Summary Report: Workshop on Survey Methodology
Prepared by:
Mariama Deen-Swarray and Broc Rademan
E: mdeen-swarray@researchictafrica.net & brademan@researchictafrica.net
Research ICT Africa

The VI NIC.br Annual Workshop on Survey Methodology was held in São Paulo, Brazil from the 11th to 14th April, 2016. Two RIA researchers attended this workshop hosted by NIC.br and organised by CETIC.br. The Brazilian Network Information Center (NIC.br) implements the decisions and projects designed by the Brazillian Internet Steering Committee (CGI.br). The Regional Center for Studies on the Development of the Information Society (Cetic.br), produces indicators and statistics on the use of information and communication technologies in Brazil.

The aim of the Workshop was to discuss and analyse concepts and problems regarding survey methodologies in the ICT sector. This year's Workshop began with a focus on Big data which spanned a 2-day period. There was no attempt to exhaust the theme in that amount of time but no less than 11 speakers addressed the participants on the nature, applications, challenges and opportunities of Big Data.

After Prof. Dr. Demi Getschko opened the first day's proceedings, Dr Karin Breitman introduced the concepts, challenges, and opportunities of Big Data. The overview looked at the growth of digital data generated by devices and sensors through a business lens. Having reliable and extensive infrastructure, logical data organisation, algorithms to produce results, as well as the applications to present them can generate value for businesses and society at large.

The participants saw how Big Data can help achieve development and social good when Dr Emmanuel Letouzé described how it should be thought of as a phenomenon with the potential to challenge existing power structures and decision-making processes. This view extended Big Data beyond being computed datasets and sought to satisfy the targets of the Sustainable Development Goals.

Big data and digital media may also be used to produce data for policy making as shown by Dr Robert Faris. In an overview of the adoption of digital media in a digitally connected society, digital media can affect public policy making by mediating, intervening and raising awareness of social and political processes. However, one must bear in mind limiting factors such as data not being easily usable and many public spheres not yet networked.

The changes being made to global value chains by data-driven innovation were highlighted by Dr. Vincenzo Spiezia. The presentation, which included findings from the OECD's "Data-Driven Innovation" report, identified the IT service and infrastructure providers, data producers and analysts, as well as entrepreneurs as the actors involved in cooperation and competition that will restructure the global value chain.

Closing the day off, Dr. Dhanaraj Thakur, shared with the participants the findings of a mobile data services survey in developing countries. The findings lament the dearth of evidence available on the effects of mobile services for policy-making and focussed on the mobile data services designed to provide access at lower costs, such as zero-rated products.

Day two of the program focused on the application and use of Big Data. In the first session, Prof. Dr. Denise Britz do N.Silva, highlighted the existing challenges in the production, analysis and dissemination of official statistics and what this will mean with the addition of big data. One of the main challenges in big data lie in the complexity in linking various data sources from which it can originate. Whereas traditional survey or census is considered 'designed' data, big data is referred to as 'organic' – data produced to record processes. The pros and cons of each form of data were presented. However, according to Prof. Grooves, incorporating these two types of data is the future: http://directorsblog.blogs.census.gov/2011/05/31/designed-data-and-organic-data/. There is however the need to follow a well-established and standard framework to ensure the quality of Big Data. Big data has been used to produce official statistics in some countries: http://unstats.un.org/unsd/trade/events/2014/beijing/default.asp.

The use of big data in Health care was presented in the second session by Dr. Marco Antoni Gutierrez. The efficient analysis and interpretation of big data is envisaged to address some of the challenges in medical and health care systems – better understanding of diseases and development of better diagnostics among others. There is, however, the lack of adequate capacity to deal with such large amount of data. He emphasised that the landscape of healthcare as well as processes in health are rapidly changing. There is a digital health revolution and a paradigm shift in medicine to which big data can contribute.

The participants then had an opportunity to hear from Prof. Dr. Bernadette Farais Loscio and Caroline Burle on the possibilities and challenges of making data available on the web. They highlighted that data on the web is a great repository of big data, emphasing the need for people to be able to share this data in a manner that will allow others to derive, add value and utilise the data to suit them. The presentation focused on the need to have this data in a standard format and applying best practices to publishing data on the web. The presenters demostrated how the Web is evolving from a 'Web of linked documents' into that of 'linked data' and the intersection between big data and open data.

In the session following, Prof. Dr. Adriano C. Machado Pereira showed how Big Data is actually part of Data Science; how they correlate and the transition to a data-empowered society in particular on the Web. Natalia Mazotte in the discussion on Data Journalism, highlighted the potential of using Big data to increase accountability of governments and improve the lives of citizens. This session focused on how data-driven journalism was used to promote public policy debates and shed light on socially controversial issues in Brazil.

Prof. Dr. Bruno Cautres delivered a short course on Quantitative Methods in Policymaking. Participants were presented with the main methodological and practical issues for policy making and evaluation using quantitative techniques. The major issues of causal inferences and the basic framework for evaluating the 'treatment' effects of policy shifts were discussed. The most important concepts and methodological challenges about multivariate analysis for public policy analysis were highlighted. He emphasised the need to not only estimate a causal

claim but also to estimate a plausible counterfactual to this claim when conducting policy evaluations.

The third day, the participants were introduced to the second part of the workshop: Web Surveys – Questionnaire design and implementation. Dr. Pamela Campanelli comprehensively covered the difficulties of performing web-based research using survey designs. Beyond the regular challenges of questionnaires, such as question wording, online surveys also suffer from design incompatibilities as alternative operating systems - web browsers and screen sizes can affect the layout of the questionnaire.

However, there are some benefits to issuing surveys online:
- Skip patterns can be automated;
- Piping can link additional information easily for respondents to find out more by simply clicking on a link;
- Editing can take place online;
- Randomising the answers is possible.

On the final day, Prof. Dr. Campanelli discussed the issues surrounding web survey implementation, highlighting the unique features of web surveys, the process of sending out invitations and reminders and whether to offer incentives to respondents. Another key issue that was discussed was the testing and designing of the right software for web surveys to be implemented on mobile devices. It was suggested that it was good practice to present web questions in simple conventional format as used on paper questionnaires. Another important issue raised was the positioning of buttons on the page to facilitate navigating through pages: http://surveypractice.org/index.php/SurveyPractice/article/view/93/pdf
Whilst Web survey is considered to be cost effective, there are challenges over the control of question order and quality of the answers received. Response rate can also be relatively low. There is therefore a need to have strategies in place to maximise participation on the Web and improve response rates.

In conclusion, while big data has the potential to contribute to advances in statistics and to different fields of studies, it is still in the experimental stage for most countries. Having a set of standardised frameworks is essential to work effectively with big data. Data obtained in this form should not be seen as a stand alone source of information, but as one that can support traditional sources of data collection to increase the repository of available data.