Outline

- Introduction
- ITU Big Data work
- Use cases
- Recommendations (national coordination, stakeholders)
- Way Forward
- Other resources

Introduction – ITU's ICT statistics work

Facts and figures

ICT prices

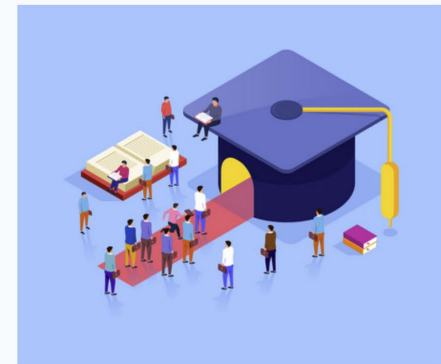**Universal and meaningful connectivity targets**

Big data

Expert groups

Partnerships

Standards and definitions

Capacity development

Overview - Information society indicators included in the SDG monitoring framework – collected by ITU

- Target 4.4: Proportion of youth/adults with ICT skills, by type of skills

- Target 5b: Proportion of individuals who own a mobile telephone, by sex
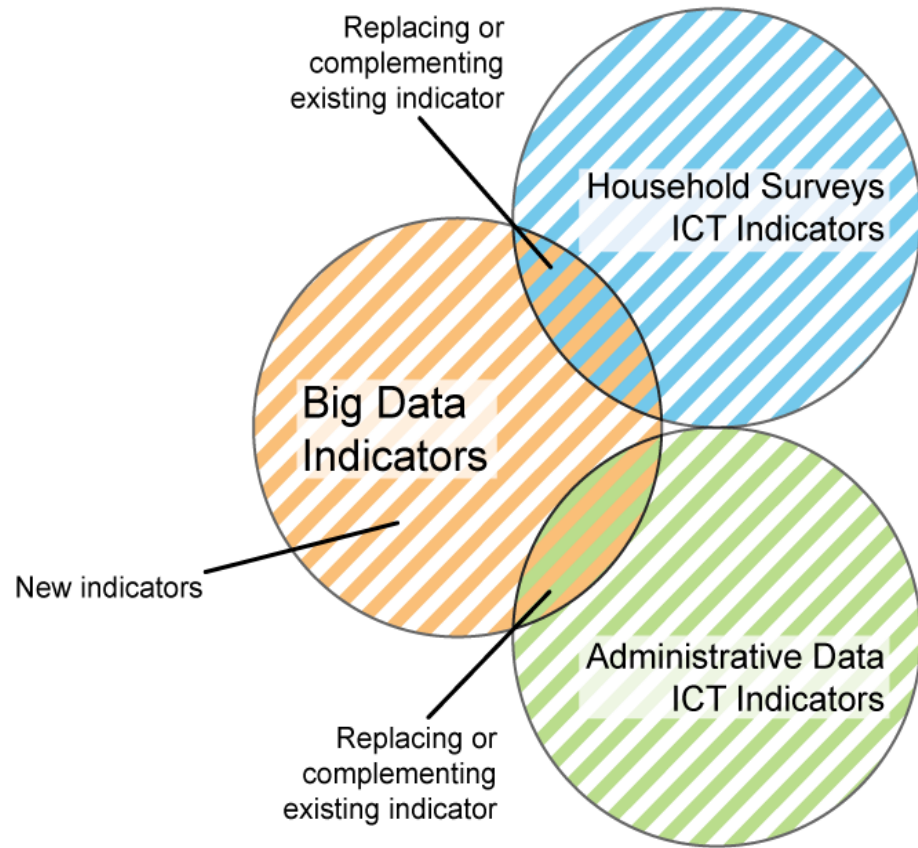
- **Target 9c: Percentage of the population covered by a mobile network, broken down by technology**

- Target 17.6: Fixed Internet broadband subscriptions, broken down by speed

- **Target 17.8: Proportion of individuals using the Internet**

ITU ICT statistics work

ITU data collection schedule 2024				
Questionnaire	**Launch**	**Open**	**Deadline**	**To be published on ITU DataHub by:**
Household Short Questionnaire	25-Mar-2024	5 weeks	26-Apr-2024	31-Jul-2024
World Telecommunication/ICT Indicators Short Questionnaire	25-Mar-2024	5 weeks	26-Apr-2024	31-Jul-2024
ICT Price Basket Questionnaire	15-May-2024	2 weeks	30-May-2024	31-Mar-2025
Household Long Questionnaire	16-Sep-2024	4 weeks	16-Oct-2024	31-Jan-2025
World Telecommunication/ICT Indicators Long Questionnaire	16-Sep-2024	4 weeks	16-Oct-2024	31-Jan-2025

https://www.itu.int/en/ITU-D/Statistics/Pages/bigdata/default.aspx

ITU Big Data work

1st pilot: 2016-2017
- ✓ 6 countries (Colombia, Georgia, Kenya, Philippines, Sweden, UAE)
- ✓ 16 ICT indicators (administrative data)

2nd pilot: 2020-2021
- ✓ Brazil, Indonesia

- ✓ 2 SDG ICT indicators
  - ✓ 9.c.1 – Percentage of population covered by mobile network: 2G, 3G and 4G and above (administrative data)
  - ✓ 17.8.1 – Percentage of population using the Internet (household survey data)
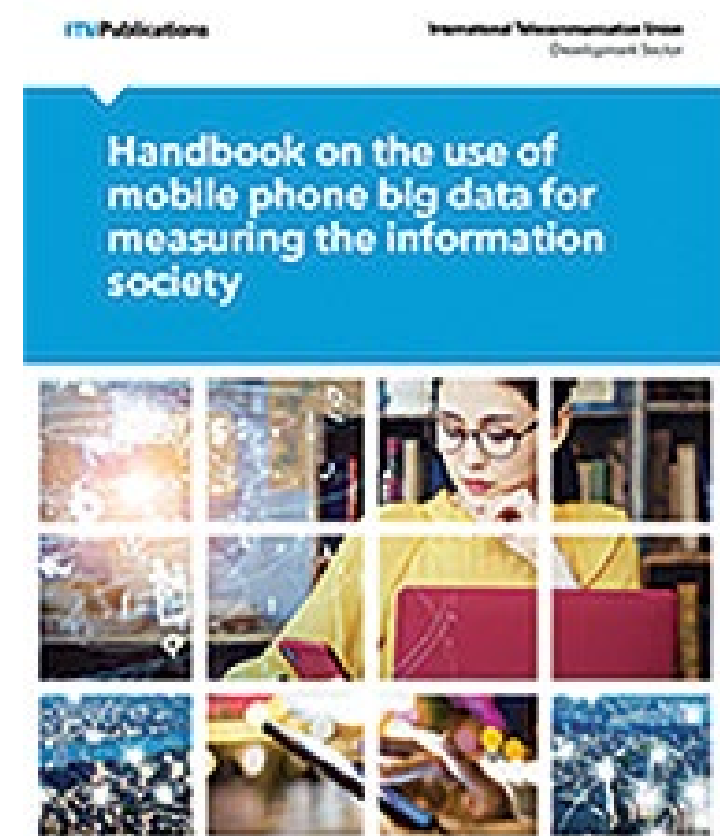
Ongoing: 2023-2024
Uganda, Malaysia, Mongolia, Liberia, Uruguay, Tunisia, Botswana

Example:
# Guideline on Big Data for measuring the SDG Information society indicators (Lead: ITU)

1. Introduction
2. Background
3. Access and preparations
4. Data sources (description of mobile operator data, quality assurance of raw data)
5. Reference data (local admin units, world population, cell data, digital elevation, household survey data)
6. Data processing (models, data protection guidelines)
7. Calculating the indicators (rationale, definition, indicators calculation, quality assurance)
8. Quality assurance
9. Conclusions

- with experiences and examples from country pilots

ITU Publications   International Telecommunication Union
                                    Development Sector

Handbook on the use of mobile phone big data for measuring the information society

Example: Information society indicators

SDG indicator 17.8.1: Percentage of the population using the Internet, Rio de Janeiro, Brazil, 2021
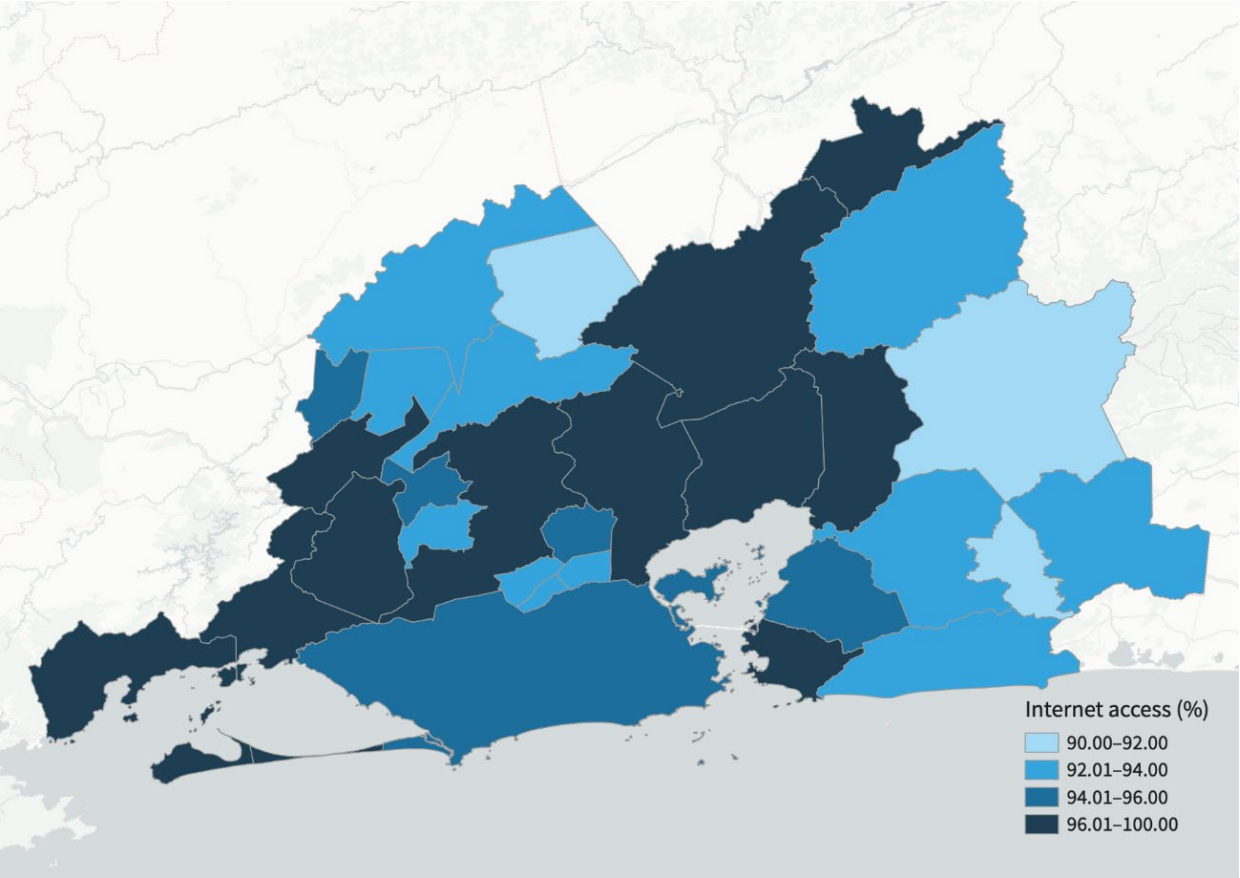
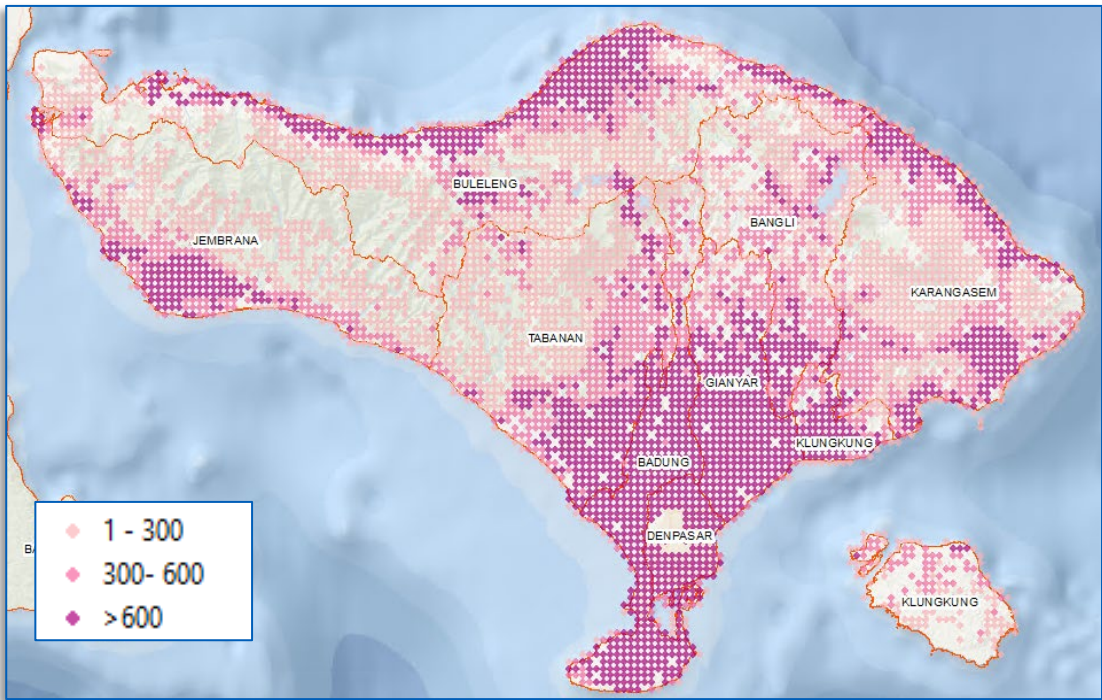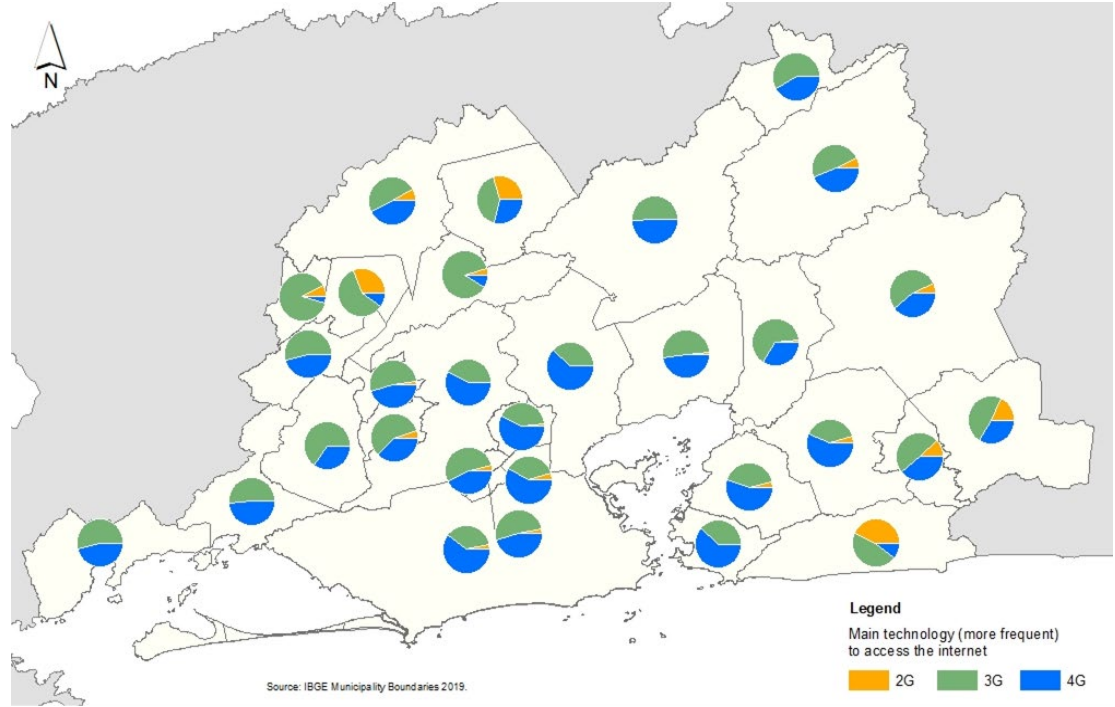Mobile phone data in line with household survey results

Internet access (%)
- 90.00–92.00
- 92.01–94.00
- 94.01–96.00
- 96.01–100.00

■ MPD ■ Survey

Results: Percentage of population covered by mobile signal (SDG 9.c.1)

Indonesia

Brazil

National coordination: Roles of different stakeholders

1. Telecommunication Regulator / ICT ministry
   - ✓ Request mobile phone data as a requisite for monitoring licensing condition
   - ✓ Invested in equipment and expertise to store and process these records
   - ✓ Frequently interact with operators in the course of their regulatory work
   - ✓ Positioned to negotiate and mediate access to mobile phone data
2. National Statistics Office
   - ✓ Statistical Act
   - ✓ Mandate to produce official statistics and collect the data
   - ✓ Skills to analyse statistical information
3. Mobile Phone Operators and service providers
   - ✓ Custodian of mobile phone data
   - ✓ Invested in equipment and expertise to store and process these records
   - ✓ Required to submit records to a regulatory agency as a condition of their licence or franchise
   - ✓ Have staff that have big data skills to analyse MPD
4. Data Protection Authority
   - ✓ Provides guidance and oversight for lawful data processing
   - ✓ Ensures safeguards are in place to ensure privacy (pseudonymisation or anonymisation)

Here are some summarized tips for convincing mobile operators to share data for official statistics:

1. Build Trust and Relationships
- **Partnerships:** Develop long-term partnerships to ensure ongoing data sharing. **Transparency:** Be clear about data usage and benefits to build trust.

2. Legal and Regulatory Framework
- **Compliance:** Ensure data sharing complies with national privacy laws. **Formal Agreements:** Use MoUs to define responsibilities and legal obligations.

3. Data Anonymization and Security
- **Anonymization:** Assure that data will be anonymized to protect privacy. **Security Measures:** Highlight strong data security protocols.

4. Show the Value of Data Sharing
- **Case Studies:** Use successful examples to illustrate benefits. **Insights:** Share useful insights derived from the data.

5. Simplify the Process
- **Standardization:** Use standardized formats and protocols. **Technical Support:** Provide support for data extraction and processing.

6. Collaborative Frameworks
- **Working Groups:** Create groups with representatives from both sides to address concerns. **Public-Private Partnerships:** Foster collaboration through partnerships.

7. Incentives and Recognition
- **Incentives:** Benefits like tax breaks or regulatory leniency. **Recognition:** Publicly recognize and commend participating operators.

8. Communication and Advocacy
- **Workshops:** Organize educational workshops and seminars. **Advocacy Campaigns:** Raise awareness through campaigns about the importance of data sharing

Way forward

Synthetic mobile phone data

- to create a simulated dataset that mimics real-world mobile device movements and positions

- purpose of testing out codes related to mobile positioning

- addresses privacy, overcome data access issues, enable testing and development of algorithms, reduce data collection costs, innovation and development

- task team sub-group that works on developing methods

ITU Python codes for information society

Field Name	Type	Mode	Description
msisdn	String		Hashed subscribers identifier
datetime	Timestamp		Transaction date (date and hour)
cell_id	String	NULLABLE	Hashed cell identifier
latitude	Float		Latitude of Base Transceiver Station (BTS)
longitude	Float		Longitude of Base Transceiver Station (BTS)
data_type	String		Data source, can be CDR/CHG or IPDR/UPCC
service	String		Transaction service (4G/ 3G/ 2G)

Field name	Type	Mode	Description
msisdn	String		Hashed subscribers identifier
age	Int		Subscribers age from registration data
gender	String		Subscribers gender (M/F) from registration data

```python
# detect duplicate rows
df_duplicates = df.groupBy(df.columns).count().filter("count > 1")
print(f"number of duplicate rows: {df_duplicates.count()}")
df_duplicates.show()
```

```
number of duplicate rows: 539
[Stage 11:>                                              (0 + 3) / 3]
+------+-------------------+-------+--------+---------+---------+-------+----------+-----+
|msisdn|           datetime|cell_id|latitude|longitude|data_type|service|      date|count|
+------+-------------------+-------+--------+---------+---------+-------+----------+-----+
|     8|2024-08-09 03:10:00|  164.0|  43.277|   -3.163|     IPDR|     3G|2024-08-09|    2|
|     8|2024-10-28 05:19:00|  164.0|  43.277|   -3.163|     IPDR|     3G|2024-10-28|    2|
```

```python
# Print the number of records in the DataFrame
print("Number of records before deduplication: {}\n".format(df.count()))

# Drops the duplicate rows from the dataframe
df = df.dropDuplicates()

# Display the first five rows of the DataFrame in a tabular format
df.show(5)

# Print the number of records in the DataFrame
print("Number of records after deduplication: {}".format(df.count()))
```

```
Number of records before deduplication: 536851
```

```
+------+-------------------+-------+--------+---------+---------+-------+----------+
|msisdn|           datetime|cell_id|latitude|longitude|data_type|service|      date|
+------+-------------------+-------+--------+---------+---------+-------+----------+
|     0|2024-05-11 17:19:00|  419.0|  43.217|   -3.122|     IPDR|     3G|2024-05-11|
|     0|2024-05-19 03:01:00|  746.0|  43.327|   -3.089|     IPDR|     4G|2024-05-19|
|     0|2024-05-25 07:00:00|  873.0|  43.321|   -3.077|     IPDR|     3G|2024-05-25|
|     0|2024-05-29 11:13:00|  655.0|  43.327|     -3.1|     IPDR|     3G|2024-05-29|
|     0|2024-06-01 23:11:00|  786.0|  43.318|   -3.084|      CDR|     2G|2024-06-01|
+------+-------------------+-------+--------+---------+---------+-------+----------+
only showing top 5 rows
```

```
[Stage 20:====================>                          (1 + 2) / 3]
Number of records after deduplication: 536310
```

ITU-WB collaboration on using MPD to work for policy

- **Aspiration:** Operationalize the responsible use of MPD for policy and Official Statistics in 30 countries. Put this data to work for dynamic insights and actions to tackle migration, disaster, epidemics, information society and more.

- **Where to play:** in the local national mobile phone data ecosystems with public and private stakeholders to foster ownership, shared value, sustainable capabilities and responsible use

- **How to win:** Integrated and comprehensive approach to develop local capacity through phased engagements that invest in public private partnerships, and institutional enablers and safeguards.

- **Capabilities/Systems Needed:** WB grant window management; Partnership management; Training Content, Delivery Management; Collaborative Code and Learning Management Systems; etc

- **Timelines:** Kick-off in June 2024, Bilbao conference

8th International Conference on Big Data

- 10-14 June in Bilbao, Spain
- Themes: Sustainability of tourism, climate change, biodiversity
- Sessions on MPD:
  - High-level on MPD for tourism statistics
  - Demo of ITU Python codes
  - Workshops on MPD developments (use cases, synthetic data, bias, etc)
  - Workshops on maturity assessment, theory of change, roadmap for an MPD project

https://unstats.un.org/bigdata/events/2024/conference/

Additional resources

➤ [UN Big data task team on mobile phone data](#)

➤ [ITU Big Data pilots](#)

➤ [Online training course on mobile phone data](#)

Or contact us at: [indicators@itu.int](#)

Thank you!

indicators@itu.int