National adaptation of the GSIM at the Hungarian Central Statistical Office

Ildikó Györki & Csaba Ábry & Zoltán Vereczkei
Methodology Department
Architectural Services Directorate
Hungarian Central Statistical Office

zoltan.vereczkei@ksh.hu
modszertan@ksh.hu

UNECE ModernStats World Workshop 2019
Geneva, 26th-28th June 2019
Outline

- GSIM adaptation in the HCSO: motivation, state-of-the-art
- Challenges (so far) with GSIM adaptation
- Way forward with adaptation
GSIM adaptation in Hungary

- Adaptation of the ModernStats models started with GSBPM (as in most countries)
- Hungary is dedicated to implement and use ModernStats models in practice („Statement of Intent”)
- Presentation is focusing more on challenges (problems) with the GSIM adaptation (we are big supporters of the model!)
- Lessons learnt & challenges with GSIM adaptation
GSIM adaptation – starting setup

• Hungary has been developing and using metadata-driven integrated IT systems since the 1970s.
• This resulted in the situation that our own information architecture has been in place since that time.
• The backbone of those systems are present in the form of our integrated Metainformation system, standard database structure, naming conventions, etc.
Main subsystems of the metainformation system

- Data sources
  - Data collections
  - Data transfers in the National Statistical System
  - Data transfers within the HCSO
  - Registers

- Statistical domains (SzOR)
  - Methodological documentation of statistical domains

- Metadata of data warehouse
  - Concepts
  - Variables
  - Legal acts
  - Nomenclatures

- Metadata of database
- Metadata of data collection
- Metadata of data preparation
- Metadata of processing data

GSIM adaptation – the challenge

- As Hungary already has a mature information model in place, the main challenges is (how) to map the existing models to GSIM and to make sure that the „GSIM-ised” version of the Hungarian model is suitable for gap analysis against the GSIM model
- GSIM to be adapted on 3 levels:
  - Conceptual level
  - Business level
  - Technical level
Challenge 1: Conceptual framework of GSIM #1

- GSIM is meant to be a common language for countries (statisticians) using the model. Common language should have common vocabularies (concepts).
- GSIM introduced new concepts that have not been used before in official statistics in such frameworks as the GSBPM (difficulties to link the two frameworks).
  - Examples: „statistical program” during „Specify needs” and „Design” is nowhere to be found in GSBPM [GSBPM operates with terms „statistics” or „statistical outputs, products and services”]
- GSIM is basically using different concepts as used before GSIM: unique challenge for countries that built information systems before the GSIM
Challenge 1: Conceptual framework of GSIM #2

- In order to be a common language, the connection of the concepts used by GSIM to other concepts used should be given and should be clear (many of these are missing in the current version).
  - Example #1: "node". Readers of GSIM should check the exact types of term "node" to have an approximation on how they are linked to other concepts: GSIM does not give the direct link, the user needs to work on the interpretation to link the concepts on their own.
  - Example #2: "represented variable": clear as a definition but when you go into the details (examples given) it will be unsure to the reader whether he/she understood it right in the first place. Example in the GSIM description: "industry": in our understanding it is more a node type but not an example for a variable?!
Challenge 2a: GSIM model framework

Different approach???

GSIM 1.0 – Information object groups
Challenge 2b: GSIM model framework

- Interoperability of GSBPM and GSIM is now in the highlight of the work within the HLG.
- There are some anomalies in the GSIM where more alignment to the GSBPM should be made
  - Example: sometimes when the sub-processes of the GSBPM are referred to in GSIM, the sub-processes are grouped, structure is a bit different and sometimes and process steps are also different from GSBPM. Is this a problem?
Challenge 3: How to start with GSIM adaptation?

• Any mapping activity against GSIM can start only when the current information architecture of the organisation is described according to the GSIM logic, but using the current practices on objects and their connections (current state of the Hungarian adaptation).

• Dilemma: what is the needed level of preparing this mapping? (current state of the Hungarian adaptation: entity level) We are still not preparing this mapping on attribute or on UML level (problems solved at deeper levels?!).

• Compromise: describe the current state-of-the-art according to our current practice but according to the logic of the GSIM (challenge).
Challenge 4: Level of detail

• Sometimes our national information model is deeper and more integrated than GSIM but there are also some elements completely missing from the Hungarian practice from the GSIM.

• Example: we have thesaurus-like system (connections between concept and type of connection: synonym, broader, narrower, other) of describing concepts but GSIM has different approaches (providing list)?
Next steps (for us)

- Complete the current description of our information architecture according to the GSIM
- Mapping of the GSIM model to the Hungarian „GSIM-ised” description
- Result of the gap analysis: actions
- When do we say it is adopted? What does it really mean?
Use of GSIM (2018)

(Ultimate) Challenge 4: Using GSIM

- What does it mean that countries use GSIM already?
- How did you start and go through GSIM implementation?
- What are the differences between adaptation practices between countries with and without already existing information architecture elements?
Thank you for your attention!

Ildikó Györki & Csaba Ábry & Zoltán Vereczkei
Methodology Department
Architectural Services Directorate
Hungarian Central Statistical Office

zoltan.verezkei@ksh.hu
modszertan@ksh.hu