International
Telecommunication
Union

ICT Regional Forum on ICT Measurement
Dubai, United Arab Emirates
13-15 December 2016

Session 10:
Big Data for IT measurement

José CERVERA, Consultant to ITU
DevStat

Contents

- Experiences
- Technical issues: integration of Big Data and oficial statistics
- Institutional issues
- The Human Factor

Big Data in the international Statistical System

- European Statistical System:
  - ➤ Scheveningen Memorandum Sept 2013
  - ➤ Several conferences since 2014
  - ➤ Big Data Road Map (June 2014)
  - ➤ ESSnet Pilot projects (https://ec.europa.eu/eurostat/cros/content/programme-ess-big-data-workshop-2016_en)

- UN ECE "Big Data in Official Statistics" project
  - ➤ "SandBox"

- UN Statistical Commission 45th session:
  - ➤ Global Working Group (GWG) on Big Data for Official Statistics

Application of <u>Big Data</u> for other statistics

- Sources: mobile phone data, web pages, smart meters, IoT

  - demographic statistics (e.g. daily population, migration)
  - labour statistics (job portals)
  - price statistics (price scrapping)
  - innovation statistics
  - trade and balance of payment statistics
  - transport statistics (mobile, sensors)
  - energy statistics (smart meters)

Applications in ICT statistics

- Infrastructure coverage
  - Internet traffic data
- Households and individuals:
  - Uses of Internet (web tracking)
  - Use of mobile phone (mobile data)
- Businesses:
  - Use of Internet by businesses (website scrapping)
  - e-commerce (price scrapping)
  - IT jobs (Job portals scrapping)

- Source: Big Data Project Inventory
- http://unstats.un.org/bigdata/inventory/

Integration of Big Data and official statistics sources

- Representativeness of Big Data sources

- Single source or combination of sources?
  - Record linkage
  - Key socio-economic variables
  - Classifications and nomenclatures
  - Geolocation

- Other Big Data quality issues (completeness, coherence)
  - Need for a framework for assessing the quality of Big Data sources for Official Statistics

Institutional issues

- Private-public partnership
  - Telecom + NSOs
  - Banking and credit institutions + NSOs
  - E-commerce sites + NSOs

- Sustainability of access

- "Data custodian" role of NSOs?

- Cost issues

The Human Factor

- Skills
  - "Can a Statistician be a data scientist?"
    - From *design-based* to *model-based* to *algorithm-based* inference
    - Rapid evolution of data manipulation technologies (distributed storage and computation)

- Salaries in NSOs vs business sector
  - Can Official Statistics attract (and retain) the talent necessary for Big Data use?

Thank you

José CERVERA (jcervera@devstat.com)