

JIDE 2017

VIII Jornadas Ibéricas de Infraestructuras de Datos Espaciais

Lisboa | 15 - 17 novembro 2017

GEOSTAT 3 - A Statistical Geospatial Framework for Sustainable Development

INE

DMSI/GEO

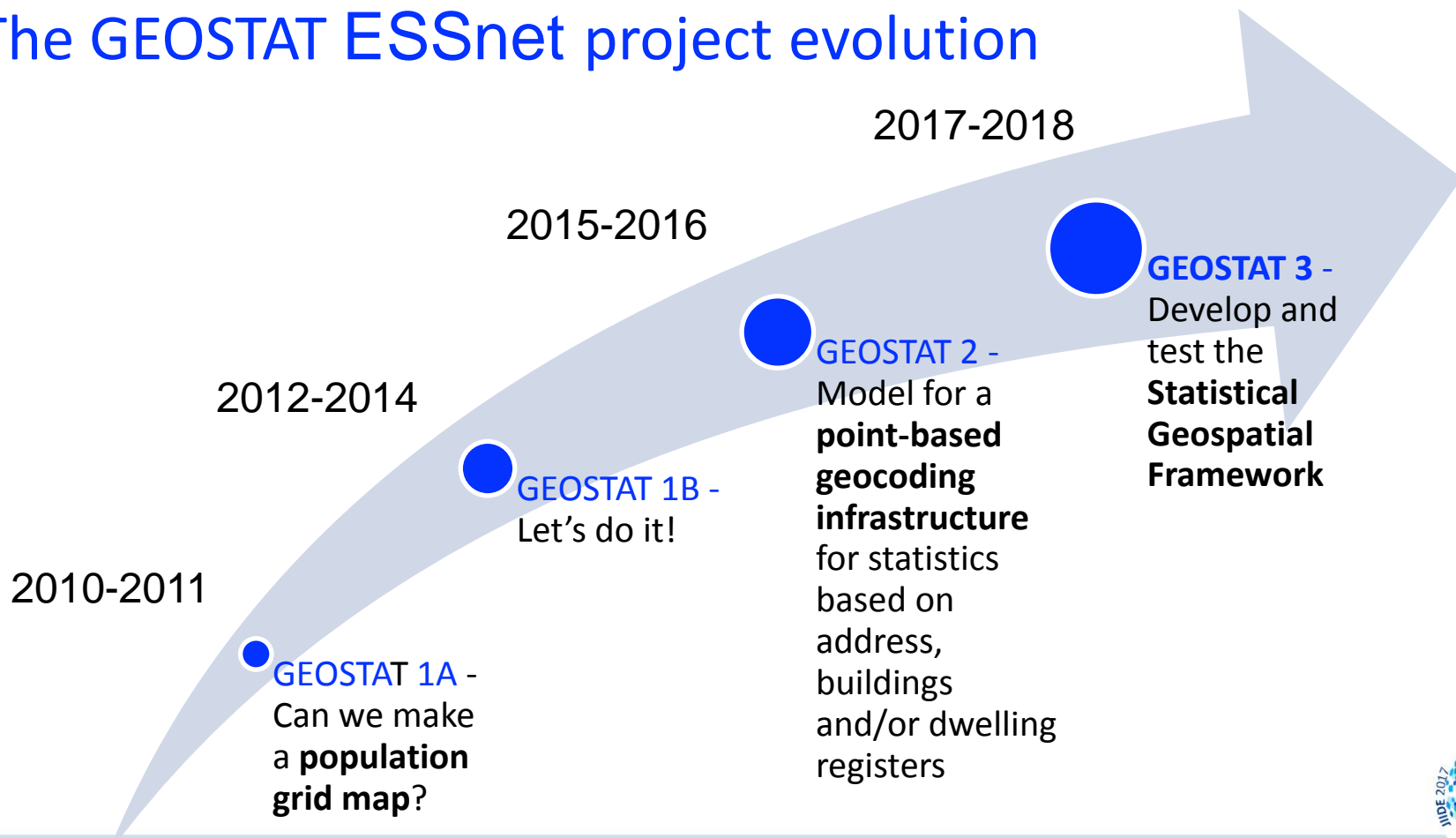
Ana Santos



INSTITUTO
GEOGRÁFICO
NACIONAL



The GEOSTAT ESSnet project evolution



The consortium

- Statistics Sweden (Coordinator)
- Statistics Austria
- Statistics Estonia
- Statistics Finland
- Statistics Netherlands
- Statistics Norway
- Statistics Poland
- **Statistics Portugal**

Sub-contractors:

- BKG (Germany)
- Kartverket (Norway)
- MD (Sweden)



GEOSTAT 3 overall scope

“To develop and test the **Statistical Geospatial Framework (SGF)** for the ESS, taking into account existing conditions, initiatives and European and national frameworks”

Why?

- To harmonise methods for the integration of statistical and geospatial information within the ESS
- To modernise the statistical system and increase efficiency and flexibility
- To provide a better foundation for collaboration between NSIs and geospatial agencies in providing society with more and better data for evidence based decision-making
- **The main drivers:** the goal of a fully geocoded population census 2021 and provision of data for the UN SDG indicator framework

By Jerker Mostrom from EFGS Conference – 2, 3 November 2017, Dublin 2017



But there is already a Global Framework!

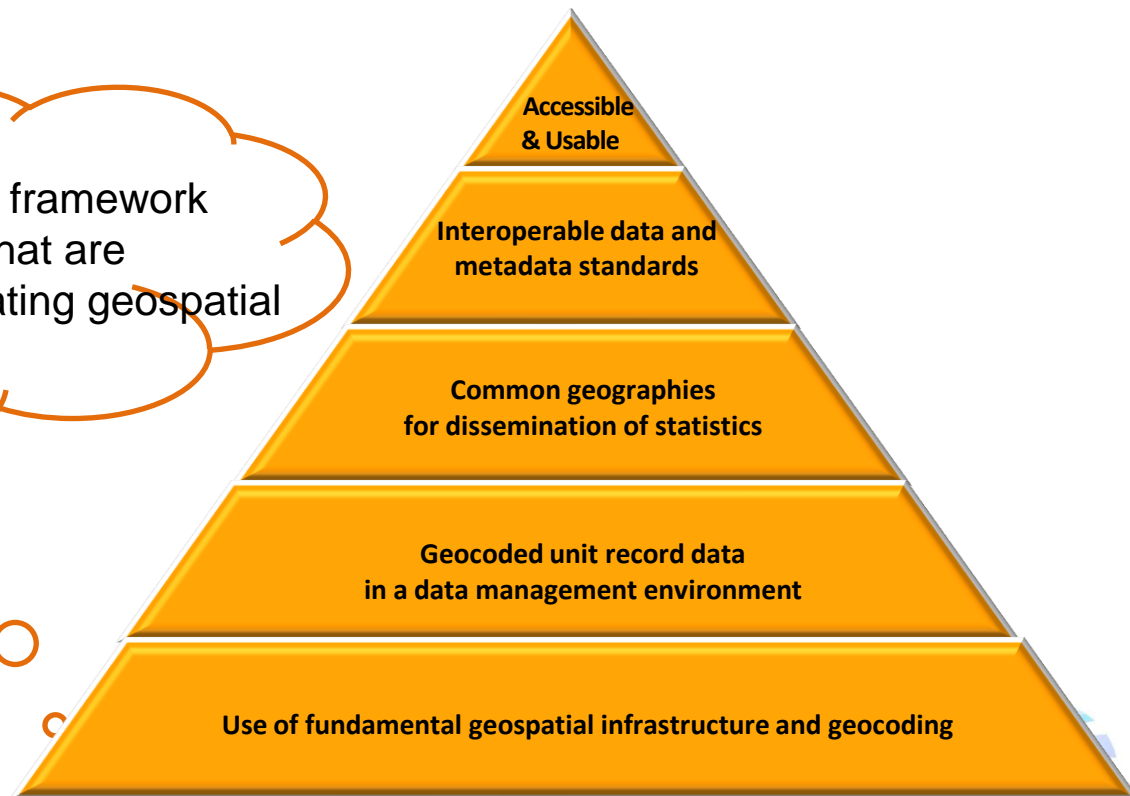
- The ESS-SGF builds on the 5 principles of the **Global Statistical Geospatial Framework (GSGF)** but is to be less generic
- Specific regional conditions to build on:
 - **INSPIRE** > Legal framework for National Spatial Data Infrastructures
 - The **European Statistical System (ESS)** > Legal obligations and mechanisms for statistical data + funding opportunities
 - **EFGS** > Voluntary collaboration and harmonisation for geospatial statistics
 - **EuroGeographics** > Voluntary collaboration for the development of the European Spatial Data Infrastructure
 - **UN GGIM Europe** > A very active regional committee!

By Jerker Mostrom from EFGS Conference – 2, 3 November 2017, Dublin 2017

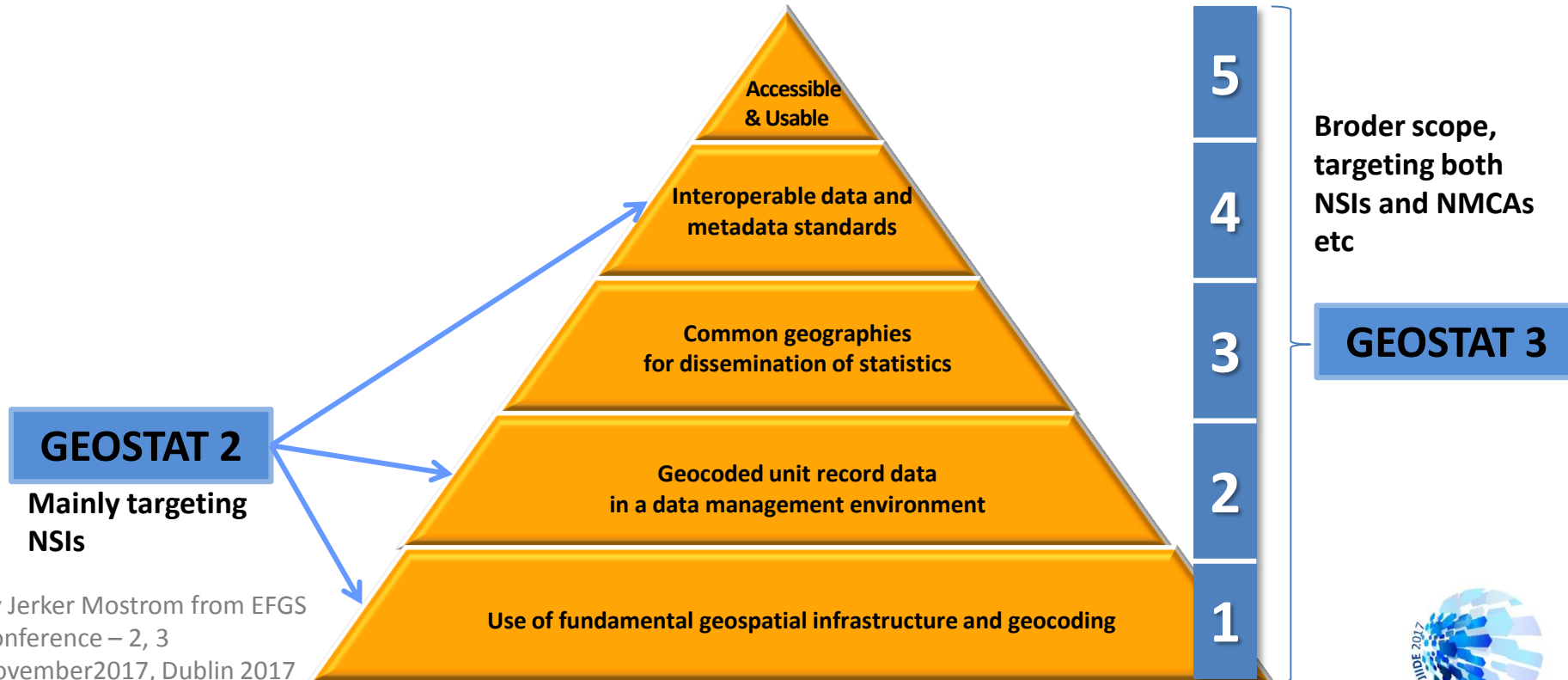


Global Statistical Geospatial Framework (GSGF)

The **GSGF** is ...”is a high-level, framework that consists of five principles that are considered essential for integrating geospatial and statistical information”



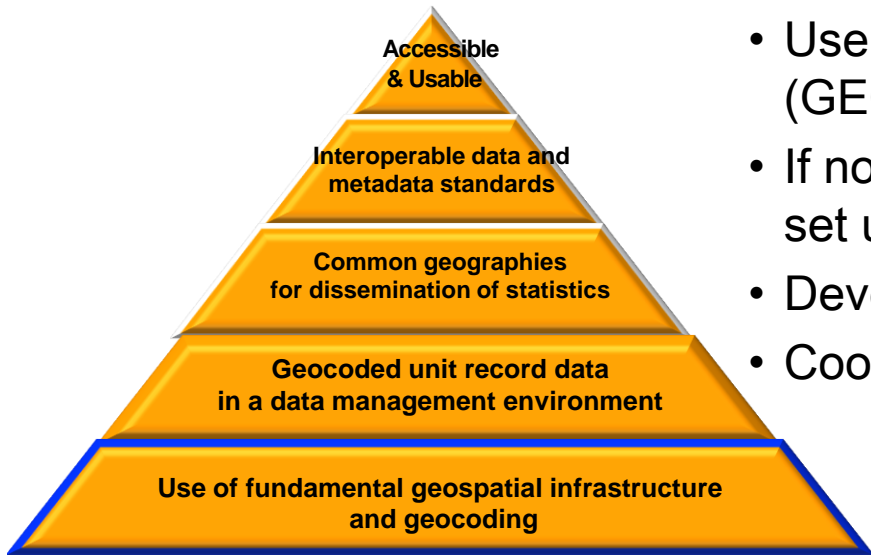
Statistical Geospatial Framework



By Jerker Mostrom from EFGS
Conference – 2, 3
November 2017, Dublin 2017



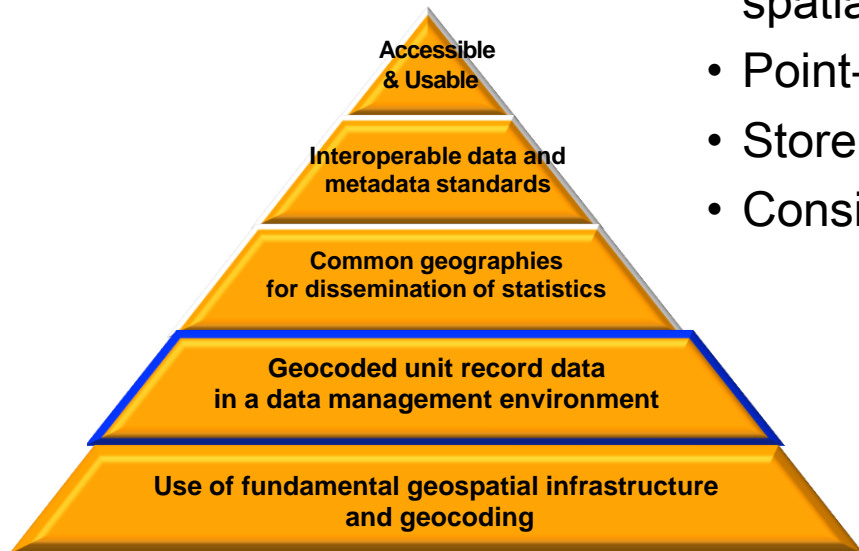
Principle 1: Use of fundamental geospatial infrastructure and geocoding



- Build on data from National Spatial Data Infrastructures (INSPIRE)
- Use point-based location data for geocoding (GEOSTAT 2)
- If no point-based infrastructure exists, it needs to be set up!
- Development of addressing and geocoding services
- Cooperation supported by institutional arrangements



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



- Statistical objects in unit record data correspond with spatial objects in location data
- Point-of-entry validation applied in data collection
- Store location only once!
- Consistent management of temporality of data

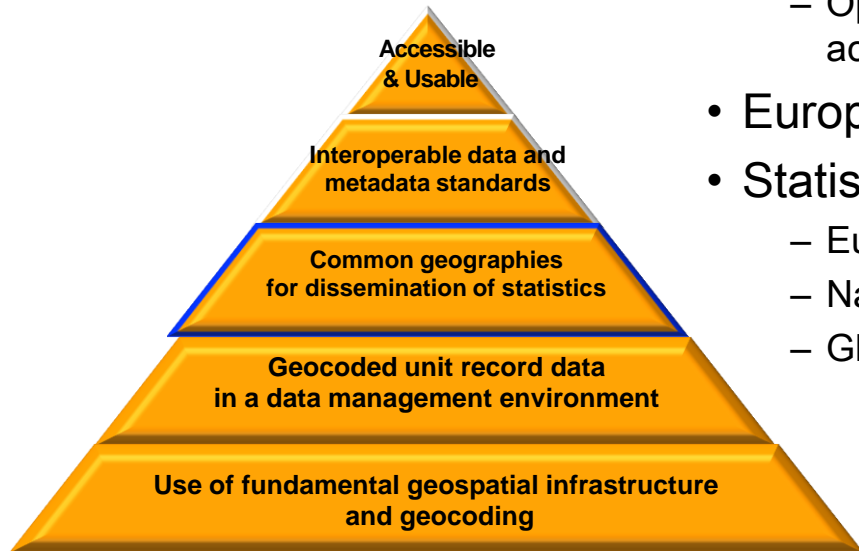
GAs



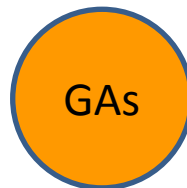
NSIs



Principle 3: Common geographies for production and dissemination of statistics

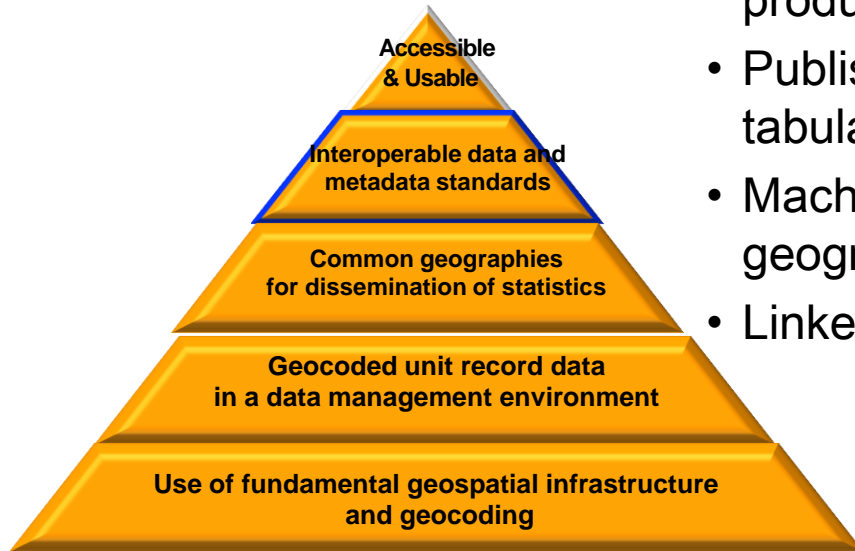


- National statistical and administrative geographies
 - Open access, services, maintenance, coordination, scale & accuracy
- European statistical geographies
- Statistical grids
 - European (INSPIRE)
 - National grid systems
 - Global (OGC Discrete Global Grid System)?



Principle 4: Statistical and geospatial interoperability

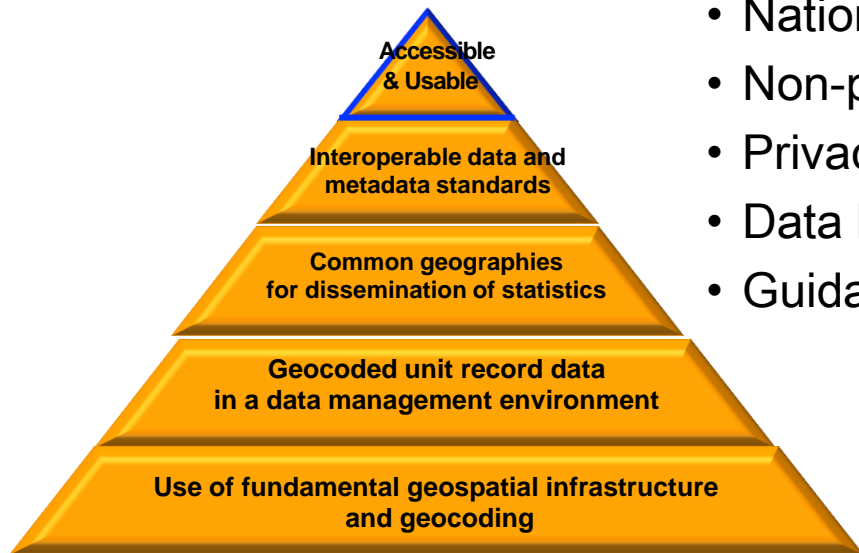
– standards, processes



- Improve geospatial workflows within statistical production (GSBPM)
- Publish once and leave data at its source! (SDMX/ tabular data, OGC services/geospatial data)
- Machine-to-machine services for merging geographies and statistical data
- Linked data



Principle 5: Accessible and usable geospatially enabled statistics



- Map services for pan-European data
- National portals and dissemination platforms
- Non-proprietary solutions for data access
- Privacy issues
- Data licensing
- Guidance on use and analysis



Priority data for spatial reference framework for statistics

First category

Topographic data

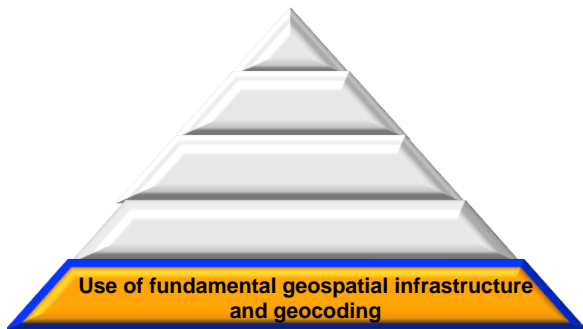
- Detailed transport networks including public transport stops
- Hydrographic network
- Ortho-imagery
- DEM

Administrative data sources

- Administrative boundaries
- Statistical regions
- Census enumeration areas
- Integrated geocoded address, building, dwelling register
- Land parcels (agriculture and estate)
- Cadastral maps

Other data

- Postal code areas



Priority data for spatial reference framework for statistics

Second category

Using the category 1 framework, other data sources may be directly or indirectly geocoded and this way used to produce spatial statistics:

- Sample frame for surveys geocoded to the above address register
- Person register
- Workplace points
- Public services points
- Real Estate Tax registers associated to buildings
- Traffic information
- Other types of public files (tax, registrations, social security files, ...)



Priority data for spatial reference framework for statistics

Third level

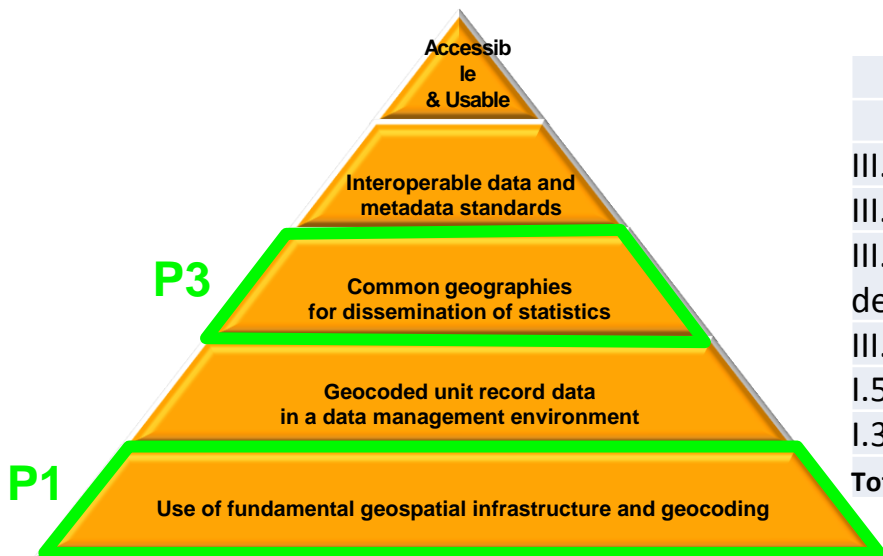
These thematic geospatial data can be used to directly create spatial statistics but also in combination with category 1 and 2.

- Land cover maps
- Protected areas
- Statistics referenced to functional areas (non-administrative or administrative)

For a complete range of spatial statistics products NSIs need access to all three categories of geospatial data sources. However as a first priority access to category 1 should be ensured



How does INE SDI respond to the principles of the Statistical Geospatial Framework?



	2017		
	WMS	WFS/ATOM	CDG's
III.1 Statistical units	40	1	40
III.5 Human health and safety	8	0	8
III.10 Population distribution and demography	0	0	37
III.2 Buildings	4	4	4
I.5 Addresses	1	1	1
I.3 Geographical names	1	1	1
Total	54	7	91



Integration of statistical and geospatial information means:

GSBPM

Specify Needs	Design	Build	Collect	Process	Analyse	Disseminate	Evaluate
1.1 Identify needs	2.1 Design outputs	3.1 Build collection instrument	4.1 Create frame & select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Gather evaluation inputs
1.2 Consult & confirm needs	2.2 Design variable descriptions	3.2 Build or enhance process components	4.2 Set up collection	5.2 Classify & code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design collection	3.3 Build or enhance dissemination components	4.3 Run collection	5.3 Review & validate	6.3 Interpret & explain outputs	7.3 Manage release of dissemination products	8.3 Agree an action plan
1.4 Identify concepts	2.4 Design frame & sample	3.4 Configure workflows	4.4 Finalise collection	5.4 Edit & impute	6.4 Apply disclosure control	7.4 Promote dissemination products	
1.5 Check data availability	2.5 Design processing & analysis	3.5 Test production system		5.5 Derive new variables & units	6.5 Finalise outputs	7.5 Manage user support	
1.6 Prepare business case	2.6 Design production systems & workflow	3.6 Test statistical business process		5.6 Calculate weights			
		3.7 Finalise production system		5.7 Calculate aggregates			
				5.8 Finalise data files			



Integration of statistical and geospatial information means:

1. The process of geocoding statistical and administrative information (micro or aggregated) using spatial reference frameworks
2. Exploitation of geospatial data sources for the calculation of new statistics
3. Processing and manipulation of statistical information using spatial analysis techniques (distances, spatial selection, intersection, aggregation) with the purpose to select information or derive new information with a focus on their spatial characteristics





Integration of statistical and geospatial information means:

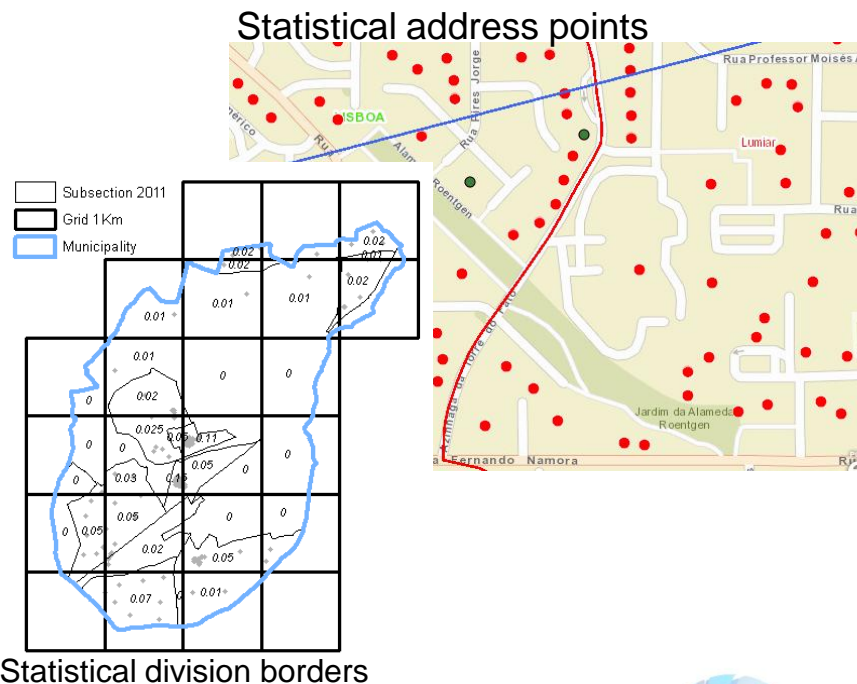
- 4) Supporting a more efficient and flexible statistical production process with geospatial information e.g. for surveying and sampling, field operation
- 5) Combination of statistical end products with geospatial information for dissemination (statistical mapping)
- 6) Improving the quality of existing statistical products

A complete integration of statistical and geospatial information is achieved if location and statistics are just attributes of information objects



Harmonization of statistical and geodesy reference framework

Geodetic System	Layers (suitable for geocoding)	Statistical System
+	NUTS1 - Administrative level 1	+
+	NUTS2 - Administrative level 2	+
+	NUTS3 - Administrative level 3	+
+	LAU1 - Administrative level 4	+
+	LAU2 - Administrative level 5	+
 Cadastral units Cadastral parcels	INDIVIDUAL UNITS level 6	 Statistical regions Enumeration areas



Source: Adapted from Anna Sławińska, from CSOPoland at UNECE/UN-GGIM Workshop on Integrating Geospatial and Statistical Standards, Stockholm, 6-8 November 2017

Conclusions

Follow the project
on EFGS.info

- GEOSTAT 3 Work in progress – final proposal by the end of 2018
- Testing phase during 2018 (SDG indicators and population grid)
- Where desired and possible, the work on the ESS-SGF will be aligned with the work on the Global SGF
- Promote the application of geospatial statistics and the integration of geospatial information into the statistical production process
- Encouraging collaboration between INE and geospatial Community



Conclusions

- A major achievement for the availability of geospatial information needed for geocoding statistics is the INSPIRE Directive, the legal framework that regulates geospatial information in European Union (EU)
- A number of INSPIRE Annex data themes are particularly relevant for the Statistical Geospatial Framework (SGF) such as such as Addresses, Geographical grid, Geographical names, Buildings, Population distribution and demography and Statistical units
- The role of INSPIRE will have to be specifically addressed in the European version of the SGF, which is partially the subject for future developments within the GEOSTAT project.



GEOSTAT 3 - A Statistical Geospatial Framework for Sustainable Development

Obrigada
ana.msantos@ine.pt

