

**Economic and Social Council**Distr.: General  
12 February 2016

Original: English

**Economic Commission for Europe**

## Conference of European Statisticians

**Sixty-fourth plenary session**

Paris, 27-29 April 2016

Item 6 of the provisional agenda

**Geospatial information services based on official statistics****Building an infrastructure for European geospatial statistics****Note by Statistics Sweden in collaboration with Eurostat***Summary*

This paper presents the concept of geospatial statistics as the result of a close integration of statistical and geospatial information. Geospatial statistics are statistics with a high degree of spatial resolution. While the methodology of producing geospatial statistics as such are well established and described, operational infrastructure similar to the infrastructure for regular official statistics is missing. The paper introduces the Statistical Geospatial Framework currently under development internationally. As a national best practice, Statistics Sweden has already successfully created such an infrastructure and is planning on enhancing it based on the results of the two above initiatives. The paper then addresses a number of organisational issues mainly concerning the cooperation between statistical offices and mapping authorities that hamper the building and operation of the infrastructure. The paper concludes with the recommendation to develop a joint implementation plan of statistical offices and mapping authorities under the auspices of the UN initiative on Global Geospatial Information Management and coordinate the actual implementation with the 2020 round of censuses.

This paper is presented for discussion to the Conference of European Statisticians seminar on “Geospatial information services based on official statistics”.



## I. Introduction

1. Statistical offices have linked statistical data to a location for many years and used this link for producing sub-national statistics, high resolution geospatial statistics and other types of location based information. Still, the majority of sub-national statistics are disseminated for administrative areas. On the user side though, policy makers and businesses demand more and higher resolution geospatial statistics products and location based information. Until recently this demand could not be met in most countries due to scarcity of georeferenced input data and essential geospatial information.

2. Several major trends and initiatives over the last 10 years have substantially improved the condition for geospatial statistics in Europe, though, and have helped to better answer user needs: the proliferation of location-enabled devices (GPS) producing large amounts of geo-enabled data on the data side and the adoption of the directive on the Infrastructure for Spatial Information in the European Community (INSPIRE) to mention a few.

3. Forthcoming major statistical operation such as the next census 2021 will very likely increase the use of both smart GIS technology for data collection and georeferencing of the data for more detailed products.

4. The establishment of the United Nations initiative on Global Geospatial Information Management (UN-GGIM) in 2011 with the goal to enhance the use of geospatial information for policy making addresses the integration of statistical and geospatial as one of its key priorities. In the coming years the United Nations 2030 agenda for sustainable development will call for a data revolution to monitor the progress and achievements. The recently adopted Sustainable Development Goals and the European Statistical System (ESS) Vision 2020 make explicit reference to geospatial information as a valuable source for more relevant information. Hence the geospatial information community and the statistical community together are at the centre of this data revolution.

5. The on-going modernisation of statistics has already resulted in much better conditions for geospatial statistics, mostly in business and social statistics. For instance administrative data or geocoding of unit record information are now acknowledged as a means to create more relevant information and to be more efficient.

6. These trends have helped to raise the awareness for location based information and as such set the scene for more geospatial statistics. However there are many overlaps and redundancies between these trends and initiatives while also gaps remain and a single European geospatial reference framework for statistics is missing. Several study groups have already addressed the topic of statistical-geospatial data integration and have put forward suggestions on how to improve the situation and merge all ongoing activities into a single coherent framework<sup>2</sup>. This calls for a concerted approach and a close cooperation among the public information providers led by the National Mapping and Cadastral Authorities (NMCA) and National Statistical Offices (NSO). This should result in a harmonised European infrastructure for geospatial statistics that is comprehensive and can be implemented in all countries in a similar way to achieve full interoperability across country and domain borders.

---

<sup>2</sup> GEOSTAT 1 project <http://www.efgs.info/geostat>  
ESS Task Force on the integration of statistical and geospatial information  
[https://circabc.europa.eu/d/a/workspace/SpacesStore/fd349927-3c7d-435a-8b80-4ace48daf646/D\\_GIS\\_105%20GISCO-TF-Report-V\\_3.doc](https://circabc.europa.eu/d/a/workspace/SpacesStore/fd349927-3c7d-435a-8b80-4ace48daf646/D_GIS_105%20GISCO-TF-Report-V_3.doc)

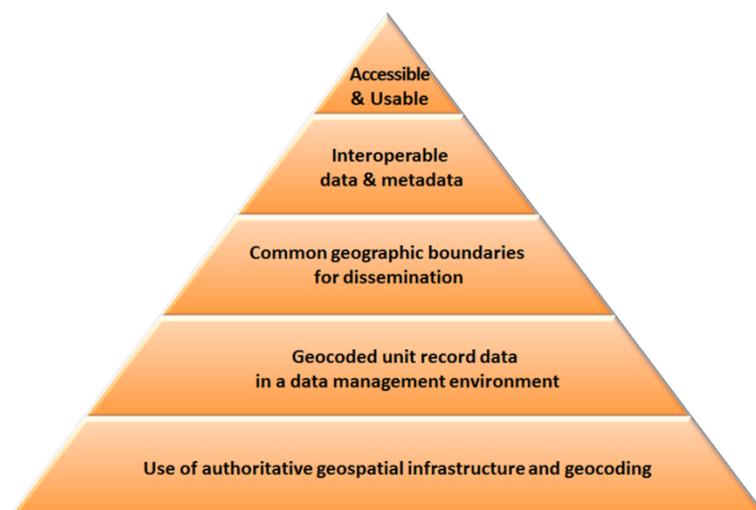
7. This paper aims at presenting a proposal for such a framework building on several recent projects at the European and international level. In the second part the paper sets out a number of issues that currently prevent the full implementation of such an infrastructure in all countries and thus the smooth integration of statistical and geospatial information for the purposes of evidence based policy making.

## II. What is an infrastructure for geospatial statistics?

8. Geospatial information services based on official statistics require an infrastructure for European geospatial statistics. The main features of the infrastructure are covered in the Statistical Geospatial Framework proposed by the UN-GGIM Expert Group on Integration of Statistical and Geospatial Information<sup>3</sup>. This proposal aims at encompassing all aspects of geospatial statistics and statistical-geospatial data integration and provides a solid yet generic framework for all providers of spatially enabled information.

Figure 1

### The Statistical Geospatial Framework



9. The main benefit of the framework is to establish common methods of geospatially enabling statistical and administrative data and integrating this with geospatial information. This will enable:

- new, better and more integrated information for decision making processes;
- comparisons within and between countries;
- increased information on smaller geographic areas;
- the development of common tools/applications to support the integration of data and data sharing;
- commercial development and adoption of geospatial tools; and
- more efficient production of official statistics.

10. The Statistical Geospatial Framework should be supplemented by various generic elements to be operational, the most important being an extended version of the Generic

<sup>3</sup> [http://ggim.un.org/UN\\_GGIM\\_Expert%20Group.html](http://ggim.un.org/UN_GGIM_Expert%20Group.html)

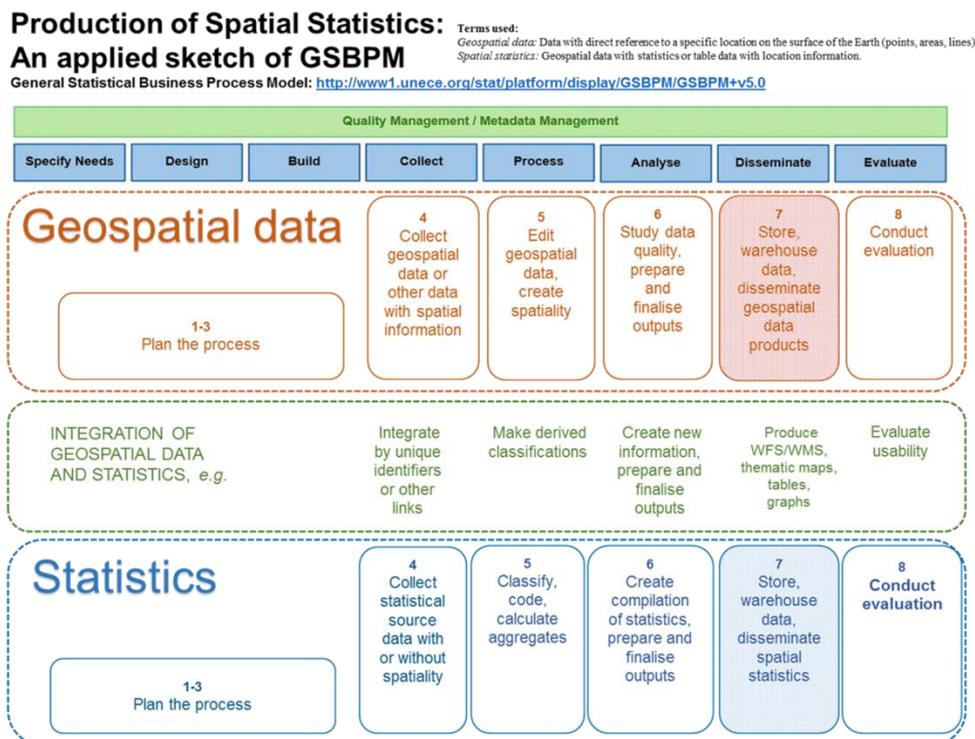
Statistical Business Process Model (GSBPM). Interoperable metadata standards and common terminology are examples of other elements that need to be developed. The extension of the GSBPM model to the integration of geospatial information into the statistical production process provides a link to internationally agreed statistical processes and facilitates the communication between the statistical and the geospatial communities. In addition, including geospatial information management in the GSBPM (see figure 2) would extend the infrastructure for geospatial statistics to the actual data production and thus to the core business process of statistical offices. In fact including geospatial information management into the GSBPM has the advantage that the infrastructure for geospatial statistics would merely represent an enhancement of the existing statistical infrastructure and would not result in a separate infrastructure, possibly incompatible infrastructure.

11. The GEOSTAT 2 project<sup>4</sup> will have a corresponding proposal ready by the end of 2016, which together with the framework will constitute the infrastructure for European geospatial statistics.

12. Without explicitly making reference to the main building blocks of the future European infrastructure there are already many examples for good progress in the direction of a better integration of statistics and geospatial information at the national and European level: the censuses and population statistics are the most noteworthy areas but also transport statistics. As an example the recently published UNECE census recommendations<sup>5</sup> advocate the creation of 1km<sup>2</sup> population grids and require that persons should be geocoded to a geographical point.

Figure 2

**An early draft of how to include geospatial processes in the GSBPM by the GEOSTAT2 project in 2015**



<sup>4</sup> <http://www.efgs.info/geostat/2>

<sup>5</sup> <http://www.unece.org/publications/2020recomm.html>

13. The national implementation of the framework and process model should aim to create effective and streamlined production of geospatial statistics of high relevance for users. Depending on the national preconditions this might be achieved in a number of different ways. The setup at Statistics Sweden can serve as one example, where the production process relies entirely on administrative sources and geospatial information provided by the National Mapping and Cadastral Agency.

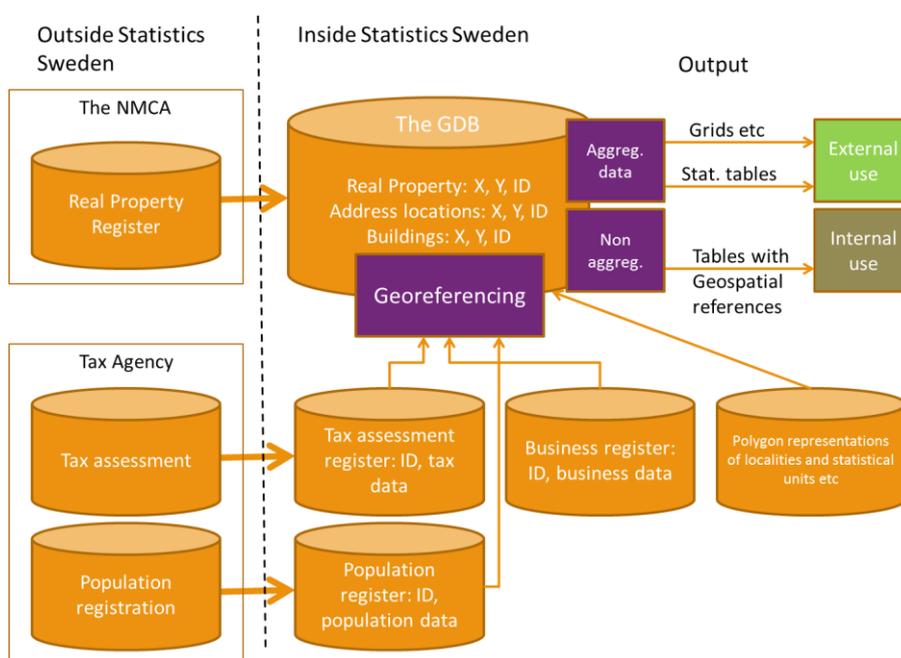
14. Europe is in a good position to develop geospatial information services due to the fact that access to geospatial information gradually increases when the EU countries put up INSPIRE services. There is also an increasing national expertise on how to disaggregate data at high levels down to small area geographies as techniques has evolved in the GEOSTAT 1 project<sup>6</sup> – both on how to disaggregate (top-down) and how to use administrative sources (bottom-up).

15. The need for geographic detail in official statistics will increase in the coming years: to answer to new needs in relation to the United Nations sustainable development goals, or needs to be more efficient in policy making proving that evidence based decisions can be backed up with relevant statistics. Relevance includes both the right level of geographic detail and timeliness in the data.

16. There are already a number of good examples of national cooperation between authorities responsible for geospatial information for instance in the Nordic countries. There is a rapid transformation going on making more and more publically financed data open and accessible, but also in countries where geospatial information still comes with a fee it is possible to increase the use within the public sector through agreements – such as in Sweden for example.

Figure 3

### Production setup for geospatial statistics at Statistics Sweden



<sup>6</sup> <http://www.efgs.info/geostat>

17. The requirements on data sharing in the INSPIRE directive was the starting point for the spatial data cooperation which was put in place in 2011. Since then all authorities in Sweden by paying an annual fee get access to all geospatial information they need for their public responsibilities. For Statistics Sweden this has led to a whole range of new possibilities mainly in the production of official statistics on land use. The statistical production is now more advanced in the throughput phase of the GSBPM, as geographic information system (GIS) analysis is carried out to measure all kinds of geospatial phenomena. Linking geospatial objects with statistical registers has proven to give new and interesting results, such as land accounts for biodiversity and ecosystem services accounting.

18. Another positive effect has been the timeliness and accuracy in real estate, address and building data – essential for the whole register based statistical production at Statistics Sweden. Before 2011 the updates were too expensive, which meant that while the total population register or the business register were updated every day, the real estate register were updated only a few times a year.

### **III. Issues and challenges**

19. With INSPIRE and the proposals for the Statistical and Geospatial Framework, the main elements of a European infrastructure for geospatial statistics are in place or fairly advanced as a concept.

20. However, all steps in the direction of geospatial statistics have so far been made in an ad-hoc manner for specific statistical operations and without an underlying strategy for geospatial statistics in mind. A more systematic and strategic approach to data integration building on a common model for this infrastructure is desirable.

21. While technical challenges remain, and full technical interoperability of statistical and geospatial data is still an issue in the day to day operations, this has been already addressed under various interoperability actions such as INSPIRE and for most of them there is a responsible party.

22. However the success stories from national infrastructures such as Sweden underline that the main challenges come from a lack of cooperation between different institutions and between countries. This applies for the European level but also for the internal situation in most countries. This overall challenge can be structured into a number of organisational and institutional issues that mainly concern the cooperation between NSO and NMCA within countries but also touch on the international cooperation among these institutions.

#### **A. Issue 1: Awareness of senior managers of the benefits of data integration and cooperation**

23. Building and maintaining the infrastructure for geospatial statistics will require investments that will often only pay off in the longer term. As an example building the infrastructure for the census based on geocoded registers might be higher initially than for a traditional census, but once the infrastructure is in place, this will generate substantial savings in future censuses. Senior managers who have to take these strategic investment decisions have not only to understand the benefits but also have to be rewarded for these investments. This is often not the case.

24. The lack of awareness or benefits on the producer side is often paired with a lack of awareness of the power of geospatial statistics on the demand side, in particular among

decision makers and politicians. This lack of a concrete demand makes investments even more difficult to justify.

25. As a result the cooperation between the two communities is often not beneficial for either side. Currently only if the benefits are immediate, investments are made but often restricted to an individual project. These projects then risk producing a fragmented data infrastructure.

26. Action to be considered: The political level has to request more detailed information, and heads of public authorities have to develop and enforce a strategy to meet these requirements in a systematic way. The political level and senior managers in NSOs need to create the conditions that investments into the infrastructure for geospatial statistics are rewarded.

## **B. Issue 2: Missing institutional arrangements between NSOs and NMCAs**

27. Cooperation between NSOs and NMCAs is essential to develop new geospatial statistics. The INSPIRE directive has pushed for cooperation between all authorities with geospatial information of importance for environment purposes. However the current level of implementation of the directive – mainly meeting very technical demands on web services that allows for discovery, view and download of geospatial information – is not enough as a foundation for the integration of geospatial and statistical information in most countries.

28. Achieving semantic interoperability and access to the essential geospatial information can only be achieved by technical coordination and harmonisation of data that goes beyond the requirements of INSPIRE. First examples of such arrangements have emerged in a few countries such as Sweden but are not common practice in other countries.

29. The statistical community in the form of the ESS has established solid formal governance internationally for this coordination and harmonisation work. Specifically for geospatial statistics ESS initiatives such as GEOSTAT and the European Forum for Geography and Statistics EFGS have already achieved a lot in terms of better coordination among NSOs, but the involvement of NMCAs is relatively low.

30. The geospatial community on the other side outside EuroGeographics lacks a firm governance structure, and coordination and harmonisation among NMCAs in addition to INSPIRE is only emerging.

31. UN-GGIM and its European branch UN-GGIM: Europe are addressing these issues. As an example UN-GGIM: Europe which is chaired by the Swedish NMCA has formulated an ambitious work programme, focussing on core data (Work Group A) and the integration of geospatial data and statistics (Work Group B). However participation in both initiatives is still on a voluntary basis and formal and binding agreements between NSOs and NMCAs on their mutual responsibilities are missing. Also UN-GGIM does not yet provide a working platform for the actual creation of information or take care of building the infrastructure for combining data.

32. Action to be considered: Under UN-GGIM a set of formal institutional arrangements for instance in the form of Memoranda of Understanding and a common code of practice need to be put in place which favour long term and systematic cooperation between NSOs and NMCAs.

### **C. Issue 3: Access to geospatial information and geocoded administrative and big data**

33. The main purpose of institutional arrangements between NMCAs and NSOs is to have clearly regulated access conditions to geospatial data. Currently, the situation regarding access and use conditions to essential geospatial data is extremely diverse across countries but a full free and open access of statistical offices to geospatial information is still the exception rather than the rule.

34. For geospatial statistics alternative data source such as Big Data, e.g. from mobile phone positions represent a huge potential to fill data gaps. However, most of these data are owned by private companies and access to these data is subject to agreements or data procurement. So far most providers are reluctant to transfer Big Data to statistical offices for various reasons, including privacy concerns and own business interests.

35. In terms of data requirements, future data specifications need to pay more attention to the time dimension of geospatial information (timeliness, periodicity, and time stamping), which is very important to follow trends and changes in the environment /SDGs.

36. Action to be considered: A formal framework for data exchange among public authorities needs to be established in all countries granting comprehensive access to all public geospatial information in a country. NMCAs should make strong and binding commitment to provide harmonised geospatial information.

37. In addition to these more global issues regarding the cooperation between providers of geospatial information and NSOs there are several issues that are specific to statistical offices.

### **D. Issue 4: Conflicting data policies and business models for geospatial statistics**

38. Open data is an increasing trend in most European countries, and official statistics have been open data for many years. However, geospatial statistics are not yet official statistics and represent an important source of income for several NSOs. This represents a barrier to their more widespread uptake from communities that generally cannot afford expensive data sources including third party public authorities. At this stage the benefits of more open data are hard to quantify and as long as NSOs are not compensated for the loss of revenues, they are reluctant to make their geospatial statistics available as open data.

39. Action to be considered: NSOs are challenged to find alternative funds for making available geospatial data as open data.

### **E. Issue 5: Privacy and data protection**

40. In Europe, where vast sparsely populated areas alternate with densely populated cities, often within one country, introducing the territorial dimension into statistical products dramatically increases the risk of disclosure. Finding a balance between data protection and the usefulness of the information for the user is not easy. Also privacy is not always an absolutely defined concept and rules and restrictions vary across countries. NSOs are aware of the problem and each of them has developed their own standards and practices. However this diversity makes it difficult for the user of cross-border geospatial statistics to understand the limitations of data. This calls for a harmonised approach.

41. Action to be considered: NSOs and international statistical organisations should address this issue as a priority and define a set of harmonised data protection rules and practices, with a view to the next round of censuses.

## **F. Issue 6: Resources and skills among staff**

42. Based on a survey conducted as part of the Eurostat GEOSTAT 2 project, many experts in NSOs see a need for improved geospatial skills in their organisation. This issue, together with the natural inertia of large communities regarding change and the lack of management support leads to sluggish uptake of new geospatial data sources and technologies and data sources into statistical production processes.

43. The ESS has already partially addressed this issue by including an ESTP training course on basic geospatial information management in their training offer, but the capacity is small and the level or expertise obtained from such short term trainings is limited.

44. Action to be considered: To improve the situation, geospatial skill should get much more attention in the future recruitment and training policies of NSOs. NSOs and the ESS should intensify their training efforts for geospatial information management and increase the number of geospatial experts among their staff.

## **IV. Conclusions and recommendations**

45. Any statistical geospatial infrastructure that is built to support the production of geospatial statistics should include the use of geospatial workflows and technology, as a key to advance on the integration of geospatial and statistical information. These workflows should be incorporated into the GSBPM following the proposal from the GEOSTAT 2 project.

46. The 2020 round of censuses represents a huge opportunity to start building the infrastructure for geospatial statistics based on the Statistical Geospatial Framework and the enhanced GSBPM in a systematic way. As a minimum the requirement to geocode all unit record information to an address point as set out in the Framework should be implemented for the census, not least to fulfil this requirement from the UNECE census recommendations. As the lead-in time to deliver any census is significant, these challenges need to be considered now if standardised approaches are to be put in place in time.

47. Countries should initiate a process to increase the number of national, authoritative geospatial datasets (addresses and others) meeting stakeholders (such as statistics) requirements within Member States. This should be incorporated into a geospatial infrastructure maintenance process including data, metadata, services, architectures and business models. The content (data, metadata and services) should be accessible to all stakeholders (authorities). The report<sup>7</sup> of the Eurostat task force, on the integration of statistical and geospatial information, states which data that are needed from the NSOs and should be used when deciding which data to give priority.

48. The data sharing principles of Service-Oriented Architectures (like INSPIRE), the Statistical Geospatial Framework and the extended GSBPM should form the core elements of the European infrastructure for geospatial statistics. A coherent description of the infrastructure including an implementation plan similarly to the ESS Vision 2020 would

---

<sup>7</sup> [https://circabc.europa.eu/d/a/workspace/SpacesStore/fd349927-3c7d-435a-8b80-4ace48daf646/D\\_GIS\\_105%20GISCO-TF-Report-V\\_3.doc](https://circabc.europa.eu/d/a/workspace/SpacesStore/fd349927-3c7d-435a-8b80-4ace48daf646/D_GIS_105%20GISCO-TF-Report-V_3.doc)

allow all stakeholders to work towards a common and shared goal and to communicate the key characteristics of the infrastructure to external stakeholders, in particular policy makers. Countries should start such an effort under the umbrella of an international statistical organisation.

49. To tackle the above and other more technical issues requires a strategic vision that is supported by all levels of management and shared by all countries. Countries in Europe should therefore launch the development of a European Geospatial Data Strategy based on comprehensive National Geospatial Data Strategies. Countries are also invited to decide on institutional arrangements, including legal arrangements, needed to enable and increase the cooperation between NMCAs and NSOs as well as commercial, scientific and public domains. Clear responsibilities for improving the conditions for geospatial statistics are essential as currently neither NSOs nor NMCAs feel responsible for advancing the topic of geospatial statistics. International cooperation mechanisms such as UN-GGIM should take the lead in the creation of the infrastructure and obtain a strong mandate to enforce the governance of the whole process.

50. The timeline for designing the implementation road map for the European infrastructure for geospatial statistics should be aligned with the implementation roadmap of the 2020 round of censuses. As the previous one the next census as one of the largest single statistical operation represents a huge opportunity to deeply transform statistical production processes and tap new data sources, as it typically leads to the mobilisation of resources and a level of coordination among public authorities that otherwise cannot be achieved. Thus the statistical community together with its geospatial partners should grasp this opportunity and tackle the above issues as part of the census roadmap with the overall infrastructure in mind. This will create huge synergies with the census implementation and help to minimise the resource impact of building of the European infrastructure for geospatial statistics.

---