Seventh SG13 Regional Workshop on "Standardization of Future Networks towards Building a Better Connected Africa"

Big Data in Ghana

Isaac Kobina Kwarko, PMP.

Abuja, Nigeria, 3-4 February 2020

Presentation Outline

- Introduction to Big Data

- National Telecommunication Policy 2005

- Regulatory Framework in Ghana

- Evolution to Big Data

- Big Data Opportunities for Ghana

- Challenges

- Big Data Strategic Plan

Introduction to Big Data

- Big Data definition:
  - It is the dataset which traditional database systems cannot store, analyse or manipulate.

  - [ITU-T Y.3600] - "A paradigm for enabling the collection, storage, management, analysis and visualization, potentially under real-time constraints, of extensive datasets with heterogeneous characteristics;

Introduction to Big Data Cont'd

Volume — SCALE OF DATA

40 ZETTABYTES
[43 TRILLION GIGABYTES]
of data will be created by 2020, an increase of 300 times from 2005

It's estimated that
2.5 QUINTILLION BYTES
[2.3 TRILLION GIGABYTES]
of data are created each day

6 BILLION PEOPLE
have cell phones

WORLD POPULATION: 7 BILLION

Most companies in the U.S. have at least
100 TERABYTES
[100,000 GIGABYTES]
of data stored

The FOUR V's of Big Data

From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day. But what exactly is big data, and how can these massive amounts of data be used?

As a leader in the sector, IBM data scientists break big data into four dimensions: **Volume, Velocity, Variety and Veracity**

Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.

By 2015
4.4 MILLION IT JOBS
will be created globally to support big data, with 1.9 million in the United States

Variety — DIFFERENT FORMS OF DATA

As of 2011, the global size of data in healthcare was estimated to be
150 EXABYTES
[161 BILLION GIGABYTES]

30 BILLION PIECES OF CONTENT
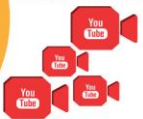are shared on Facebook every month

By 2014, it's anticipated there will be
420 MILLION WEARABLE, WIRELESS HEALTH MONITORS

4 BILLION+ HOURS OF VIDEO
are watched on YouTube each month

400 MILLION TWEETS
are sent per day by about 200 million monthly active users

Velocity — ANALYSIS OF STREAMING DATA

The New York Stock Exchange captures
1 TB OF TRADE INFORMATION
during each trading session

Modern cars have close to
100 SENSORS
that monitor items such as fuel level and tire pressure

By 2016, it is projected there will be
18.9 BILLION NETWORK CONNECTIONS
– almost 2.5 connections per person on earth

Veracity — UNCERTAINTY OF DATA

1 IN 3 BUSINESS LEADERS
don't trust the information they use to make decisions

Poor data quality costs the US economy around
$3.1 TRILLION A YEAR

27% OF RESPONDENTS
in one survey were unsure of how much of their data was inaccurate

Sources: McKinsey Global Institute, Twitter, Cisco, Gartner, EMC, SAS, IBM, MEPTEC, QAS

IBM

National Telecommunications Policy 2005

- Liberalisation started in 1994 with the authorization of four network operators;

- Every citizen and resident of the Republic of Ghana to have affordable access to information and communications services – universal access for all communities and population groups by the year 2010;

- Moderate in ambition – 25% of population to be directly connected, with at least 10% in rural areas by 2010;

National Telecommunications Policy 2005 Cont'd

- Connection of all second cycle schools, universities and polytechnics; medical facilities, and government offices;

- Reduce the cost of communications services, particularly for low income citizens;

- Deepen the competition in the communications market for more investment opportunities through a transparent and technology neutral regulatory regime;

Regulatory Framework

- Legislations –
    - National Communications Authority Act, 2008 (Act 769)
    - Electronic Communications Act, 2008 (Act 775)
    - Electronic Transactions Act, 2008 (Act 772)
    - National Information Technology Agency, 2008 (Act 771)
    - Data Protection Act, 2012 (Act 843)

Regulatory Framework Cont'd

- Regulations –
  - Electronic Communications Regulations, 2011 (L.I 1991);
  - Electronic Communications (Rules of Procedure of the EC Tribunal) Regulations, 2016 (L.I. 2235);
  - Subscriber Identity Module Regulations, 2011 (L.I. 2006) ;
  - Mobile Number Portability Regulations, 2011 (L.I. 1994)

Regulatory Framework Cont'd

- Guidelines –

  – Mast & Towers Guidelines

  – Unsolicited Electronic Communications Guidelines

  – Special Numbering Resources Guidelines

  – Type Approval Guidelines

  – Guidelines for Television White Spaces (TVWS) Data Services

Outcome

- Many of the objectives were achieved within three years;
- Currently:
  - Four (4) major network operators have created a total subscriber base of 24,209,027 out of 30,069,682 making a penetration of 80.5%;

  - Thirty Eight (38) Internet Service Providers (ISPs) and Four (4) Broadband Wireless Access (BWA) operators providing broadband data services;

Outcome Cont'd

- Five (5) International fibre optic landing cables with a total capacity of 1069.97 Gbit (21% capacity utilization);

- Ten inland fibre optic network 14,127.52 KM

- Hundreds of Value Added Services – creating ICT applications;

Evolution To Big Data

- Increased connectivity ⟹ Increased in demand for higher capacity:

  – Six 4G operators – registered almost 2million subscribers with average data speed above 4Mbit/s;

  – Increase "smart phones" (it is estimated that there are more than 20million smart phones on mobile networks);

  – Fibre-to-Homes spreading across the country;

- Individuals and enterprises creating huge datasets;

- Called for a Big Data strategy.

Big Data Awareness Scale

- Gauge the awareness level:
  – Questionnaire administered in June 2019;
- Over 20 organisations, we obtained five responses:
  – There is a plan to extend the survey to cover larger population in 2020.

Big Data Awareness Scale

- All respondents agreed that they need to use big data in their organization:
  - At least one out of the 5 (20%) says it is critical;
  - Three out of the 5 (60%) says it is important;
  - Not a single one said it is least important.
  - Three out of the 5 are already making use of big data technologies and one is planning on using big data technologies;
- Only one (20%) has no plans to use Big Data.

Big Data Opportunities for Ghana
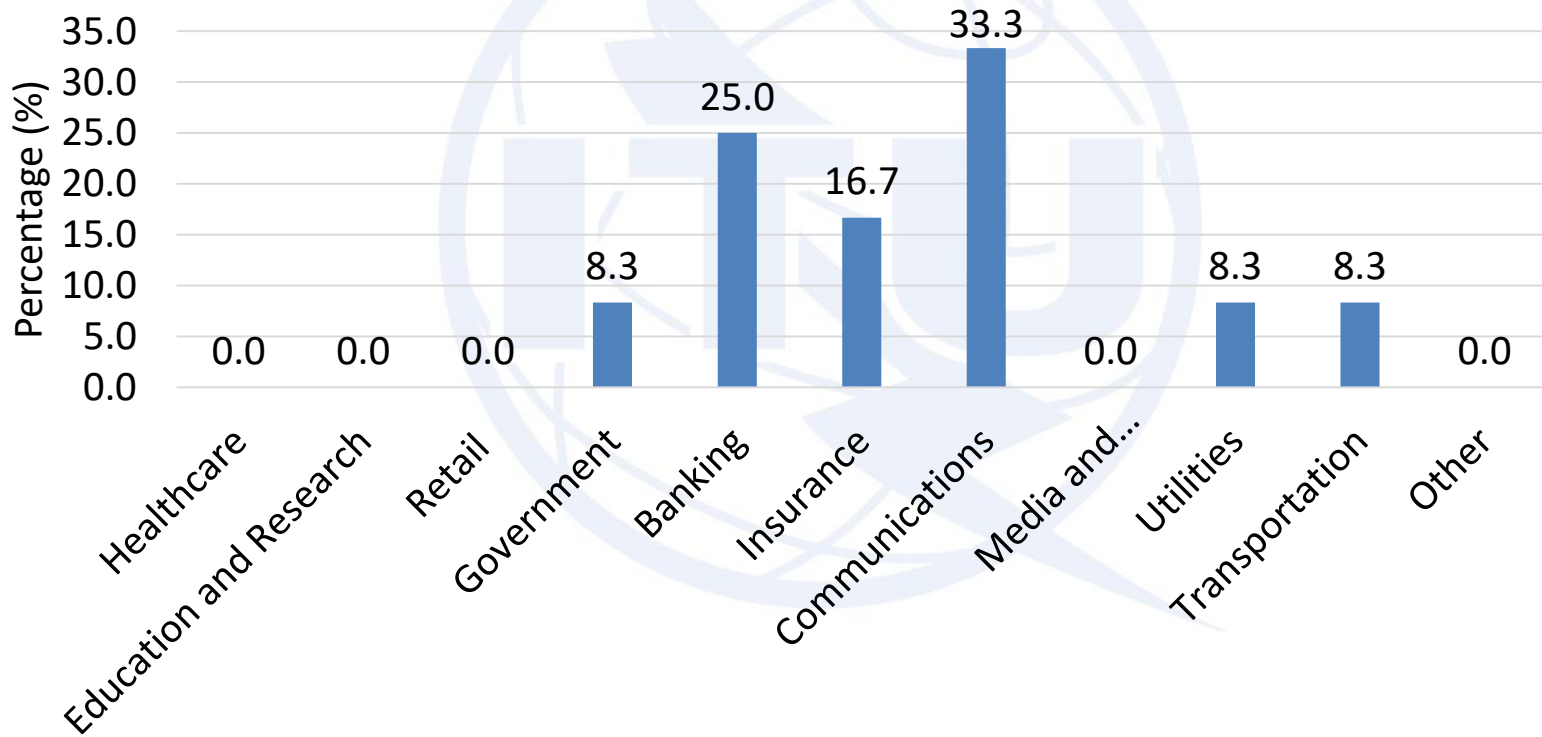
Why Use Big Data
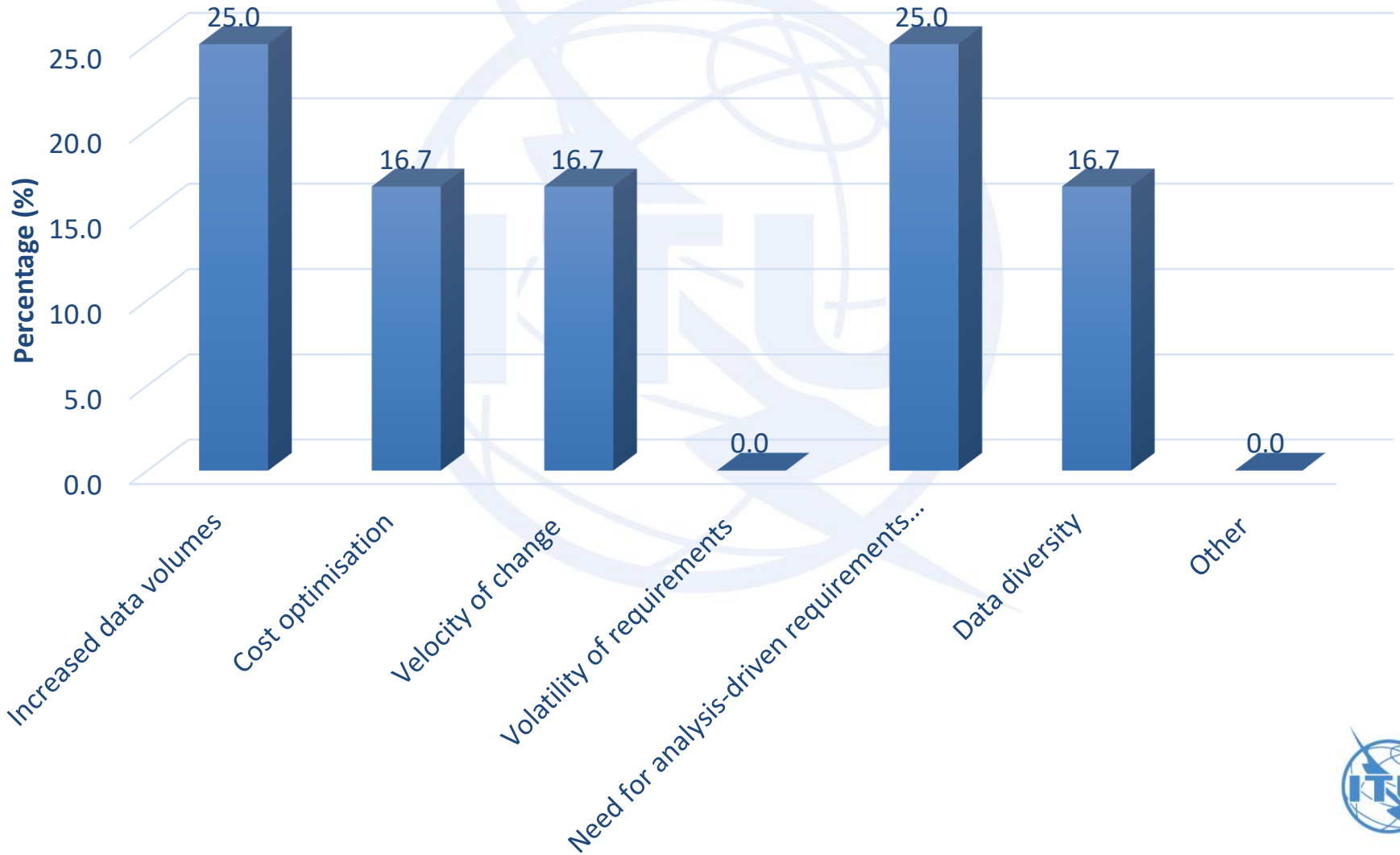
Opportunities

- Agriculture;
  - Esoko Ghana Limited – Provides digital services for farmers and agribusinesses. Specialises agricultural data management analytics to empower rural populations; Insyt Application – http://www.esoko.com

- Government services
  - Ghana Statistical Services (GSS): ata collection and analyses of huge datasets for and on behalf of the GoG;

- Banking and Insurance – Bigdata Ghana Limited – http://www.bigdataghana.com

Opportunities Cont'd

- ICT – Efficient use of spectrum utilization by MNOs:
  - E.g. Huawei and Nokia Ghana both employs data analytics in their services to network operators;

    Thousands of base stations collecting data and being sent to analytics for faster and efficient decision makings such as resource provisioning;

Opportunities Cont'd

- Intelligent Transport System (ITS) – Many ITS deployed in Ghana are based on Big Data technology . E.g. Traffic management system on the Accra to Kumasi road

- Log Management – Managing IT logs, sensor information, and connected devices, IoT, M2M.

- From all sectors of the economy – once data can be collected, stored and analysed, we can create value out it.

How Much Data Created Per Day

- Sources of data – National Communications Authority, Ghana:
  - Automated Spectrum Management System with a Spectrum Monitoring System (relational database) – Structured;
  - Network Monitoring System – Semi-structured;
  - Broadcasting Management System – Semi-structured;
  - Cyber Security (CERT) – Structured and Unstructured;
  - Customer Relationship Management System, CRM – structured;

How Much Data Produced Per Day

- Quality of Service monitoring – generates more than 6 gigabyte of log data for one Region (16 Regions)– semi-structured;

- System logs from IT systems (17 MS Windows Servers, 10 Linux Servers and more than 25 Access Points) sensors for the office building, etc. E.g. One UTM system produces more than 3gigabyte of data every day;

- Common Platform Monitoring – semi-structured.

Challenges

- Regulatory deficiency – the current regulatory framework is limited;

- The Big Data infrastructure: In-house / shared hosting or cloud:

  – Many companies using Big Data are not in the IT infrastructure business;

  – Companies still not too sure about cloud services:

  – Internet cost;

  – Security of data;

Big Data Strategic Plan

- Ability to provide the right data at the right time;

- Building the Big Data Infrastructure:
    - The plan is to consolidate all data systems
    - Plans to start small, move to NoSQL first;
    - Local Cloud Hosting/Shared infrastructure – Make use of GoG datacenter or private hosting;
    - Amazon Web Services.

Big Data Strategic Plan

- Building capacity for data scientists;

- Privacy – strengthen the privacy laws of the country;

- Begin the process to set up a regulatory framework for Big Data – Make use of ITU-T and other SDO Recommendations;

- Produce guidelines and/or regulations;

- Creating the Big Data Ecosystem – authorizing Data Providers (data suppliers and data brokers) and Big Data Service Providers, etc;

Thank You!!!

References

- NCA, Ghana website: http://www.nca.org.gh;
- ITU-T Recommendation Y.3500 – Cloud Computing;
- ITU-T Recommendation Y.3600 – Cloud Computing based Requirements and Capabilities;
- Series Y Supplement 40 – Big Data standardization roadmap;