



Food and Agriculture  
Organization of the  
United Nations



# Using Big Data to promote gender equality in agriculture

**Expert meeting on Statistics on Gender and the Environment  
2-4 September 2019, Vie Hotel, Bangkok, Thailand**

**Sangita Dubey**  
FAO Regional Statistician for Asia Pacific

# CONCLUSIONS

1. Big Data in Agriculture  $\approx$  EO Data

- **Earth Observation data** is the gathering of information about the physical, chemical, and biological systems of the planet via remote-sensing technologies, supplemented by **Earth**-surveying techniques, which encompasses the collection, analysis, and presentation of **data**. (Wikipedia)
- EO data *does not* collect information on demographics

2. Add gender to EO data via data integration

- Agriculture data with demographics (e.g. sex/gender) comes from surveys, censuses, and administrative data (private *and* public)
- Respondent matters
- Integrating with EO data requires data interoperability (consistent granular location identification)

Open (micro) data improves access/ use/ interoperability

- Agriculture data published is a fraction of data collected (usually in tabular form)
- Anonymized micro-data expands potential data use, as does open data

# BIG DATA DEFINITIONS

- Definition (Gartner): “*Big data* is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation.”
- Two additional V’s often added: veracity and value
- Big data brings in the private sector as a new source of information for official statistics

1. **Volume:** The amount of data, including high volume unstructured data (e.g. text, audio, video, twitter feeds, photos, clickstreams of web pages, sensor-enabled equipment, satellite images).
  2. **Velocity:** the rate at which data is received; often in real-time or near real-time. (Timeliness)
  3. **Variety:** types of data, including traditional data structured into relational databases; and unstructured data. Additional data processing often required.
- **Veracity:** extent to which data is accurate and reliable. (Accuracy/precision)
  - **Value:** ability to transform data into valuable analytics/evidence for decision making (Relevance)

# TYPES OF BIG DATA USED IN AGRICULTURE

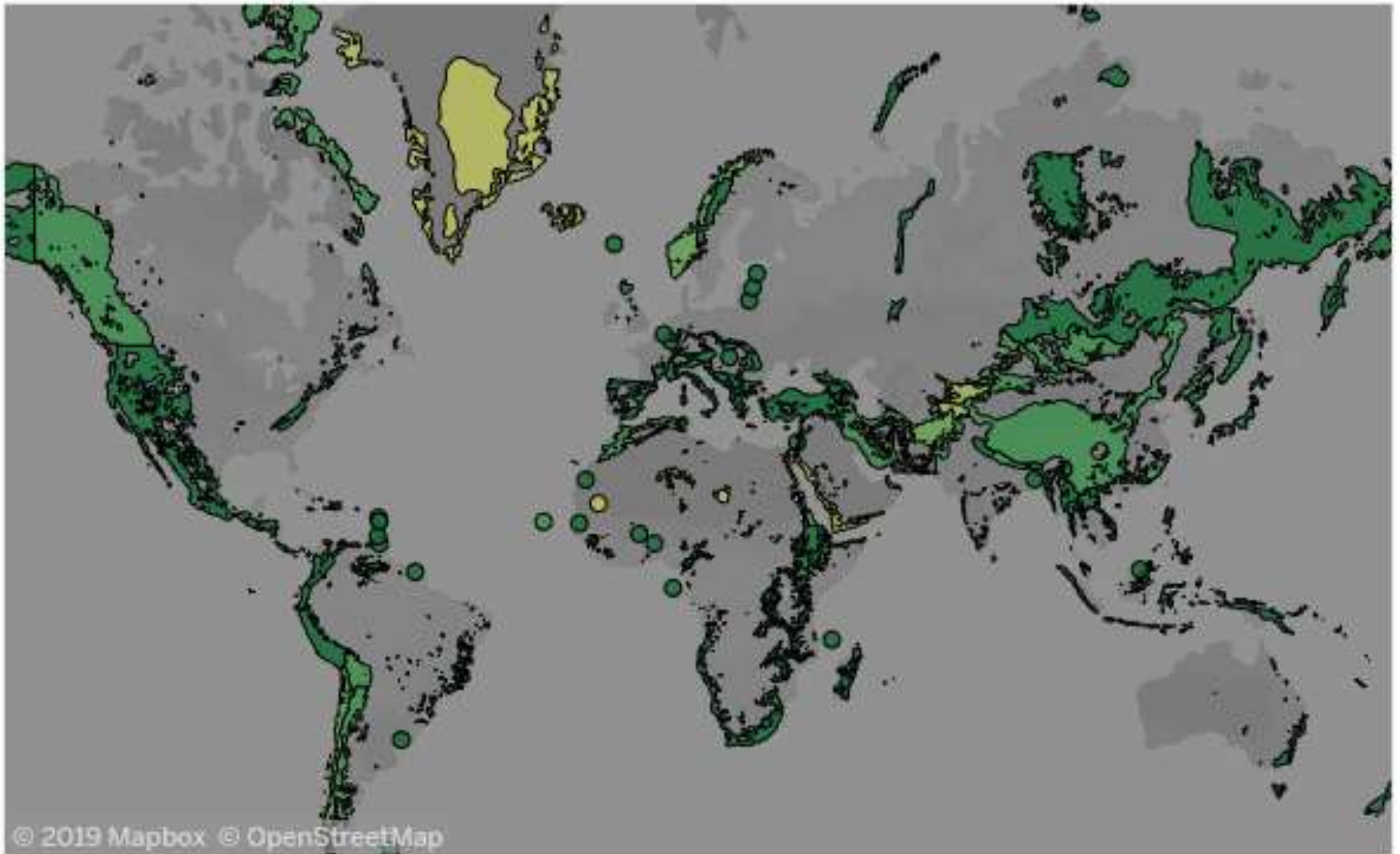
- **Earth Observation (EO) data:** *EO data is the gathering of information about the physical, chemical, and biological systems of the planet via remote-sensing technologies, supplemented by **Earth**-surveying techniques, which encompasses the collection, analysis, and presentation of **data**. (Wikipedia)*
  - This includes data from satellite images, drone images, GPS coordinates
- **Non-EO Big Data in Agriculture:** sensor data (on agriculture machines for soil measurement; on livestock for traceability); photos (for pest identification); mobile data (voice, text, etc.)

## KEY CHALLENGES IN DISSEMINATING/USING BIG DATA:

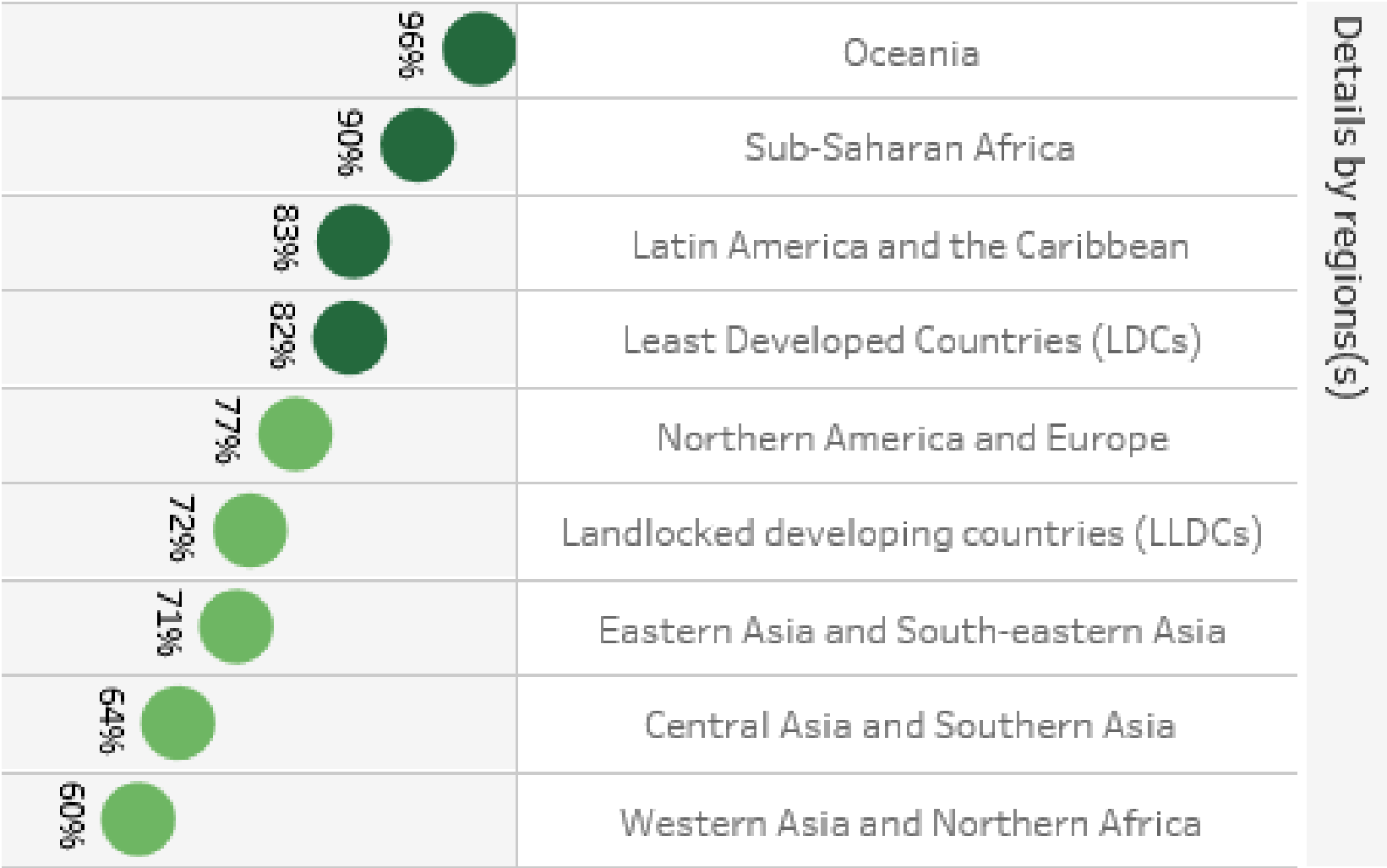
1. **Data Management** (storage, archiving, retrieval, access)
2. **Confidentiality/Privacy**
3. **National Security** (new Government players in data nexus)

# EO DATA: SDG 15.4.2 – MOUNTAIN GREEN COVER

Mountain Green Cover Index – baseline map (2017 values for all countries)



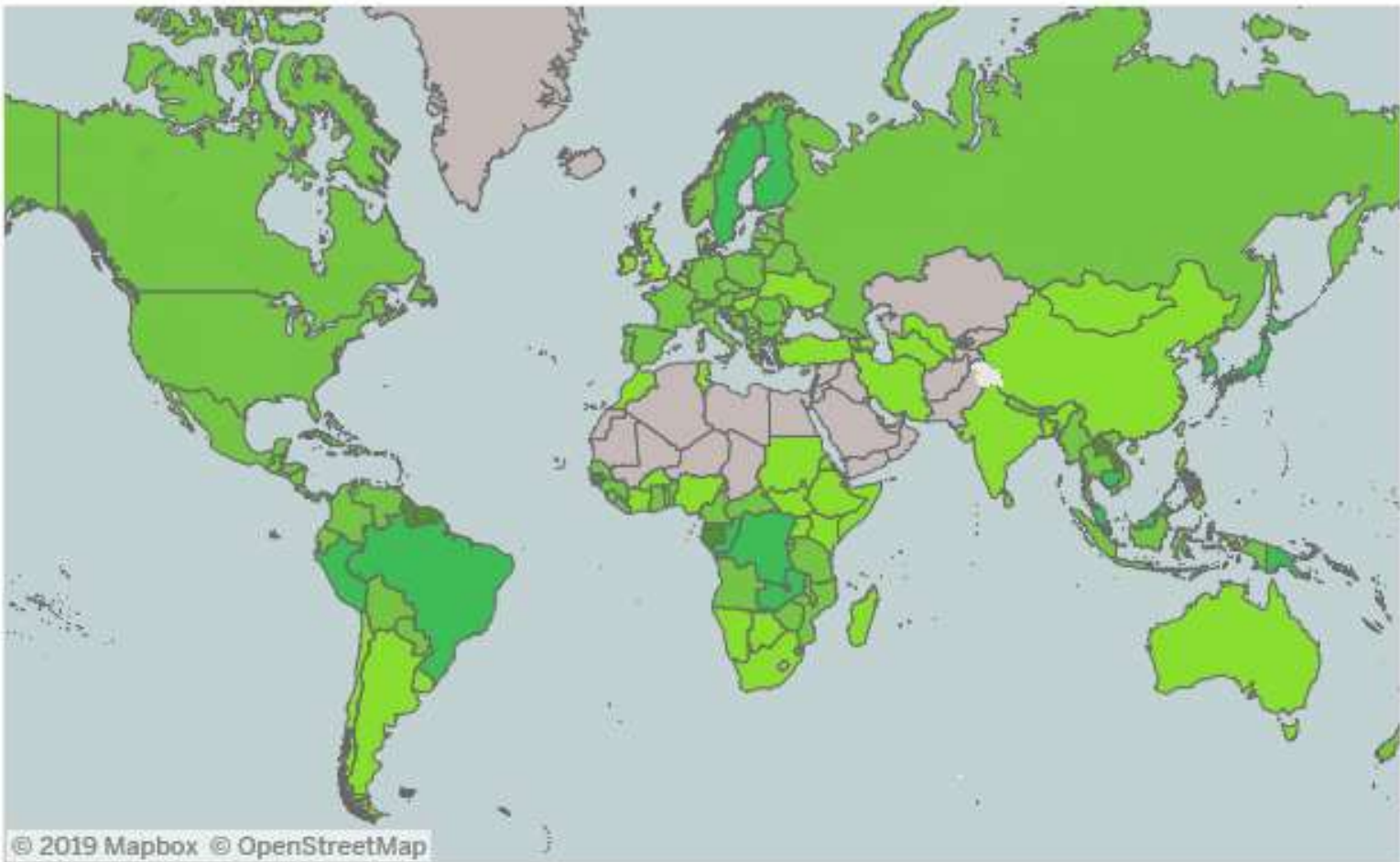
# EO BIG DATA & SDG 15.4.2 – MOUNTAIN GREEN COVER





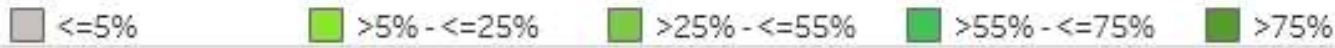
# EO DATA: SDG 15.1.1 – FOREST AREA AS % OF TOTAL LAND AREA

Forest area as a proportion of total land area (%) in 2010



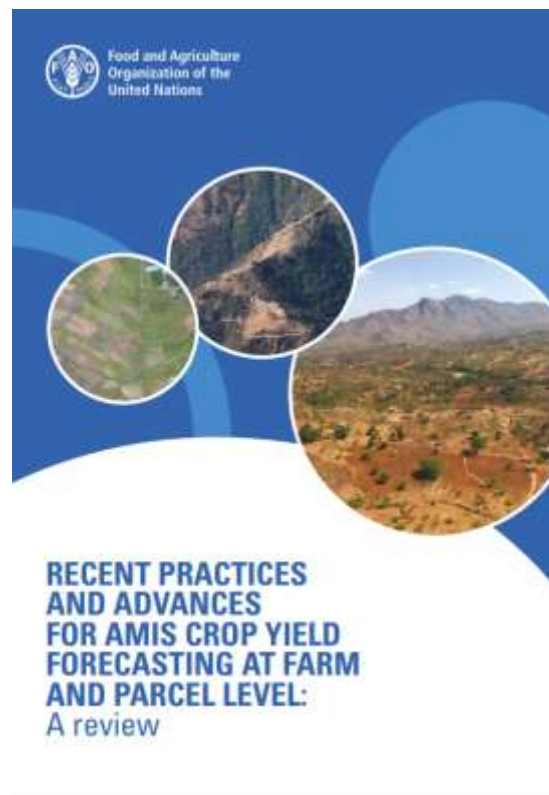
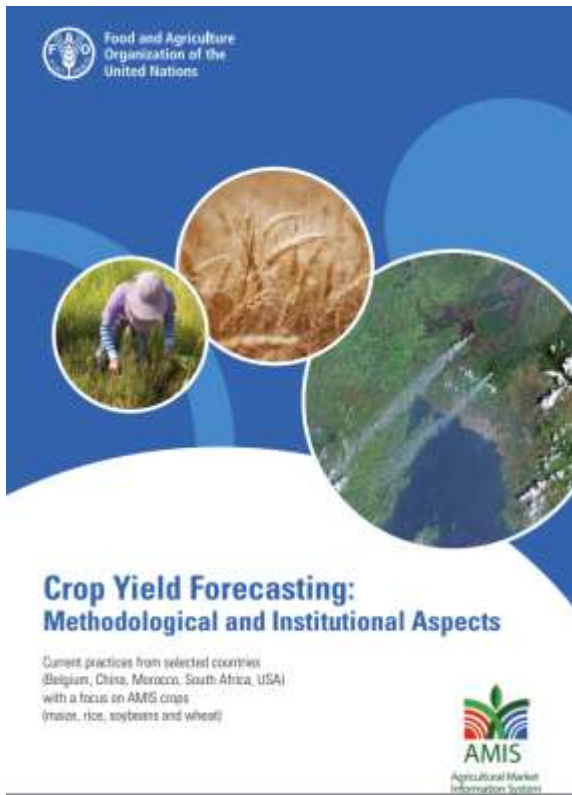
*The designations employed and the presentation of material in the maps do not imply the expression of any opinion whatsoever on the part of FAO concerning the legal or constitutional status of any country, territory or sea area, or concerning the delimitation of frontiers.*

Map Legend



# Other uses of EO data in Agriculture:

1. Crop yield forecasting; land area estimation
2. Data presentation





# GENDER IN THE SDG INDICATORS UNDER FAO CUSTODIANSHIP

Indicator	Gender statistics?	Respondent
<a href="#">2.1.1 Hunger</a>	Sex-disaggregatable	Individual traditionally; can be household
<a href="#">2.1.2 Severity of food insecurity</a>	Sex-disaggregatable	Household consumption (2nd best); rarely individual (1st best)
<a href="#">2.3.1 Productivity of small-scale food producers</a>	Sex-disaggregatable	Agriculture household
<a href="#">2.3.2 Income of small-scale food producer</a>	Sex-disaggregatable	Agriculture household
<a href="#">2.4.1 Agricultural sustainability</a>	Sex-disaggregatable	Agriculture household
<a href="#">2.5.1.a Conservation of plant genetic resources</a>	X	Gene Banks
<a href="#">2.5.1.b Conservation of animal genetic resources</a>	X	Gene Banks
<a href="#">2.5.2 Risk status of livestock breeds</a>	X	Measured in gene banks
<a href="#">2.a.1 Public Investment in agriculture</a>	X	Governments
<a href="#">2.c.1 Food price volatility</a>	X	Wholesalers (market prices); retailers (food CPI)
<a href="#">5.a.1 Women's ownership of agricultural land</a>	Gender specific	Agriculture Households
<a href="#">5.a.2 Women's equal rights to land ownership</a>	Gender specific	Government - Assessment of laws and policies
<a href="#">6.4.1 Water use efficiency</a>	X	Enterprises (ISIC sectors)
<a href="#">6.4.2 Water stress</a>	X	Enterprises (ISIC sectors)
<a href="#">12.3.1 Global food losses</a>	Sex-disaggregatable (?)	Agriculture households (harvest; early post harvest loss only)
<a href="#">14.4.1 Fish stocks sustainability</a>	X	Replaceability of marine fish stocks
<a href="#">14.6.1 Illegal, unreported unregulated fishing</a>	X	Government - compliance with international agreements
<a href="#">14.7.1 Value added of sustainable fisheries</a>	X	National Accounts
<a href="#">14.b.1 Access rights for small-scale fisheries</a>	X	Government - enabling policies, regulations, institutions
<a href="#">15.1.1 Forest area</a>	X	Big (EO) Data
<a href="#">15.2.1 Sustainable forest management</a>	X	Government for several of the 5 subindicators
<a href="#">15.4.2 Mountain Green Cover</a>	X	Big (EO) Data

# GENDER IN THE SDG INDICATORS UNDER FAO CUSTODIANSHIP

Indicator	Gender statistics?	Respondent
<a href="#">5.a.1 Women's ownership of agricultural land</a>	Gender specific	Agriculture Households
<a href="#">5.a.2 Women's equal rights to land ownership</a>	Gender specific	Government - Assessment of laws and policies
<a href="#">2.1.1 Hunger</a>	Sex-disaggregated	Individuals traditionally; can be household
<a href="#">2.1.2 Severity of food insecurity</a>	Sex-disaggregatable	Household consumption (2nd best); rarely individual (1st best)
<a href="#">2.3.1 Productivity of small-scale food producers</a>	Sex-disaggregatable	Agriculture household
<a href="#">2.3.2 Income of small-scale food producer</a>	Sex-disaggregatable	Agriculture household
<a href="#">2.4.1 Agricultural sustainability</a>	Sex-disaggregatable	Agriculture household
<a href="#">12.3.1 Global food losses</a>	Sex-disaggregatable (?)	Agriculture households (harvest; early post harvest loss only)
<a href="#">15.1.1 Forest area</a>	X	Big (EO) Data
<a href="#">15.4.2 Mountain Green Cover</a>	X	Big (EO) Data
<a href="#">6.4.1 Water use efficiency</a>	X	Enterprises (ISIC sectors)
<a href="#">6.4.2 Water stress</a>	X	Enterprises (ISIC sectors)
<a href="#">2.5.1.a Conservation of plant genetic resources</a>	X	Gene Banks
<a href="#">2.5.1.b Conservation of animal genetic resources</a>	X	Gene Banks
<a href="#">14.6.1 Illegal, unreported unregulated fishing</a>	X	Government - compliance with international agreements
<a href="#">14.b.1 Access rights for small-scale fisheries</a>	X	Government - enabling policies, regulations, institutions
<a href="#">15.2.1 Sustainable forest management</a>	X	Government for several of the 5 subindicators
<a href="#">2.a.1 Public Investment in agriculture</a>	X	Governments
<a href="#">2.5.2 Risk status of livestock breeds</a>	X	Measured in gene banks
<a href="#">14.7.1 Value added of sustainable fisheries</a>	X	National Accounts
<a href="#">14.4.1 Fish stocks sustainability</a>	X	Replaceability of marine fish stocks
<a href="#">2.c.1 Food price volatility</a>	X	Wholesalers (market prices); retailers (food CPI)

# QUESTION ON GENDER INDICATORS

- **Respondent = individual, gender statistics can be available.**
- **Respondent = household, how to sex-dissagregate?**
  - By head of household?
  - By inclusion of all household members? (costly)
  - Who collects data and from whom matters
- **Respondent = enterprise, how to sex-dissagregate?**
  - By owner/manager?
  - By proportion of female employees?
  - Who collects data and from whom matters

# INTEGRATING “TRADITIONAL” STATISTICS WITH EO DATA

- **Requirements for Integration/Interoperability:** Geographic coordinates (the more detailed the better)
  - Digital/CAPI data collection enables use of GPS coordinates
  - Processed satellite images with geo-political boundaries and infrastructure (roads, schools) may help answer:
    - Are subsistence producers farther from roads?
    - Are female-headed agriculture households more prevalent in disaster prone areas?
    - Where are sustainable farms?

## **KEY CHALLENGES IN DISSEMINATING/USING BIG DATA:**

- 1. Data Management** (storage, archiving, retrieval, access)
- 2. Confidentiality/Privacy**
- 3. National Security** (new Government players in data nexus)

# INTEGRATING “TRADITIONAL” STATISTICS WITH EO DATA

- **Tools to increase users ability to integrate Big Data with Gender or Sex-disaggregated Statistics**
  - **Open Data**
    - Legally open; Technically Open; Clear terms of use
    - Known to increase (free) research and data use, particularly if micro data are available
  - **Anonymized Microdata**
    - Mechanisms employed in anonymization to ensure data confidentiality/privacy and produce public use micro-data files (PUMFs)

# WHAT IS OPEN DATA ?

## Open Data **is Legally Open**

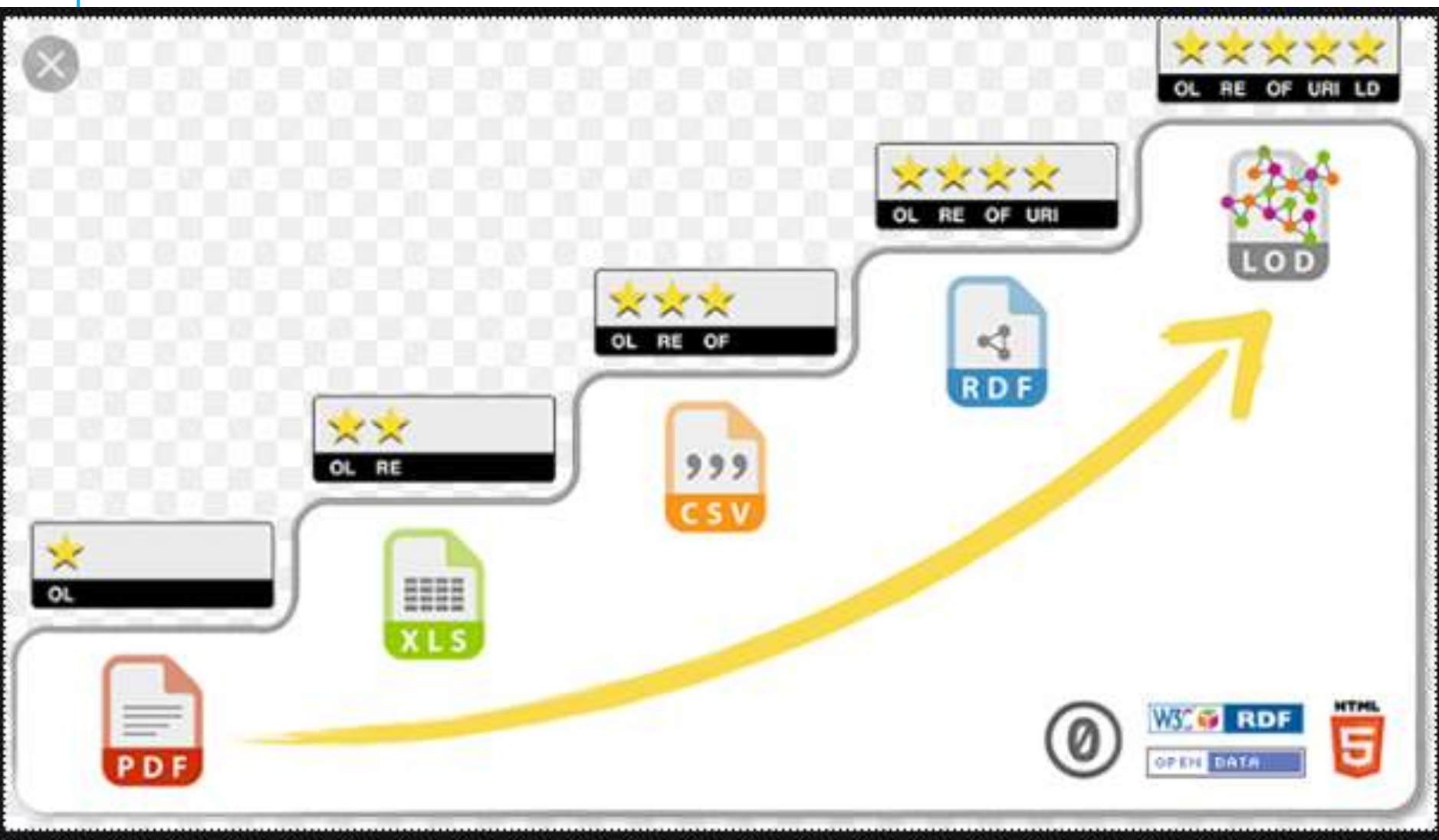
- Free to use; there can be share-alike and commercial restrictions

## Open Data **is Technically Open**

- Easy to find / searchable
- Machine-readable
- Downloadable in raw form / open formats
- Well documented (metadata)
- No registration or pay-walls
- Terms of use are clear
  - Protects producer from liabilities incurred in misuse



# WHAT IS TECHNICALLY OPEN?



# WHY IS OPEN DATA IMPORTANT?

Increases data use and value-addition:

- Enhances government transparency
- NASA, South African gold mines on value generated by users
- Data aggregators: Booking.com; Monster.com

Can be an effective data management/archiving tool

Protects users from mis-use

Examples of Open Data:

- OECD, World Bank, ISTAT, NASA, EO Data, International Aid Transparency Initiative (Aid data)



*THANK YOU!*