

Using innovative approaches to fill data gaps

Faryal Ahmed

UN Statistics Division | Development Data and Outreach Branch

Workshop on supporting evidence-based VNRs and SDG reporting: data and statistics innovations
Ankara, Türkiye | 6-7 December 2023



Why innovative approaches?

- Utilizes existing sources of data to produce and disseminate better, more timely and disaggregated data
- Allows integration of non-traditional and traditional data from various entities
- Capitalizes on resources from various entities and new partnerships



[This Photo](#) by Unknown Author is licensed under [CC BY-SA](#)



United Nations

DESA
Statistics Division

Link with official statistics

- Implementation to be in line with the **UN Fundamental Principles of Official Statistics**, following key data quality, privacy and confidentiality standards

Innovation to address policy priorities

- More granular data to plan targeted programs and policies
- Enables measurement of hard-to-count populations (such as those missed out in traditional data collection methods)
- Provides data with less time lag for timely decision-making (it may be costly to conduct more frequent traditional surveys)

Innovative Approaches

Partnerships

- CSOs
- Academia
- Development partners
- Private sector...

Data Sources

- Social media data
- Geospatial data
- Telecom data
- Citizen generated data
- Admin data....

Methods

- Probabilistic Record Linkage
- Small Area Estimation
- De-Identification, PET
- Web Scraping
- Natural Language Processing
- Machine Learning....

Tools

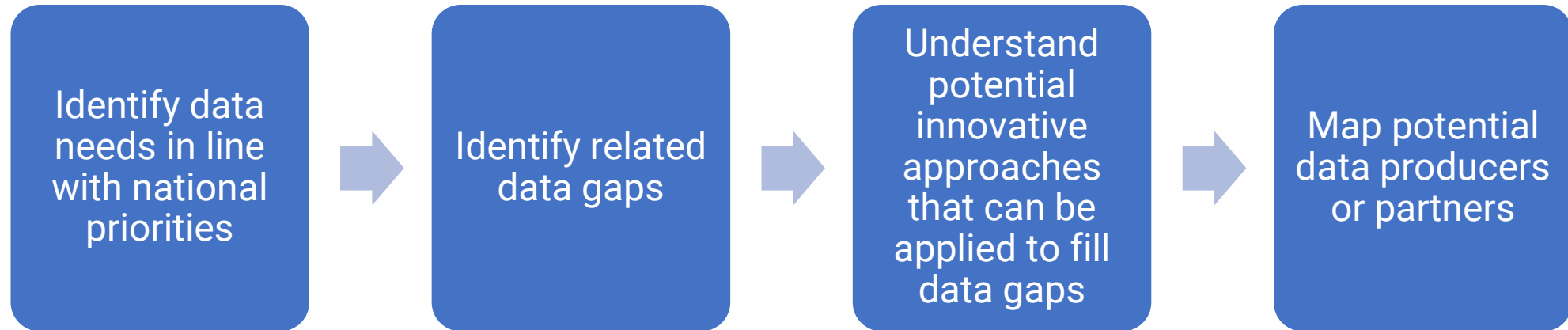
- Data Collection: Pentaho, R, Python, Batch/Stream Processing...
- Preparation: data warehouse, data-lake, Apache Parquet, R, Python...
- Analysis: Google Earth, QGIS, ArcGIS, R & Python libraries (NumPy, Pandas, BeautifulSoup...)...
- Presentation: Dashboards, DataMarts, DataAPI, ...
- Dataflow Automation....



United Nations

DESA
Statistics Division

Mapping of data stakeholders



United Nations

DESA
Statistics Division

Addressing data gaps through innovation: Examples



Small Area Estimation using nighttime lights for poverty estimation (Philippines)

Satellite imagery used to produce estimates of livestock population (South Sudan)
Small Area Estimation used for food insecurity experience scale (Chile)



Census and satellite imagery used for accessibility of roads for rural population (Japan, Colombia, Canada, China)



United Nations

DESA
Statistics Division

Addressing data gaps through innovation: Examples



Use of geospatial information for education-related indicators on home-to-school distance and drop-out (Colombia)

Geolocation data from cell phones used to track COVID-19 outbreaks (Bangladesh)

Early warning of norovirus outbreaks through predictive modelling from social media data (UK)



Marine litter monitored through citizen science data (Ghana)

Mobile positioning data used to increase coverage & timeliness of tourism data (Indonesia)



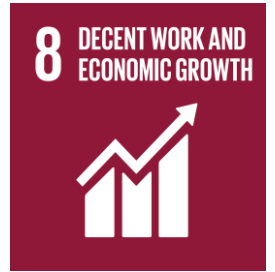
United Nations

DESA
Statistics Division

Addressing data gaps through innovation: Examples

Web-scraping of prices used to improve inflation data (Norway)

Natural language processing used of online job adverts used to provide weekly estimates of job vacancies (UK)



Producing experimental estimates of SDG 16 perception indicators through social media data (Colombia)

Use of satellite imagery, OpenStreetMap and statistical information to produce SDG 11.7.1 on access to open spaces (Palestine)



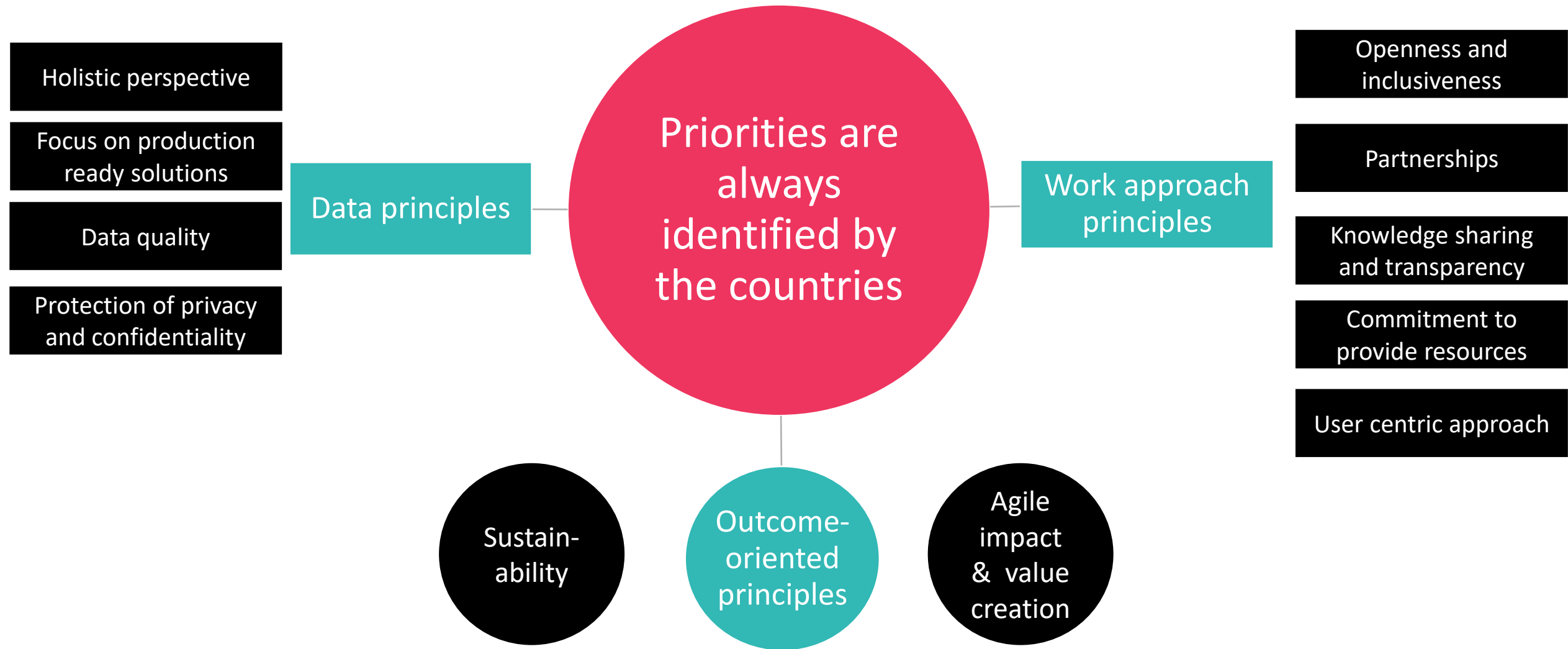
Earth observation data used to develop land and ecosystem extent accounts (Philippines)

The Data for Now initiative

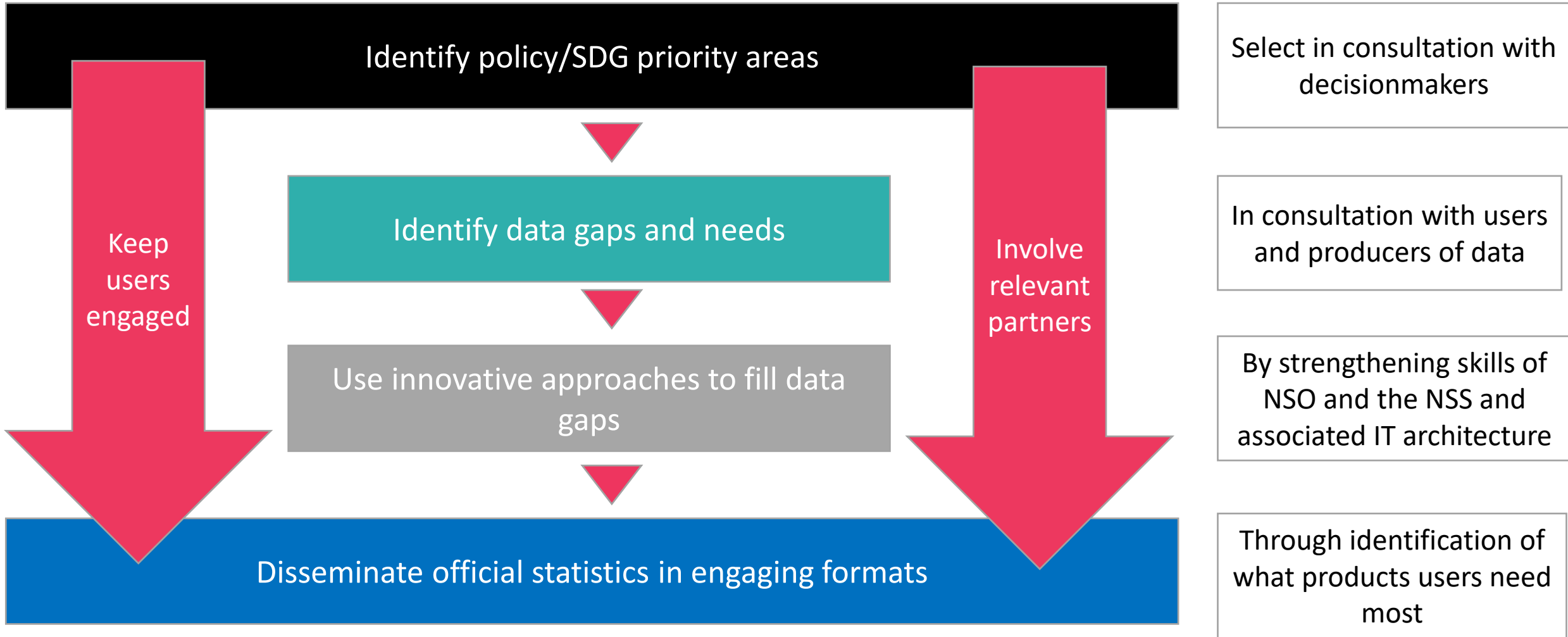
- Launched by UN Deputy-Secretary General Amina Mohammed on the sidelines of the UN General Assembly, in September 2019
- Supports countries in the use of innovative sources, technologies and methods for the streamlined production and dissemination of better, more timely and disaggregated data for sustainable development
- Is co-led by the United Nations Statistics Division (UNSD), United Nations Development Program (UNDP), the World Bank, the Global Partnership for Sustainable Development Data (GPSDD), and the Sustainable Development Solutions Network (SDSN)
- National Statistical Offices (NSOs) are the main partners in countries



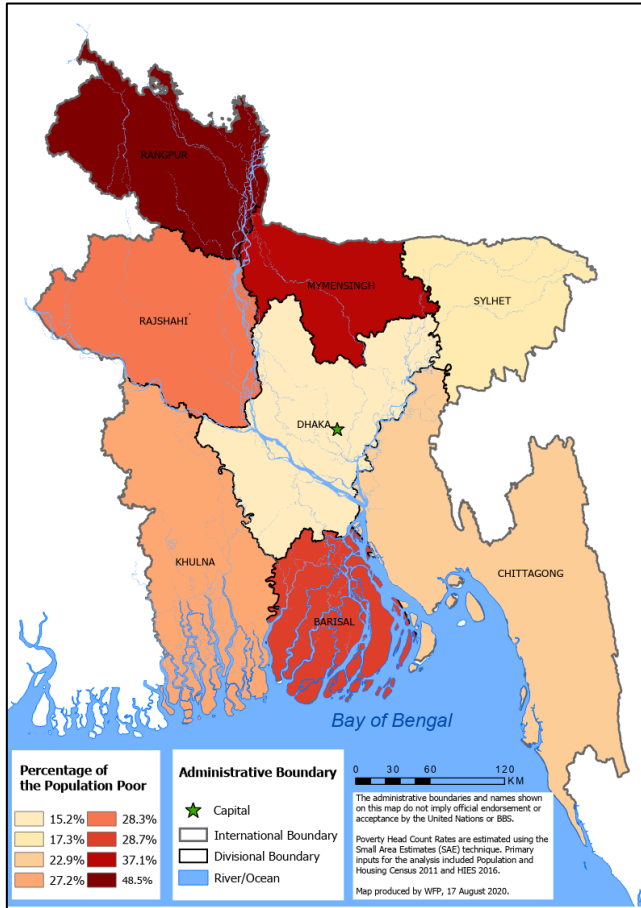
Guiding principles



Implementation process



Poverty estimation in Bangladesh



• The data need:

- Bangladesh Government identified reducing % of population living below poverty line as a National Priority Target
- Poverty measured using 5-yearly Household Income & Expenditure Survey, providing direct estimates at Division level
- More frequent disaggregated data at district/sub-district level deemed essential for better developing & targeting policy

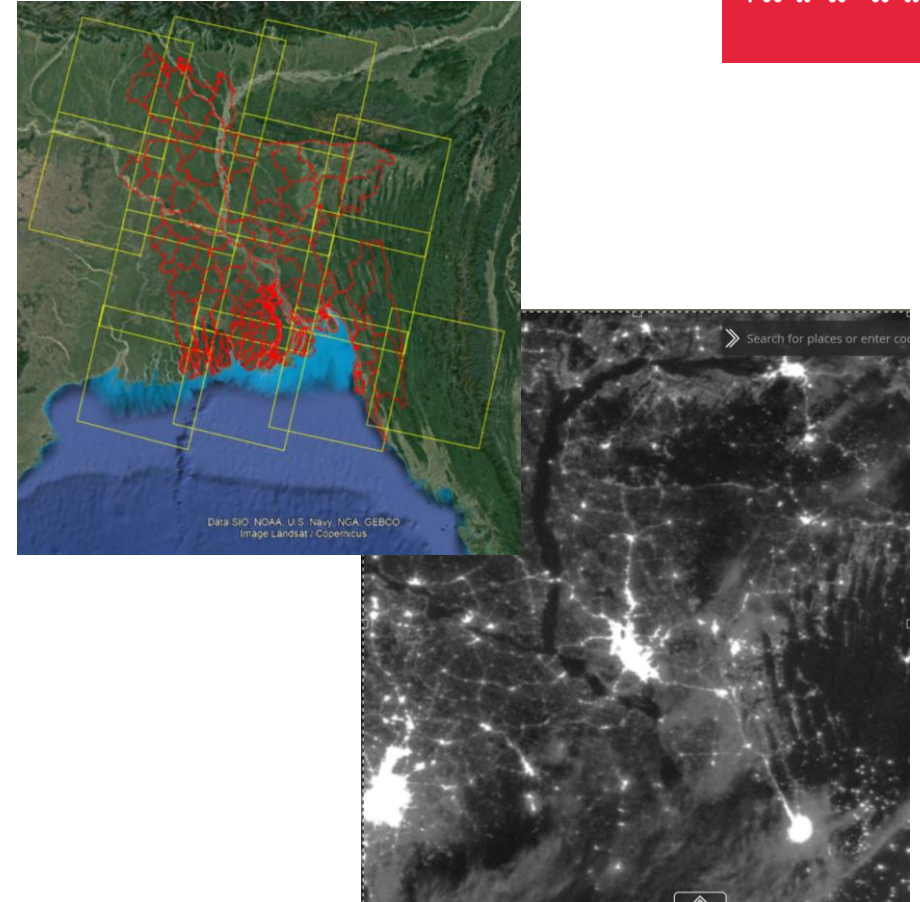
“The poverty rates in Bangladesh are changing frequently, because of many reasons including globalization and the COVID19, so if we use the data collected five years ago it would not reflect the reality. Bangladesh needs more frequent estimates”.

Poverty estimation in Bangladesh

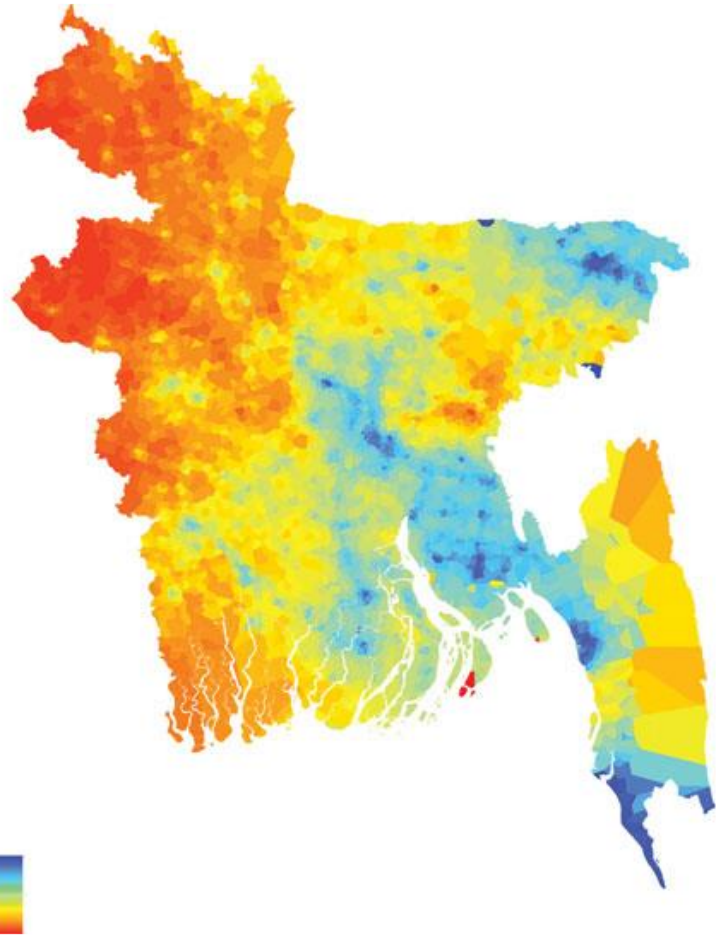


- **The response:**

- Through Data For Now, workshop held with line ministries & civil society to identify potential data sources to support more frequent estimates, including earth observation, telecoms & administrative data
- BBS led establishment of high-level steering group on poverty statistics to guide data acquisition & production
- Training workshops held for BBS & NSS partners on small-area estimation (SAE) methods & use of geospatial & other non-traditional data sources, applying techniques using real data



Poverty estimation in Bangladesh



- **The outcome:**

- Initial poverty estimates produced using sources including night time lights, land use & land cover & road density data.
- BBS continuing to refine models, bringing in additional data sources where available.
- Continued engagement with ministers and policy makers, securing high-level buy-in for continued development.
- Opportunities identified to apply techniques learned to other priority unmet data needs.
- In 2022, the new methodology was formally endorsed by the Secretary of the Statistics and Informatics Division, declaring it as the national statistical methodology for poverty measurement in non-survey years

Countries and selected activities currently supported by UNSD

- Funded by Italian and Norwegian Governments
 - **Morocco:** SDG 11 (geo-stats information), CPI (web scraping)
 - **State of Palestine:** SDG 11 (geo-stats information)
 - **Jordan:** SDG 9, 11 (geo-stats information)
 - **Kyrgyzstan:** SD9, 11 (geo-stats information and citizen generated data)
 - **Tunisia:** Insights into internal migration patterns using mobile network data; CPI (web scraping and scanner data)

Countries and selected activities currently supported by UNSD

- Funded by Italian and Norwegian Governments
 - **Kenya:** Support on development of user engagement strategy
 - **Viet Nam:** Poverty in children; agriculture (small area estimation/geospatial information)
 - **Sierra Leone:** Marine litter data (citizen science), SDG 15, 16 (admin data)
 - **Ethiopia:** Poverty; agriculture (small area estimation/satellite imagery)

Cross-cutting themes

- Integration of **statistical and geospatial information**
- Modernization of **IT architecture/infrastructure** to effectively manage and analyze a diverse range of data types, encompassing both traditional and unconventional sources
- Strengthen capacity of national staff for improved and institutionalized **engagement with users**
- Development of the **conceptual framework for citizen data** and the roadmap for its implementation
- Development of a road map for **NSO-academia collaboration**



THANK YOU!

For resources, visit:

<https://unstats.un.org/UNSD/Website/capacity-development/data-for-now/>

Contact:

Faryal.ahmed@un.org

