Panel Session A: 
Integrating Location in Statistical Production

Marie Haldorson
Director, Statistics Sweden
Agenda

- Introduction
- The Statistical Geospatial Framework
- GEOSTAT 2 – Testing GSBPM
- Need for more work on standards!
Introduction
Integrating statistical and geospatial information

1. Use of fundamental geospatial infrastructure and geocoding
2. Geocoded unit record data in a data management environment
3. Common geographies for dissemination of statistics
4. Interoperable data & metadata standards
5. Accessible & Usable
Principle 4:
Interoperable data and metadata standards

**Statistical**
- GSIM
- GSBPM
- SDMX
- DDI

**Geospatial**
- General Feature Model (GFM)
- ISO 19115 Metadata standard
- Application specific standards
Contents of GEOSTAT 2
Production of Spatial Statistics: An applied sketch of GSBPM

General Statistical Business Process Model: http://www1.unice.org/stat/platform/display/GBPM/GBPM+v5.0

Terms used:
- Geospatial data: Data with direct reference to a specific location on the surface of the Earth (points, areas, lines)
- Spatial statistics: Geospatial data with statistics or table data with location information.

**Quality Management / Metadata Management**

<table>
<thead>
<tr>
<th>Specify Needs</th>
<th>Design</th>
<th>Build</th>
<th>Collect</th>
<th>Process</th>
<th>Analyse</th>
<th>Disseminate</th>
<th>Evaluate</th>
</tr>
</thead>
</table>

**Geospatial data**

1-3 Plan the process

4 Collect geospatial data or other data with spatial information

5 Edit geospatial data, create spatiality

6 Study data quality, prepare and finalise outputs

7 Store, warehouse data, disseminate geospatial data products

8 Conduct evaluation

**INTEGRATION OF GEOSPATIAL DATA AND STATISTICS, e.g.**

Integrate by unique identifiers or other links

Make derived classifications

Create new information, prepare and finalise outputs

Produce WFS/WMS, thematic maps, tables, graphs

Evaluate usability

**Statistics**

1-3 Plan the process

4 Collect statistical source data with or without spatiality

5 Classify, code, calculate aggregates

6 Create compilation of statistics, prepare and finalise outputs

7 Store, warehouse data, disseminate spatial statistics

8 Conduct evaluation
Design

Geographical survey frames: Adding coordinates to households and businesses allows more flexible surveys.
Process

Screening with GIS tools: Identify outliers in geography or detect geographical patterns in the collected data.
Production of Spatial Statistics: An applied sketch of GSBPM

General Statistical Business Process Model: [http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0](http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0)

Terms used:
- **Geospatial data**: Data with direct reference to a specific location on the surface of the Earth (points, areas, lines).
- **Spatial statistics**: Geospatial data with statistics or table data with location information.

**Quality Management / Metadata Management**

**Specifying Needs**

**Delivery**

**Integrate Geospatial Data and Statistics**

**Statistics**

**Plan the processes**

**Disseminate**

**Evaluate**

**7 Store, warehouse data, disseminate geospatial data products**

**Conduct evaluation**

**Produce WFS/WMS, thematic maps, tables, graphs**

**Evaluate usability**

**ESS Eurogrid Population Map 2009**

Population by 1 square km grids. A hybrid grid map based on population register and estimated data.

**Disseminate**

| Inhabitants or estimated number of inhabitants per individual square kilometre |
|---|---|---|---|---|
| 1 - 4 | 5 - 19 | 20 - 199 | 200 - 999 | 1000 - 4999 | 5000 - 10000 | Unreported |

**Produced for ESS Coordination Committee in Luxembourg, May 2009**

[Map Description](http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0)

**Notes:**
- The map uses a hybrid grid approach.
- It integrates population register data with estimated data.
- The color coding indicates population density levels.
Story mapping

- The GEOSTAT 2 team identifying tasks when producing grid statistics!
Lessons learned so far (SE)

- Process modeling show that GSBPM is useful for processes involving geospatial data.
- GEOSTAT 2 has enabled increased collaboration between process owners and geospatial experts at Statistics Sweden.
- The GSBPM should include more concepts and a terminology that links to geospatial activities.
- The GSBPM could also benefit from including examples of ”Geospatial statistical output”!
Lessons learned so far (FR)

- The GSBPM is useful.
- The GSBPM should include specific processes or sub-processes dedicated to the integration of geospatial data in statistics.
- Process needed to maintain, update, ensure consistency of authoritative geospatial data and statistics over time?
- Managing confidentiality is one of the big issues to manage geospatial statistics, it might need a specific process?
- Specific attention should be paid on terminology!
Need for more work on standards!

- Include descriptions and terminology for **integration** of statistical and geospatial data in the GSBPM!
- Describe geospatial processes?
- Principle 4 in the Statistical Geospatial Framework: Initiatives that can advance interoperability?
There is a distinction between:

- Geospatially enabled data or statistics - statistical or administrative data linked to a geospatial object
- Geospatial infrastructure data used to geospatially enable statistics
- Geospatial datasets used to create statistical content

Suggested terminology currently up for discussion (Australian Bureau of Statistics)
Geospatially enabled data

- **Fine level** - unit record data linked to a coordinate, small linear object or small area geography building block
- **Mid level** - unit record data linked to a large geographic unit or aggregate data linked to a small area geography building block
- **Course level** - aggregate data linked to any medium or large geographic unit
Geospatially enabled data
Geospatial infrastructure data

- Address or Building registers with coordinate references
- Parcel or property databases
- Topographic datasets
- Administrative, statistical and grid geographies
Geospatial infrastructure
Geospatial datasets

- Datasets used to create statistical content (geospatial statistics)
- To enhance (i.e. add variables to) existing statistical and administrative datasets (e.g. distance to green space: road network, parklands, etc.)
- More elaborate production of new statistical content (e.g. land accounts: cadastral information - land use and value, gridded land cover from Earth Observation, etc.)
Geospatial datasets