Session 5:
innovative methods
for statistical data production

Measuring the adoption of disruptive digital technologies by government, industry, and citizens

Mr. Ronald Jansen,
Chief, Data Innovation and Capacity Branch,
United Nations Statistics Division,
New York

Overview

1. Example: Mobile Phone Data
2. UNCEBD and UN Global Platform
3. Privacy Preserving Techniques
4. Training & Skills
5. Regional Hub in Brazil
6. Official Statistics and Data Science

Use of Mobile Phone Data for Official Statistics

Active Positioning

- Obtaining real-time location of the mobile device
- Usually requires consent from the phone owner

GPS precision (accuracy: ~10 m) and 3G MPS (~200 m using trilateration)

Passive Positioning

- Retrieving stored records of the activities of mobile devices from the mobile operator networks or app developers' databases
- Digital geographical footprint left by the mobile device users

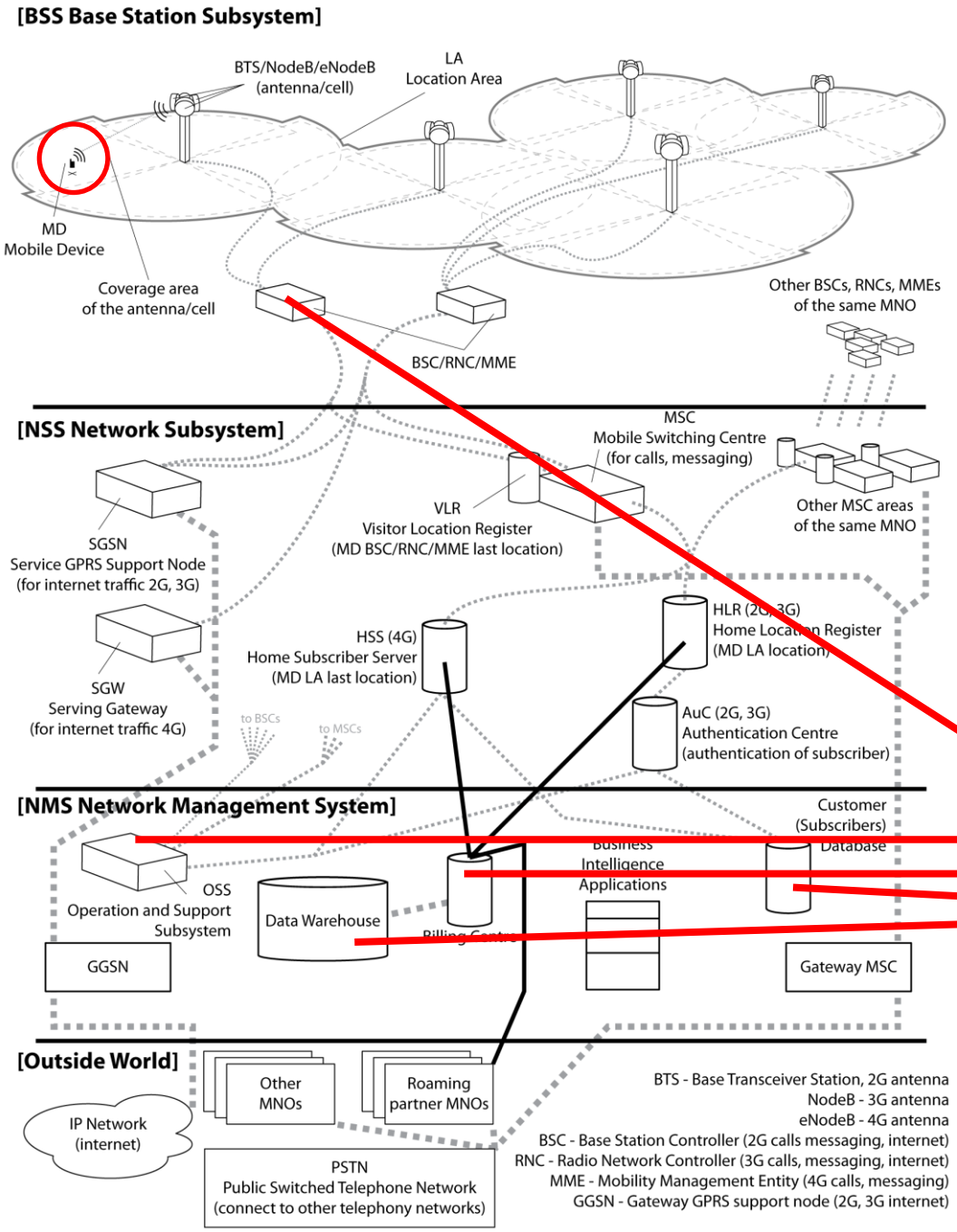
CDR – Call Detail Records:

- Outgoing calls, messages (avg 2-3 CDR/subsc/day)
- Incoming calls, messages (avg 2-3 CDR/subsc/day)
- Roaming data – the source for outbound data
- Internet traffic (DDR/IPDR) (avg 150 DDR/subsc/day)

Location Area updates

Other network data (handover data, network probes, etc.)

Technical Infrastructure of MNO



Location events collected from different registries and databases

Forms of Passive Positioning Data

- Domestic data – home subscribers
- Outbound roaming – home subscribers abroad
- Inbound roaming – foreign subscribers

Additional Data

- Geographic antennae reference data (needed for inbound roaming and domestic data)
- CRM data (demography, phone usage, customer value, average phone bill, etc.)
- Mobile banking (if SIM card connected to banking account)

Applications

Tourism statistics

- Number of trips & Number of unique travellers
- Duration of the visit in a destination country / in sub-regions
- Breakdown by the country of origin for foreign tourists

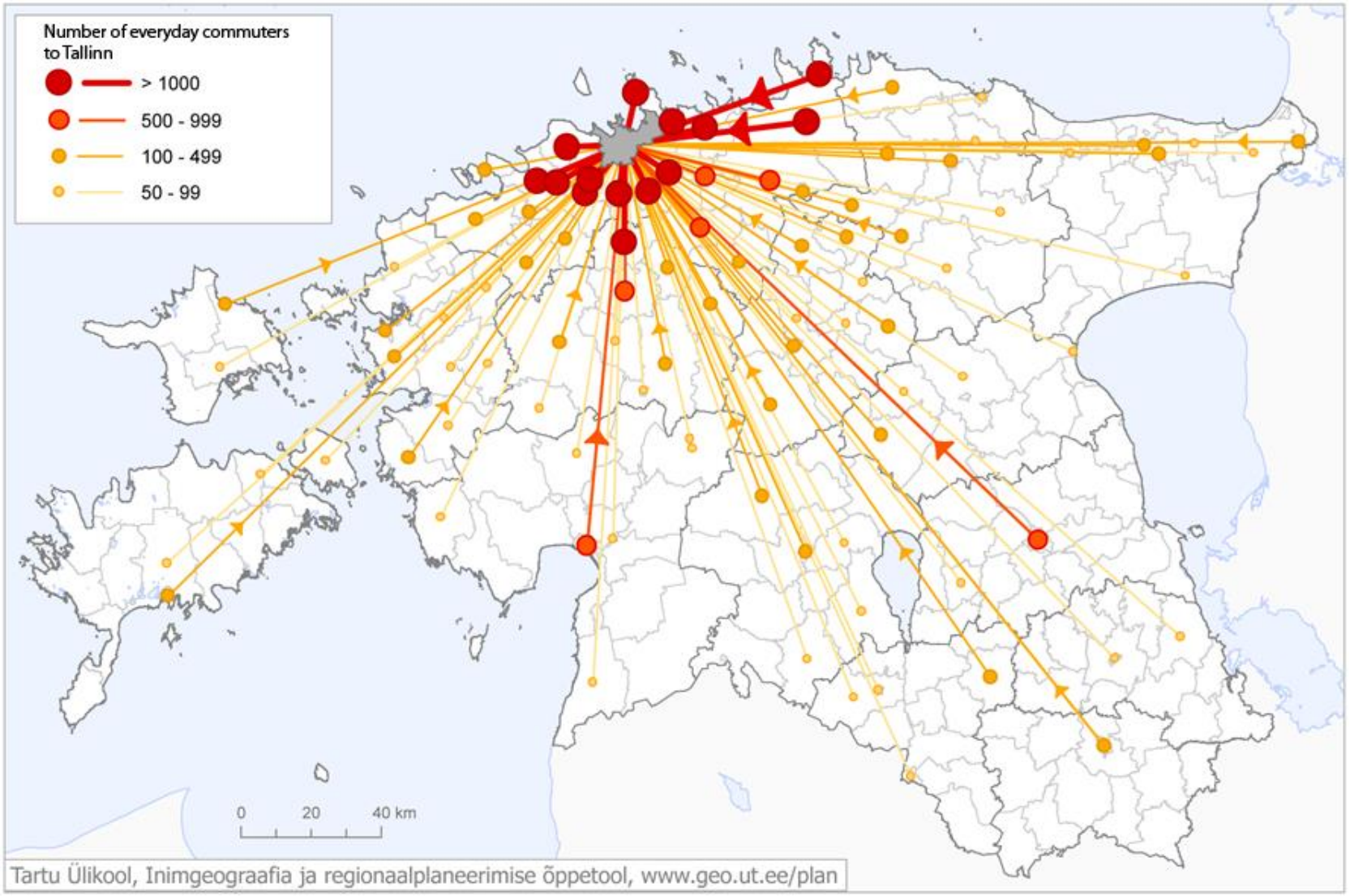
Transportation statistics

- Origin-destination matrices with hourly and daily travel numbers
- Identification of everyday commuting patterns
- Actual distance travelled

Population statistics

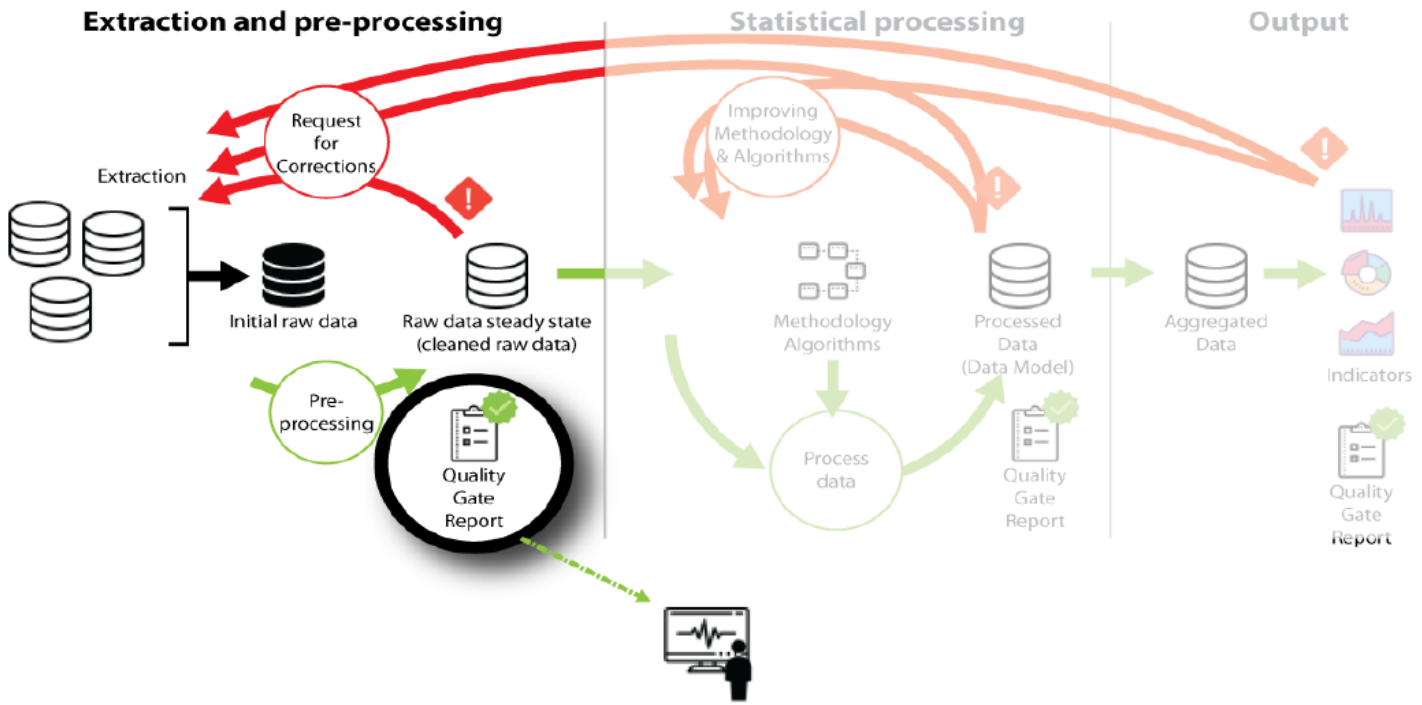
- Identification of workplace, school, secondary home, and other regular locations
- Internal migration based on the change of the residences within the country
- Temporary population (hourly, daily, weekly, monthly, etc.)

Commuters to Tallinn



Quality Gate 1 – Raw Data

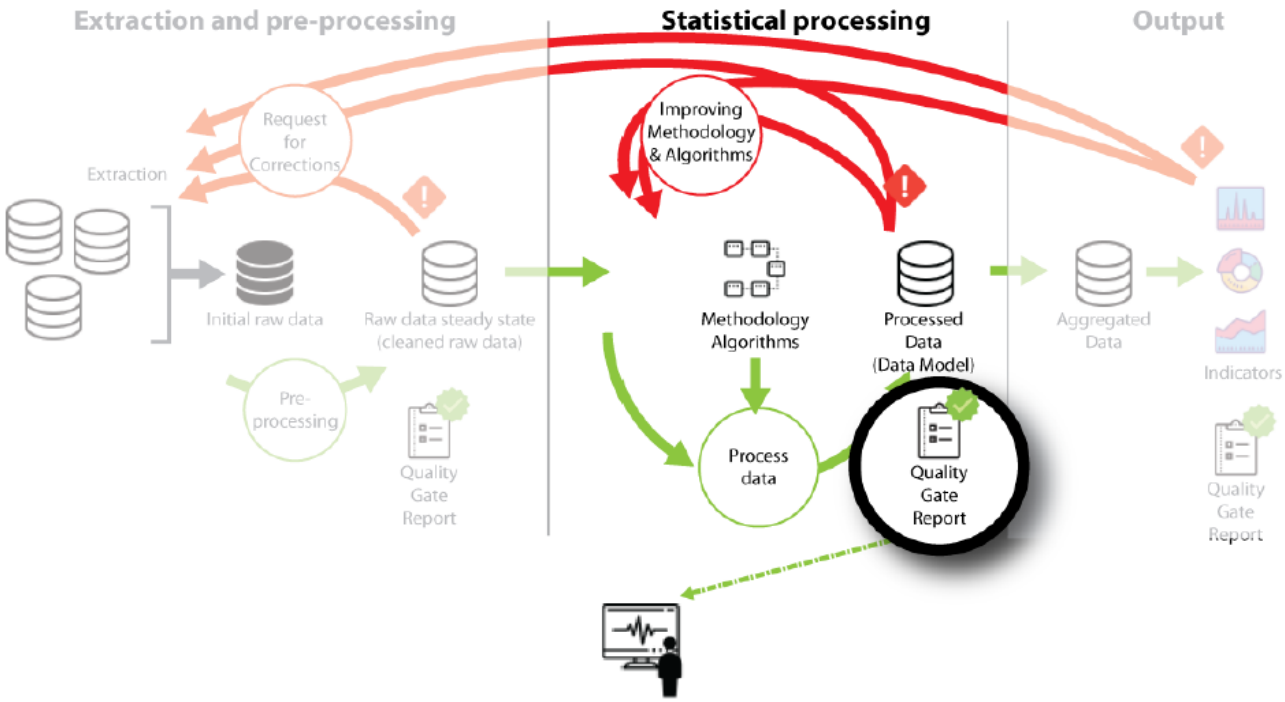
NATIONAL CENTRE FOR STATISTICS & INFORMATION
SULTANATE OF OMAN



1

Quality Gate 2 – Modelled Data

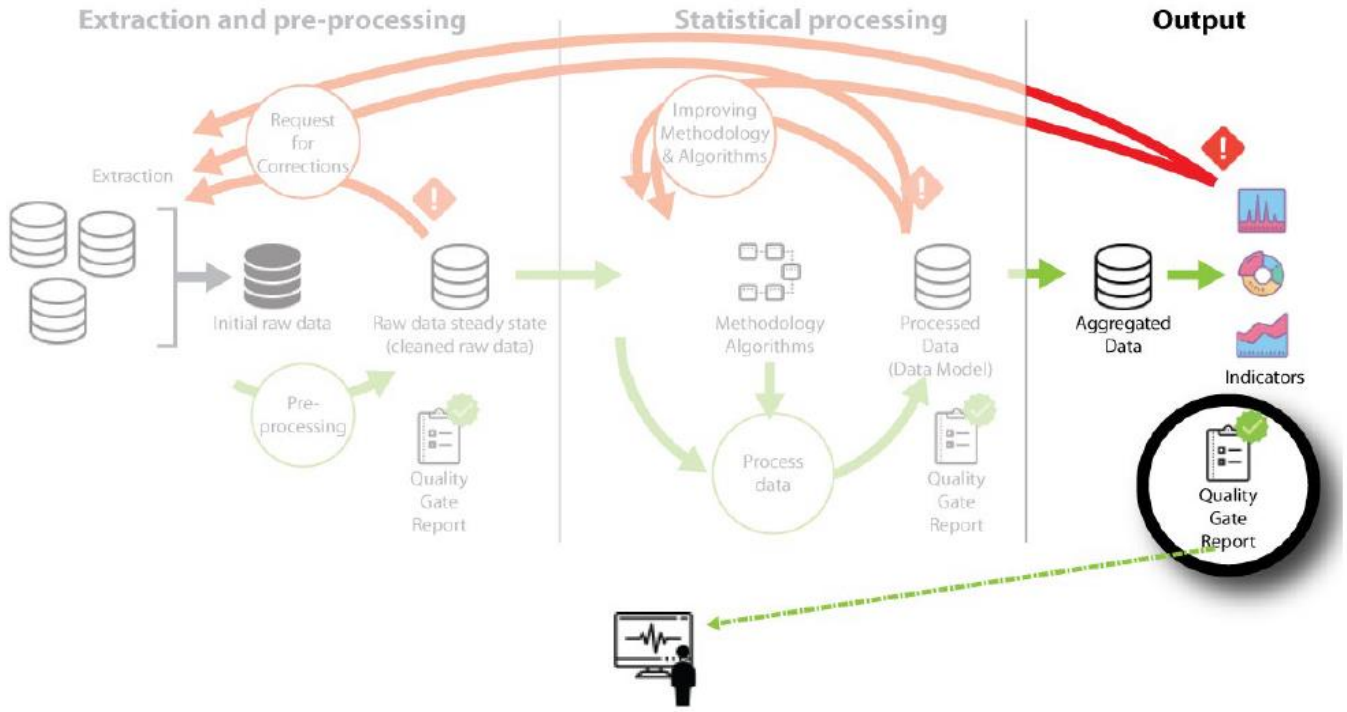
NATIONAL CENTRE FOR STATISTICS & INFORMATION
SULTANATE OF OMAN



2

Quality Gate 3 – Output Data

NATIONAL CENTRE FOR STATISTICS & INFORMATION
SULTANATE OF OMAN



3

Guiding principles to maintain public trust in the use of mobile operator data for policy purposes

Ronald Jansen [1*] Karoly Kovacs [1] Siim Esko [2] Erki Saluveer [2]
Kaja Sõstra [3] Linus Bengtsson [4] Tracey Li [4] Wole Ademola Adewole [4]
Jade Nester [5] Ayumi Arai [6] Esperanza Magpantay [7]

[1] United Nations Statistics Division

[2] Positium

[3] Statistics Estonia

[4] Flowminder

[5] GSMA

[6] University of Tokyo

[7] International Telecommunication Union

Partners in mobile phone data project

Mobile Network Operator

- Vodafone
- Telia, Elisa, Tele2

Government Mobile Network Regulator

- Public Utilities Regulatory Authority (the Gambia)

National Data Protection Authority

- Ghanaian Data Protection Commission

Intermediary Service Provider

- Positium
- Flowminder
- University of Tokyo

National Statistical Office

- Statistics Estonia
- Ghana Statistical Service
- Gambia Bureau of Statistics

Five Principles to ensure trust

Following the Fundamental Principles of Official Statistics

Necessity and proportionality (Fit-for-purpose)

Professional independence

Privacy protection

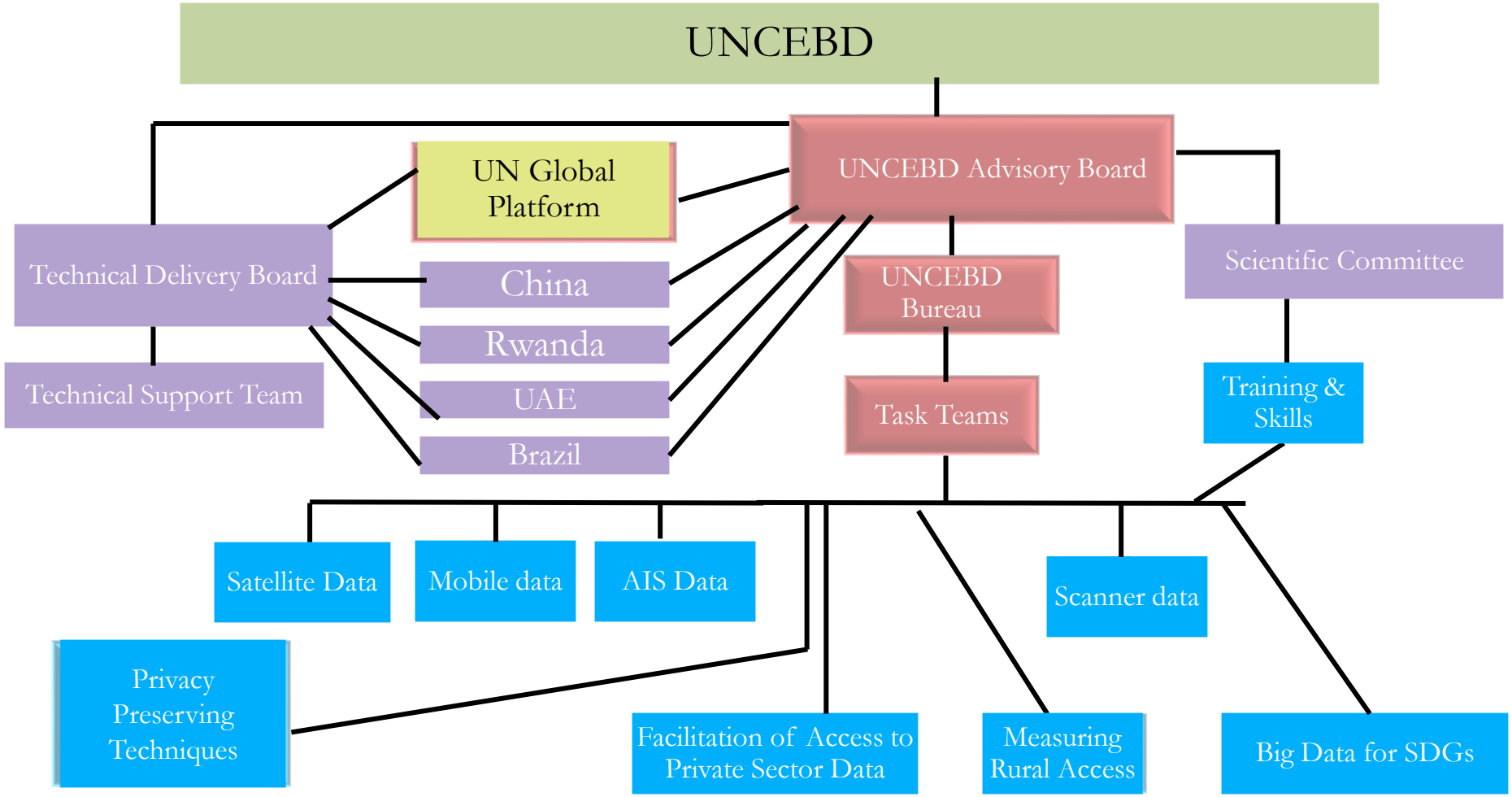
Commitment to quality

International comparability

UN Committee of Experts on Big Data and data science for official statistics

- Created in March 2014 by the UN Statistical Commission
- Mandated to give direction to the use of Big Data for Official Statistics
- Inter-governmental body with 31 countries and 16 international organizations
- Collaboration of more than 300 experts from all stakeholder communities

United Nations BigData


UNCEBD
UNCEBD Advisory Board
UN Global Platform
Technical Delivery Board
China
Rwanda
UAE
Brazil
Technical Support Team
UNCEBD Bureau
Scientific Committee
Task Teams
Training & Skills
Satellite Data
Mobile data
AIS Data
Scanner data
Privacy Preserving Techniques
Facilitation of Access to Private Sector Data
Measuring Rural Access
Big Data for SDGs

UNBigData

HOME ABOUT EVENTS TASK TEAMS UN GLOBAL PLATFORM

Mobile Phone Data

Task Team of the UN Committee of Experts on Big Data and Data Science for Official Statistics

Home › Task Teams › Mobile Phone Data

Introduction

The statistical community has the obligation of exploring the use of new data sources, such as Big Data, to meet the expectation of the society for enhanced products and improved and more efficient ways of working. Use of Big Data could also support the monitoring of the Sustainable Development Goals (SDGs) by improving timeliness, frequency, detail and relevance of indicators without compromising their impartiality and methodological soundness. The reports of the UN Committee of Experts on Big Data and Data Science for Official Statistics (UN-CEBD) to the Statistical Commission (E/CN.3/2015/4, E/CN.3/2016/6, E/CN.3/2017/7, E/CN.3/2018/8 and E/CN.3/2019/27) provide additional background to the work of the task teams.

Mobile phones are used by large parts of the population in all parts of the world, and it is thus expected that Mobile Phone data could fill data gaps worldwide. In its 2018 "Measuring the Information Society

Publications

Handbook on the use of Mobile Phone data for Official Statistics

Events

Oman's Experience in Utilizing Mobile Positioning Data for Official Statistics

Virtual Webinar 12 Apr 2021

Handbooks – areas of focus

1) Tourism statistics (lead: Indonesia)

2) Migration statistics (Lead: Georgia)

3) Dynamic population (lead: Positium)

4) Transport and commuting statistics (lead: World Bank)

5) Information society indicators (lead: ITU)

6) Displacement in disaster context (lead: University of Tokyo)

UNBigData

HOME ABOUT EVENTS TASK TEAMS UN GLOBAL PLATFORM

Privacy Preserving Techniques

Task Team of the UN Committee of Experts on Big Data and Data Science for Official Statistics

Home › Task Teams › Privacy Preserving Techniques

Introduction

The Privacy Preserving Techniques Task Team (PPTTT) is advising the UN Committee of Experts on Big Data and Data Science for Official Statistics (UN-CEBD) on Big Data on developing the data policy framework for governance and information management of the global platform, specifically around supporting privacy preserving technique.

Membership

Members have substantial experience and expertise in encryption techniques, algorithms and products/services. The membership will be reviewed annually.

Workstreams

The team's work spans multiple valuable workstreams:

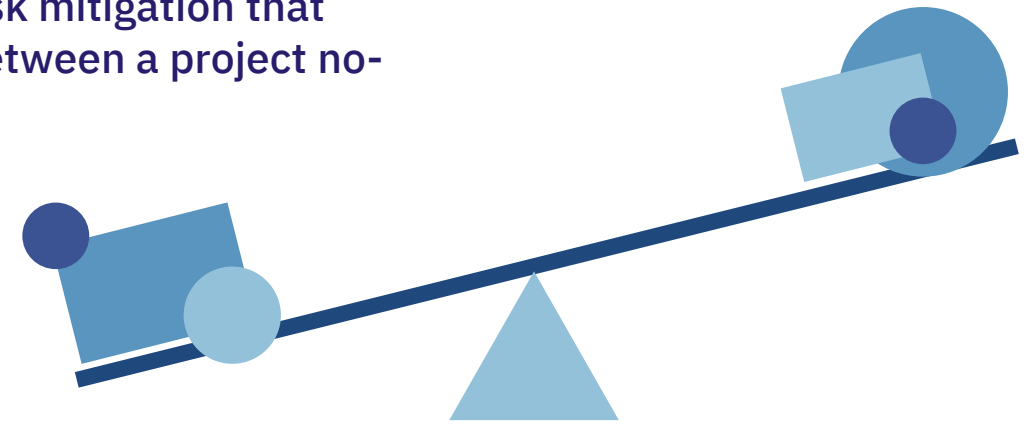
1. **UN Handbooks**
   (first published in 2018)
2. **Exchange of Experience**
   (for example application of PPT in COVID-19 response activities in 2020)
3. **Training Courses**
   (in partnership with *openmined.org)*
4. **Experimentation**
   (UN GP infrastructure + PET technologies)
5. **Promotion**
   (participation in events and other projects)
6. **Building PPT Community of Practice in Official Statistics**

United Nations BigData

The Balancing Act

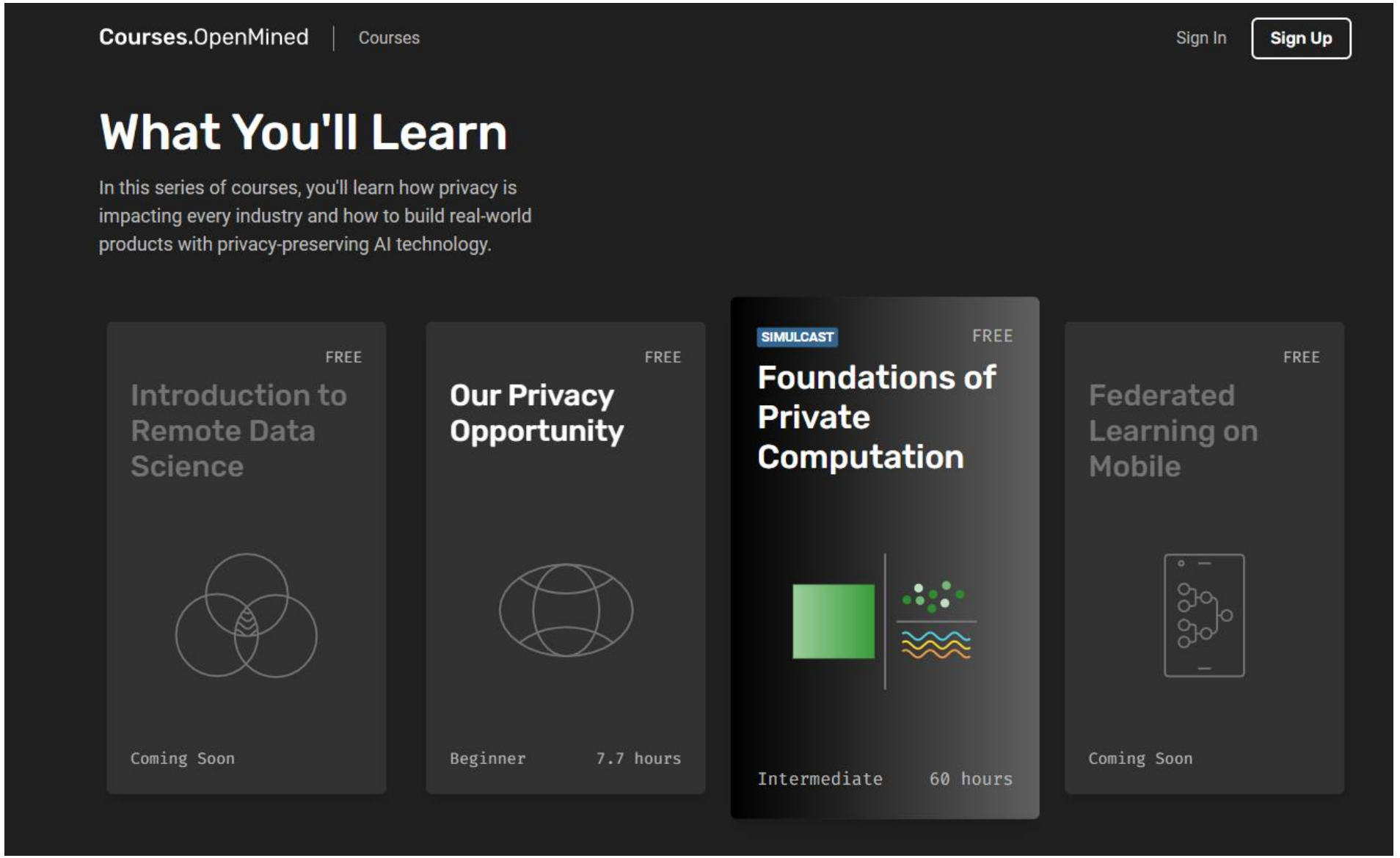
We continually play the game of balancing data usability and security:

PPTs *do not solve* the balancing act of security, privacy and data use in and of themselves, but offer risk mitigation that may be the difference between a project no-go and go.



Security & Privacy

Value Creation from Statistical Inference


Courses.OpenMined
Courses
Sign In
Sign Up
What You'll Learn
In this series of courses, you'll learn how privacy is impacting every industry and how to build real-world products with privacy-preserving AI technology.
FREE
Introduction to Remote Data Science
Coming Soon
FREE
Our Privacy Opportunity
Beginner
7.7 hours
SIMULCAST
FREE
Foundations of Private Computation
Intermediate
60 hours
FREE
Federated Learning on Mobile
Coming Soon

Training, Competencies and Capacity Development

Task Team of the UN Committee of Experts on Big Data and Data Science for Official Statistics

Introduction

Download Terms of Reference

The statistical community has the obligation of exploring the use of new data sources, such as Big Data, to meet the expectation of the society for enhanced products and improved and more efficient ways of working. Big Data could also support the monitoring of the Sustainable Development Goals by improving timeliness and relevance of indicators. This should go without compromising their impartiality and methodological soundness.

Big Data is by definition different from traditional data sources currently used by National Statistical Systems (NSSs) requiring the development of new methodologies. Big Data sources pose challenges regarding methodology, quality assurance, technology, security, privacy and legal matters. This means that new skill sets are necessary. Some of which could be hired temporarily, others will need to become an integral part of the institution. It is up to the senior management to decide what will be done by the institute itself and what will be outsourced. Most likely, the statistical institute will need to build long-term partnerships with private sector, academia and research institutes to successfully work with new data sources and new technologies.

Recent Events

6th International Conference on Big Data

Session 10 - Training in use of new data sources and new technologies - Sep 2020

- Why teaching Big Data?
  Christophe Bontemps, SIAP
- Big Data training courses in STI Korea
  Jeongran Kim, Jaemin Na, STI Korea
- Big Data Training in a Post-Covid Environment
  David Johnson, ONS UK
- UN Task Team on Training, Competencies and Capacity Development
  Ceri Regan, ONS UK; Dominika Nowak, Statistics Poland

Training in Big Data and Data Science for official statistics

- **Big Data Training Curriculum**
- **E-Learning Courses**
- **Big Data Maturity Matrix**
- **Training of data scientist in academic centers**
- **Big Data Competency Framework**
- **Mentorship**

UN Global Platform – Collaboration on projects with over 100 institutes worldwide

Using AIS (Shipping Data) to provide faster economic indicators

Producing consumer price indices using Scanner data and webscraping

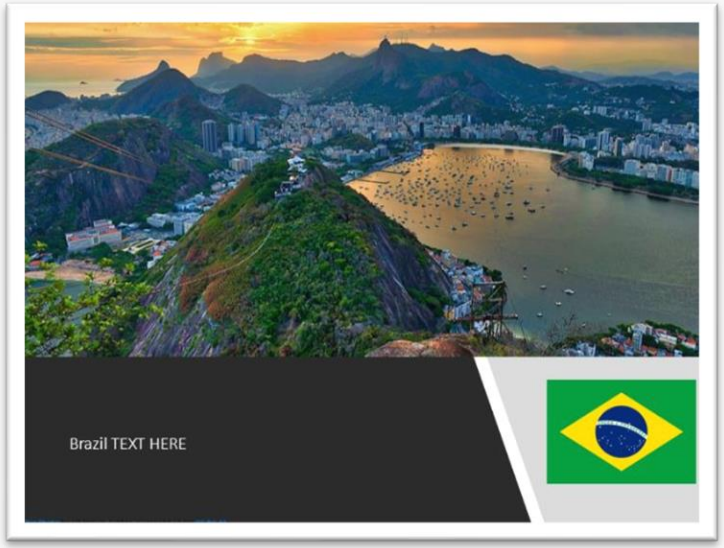
Producing crop statistics from Earth Observations and data pipelines

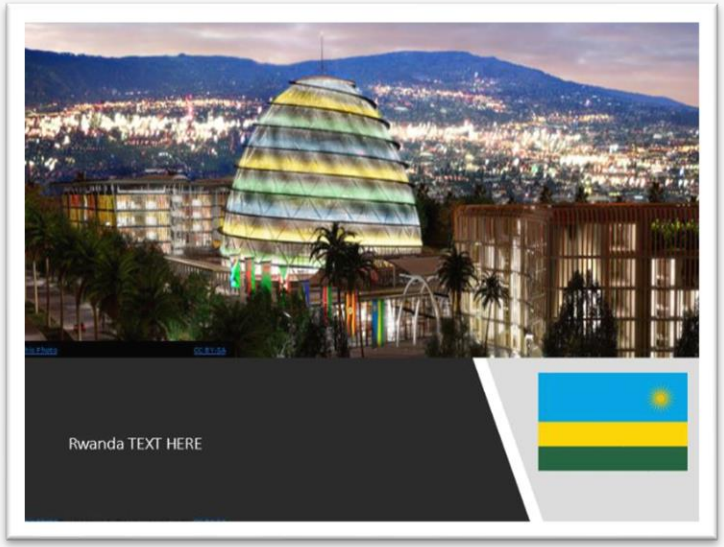
Producing tourism, migration, population and transport statistics using mobile phone data

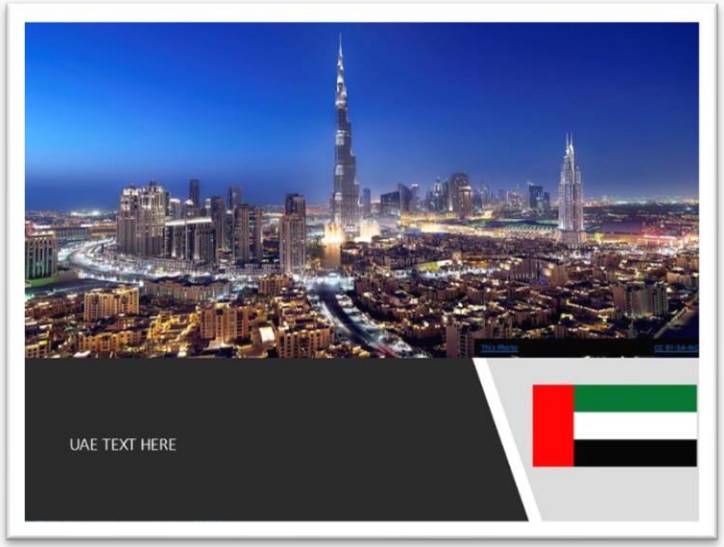
Developing techniques to preserve the privacy of sensitive data

Machine Learning projects

Activities of Regional hubs







Regional Hub for Big Data in Brazil
Rio de Janeiro, ENCE / IBGE

- Strengthening ties and promoting cooperation between producers of official statistics in the Region
- Training and fostering the interest of young statisticians on the use of Big Data in Official Statistics
- Supporting research on the use of Big Data and Data Science
- Organizing and hosting seminars and conferences

HDSR

HARVARD DATA SCIENCE REVIEW

A Telescopic, Microscopic, and Kaleidoscopic View of Data Science

Data Science and Official Statistics: Toward a New Data Culture

by Stefan Schweinfest and Ronald Jansen

Full article forthcoming in December

Toward a New Data Culture

NSO as Data Steward – making more data sources available to society

New data governance – fundamental principles for the whole data community

Data privacy by design

Data integration for a holistic approach to data and policies

(statistics / geospatial information)

Publication of experimental data with appropriate communication

Training / labor force of statisticians and data scientists

References

- Home — UN-CEBD
- Task Teams — UN-CEBD
- Mobile Phone Data Handbook
- Margus Tiru - Mobile Positioning Data Paper
- Global Network Webinar: Oman's Experience in Utilizing Mobile Positioning Data for Official Statistics
- Guiding principles to maintain public trust in the use of mobile operator data for policy purposes | Data & Policy | Cambridge Core
- UN Handbook for Privacy-Preserving Techniques
- Road to EXPO2020: Methods and applications
- Road to EXPO2020: UN Global Platform

Thank You

Ronald Jansen | Assistant Director
Chief, Data Innovation and Capacity Branch
Statistics Division | Department of Economic and Social Affairs
United Nations
Email: jansen1@un.org | BigData@un.org
http://unstats.un.org/bigdata