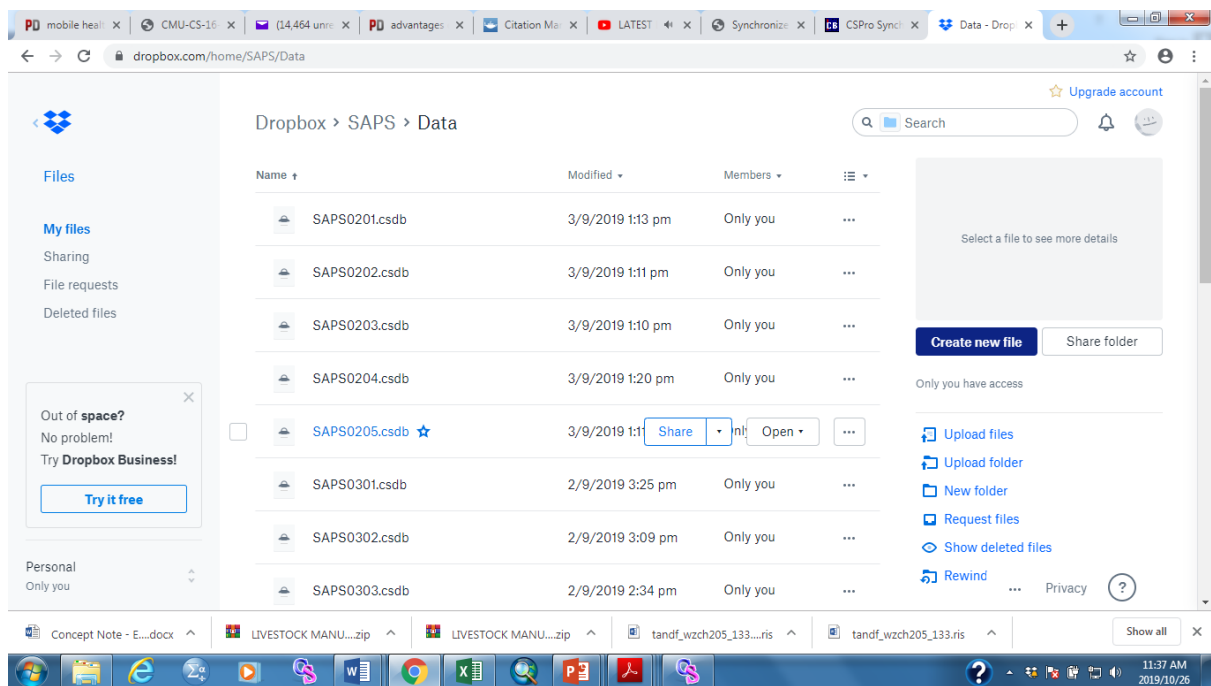


Figure 4.1-23: Data Retrieval from the Server

4.11 Tabulation of Data and Report Writing

The CPro export data tool is available from the Tools menu. When you first start Export Data you are prompted to provide a data dictionary. From the main export screen, you can choose which records/variables to export using the checkboxes next to each one. To run the export, click on the traffic light. You are prompted to choose the data file. Once the export completes the exported files are displayed in Text Viewer. The following table is data exported into SPSS ready to be analysed.

erb0710finalEDITTED.sav [DataSet1] - IBM SPSS Statistics Data Editor													
File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help													
Visible: 92 of 92 Variables													
	ENU_ID	QNUM	CLUSTER	PROV	DIST	CONST	WARD	REGION	CSA	SEA	HHN	ZESCO_ELECTRICITY	NAME_HEAD
1	201	3	Cluster 204	COPPE...	KITWE	KAMFINS	16	Urban	7	2	3	Yes	MUTAMBO PETER
2	201	4	Cluster 204	COPPE...	KITWE	KAMFINS	16	Urban	7	2	4	Yes	MWESHI CLEMENT
3	201	6	Cluster 204	COPPE...	KITWE	KAMFINS	16	Urban	7	2	6	Yes	KASHAYA BUPE
4	201	10	Cluster 213	COPPE...	KITWE	WUSAKILE	18	Urban	9	3	3	Yes	BWALYA MOSES
5	201	11	Cluster 213	COPPE...	KITWE	WUSAKILE	18	Urban	9	3	4	Yes	LUPANDO JOYCE
6	201	12	Cluster 212	COPPE...	KITWE	WUSAKILE	19	Urban	6	1	5	Yes	CHOMBA MORGAN
7	201	16	Cluster 212	COPPE...	KITWE	WUSAKILE	19	Urban	6	1	2	Yes	KAPUKA ANNAH
8	201	18	Cluster 212	COPPE...	KITWE	WUSAKILE	19	Urban	6	1	4	Yes	CHISHIMBA MARTHA
9	201	19	Cluster 212	COPPE...	KITWE	WUSAKILE	19	Urban	6	1	5	Yes	CHANGWE FRANCIS
10	201	20	Cluster 212	COPPE...	KITWE	WUSAKILE	19	Urban	6	1	6	Yes	KAMPAMBA CLARENCE
11	201	22	Cluster 212	COPPE...	KITWE	WUSAKILE	19	Urban	6	1	8	Yes	MWILU GOLDEN
12	201	24	Cluster 201	COPPE...	KITWE	CHIMWE...	1	Urban	6	1	2	No	NTANKULA EMMANUEL
13	201	25	Cluster 201	COPPE...	KITWE	CHIMWE...	1	Urban	6	1	3	No	BWEMBYA PETER
14	201	31	Cluster 203	COPPE...	KITWE	CHIMWE...	2	Urban	9	1	2	No	KAKWENDE FREDERICK
15	201	32	Cluster 203	COPPE...	KITWE	CHIMWE...	2	Urban	9	1	3	Yes	NKUMBULA BORNFACE
16	201	36	Cluster 203	COPPE...	KITWE	CHIMWE...	2	Urban	9	1	7	No	SAMAKAJ MARJORY
17	201	37	Cluster 210	COPPE...	KITWE	NKANA	27	Urban	3	1	1	Yes	MWAMBA EMMANUEL
18	201	38	Cluster 210	COPPE...	KITWE	NKANA	27	Urban	3	1	2	Yes	CHILUFYA MUMBA EMMANUEL
19	201	39	Cluster 210	COPPE...	KITWE	NKANA	27	Urban	3	1	3	Yes	CHIBAKA GIBSON
20	201	40	Cluster 210	COPPE...	KITWE	NKANA	27	Urban	3	1	4	Yes	SICHILYA KENNEDY
21	201	42	Cluster 210	COPPE...	KITWE	NKANA	27	Urban	3	1	6	No	CHALWE CHOLA
22	201	43	Cluster 210	COPPE...	KITWE	NKANA	27	Urban	3	1	7	Yes	SIMUNENE LUCKY
23	201	44	Cluster 202	COPPE...	KITWE	CHIMWE...	3	Urban	17	3	1	Yes	SINKALA BORNFACE

Figure 4.1-24: Data Ready for Tabulation and Analysis

4.12 Testing

Testing is a phase in software development that shows that a program does what it is intended to do. It aids in the discovery of a program's defects before it is put into use (Boston:Addison-Wesley, 2011).

Table 4.1.2: Implementation of Testing Evaluation of the Prototype.

Actor	Test ID	Scenario	Actual Results	Expected Results	Test Data	Comments
User	1	The user authenticates him/herself before accessing system functionality.	The user was logged in to the system upon supplying the correct login credentials	The user should be logged in to the system when they provide the correct login credentials and when an incorrect username and password is entered a message should be given When they fail to do so.	Username: 0101Chenela Password: Chenela@ERB	The login module worked as expected.
		The user connects to the server to enable communication between the hardware and the software application	The user successfully Connected to the communication port provided the hardware was connected to the internet.	The user should be able to Connect to the communication port provided the hardware was connected to the internet.		The user successfully connected to the communication port .

Actor	Test ID	Scenario	Actual Results	Expected Results	Test Data	Comments
		The user enter data on the main application on the tablet	Data was successfully displayed on the screen in a text box.	Data should be seen on the tablet application in a text box.		Data was successfully entered.

4.13 Summary

In this chapter, the data that was collected from the survey was analysed and presented in form of Bar and Pie Charts. The researcher established that the data collection system currently experiences difficulties, they physically collect data using a paper-based manual system for storage of records and derivation of reports. Under system implementation, the designs from the interviews conducted were derived and illustrated, and the proposed designs for automation of the current manual processes were presented. Finally, the screenshots from the system prototype implementation were presented and discussed.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter is a summary of the conclusions as well as the recommendations drawn from the key findings of the study. The purpose of this study was to establish a mechanism from which mobile data collection could be used to improve speed, accuracy, productivity gain, prevention of errors and immediate access to data. The purpose of this study was further broken down into the following specific objectives namely:

- (i) To investigate the challenges faced by the Zambia Statistics Agency in the process of data collection using PAPI.
- (ii) To develop the benefits of using an integrated and automated mobile data collection system
- (iii) To establish how an integrated and automated mobile data information system could be used to improve the data collection process.

5.2 Conclusions

The study that was conducted identified challenges in the current data collection systems such as delayed time interval between data collection and data checking which resulted in erroneous data at a point when it was too late to make corrections. The main errors detected were omissions and illogical data. This current data collection system is very costly in terms of time and resources which make the whole process to become challenging when it comes to implementation.

With the challenges identified, a new system was developed which offers direct data entry at the initial point of contact and addresses the problems identified above which are mostly associated with the current data collection system. The developed system has also proved to

be more accurate and complete. The system also helped in the collection of data from eligible respondents. Moreover, the developed system improves on time management in the collection of data and is very much cost-effective.

5.3 Recommendation

The study has revealed that the automated mobile data collection system is desirable. The researcher then recommends that:

- (a) Zambia Statistics Agency adopts the proposed model in order to improve on the efficiency of data collection.
- (b) That the users should be educated on the importance of data collection applications and how to use the electronic mobile data devices and this is to be done by the master trainers, IT personnel and the subject matter specialist so that the field staff understand both the concepts and the use of the devices.
- (c) Capacity is built in the programmers to equip them with the knowledge for developing these applications because the success of CAPI depends substantially on the effort spent programming, piloting and testing the application, as well as on careful consideration to the underlying data management and transfer systems.

REFERENCES

- Chen H, Hailey D, Wang N, Yu P.(2014) A Review of Data Quality Assessment Methods for Public Health Information Systems.
- Basch E, Goldfarb S. (2009)Electronic patient-reported outcomes for collecting sensitive information from patients.
- Garritty C, Emam K.(2006) Who's using PDAs? Estimates of PDA use by health care providers.
- Welker J.(2007) Implementation of electronic data capture systems.
- Brandt C, Argraves S, Money R, Ananth G, Trocky N, Nadkarni P.(2006)
- Caban-Martinez A, Clarke T, Davila E, Fleming L, Lee D.(2011) Application of handheld devices to field research among underserved construction worker populations.
- Walther B, Hossin S, Townend J, Abernethy N, Parker D, Jeffries D.(2011)
- Asare-aboagye Y.(2012) Setting Up Electronic Data Capture Capabilities. In: Global Clinical
- Lauraitis, R. Maskeliūnas, R. Damaševičius, D. Połap and M. Woźniak (2019), "A Smartphone Application for Automated Decision Support in Cognitive Task Based Evaluation of Central Nervous System Motor Disorders.
- Bet Caeyers (2010) A Comparison of CAPI and PAPI through a Randomized Field.
- K. Biron, W. Mansoor, S. Miniaoui, S. Atalla, H. Mukhtar and K. F. Bin Hashim(2019), "Data Science Tools for Crime Investigation, Archival, and Analysis.
- A.Dankar and P. P. Kundapur(2019) Automated Mobile Attendance System

Konstantinos Rammos, Spyros Papadimitriou, Maria Virvou, Efthimios Alepis(2018), "A Web-Based Application for Innovative Hospital Appointment Scheduling Using Neural Network.

J. Schobel, R. Pryss, M. Schickler and M. Reichert(2016) Towards Flexible Mobile Data Collection in Healthcare.

B. R. Hermanto, A. Indrayanto and A. Sasongko(2013)Smart card mobile data collection system concept for health and medical data collecting activities in rural area.

M. Katarahweire, E. Bainomugisha and K. A. Mughal(2017) Authentication in Selected Mobile Data Collection Systems:

H. Zhong et al(2010) A solution for the data collection in the field survey based on Mobile and Wireless GIS.

J. Schobel, R. Pryss and M. Reichert(2015) Using Smart Mobile Devices for Collecting Structured Data in Clinical Trials.

Y. Anokwa, C. Hartung, A. Lerer, B. DeRenzi and G. Borriello(2009) A new generation of open source data collection tools.

[59]ITU. ICT facts and figures. 2013. Available from: <http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013-e.pdf>.

David G. Dillona(2014), "Open-source electronic data capture system offered increased accuracy".

Anokwa Y, Hartung C, Brunette W, Borriello G, Lerer A. Open Source Data Collection

Bertino, E., Ferrari, E., & Perego, E. (2002). Max: An access control system for digital libraries and the web.

Trnka, Michal, Černý, Tom , Stickney, Nathaniel (2018) .Survey of Authentication and Authorization for the Internet of Thing.

Zhang S, Wu Q, van Velthoven MH, Chen L, Car J, Rudan I, Zhang Y, Li Y, Scherpbier RW. Smartphone versus pen-and-paper data collection of infant feeding practices in rural China.

Sommerville, Software Engineering, 9th ed., Boston: Addison-Wesley, 2011.

Singleton, K. W. et al.(2011), Wireless Data Collection of Self Administered Surveys Using Tablet Computers.

Shao D. A(2012) Proposal of a Mobile Health Data Collection and Reporting System for the developing countries.

Jasmeet, Singh P(2014) Review of smartphone operating systems like Symbian OS and Windows mobile.

Adama Baguiya(2016) An offline mobile data capture module for health and demographic surveillance system

W3C. Extensible mark-up language (XML) 1.0. 5th ed;

Ahmed, M. Y., (2001), “Welcome Address at the Opening Ceremony of the National Assisted Psychological and Psychophysiological Methods in Monitoring and Field Studies.

Baker, R.P. (1992). New technology in survey research: Computer assisted personal

Baker, R.P. (1998). Computer assisted survey information collection.

Bates, N. (2001). Internet versus mail as a data collection methodology .

Beierle, F., Tran, V. T., Allemand, M., Neff, P., Schlee, W., Probst, T., et al. (2018b).

Context data categories and privacy model for mobile data collection apps.

Blaya JA, Fraser HSF, Holt B(2010) E-health technologies show promise in

Bleidorn, W., and Hopwood, C. J. (2018). Using machine learning to advance.

Bowers, K. D., Juels, A., Oprea, A. HAIL(2009) a high-availability and integrity layer for cloud storage Proceedings of the 16th ACM conference on Computer and Communications Security November 2009.

Bradburn, N.M., Frankel, M.R., Baker, R.P. & Pergamit, M.R. (1992). A comparison of CAPI with PAPI.

C.R . Kothari(2004) Research Methodology.

Cho, Ying-Chiang, Pan, Jen-Yi (2013).Vulnerability Assessment of IPv6 Websites to SQL Injection and Other Application Level Attacks.

CloudComputingArchitecture,<https://www.google.com/search?q=Cloud+Computing+Architecture>, [Accessed 20 March 2020].

Walther B et al (2011), Comparison of electronic data capture (EDC) with the standard data Capture method for clinical trial data.

Davis FD and Venkatesh V(2004)Toward Preprototype User Acceptance testing of New Information Systems.

De leeuw, Edith(2008) The effect of computer-assisted interviewing on data quality.

Carina King et al(2014) electronic data collection process in clinical trials experiences with different systems in Malawi.

Fahrenberg, J. and Myrtek, M. eds. (1996). Ambulatory Assessment:

Furukawa MF, King J, Patel V, Hsiao C-J, Adler-Milstein J, Jha AK.(2014) Despite substantial progress In EHR adoption, health information exchange and patient engagement remain low in office settings .

G. Jing, H. Wei, W. Ziyang, Z. Chao and J. Haitao (2019), "Research on Android Component Security Testing Method for Mobile Application in Power System.

Gustarini, M., Wac, K., and Dey, A. K. (2016). Anonymous smartphone data collection

Harrison R (2014) Usability of mobile applications: literature review and rationale for a new usability model.

iConverse Whitepaper, (March 2001) Zambia's mobile phone sector in near saturation – Zicta

J. Schobel, R. Pryss, M. Schickler and M. Reichert (2017), "Towards Patterns for Defining and Changing Data Collection Instruments in Mobile Healthcare Scenarios.

Johnson SD,(1998) “Productivity, the Workforce, and technology Education.

King, Carina (2014). Electronic data capture in a rural African setting

Ladouce, S., Donaldson, D. I., Dudchenko, P. A., and Ietswaart, M. (2017).

Luna D, Almerares A, Mayan JC 3rd, Gonz_alez Bernaldo de Quir_os F, Otero C.(2014)

Health informatics in developing countries

Olst, R. v., n.d (2004) Electronic data collection using mobile computing.

Pavlovic I, Kern T, Miklavcic D.(2009) Comparison of paper-based and personality assessment and theory.

Rosero-Bixby, Luis Hidalgo-CÃ©spedes, Jeisson Antich-Montero, Daniel Seligson, Mitchell. Improving the Quality and Lowering Costs of Household Survey Data Using Personal Digital Assistants (PDAs). An Application for Costa Rica.

Schiffman, J., Moyer, T., Vijayakumar, H., Jaeger, T., McDaniel, P.(2010) Seeding clouds with trust anchors Proceedings of the ACM workshop on Cloud computing security workshop.

Schuster C, Perez Brito C.(2011) Cutting costs, boosting quality and collecting data real-time. Security and Privacy in Cloud Computing.

Yunchuan Sun et al. (2014) Data security and privacy in cloud computing.

Tim Rochford (2001), “The Impact of Mobile Application technology on Today’s Workforce”.

Tomlinson M, Solomon W, Singh Y, Doherty T, Chopra M, Ijumba P, et al.(2009) The use of mobile phones as a data collection tool.

Weiskopf NG, Weng C.(2013) Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research.

WHO (World Health Organization). Framework and Standards for Country Health Information Systems (2008) 2nd Editio(January):72.

Z. Zhou, C. Sun, J. Lu and F. Lv,(2018) "Research and Implementation of Mobile Application Security Detection Combining Static and Dynamic.

Gangadharan, S. P., et al. (2017). Ten simple rules for responsible big data research.
www.w3.org/TR/REC-xml/#sec-origin-goals

APPENDIX A: Data collection permission letter



UNIVERSITY OF ZAMBIA

SCHOOL OF ENGINEERING

INTEGRATED AND AUTOMATED MOBILE DATA INFORMATION SYSTEMS IN STATISTICS - A CASE OF CENTRAL STATISTICAL OFFICE.

QUESTIONNAIRE

Dear Respondent,

My name is **CHENELA NKHOWANI (0977442197)** and I am a student at the University of Zambia pursuing a Master of Engineering degree in Information and Communication Technology (ICT). I am carrying out a baseline study to profile the data collection practices by central statistical office officers. The research is aimed at designing a computerized mobile data collection system to improve data quality through logical validation rules in the questionnaire. In order to accomplish this study, I need input from the **central statistical office** on the experiences and knowledge about data collection to assist with information on the research subject.

Kindly help by providing answers to the questions below and provide attachments where possible. Your responses will be treated with the highest level of confidentiality and will strictly be used for academic purposes.

Thank you in anticipation

Research Supervisor: Abraham Ngoyi- Email – abrahamngoyi@yahoo.com

Assistant Dean-Postgraduate: Dr. C. Kahanji – Email – Charles.kahanji@unza.zm

<u>SECTION A – BACKGROUND INFORMATION</u>		
1.Station	1= CSO HQ 2= Provincial office	☐
2.Department	1= IT 2= Demography 3= Administration 4= Accounts 5=Labour 6= Dissemination 7= GIS 8 = LCMS	☐
3.Sex	1= Male 2= female	☐
4. What is your age	1= 18 – 29 2= 30 – 45 3= 46 – 55 4= Above 55	☐
5.Highest Education Level completed	1= School Certificate 2= College Certificate 3= Diploma 4= Degree 5= Master's Degree 6= PHD	☐

6. How long have you worked for this institution?	1. Less than 1 year 2. 1-5 years 3. 6-10 years 4. 11-15 years 5. 16-20 years 6. Above 20 years	☐
<u>SECTION B – DATA COLLECTION SYSTEM</u>		
7. Do you use any ICT equipment in doing your day to day work?	1= Yes 2=No	☐
8. What is your level of knowledge in computers?	1= Basic 2= Intermediate 3= Excellent	☐
9. What Operating System are you mainly conversant with?	1= Android 2= Apple 3= Windows	☐
10.Does your Institution have Computers	1 = Yes 2= No	☐
11. Does your institution have a system for data collection?	1 = yes 2= No >> .skip to 13	☐
12. What kind of a system do you mostly use for data collection?	1= Paper Questionnaire 2= Electronic	☐

13. What are the advantages of the system you chose in Q12?		
14. What do you think are the drawbacks of using paper questionnaire?		
15. Do you think using electronic questionnaire can reduce some of these drawbacks mentioned above?	1 = Yes 2= No	☐
16. Do you agree that dissemination period is shortened by using Electronic questionnaire?	1= strongly agree 2= agree 3= Neutral 4 = disagree 5 = strongly disagree	☐
17. Skip patterns can be difficult for enumerators to follow when using paper questionnaire	1= strongly agree 2= agree 3= Neutral 4 = disagree 5 = strongly disagree	☐
18. Using paper questionnaire is more expensive than using electronic questionnaire for example when printing.	1= strongly agree 2= agree 3= Neutral 4 = disagree 5 = strongly disagree	☐
19. Data is instantly digitized as it is entered in to the tablet.	1= strongly agree 2= agree 3= Neutral 4 = disagree 5 = strongly disagree	☐

END OF INTERVIEW – THANK YOU

APPENDIX B: System prototype code

Logging in to the application

```
PROC GLOBAL
```

```
file SAPSFile;//,SAPSREPORT; numeric egoods, igoods, eservices, iservices, totalG, totalS;  
string user, boarder; function changeUser()
```

```
LOGIN = notappl;savesetting(clear);savesetting("user", "");if getsymbol() = "LOGIN" then  
reenter; else reenter LOGIN; endif;end;
```

```
PROC MENU_FF PROC MENU_ID preproc $=1; onfocus noinput;
```

```
PROC INTERV_YEAR preproc $ = sysdate("YYYY"); onfocus noinput;
```

```
PROC INTERV_MONTH
```

```
Preproc $ = sysdate("MM"); onfocus noinput; PROC LOGIN
```

```
Onfocus if loadsetting("userid") <> "" then LOGIN = tonumber(loadsetting("userid"));  
noinput; endif; if visualvalue(LOGIN) then noinput; endif; postproc if  
loadcase(LOGIN_DICT, LOGIN) then //Warning("User %s has been found at %s  
Station",strip(USERNAME),strip(STATION));
```

```
    user = strip(USERNAME);
```

```
    boarder = strip(STATION);
```

```
else
```

```
Warning("This user ID is not in the database, please contact HQ for your correct user  
ID");
```

```
reenter;
```

```
endif;
```

```

savesetting("userid", maketext("%04d",LOGIN));

string mystring = "Welcome " + user + ", enter your password to continue:";

string pwd = prompt(mystring, password);

if strip(PASSWORD) = strip(pwd) then

//Warning("Password accepted");

else

Warning("Incorrect password");

reenter;

endif;

PROC MENU

if MENU = 1 then

//BORDER

//write pff for opening the data entry program

setfile(SAPSFile,"SAPS.pff", create);

filewrite(SAPSFile,"[Run Information]");

filewrite(SAPSFile,"Version=CSPro 7.1");

filewrite(SAPSFile,"AppType=Entry");

filewrite(SAPSFile,"");

filewrite(SAPSFile,"[DataEntryInit]");

filewrite(SAPSFile,"Interactive=Ask");

filewrite(SAPSFile,"");

```

```

filewrite(SAPSFile,"[Files]");

filewrite(SAPSFile,"Application=.\SAPS.ent");

filewrite(SAPSFile,"InputData=..\Data\SAPS%s.csdb",edit("9999",LOGIN));

filewrite(SAPSFile,"Paradata=.\SAPS.cslog");

filewrite(SAPSFile,"");

filewrite(SAPSFile,"");

filewrite(SAPSFile,"[Parameters]");

filewrite(SAPSFile,"XSTATION=%s",strip(STATION));

filewrite(SAPSFile,"XUSERNAME=%s",strip(USERNAME));

filewrite(SAPSFile,"XLOGIN=%04d",LOGIN);

filewrite(SAPSFile,"XINTERV_YEAR=%04d",INTERV_YEAR);

filewrite(SAPSFile,"XINTERV_MONTH=%02d",INTERV_MONTH);

close(SAPSFile);

execpff("SAPS.pff", stop);

connecting to the dropbox

elseif MENU = 2 then

// Data upload to dropbox

// connect to Dropbox

if connection() then

if syncconnect(Dropbox)= 1 then

//if      syncconnect(csweb,      "http://154.120.217.101/csweb/api",      "admin",
"Comstat@2019")= 1 then

```

```

string datafile = maketext("../Data/SAPS%s.csdb", edit("9999",LOGIN));

if fileexist(datafile) then

setfile(SAPS_DICT, datafile);

datafile = maketext("../Data/SAPS%s.csdb", edit("9999",LOGIN));

if fileexist(datafile) then

setfile(SAPS_DICT, datafile);

endif;

if      syncfile(PUT,      datafile,      maketext("/SAPS/Data/SAPS%s.csdb",
edit("9999",LOGIN))) = 1 then

//and syncdata(PUT,SAPS_DICT) = 1 then

//if syncfile(PUT,"/SAPS/Data/*.csdb") = 1 and syncdata(PUT,SAPS_DICT) = 1 then

//Commented the updating of program

syncfile(GET,"/SAPS/App/*.pff");

syncfile(GET,"/SAPS/App/*.pen");

syncfile(GET,"/SAPS/App/*.usr");

//syncfile(GET,"/SAPS/Docs/SAPSS Manual.pdf","../Docs/SAPSS Manual.pdf");

//syncfile(GET,"/SAPS/Docs/SAPS   CAPI   Manual.pdf","../Docs/SAPS   CAPI
Manual.pdf");

errmsg("Data send and application update successful");

else

errmsg("Uploading data failed, try again");

endif;

```

```

else

errmsg("No data to send for Interviewer %04d", USERID);

endif;

syncdisconnect();

else

errmsg("Connection to Dropbox server failed");

endif;

reenter MENU;

syncdisconnect();

else

errmsg("There is no internet connectivity");

endif;

elseif MENU = 3 then

changeUser();

elseif MENU = 4 then

//errmsg("Are you sure you want to exit?");

//select("Yes",endlevel,"No",MENU);

stop(1);

endif;

```

System prototype code to run the back-end of the CAPI system.

PROC GLOBAL

PROC SAPS_FF

PROC QNUM

```

preproc
noinput;
PROC ENU_ID
preproc
$= tonumber (sysparm("XLOGIN"));
onfocus
noinput;
PROC CLUSTER
preproc
if ENU_ID in 201:299 then setvalueset (CLUSTER, CLUSTER_VS1);
elseif ENU_ID in 301:399 then setvalueset (CLUSTER, CLUSTER_VS2);
elseif ENU_ID in 501:599 then setvalueset (CLUSTER, CLUSTER_VS3);
elseif ENU_ID in 701:799 then setvalueset (CLUSTER, CLUSTER_VS4);
elseif ENU_ID in 901:999 then setvalueset (CLUSTER, CLUSTER_VS5);
endif;
PROC PROV
preproc
if CLUSTER in 201:214 then      $ =2;
elseif CLUSTER in 301:308 then  $ = 3;
elseif CLUSTER in 501:528 then  $ = 5;
elseif CLUSTER in 701:706 then  $ = 7;
elseif CLUSTER in 901:908then   $ = 9;
endif;
PROC DIST
{PreProc
if PROV = 1 then setvalueset(DIST, DIST_VS2); // Central
elseif PROV = 2 then setvalueset(DIST, DIST_VS3); // Copperbelt
elseif PROV = 3 then setvalueset(DIST, DIST_VS4); // Eastern
elseif PROV = 4 then setvalueset(DIST, DIST_VS5); // Luapula
elseif PROV = 5 then setvalueset(DIST, DIST_VS6); // Lusaka
elseif PROV = 6 then setvalueset(DIST, DIST_VS7); // Muchinga
elseif PROV = 7 then setvalueset(DIST, DIST_VS8); // Northern
elseif PROV = 8 then setvalueset(DIST, DIST_VS9); // Northwestern

```

```

elseif PROV = 9 then setvalueset(DIST, DIST_VS10); // Southern
elseif PROV = 10 then setvalueset(DIST, DIST_VS11); // Western
else
errmsg("Province code %d is not a valid province cod", PROV);
endif;
}

preproc
if CLUSTER = 201 then $ = 204 ;
elseif CLUSTER = 202 then $ = 204 ;
elseif CLUSTER = 203 then $ = 204 ;
elseif CLUSTER = 204 then $ = 204 ;
elseif CLUSTER = 205 then $ = 204 ;
elseif CLUSTER = 206 then $ = 204 ;
elseif CLUSTER = 207 then $ = 204 ;
elseif CLUSTER = 208 then $ = 204 ;
elseif CLUSTER = 209 then $ = 204 ;
elseif CLUSTER = 210 then $ = 204 ;
elseif CLUSTER = 211 then $ = 204 ;
elseif CLUSTER = 212 then $ = 204 ;
elseif CLUSTER = 213 then $ = 204 ;
elseif CLUSTER = 214 then $ = 204 ;
elseif CLUSTER = 301 then $ = 302 ;
elseif CLUSTER = 302 then $ = 302 ;
elseif CLUSTER = 303 then $ = 302 ;
elseif CLUSTER = 304 then $ = 302 ;
elseif CLUSTER = 305 then $ = 302 ;
elseif CLUSTER = 306 then $ = 302 ;
elseif CLUSTER = 307 then $ = 302 ;
elseif CLUSTER = 308 then $ = 302 ;
elseif CLUSTER = 501 then $ = 506 ;
elseif CLUSTER = 502 then $ = 506 ;
elseif CLUSTER = 503 then $ = 506 ;
elseif CLUSTER = 504 then $ = 506 ;

```

elseif CLUSTER = 505 then \$ = 506 ;
elseif CLUSTER = 506 then \$ = 506 ;
elseif CLUSTER = 507 then \$ = 506 ;
elseif CLUSTER = 508 then \$ = 506 ;
elseif CLUSTER = 509 then \$ = 506 ;
elseif CLUSTER = 510 then \$ = 506 ;
elseif CLUSTER = 511 then \$ = 506 ;
elseif CLUSTER = 512 then \$ = 506 ;
elseif CLUSTER = 513 then \$ = 506 ;
elseif CLUSTER = 514 then \$ = 506 ;
elseif CLUSTER = 515 then \$ = 506 ;
elseif CLUSTER = 516 then \$ = 506 ;
elseif CLUSTER = 517 then \$ = 506 ;
elseif CLUSTER = 518 then \$ = 506 ;
elseif CLUSTER = 519 then \$ = 506 ;
elseif CLUSTER = 520 then \$ = 506 ;
elseif CLUSTER = 521 then \$ = 506 ;
elseif CLUSTER = 522 then \$ = 506 ;
elseif CLUSTER = 523 then \$ = 506 ;
elseif CLUSTER = 524 then \$ = 506 ;
elseif CLUSTER = 525 then \$ = 506 ;
elseif CLUSTER = 526 then \$ = 506 ;
elseif CLUSTER = 527 then \$ = 506 ;
elseif CLUSTER = 528 then \$ = 506 ;
elseif CLUSTER = 701 then \$ = 703 ;
elseif CLUSTER = 702 then \$ = 703 ;
elseif CLUSTER = 703 then \$ = 703 ;
elseif CLUSTER = 704 then \$ = 703 ;
elseif CLUSTER = 705 then \$ = 703 ;
elseif CLUSTER = 706 then \$ = 703 ;
elseif CLUSTER = 901 then \$ = 906 ;
elseif CLUSTER = 902 then \$ = 906 ;
elseif CLUSTER = 903 then \$ = 906 ;


```

elseif CLUSTER = 904      then $ = 906  ;
elseif CLUSTER = 905      then $ = 906  ;
elseif CLUSTER = 906      then $ = 906  ;
elseif CLUSTER = 907      then $ = 906  ;
elseif CLUSTER = 908      then $ = 906  ;
endif;
PROC CONST
preproc

```

```

      If CLUSTER = 201 then $ = 020    ;
elseif CLUSTER= 202      then $ = 020  ;
elseif CLUSTER= 203      then $ = 020  ;
elseif CLUSTER= 204      then $ = 021  ;
elseif CLUSTER= 205      then $ = 021  ;
elseif CLUSTER= 206      then $ = 022  ;
elseif CLUSTER= 207      then $ = 022  ;
elseif CLUSTER= 208      then $ = 022  ;
elseif CLUSTER= 209      then $ = 022  ;
elseif CLUSTER= 210      then $ = 023  ;
elseif CLUSTER= 211      then $ = 023  ;
elseif CLUSTER= 212      then $ = 024  ;
elseif CLUSTER= 213      then $ = 024  ;
elseif CLUSTER= 214      then $ = 024  ;
elseif CLUSTER= 301      then $ = 039  ;
elseif CLUSTER= 302      then $ = 039  ;
elseif CLUSTER= 303      then $ = 040  ;
elseif CLUSTER= 304      then $ = 040  ;
elseif CLUSTER= 305      then $ = 040  ;
elseif CLUSTER= 306      then $ = 041  ;
elseif CLUSTER= 307      then $ = 041  ;
elseif CLUSTER= 308      then $ = 042  ;
elseif CLUSTER= 501      then $ = 079  ;
elseif CLUSTER= 502      then $ = 079  ;

```

elseif CLUSTER= 503	then \$ = 079 ;
elseif CLUSTER= 504	then \$ = 078 ;
elseif CLUSTER= 505	then \$ = 078 ;
elseif CLUSTER= 506	then \$ = 078 ;
elseif CLUSTER= 507	then \$ = 080 ;
elseif CLUSTER= 508	then \$ = 080 ;
elseif CLUSTER= 509	then \$ = 080 ;
elseif CLUSTER= 510	then \$ = 080 ;
elseif CLUSTER= 511	then \$ = 080 ;
elseif CLUSTER= 512	then \$ = 080 ;
elseif CLUSTER= 513	then \$ = 077 ;
elseif CLUSTER= 514	then \$ = 077 ;
elseif CLUSTER= 515	then \$ = 082 ;
elseif CLUSTER= 516	then \$ = 082 ;
elseif CLUSTER= 517	then \$ = 082 ;
elseif CLUSTER= 518	then \$ = 082 ;
elseif CLUSTER= 519	then \$ = 082 ;
elseif CLUSTER= 520	then \$ = 082 ;
elseif CLUSTER= 521	then \$ = 081 ;
elseif CLUSTER= 522	then \$ = 081 ;
elseif CLUSTER= 523	then \$ = 081 ;
elseif CLUSTER= 524	then \$ = 081 ;
elseif CLUSTER= 525	then \$ = 076 ;
elseif CLUSTER= 526	then \$ = 076 ;
elseif CLUSTER= 527	then \$ = 076 ;
elseif CLUSTER= 528	then \$ = 076 ;
elseif CLUSTER= 701	then \$ = 097 ;
elseif CLUSTER= 702	then \$ = 097 ;
elseif CLUSTER= 703	then \$ = 097 ;
elseif CLUSTER= 704	then \$ = 097 ;
elseif CLUSTER= 705	then \$ = 098 ;
elseif CLUSTER= 706	then \$ = 098 ;
elseif CLUSTER= 901	then \$ = 127 ;

```

elseif CLUSTER= 902      then $ = 127  ;
elseif CLUSTER= 903      then $ = 127  ;
elseif CLUSTER= 904      then $ = 127  ;
elseif CLUSTER= 905      then $ = 127  ;
elseif CLUSTER= 906      then $ = 127  ;
elseif CLUSTER= 907      then $ = 127  ;
elseif CLUSTER= 908      then $ = 127  ;
endif;

```

PROC WARD

preproc

```

        if CLUSTER = 201  then $ = 1    ;
elseif CLUSTER= 202      then $ = 3    ;
elseif CLUSTER= 203      then $ = 2    ;
elseif CLUSTER= 204      then $ = 16   ;
elseif CLUSTER= 205      then $ = 14   ;
elseif CLUSTER= 206      then $ = 8    ;
elseif CLUSTER= 207      then $ = 9    ;
elseif CLUSTER= 208      then $ = 12   ;
elseif CLUSTER= 209      then $ = 11   ;
elseif CLUSTER= 210      then $ = 27   ;
elseif CLUSTER= 211      then $ = 23   ;
elseif CLUSTER= 212      then $ = 19   ;
elseif CLUSTER= 213      then $ = 18   ;
elseif CLUSTER= 214      then $ = 21   ;
elseif CLUSTER= 301      then $ = 15   ;
elseif CLUSTER= 302      then $ = 14   ;
elseif CLUSTER= 303      then $ = 21   ;
elseif CLUSTER= 304      then $ = 22   ;
elseif CLUSTER= 305      then $ = 19   ;
elseif CLUSTER= 306      then $ = 12   ;
elseif CLUSTER= 307      then $ = 11   ;
elseif CLUSTER= 308      then $ = 2    ;

```

elseif CLUSTER= 501	then \$ = 2	;
elseif CLUSTER= 502	then \$ = 3	;
elseif CLUSTER= 503	then \$ = 1	;
elseif CLUSTER= 504	then \$ = 8	;
elseif CLUSTER= 505	then \$ = 9	;
elseif CLUSTER= 506	then \$ = 5	;
elseif CLUSTER= 507	then \$ = 11	;
elseif CLUSTER= 508	then \$ = 11	;
elseif CLUSTER= 509	then \$ = 11	;
elseif CLUSTER= 510	then \$ = 10	;
elseif CLUSTER= 511	then \$ = 10	;
elseif CLUSTER= 512	then \$ = 10	;
elseif CLUSTER= 513	then \$ = 14	;
elseif CLUSTER= 514	then \$ = 16	;
elseif CLUSTER= 515	then \$ = 20	;
elseif CLUSTER= 516	then \$ = 23	;
elseif CLUSTER= 517	then \$ = 19	;
elseif CLUSTER= 518	then \$ = 22	;
elseif CLUSTER= 519	then \$ = 22	;
elseif CLUSTER= 520	then \$ = 17	;
elseif CLUSTER= 521	then \$ = 25	;
elseif CLUSTER= 522	then \$ = 26	;
elseif CLUSTER= 523	then \$ = 28	;
elseif CLUSTER= 524	then \$ = 27	;
elseif CLUSTER= 525	then \$ = 29	;
elseif CLUSTER= 526	then \$ = 31	;
elseif CLUSTER= 527	then \$ = 30	;
elseif CLUSTER= 528	then \$ = 30	;
elseif CLUSTER= 701	then \$ = 2	;
elseif CLUSTER= 702	then \$ = 8	;
elseif CLUSTER= 703	then \$ = 3	;
elseif CLUSTER= 704	then \$ = 7	;
elseif CLUSTER= 705	then \$ = 14	;

```

elseif CLUSTER= 706      then $ = 9      ;
elseif CLUSTER= 901      then $ = 13     ;
elseif CLUSTER= 902      then $ = 1      ;
elseif CLUSTER= 903      then $ = 6      ;
elseif CLUSTER= 904      then $ = 8      ;
elseif CLUSTER= 905      then $ = 7      ;
elseif CLUSTER= 906      then $ = 16     ;
elseif CLUSTER= 907      then $ = 12     ;
elseif CLUSTER= 908      then $ = 14     ;
endif;
PROC REGION
preproc

```

```

    if CLUSTER = 201 then $ = 2      ;
elseif CLUSTER = 202      then $ = 2      ;
elseif CLUSTER = 203      then $ = 2      ;
elseif CLUSTER = 204      then $ = 2      ;
elseif CLUSTER = 205      then $ = 2      ;
elseif CLUSTER = 206      then $ = 2      ;
elseif CLUSTER = 207      then $ = 2      ;
elseif CLUSTER = 208      then $ = 2      ;
elseif CLUSTER = 209      then $ = 2      ;
elseif CLUSTER = 210      then $ = 2      ;
elseif CLUSTER = 211      then $ = 2      ;
elseif CLUSTER = 212      then $ = 2      ;
elseif CLUSTER = 213      then $ = 2      ;
elseif CLUSTER = 214      then $ = 2      ;
elseif CLUSTER = 301      then $ = 1      ;
elseif CLUSTER = 302      then $ = 1      ;
elseif CLUSTER = 303      then $ = 1      ;
elseif CLUSTER = 304      then $ = 2      ;
elseif CLUSTER = 305      then $ = 2      ;
elseif CLUSTER = 306      then $ = 1      ;

```

elseif CLUSTER = 307	then \$ = 1	;
elseif CLUSTER = 308	then \$ = 1	;
elseif CLUSTER = 501	then \$ = 2	;
elseif CLUSTER = 502	then \$ = 2	;
elseif CLUSTER = 503	then \$ = 2	;
elseif CLUSTER = 504	then \$ = 2	;
elseif CLUSTER = 505	then \$ = 2	;
elseif CLUSTER = 506	then \$ = 2	;
elseif CLUSTER = 507	then \$ = 2	;
elseif CLUSTER = 508	then \$ = 2	;
elseif CLUSTER = 509	then \$ = 2	;
elseif CLUSTER = 510	then \$ = 2	;
elseif CLUSTER = 511	then \$ = 2	;
elseif CLUSTER = 512	then \$ = 2	;
elseif CLUSTER = 513	then \$ = 2	;
elseif CLUSTER = 514	then \$ = 2	;
elseif CLUSTER = 515	then \$ = 2	;
elseif CLUSTER = 516	then \$ = 2	;
elseif CLUSTER = 517	then \$ = 2	;
elseif CLUSTER = 518	then \$ = 2	;
elseif CLUSTER = 519	then \$ = 2	;
elseif CLUSTER = 520	then \$ = 2	;
elseif CLUSTER = 521	then \$ = 2	;
elseif CLUSTER = 522	then \$ = 2	;
elseif CLUSTER = 523	then \$ = 2	;
elseif CLUSTER = 524	then \$ = 2	;
elseif CLUSTER = 525	then \$ = 2	;
elseif CLUSTER = 526	then \$ = 2	;
elseif CLUSTER = 527	then \$ = 2	;
elseif CLUSTER = 528	then \$ = 2	;
elseif CLUSTER = 701	then \$ = 1	;
elseif CLUSTER = 702	then \$ = 2	;
elseif CLUSTER = 703	then \$ = 1	;

```

elseif CLUSTER = 704      then $ = 2      ;
elseif CLUSTER = 705      then $ = 1      ;
elseif CLUSTER = 706      then $ = 1      ;
elseif CLUSTER = 901      then $ = 2      ;
elseif CLUSTER = 902      then $ = 2      ;
elseif CLUSTER = 903      then $ = 2      ;
elseif CLUSTER = 904      then $ = 2      ;
elseif CLUSTER = 905      then $ = 2      ;
elseif CLUSTER = 906      then $ = 2      ;
elseif CLUSTER = 907      then $ = 2      ;
elseif CLUSTER = 908      then $ = 2      ;
endif;
PROC CSA
preproc

```

```

        if CLUSTER = 201  then $ = 6      ;
elseif CLUSTER= 202      then $ = 17     ;
elseif CLUSTER= 203      then $ = 9      ;
elseif CLUSTER= 204      then $ = 7      ;
elseif CLUSTER= 205      then $ = 13     ;
elseif CLUSTER= 206      then $ = 6      ;
elseif CLUSTER= 207      then $ = 10     ;
elseif CLUSTER= 208      then $ = 3      ;
elseif CLUSTER= 209      then $ = 10     ;
elseif CLUSTER= 210      then $ = 3      ;
elseif CLUSTER= 211      then $ = 2      ;
elseif CLUSTER= 212      then $ = 6      ;
elseif CLUSTER= 213      then $ = 9      ;
elseif CLUSTER= 214      then $ = 7      ;
elseif CLUSTER= 301      then $ = 18     ;
elseif CLUSTER= 302      then $ = 23     ;
elseif CLUSTER= 303      then $ = 5      ;
elseif CLUSTER= 304      then $ = 3      ;

```

elseif CLUSTER= 305	then \$ = 6	;
elseif CLUSTER= 306	then \$ = 3	;
elseif CLUSTER= 307	then \$ = 6	;
elseif CLUSTER= 308	then \$ = 1	;
elseif CLUSTER= 501	then \$ = 8	;
elseif CLUSTER= 502	then \$ = 4	;
elseif CLUSTER= 503	then \$ = 10	;
elseif CLUSTER= 504	then \$ = 6	;
elseif CLUSTER= 505	then \$ = 1	;
elseif CLUSTER= 506	then \$ = 11	;
elseif CLUSTER= 507	then \$ = 8	;
elseif CLUSTER= 508	then \$ = 26	;
elseif CLUSTER= 509	then \$ = 44	;
elseif CLUSTER= 510	then \$ = 11	;
elseif CLUSTER= 511	then \$ = 32	;
elseif CLUSTER= 512	then \$ = 51	;
elseif CLUSTER= 513	then \$ = 5	;
elseif CLUSTER= 514	then \$ = 29	;
elseif CLUSTER= 515	then \$ = 5	;
elseif CLUSTER= 516	then \$ = 7	;
elseif CLUSTER= 517	then \$ = 3	;
elseif CLUSTER= 518	then \$ = 3	;
elseif CLUSTER= 519	then \$ = 21	;
elseif CLUSTER= 520	then \$ = 15	;
elseif CLUSTER= 521	then \$ = 12	;
elseif CLUSTER= 522	then \$ = 17	;
elseif CLUSTER= 523	then \$ = 15	;
elseif CLUSTER= 524	then \$ = 6	;
elseif CLUSTER= 525	then \$ = 8	;
elseif CLUSTER= 526	then \$ = 6	;
elseif CLUSTER= 527	then \$ = 12	;
elseif CLUSTER= 528	then \$ = 32	;
elseif CLUSTER= 701	then \$ = 13	;


```

elseif CLUSTER= 702      then $ = 17    ;
elseif CLUSTER= 703      then $ = 8      ;
elseif CLUSTER= 704      then $ = 7      ;
elseif CLUSTER= 705      then $ = 3      ;
elseif CLUSTER= 706      then $ = 4      ;
elseif CLUSTER= 901      then $ = 2      ;
elseif CLUSTER= 902      then $ = 3      ;
elseif CLUSTER= 903      then $ = 1      ;
elseif CLUSTER= 904      then $ = 3      ;
elseif CLUSTER= 905      then $ = 2      ;
elseif CLUSTER= 906      then $ = 1      ;
elseif CLUSTER= 907      then $ = 3      ;
elseif CLUSTER= 908      then $ = 4      ;
endif;

PROC SEA

preproc
if          CLUSTER          in          201,
203,207,210,211,212,305,306,307,503,506,510,512,516,522,526,704,706,905 then $ = 1;
elseif CLUSTER in 204,206,208,303,304,308,505,515,519,520,527,528,906,907 then $ = 2;
elseif          CLUSTER          in
202,205,213,301,302,501,502,507,508,511,518,523,524,701,702,901,902,903,904,908 then $
= 3;
elseif CLUSTER in 214,504,525,703,705 then $ = 4;
elseif CLUSTER in 209,513,517 then $ = 5;
elseif CLUSTER in 514,521 then $ = 6;
elseif CLUSTER = 509 then $ = 8;
endif;

PROC Q1

preproc
$ = RESPONDENT

PROC Q7
if $=2 then

```

```
endlevel;  
endif;  
PROC Q8SPECIFY  
preproc  
ask if Q8 = 8;  
PROC Q10SPECIFY  
preproc  
ask if Q10 = 8;  
PROC Q12  
if $ = 2 then skip to Q14;  
endif;  
PROC Q15  
preproc
```

```
Ask if Q14=1;  
PROC Q16SPECIFY  
preproc  
ask if Q16I = 1;
```

```
PROC Q17A  
preproc  
Ask if Q16A=1;  
PROC Q17B  
preproc  
Ask if Q16B=1;  
PROC Q17C  
preproc  
Ask if Q16C=1;  
PROC Q17D  
preproc  
Ask if Q16D=1;  
PROC Q17E  
preproc
```

Ask if Q16E=1;
PROC Q17F
preproc
Ask if Q16F=1;
PROC Q17G
preproc
Ask if Q16G=1;
PROC Q17H
preproc
Ask if Q16H=1;
PROC Q17I
preproc
Ask if Q16I=1;
PROC Q17I_SPECIFY
preproc
ask if Q17I = 1;
PROC Q18A
preproc
ask if Q17A =1;
PROC Q18B
preproc
ask if Q17B =1;
PROC Q18C
preproc
ask if Q17C =1;
PROC Q18D
preproc
ask if Q17D =1;
PROC Q18E
preproc
ask if Q17E =1;
PROC Q18F
preproc

```
ask if Q17F =1;
PROC Q18G
preproc
ask if Q17G =1;
PROC Q18H
preproc
ask if Q17H =1;
PROC Q18I
preproc
ask if Q17I =1;
PROC Q18I_SPECIFY
preproc
ask if Q18I in 1:4;
```

```
PROC Q19A_S
preproc
```

```
Ask if Q19A = 6;
PROC Q19B_S
preproc
```

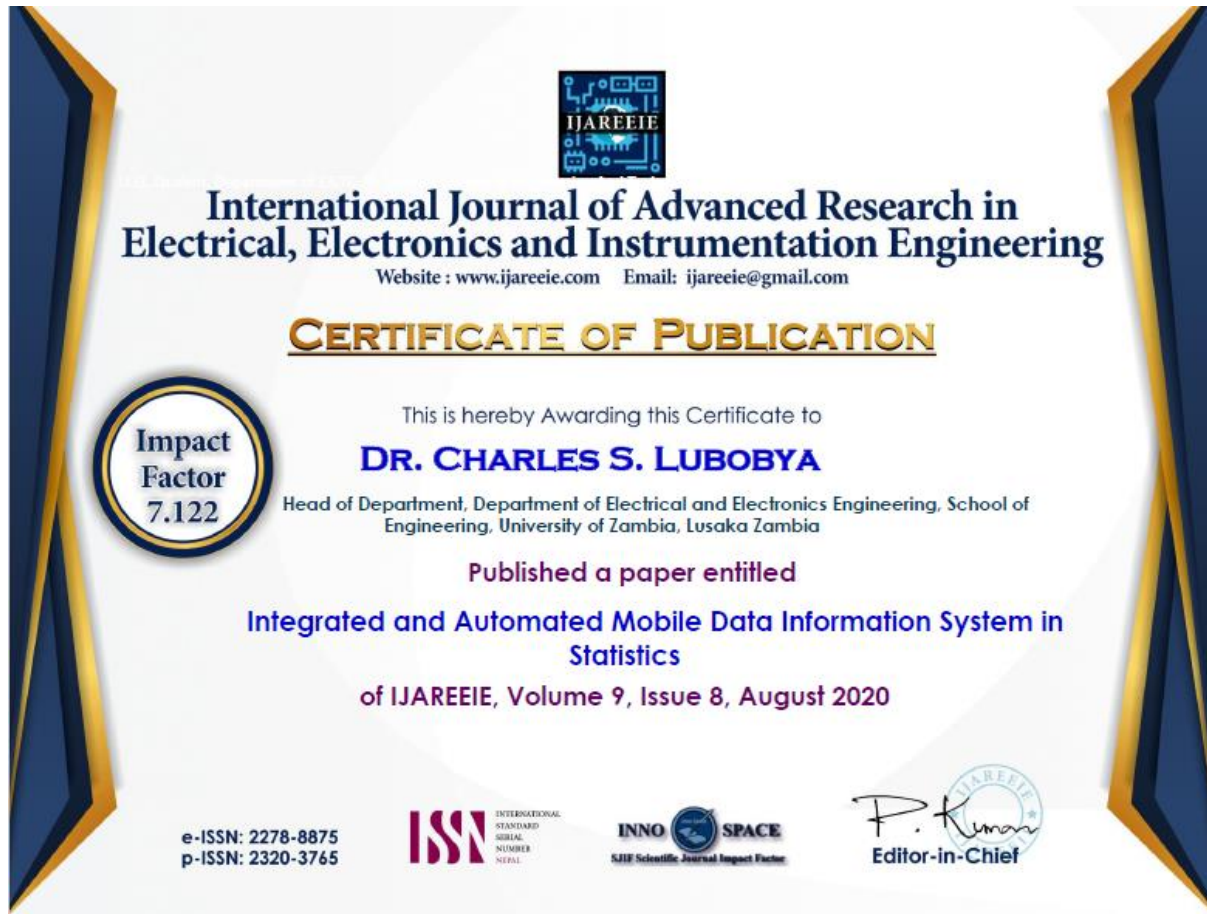
```
Ask if Q19B = 6;
PROC Q19C_S
preproc
```

```
Ask if Q19C = 6;
PROC Q20
if $ = 2 then
    skip to Q29;
endif;
PROC Q22
if $ = 2 then
    skip to Q29;
```

```
endif;  
PROC Q24SPECIFY  
preproc
```

```
Ask if Q24 = 9;  
PROC Q27  
if $ = 1 then  
skip to Q29;  
endif;  
PROC Q29  
if $ = 2 then  
skip to Q31;  
endif;
```

APPENDIX C: International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Certification





International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

Website : www.ijareeie.com Email: ijareeie@gmail.com

CERTIFICATE OF PUBLICATION



This is hereby Awarding this Certificate to

ABRAHAM DANIEL NGOYI

Department of Electrical and Electronics Engineering, School of Engineering,
University of Zambia, Lusaka, Zambia

Published a paper entitled

**Integrated and Automated Mobile Data Information System in
Statistics**

of IJAREEIE, Volume 9, Issue 8, August 2020

e-ISSN: 2278-8875
p-ISSN: 2320-3765




Editor-in-Chief



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

Website : www.ijareeie.com Email: ijareeie@gmail.com

CERTIFICATE OF PUBLICATION



This is hereby Awarding this Certificate to

CHENELA NKHOWANI

P.G. Student, Department of Electrical and Electronics Engineering, School of
Engineering, University of Zambia, Lusaka, Zambia

Published a paper entitled

**Integrated and Automated Mobile Data Information System in
Statistics**

of IJAREEIE, Volume 9, Issue 8, August 2020

e-ISSN: 2278-8875
p-ISSN: 2320-3765

