Big Data Statistics
# The Kenyan Perspective

by

Paul Kiage

Agenda

1. Background
2. Proposed Mechanism for BD coordination
3. Why Big Data?
4. Implementation Plan / Road Map
5. BD Indicators by ITU
6. Proposed BD Indicators by NSO- Kenya
7. Assessment of collaboration, Challenges
8. Way Forward

Background

- BD focuses on innovative ways to utilize big data as a new data source FOR ICT INDICATORS

- The objective of the project is to demonstrate how big data can be used in ICT measurement – to produce new and existing ICT indicators to enhance data availability and measure the information society

- Big Data sources are telecom operators, Internet service providers, or any other data providers who possess valuable residual data from the aspect of ICTs.

- Kenya is one of the 6 pilot countries where the Big Data is being piloted.

Mechanism for BD Coordination

Mechanism of Coordination cont.…
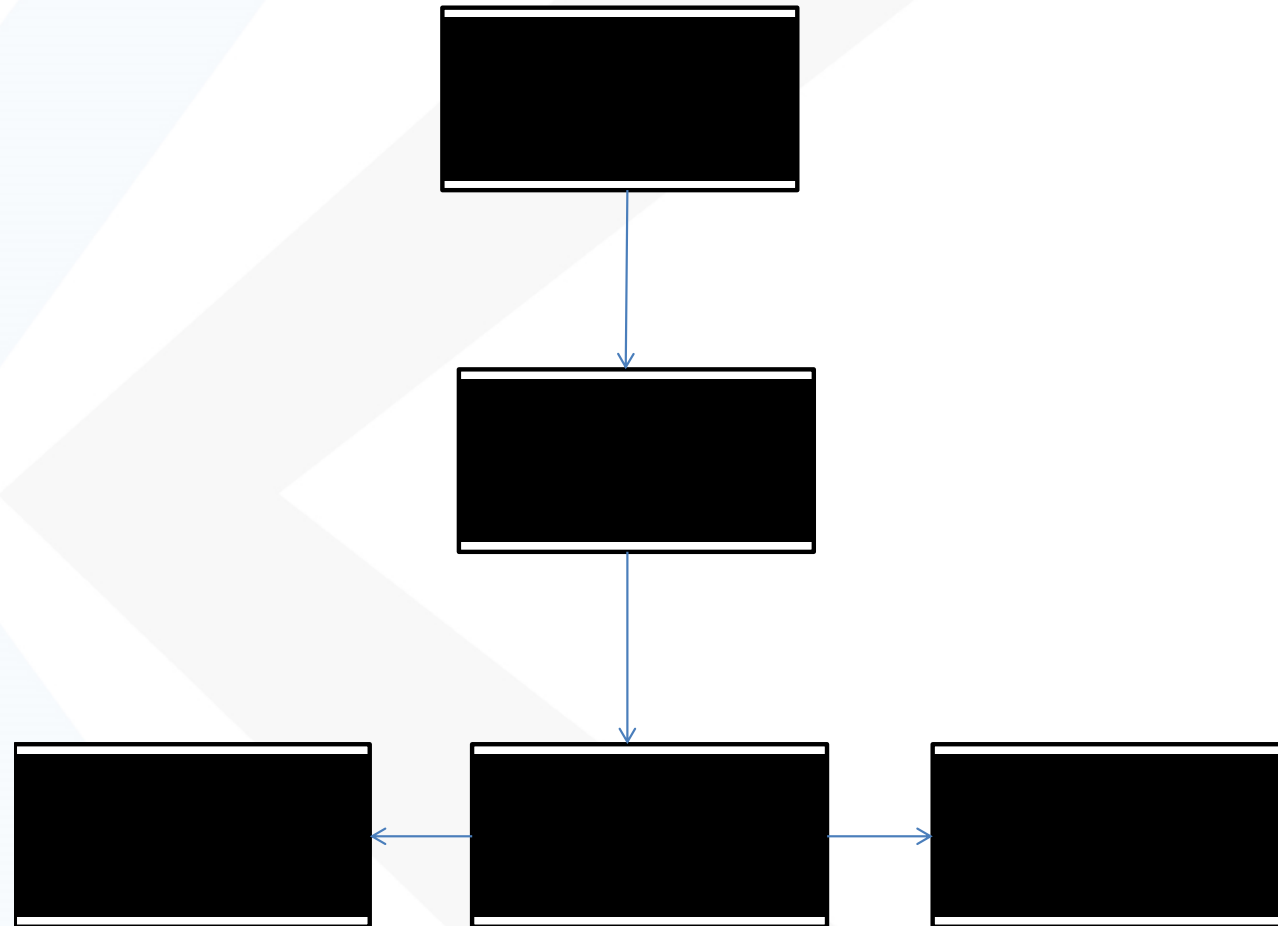
- **Telecommunication regulator - CA**
  - Country coordinator of the project, focal point, administrative, cooperation and regulatory issues.
- **National Statistical Office - KNBS**
  - Provide necessary existing reference data (geographic administrative GIS data, demographic statistics, etc.)
  - Quality assessment of the methodology used and resulting indicators
- **Telecom operators & Internet service providers – data providers**
  - Provide data and resources if necessary for the project

Importance of Big Data for Kenya

➢ Leveraging on such data to complement official statistics will enable the government and development agencies to better serve Kenyans and measure information society.

➢ Mobile network big data may be used for ex-post evaluation, real-time measurement, and future predictions and planning in areas such as health, agriculture, education etc.

➢ Big data can provide new insights into measurement of information society in Kenya.

➢ Big data will measure the use of different ICT services and applications, intensity and frequency of use and the geographic location of ICT users: urban vs rural.

4. Implementation Plan /Road map

	Activity	Progress/Status
1.	Engagement of the stakeholders, data providers, partners	Meeting with various stakeholders in September 2016 - done
2.	Getting commitment from stakeholders.	Consultation with data providers in progress
3.	Developing of NDAs	November 2016
4.	Signing of NDAs, visit by data scientist, Submission of data and processing	December 2016
5.	Data interpretation and Analysis	January 2017
6.	Reporting and comparative analysis of the results with other countries	Feb/March 2017

Big Data Indicators

- **BD 01**. Percentage of the Land Area Covered by Mobile-Cellular Network, by Technology
- **BD 02**. Percentage of the Population Covered by a Mobile-Cellular Network, by Technology
- **BD 03.** Usage of Mobile-Cellular Networks for non-IP Related Activities, by Technology
- **BD 04.** Usage of Mobile-Cellular Networks for Internet Access, by Technology
- **BD 05.** Number of Subscriptions with Access to Technology
- **BD 06.** Active Mobile Voice and Broadband Subscriptions, by Contract Type
- **BD 07.** Average Number of Active Mobile Subscriptions per Day, by Contract Type
- **BD 08.** Active Mobile Devices
- **BD 09.** IMEI Conversion Rate
- **BD 10.** Fixed Domestic Broadband Traffic, by Speed, Contract Type
- **BD 11.** Mobile Domestic Broadband Traffic, by Speed, Contract Type, Technology
- **BD 12.** Mobile International Broadband Traffic, by Contract Type
- **BD 13.** Inbound Roaming Subscriptions per Foreign Tourist
- **BD 14.** Fixed Broadband Subscriptions, by Technology
- **BD 15.** Fixed Broadband Subscriptions, by Speed
- **BD 16+** Proposed New Indicators from Pilot Countries that had not been foreseen by ITU

Proposed new indicators by the NSO

- **BD 16**. Number of active mobile money subscribers (Urban and Rural)
- **BD17**. Total amount of money sent from subscriber to other subscribers(Urban and Rural)
- **BD18** Number of transactions sent from subscriber to another subscriber(Urban and Rural)
- **BD19** Number of mobile money transactions sent from Kenya to other countries(Urban and Rural)
- **BD20** Total amount of mobile money transactions sent from Kenya to other countries(Urban and Rural)
- **BD21** Number of mobile money transactions received from outside country to Kenya(Urban and Rural)
- **BD22** Total amount of mobile money transactions received from outside country to Kenya(Urban and Rural)
- **BD23** Number of mobile money transactions received/ sent by Kenyans while outside Kenya(Urban and Rural)
- **BD24** Total amount of mobile money transactions received/ sent by Kenyans while outside Kenya(Urban and Rural)
- **BD25** Number of mobile money transactions received/ sent by foreigner's while in Kenya(Urban and Rural)
- **BD26** Total amount of mobile money transactions received/ sent by foreigner's while in Kenya(Urban and Rural)
- **BD27** Number of active mobile banking subscribers(Urban and Rural)
- **BD28** Total amount of money sent from mobile money platform to bank(Urban and Rural)
- **BD29** Number of transactions (sent) from mobile money platform to bank(Urban and Rural)
- **BD30** Total amount of money transacted from bank to mobile money platform(Urban and Rural)
- **BD31** Number of transactions from bank to mobile money platform(Urban and Rural)
- **BD32** Number of active Pay bill numbers(Urban and Rural)
- **BD33** Total amount of money received through Pay bill number (Urban and Rural)
- **BD34** Number of transactions done via Till number(Urban and Rural)

Challenges

- Currently, there is no Act of Parliament in place to regulate Data Protection although there is a **Data Protection Bill, 2013** which is yet to be passed by Parliament.

- Data privacy frameworks for most operators may require signing of NDAs with all parties handling the data including ITU

- Some of the data for the project are not currently collected or stored by the operators in the required format (Urban and Rural).

- Operators have limited human resource capacity and are unwilling to commit data scientists to the project

- Limited time to handle and process large data sets

Assessment of Collaboration

- The collaboration is easy to achieve, why?
- The Authority already have a framework of partnership with KNBS on the sharing of data.
- The licensees have an obligation to submit to the Authority data on a quarterly basis
- ITU is already working with CA & KNBS in the development, collation and analysis of ICT indicators statistics & other areas of capacity building.
- KNBS is already capturing some of the indicators through KIHBS and periodic surveys and have skilled data specialists.
- KNBS has an elaborate administrative GIS data that is needed for the project (up to urban/rural level)

Way Forward

- Engagement with stakeholders to agree on BD Methodology and data processing.

- Visitation by data scientist to oversee BD calculation methodology

- Finalization on country indicators

- Signing the relevant NDAs (CA/providers? ITU/providers?)

- Sharing of KNBS GIS administrative data to facilitate geographical and urban/rural breakdowns / reporting.

- Data submission by the operators