GSIM Task Team Report

Francine Kalonji (Statistics Canada)

Delivering insight through data, for a better Canada
What is GSIM?

• The Generic Statistical Information Model (GSIM) is a reference framework for statistical information and provides a set of standardised, consistently described information objects.

• It provides an extensive coverage of statistical information objects and is used across several statistical agencies around the world.

• It also provides standardised inputs and outputs to the Generic Statistical Business Process Model (GSBPM), a reference framework for statistical processes.

• Current version is 1.2
Why The Task Team?

• As more and more agencies use GSIM, they uncover and report inconsistencies, ambiguities, errors, etc. that arise from:
  ➢ challenges in implementing the model
  ➢ different interpretations of the model
• The Metadata Glossary task teams provided feedback for improving the model and making it consistent with the other Modern Stats Models: GAMSO (Generic Activity Model for Statistical Organizations), GSBPM, and CSPA (Common Statistical Production Architecture) documentation
• Also, the Linking GSIM and GSBPM task team has flagged some concerns
• To address all these issues, there is a need for a dedicated team hence this task team
What is our direction?

• By April 2020, statistical organisations that are part of the GSIM community voted on four suggested work areas:

1. “Soft updates” of GSIM to clarify ambiguous points and fix minor errors
   ➢ more significant errors need to be compiled for the next GSIM revision
   ➢ update it to reflect changes made in GSIM v1.2 and recent experience
3. Review of other GSIM supporting documents
   ➢ e.g. Specification, Communication
4. Extension of GSIM Group(s)
   ➢ Structure Group, Business Group
What is our agreed workload?

- Updates with inputs from Supporting Standards Group task teams
  - Metadata Glossary task team
  - Linking GSBPM and GSIM task team
- Updates from issues raised by statistical agencies in the discussion forum
  - Prioritised list
- Updates on the communication paper to align with current version of GSIM, v1.2
  - Essential, as it describes why organisations want to use GSIM
- Updates on the GSIM User Guide and User Specification
  - As they will have to align with the version that will result from the task team work
How do we proceed?

- Conference calls through WebEx
  - chaired by Francine Kalonji and Flavio Rizzolo, both from Statistics Canada
  - first meeting on June 26th, 2020
  - frequency: every three weeks, but adjusted as needed
  - 22 members have joined from Australia, Canada, Finland, France, Hungary, Italy, Mexico, Norway and USA

- Issues are compiled on the UNECE Stats Wiki by Inkyung Choi (UNECE secretariat)
  - Based on agreed priorities, and shared with the team

- Team decides on which issue to tackle first

- Discussions incur, thoughts are exchanged, members are consulted for a collective decision, the outcome is logged, action items are assigned when required

- Some interventions happen offline via email if the discussion is too long for the meeting
  - Outcomes are always discussed during meetings
<table>
<thead>
<tr>
<th>Country</th>
<th>Member</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>Francine Kalonji – Chair</td>
</tr>
<tr>
<td></td>
<td>Flavio Rizzolo - Chair</td>
</tr>
<tr>
<td></td>
<td>Farrah Sanjari</td>
</tr>
<tr>
<td>Italy</td>
<td>Giorgia Simeoni</td>
</tr>
<tr>
<td></td>
<td>Mauro Scanu</td>
</tr>
<tr>
<td></td>
<td>Laura Tosco</td>
</tr>
<tr>
<td></td>
<td>Orietta Luzi</td>
</tr>
<tr>
<td></td>
<td>Cecilia Casagrande</td>
</tr>
<tr>
<td></td>
<td>Marina Signore (observer)</td>
</tr>
<tr>
<td>Mexico</td>
<td>Juan Munoz</td>
</tr>
<tr>
<td>Norway</td>
<td>Jenny Linnerud</td>
</tr>
<tr>
<td>Hungary</td>
<td>Zoltan Verczkei</td>
</tr>
<tr>
<td></td>
<td>Csaba Ábry</td>
</tr>
<tr>
<td></td>
<td>Józsefné Ferencz</td>
</tr>
<tr>
<td></td>
<td>Éva Ender Hajósné</td>
</tr>
<tr>
<td></td>
<td>Andrea Petres</td>
</tr>
<tr>
<td>Finland</td>
<td>Essi Kaukonen</td>
</tr>
<tr>
<td></td>
<td>Mikko Saloila</td>
</tr>
<tr>
<td>Australia</td>
<td>Alistair Hamilton</td>
</tr>
<tr>
<td>France</td>
<td>Florian Vucko</td>
</tr>
<tr>
<td>USA</td>
<td>Dan Gillman</td>
</tr>
<tr>
<td></td>
<td>Christopher N Carrino</td>
</tr>
</tbody>
</table>
Communication paper

• Work led by Inkyung Choi, Jenny Linnerud (Statistics Norway) and Juan Muñoz (INEGI, Mexico)
• Based on version 1.1, it was updated to align with version 1.2
• Reviewed by the team
  ➢ Garnered very good and comprehensive feedback
• Input was discussed during the August 20th, 2020 meeting as well as offline
• Updated and published on the wiki at the beginning of October 2020
Changes from Metadata Glossary task team

- Advised to follow convention for definitions
  - UK English spelling
  - No leading articles, be able to substitute grammatically for the term, start with lowercase and no ending dot, doesn’t start with concept being defined

- Proposed changes to some definitions and/or explanatory texts, where it was unclear
  - Cases where definition was too long and did not follow the convention, part was moved to explanatory text

- Identified terms for modifications when proposals were not possible
  - This will require more work on the part of the task team to come up with meaningful definitions
  - Concrete examples were also suggested to clarify some concepts

- Review started with the Base Group
Changes from Linking GSBPM and GSIM task team

• Misinterpretations of GSIM definitions were also reported by this task team
• Started with Business Group
  ➢ e.g. In mapping Information Objects to Transformable Input, one element in the definition led to inconsistent mappings of information objects by different countries.
    ❖ A type of Process Input whose content goes into a Process Step and is changed in some way by the execution of that Process Step.
  ➢ Proposition to change the Transformable Input to Main Input
  ➢ Proposition to change its definition
    ❖ Main Input: main Process Input used to produce a Transformed Output where some or all its content may be changed by the execution.
• Details on other reported information objects can be found on the wiki
Impact

• Changes must be looked at as whole because of the connection between information objects
  ➢ To ensure the meaning/harmonisation of connected objects is not altered

• Impact on GSIM implementations
  ➢ Change in names of the information objects
  ➢ Change in definitions (the mapping may have to be altered)
    ❖ It is the case of the work in Linking GSBPM and GSIM
  ➢ Structural changes (for this revision, they are done, only if absolutely necessary)

• Aim for minimal impact, for this revision
  ➢ Whenever possible

• Input from countries effected by the modifications is very encouraged
Useful Links

- CSPA: https://statswiki.unece.org/display/CSPA
- GAMSO: https://statswiki.unece.org/display/GAMSO
- GSBPM: https://statswiki.unece.org/display/GSBPM
- GSIM Communication Paper: https://statswiki.unece.org/display/gsim/GSIM+v1.2+Communication+Paper
- GSIM Task Team: https://statswiki.unece.org/display/Standards/GSIM%3A+Home
- GSIM: https://statswiki.unece.org/display/gsim
- Metadata Glossary: https://statswiki.unece.org/display/hlgbas/Statistical+Metadata+Glossary
- Stats Wiki: https://statswiki.unece.org/display/Standards/GSIM%3A+Home