Meeting of the Group of Experts on Business Registers organized jointly by UNECE, Eurostat and OECD. 21-23 September 2015

Session No. IV:
"The role of statistical business registers in the modernization of the statistical production and services – GSBPM, GSIM, data warehouse, use of new data sources, including big data"

Lessons learned from implementing GSBPM.
Statistics Estonia, Svetlana Shutova

In 2014 Eurostat launched the grant named "Improvement of national business registers (NBR) and testing of European profiling". The scope of the Topic 2 of the grant is "Describing the NBR with Generic Statistical Business Process Model (GSBPM) and Generic Statistical Information Model (GSIM)".

Statistics Estonia is participating in the project and this paper presents the lessons learned from implementing GSBPM in business register domain.

GSBPM is a reference framework to define and describe statistical processes in a coherent way and to compare them within and between different organisations.

GSIM is a reference framework used to describe the information objects of a statistical business process. It serves to have generic descriptions of all data and metadata that are part of a statistical production process, including their definition, management and use. With GSIM, information objects that constitute the input and output of the statistical process are described according to standards.

GSBPM was developed by the UNECE and the Conference of European Statisticians Steering Group on Statistical Metadata (METIS). First release was in April 2009 and since this time many statistical organisations adopted this in their activity. But the experience in application of this system to business registers is quite new.

Use of standard models facilitates comparison of different processes in different organizations. This helps to identify good practices and areas where efficiency can be gained and costs saved. The latter can be achieved either by adopting common tools, or sharing the costs of developing new tools or methods between several processes or organisations.

Expected results of the project are following:

1. Build up the capacity for using GSBPM and GSIM to document the NBR processes,
2. Describe the NBR process by using GSBPM and GSIM.

According to the grant technical description member states are free to decide on which level of GSBPM the description will be done. However it should be considered that grants shall be used to build capacity, create knowledge and allow for future re-use of the results.

For Statistics Estonia the first experience with GSBPM was in 2008 when the implementation of the system for tracking the working hours and activities was started. In that system all our working hours had to be allocated to the processes and sub-processes described in GSBPM. This was the first time when we found that it was a challenge to match the business register activities with the GSBPM.

Since 1st of October 2013 Statistics Estonia implemented the new organisational structure which is based on GSBPM. Thus, formerly separate domain-based departments were re-organised into four departments, based on processes of the production of statistics — data collection and processing and three domain-based analysing departments. And the new question arose — what is the best place for the register in this organisational structure?

Organisational chart of Statistics Estonia.
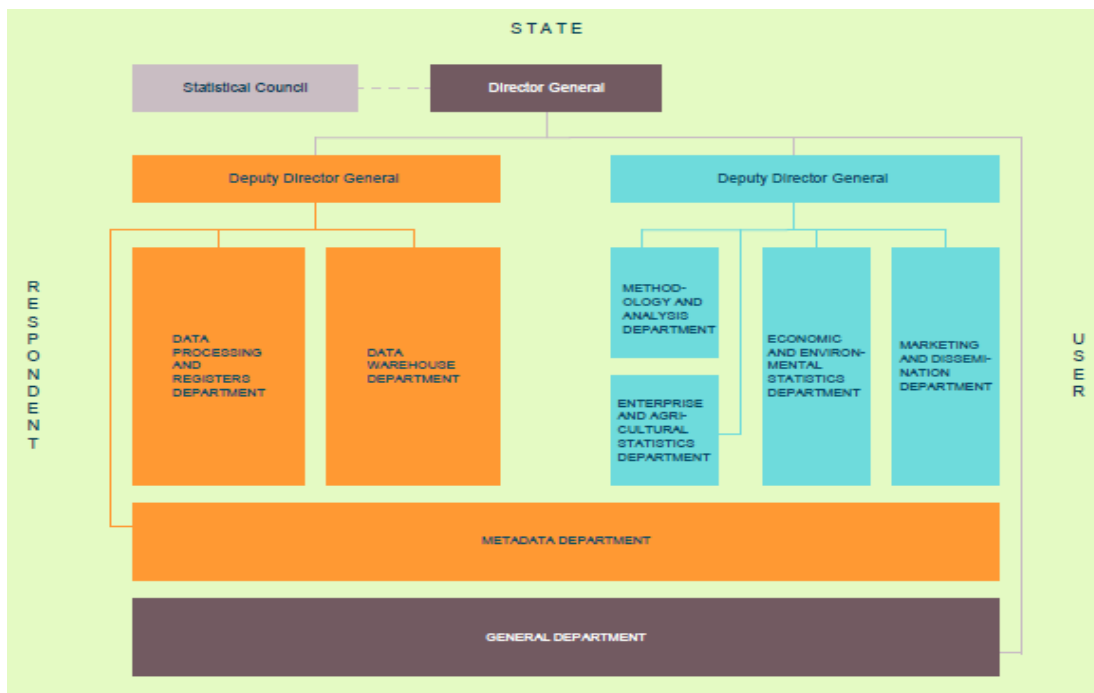
Four departments which have gone through the biggest changes are Enterprise and Agricultural Statistics Department, Economic and Environmental Statistics Department, Methodology and Analysis Department (which also deals with social statistics) and Data Processing and Registers Department. First three departments deal mainly with anaysis of statistics on aggregated level. Micro-data processing function was centralised under the new department — Data Processing and Registers. Currently Business Register is a part of Data Processing and Registers Department. GSBPM is used as a basis for main processes and IT solutions in the implementation of new organisational structure.

Responsibilities of different departments according to the GSBPM processes and sub-processes.

According to this main responsibility of our department is data collection, design of the data processing and the data processing itself.

One reason to participate in aforementioned Eurostat's grant was to understand how the business register processes can be mapped with this allocation. During this project we were hoping to improve our knowledge about GSBPM, to try to describe our everyday work activities and match our work with other processes of producing statistics.

The training course on GSBPM and GSIM was organised within the framework of the project. Mr. Steven Vale, Chief - Statistical Management and Modernisation Unit, Statistical Division, United Nations Economic Commission for Europe was invited as a trainer. The training course took place in December 2014 in Tallinn. Specialists from statistical register, metadata and data warehouse department and enterprise statistics participated in the training. The theory and practical use of GSBPM and GSIM for describing the processes of production of statistics and maintenance of the register were discussed. This was followed by practical exercises.

The main questions which arose before beginning the description phase were:

- Which register processes should be described?
- Which level of detail should be chosen to achieve better comparability with other member states?
- Which timeframes/periods should be covered?

The procedure of selection of the processes which should be described took quite significant time. On one hand we had two processes specified in the grant description — the production of the frame(s) and interaction of the NBR with the Euro Groups Register. On the other hand — we were used to allocate our statistical activities according to the official statistical program. The program is compiled by Statistics Estonia and approved by Ministry of Finance by first of July every year and is valid for the following five years. According to this program business register activities are divided into four main categories:

1. Statistical Economic Units Database (includes maintenance of NBR and production and maintenance of the frame and also register survey).
2. Economic units statistics (includes production and publishing of the aggregates of economic units population, and responding to queries).
3. Business Demography.
4. Enterprise groups (includes production and maintenance of the frames of enterprise groups and interaction with EGR).

According to the grant we were required to describe main processes and also visualise them, using standard notation BPMN. As we understand, the visualisation means the presentation of the processes as a timeline. The main challenge which arose: how to describe and visualise activities which have different character and timetable?

Activities of the production and maintenance of the frame and all preparatory activities like updating of the register from different sources and quality checks have different character and timetable. Some of these are continuous activities, some have strict timeframe with start and end dates and some — performed only in case of need. Moreover, several of these activities take place simultaneously.

The following steps were performed for selection of the processes which should be described:

> 1st step: All register activities were listed and mapped with GSBPM as a result of extensive brainstorming.
>
> 2nd step: This list was organised into Excel table according to GSBPM logic and order of the processes.
>
> 3rd step: All mapped information was supplemented with additional information like periods, when process or sub-process was usually performed, actors who perform the activity, manner of execution (manual, automatic or mixed) and whether the activity was carried out in live register or frozen frame.
>
> 4th step: All compiled information was analysed and a list of processes was finalised.

As a result four processes were selected for description:

1. Production of the frame
2. Maintenance of the frame
3. Register survey for new registered enterprises
4. Interaction of the NBR with the EuroGroups Register

Outputs to be produced by the end of the project for every selected process are following:

- Table with list of activities mapped with GSBPM relevant processes and sub-processes
- GSBPM diagram where all relevant processes and sub-processes are displayed
- Description of the processes using standard notation (BPMN)

**Example 1. GSBPM in a table form for register survey for new enterprises.**

The processes and sub-processes which are relevant for register survey were selected and information for time, periods, way of processing and actors were added.

| GSBPM sub-process	Processes and subprocesses	Time, Periodi-city	Frame or Base	Way of process ing	Roles/ actors	New survey	Existing survey	Comment
1.1	**Identify needs**							
	Creating new survey for updating the register		B		AKO, ARO	Always	—	New statistical unit, data source, variables…
1.2	**Consult and confirm needs**							
	Creating new questionnaire for updating the register		B		ARO, AKO, SOd	Always	—	
1.3	**Establish output objectives**							
	Assessment/revision of the data composition of the questionnaire	August-October	B		ARO	Always	As needed	
2.1	**Design outputs**							
	Design template of the questionnaire in cooperation with ALO	Decembe r	B		ARO, ALO	Always	As needed	
2.2	**Design variable descriptions**							
	Describing questionnaire variables in iMETA	Decembe r	B		ARO, ALO, MO	Always	As needed	
2.4	**Design frame and sample**							
	Design and approve methodology for frame and sample of the questionnaire		B		ARO, AKO, SOd	Always	—	
2.5	**Design processing and analysis**							
	Developing and approving of the processing rules, revising the existing rules for questionnaire in VAIS	Decembe r	B		ARO, ALO	Always	As needed	

It is important to specify whether we are describing an existing survey or the creation (and implementation) of completely new survey. In this table two columns were added – "New survey" and "Existing survey", which describe whether the sub-process is relevant or not. For example the processes 1.1 and 1.2 are only relevant in case of new survey.

Example 2: Production of the frame on GSBPM diagram

All relevant sub-processes are coloured.
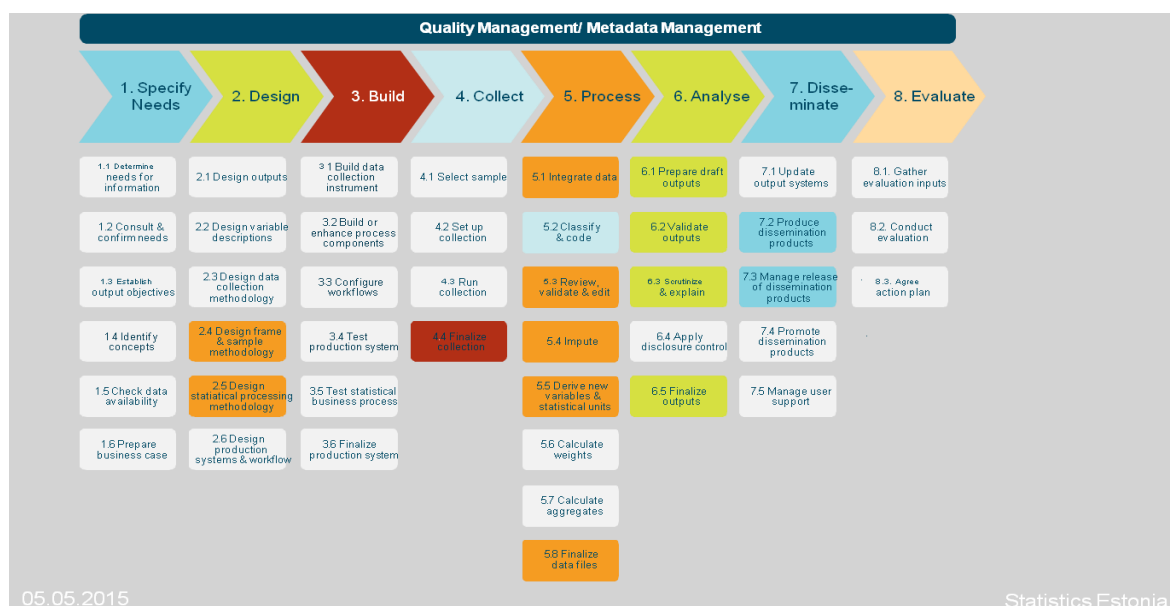
Example 3: Production of the frame on BPMN graph
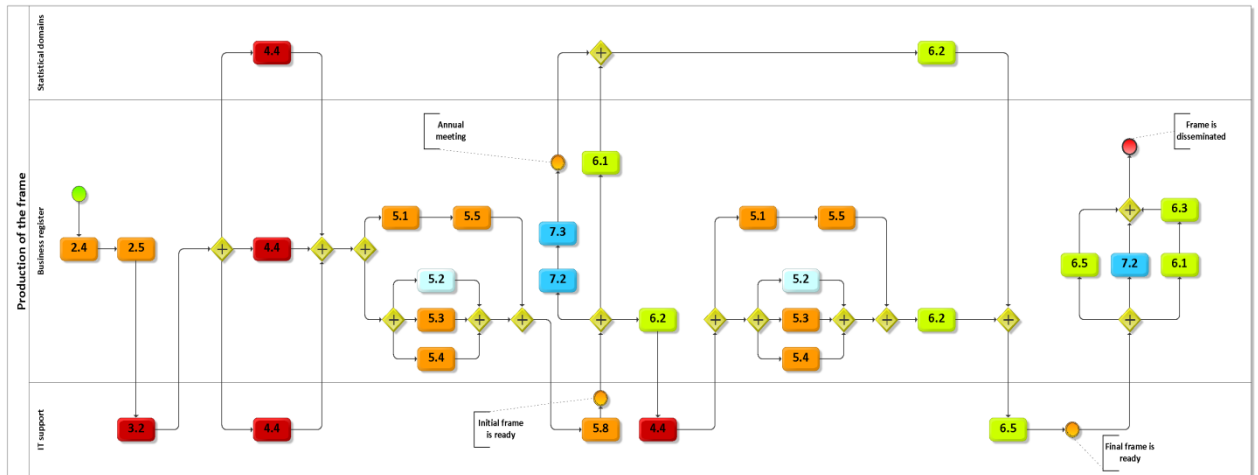
The process is presented as a timeline including the actors and the main milestones such as "Initial frame is ready" and "Final frame is ready".

The lessons learned so far:

➢ It is possible to use GSBPM for description of the register processes but in order to achieve the comparability between the MS the level of the detail of description needs to be agreed upon.

➢ There are some NBR processes which are described under more than one GSBPM sub-process.To minimize the differencies in interpretations of the register sub-processes the GSBPM model descriptions should be modified according to the BR processes (it can be a new version of GSBPM descriptions specifically created for register domain which relevant examples).

➢ How and what to present as a timeline as some activities are periodical, others are performed all year around.