UN-GGIM EUROPE

UNITED NATIONS
COMMITEE OF EXPERTS ON
GLOBAL GEOSPATIAL
INFORMATION MANAGEMENT

Core Spatial Data Theme Statistical Units Recommendation for Content

Working Group A - Deliverable of Task 1.b

Version 1.1 - 2022-05-05

Version History

Version number	Date	Modified by	Comments
1.0	2017-02-14	WG A	Consolidated draft, for review by geographic and statistic community
1.1	2022.05.05	WG A	Comments from geographic and statistic community taken into account Definitive deliverable

Warning: in the following parts of this document, the paragraphs written in grey e.g. "This document has annexes containing more detailed explanations " are common to all core spatial data themes; they aim to provide context and objectives of core data. The paragraphs written in black are specific to core spatial data themes Statistical Units.

Content

1 Executive Summary

In September 2015 the countries of the United Nations adopted the 2030 Agenda for Sustainable Development; a set of goals to end poverty, protect the planet, and ensure prosperity for all as part of a new sustainable development agenda. Each goal has specific targets to be achieved over the next 15 years. The 17 Sustainable Development Goals (SDGs) of the 2030 Agenda are supported by 169 targets and 230 indicators.

Geospatial data supports the measuring, achieving and monitoring of many of the goals and targets set by the 2030 Agenda. The 2030 Agenda demands new data acquisition and integration approaches to improve the availability, quality, timeliness and disaggregation of data. Goal 17 explicitly emphasizes the need for developing capacities and partnerships. In this context the success of the 2030 Agenda depends on senior administrators owning and leading the geospatial efforts in their respective countries.

In Europe, building on the INSPIRE Directive redirecting the focus on a cohesive spatial data infrastructure without gaps in content and discrepancies in quality, stakeholders are working on geospatial standardization and increasing richness of data through Core Data Recommendations for Content that correspond to the first phase of WGA work program. Core data is primarily meant for fulfilling the common user requirements related to SDGs in Member States and European institutions.

Certainly many Core and INSPIRE themes should be seen as statistical units, because one could add statistical information to a cadastral parcel, to a road link, to a lake or to a natural habitat. But the bulk part of statistical production is released by specific geographical areas, named in INSPIRE as Statistical Units, and the geometry of these units is the cornerstone to link geospatial and statistical information.

The Core Spatial Data Theme "Statistical Units" Recommendation for content includes a step wise approach for its adoption by the countries. As first priority, core recommendations include the selection of main categories of statistical units (1 km grid, Territorial Units for Statistics and Urban Statistical Units), the availability of statistical units at various levels of detail and the necessity of a correct topology. As statisticians have to work on time series, tracking changes of the statistical units is strongly encouraged, e.g. through the inclusion of core temporal attributes.

As second priority, **good practices** recommend mainly the capture of other categories of statistical units (smaller grids and enumeration districts units),

Considerations for future lists the main topics for technological development monitoring ; they include the importance of a global grid, statistical units harmonisation for the whole geographic Europe, better tracking of statistical units across time, correspondence between INSPIRE and Statistical classification.

2 Foreword

2.1 Document purpose and structure

2.1.1 Purpose

This document provides the main characteristics of core data for theme Statistical Units with focus on the recommendation for content. This document aims to help decision makers (from governments, data producers, national coordination bodies, etc.) to define their policy regarding the improvement of existing data and production of new geospatial data. It addresses digital data.

This document has Annexes containing more detailed explanations targeting the technical people who will be in charge of implementing or adapting core data recommendations (e.g. for production purpose, as source of other standards, etc.).

2.1.2 Structure

The executive summary synthesizes the main conclusions of the Working Group A (WG A) process and results to develop the recommendation for content. It is meant mainly for high level decision makers.

The foreword reminds the general context of core data, the first step achieved by WG A (i.e. selecting core data themes), and it explains the general principles set by WG A to develop the recommendations for content of core data specifications for all selected themes.

The 'recommendation for content' document itself includes four chapters:

- Overview: it provides the general scope of the theme and describes the main use cases addressed;
- Data content: it provides the main characteristics of the recommended content, such as the list of core features and attributes (for vector data), as well as data capture and quality rules;
- Other recommendations: e.g. Coordinate Reference System, Metadata, Delivery;
- Considerations for future: this chapter addresses some key trends or significant user requirements that cannot be considered as core today but that might be considered in future.

The 'recommendation for content' document is meant for medium level decision makers. It is written in natural and not too technical language.

The technical explanations included in annexes describe the relationship between the recommendation for content and the corresponding INSPIRE specification, and contain any other appropriate information useful for this theme.

2.2 Core data context

2.2.1 Rationale for core data

The following background of harmonised pan-European data was identified.[1]

Authoritative geospatial data are used to support both the implementation of public policies and the development of downstream services. Moreover, geospatial data are required to be homogenous to enable the implementation of public policies in a coherent and coordinated way among countries and at regional or global level. Likewise, significant opportunities exist if services developed by industry can be exploited without requiring country specific adaptation.

The INSPIRE Directive has set up the legal and technical framework for harmonisation of the existing data related to the themes in annexes I, II and III. INSPIRE specifications provide common data models that ensure a first step towards interoperability, however ensuring homogeneous content is outside their scope, as they contain no indication about levels of detail, very few recommendations about quality, and as most features and attributes are “voidable”, i.e. to be supplied if available or derivable at reasonable cost.

This background led the UN-GGIM: Europe Regional Committee to setup in 2014 the Working Group A on Core Data to deal with core data content and quality, production issues, funding and data availability.

Recommendations for content of core data will complement INSPIRE data specifications by defining the priorities on the core content that is encouraged to be made available in Europe in order to fulfil the main user requirements that are common to many countries, with focus on the SDG related ones.

Core data availability may be ensured either through upgrading of existing data when feasible or through production of new data when necessary.

2.2.2 Core data scope

In its first phase, WG A selected core data themes according to the following criteria: core data is the geospatial data that is the most useful, either directly or indirectly, to analyse, to achieve and to monitor the Sustainable Development Goals.

Among the 34 INSPIRE data themes, 14 have been considered as core including theme Statistical Units.

More information about the selection process and results may be found in document ‘Core Data Scope - Working Group A - First Deliverable of Task 1.a - Version 1.2’ on http://un-ggim-europe.org/content/wg-a-core-data

[1] Extract from the Report by the Preparatory Committee on the establishment of the UN-GGIM: Europe Regional Committee, European Commission Ref. Ares(2014)1491140 - 09/05/2014.

2.3 Document objectives and principles

2.3.1 Encouraging content availability

This deliverable provides recommendations for national governments and data producers, aiming to help them to define their priorities for enriching existing data or producing new data. This deliverable is meant mainly for data producers, however it defines the recommended result and target but not the production process.

2.3.2 Complementing INSPIRE

Core data specifications are built upon INSPIRE data specifications. On one hand, they often simplify INSPIRE by selecting core feature types and attributes and by restricting or clarifying the scope; On the other hand, they enrich INSPIRE by recommending specific levels of detail, quality rules and sometimes data model extensions. Besides, the INSPIRE common terminology is thoroughly used for naming core features and attributes.

Regarding the levels of detail, the ELF (European Location Framework) project terminology has been used. The ELF levels of detail are the following: Global, Regional, Master level 2, Master level 1, Master level 0. These terms are defined in the glossary.

Regarding delivery, core data may be supplied according to several ways. It is expected that, very often, the core data recommendations will be used to enrich and upgrade existing products. In this case, core data will be available through these improved products. Core data may also be delivered through INSPIRE conditions (specifications and services).

2.3.3 Status of core data recommendations

This document contains recommendations that are not legally binding. However, some recommendations are more important than others. This order is indicated as follow:

Core Recommendation X

It is first priority recommendation, considered as both necessary and achievable in principle. Ideally, it should encourage involved stakeholders to launch short-term actions (typically within a couple of years).

Core recommendations are usually addressing only technical aspects and are meant for the organisations in charge of producing this theme. The set of core recommendations defines the basic expectations on core data.

Good Practice X
It is second priority recommendation; if adopted, it will provide significant added value to core data; it indicates a relevant trend to be adopted as much as possible. It encourages involved stakeholders to take these recommendations into account in long term, if not possible in short term.

NOTE: some of these good practices may be quite easy to achieve and are already effective in some countries whereas some others may be more difficult to achieve. This is typically the case when these good practice recommendations involve other stakeholders in addition to the organisations in charge of producing this theme, and when they address not only technical aspects but also legal or organisational ones.

A "core data set" should contain the minimum data defined by the core recommendations (and ideally also by the good practices) of this deliverable but may of course contain more and/or better information.

2.4 Abbreviations

CRS	Coordinate Reference System
DGGS	Discrete Global Grid System (Domain Working Group at OGC)
EFTA	European Free Trade Association
ELF	European Location Framework
GSGF	Global Statistical Geospatial Framework.
ITRF	International Terrestrial Reference Frame
ITRS	International Terrestrial Reference System
LAEA	Lambert azimuthal equal-area projection
LAU	Local Administrative Units usually corresponding to the lowest administrative level of a country.
FUA	Functional Urban Area (supra-municipal urban statistics area defined by Eurostat.)
NUTS	Nomenclature of Territorial Units for Statistics of European Commission (Eurostat)
OGC	Open Geospatial Consortium
SDG	Sustainable Development Goal
SU	INSPIRE theme Statistical Units
UNECE	United Nations Economic Commission for Europe
UN-GGIM	United Nations initiative on Global Geospatial Information Management
WG A	(UN-GGIM: Europe) Working Group on Core data

2.5 Glossary

2.5.1 Levels of detail

Global	Level of detail defined by ELF: data to be used generally at scales between 1: 500 000 and 1: 1 000 000, i.e. mainly at international level
Regional	Level of detail defined by ELF: data to be used generally at scales between 1: 100 000 and 1: 500 000; data mainly for national or regional (European or cross-border) actions.
Master level 2	Level of detail defined by ELF: data to be used generally at scales between 1: 25 000 and 1: 100 000); data mainly for regional (sub-national) actions.
Master level 1	Level of detail defined by ELF: data to be used generally at scales between 1: 5 000 and 1: 25 000; data mainly for local level actions.
Master level 0	Level of detail defined by ELF: data to be used generally at scales larger than 1: 5 000; typically, data at cadastral map level, mainly for local level actions.

NOTE: the above definitions are indicative; in practice, detailed data (Master levels) may also be required also by national, European or international users.

2.6 Reference documents

United Nations Statistical Classifications and Standards:
Country and Area Codes:
https://unstats.un.org/unsd/methodology/m49/

Eurostat Standards:
Nomenclature of Territorial Units for Statistics (NUTS):
http://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_CLS_DLD&StrNom=NUTS_2013L&StrLanguageCode=EN&StrLayoutCode=HIERARCHIC

City (Urban Audit) statistics:
http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-17-006

Territorial Typologies:
http://ec.europa.eu/eurostat/web/nuts/tercet-territorial-typologies

INSPIRE Data Specification on SU– Technical Guidelines 3.0:
http://inspire.ec.europa.eu/id/document/tg/su

3 Overview

3.1 General scope

Definition:

Statistical units are the spatial features used for dissemination of statistics. They encompass any geographical aggregation of statistical information.

NOTE 1: The statistical information considered in the scope of this document is a quantitative representation of any fact or phenomenon, according with a documented collection of data, broken down by qualitative or quantitative standard classification, and made by an official or authoritative source or institution.

Examples of this statistical information related to SDGs include: population below an economic level, mortality rate for specific causes, children that attend different levels of school education, material footprint per capita, research and development expenditure in GDP (Gross Domestic Product), urban population in slums, sea floating plastic debris density, surface of forest, victims of intentional homicide or proportion of domestic budget funded by domestic taxes.

NOTE 2: Due to usual practices and confidentiality rules of official statistics, the statistical units are mainly grid cells and polygons. Points and lines could also be used as statistical units, but in very few cases; as a consequence, the point and line statistical units are not considered as core data.

NOTE 3: The Statistical Units are mainly presented in a hierarchical way with different levels, all of them covering the whole area under study and with a relationship one-to-many between a level and its immediate lower one. Most of the Statistical Units at higher level are also Administrative Units, but the lowest levels of Statistical Units may be smaller than the lowest level of Administrative Units (corresponding generally to municipalities).

NOTE 4: The statistical units are geometrical representations of the locations on which statistics are reported. The Statistical Units not properly contain any statistical data or figure, but they are also the bridge to link all the published statistical information to their pertinent geometry. This output link is mainly made by the standard identifiers of statistical units, that are consequently the critical part for a successful association between statistical and geospatial information.

NOTE 5: This definition of statistical units in INSPIRE data specifications and in this document is very different from the usual definition of statistical units in statistical works by National Statistical Institutes, like census, surveys or register based statistics, which focusses more on the unit record on which information is gathered. The classical definition of statistical units in statistical work would be a person or dwelling or a unit of economic activity, which would have a geometry associated to describe their location and is defined at micro-data level. This information could then be aggregated to geospatial output areas, which are defined here as a statistical unit.

3.2 Use cases

'Statistical Units' are the geographic part of a wide range of statistical data. The geography of statistical units is the mandatory bridge that connects the territory and statistical data. Simply combined with basic population information, theme SU provides a location associated with a number of persons and possibly with their characteristics (gender, age, etc.).

Therefore this theme is necessary to assess the number of persons within an area of interest, such as a risk area, a polluted area, the catchment area of a public service and knowing (more or less) the population in any area of interest is key information to make accessibility studies, to assess the population requirements when preparing spatial planning or just when managing a populated area, to report about the number of persons submitted to noise, air pollution, flood risk etc. to the European Commission and it will be also necessary to compute some of the SDG indicators. It is also quite useful for assessing the human pressure on environment and natural resources.

Statistical Units may also be combined with more specialised statistics, such as socio-economic data or human health data, allowing various analyses about poverty, employment, education, health etc.

As Administrative units, NUTS/LAU may be used for money allocation.

Statistical information is a key part of administrative decisions of governments from Local to National and international institutions, using them to analyse problems, taking the decisions, monitoring action's plan and supervise results by all societal and political instances.

The presentation of statistical information by Statistical Units is highly relevant for settlement, urban and regional planning, analysing poverty and wealth inequalities, monitoring health conditions, resources and plans, coping with environmental and social issues, being aware of pressure on natural resources, managing accessibility at public centres, choosing location or change of community resources and also for simulation of different scenarios due to hypothetical conditions or changes.

Statistical information is involved in most of the governments', social and firms' activity and the citizen's overview of any socio-economic decision is plenty of statistical facts and arguments.

In conclusion, the analysis of statistical information by Statistical Units shows the key data to understand nearly any phenomenon at every stages of knowledge approach: historical overview, analytical search, tracking evolution and forecasting future trends. The next image drafts a short summary overview of the main variety of uses cases involved with this theme.

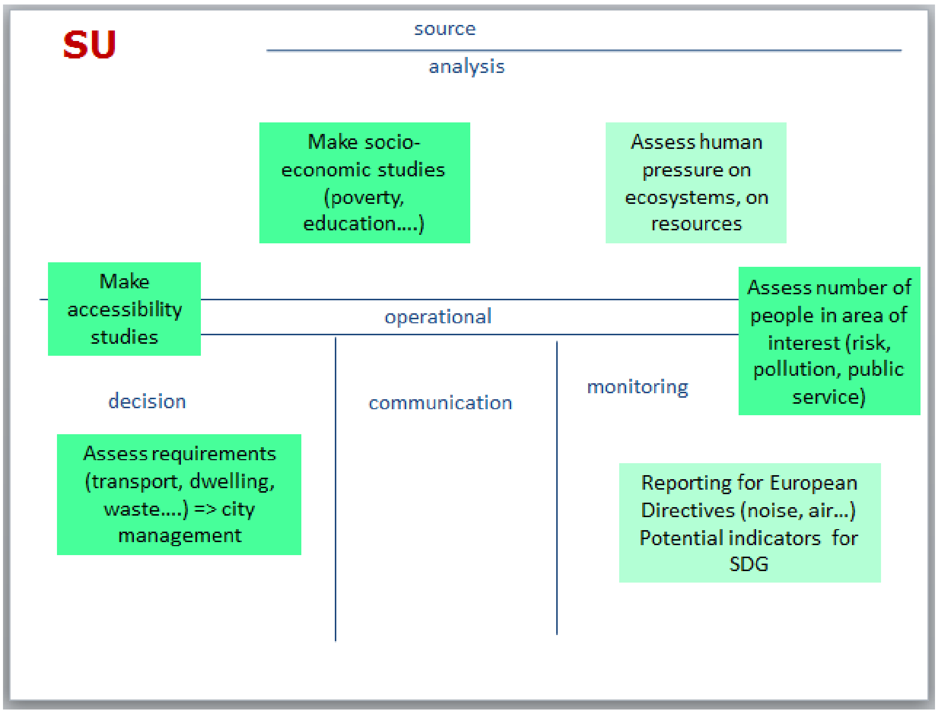


Figure 1: the use case map for theme "Statistical Units"

The GSGF (Global Statistical Geospatial Framework) is a high-level framework that consists of five broad principles that are considered essential for integrating geospatial and statistical information. The GSGF recognizes also the importance of theme "Statistical Units" as one of its principles is related to "Common geographies for dissemination of statistics".

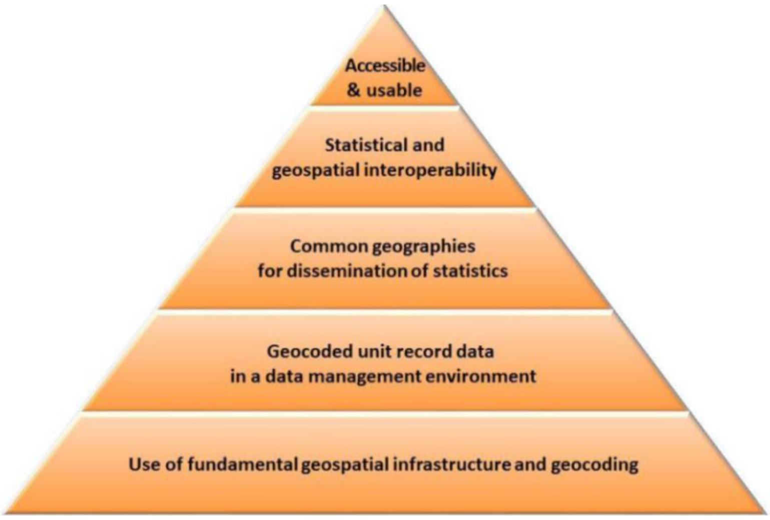


Figure 2: the five principles of the GSGF (the Global Statistical Geospatial Framework)

NOTE: Until now, the location of statistical information is mainly stored by statistical units. However, the Global Statistical Geospatial Framework (GSGF) principles encourage the link between statistical data and geography to be made at the unit record level by storing geometry against the unit record, typically on a point-based framework, at address or building or cadastral parcel level. In this way the data can be flexibly aggregated to any geographic output regions (such as statistical units) or integrated with other geospatial information such as greenspace, sensor information, road networks etc. to compute new added value geostatistical data.

4 Data content

4.1 Features types and attributes

4.1.1 Grid statistical units.

The current trend among the statistical community is to use more and more grid statistical units. The grids offer persistent units across time, enabling to display, on same geometry, long row series of statistical data and so, to show in a reliable way, the trends of any phenomena and the grids could also overcome the Modifiable Areal Unit Problem. In addition, if chosen in an equivalent projection, the grid cell units have the same area, making easier the comparison of territories for a given phenomenon.

Core Recommendation 1
Core content should include the pan-European INSPIRE standard LAEA square grid of 1000 meters side.
Core feature type is StatisticalGridCell with following attributes:
- Geometry
-Unique and persistent identifier

NOTE 1: The optimum grid size depends on the statistical variable and the required confidentiality level. 1km resolution is for example considered an optimal choice for population statistics. In addition, the success of the GEOSTAT 2 project has shown the feasibility of displaying population distribution on a $1km^2$ grid.

NOTE 2: The Core Recommendation 1 applies on the geographic scope of INSPIRE LAEA coordinate reference system, i.e. on continental Europe. For the over-sea territories, it is recommended to use a grid based on an equal-area projection and on an ITRS based datum.

Good Practice 1
Statistical information of urban and high density zones need a grid cell with more detailed resolution square cell, with the following attributes:
- Geometry
- Unique and persistent identifier

NOTE 1: The size of square cell with more detailed resolution is today an open issue, due to the huge gap (in terms of confidentiality issues) between the 100 m size of INSPIRE implicit recommendation and the next step of 1.000m standard. Many statistical offices are producing grids of 200, 250 and 500 meters, and there is not a consensus about the more pertinent standard, if it could be defined for the most useful or even all purposes. In practice, decision about grid size may be taken for a Member State, for a group of countries or as a recommendation from international bodies.

4.1.2 Vector statistical units

Core Recommendation 2
On whole territory, there should be a hierarchical set of Territorial Units for Statistics, under a standard classification.

NOTE 1: Territorial units for statistics are generally based on administrative units and under a standard classification.

NOTE 2: The advantage of territorial units based on administrative levels is that most of these administrative levels are publicly well-known and familiar for the dissemination of national statistics. Territorial units may also be used for operational purposes (e.g. funding of local level government by higher level government).

NOTE 3: There is an existing system of standardised classification of statistical units in most European countries.

- In the case of the European Union, the standard classification is the Nomenclature of territorial units for statistics (NUTS) that includes NUTS1, NUTS2, NUTS3 and LAU levels. Principle of NUTS classification is to have a common set of Statistical Units by level, with the best compromise of equilibrium between maintain the administrative and statistical units of the country and have similar (wide) ranges of population.
- Similar system exist in candidate countries (Turkey, Serbia, Montenegro, Albania and may be Kosovo and Bosnia-Herzegovina that are potential candidates) and in EFTA – European Free Trade Association – (Norway, Iceland, Switzerland, Liechtenstein).
- Outside European Union and candidate countries each country has one or many standard territorial classifications and the objective to establish a common standard territorial classification is marked as a future task, and it’s mentioned in considerations for future, 6.2.1 below.

Core Recommendation 3
In urban areas, there should be a set of urban standardised statistical units.

NOTE 1: As most people live and work in cities, many phenomena have to be analysed in urban areas. This is why it is necessary to maintain a set of urban statistical units.

NOTE 2: In European Union, the set of urban standardised statistical units includes cities, greater cities and functional urban areas [FUA] (see referenced document in chapter 2.6).

Core Recommendation 4
The vector core feature types should have the following core attributes:
- Geometry (as surface or as multisurface)
- Unique and persistent identifiers
- Level of the unit.
- Geographical name (if any)
- Version Id:
- Reference Period Begin
- Reference Period End

NOTE 1: The vector core feature types addressed in Core Recommendation 4 are "Territorial Units for Statistics" of Core Recommendation 2 and the set of urban standardised statistical units of Core Recommendation 3.

NOTE 2: In the INSPIRE data model, there are two identifiers: the thematic identifier and the INSPIRE identifier. Both are useful. For more details, see annex A (7.1.1). The thematic identifier includes the code level and all the codes of previous levels in the hierarchy.

<u>Good Practice 2</u>
The availability of the Enumeration Districts, or people based statistical areas, are considered a good practice with the following attributes:
- Geometry (as surface or as multisurface)
- Unique and persistent identifiers
- Version Id:
- Reference Period Begin
- Reference Period End

NOTE 1: Enumeration Districts are people based statistical units because they are generally inside one stablished bracket of inhabitants.

NOTE 2: The Enumeration Districts, or people based statistical areas, are the most detailed statistical units used by statistical offices for census fieldwork, sample frames, electoral tasks and others, and they also are generally used in the case of statistics based on registers. The Enumeration Districts, or people based statistical areas, offer an optimal balance between respect of privacy rules and display of detailed statistical information.

4.1.3 Temporal aspects

Any description of statistical units at a particular level represents a snapshot of the breakdown of statistical information in a precise moment of time, and so the information about Statistical Units needs to consider the time tracking of changes as an essential part of Statistical Units content. The exception is the Statistical Unit's Grid class; by definition, grid cells have no changes over time.

Therefore, the recommendations below apply for all vector features classes of statistical units, but of course, not for grids.

Core Recommendation 5
Currently valid features are considered as core data.

NOTE: in other words, statistical units of the past are not considered as core data. See 6.4 for additional explanation of tracking the change of statistical units across time.

Core Recommendation 6
Statistical Units should have a harmonised reference date, ideally 31/12/ of any given year. It is recommended for geospatial authorities or other authorities responsible for the definition of statistical units to provide them at the latest 3 months after this reference date to the public.

Core Recommendation 7
Once features have been produced, it is necessary to track the history of all the units and the versioning system is recommended for all statistical units of the territory beyond 2020.

NOTE: The key attributes for tracking the statistical unit over time are "reference period" begin and end, as dates that are the frame in which the statistical units must be used for statistical production and dissemination, linked with statistical information related to the same time frame period. In other words, referencePeriod should ensure consistency between geometry (statistical units) and semantic (statistical information).

Good Practice 3
For a proper tracking of statistical unit's versions the following attributes are recommended:
- beginLifespanVersion
- endLifespanVersion

NOTE 1: Maintaining the version history of statistical units together with the life-cycle attributes in the database allows the interchange of only the changes of these units across any interval of time.

NOTE 2: These attributes are needed for a better tracking of the insertion of changes in the databases or system repositories and they may be different from the begin-end span of the reference of these statistical units. A particular statistical unit could be inserted in the repository preferably before but sometimes after a particular date of official validity. An efficient system of query services has to use the begin/end life span version for tracking the temporal changes from a particular source and, consequently, is a good practice for temporal issues.

4.2 Levels of detail

Due to their definition, the 1 km^2 Grid Statistical Units may be captured only once independently from the level of detail and may be used at a wide range of scales. In opposite, Vector Statistical Units should be provided at various scales, in order to enable an easy use by all levels of governments, from local to global. See 6.3 for related issues.

Core Recommendation 8
Core data on vector statistical units should be captured at large scale (master level 1). When relevant, other levels of detail should be derived from the large scale core data.

Category of Statistical Units	*Master level 1*	*Regional*	*Global*
Territorial Units for Statistics	*X*	*X*	*X*
Urban areas	*X*	*X*	*X*
Enumeration Districts	*X*	*X*	

Table 1: core levels of detail for vector statistical units

NOTE 1: The derivation process consists mainly in the generalisation of the geometry. The generalisation process should respect the topological and hierarchical relations between statistical units and ideally with other themes. A possible way is to use the generalised representation, at the required Level of detail, of the building components of statistical units, i.e. in many cases the administrative units.

NOTE 2: in addition to the generalisation process, it is advised to agree on a common representation of international boundaries at medium and small scales (see core recommendation 9). The edge-matching of statistical units generally depends on the edge-matching of administrative boundaries. It is recognised that getting technically agreed international administrative boundaries is more easily achievable at Regional and global levels than at large scale levels.

Core Recommendation 9
For Regional and Global data, a seamless European dataset of statistical units should be available, based on technically agreed administrative boundaries (with exception of areas under political dispute).

NOTE 1: Seamless pan-European dataset of Spatial Units will be required to display the monitoring of SDG indicators.

NOTE 2: Regarding Territorial Units for statistics, this core recommendation has already been (more or less) achieved through the "EuroBoundaryMap" pan-European product of EuroGeographics. The efforts to maintain such or better product should be continued in future.

4.3 Geographical extent

Core Recommendation 10
Except for urban units, the whole land territory should be covered by the statistical units at each level of detail.

NOTE: For legal or operational reasons there might be few exceptions to the rule of no-gaps, and these exceptions should be well documented in the metadata or in the data specifications of the statistical unit's product.

4.4 Data capture

4.4.1 Geometry

The following rules apply only for vector statistical units; they are not relevant for grid statistical units.

> <u>Good Practice 4</u>
> Great care has to be taken to ensure that the positional accuracy of geographic data reflects the relative position of Administrative Units and of Statistical Units in the real world.

NOTE 1: All administrative and data capture processes must ensure that there is unambiguity between Administrative Units and Statistical Units. Typically, in most countries, several categories of statistical units (e.g. NUTS, LAU, Functional Urban Zones, Enumeration Districts) should share all or part of their geometry with administrative units.

> <u>Good Practice 5</u>
> Great care has to be taken to ensure that geographic data reflects the relative position of cadastral and topographic data (such as roads, rivers, buildings) and of statistical units in the real world.

Example: in real world, the limits of enumeration districts are usually defined by limits or axis of linear elements (like streets, rivers, channels or coastline) or by the limits of polygon features (like forests, lakes, parcels or others). Therefore, in the data, at same level of detail, the boundaries of enumeration districts of theme Statistical Units should share the geometry or should be coherent with the appropriate geographical features, as in the case of the links with Transport Networks, Cadastral Parcels or other INSPIRE themes.

The appropriate coherence and relationship between the Statistical Units and the other geospatial themes (mainly administrative Units) should be produced and maintained by the institution in charge of the geometries of statistical units inside each country.

4.4.2 Identifier

The Statistical Units not properly contain any statistical data or figure, but they are the bridge to link all the published statistical information to their pertinent geometry. This output link is mainly made by the standard (thematic) identifiers of statistical units, that are consequently the critical part for a successful association between statistical and geospatial information.

> <u>Good Practice 6</u>
> Great care has to be taken to ensure that all statistical units have convenient and unique thematic identifiers.

NOTE: More accurately, data producers should check that all statistical units have an identifier (completeness), that this identifier is unique in order to ensure unambiguous identification and efficient link with statistical information and that relevant procedures to ensure the identifier persistency across time, will be put in place by Member State's Statistical Offices. An agreement between statistical and geospatial data producers about coding systems and its custodianship may be a good way to achieve this recommendation.

4.5 Quality

4.5.1 Topologic consistency

Core Recommendation 11
The coverage of the relevant territory by statistical units will be fully topological consistent, with no overlaps between units and with no gaps.

NOTE 1: For urban units, relevant territories are obviously limited to urban areas. For other statistical units, relevant territories are the whole national land territory.

NOTE 2: Some exceptions may occur regarding gaps. As stated in 4.3, these exceptions should be documented. See 6.3 for edge-matching issues.

NOTE 3: In other words, each level of grids, of Territorial Units for statistics and of Enumeration Districts should form a partition of national territory. For urban areas, this recommendation applies – of course - only to the urban part of the national territory.

4.5.2 Temporal consistency

There should be temporal consistency between the statistical unit in the spatial data set and the statistical information collected or disseminated on it. This consistency is ensured by the attributes Reference Period Begin and Reference Period End.

Core Recommendation 12
Great care has to be taken to ensure the reliability of the attributes ‘Reference Period Begin’ and ‘Reference Period End’ (no omission, no error).

5 Other recommendations

5.1 Coordinate Reference System (CRS)

5.1.1 Case of 2D data

Good Practice 7
Core data should be stored and managed in a CRS based on datum ETRS89 in areas within its geographical scope, either using geographic or projected coordinates.

NOTE 1: The geographical scope of ETRS-89 excludes overseas territories, such as Canary Islands or French Guyana or Madeira Islands and Azores Islands. In these cases, it is recommended to use a CRS based on ITRS (International Terrestrial Reference System).

NOTE 2: storing and managing data in CRS based on international datum facilitates the import of measures from modern sensors, ensures that data is managed in a well-maintained geodetic framework and of course, facilitates the export of data into international CRS (e.g. those mandated by INSPIRE).

NOTE 3 : if core data at regional and global levels has to be provided as a single data set on an area including overseas territories, it is recommended to use as CRS geographic coordinates with any realisation of the International Terrestrial Reference System (ITRS), known as International Terrestrial

Reference Frame (ITRF). At small or medium scales, all ITRS realisations can be considered as equivalent, as deviations between them are negligible compared to data accuracy.

NOTE 4: regarding grids, it is recommended to use only projected coordinates, in an equivalent projection (see 4.1.1).

5.1.2 Case of 2.5D or 3D data

Statistical Units for dissemination are expected to be produced as 2D data.

5.2 Metadata

Good Practice 8
Core data should be documented by metadata for discovery and evaluation, as stated in the INSPIRE Technical Guidelines for metadata and for interoperability.

NOTE: This is an INSPIRE recommendation (only the INSPIRE Implementing Rules are legally binding for the Member States belonging to the European Union, but the Technical Guidelines are considered necessary to make the European Spatial Data Infrastructure work in practice). For the other countries, this is a way to make their data easily manageable by transnational users.

5.3 Delivery

It is expected that core data will be made available through improved existing products (or new products) or as INSPIRE data, and perhaps as specific core products (delivery issues still have to be investigated by the working group).

Good Practice 9
Core data corresponding to INSPIRE theme Statistical Units should be made available according to the INSPIRE Technical Guidelines for interoperability, for metadata and for services.

NOTE: this is an INSPIRE recommendation (only the INSPIRE Implementing Rules are legally binding for the Member states belonging to the European Union, but the Technical Guidelines are considered necessary to make the European Spatial Data Infrastructure work in practice). For the other countries, this is a way to make their data easily manageable by transnational users.

NOTE 2: It is reminded that, if there are gaps in the coverage of statistical units, these gaps must be documented. There may be several ways to publish this information. For instance, the recommended information may be integrated in the relevant elements of metadata (such as abstract); it may also be explained in a national statistical unit's data specification document. Ideally, the documentation should be available both in national language and in English.

6 Considerations for future

6.1 Linked data

In order to maximise the usefulness of core data on Statistical Units, it may be advisable to publish it as linked data. The delivery of statistical units, using the technologies of Linked data, would facilitate re-use of statistical data in general by the community of Web developers and so, the development of applications using statistical units.

However, as this technology is still relatively new, more experience and more feedback on costs and benefits of such practice would be useful to support a potential future recommendation.

6.2 Global grid

1km LAEA grid is suitable mainly for European scale applications, but for global applications, it is pertinent to define another grid at global level or geographical grid based systems.

Recent global efforts have culminated in the development of a Discrete Global Grid Systems (DGGS) standard, which has been developed under the auspices of Open Geospatial Consortium (OGC). This system offers further options in the use of grids within the context of the principle of common geographies and in geospatially enabled statistics.

The goal of DGGS is to enable rapid assembly of spatial data without the difficulties of working with projected coordinate reference systems. The OGC DGGS Abstract Specification standard defines the conceptual model and a set of rules for building highly efficient architectures for spatial data storage, integration and analytics. In other words, the OGC DGGS describes a universe of valid grid systems. Valid grids could be used as “interlingua” between grids at different geographical levels.

However, there are 18 topics that have to be carefully investigated before defining DGGS conformant grids. Among them, for instance, the principle of cells of equal area that needs to be under scrutiny. In addition, many use cases are combining grid data with vector data, such as administrative units, transport networks, etc. Will the use of a DGGS conformant grid make this data combining easier or in opposite more difficult or even impossible?

For the future recommendations in this field, it is very important to follow the current OGC activities on the topic of Discrete Global Grid Systems and to assess the benefits and drawbacks of DGGS grids.

6.3 Improving the data harmonisation in whole geographic Europe

6.3.1 European standard classification of Territorial Units for statistics and urban areas

This document recommends a set of hierarchical levels of Terrestrial units for statistics that should be under a standard classification. As stated in 4.1.2, there is a pan- European classification applying in the Member States of European Union and some similar system in candidate and EFTA countries. This standard classification ensures that Territorial Units of same level represent the best compromise of equilibrium between maintain the administrative and statistical units of the country and having units with similar (wide) ranges of population. However, the recommendations for content of this document apply for whole geographic Europe.

Regarding the countries that are not under the NUTS or similar classification system, it should be convenient, under the UNECE umbrella and with the support of EUROSTAT, to establish a comparative standard classification of Territorial Units for the geographic Europe, and maybe also for the UNECE region. This future task needs to build the correspondence table between national classifications and NUTS. The same aim should be achieved for urban and city territorial units.

The main objective of these enlarged territorial statistical units classifications is to produce, display and track the SDG statistical indicators, by facilitating comparison between countries, sub-national levels and urban areas.

6.3.2 Other topics

In its report about "A European Statistical Geospatial Framework", the GEOSTAT 3 project has identified harmonisation issues between Member States. These issues are mainly related to inconsistencies in cases where the LAU level is not the lowest administrative level for building statistical geographies and to unclear handling of water areas (inland, estuaries and sea).

These issues might be solved just by Member States documenting their practices (e.g. in metadata) or by Member States using the same practices. The first solution looks the most appropriate regarding the methodology to build LAU units as it is mainly depending on the administrative structure of each country. The second solution might be relevant for handling of water areas; however, there should be some coordination – ideally on whole geographic Europe- to agree on common rules and/or on a common methodology to compute the land area of a statistical unit.

It might also be useful to have on the various categories of core statistical units a set of attributes indicating their typology (e.g. coastal/ not coastal, degree of urbanisation) as this would facilitate the disaggregation of most statistical variables. Once again, this information might be added in future to core content under condition that there is an agreement – ideally on whole geographic Europe - about the most useful typologies and about clear and persistent rules to define them.

In summary, this paragraph encourages coordination of relevant stakeholders to enhance the harmonisation of statistical units on whole geographic Europe.

6.4 Getting pan-European data – edge-matching and geometrical harmonisation.

This document recognises the need, at least for Global/Regional levels, of a seamless European data set of statistical units. Regarding Territorial Units for statistics, this recommendation is more or less achieved through the EuroBoundaryMap (EBM) product of EuroGeographics that ensures seamless mosaicking of national contribution and consistency with theme Administrative Units.

However, the production process of EBM results in long delays in data availability, this delay being considered by users as a significant issue. Therefore, data producers of this product are encouraged to investigate the possible ways to speed up the process. In addition, there is no similar initiative regarding the other types of vector statistical units, such as urban areas or enumeration districts. Significant improvements could be achieved by adopting an incremental updating approach. In such approach, a central data repository is accessed by each member state to update (maybe continuously and automatically) the data related to their territories. This production model is already in place in several Member States between local and national levels, and should also be applied between national and European level.

Regarding grids, there is no issue on the geographic data of statistical units: it is quite easy to edge-match the national recommended 1km^{2}grids or to compute directly a European grid. The users must be aware with the attribution of statistical information coming from two countries on the same grid cell, because the definition or production process, the confidentiality controls or suppression of values may be slightly different and should be clearly documented.

6.5 Historical data on statistical units

This historical data may be required in many use cases. The first one is just to get statistical units data at a given period of the past in order to be able to combine it with any statistical information of the same period of time. Core recommendation 7 encourages data producers to track all features beyond 2020 but there may be use cases requiring data from the past. For instance, a user may require the statistical units of year 2010 in order to make a thematic map displaying the population or employment information of 2010.

Costs-benefits of keeping the historical data of statistical units available have to be assessed and if the results are favourable, data producers should be encouraged to ensure this availability for past units across a length of time back as long as possible and useful.

The second case for general users is displaying statistical data on fixed geometries across time. Its purpose would be to display on a fixed unique geometry the statistical information of a long time period. For instance, a user might be willing to display a statistical variable (e.g. employment rate) for years from 2000 to 2020 on the statistical units of year 2000. The use of unique geometry across time would of course facilitate analyse of evolution of the statistical variable and the comparison between several years, but in some cases is extremely difficult or simply impossible. This task needs to be achieved in two steps.

First step, the user should be informed about the differences between statistical units across time, including changes of shapes for the same unit, changes of identifiers if any (the recommendation about persistent identifiers avoids them), splitting units, aggregate units and, even more important, mixed combined changes that combines some of these simple changes. Core Recommendations 4 and 7 and Good Practices 2 and 3 are the base for the purpose of a precise time tracking. However, to enable users to understand the changes in statistical units across time, it might be required to include, in the data management system, information about the evolutions.

The second step is about finding relevant procedures to transfer statistical information, taking into account the differences of geometries across time (adding areas with changes in common polygons or estimating the share parts for different units, or mixed procedures). This second step is outside the present SU Recommendations, because this step requires to develop methods of estimation to transfer the statistical variable according to its intensity inside any statistical unit. This task nearly always needs the use of some auxiliary variables, the appropriate statistical or geostatistical methods and a clear knowledge of the particular behaviour of any phenomena and corresponding statistical indicators across space. Due to that circumstances, this second aspect is out of the scope of this SU Recommendations, but is a clear consideration for the future that users comparison across time will be easier if the statistical offices publish geographic correspondences which use the distribution of common dwellings, population, employment or other variables to measure and quantify the changes.

6.6 Bridges and convergence between INSPIRE and Statistical Standard classifications

The general scope of core theme Statistical Units is broadly the same as the scope of the INSPIRE theme Statistical Units; the difference is that, unlike INSPIRE, it is not limited to existing data.

The conceptual model of INSPIRE objects does not allow the full linking of statistical information and geospatial data, because the INSPIRE definition of real world objects and activities is very disparate with the standard statistical classification of economic activities and statistical units of information. Even when the relationship between INSPRIRE and Statistics Classifications could appear as slightly similar, as in Agricultural and Aquaculture Facilities [AF], Production and Industrial Facilities [PF], Transport Networks [TN] or Utility and governmental services [US], the INSPIRE definition is very narrative and with no mention to any statistical classification. On the other hand, the statistical units of economic activity also have critical difficulties to locate at a particular place their conceptual approach, mainly linked with the classification with a firm's production and the complexity associated to the groups of enterprises or conglomerates.

Those issues are very well known methodological differences, as a particular case of differences of cartographic and statistical objects, and some effort have been spent during past years for narrowing the gap and develop the appropriate interlinks between them. Anyway more effort will be needed in the future in this area. May be a first attempt of a table of correspondence between INSPIRE themes AF, PF, TN and US and the statistical classification of economic units and activities could be a good first step in this path.

Another issue is the contradictory status of Population and Density (PD) theme of INSPIRE, due to the redundancy with Statistical Units, deficient treatment of time variable and limited available statistical information. May be, the INSPIRE services could take advantage of statistical information services with a deeper use of SDMX available statistical services, with the use of Table Joint Services tools and/or with other statistical services facilities. It could drop the additional burden of developing limited and ad-hoc INSPIRE services (like PD) with data referred to previously defined Statistical Units. More efforts will be needed in this field in the future.

7 Annex A: Relationship with INSPIRE

7.1 Data model

The UML models provided in this annex are only graphical illustrations of the core recommendations and of the good practices present in this document.

The recommendations for content are represented by highlighted the selected attributes in the following way:

Core recommendation

Good practice

7.1.1 Comparison between Core Data and INSPIRE content

7.1.1.1 Grid statistical Units

Core recommendation 1
Core feature type is StatisticalGridCell with following attributes:
- Geometry
- Unique and persistent identifier

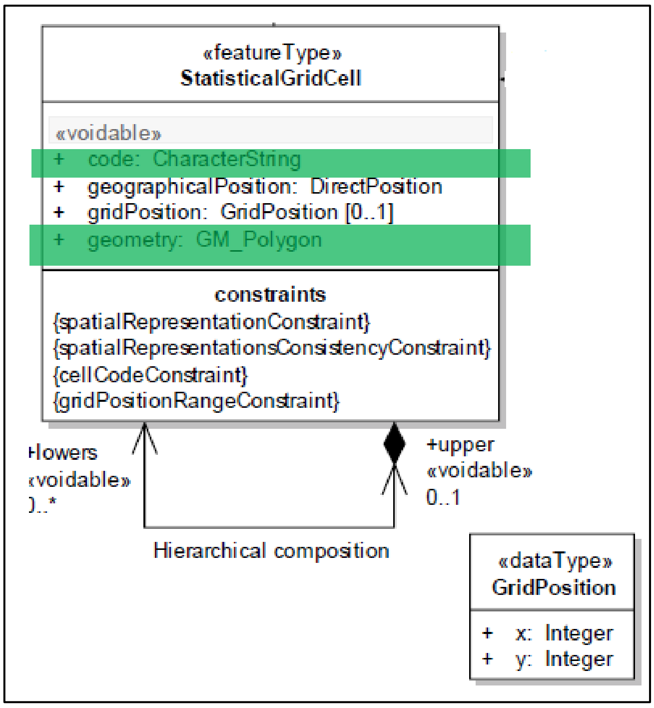


Figure 3: The statistical grid data model

7.1.1.2 Vector Statistical Units

Core recommendation 4

Core Data of feature type of Vector Statistical Units have the following attributes:
- Geometry (as surface or as multisurface)
- Unique and persistent identifier(s)
- Level of the unit.
- Geographical name
- Version Id
- Reference Period Begin
- Reference Period End

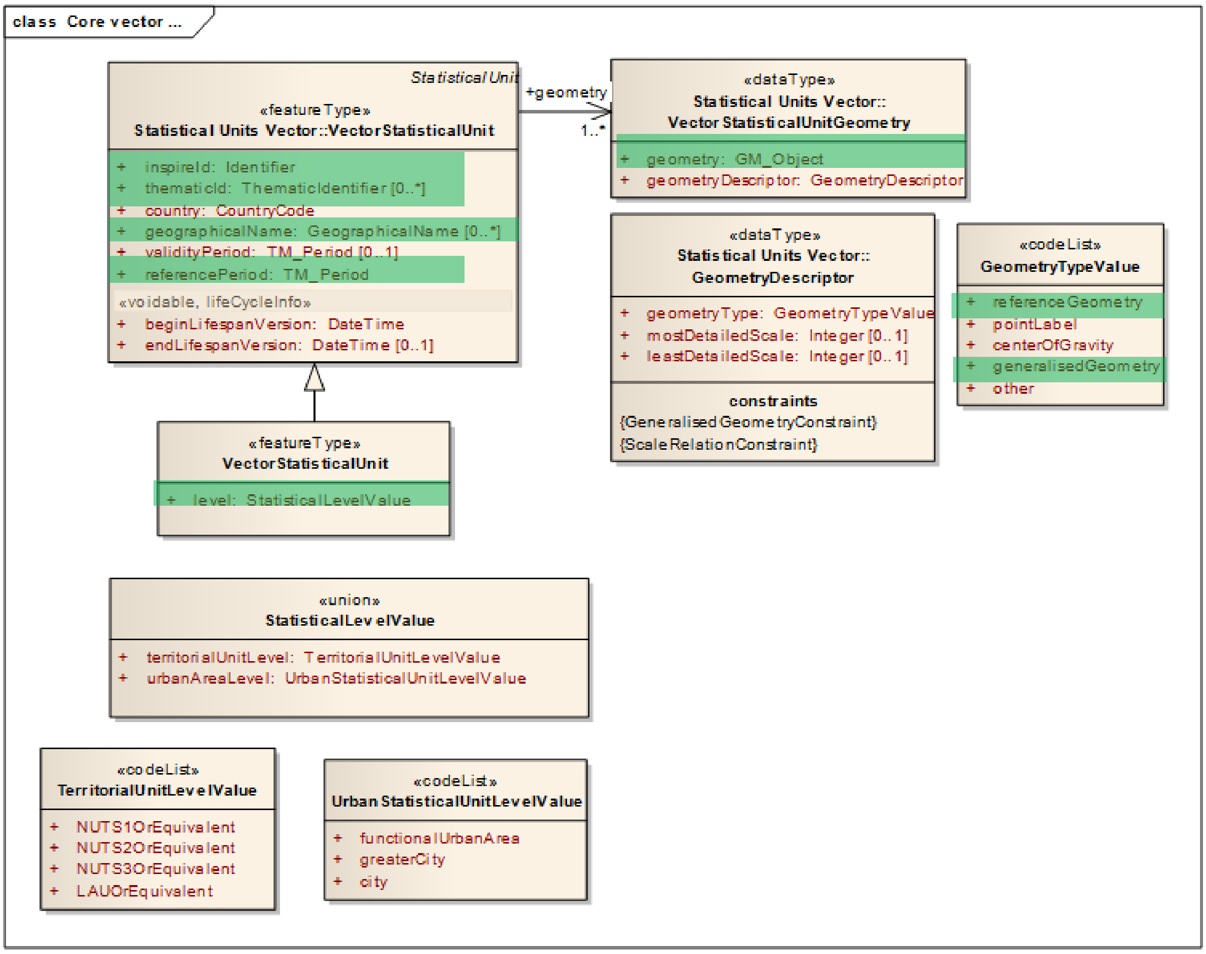


Figure 4: UML data model for core vector statistical Units (extended from INSPIRE data model)

NOTE 1: country code

This is a mandatory attribute in the INSPIRE data model but there is no strong need for Member States to capture it at feature level. It may be generated automatically when publishing national data for INSPIRE, by attributing a constant value (e.g. FR, NL, UK… respectively for France, Netherlands, United Kingdom …) to all features of the national data set. Of course, it is also of interest to generate this attribute for dissemination in other cases, especially for use at European or Global level.

NOTE 2: level of the unit

This attribute is not present in INSPIRE data model; it is considered as core attribute, because the level of the unit is a key piece of information in the case of Territorial and Urban units: producers could disseminate more consistent information and users could easily understand the structure of territorial units. The INSPIRE data model has been extended.

NOTE 3: temporal attributes in real world

INSPIRE Data Specification of Statistical Units (p.40) includes 2 mandatory attributes:

- The "referencePeriod" defined as "The period when the data is supposed to give a picture of the territorial division in statistical units".
- The "validityPeriod" defined as "The period when the statistical unit is supposed to be preferably used and not".

The definition of "referencePeriod" is not very clear, but it's understandable. We explain it as the time period of a particular statistical unit shape that must be used for statistical production and dissemination, linked with statistical information related to the same time frame period. In other words, referencePeriod should ensure consistency between geometry (statistical units) and semantic (statistical information).

In opposite, the INSPIRE definition of "validityPeriod" is difficult (or impossible) to understand, at least for the statistical community and therefore has not been selected for core data content. From the statistical point of view once a version of statistical unit is being used, it will always exist for a precise period of reference of data. The concept of "validityPeriod" is difficult to catch and all statistical units are "valid" in a particular reference period for ever, because any statistical producer could publish "old" data with their correspondent "old" statistical units without any "validity" constraint.

Anyway, in some parts of the same document (and mainly in the informative Annex F, p. 114-116) it seems to use Validity period as the most common and clear term instead of reference period. Annex haven't the same priority as core text, but the internal comprehension difficulties of this INSPIRE document in a key point of the Data Specifications of Statistical Units, don't allow us to know if our proposals are fully in line with the INSPIRE specifications or not.

This is why only the referencePeriod is considered as relevant information and as core attribute.

7.1.1.2.1 Identifier

INSPIRE makes a distinction between the external (or INSPIRE) identifier of the feature that has to be unique and persistent (called also primary key in database syntax) and the thematic identifier that is used to ensure the link to another dataset (called also secondary key in database syntax). Typically, for the statistical units, their full code is the thematic identifier ensuring link to statistical information whereas the INSPIRE identifier is a database identifier. In other words, for statistical units, the thematic identifier aims to identify the real-world object: a given statistical unit will have the same thematic identifier in different databases, typically in geographical datasets at different scales and in datasets of statistical information. In opposite, the INSPIRE identifier identifies the feature in the data set; as a consequence, different representations of the same statistical unit in different datasets will have different identifiers.

The thematic identifier is clearly the key attribute for statistical units as it enables the link with statistical information and a wide set of use cases. Without this attribute, statistical units would just be useless.

However, the external (or INSPIRE) identifier has also its interest as it enables to uniquely identify each representation of vector statistical units, at various levels of detail (e.g. Master level 1, Regional, Global). This unique identification of each feature is a mandatory requirement of INSPIRE and should be provided. In addition, in the INSPIRE data model, it is the attribute carrying the versioning mechanism.

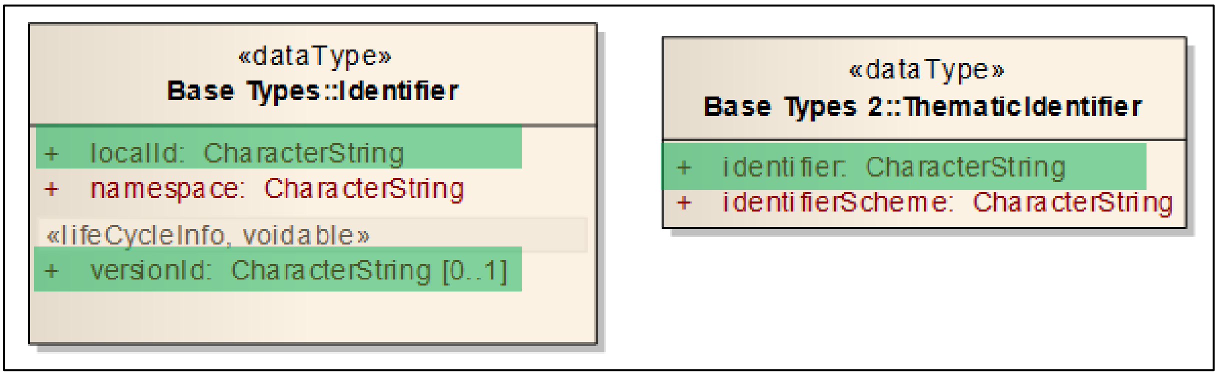


Figure 5: The two types of identifiers in the INSPIRE data model

NOTE 1: As for the country code, on the territory of a Member State, the namespace and identifierScheme are constant values that don't need to be captured at feature level and may be added automatically for dissemination of data, especially if data is dedicated to European use.

NOTE 2: Definition of "persistence" may differ according to various points of view. As a general rule, a minor change should conduct to consider it is still the same object (keep the identifier and change the version) and a major change should conduct to consider it is a new object (change the identifier) but what is minor or major change is subject to various interpretations.

The traditional way of work in statistical systems, not usually considering the geospatial measurements of statistical objects, is that a statistical code is persistent if its textual description (code itself and may be name of the unit) is the same.

As a general rule, National Statistical Institutes should be responsible to define and apply the persistency rules of Statistical Units thematic identifiers whereas the external (INSPIRE) identifier and its versioning mechanism should be used (mainly) to manage the changes in the geographic representation of Statistical Units.

7.1.1.2.2 Temporal aspects

<u>Good practice 3</u>

Core Data of feature type of Vector Statistical Units additional recommended attributes:

- Begin life-span version
- End life-span version

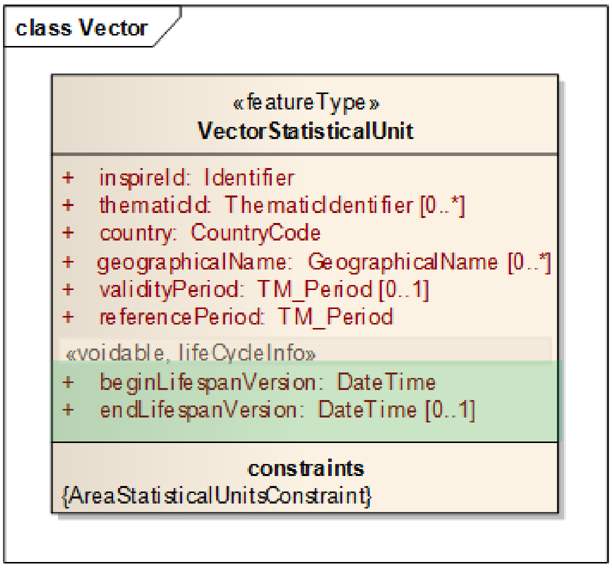


Figure 6: life-cycle attributes from the INSPIRE data model

7.1.2 Alternative implementation data model

7.1.2.1 Level as code list or as children feature types

For each category of vector statistical units, it is recommended to document the hierarchical level of the unit. This may be done by using an attribute (as in figure 4). This might also be done using children feature types. Find below an example in the case of Territorial Units for statistics.

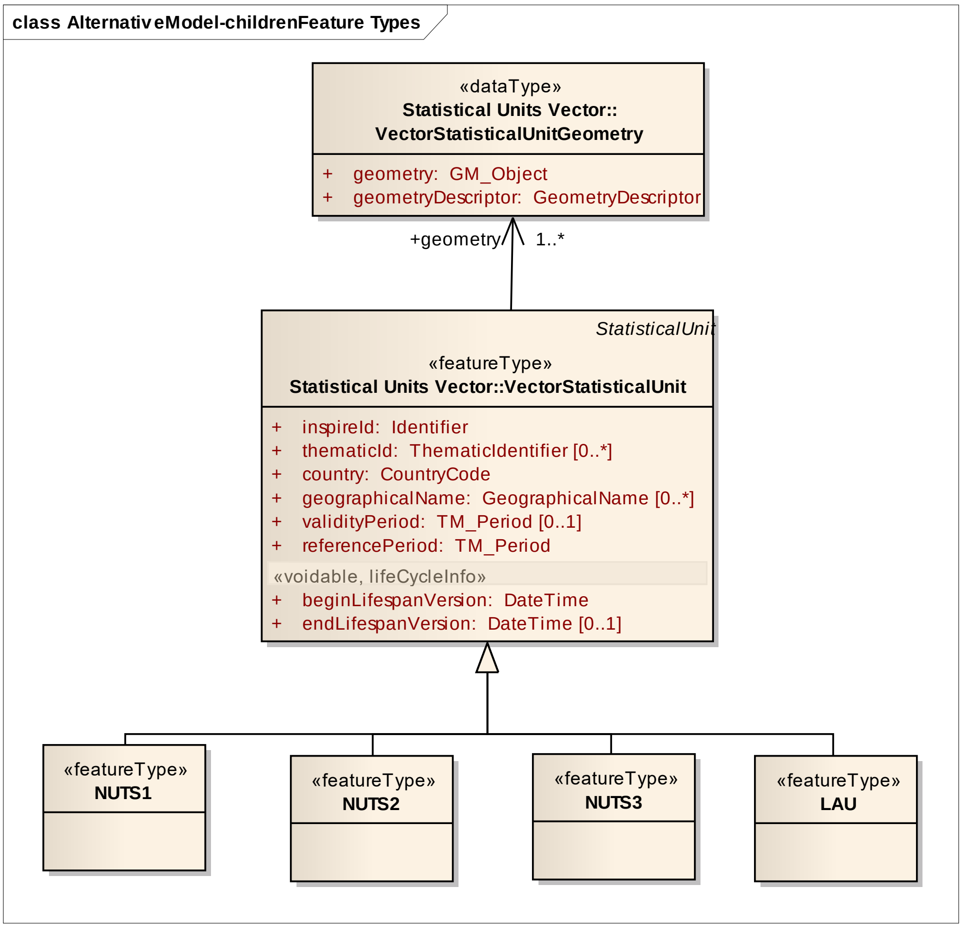


Figure 7: alternative data model to document level of Territorial Units for statistics

7.2 Other

7.2.1 Scope

The scope of INSPIRE is quite wide as it includes all statistical units with the exception of the features that are already included in other themes. In opposite, the core recommendations for content have limited the scope to a few categories of statistical units, considered as the most useful: grid cells, Territorial Units, Urban Zones and Enumeration Districts.

7.2.2 Levels of detail

There is no indication in INSPIRE regarding at which scales statistical units should be captured. Core data recommendations clarify which levels of detail should be produced in priority.

7.2.3 Quality

In the INSPIRE data specifications, there are a few rules related to generalisation and to topology, including edge-matching. The “Recommendations for content” of core spatial data theme Statistical Units have selected the most relevant rules, according to the type of statistical units and to the levels of detail.

In addition, the current document also recommends some other rules, such as taking care of identifiers and of reference dates.

8 Annex B: Methodology

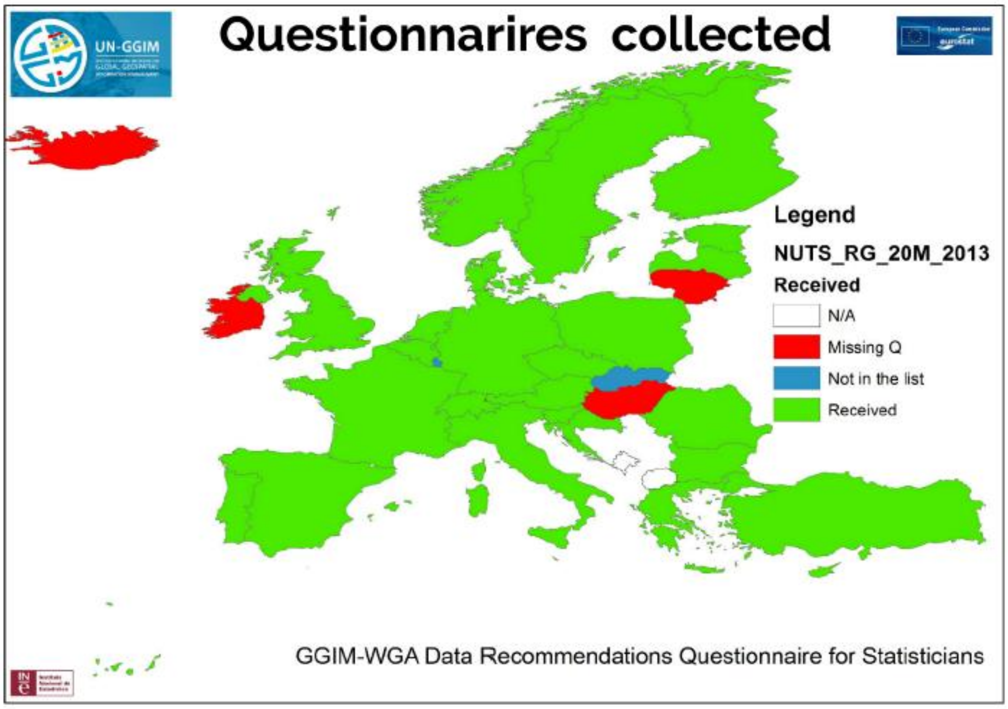
Core data specifications have been elaborated based on one hand on user requirements (with focus on the ones related to SDG) and on the other hand on INSPIRE data specifications.

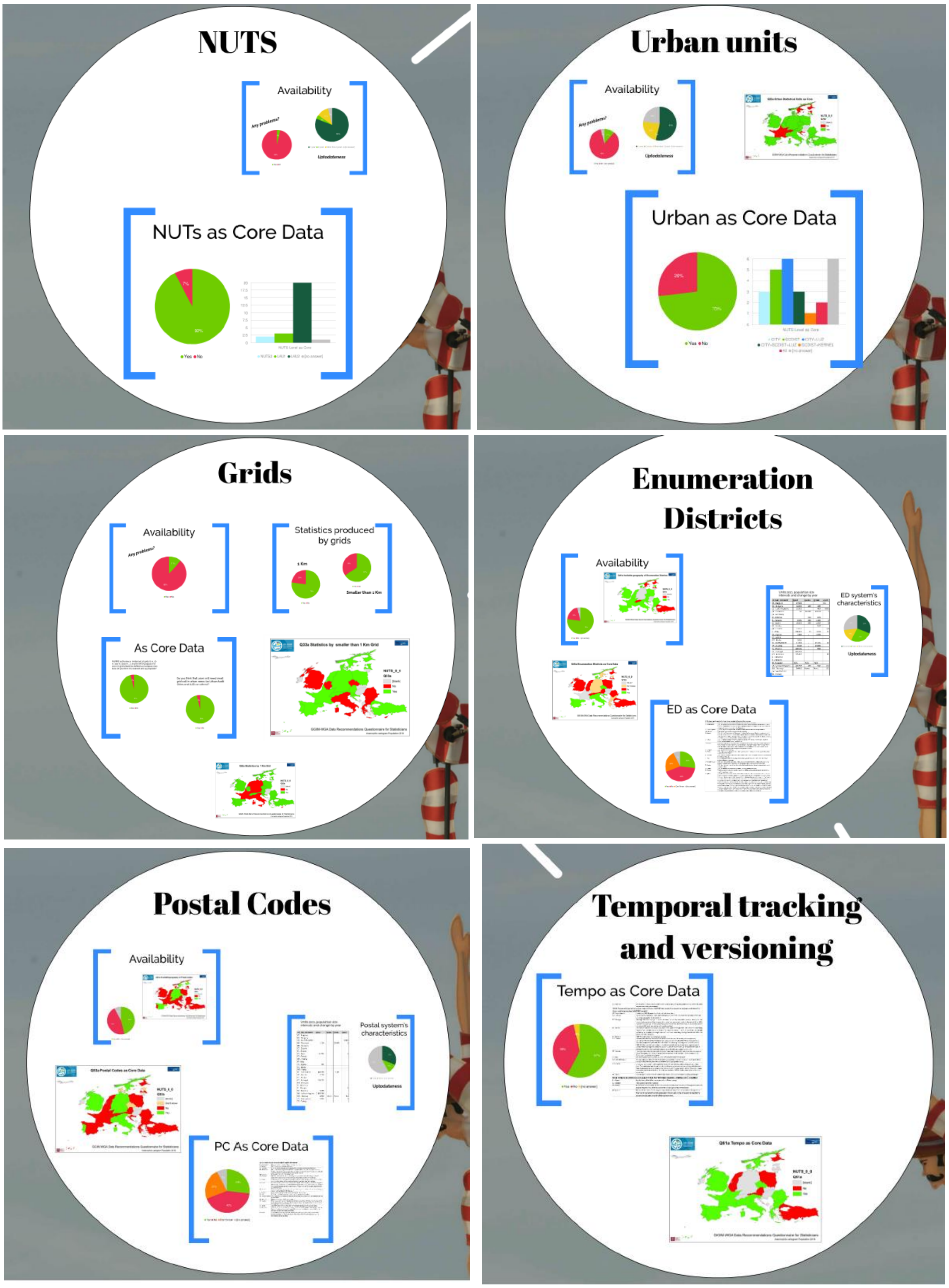
The first step was a deep review of INSPIRE data model and a consultation with the statistical experts in this field.

Due to the specific subject and the background of WGA Members a consultation to geostatistical experts of National Statistical Institutes was decided. The Questionnaire for Statisticians was designed during June-September of 2016 and presented at the European Forum of Geography and Statistics 2016 Conference in Paris (15-17 of November 2016). The questionnaire collected information about the main statistical units: their availability, up-to-dateness, scale and quality issues and if each national expert considered any statistical units as core data, according the GGIM context and aims. NUTS, Urban Areas, Grid of 1Km side and smaller, Enumeration Districts and Postal Codes were surveyed in detail. Temporal aspects, metadata and quality issues were also included and other proposals or comments were asked for.

The Questionnaire was sent to the Eurostat GISCO contact point of Member States and other countries and the answer was collected from December of 2016 to February of 2017. After the collection stage a document of Synthesis was compiled with the aggregate results of pre-defined answers and the list of the qualitative answers and other comments. The main results were presented at GISCO 2017 meeting (Luxemburg, 30-31 of March 2017). The information collected is very rich from the qualitative point of view and the total results were very useful for the present recommendations.

The following images show some results of this key point for drafting the Recommendations for Statistical Units.





Postal codes were viewed in principle as a good candidate for good practice. After the review of current practice in EU Member states, the main difficulty arose due to the strong disparate size between countries. Secondly sometimes the access to the information and updates is difficult, considering they are managed in some countries by semi-public bodies or the private sector companies. Finally Postal Codes was not considered as manageable for a pan European good practice of statistical units and this information might be added as an attribute to the Address Core data.

In addition WGA have collected user requirements from other sources (Eurostat geospatial projects, GSGF) and activities (meetings at Eurostat, European Forum for Geography and Statistics Conferences, etc.) that have provided useful inputs for good practices and considerations for future.

9 Annex C: statistical information in historical changes and context

From the point of view of the history of statistical production, authoritative and reliable data have been always associated with the main transformations of societies and countries in the field of population, wealth, health improvements and economic development. Despite different crises and troubles the production of new, more complex or more detailed data was the permanent companion of any new effort to change societies or improving economic performance. After the Second World War, with more mature statistical systems this tradition was mainly developed under different projects of social and statistical indicators, as the final product of a strong statistical production with pertinent measures that encompass many relevant aspects of the social, economic and environmental reality.

The past decades have seen many different attempts to build a knowledge framework based on a common agreement from techno-academic, trans-political and trans-governmental goals which allows the development of a set of socio-economic indicators. Due to the international arena in which the discussion is developed, the goals established should not to take sides between the contending theoretical approaches and the difficult debate over the historical socioeconomic causes of current problems, such as poverty, wealth inequalities and the overconsumption of natural resources. But this is not the only deliberate oversight that the goals must cope with. Some of the issues not taken in depth are: the hypothetical contradiction between different goals, mainly between natural resources conservation and development measured by monetary figures like GDP, but also between socioeconomic performance and an enlarge approach to people freedom and capabilities, or the critical analysis of the reasons for unsuccessful past initiatives about reducing social and economic disparities between countries and people.

Nevertheless the United Nations Goals are the lone transnational initiative that could take the attention about our common destination as human beings, and could remember to lead governments, societies, civil groups and regular citizens towards the importance of agreeing affordable targets for caring for natural resources, improving living conditions of the people and developing strong institutions to face the big challenges of our era.

The UN 2030 Agenda for Sustainable Development in September 2015 is presented as the latest ambitious, integrated, indivisible and transformational global agenda with 17 Sustainable development Goals, 169 associated targets with 232 indicators promising to achieve sustainable development in its three dimensions – economic, social and environmental – in a balanced way. Geospatial and statistical data need to support, measure, achieve and monitor several if not all goals and targets set by the 2030 Agenda.

The difficulties come mainly from the fact that statistics has to continue the current production, face the increased difficulties of accessing to informers and data sources, and the fact that it is governed by separate institutions and tools, mainly by the statistical plans at country, region and global level. Any new or more detailed data is costly in terms of resources and time for implementation at international level, and has to compete with a lot of other well stablished demands for information.

The GGIM initiative has pointed out the relevance of core geospatial and statistical data for the UN Goals. The goal of the related UN-GGIM-Europe work group on core data was to select which geospatial data are the most important for supporting the Sustainable Development Goals, choosing from the list of the INSPIRE annexes, which is the most relevant and only binding framework for the sharing of geospatial data across borders in Europe. Statistical Units definitely belong to this list of Core Data, because the statistical units are the necessary hinge between the geospatial and statistical world of information.

Anyway the history of statistical production shows us that the improvements of wide and pertinent statistical indicators should be associated to the progress of wellbeing and equitable wealth of people and societies, throughout an open discussion about evidence-based practices.