Using Tablets for the 2018 Algerian Census: Mobile Census Application Quality Assessment

Tarik BOUREZGUE
Office for National Statistics, Algiers, Algeria
bourezgue@ons.dz, tarikbourezgue@gmail.com

Abstract
Nowadays, to achieve progress in this era of data revolution, it is necessary to make use of the rapid emergence of new technologies. This is, once more, a challenge for official statistics. Algeria is planning to carry out its sixth General Census of Population and Housing (RGPH) in 2018. ONS intends to use, for the first time, tablets (equipped with mobile chips) as a collection tool instead of the paper questionnaire. This paper mainly concentrates on ONS' vision of census data collection through the use of mobile technology, namely 3G and 4G LTE technologies. So we need a system that segregates data as and when data is entered. Another challenge in this method is the accuracy of data entered: the mobile applications environment has a multitude of particularities. This requires an adaptable approach for performance analysis in terms of total quality management for software projects. In this context, quality is analyzed from the viewpoints of developers, of users and, respectively, of one who aims to recover an investment. Thus, our paper will address the issues related to the implementation of this mobile solution so as to present the census data meticulously with minimum hardship through a mobile application quality estimation model that requires the identification of features (e.g. interaction time, volume of provided information, error management, data security, transaction security, …) and which properties should be highlighted for ensuring usability in a practical environment. The paper proposes to look at a set of indicators normalized on the [0; 1] interval as to measure the application quality level.
Keywords: Census Mobile Application, Total Quality Management, Aggregate Indicator.

1. Introduction

The rapid emergence of new technologies brings about some concern. What the world needs now is not the aptitude in the production of new technologies but the capacity of understanding the impact of technology on society and individuals as well as the capability to implement ICT as to positively impact human development. This is, once more, a challenge for statistics. We need to produce statistics with a certain depth of

understanding of the environment and the difference between technological development and human development.

The main feature of the first 2011 Economic Census (RE) was the use, for the first time, of the optical scanning technology for the exploitation of questionnaires. Moreover and through The Office for National Statistics (ONS), Algeria is planning to carry out its sixth General Census of Population and Housing (RGPH) in 2018. As part of the framework of the United Nations recommendations for the 2020 round of censuses of population and housing, advocating greater use of Information and Communication Technology (ICT) in statistical operations, Algeria's plans to use for the first time, tablets as a collection tool instead of the paper questionnaire, namely tablets equipped with mobile chips.

The usual census data collection system involves an enumerator, who collects the census data manually on a paper questionnaire. So in today's world of data communication, a mobile tool for a portable system becomes important to access and update a database.

Innovative technology helps us to improve the existing system. So, we thought the finest solution for this handicap would be using a tablet. We are in the process of designing and developing an ICT solution using mobile technology, mainly 3G, as a communication tool. Our idea is to implement this by providing a tablet containing the census application to every authorized enumerator through which they can collect census data and update the collected data to the census database. A second and important module of the solution is the supervision and monitoring platform. We are motivated by the fact that "*Technology"* allows statistical processes to meet important society requirements and expectations, namely:

- ✓ Timeliness of data;
- ✓ Accuracy;
- ✓ Relevance;

so, as to achieve "*Quality"* improvements as it is recommended by all international guidelines on the subject.

2. Problem description

2.1 Census preparation

In the RGPH census preparation, generally these steps are to be followed:

- Identifying the houses i.e., dwelling places and places that are usually used for living. This is the first step.
- Then the municipal delegate prepares the construction list. Once the list is prepared, he prepares sketches of blocks of houses that give the primary information about the type of houses and facilities that are being used by the population in that area.

2.2 Census execution

With this information in hand census enumerators go to the census houses and collect data. The collected data would be used for analysis of the population in various aspects like finding population ratio, employment reports, and facilities available to people etc.

Nowadays, in this fast world of technology, it is very important to complete a work in an efficient manner. In the traditional method, collecting data takes a long time, because enumerators have to manually fill in the census form then again sort out the data, etc. So we need a system that segregates data as and when data is entered. Another test for this method is the accuracy of data entered. In a census, age of people should be entered accurately because these details would be used in employment analysis, etc. Efficiency is another issue; efficient data collection is needed because this data will be used for further analysis.

3. Proposed solution

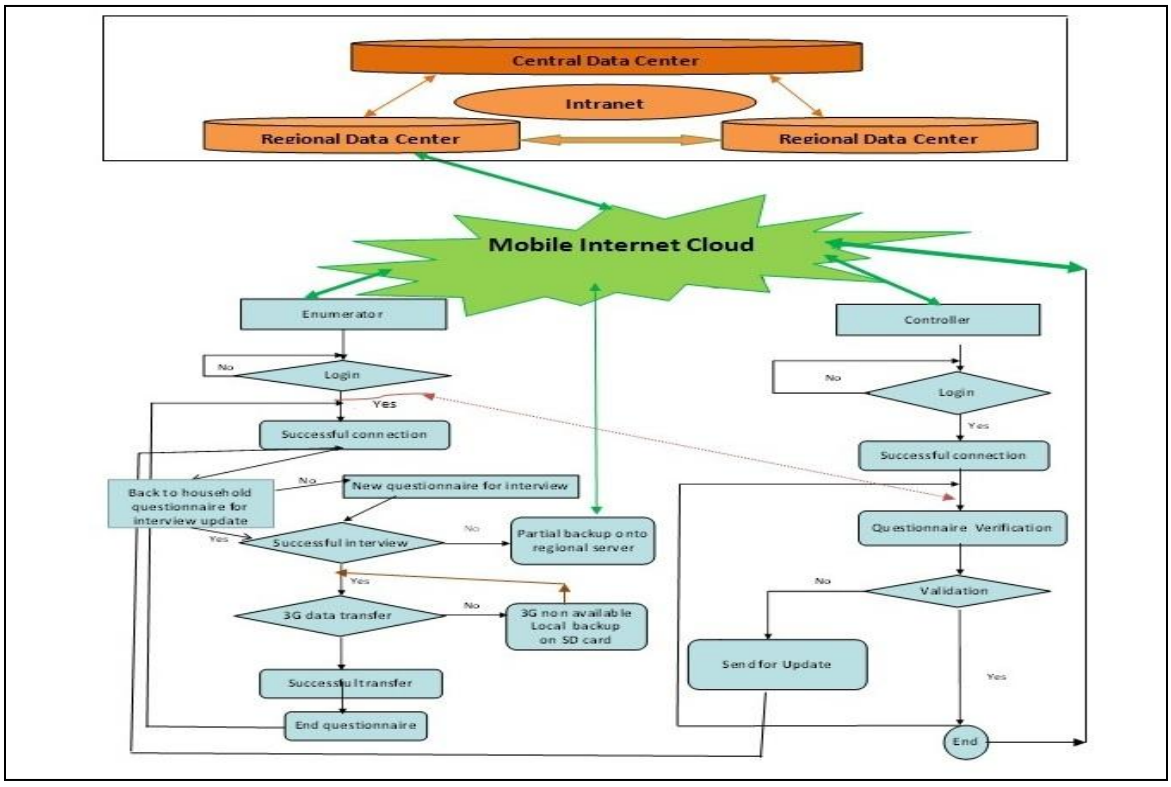
For RGPH 2018, 50,000 agents will be recruited and trained as enumerators and 13,000 agents will be recruited and trained as controllers, in addition to the municipal delegates used in the preparation.

3.1 Census agents

Each enumerator will be equipped with a tablet for data entry and a paper district notebook that is provided by the municipal delegate (The first five district notebook columns are already filled by the municipal delegate, the remaining columns will be filled in throughout the execution of the census). Each controller is also equipped with a Notebook tablet (for an overall vision of the questionnaire) to verify the quality of work of the five enumerators, under his supervision, and have a status of the collection in the field and daily statistics on households surveyed and enumerator performance.

3.2 Authentication of enumerators and controllers

For the purposes of authentication of enumerators and controllers, login lists (username and password) will be defined following the use of mobile chips' PINs affected by the mobile operator: the login could be defined as the concatenation of the wilaya (territorial administration), municipality and district codes (given at the conclusion of cartographic updates), which is an exact identification of the enumerator. The flow of information is as follows:



3.3 Technical requirements

In order to insure that the process of data collection goes as planned, a few technical elements need to be in place such as:

- A relational database management system such as SQL Server is needed to administer and manage the RGPH database to guarantee quality, continuity and confidentiality of information.
- A global database at the central site to store and retrieve real-time data entirety.
- An implementation of database levels at regional collection sites.
- Software installed at regional levels for supervision, control and monitoring of the implementation on the field.
- Entry applications for entering data collected on tablets as well as the transcript of paper questionnaires[1].
- Establishment of specialized interconnection links between the central site (headquarters) and various regional sites.

3.4 Information flow estimates

We base our preliminary estimates on what follows:

i. 50 000 enumerators spread over the four regional annexes.

ii. 150 questionnaires per enumerator during the period of execution (average of 10 daily).

As per our benchmark, the size of a questionnaire is about few Kilobytes (we assume 50 Ko per questionnaire). As bandwidth is the rate of a connection expressed in bits per second, we can estimate that we will face a rate of transfer of 100Ko/s and the necessary bandwidth should be around 2Mo.

[1] As per international recommendations, 10% of the questionnaires will be on paper as to take charge of the specific cases on the field.

4. Quality assessment

In software engineering literature (Aksit, 2004 and Pressman, 2009), software quality and software quality management are treated separately and the important role that they play in successfully completing software applications projects is highlighted.

Mobile applications are a special category of software applications as mobile devices and mobile applications are seeing continuous development. Hence, their quality management processes are influenced by their particular specifications. It is considered that the implementation of specific total quality management (TQM) methods and techniques will help developers meet the objectives related to the continuous quality improvement of the processes and results (Hoyle, 2007).

The mode of interaction (MI) represents the use of appropriate controls to build the application user interface so that the user will have a natural interaction with it. The mode of interaction is influenced by the number of displayed controls at a time and their complexity. The measurement is a score received from the user after he/she interacts with the application. The mark, MImark, is a value in the range $[A_{inf}; A_{sup}]$, where:

$A_{inf} < A_{sup}$; $A_{inf}, A_{sup} \in \mathbf{IN}$; for this case we have $A_{inf} = 0$, $A_{sup} = 10$. The mode of interaction indicator, MII, is determined as follows:

$$MII = \frac{MImark}{10}. \qquad (1)$$

Error management (EM) is the ability of mobile applications to function even in cases where operational errors appear (ISO/IEC 25010:2011). Measuring is done by observing the number of instructions successfully fulfilled compared to the total number of instructions given by the user. Thus, the error management indicator (EMI) is calculated using the following relation:

$$EMI = \frac{NT}{Ntotal}, \qquad (2)$$

where NT is the number of successful instructions and N_{total} is the number of total instructions.

Data security (DS) ensures data privacy in the portable device. When used by an unauthorized person, the accessing of confidential data on the device may harm the operation, and thus it is necessary to secure access to this information if the mobile device is lost or stolen, (Popa, 2012). The data security indicator (DSI) is given by the number of cases where the application does not use the mobile device resources correctly, thus allowing uncontrolled events to occur per unit time (UT):

$$DSI = 1 - \frac{TNEU}{UT}, \qquad (3)$$

where TNEU is the total number of events that are unpredictable and uncontrolled.
Hence, taking into account file transfer issues and using (1), (2) and (3), we consider the following 2018 Census specific total quality management indicator (TQMI):

$$TQMI = TQI*MII*EMI*DSI, \qquad (4)$$

where TQI = (successful transfers per UT)/(total transfers per UT), transfers expressed in Ko/s, represents the transfer quality indicator.

Since TQMI, clearly, belongs to the interval [0; 1], we will be able to assess the quality of our mobile application by the closeness of its TQMI to 1.

5. Conclusion

As we are well aware of the mains risks and challenges in implementing such a solution, we always keep in mind the fact not underestimate how much work it takes to properly implement a census, even with the use of the latest technology (IT or ICT). Thus, our paper comforts us in knowing that this solution will make it possible for ONS to enumerate and present the census data meticulously with minimum hardship. A key ingredient also, is our ongoing consultation with the different users and producers of statistical information as

well as our national technology partners, namely mobile operators, internet providers and IT manufacturers. In order to test and validate these indicators, the following activities will be conducted:

- development of the census mobile application;
- consultation with specialists in order to determine the different values;
- application of these indicators to the developed mobile application.

Because use and evolution of such technology is growing rapidly, it may be helpful to carry out, post census, further detailed research on improved data quality through the use of technology.

Acknowledgements

This work was done in collaboration and with valuable input from the technical departments at ONS and unconditional support from our Director General Mounir Khaled Berrah; with a special mention to Zineb Hentabli, Amel Lakehal, Soraya Saadi, Samia Salmi and Hamid Zidouni.

References

Aksit M. (2004), The 7 C's For Creating Living Software: A Research Perspective For Quality-Oriented Software Engineering, Turk J Electr Eng Co, 12, pp. 61-96.

Arthur Dan, Overview of Data Collection using Mobile Devices, LSHIM seminar, 2013.

Bellavista P, Corradi A, Fanelli M and Foschini L. (2012), A Survey of Context Data Distribution for Mobile Ubiquitous Systems. ACM Comput Surv, 4, pp. 1-49.

Dunlop MD, Durga N, Motaparti S, Dona P and Medapuram V. (2012), QWERTH: An Optimized Semi-Ambiguous Keyboard Design. In: 14th International Conference on Human-Computer Interaction with Mobile Devices and Services, San Francisco, CA, USA. New York, NY, USA, ACM, pp. 23-28.

Ferguson N, Schneier B and Kohno T. (2010), Cryptography Engineering: Design Principles and Practical Applications. Indianapolis, IN, USA, Wiley Publishing.

Gao K, Khoshgoftaar TM, Wang H and Seliya N. (2011), Choosing Software Metrics for Defect Prediction: An Investigation on Feature Selection Techniques. Software-Pract Exp, 41, pp. 579-606.

Gomaa H, Hashimoto K, Kim M, Malek S and Menasc DA. (2010), Software Adaptation Patterns for Service-Oriented Architectures. In: Symposium on Applied Computing; 22-26, Sierre, Switzerland. New York, NY, USA, ACM, pp. 462-469.

Hoyle D. (2007), Quality Management Essentials. Burlington, MA, USA: Butterworth-Heinemann, 2007.

International Organization for Standardization. ISO/IEC 25010:2011, Systems and Software Engineering - Systems and Software Quality Requirements and Evaluation (SQuaRE) - System and Software Quality Models. Geneva, Switzerland.

Junglas IA. and Watson RT. (2008), Location Based Services Evaluating User Perceptions of Location-Tracking and Location-Awareness Services. Commun ACM, 3, pp. 65-69.

Plaza I, Martin L, Martin S and Medrano C. (2011), Mobile Applications in an Aging Society: Status and Trends. J Systems Soft, 84, pp. 1977-1988.

Popa M. (2012), Requirements of a Better Secure Program Coding, Inform Econ, 4, pp. 93-104.

Pressman S. (2009), Software Engineering: A Practitioner's Approach. 7th ed. New York, NY, USA: McGraw-Hill.

Redd JB. and Schmidt D. (2011), Using Mobile Devices and Gaming as a Means of Building Vocabulary. Intl J Interact Mob Tech, 4, pp. 30-38.

Salehie M. and Tahvildari L. (2009), Self-Adaptive Software: Landscape and Research Challenges. ACM Trans Auton Adapt Syst, 2, pp. 1-42.

Sommerville I. (2011), Software Engineering. 9th ed. Boston, MA, USA: Addison-Wesley.

Vijayaraj A. and Dinesh Kumar P. (2010), Design and Implementation of Census Data Collection System using PDA, International Journal of Computer Applications, 9, pp. 28-32.

Yacob Zewoldi (2011), Seminar on New Technologies in Population and Housing Censuses: Country experiences, United Nations Statistics Division, Side event of the 42nd session of the, Statistical Commission, New York.