

Global Initiatives for accessing alternative data sources, processing and storage

Ralf Becker

United Nations Statistics Division

Regional Thematic Conference on Managing a
Statistical Organization in Times of Change
24-26 June 2024, Nadi, Fiji



Overview

- UNSD activities on:
 - Administrative data sources
 - Big data
 - The UN Regional & Global Hubs
 - Latest developments

Use of administrative data (1)

- Has been in discussion at UNSD for a long time
- No one-size-fits all solution for all countries and applications
- DA project is ongoing to assist some countries, but not in the Pacific region

Use of administrative data (2)

- A separate project “Collaborative on admin data” started in 2020 with GPSDD
- Outputs:
 - MoU guide and template
 - Self-assessment tool (legal framework)
 - Quality assessment of admin data sources (Toolkit)
 - Inventory of resources (108 items so far)



The aim of this inventory is to make materials on the use administrative data for statistical purposes more readily available and easy to find. The inventory includes a variety of materials, ranging from recommendations to practical examples in different contexts. Not all materials may be fully in line with international guidelines, so please take this into consideration.

To submit resources and use cases, please contact: statistics@un.org

Inventory listing

Select Filters

Thematic group
 Select thematic group

Cross cutting area
 dissemination

- coordination and collaboration
- data quality
- governance
- information technology
- interoperability
- linking
- other
- privacy and security
- standards and metadata

Sort By

1-6 of 12 items

6 / page



Eurostat
Guidelines on the use of estimation methods for the integration of administrative sources

Jan 27, 2021

International Guidelines

These guidelines on the use of estimation methods for the integration of administrative sources should help to:

- Promote the use of administrative data in the statistical production process;
- Promote the use of advanced methods in statistical data integration;

[Read full description](#)

data quality standards and metadata +3 more

Share via



Mexico
Public Sector Administrative Records - INEGI

Jan 14, 2023

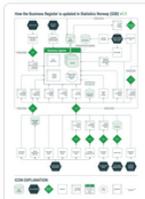
Case Study

Public sector administrative records are used for the generation of various statistical programs with economic information.

[Read full description](#)

coordination and collaboration data quality +14 more

Share via



Norway
Business register in Statistics Norway

Jan 17, 2023

Case Study

In this case study we try to explain how the business register in Statistics Norway is set up and how it's kept updated with various other administrative data sources.

[Read full description](#)

coordination and collaboration data quality +14 more

Big data

- Work on Big Data is now carried out by the UNCEBD
 - (formerly known as GWG on Big Data)
- Develop methodology
- Support applications
- Support development work through the UN Global Platform
- Capacity development



Big Data (2)

- Work is organized in Task Teams:
 - Access to Privately-held Data
 - AIS Data
 - Big Data and the SDGs
 - Earth Observation Data
 - Measuring Rural Access
 - Mobile Phone Data
 - Privacy-Enhancing Technologies
 - Scanner Data
 - Training, Competencies and Capacity Development

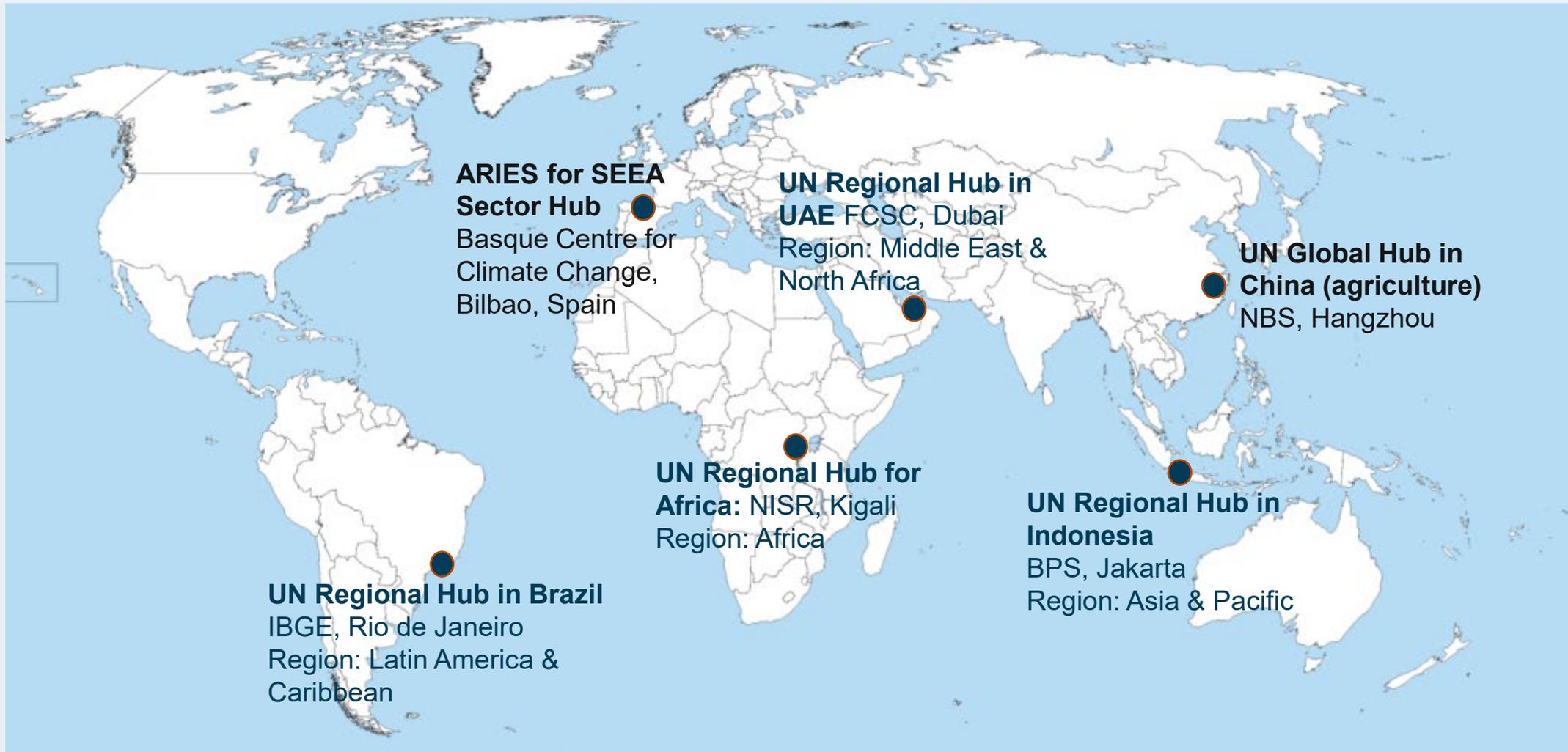
Big Data (3)

- Work on methodologies etc. is accompanied by training activities
- What are the most efficient options?
 - E.g. Country projects vs. Training courses
- Task Team on Training works with other task teams to develop curricula for their area and assist in developing training courses
 - Courses available

Big Data (4)

- Training and cooperation is managed in multiple ways:
 - Country projects
 - E-learning courses on our E-learning platform
 - Work via the Regional hubs

UN Regional & Global Hubs



UN Regional Hubs – what are they?

Identification of & focus on region specific big data needs

UNGP – Provides environment & shared platform for regional collaborations & developments

Focus on the training & development needs for the region – workshops, webinars, Newsletters, etc

Provide advice, consultancy & support role for NSOs in region

Collaboration, knowledge exchange & shared challenges – across borders

More than access & use of the UN Global Platform...

Strategic Working Areas

- Data Tech for People and Planet for Social Impact
- Knowledge Hub for the Mena Region
- Data Science Accelerator
- Creating the Largest Data Collaborative in the Mena Region



Action Areas

- Access to Data
- Access to Technology Services
- Access to Expertise
- Data Science & AI Practices

WORK STREAMS

Working with Government Sector

Support and enable government practices and

Strategic Partnerships with Private Sector

Working on building strategic partnerships with

Empowering Academics and Talents

Develop a knowledge center for data science

Engaging with Society and Youth

Exchange and transfer knowledge globally



GLOBAL PLATFORM
SNAPSHOT



TASK TEAMS



GOVERNANCE
FRAMEWORK



UPCOMING EVENTS &
PROJECTS

icon
PUBLICATIONS



AREAS OF
COLLABORATION

UNGP GLOBAL SNAPSHOT

What is the UN Global Platform?

Under the governance of the [UN Committee of Experts on Big Data and Data Science for Official Statistics \(UN-CEBD\)](#) the Global Platform has built a cloud-service ecosystem to support international collaboration in the development of Official Statistics using new data sources and innovative methods and to help countries measure the Sustainable Development Goals (SDGs) to deliver the 2030 Sustainable Development Agenda. The United Nations Global Platform was established as a collaborative environment to work together as a global statistical community and to learn together sharing knowledge, data and methods for all countries in the world. As the platform has matured and the value it delivers has increased, the demand for access to the tools, data and methods from Task

Big Data e-learning

- Courses are being developed by task teams, supported by the Task Team on Training (TTT)
- TTT has worked with others to develop curricula (for different audiences etc.) before actual courses are started
- So far, all are self-paced
- All courses are hosted on the UNSD/Big Data e-learning platform

UN Global Platform - Learning Management System

Learning Hub

UN Global Platform Learning Management System

Supporting a series of e-learning courses on various statistical and Big Data topics

Search courses Go 

Course categories

▼ Collapse all

▼ Big Data

- What is Big Data? ⁽¹⁾
- Automatic Identification System (AIS) ⁽²⁾
- Scanner data ⁽¹⁾
- Mobile phone data ⁽¹⁾
- Privacy preserving techniques ⁽³⁾

▸ Energy Statistics ⁽⁵⁾

▸ System of Environmental Economic Accounting (SEEA) ⁽¹⁶⁾

▸ Social and Demographics Statistics ⁽³⁾

▸ Sustainable Development Goals (SDGs) ⁽⁷⁾

Big Data (5)

- Many of the task teams and the Regional Hub need more support
 - For methodological work
 - For case studies
 - For training
- Countries in the region can join – just ask me for contacts

UN Regional Hubs - highlights

UN Regional Hub in **Indonesia** (established Nov 2023)

- Has developed dedicated [website](#)
- Hosted a workshop on remote sensing already (Nov 2023)

Priorities:

- Coordination across Asia and Pacific
- Research on use of big data
- Provide capacity building and training of statisticians and data scientists of NSOs
- Advocate for big data and data science in official statistics

UN Regional Hubs - Summary

- ✓ The UN Regional Hubs strengthen support for regional big data priorities

Forging strong relationships between the UN Regional Hubs:

- ✓ Seek synergies
- ✓ Share challenges, knowledge, expertise and resources
- ✓ Further development of the UN Global Platform
- ✓ Strengthening global statistical systems

New developments (1)

- Use of Scanner data for Consumer price index (CPI) has been priority for many countries, as well as web scraping for CPI
- Big Data Task Team on Scanner data now enlarges its work to include web scraping, covering “Alternative data sources for CPI”
 - New e-learning course already developed

New developments (2)

- The work of the task team will be linked to the work of the ISWGPS (Inter-secretariat Working Group on Price Statistics)
 - ISWGPS has started discussing alternative data sources, but not taken concrete actions yet
 - Coordination between both groups avoids overlap and conflict
 - Methods developed by the task team get better integrated in the overall approval process in the international statistical system



New developments (3)

- E-learning courses are being developed
 - AIS, Scanner data, PET are in progress
 - Mobile phone courses are being considered based on new handbooks
 - General Big Data courses will be expanded
- We also work on translations

- Home
- Dashboard
- Calendar
- Private files
- Content bank
- My courses
- Energy statistics
- Compiling Metadata for SDGs (English)
- MOOC Format
- Site administration

The automatic identification system (AIS) is a tracking system for ships, originally developed for collision avoidance. In the recent years, it is also used for analyses from various fields. The data is automatically transmitted every few seconds over very high frequency (VHF) radios from approximately 100,000 vessels worldwide.

If you have any questions regarding the AIS courses, please do not hesitate to contact us at support@officialstatistics.org.

Search courses Go ?

What is AIS Data?



This introductory course aims to raise awareness of what AIS data is and shows how it can be applied at a National Statistical Organisation (NSO). Information will be shared on how to obtain AIS data, plus the opportunities and challenges it presents for incorporating its use at the NSO.

Acquiring AIS data via the UNGP



This course aims to build the skills for acquiring and performing simple analysis of AIS data via the UN Global Platform. It includes an introduction to the UN Global Platform AIS data, procedures to request access to the platform, methods of acquiring both AIS and Ships Register data, an introduction to their data structures, and samples of reading and filtering AIS data using PySpark.

Note: Access to this course is restricted. If you are interested, send a request to support@officialstatistics.org.

Estimating shipping Greenhouse gases via the UNGP



This course aims to build the skills to generate emissions statistics, with a focus on the three main contributors of GHG from shipping (Carbon Dioxide, Methane and Nitrous Oxides) according to the International Maritime Organization. The approach of the proposed method, for this module, is to estimate emissions for vessels within proximity of a country of choice.

New developments in the region

- The Big Data Regional Hub for the Asia-Pacific region, managed by Indonesia plans to start a training programme
 - For countries in the region
 - Based on work being done in the UNCEBD task teams
 - Starting with introduction to Scanner (+), Earth observation, AIS and others
 - Extensions to more technical aspects



Thank you!

- More info:

- <https://unstats.un.org/UNSDWebsite/capacity-development/admin-data>

- <https://unstats.un.org/bigdata/>

- <https://learning.officialstatistics.org/>

- Contact: BeckerR@un.org

