



Integrated and Automated Mobile Data Information System in Statistics

Chenela Nkhowani¹, Dr. Charles S. Lubobya², Abraham Daniel Ngoyi³

¹P.G. Student, Department of Electrical and Electronics Engineering, School of Engineering, University of Zambia, Lusaka, Zambia¹

Head of Department, Department of Electrical and Electronics Engineering, School of Engineering, University of Zambia, Lusaka Zambia²

Department of Electrical and Electronics Engineering, School of Engineering, University of Zambia, Lusaka, Zambia³

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 Research is to create an Electronic Mobile Data Application to be used in collecting data electronically via Cspiro 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, Cspiro Application, Real Time Monitoring, Monitor Performance.

I. INTRODUCTION

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 onoka2017). H Carol (2013) in his journal says 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 mobile market in the country has maintained an upward trend with an increase in the number of mobile subscriptions of 7.7 million from 701 million representing a 47.08 percent penetration rate (Zambia's mobile phone sector in near saturation –Zicta). Technology has allowed many organizations to monitor our performance, make decisions about our programs, and report on our progress cheaper, faster, and better.



The main purpose of this study is to develop an integrated and automated mobile data collection application system that would improve on the quality of data in good time. This study will be guided by the following objectives; To investigate the challenges faced by Zambia Statistics Agency in the process of data collection using PAPI; To establish the benefits of using an integrated and automated mobile data collection system; and To establish how an integrated and automated mobile data information system could be used to improve the data collection process.

II. RELATED WORK

Most statistical offices and agencies in developed countries like USA and Netherlands already migrated from traditional paper and pencil interviews to computer assisted personal interview [18]. The first national household survey that used CAPI for all of its data collection was the Netherlands Labour Force Survey in 1987 [17]. 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 ([19]). Central statistical offices 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 [14] with those in North Africa (Algeria, Morocco and Tunisia) and Southern Africa (South Africa, Mauritius and Botswana) having relatively higher scores. According to [16] 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, [16].

Data collection using PDAs was more accurate and complete than paper-based data collection [6]. 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.

[15] Points out that the benefits in 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.

III. 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 C# 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.

Current System



Figure 8: Example of a Paper Questionnaire



Figure 9: Flowchart of a Paper Questionnaire

Integrated System

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, 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.

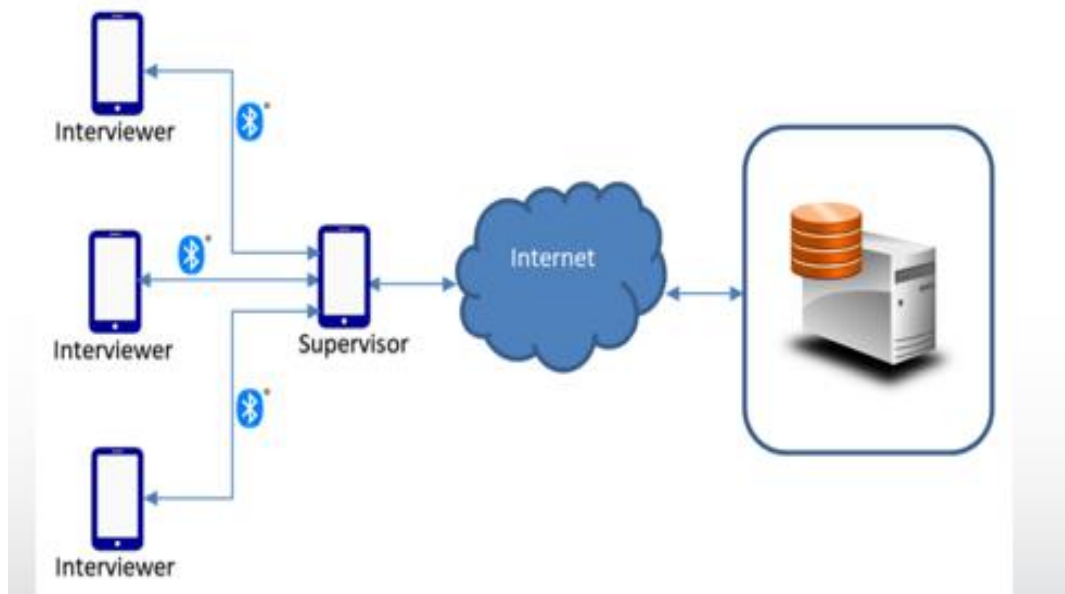


Figure 10: Integrated System

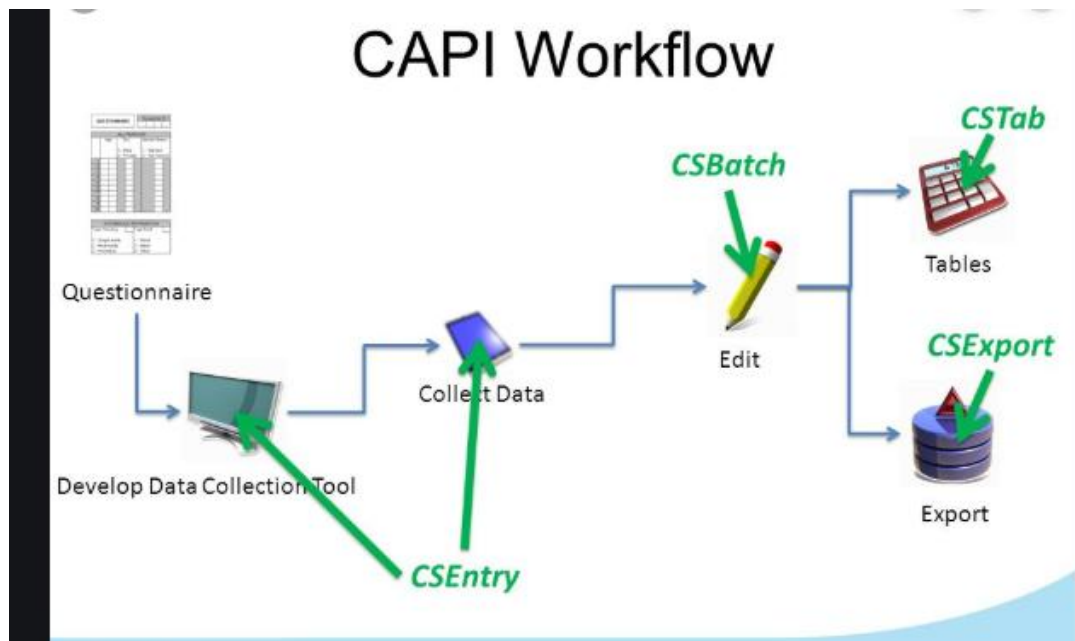


Figure 11: Workflow of an Integrated System

IV.METHODOLOGY

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 [20]. This design was chosen because there is little studies done on the subject matter in Zambia. The purposive sampling technique used was 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.

The study population included members of staff who were working under research and data processing branch in the IT Department of the Zambia statistics Agency and statisticians at the head office and the provincial office.

Questionnaires were generated in line with the required research information. The statistical values of the obtained results were analyzed, processed and verified using SPSS vs23. Thereafter, The second method of analysis involved designing the modules that made up the system that is the automated mobile data collection system This was based on the second 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:

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

IV. RESULTS AND DISCUSSIONS

The findings of this study was organized according to the objectives: To investigate the challenges faced by Zambia Statistics Agency in the process of data collection using PAPI; To establish the benefits of using an integrated and automated mobile data collection system; and To establish how an integrated and automated mobile data information system could be used to improve the data collection process.



Challenges of PAPI

The study was interested in establishing the challenges that come with 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.

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 Fig. 4.2 below

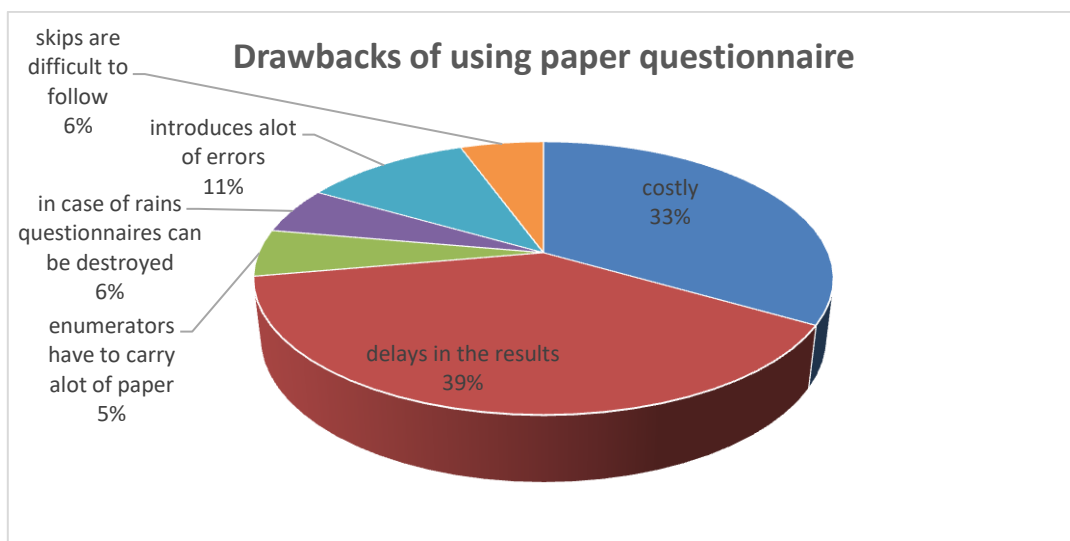


Fig 4.2.1 shows drawbacks of using paper questionnaire

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 the figure below.

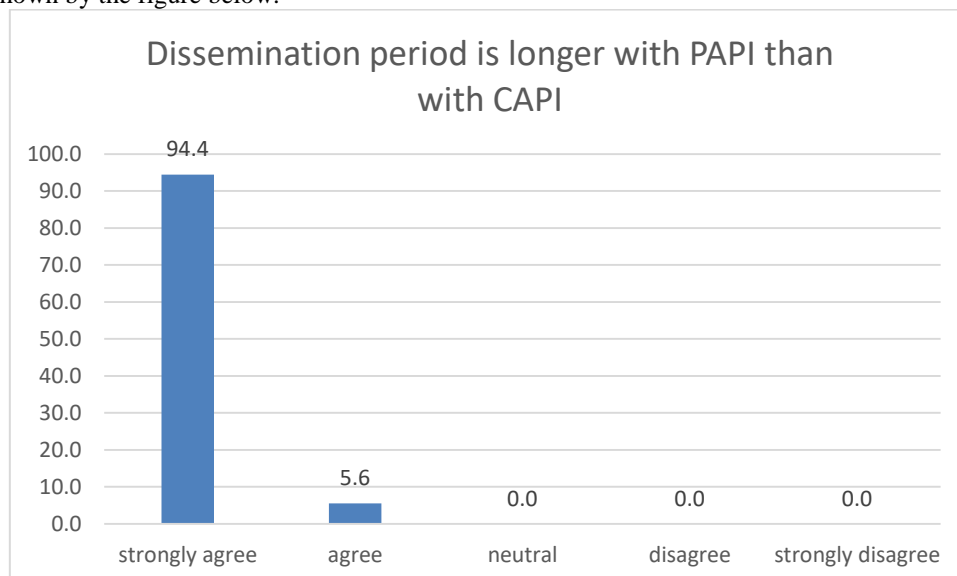


Fig.4.2.2. Dissemination period is longer with CAPI



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

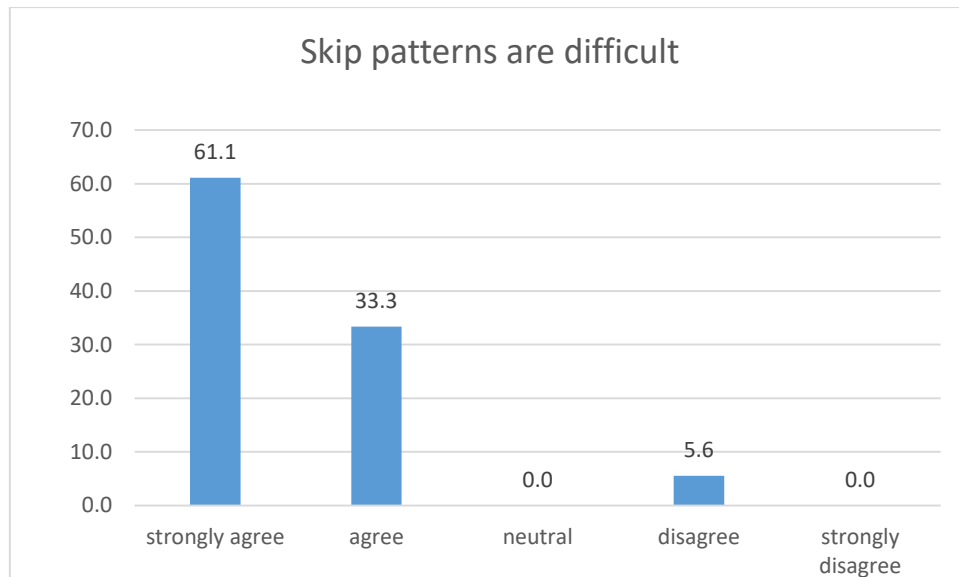


Figure 4.2.3. Skip patterns are difficult with PAPI

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:

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.

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.

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.

V. CONCLUSION AND RECOMMENDATIONS

Recommendation

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

- 1) Zambia Statistics Agency adopts the proposed model in order to improve on the efficiency of data collection.



- 2) 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.
- 3) Capacity is built in the programmers so as 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.
- 4) Data security is seriously looked into because respondents have a right to confidentiality. If respondent's data would be leaked to non-authorised personnel that would be a breach of that confidence.
- 5) Further research needs to be done with a broader sample as studies with smaller samples like this one may not be generalised.

Conclusion

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.

REFERENCES

1. Chen H, Hailey D, Wang N, Yu P.(2014) A Review of Data Quality Assessment Methods for Public Health Information Systems.
2. WHO (World Health Organization). Framework and Standards for Country Health Information Systems. World Heal. 2008;2nd Edition(January):72
3. Weiskopf NG, Weng C.(2013) Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research.
4. Ahmed, M. Y., (2001), "Welcome Address at the Opening Ceremony of the National Workshop on Pension Reforms in Nigeria", Abuja.
5. Bet Caeyers (2010) A Comparison of CAPI and PAPI through a Randomized Field Experiment 1.
6. David G. Dillona, b. F. P. S. R. C. P. M. S. S., 2014. Open-source electronic data capture system offered increased accuracy.
7. Singleton, K. W. et al., 2011. Wireless Data Collection of Self Administered Surveys Using Tablet Computers.
8. Shao D. A(2012) Proposal of a Mobile Health Data Collection and Reporting System for the developing countries
9. Jasmeet, Singh P(2014) Review of smartphone operating systems like Symbian OS and Windows mobile.
10. Basch E, Goldfarb S. (2009)Electronic patient-reported outcomes for collecting sensitive information from patients.
11. Garrity C, Emam K.(2006) Who's using PDAs? Estimates of PDA use by health care providers.
12. Welker J.(2007) Implementation of electronic data capture systems
13. Brandt C, Argraves S, Money R, Ananth G, Trocky N, Nadkarni P.(2006) Informatics tools to improve clinical research study implementation.
14. ITU. ICT facts and figures. 2013. Available from: <http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013-e.pdf>
15. De leeuw, Edith(2008) The effect of computer-assisted interviewing on data quality
16. Holmner, Marlene A.(2015)ICT in Africa: Enabling a Better Life for All
17. van Bastelaer, Alois, Frans ~essemakers, and Dirk Sikkel (1988), "Data Collection with Hand-Held Computers: Contributions to Questionnaire Design," Journal of official Statistics, Vol. 4, No. 2, pp. 141-154.
18. L Danielsson, PA Maarstad - Orebo: Statistics Sweden, 1982
19. Rothschild and Wilson 1988 nationwide food consumption survey
20. Ayankoya, K.A.2013. A framework for implementation of social media marketing strategies.