

GEOSTAT 3 project – Minutes from coordination meeting

Date: November 1, 2017

Venue: Dublin Castle, Dublin

Participants

Coordinators

Jerker Moström, Statistics Sweden

Karin Hedeklint, Statistics Sweden

Project team

Ana M Santos, Statistics Portugal

Vilni Verner Holst Bloch, Statistics Norway

Erik Engelién, Statistics Norway

Marianne Dysterud, Statistics Norway

Ingrid Kaminger, Statistics Austria

Ylle Valgma, Statistics Estonia

Marja Tammilehto-Luode, Statistics Finland

Rina Tammisto, Statistics Finland

Anna Sławińska, Central Statistical Office of Poland

Pieter Bresters, Statistics Netherlands

Niek van Leeuwen, Statistics Netherlands

Consultants

Marina Backer Skaar, EFGS secretary

Pier-Giorgio Zaccheddu, Federal Agency for Cartography and

Geodesy (BKG), Germany

Eurostat

Ekkehard Petri

Agenda/index of minutes

1. Welcome
2. WP 1; ESS-SGF
3. WP 2; Testing the ESS-SGF
4. WP 3; EFGS website
5. WP 4; EFGS conference in Helsinki 2018
6. Upcoming meetings
7. Plans for future Geostat projects
8. Summing up
9. List of actions

Slides from presentations are annexed to the minutes.

1 Welcome

Jerker Moström gave an introduction to the meeting and welcomed all participants. Presentations from Vilni and Niek were added to the agenda.

Ekkehard sent us greetings from Gunther Schaefer, who could not participate in this meeting. He send us a message that the model should be as concrete as possible, by letting every principle be exemplified and if possible enhanced with indicators to assess the level of implementation.

2 WP 1; ESS-SGF

2.1 Name of framework

We had a discussion about the name of the framework. The working title *ESS-SGF* might be excluding for countries outside of the ESS. The model rests on two cornerstones; the INSPIRE directive and the ESS system. Those countries that are affected by them, should be addressed by the name of the model. Suggestions for a new name was *EU-SGF* or *SGF-Europe*.

Conclusions:

- The name *EU-SGF* could be even more excluding as it addresses fewer countries than ESS.
- The project didn't make any decision about a new name.
- It was agreed that the report must clearly describe who the model is addressing. It should be put in an introduction to the report.

2.2 Content so far, work through ESS-SGF draft

Overall structure

We now have a draft for the model and the final report. The main part of the report will consist of a brief description of every SGF principle, supplemented by practical examples.

As part of the main report, reflections on the future governance of the ESS-SGF (ownership, maintenance) and other requirements for its development are needed.

Conclusions:

- To keep the text short, most of the practical examples should be put in an annex, as was done in the Geostat 2 final report.
- We should allocate most of our time on the first three principles.
- We should not invent new methods, only summarize and, where needed, adapt existing ones, especially the ones concerning the global framework.
- We must be specific about whom the recommendations address, if it is NSI's, NMCA's or other communities with shared responsibility relying on a national spatial data infrastructure.

Principle 1: Use of fundamental geospatial infrastructure and geocoding

We discussed the recommendation Cooperation supported by institutional arrangements. This is an important part that cuts through the framework and concerns several of the principles. To make the recommendation less abstract, it needs more good practice examples. We could use examples presented in the UN-GGIM Europe report.

We discussed if the text about non-existing spatial data infrastructure is too basic for the institutions that this framework concerns. We agreed that it is not, since the framework should reach out to more than just NSI's. Even if there is an infrastructure in place, it could have missing pieces.

Conclusions:

- There is no need to do any large changes of the text, but we need more use cases.
- The recommendation to have one single national geocoding infrastructure for geocoding all other public and potentially private data files must be very strong.

- We must be more clear in our recommendations, by explaining consequences if you do or do not use them. We cannot assume that there is a common understanding.
- The recommendation *Cooperation supported by institutional arrangements* should be put in the front of the report.

Principle 2: Geocoded unit record data in a data management environment

We discussed the overlapping of principle 1 and 2. To get it more clear, we concluded that principle 1 concerns the building of an infrastructure and services, while principle 2 concerns the process of geocoding, i.e. the usage of the infrastructure.

We discussed the correspondence of spatial objects and statistical objects, that makes the geocoding difficult. To get an understanding of the problem, we need to explain the reason for it. As a use case we could map statistical objects with spatial ones, to see how they match. An attempt from Sweden will be shared. There was also an example from Switzerland in the UNECE workshop in Stockholm.

In some cases it is differences in definitions that causes problems in geocoding. We have an example from Portugal to show this. We need more examples though, including good practice of how to avoid the problem.

The recommendation to store location only once, is described in the Geostat 2 report, together with use cases from several countries. It is important to include this in the framework.

The history of the data is important for statistical institutes, who need to compare data in time series. It is a matter for both NSI's and NMCA's, depending on who manages the data. In some of the Inspire specifications it is recommended that the lifespan of every object is put in the data.

Statistics Norway will elaborate the text about geocoding quality declaration at object level. They will try to find examples from both Norway and other European countries. The examples could point to having a standard for this. We need a coordinated approach for this.

We discussed the alternative to document the accuracy in distance. That could be useful information for some, but it would be hard to receive that information.

An important statement in principle 2 is the use of point-of-entry validation-recommendation. Estonia will provide us with an example.

We should also briefly touch on the need for cross-border geocoding, e.g. for Business registers or Census where people work abroad.

Conclusions:

- Principle 1 concerns the building of an infrastructure, while principle 2 concerns the usage of it.
- Correspondence of spatial and statistical objects needs to be better explained and tested.
- Provide more examples about differences in definitions.
- The recommendation to store location only once is important.
- Elaborate the text of data history. Our recommendation will be that registers should include every change in objects, but should also include alternatives to it.
- Coordinated recommendations for quality declarations are needed.
- We need recommendations of how to handle cross-border addresses.

Principle 3: Common geographies for production and dissemination of statistics

Vilni is following the project of building a global grid system, initiated by OGC. It could be tested in WP 2, to see if the standard could be implemented in the infrastructure of a NSI. It could also be compared to INSPIRE's grid system. Ekkehard will check if this has already been tested by JRC. Vilni will keep covering this topic.

We need to discuss the pros and cons with keeping parallel grid system. Users are often asking for national grid systems, instead of those that are international harmonized. We have examples from Sweden and UK of this. There could be disclosure problems when parallel grid systems are compared to each other.

In the 2021 census, the member states have to deliver data in the European grid system. Our recommendation will be that it should be the primary grid system, i.e. if you do not yet have a national system, you should only use the European. Data quality (e.g. due to SDC) in the European grid should not be lower than in parallel national grids.

Conclusions:

- The ESS-SGF should include OGC's global grid system, in some way. We need to show that we are aware of this project.
- Discuss pros and cons with keeping parallel grid system.

- Recommend that the European grid system should be prioritized.
- Describe the process of the NUTS areas.
- Portugal will provide us with an example of cooperation between authorities. We need more good examples of this.
- Make references to the work of core data, that has been done by UN-GGIM; Europe WG A.

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

This principle is about aggregated data and has already been dealt with in the Geostat 2 project. Do we need any additional recommendations now? We will get more input from the UNECE workshop in Stockholm in November. If NMCA's have followed our recommendations in the previous principles, there should not be much in this and the next, that is of their concern.

If the revision of GSBPM is included in our recommendations, then we have reached our goal within this principle. That will be enough, if we want to make it easy. We could limit our recommendation to solutions that are already in place, to make the workload smaller.

In Principle 4, the recommendation Leave data at its source only includes published data. In principle 1 we have the same recommendation, but then it refers to unaggregated microdata. Ekkehard will provide a short description of this problem, i.e. that a unit could mean different things depending on the context.

There are several examples of frameworks to test and recommend. We could look at the European interoperability framework¹. Another example is the match of INSPIRE to SDMX for the 2021 Census. We should use Table joining service as an example, together with comments of other methods. Then we can argue that it is the best concept to meet the demands of the SGF.

Finland argued that we should recommend area classification as a coding system.

Vilni presented conclusions from his survey of linked open data.

Conclusions:

- Make a short reference to the GSBPM revision.
- Use TJS as an example of framework, together with comments of other methods.

¹ https://ec.europa.eu/isa2/eif_en

- We need a more generic text about data integration based on existing national statistical dissemination platforms. The OGC standard could provide us with that.
- PX web is an example that could be put in the annex.

Principle 5: Accessible and usable geospatially enabled statistics

In this principle our recommendations should be guidance for what is a meaningful output of the statistics. That could be complemented with examples of output that is not meaningful and that we don't recommend. We can only provide guidance, not restrict the usage.

The issue of privacy and disclosure control is an important part of this principle. It is complicated to recommend a certain guidance, as the choice of a specific method depends on the specific type of statistics. Instead we should focus on describing the problem. Statistics Netherlands is leading an European project to create harmonized methods for this. But the result will not be ready for us to use. Instead we could recommend that countries should use existing national standards. INSEE's handbook is one good example.

Conclusions:

- We can only provide guidance, not restrict the usage. We should describe what is a meaningful output.
- The level of geodata used for presenting statistics should be included in the recommendations. One such level is the NUTS regions.
- Make more general recommendations of disclosure control.

2.3 GSBPM under revision, how to handle this in Geostat 3

There are two proposals that concerns the content of Geostat 3. The first is to add new sub processes to the model. The other is to keep the existing sub processes, but add more recommendations to them. That also includes renaming the existing processes. Martin Brady at Statistics Australia is in the revision group.

Conclusion:

- We should keep the text in our report short about this, with just a reference to the revision.

2.4 Revision of code of practice

Eurostat's code of practice is currently under revision. The plan is to complete it with references to geospatial methods. The current proposal has included the word *geospatial* in several of the principles. The purpose is to institutionalize the concept of geospatial activities.

Comment after the meeting: The proposal was rejected for principle reasons, as no data source could be specified.

3 WP 2; Testing the ESS-SGF

Marianne presented three indicators that have been chosen for testing, by UN GGIM Europe. At least one of them will be tested in the Geostat 3 project, but it might be too much work to do them all.

The indicators are:

11.2.1 Access to public transport (Tier 2):

It is a challenge that data that covers all parts of the definition is not available, for example data of what is safe and comfortable environments. Thus, the definition has to be limited. For example by only including dwellings.

11.3.1 Land consumption by population growth (Tier 2):

It has been difficult to find a good definition for this indicator. It is not clear if time series are included, when measuring growth. We also need to define the concept *land consumption*, if it should include roads, gardens etc.

Data available for the testing is earth observation data from the Global Human Settlement Layer. That could be compared to national data of land consumption, for example locality borders, when those are available.

11.7.1 Built up areas of cities, open space for public use (Tier 3):

This indicator also lacks a clear definition, since it isn't specified what types of open space that should be included. We also need to define the concept *cities*, and investigate if we should use grid data or LAU2 areas. TJS could be an option when testing this indicator.

Conclusions:

- It is an important part of the tests, to investigate what data are available and in what way the definitions need to be adjusted to that.
- To make it more easy for us, we should look at the outcome of the UN GGIM project. Our time schedule fits very well to that project, since it is a bit ahead of us.
- Our testing period will start after our meeting in February, with an end during July 2018.

4 WP 3; EFGS website

4.1 Feasibility study on migrating the EFGS website to the CROS portal

Anna presented pros and cons with having an independent EFGS web site compared to migrating the content to the CROS portal. The disadvantages with the migration seem to outweigh the advantages by far.

The financing from Eurostat is uncertain and new sources for funding should be investigated. Maintenance is a bigger cost than hosting the website.

Conclusions:

The study should now be completed with the following:

- Examples to illustrate and develop the list of pros and cons.
- Some sort of scoring/rating of the list would make it more objective.
- Ideas of independent financing.
- Results from testing the CROS portal from an administrative point of view. Anna will receive a testing space from Eurostat.

The study needs to be ready by the end of the Geostat 3 project, i.e. in January 2019.

4.2 Implementation of ESS-SGF on the EFGS website

When the ESS-SGF report is complete, it is going to be published at the EFGS website. As the framework will contain a lot of references to examples and other documents, it could be published with hyperlinks that leads to the references. The structure could be similar to what is used by Wikipedia. This could make the framework more available to those that do not like to read it as a report.

One suitable place to put this is at the tab *Information base – Statistical geospatial frameworks*. Alternatively at the tab *Geostat – Geostat 3*.

It needs to be investigated if this approach fits within the project's budget.

4.3 Other businesses concerning the EFGS website

The terminology page (*Information base – Introduction to spatial statistics – Terminology*) should have more hyperlinks, for cross-referencing. The structure could be similar to that used by Wikipedia.

The website constantly needs more case studies and best practice examples. Even though a lot of requests are sent to members of the community, not much are sent to the web editors. It is a problem that little of the national examples are translated into English.

5 WP 4; EFGS conference in Helsinki 2018

Marja presented the plan so far, for the 2018 EFGS conference in Helsinki. The conference will be for 3 days, October 16-18. Rooms for parallel sessions, during one day, have been reserved.

The draft budget is sufficient for 130 participants. But as the number of participants in Dublin was about 200, there should perhaps be room for more. As Eurostat cannot provide any additional funds, Statistics Finland is considering using sponsors. The EFGS conference 2015 in Vienna, had two sponsors. Another alternative is to limit the number of participants.

Marja would like to receive all valuable information from the previous organizers, as checklists, contacts etc.

6 Upcoming meetings

6.1 Geostat 3 project meetings

The next project meetings are:

- November 22, webex meeting, 13:00-14:00 CET
- December 14, webex meeting, 10:00-11:00 CET
- January 17, webex meeting, 13:00-14:00 CET
- February 14-15, Stockholm, 2 full days

6.2 Meetings with related projects

The work within UN GGIM Europe will continue. Jerker and Ingrid are involved in that.

Next year's Quality conference, organized by Eurostat, will be in Poland, on June 26-29. The ESS-SGF could be presented there, as geospatial methods will be one of the topics in the call. There will also be one session about the 2021 census.

The GISCO meeting will be in April next year. Linked open data and Geostat 3 will be on the agenda.

7 Plans for future Geostat projects

There will not be a launch for a new Geostat project during 2018, since it is too stressful to start a new one right after the previous. The launch will be in 2019 instead, with a project start in early 2020. To have a two years project followed by one year of break is a project cycle that will probably continue.

However, Eurostat already need some first ideas of content for the Geostat 4 project. Some short bullet points is enough. Send input to Ekkehard by the end of 2017.

During 2019 the website and conference will probably be financed with call for tenders. These two parts might be separated from the Geostat projects in the future.

NSI's could use regular call for proposals for 'Merging Statistics and Geospatial Information' to work on the implementation of the ESS-SGF principles on a national level.

8 Summing up

We ended the meeting with summing up what needs to be done next.

Jerker will continue working with the ESS-SGF document. Before we go into the testing phase we need a quite ready draft of the ESS-SGF. All additional contributions must be sent to him by December 2017.

A new draft to review should be ready by the end of 2017. That one should be circulated to a larger group, within each country. NMCA's should be included in the reviewing process.

Gunther Schaefer will need a draft of the framework before the administrative interim report is ready. The interim report will be written by Karin. It must be ready by February 2018.

We should start using the GISCO wiki for sharing documents within the project. To test the wiki, Karin will upload a document that everyone in the project will try to sign. Karin will send us an e-mail about that.

9 List of actions

Summary of actions, from the previous text.

WP	Activity	Performer
WP 0	Interim administrative report, ready by February 2018	Sweden/Karin
WP 0	Initiate test of the usability of the GISCO wiki.	Sweden/Karin
WP 1	Clearly describe who the model is addressing. Put it in an introduction to the report.	Sweden/Jerker
WP 1	All additional contributions to the report must be sent to Jerker by December 2017.	All
WP 1	A new draft for reviewing will be ready by the end of 2017. That should be circulated to NMCA's.	All
WP 1	A quite ready draft of the ESS-SGF must be done by January 2018.	Sweden/Jerker
WP 1 Pr 1	Provide examples of consequences if you do not follow the recommendations in Principle 1.	Netherlands, Poland
WP 1 Pr 2	Map statistical objects with spatial objects, to see how they match. That will illustrate the problems of correspondence of spatial and statistical objects.	Sweden
WP 1 Pr 2	Show how differences in definitions might cause problems in geocoding. Make suggestions of how to avoid the problem.	Portugal
WP 1 Pr 2	Elaborate the text of data history, with the recommendation that registers should include every change in objects. Alternatives should also be put in the text.	Austria
WP 1 Pr 2	Elaborate the text about geocoding quality declaration at object level. Find examples from both Norway and other European countries.	Norway/ Marianne
WP 1 Pr 2	Provide an example of point-of-entry validation.	Estonia

WP 1 Pr 3	Keep covering the project of building a global grid system.	Norway/Vilni
WP 1 Pr 3	Check if the global grid system has already been tested by JRC.	Ekkehard
WP 1 Pr 3	Describe the process of the NUTS areas.	Ekkehard
WP 1 Pr 3	Discuss pros and cons with keeping parallel grid system. Collect input from UK.	Sweden
WP 1 Pr 3	Provide an example of cooperation between authorities.	Portugal
WP 1 Pr 4	Provide a short description of the problem that a unit could mean different things depending on the context. Example from <i>Leave data at its source</i> .	Ekkehard
WP 1 Pr 4	Describe the idea of using area classification as a coding system.	Finland
WP 2	The testing period will be February - July 2018	Norway
WP 3	Complete the feasibility study of the CROS portal.	Poland
WP 3	Investigated if it fits within the project's budget that we make new web pages that includes the ESS-SGF with hyperlinks.	Poland
WP 3	Examine if it is possible to use more hyperlinks at the terminology web page.	Poland
WP 4	Marja would like to receive all valuable information from the previous organizers of EFGS conferences, as checklists, contacts etc.	Austria, France, Ireland etc
Geostat 4	Send ideas of content for the Geostat 4 project. Eurostat will need that by the end of 2017.	All