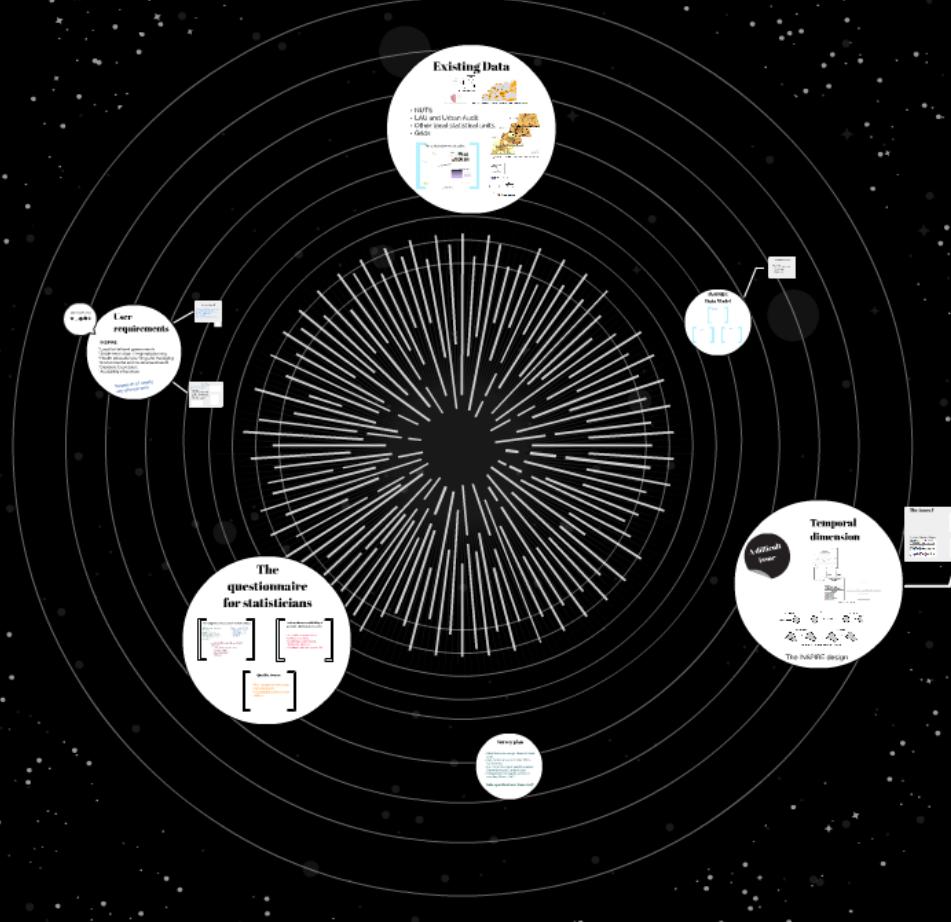


# A questionnaire for collecting opinions for Statistical Units Data Specifications

Questionnaire  
for statisticians



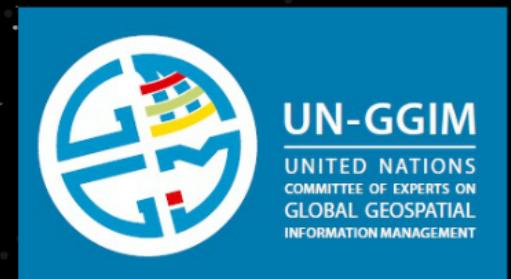
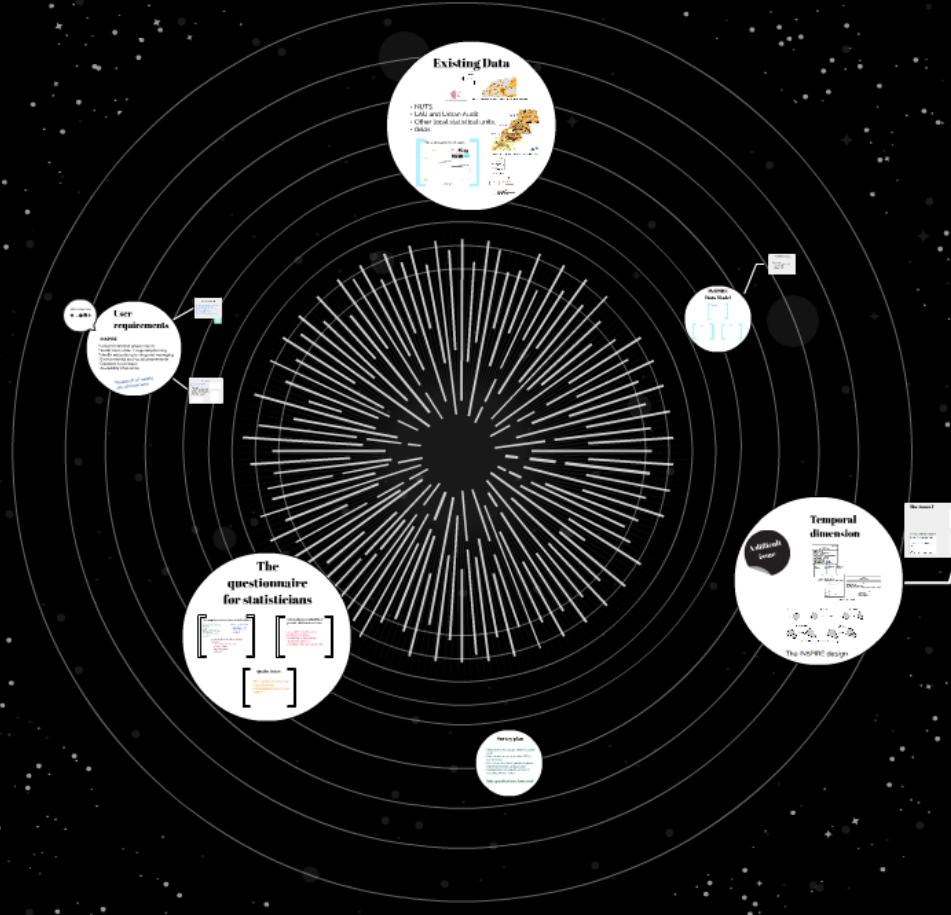
**GGIM WG-A Core Data  
Data Specifications**



**UN-GGIM**  
UNITED NATIONS  
COMMITTEE OF EXPERTS ON  
GLOBAL GEOSPATIAL  
INFORMATION MANAGEMENT

# A questionnaire for collecting opinions for Statistical Units Data Specifications

Questionnaire  
for statisticians



**GGIM WG-A Core Data  
Data Specifications**



WGA Core Data:



# User requirements

## INSPIRE

- \* Local to National governments
- \* Settlement urban / regional planning
- \* Health education planning and managing
- \* Environmental and social assessments
- \* Exposure to pressure
- \* Availability of services

Research of nearly  
any phenomena

Units for dissemination  
of statistical information

- Genericity
- Type of geographical entities
- Polygons and lines
- Multi-scale representation
- Hierarchical structure
- Temporal dimension

# **Use cases (Annex B)**

- B.1 Statistical data georeferencing
- B.2 Risk exposure analysis
- B.3 Phenomenon spread analysis
- B.5 New facility location
- B.6 Simulation
- B.7 Human health impact analysis

# SU - PD

source

analysis

Make socio-economic studies  
(poverty, education....)

Assess human pressure on ecosystems, on resources

Make accessibility studies

operational

Assess number of people in area of interest (risk, pollution, public service)

decision

communication

monitoring

Assess requirements (transport, dwelling, waste....) => city management

Reporting for European Directives (noise, air...) Potential indicators for SDG

# INSPIRE description

Units for dissemination or use  
of statistical information

- Genericity
- Type of geometries: grids, polygons and tessellations
- Multi-scale representation
- Hierarchical structure
- Temporal dimension



Figure 10 - Each area statistical unit can compose several upper area statistical units

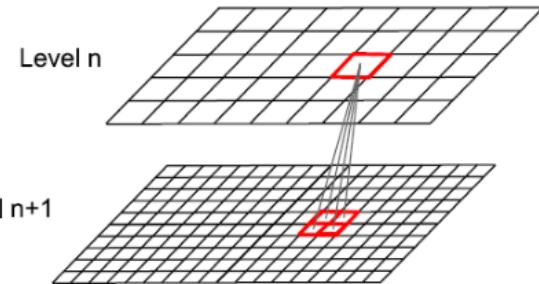


Figure 5 - Example of multi-resolution grid: A quad tree structure

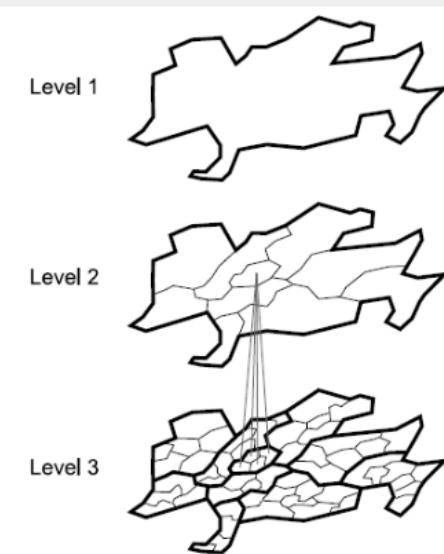


Figure 9 – A hierarchy of area statistical units

# Existing Data

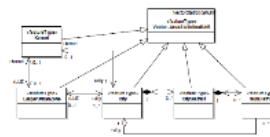


Figure 20 – Overview of the Urban Audit 2009 dataset in Western Europe



- NUTS
- LAU and Urban Audit
- Other local statistical units.
- Grids

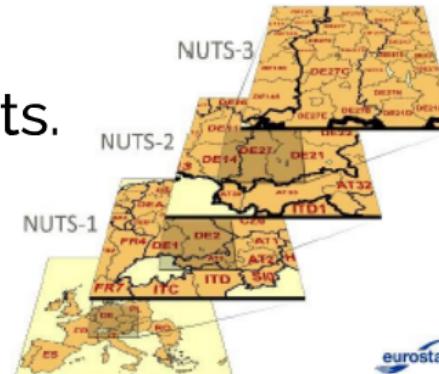
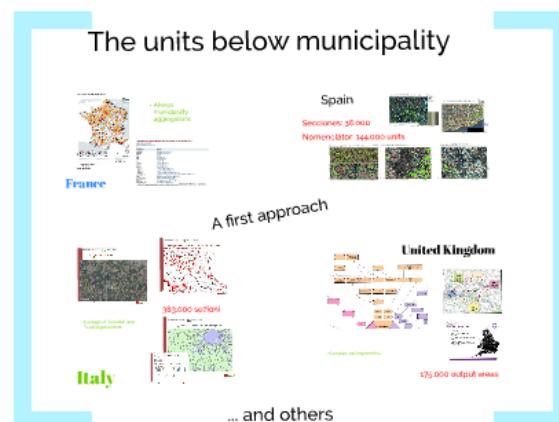
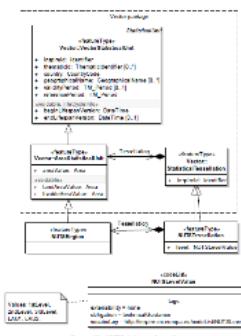


Figure 16 - The NUTS hierarchical structure.



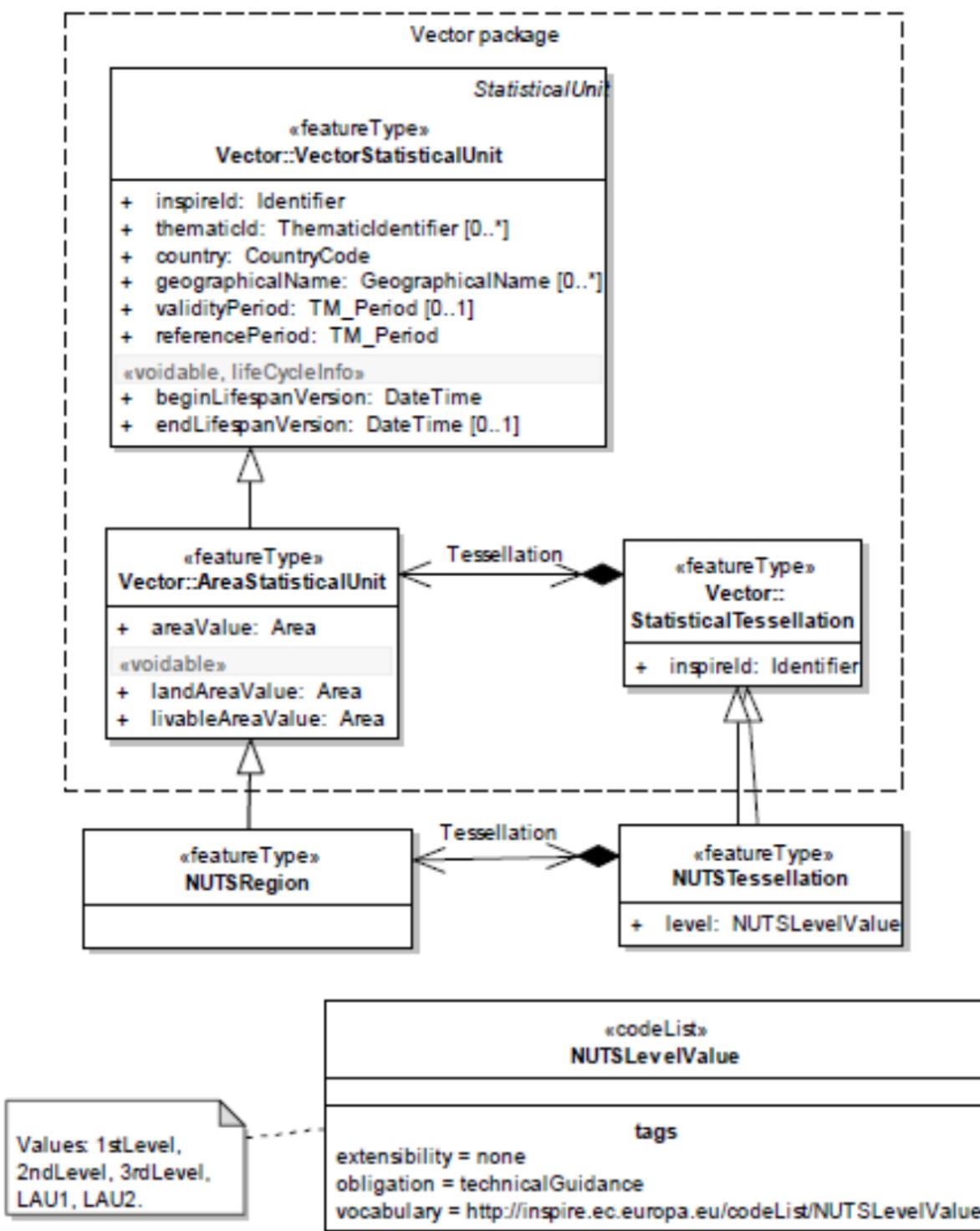


Figure 19 - NUTS dataset class diagram

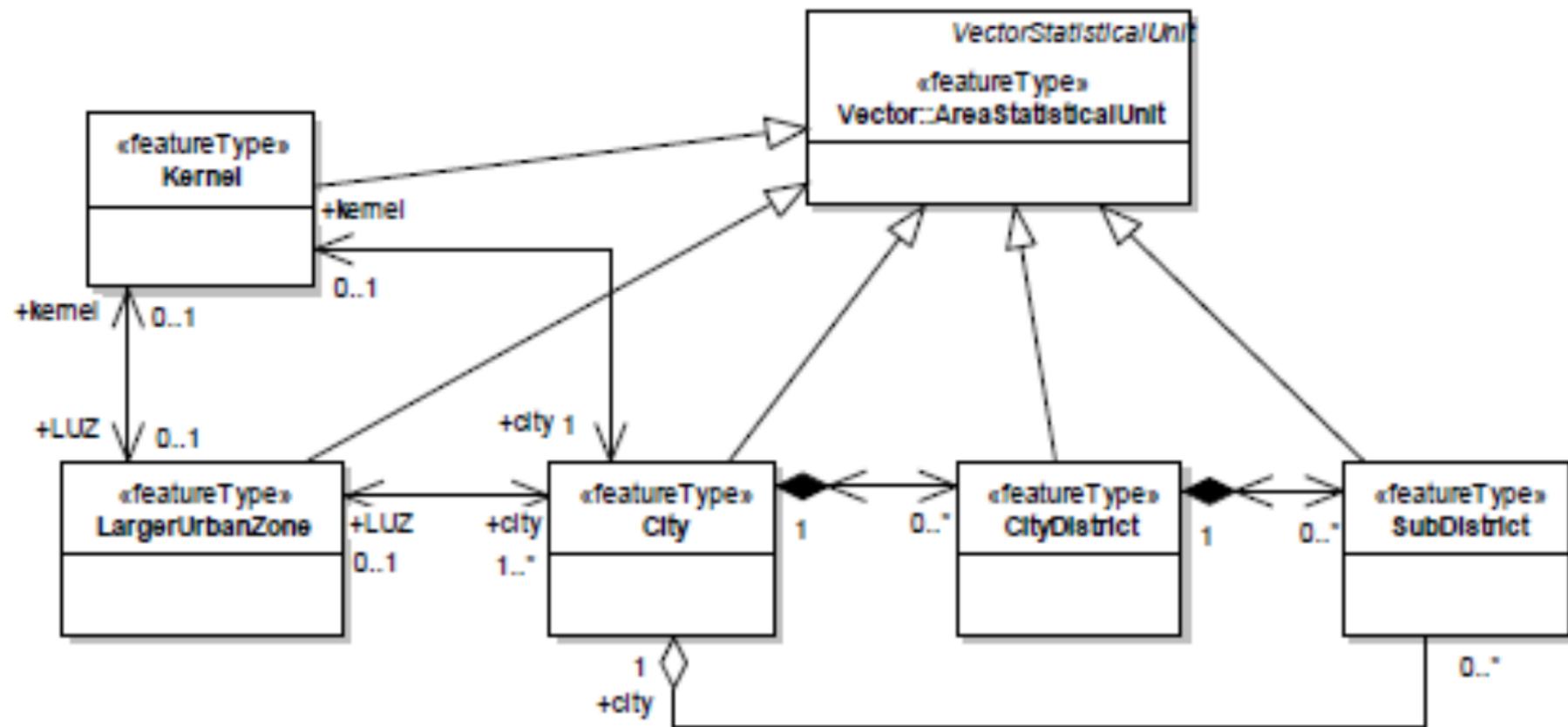
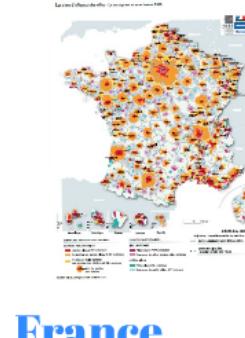


Figure 23 - Urban Audit 2009 dataset class diagram

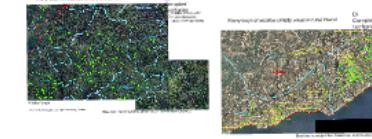


# The units below municipality



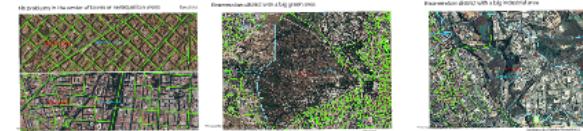
- Always municipality aggregations

France



Secciones: 36.000

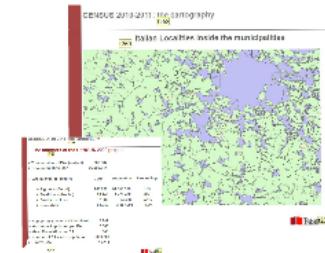
Nomenclátor: 144.000 units



- Concept of "Località", and "località produttive"



383.000 sezioni

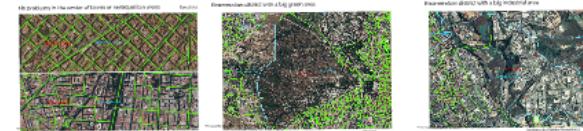


Italy

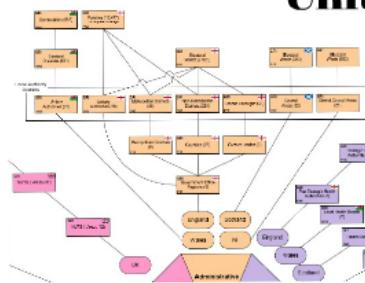
Spain

Secciones: 36.000

Nomenclátor: 144.000 units



United Kingdom

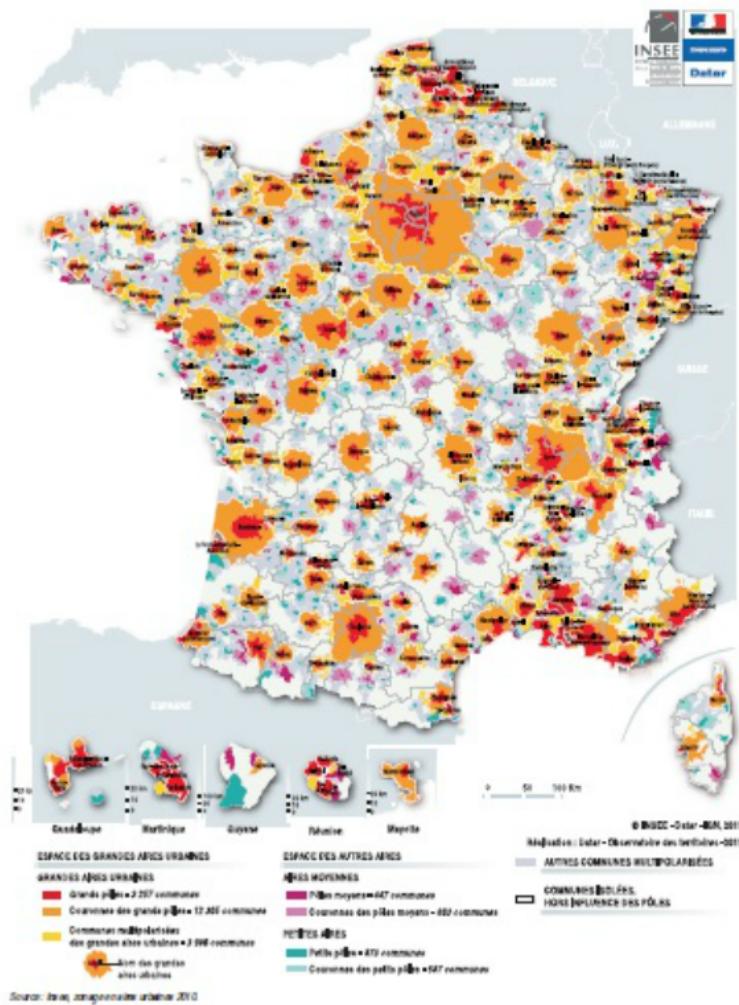


- Complex and asymmetric.

175.000 output areas

... and others

Les aires d'influence des villes - Le zonage en aires urbaines 2010



- Always municipality aggregations

## Appartenance géographique des communes au 1er janvier 2011

### Liste des variables

[Retour à la liste des communes](#)

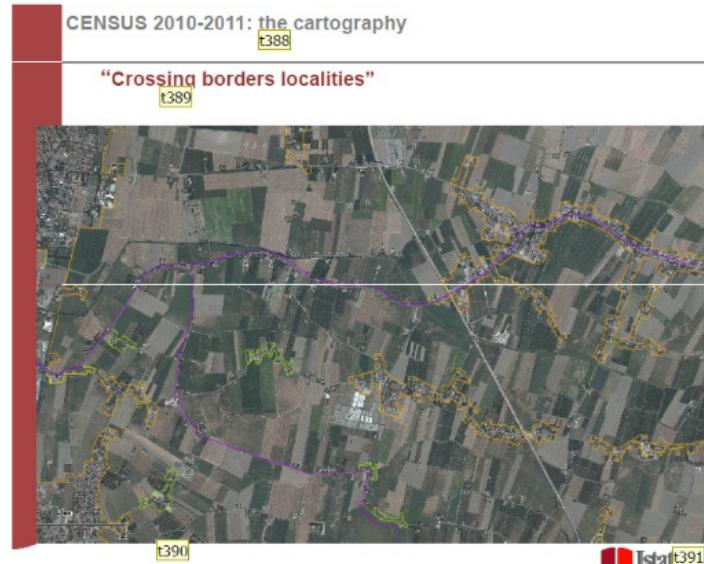
© Insee

Identifiant	Libellé
CODGEO	Département - commune
LIBGEO	Libellé de la commune
REG	Région
DEP	Département
ARR	Arrondissement
CV	Canton ville
ZE2010	Zone d'emploi 2010
UU2010	Unité urbaine 2010
AU2010	Aire urbaine 2010
TAU2010	Tranche d'aire urbaine 2010
TUU2010	Tranche d'unité urbaine 2010
TDUU2010	Tranche détaillée d'unité urbaine 2010
CATAEU2010	Catégorie de communes dans le zonage en aires urbaines 2010
EPCI	Établissement public à fiscalité propre
NATURE_EPCI	Nature d'établissement public
PSDC99	Population sans doubles comptes 1999
POP_MUN_2009	Population municipale 2009

Tous les découpages géographiques proposés ont été mis à jour dans la géographie de référence **au 1<sup>er</sup> janvier 2011**.

# France

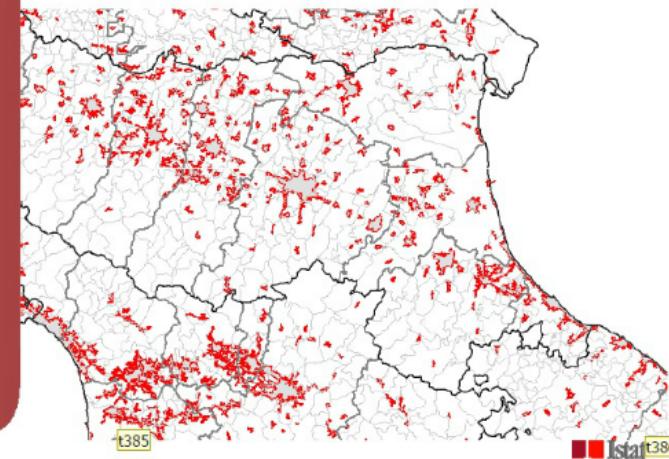
# Italy



CENSUS 2010-2011: the cartography

t383

"Crossing borders localities": the creation of morphological urban aggregates. (UNECE specifications)



Istat t386

383.000 sezioni

- Concept of "Località", and "località produttive"

CENSUS 2010-2011: the cartography

t24

The numbers of the CENSUS 2000 project

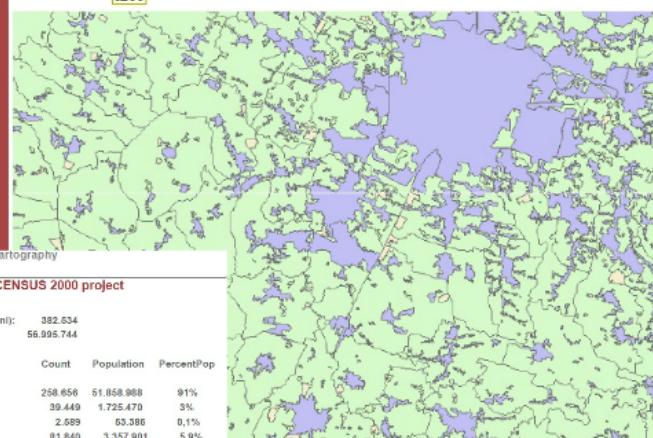
t24

	Count	Population	PercentPop
> EA by type of locality:			
> Big Loca (Centri)	258.656	51.858.968	91%
> Small Loca (Nuclei)	39.449	1.725.470	3%
> Productive Locs	2.589	63.386	0,1%
> Non-urban	81.840	3.357.901	5,8%
> Number of EAs with 1 inhabitant:	3.289		
> Maximum Population per EA:	3.386		
> Mean Population per EA:	149		
> Number of EAs with population:	336.788		
> Empty EAs:	45.748		

CENSUS 2010-2011: the cartography

t268

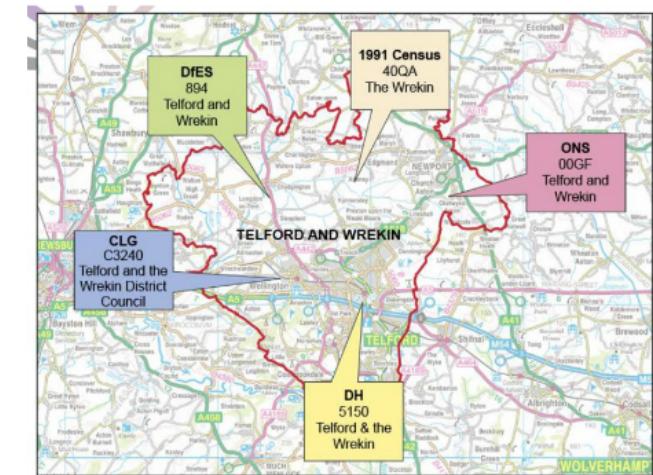
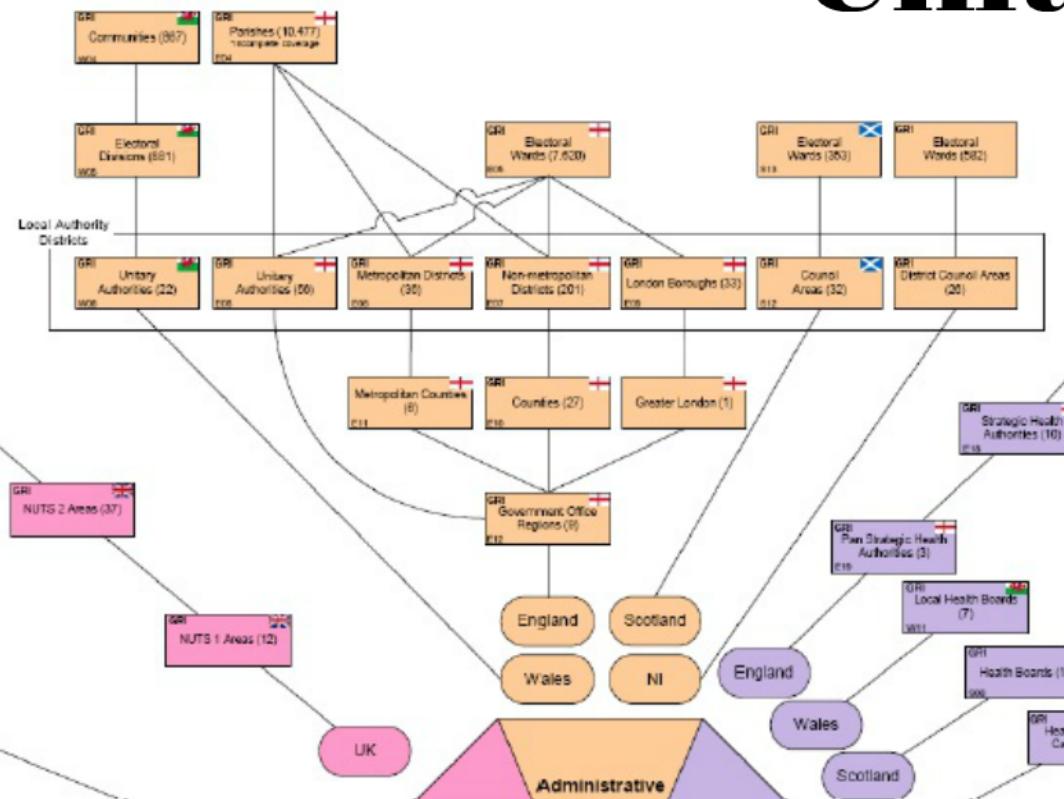
Italian Localities inside the municipalities



Istat t271

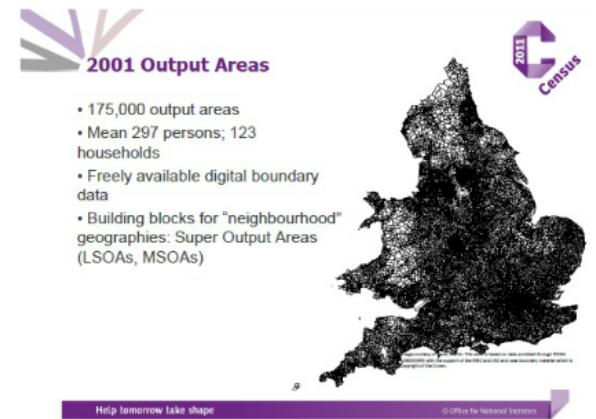
Istat t24

# United Kingdom



- Complex and asymmetric.

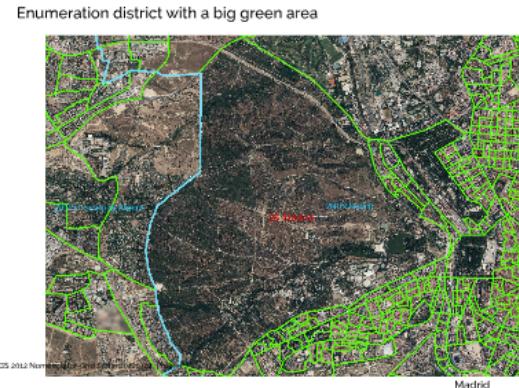
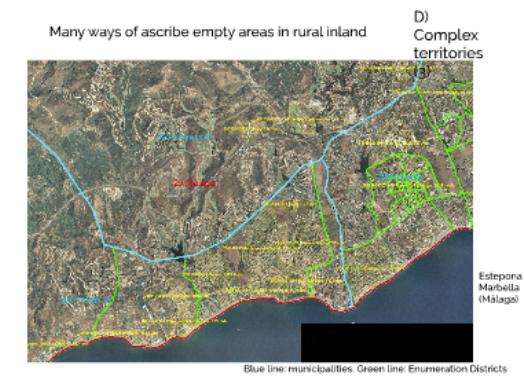
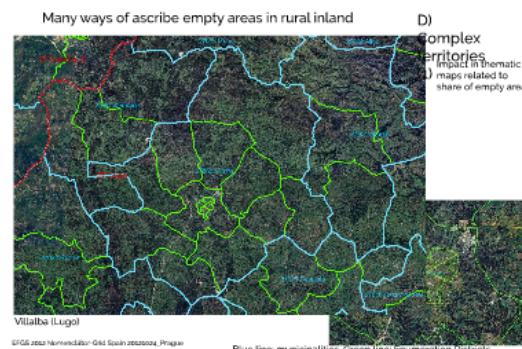
175.000 output areas



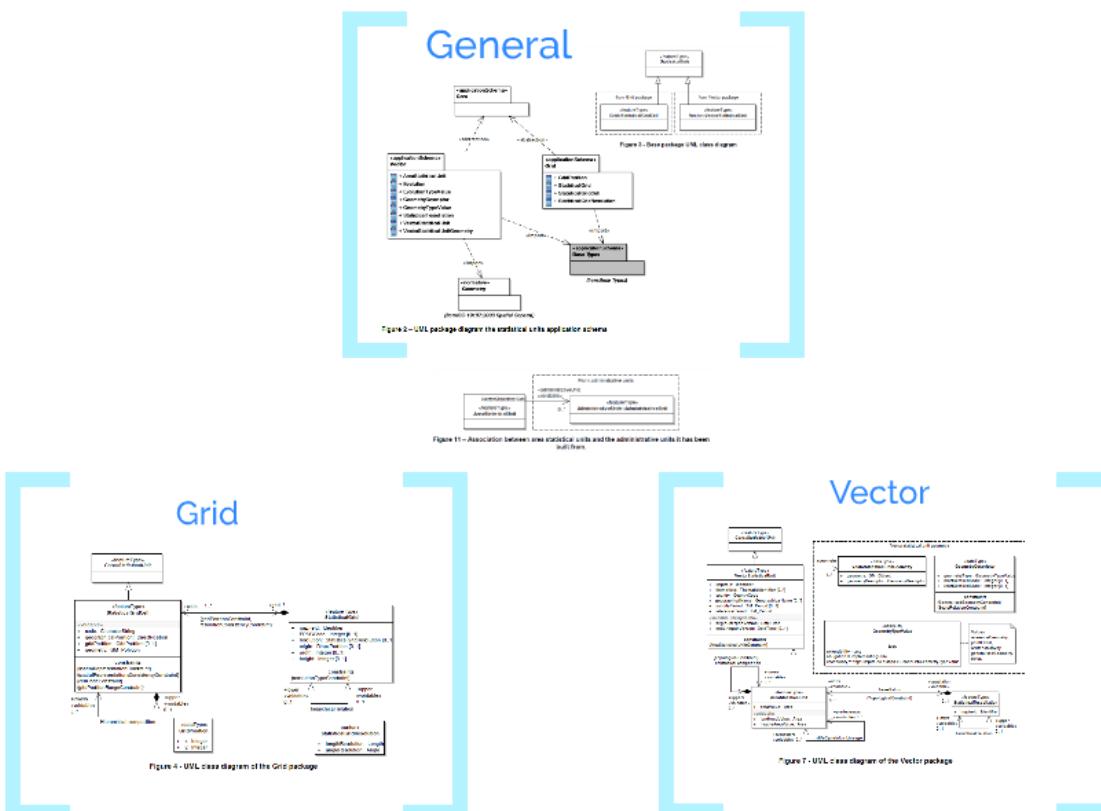
# Spain

Secciones: 36.000

Nomenclátor: 144.000 units



# INSPIRE Data Model



# General

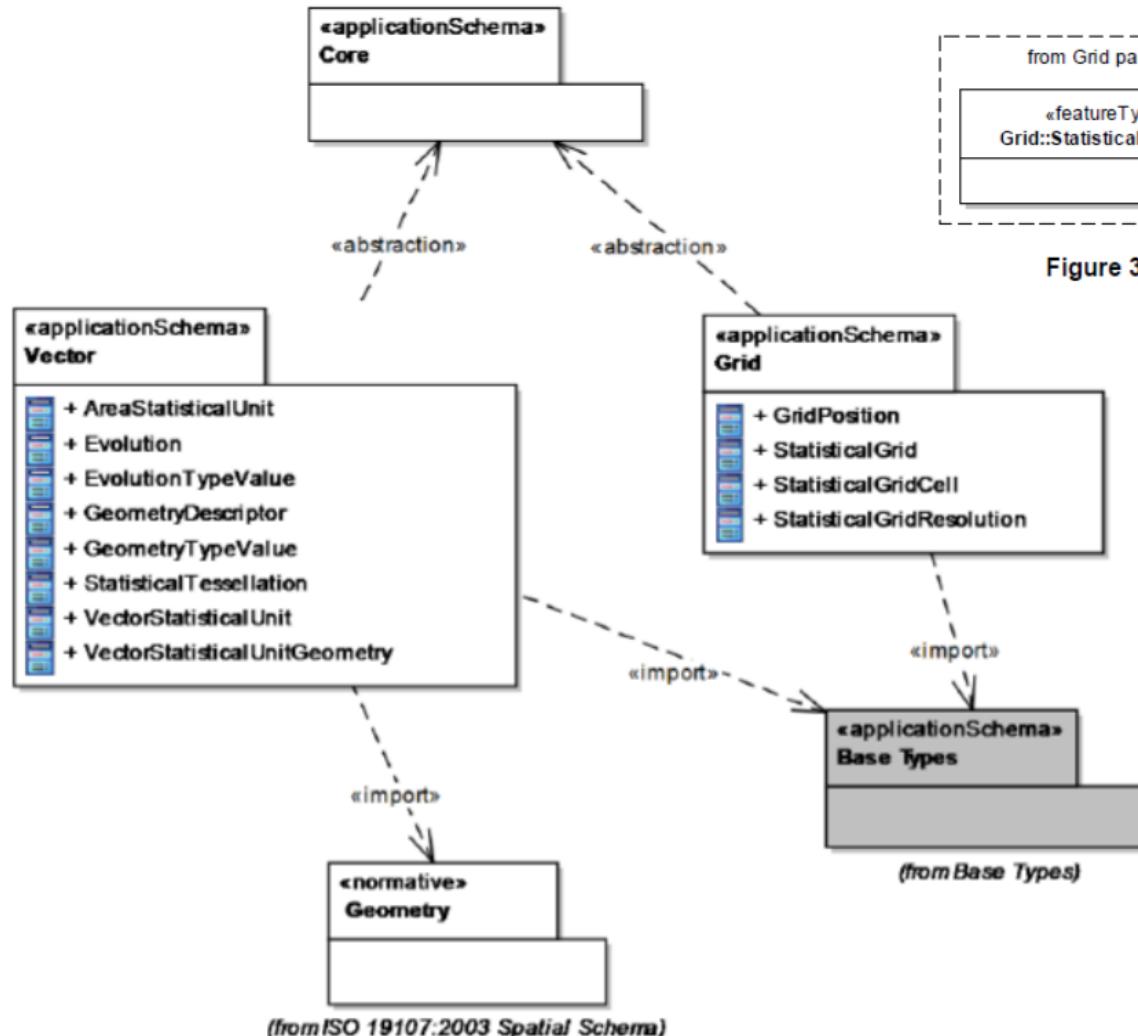


Figure 2 – UML package diagram the statistical units application schema

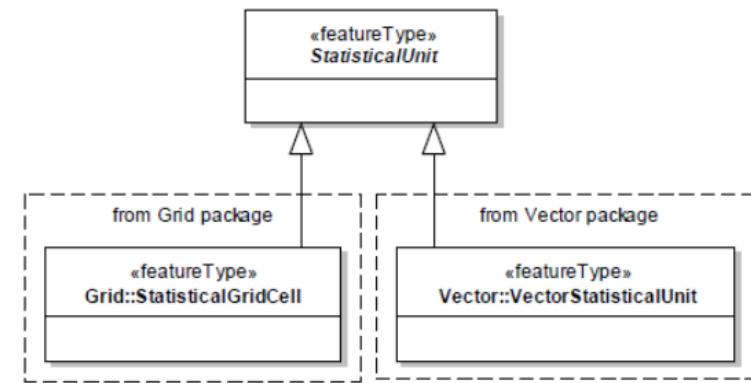


Figure 3 - Base package UML class diagram

# Grid

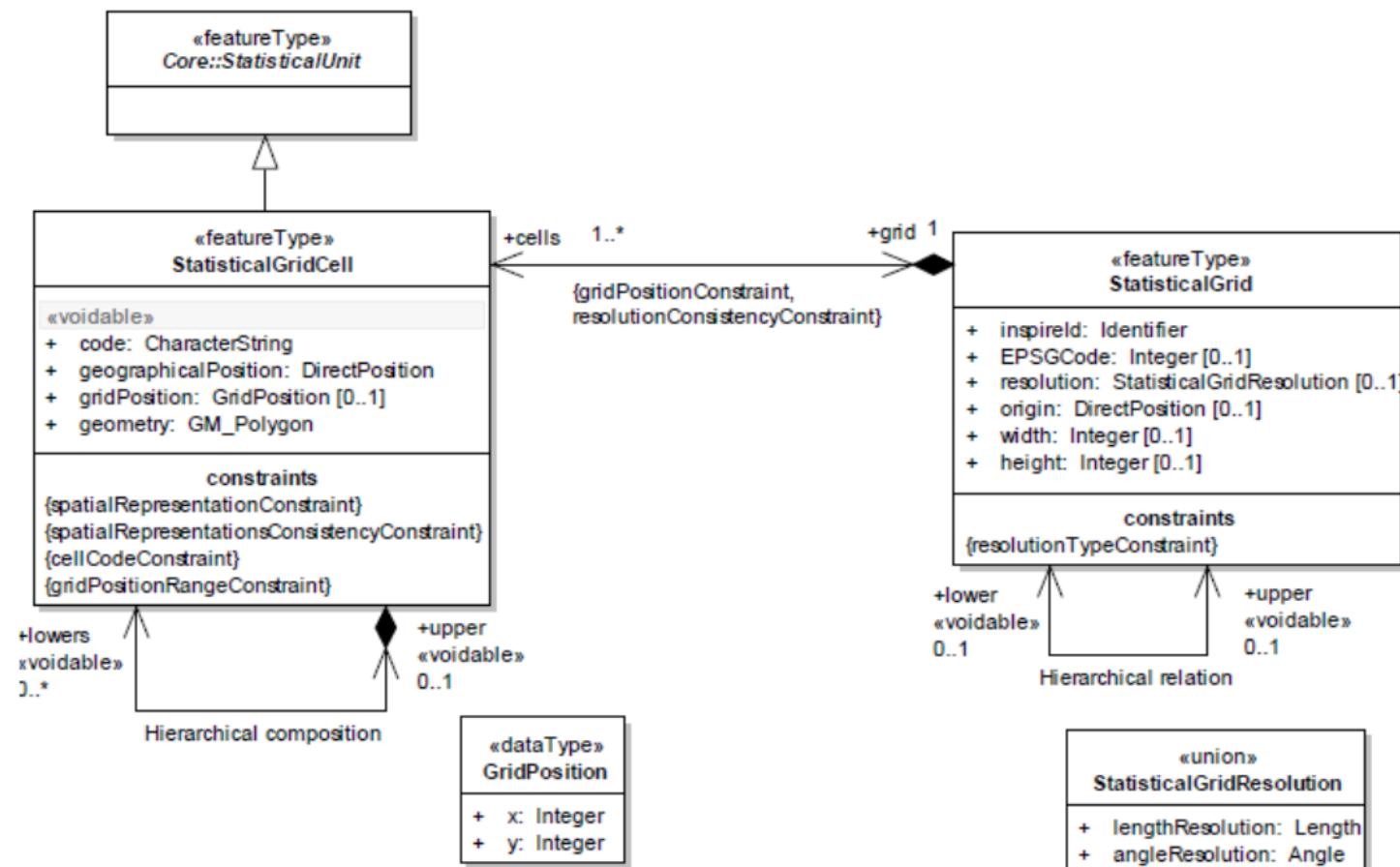


Figure 4 - UML class diagram of the Grid package

# Vector

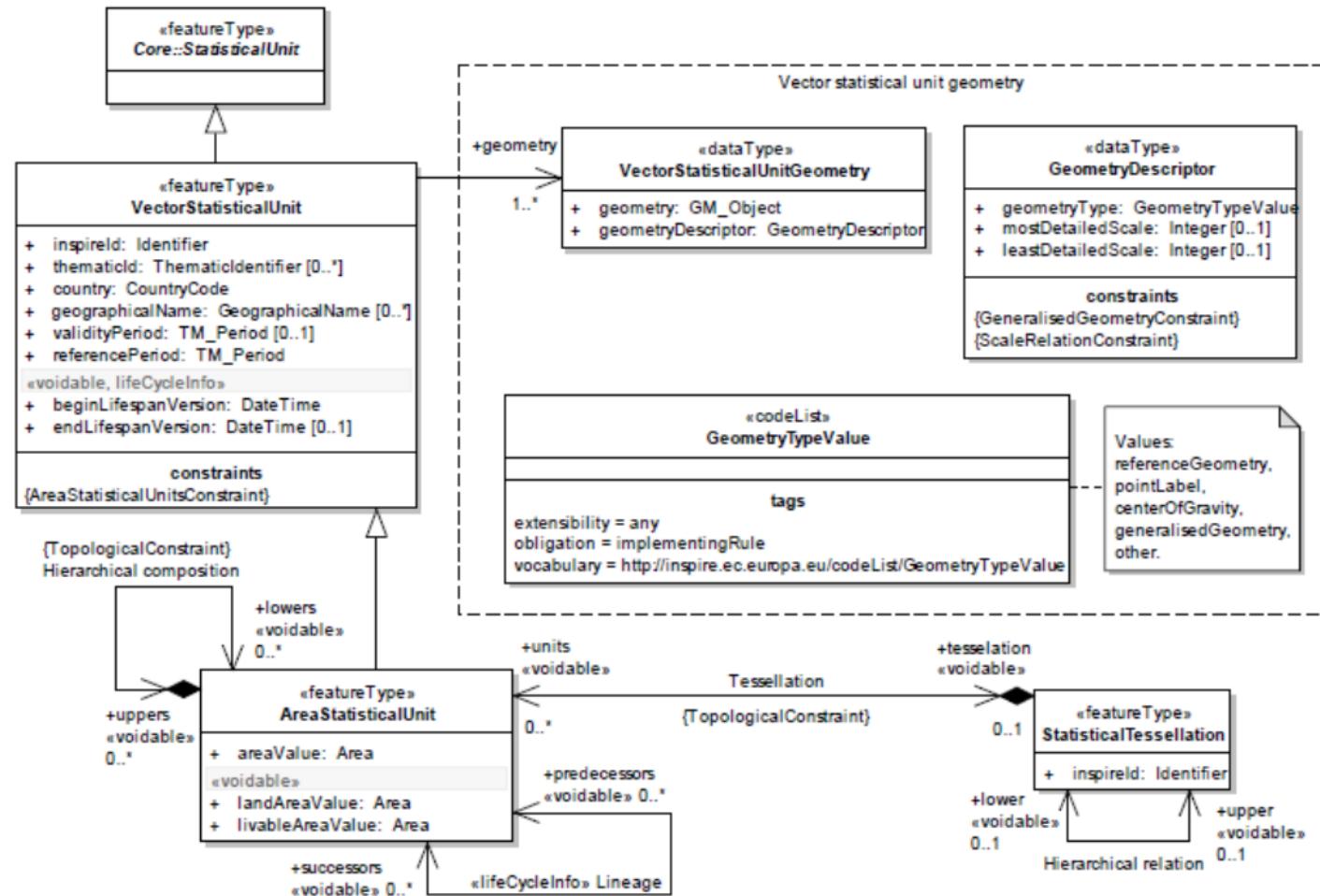


Figure 7 - UML class diagram of the Vector package

# Temporal dimension

A difficult issue

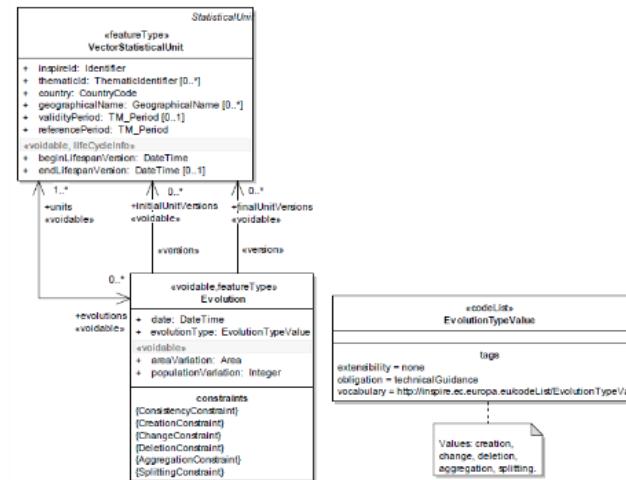


Figure 13 - Evolutions

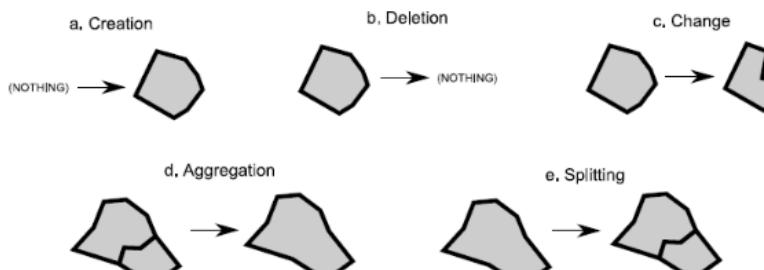


Figure 12 - Typical statistical unit changes

## The INSPIRE design

constraints  
  {ConsistencyConstraint}  
  {CreationConstraint}  
  {ChangeConstraint}  
  {DeletionConstraint}  
  {AggregationConstraint}  
  {SplittingConstraint}

Values: creation,  
change, deletion,  
aggregation, splitting.

Figure 13 - Evolutions

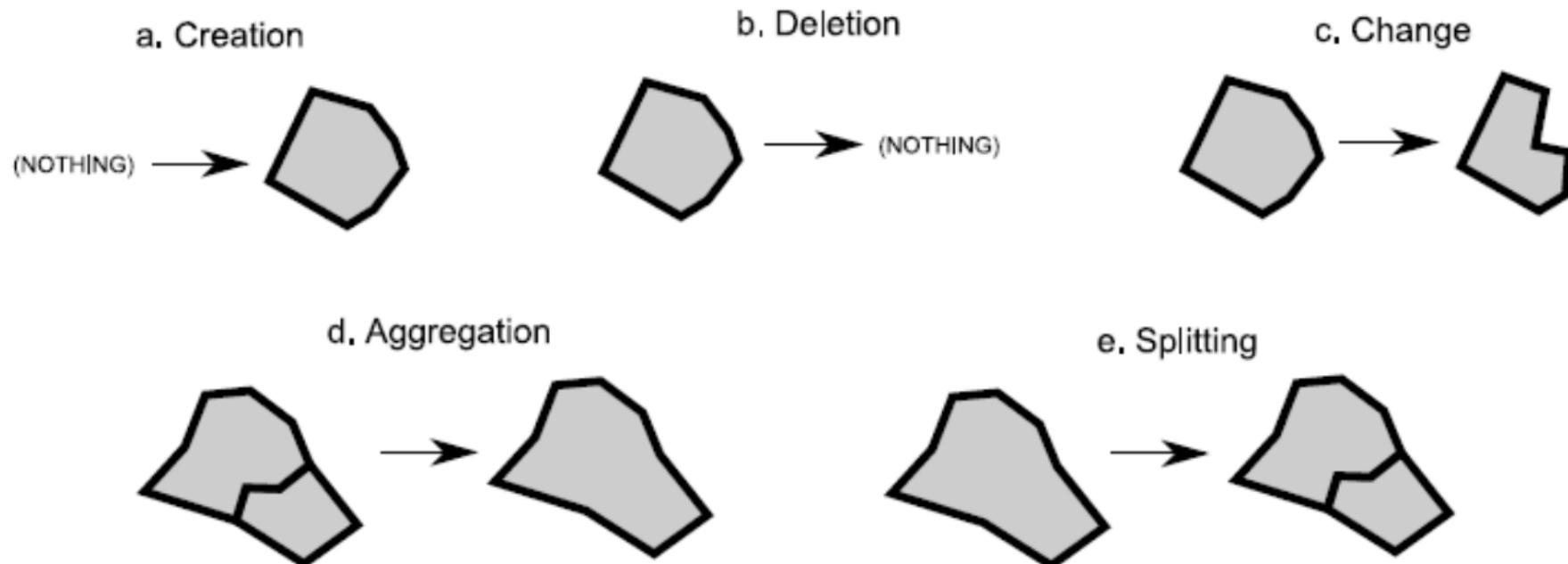


Figure 12 - Typical statistical unit changes

# The INSPIRE design

# The Annex F

## Aggregation

Simple aggregation: We consider the two following statistical units:

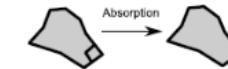
<b>Id</b>	<b>VersionId</b>	<b>validityPeriod.begin</b>	<b>validityPeriod.end</b>	<b>beginLifespanVersion</b>	<b>endLifespanVersion</b>	<b>predecessors</b>	<b>successors</b>
169	4	31-03-2010		21-02-2010			
182	2	12-03-2010		14-02-2010			

These statistical units are aggregated into a new one, with an Id 216. This aggregation is planned for the 01-01-2011, and this change is inserted into the dataset the 30-01-2010:

<b>Id</b>	<b>VersionId</b>	<b>validityPeriod.begin</b>	<b>validityPeriod.end</b>	<b>beginLifespanVersion</b>	<b>endLifespanVersion</b>	<b>predecessors</b>	<b>successors</b>
169	4	31-03-2010	01-01-2011	21-02-2010	30-01-2010		216
182	2	12-03-2010	01-01-2011	14-02-2010	30-01-2010		216
216	1	01-01-2011		30-01-2010		169, 182	

This change is equivalent to a deletion of the two initial statistical units with a creation of the new one. The Lineage association roles predecessors and successors allow the additional information to be stored. This example involves only two statistical units that are aggregated into a new one – it can be extended to the cases of aggregation of more than two statistical units.

The “absorption” case: A specific case can be considered when one of the initial statistical units absorbs the other one.



If in the previous example 182 absorbs 169, this change can be represented. It is equivalent to a deletion of the absorbed statistical units and a change of the absorbent one.

<b>Id</b>	<b>VersionId</b>	<b>validityPeriod.begin</b>	<b>validityPeriod.end</b>	<b>beginLifespanVersion</b>	<b>endLifespanVersion</b>	<b>predecessors</b>	<b>successors</b>
169	4	31-03-2010	01-01-2011	21-02-2010	30-01-2010		
182	2	12-03-2010	01-01-2011	14-02-2010	30-01-2010		
182	3	01-01-2011		30-01-2010		169	

## Creation / Deletion / Change

### Creation:

The creation of a statistical unit is decided. This statistical unit should become an official reporting unit from 01-04-2008. The new statistical unit is inserted into the dataset the 22-03-2008 with an Id 816.

An object with the following attributes is inserted:

<b>Id - VersionId</b>	<b>validityPeriod.begin</b>	<b>validityPeriod.end</b>	<b>beginLifespanVersion</b>	<b>endLifespanVersion</b>
816 - 1	01-04-2008		22-03-2008	

### Change:

An attribute value or the geometry of the statistical unit is changed. This change is inserted into the database the 25-12-2008 and is supposed to be official from the 01-01-2009. A version second of the object with the following attributes are inserted:

<b>Id - VersionId</b>	<b>validityPeriod.begin</b>	<b>validityPeriod.end</b>	<b>beginLifespanVersion</b>	<b>endLifespanVersion</b>
816 - 1	01-04-2008	01-01-2009	22-03-2008	25-12-2008
816 - 2	01-01-2009		25-12-2008	

### Deletion:

The statistical unit will be officially deleted from 04-09-2011. This information is inserted into the dataset the 03-10-2011. The attributes validityPeriod.end and endLifespanVersion are filled:

<b>Id - VersionId</b>	<b>validityPeriod.begin</b>	<b>validityPeriod.end</b>	<b>beginLifespanVersion</b>	<b>endLifespanVersion</b>
816 - 1	01-04-2008	01-01-2009	22-03-2008	25-12-2008
816 - 2	01-01-2009	04-09-2011	25-12-2008	03-10-2011

## Splitting

Simple splitting: We consider the following statistical unit:

<b>Id</b>	<b>VersionId</b>	<b>validityPeriod.begin</b>	<b>validityPeriod.end</b>	<b>beginLifespanVersion</b>	<b>endLifespanVersion</b>	<b>predecessor</b> s	<b>successor</b> s
179	3	28-08-2010		17-08-2010			

The statistical unit is split into two new ones. These new statistical units are represented with two new objects with an Id 217 and 218. This splitting is planned for the 01-01-2011, and this change is inserted into the dataset the 17-12-2010:

<b>Id</b>	<b>VersionId</b>	<b>validityPeriod.begin</b>	<b>validityPeriod.end</b>	<b>beginLifespanVersion</b>	<b>endLifespanVersion</b>	<b>predecessor</b> s	<b>successor</b> s
179	3	28-08-2010	01-01-2011	17-08-2010	17-12-2010		217, 218
217	1	01-01-2011		17-12-2010		179	
218	1	01-01-2011		17-12-2010		179	

This change is equivalent to a deletion of the initial object, and a creation of the new ones. The Lineage association roles predecessors and successors allow the additional information to be stored. This example involves the splitting of one statistical unit into only two new ones – it can be extended to the cases of splitting into more than two statistical units.

The “exclusion” case: As for the absorption case (see example 2), a specific case can be considered when one of the final statistical units is the same as the initial one. In this case, the splitting can be seen as an exclusion of one part of the statistical unit as a new one.



If in the previous example 179 excludes 217, this change can be represented. It is equivalent to a creation of the excluded statistical unit and a change of the initial one.

<b>Id</b>	<b>VersionId</b>	<b>validityPeriod.begin</b>	<b>validityPeriod.end</b>	<b>beginLifespanVersion</b>	<b>endLifespanVersion</b>	<b>predecessor</b> s	<b>successor</b> s
179	3	28-08-2010	01-01-2011	17-08-2010	17-12-2010		
179	4	01-01-2011		17-12-2010			217
217	1	01-01-2011		17-12-2010		179	

# INSPIRE to Core Data

- Open issues
  - Sets of units considered
  - Territorial level
  - Time evolution
  - Quality aspects

# The questionnaire for statisticians

## User requirements and core statistical data

All users:  
• Local, National governments and administrations.  
• Statistical offices (Eurostat, national statistical agencies).  
• Private, general users.  
• Academic institutions.

All UN member states:  
• Settlement editor / regional planning.  
• Environment and climate change management.  
• Land environment and social (labor market, education, poverty).  
• Social and labor demography (life assessments).  
• Statistical analysis of science and technology fields.  
• Statistical analysis of environment.  
• Availability of cameras.  
• Sustainability.

- ¿Adding other statistical units to geodata?:
- Enumerations Districts?
  - Postal codes?
  - Smaller grids?
  - Others?

Confirming the INSPIRE statements and level:

- NUTS-LAU units
- Urban Audit units
- Grid 1km

## Information on availability of possible additional core data

- Available features class
- Units per country.
- Updating cycles (yearly, decennial, others).
- Average changes by year (%)

## Quality issues

- Pan-european coverage
- Uptodateness
- Geographical precision.
- Others

# User requirements and core statistical data

## A) USERS:

- a) Local to National governments and administrations.
- b) Statistical offices (from local to national and European).
- c) Private, general users.
- d) Research, academics.

## B) USE CASES:

- a) Settlement urban / regional planning.
- b) Health and education planning and management.
- c) Environmental and social (labor market, education, poverty/social exclusion, demography etc.) assessments.
- d) Economical assessments.
- e) Spatial analyses of science and technology hubs.
- f) Exposure to pressure.
- g) Availability of services.
- h) Transportation.

Confirming the INSPIRE statements and level:

- NUTS-LAU units
- Urban Audit units
- Grid 1km

¿Adding other statistical units to geodata?:

- Enumerations Districts?
- Postal codes?
- Smaller grids?
- Others?

# **Information on availability of possible additional core data**

- Available features class
- Units per country.
- Updating cycles (yearly, decennial, others).
- Average changes by year (%)

# Quality issues

- Pan-european coverage
- Uptodateness
- Geographical precision.
- Others

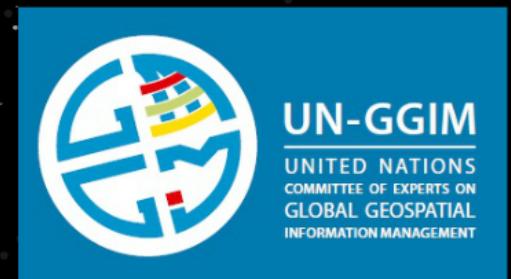
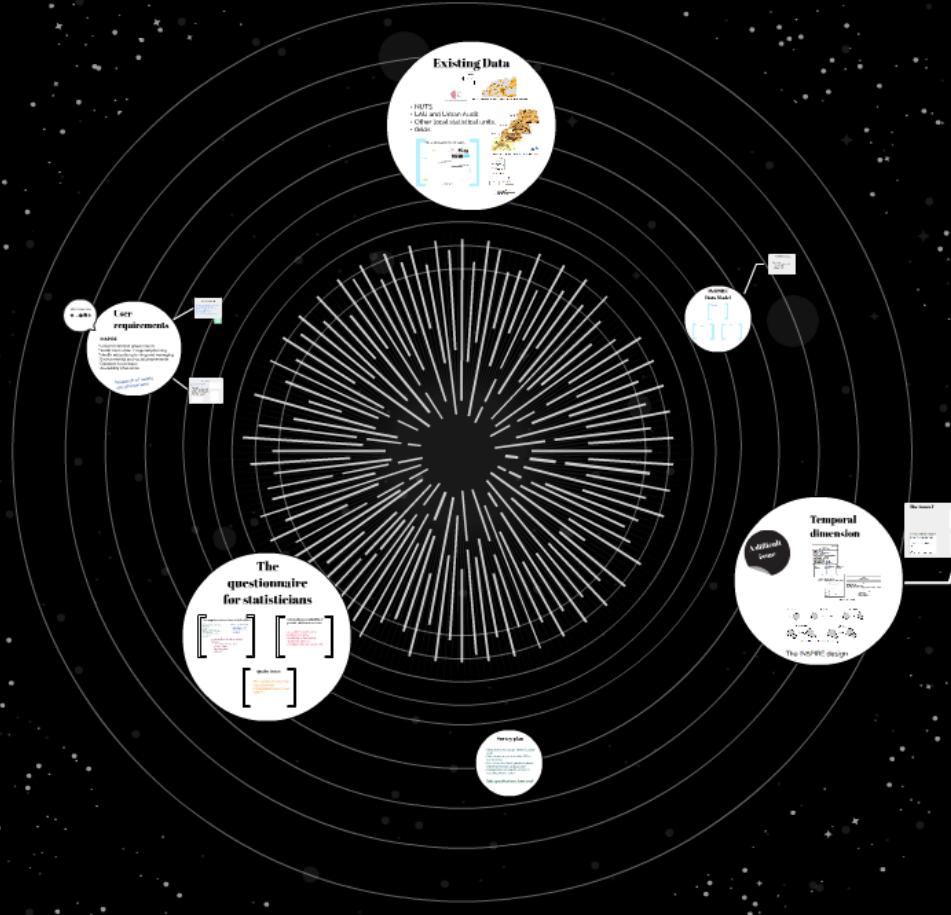
# **Survey plan**

- Questionnaire design: June-October 2016
- Questionnaire presentation EFGS Conference.
- Launch and collect questionnaires: December 2016-January 2017
- Presentation of results at GISCO meeting: March 2017

**Data specifications: June 2017**

# A questionnaire for collecting opinions for Statistical Units Data Specifications

Questionnaire  
for statisticians



**GGIM WG-A Core Data  
Data Specifications**

