Progress report principle 4 working group – Statistical and geospatial interoperability – data, standards, processes … and more?

UNECE/Eurostat

Steven VALE, Ekkehard PETRI, Eurostat

Relevant activities of the sub-group

- UNECE – UN-GGIM workshop on the integration of statistical and geospatial standards
- European Commission study on the integration of SDMX and INSPIRE for the 2021 Census
- Contribution to the revision of the GSBPM (GSIM …)
- GEOSTAT 3 project on EU/ESS – SGF
- …

Principle 4 - Statistical and geospatial interoperability – Data, Standards and Processes

Both the statistical and geospatial data communities operate their own general data models and metadata capabilities; however, often these are not universally applied. The statistical community use the Generic Statistical Information Model (GSIM), the Statistical Data and Metadata Exchange (SDMX), and Data Documentation Initiative (DDI) mechanisms. The geospatial community use the General Feature Model (GFM) and developed the ISO19115 metadata standard, plus a number of application specific standards.
The Expert Group recognises the benefits of greater interoperability statistical and geospatial data and metadata standards, from cataloguing to data interchange. Overcoming structural and syntactic barriers between data and metadata from different communities and providers will enhance the efficiency of discovery, access, and use of geospatially enabled data. …
Within the statistical community there is a need to build geospatial processes and standards into statistical business processes in a more consistent manner. The Expert Group has recognised that, to ensure this occurs, a top down approach of incorporating geospatial frameworks, standards and processes more explicitly into the CSPA and its components is required. In particular, GSBPM needs to make greater reference to the use of geospatial data and methods in the statistical production process, particularly the data, standards and methods that are incorporated into the Global Statistical Geospatial Framework…

Relevant areas of activity

- Technical standards
  - ISO, SDMX, DDI, DCAT-AP, StatDCAT-AP, GeoDCAT-AP, LOD
- Frameworks
  - GAMSO, CSPA, GSIM, GSBPM
- Data, metadata, processes

Purpose of Principle 4

- *"Greater efficiency and simplification of the creation, discovery, integration and use of geospatially enabled statistics and geospatial data."*
- *"Increasing the potential application of a larger range of data and technologies."*
- *"A wider range of data available and accessible for use in comparisons and analysis in decision making."*

5

Eurostat

Where do we need interoperability?

- Interoperability of data is relevant during almost all stages of the statistical production process.
- Interoperability according to principle should therefore cover:
  - Producer related interoperability
  - (Intermediate and end-) User related interoperability
- Producers mostly have technical interoperability issues.
- Users have technical and non-technical interoperability issues.

6

Eurostat

Interoperability barriers and enablers

Technical interoperability

- Exchange interoperability
  - Data exchange (file formats, availability of services)
  - Data structure (syntax, encodings)
- Semantic interoperability
  - Common understanding of concepts
  - Metadata

Technical interoperability

- Process interoperability
  - Harmonisation of production processes
- Temporal interoperability
  - Consistency of statistics of geographies and statistics
- Tool interoperability
  - Same results if different tools used?

Technical interoperability – special - spatial

- Projections etc.
- Information available for the same geographical detail NUTS level (need for aggregation or disaggregation)
- Geocoding interoperability (objects to be geocoded)

Non-technical interoperability

- Legal interoperability (legal barriers, licenses, …)
- Organisational interoperability/cross-domain interoperability (network of contacts, clear responsibilities, organisation for cooperation)

Human interoperability

- ***Human interoperability – Terminology***

- Terminology (definitions and jargon) http://www.efgs.info/information-base/introduction/terminology/
- General understanding, culture, ways we work and are organised

Available input

GEOSTAT 3 project interoperability recommendations

- Leave data at its source
- Publish only once
- Data integration based on existing national statistical dissemination platforms
- Use of Linked Open Data

OGC activities

- Geocoding API working group (under creating)
- Discrete Global Grid System Standards Working Group

UNECE UN-GGIM workshop

- Communication and capacity building
- Projects pitches
- Guidelines (e.g. on metadata?)
- Other follow-up activities
- …

POSSIBLE NEXT STEPS

- All work needs to be synchronised with follow-up of Standards workshop and the other activities.
- More detailed formulation of generic interoperability principles, and a generic action plan covering all aspects
- Select most relevant areas and start synthesising recommendations from existing and forthcoming material (GEOSTAT, UNECE follow-up, ModernStats)
- In concrete terms integrate principle 1 into you national process model

For this meeting

- Structure and scope of principle 4 (more detailed document, repository of best-practices, specific recommendations???)
- Define working method for terminology
- Formulate expectations towards GEOSTAT 3
- Assign tasks to possible contributors: ISO TC211, OGC, SDMX group, UNECE, UNSD, EU/Eurostat and UN-GGIM/ Member States.