

# ESPON 2013 DATABASE

## Annex to Inception Report on Data Flows

Authors : Geoffrey Caruso, Jerome Gensel, Claude Grasland, Bogdan Moisuc, Nicolas Lambert, Moritz Lennert, Christine Zanin.

Date : 29th November 2008

Claude Grasland 29/11/08 17:54

Supprimé: 28th

### CU REQUIREMENT:

*“Definition an internal data/information flow for the ESPON 2013 programme. The Lead Partner is asked to produce a short note, to be disseminated to ESPON 2013 projects under implementation, including information on the deliveries of data expected, formats, maps and metadata required.”*

### INTRODUCTION

The organisation of data circuit and data flow has to take into account two complementary (but also partly contradictory) missions that make relatively difficult the task of the project ESPON DB.

Support ESPON Projects from Priority 1 and 2 in the scientific elaboration of their research. The different projects from priority 1 that have started in summer 2008 (Energy, Climate Change, Demography, Urban, Rural, TIA) and the future projects from priority 2 that will start in spring 2008 create an immediate need for harmonised statistical data and geographical maps. On the basis of this data, the projects will then elaborate new indicators, new maps, etc. The first objective of the project ESPON DB is therefore to **be an efficient interface between ESPON projects and statistical or cartographic institutions**. For this task, we don't have to produce directly data but simply to insure an efficient management of data flows (import and export) between EUROSTAT, EEA, Eurogeographics on the one hand, and the ESPON projects on the other hand.

#### Create new research possibilities that are actually not available for technical reasons.

This objective is completely different as it is not driven by the timetable of ESPON projects that are launched by ESPON CU. On the contrary, it is a pro-active action that make possible the launching of new ESPON projects. For example, the enlargement of the regional database NUTS2/3 to neighbouring countries (Balkans, Ukrainia, Belarus, Moldova, Maghreb, Turkey, ...) is typically an action that is not driven by a specific request of ESPON projects but that will make possible the launching of new studies of high political interest about EU borders. In the same spirit, the successful elaboration of a long term database 1960-1980-2000 for a selected number of indicators (even limited to demography) could be the basis for new research on long term drivers of regional change that is currently not possible (at less at regional level).

Based on this crucial distinction, we have to propose a data circulation model that should be able at the same time to (1) answer quickly and efficiently to urgent requests of running projects (2) allow long term improvements of ESPON possibilities. Both challenges should be

addressed simultaneously. But two additional constraints have to be taken into account

**Data and Figures:** It is important to observe also that the flows described above are not limited to data (in the narrow sense of statistical table) but should also consider the production of figures in a general sense (i.e. maps and any other type of graphic materials).

**Internal and external diffusion:** Finally, it is important to make a distinction between circulation of data and figures inside the ESPON program (subject to specific rules) and the dissemination strategy of ESPON toward the external world through the ESPON website (subject to decision of the ESPON Coordination Unit Monitoring Committee).

## I. OVERVIEW OF DATA FLOW

The data flow within the ESPON 2013 Database Project defines the itinerary of data coming from external sources, within the project and then back again for availability to the other ESPON projects and to the external world.

### I.1 Data flow for the first version of ESPON 2013 Database

The first data flow elaborated within the ESPON 2013 Database Project is depicted on Figure 1. This first version of the data flow is to be used until the integration pool becomes fully operational. In this version, external data (data files or databases) are centralized by the lead partner RIATE.

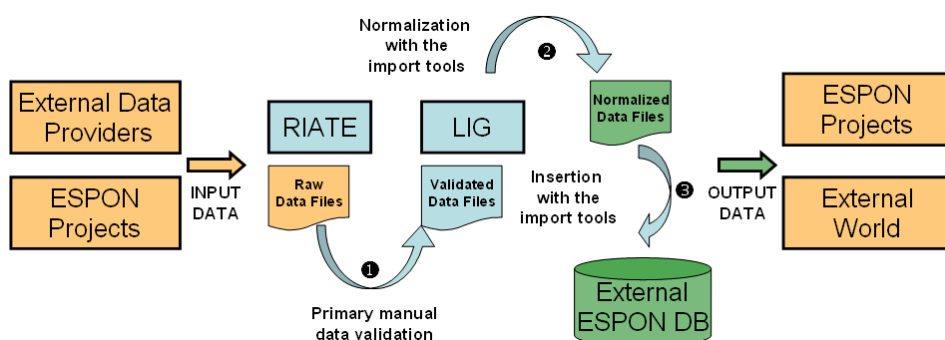


Figure 1. Data flow of the first version of the ESPON 2013 Database.

Such external data can be provided as a statistical table downloaded from the Eurostat Web site, or as a data file transmitted by any ESPON project, or as a database received from a non ESPON project (e.g. study from European Parliament on Shrinking Regions) or from different EC services (e.g. a database provided by the DG AGRI). External data can be as well national statistical figures provided by non EU countries (e.g. a database received from the ministry of employment of Switzerland or a publication named « Liechtenstein in Figures »...) External

data coming into the import pool have the following characteristics: (1) they are free from any previous transformation, (2) they have not been yet validated by a quality control, (3) they are not necessarily compatible with the standard of ESPON DB and, in particular with the cartographic standards.

Data undergo a primary, manual data checking process (see step 1 on the Figure 1) and are transmitted to the lead partner LIG. Then, data files undergo a formatting process (see step 2 on the Figure 1), in order to make more easy their importation into the first version of the external database of the ESPON project (see step 3 on the Figure 1).

In this first version of data flow, data integration is made by some experts that produce technical reports describing how integration problems have been solved. These technical reports can be transmitted to another expert (who can use and improve it). In the future, it is expected that this expert knowledge will be partially or fully automated through computer applications. As an example, the production of data at level NUTS2 through aggregation of data at level NUTS3 is typically a procedure that can be introduced in the data model (even if it is not as obvious as it looks ...).

Data outputs for the other ESPON projects will be made available in the shape of normalized data files (diffused via the ESPON site and/or directly by the lead partner) and as a database server that will answer queries to be formulated through a simple Web interface. The external Espo Database, placed on a Web server hosted by ESPON, will grant different access rights for different groups of users (ESPON projects, EC members, general public, etc.), in order to cope with data rights issues.

As a side note, it is important to point out that, in order to fully automate the data validation and importation process, it is desirable that data inputs are provided directly by other ESPON projects using a normalized format (see section on data input). Otherwise, the process of formatting heterogeneous file structures dynamically could prove to be highly time consuming and could slow the overall progress of the ESPON 2013 Database Project.

## 1.2 Data flow for the next version of ESPON 2013 Database

In the future, once the integration pool becomes operational, the next data flow elaborated within the ESPON 2013 Database Project will be the one depicted on Figure 2.

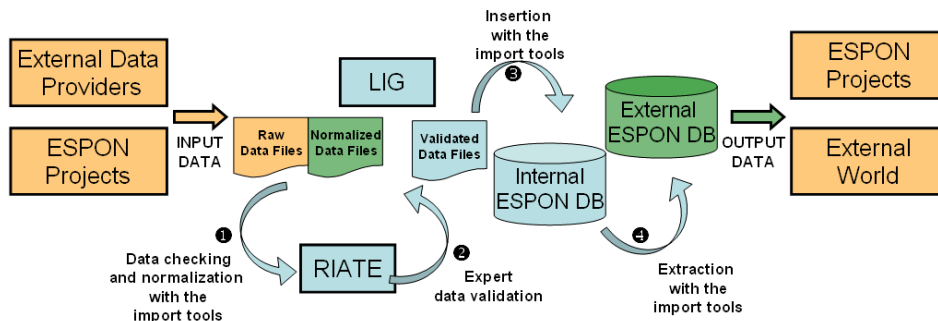


Figure 2. Data flow of the next version of the ESPON 2013 Database

Input data will arrive directly at the lead partner LIG, where they will undergo an automatic checking process (see step 1 on Figure 2). Records flagged as doubtful will be sent to the lead partner RIATE for expert validation and assessment (see step 2 on Figure 2). If necessary, communication with the external data provider will be initiated in order to eliminate doubts or to correct erroneous values. Datasets passing the quality and coherence tests will be inserted in the internal ESPON 2013 database (see step 3 on Figure 2). The internal database is a complex structure relying on PostgreSQL/Postgis DBMS, on geographical and thematic ontologies and harmonization and estimation methods. Ready to use datasets will be extracted (see step 4 on Figure 2) from the internal database and made available through the external database. From this moment on, the other ESPON projects and the external world will be given access to harmonized datasets through a Web interface.

### I.3 Internal and external database

Data outputs of the ESPON 2013 Database Project will be made available through either normalized data files (data tables, printed or not), or through a simplified database (called the external database) accessible through a Web interface.

The external database is a database with a relatively simpler schema (compared to the internal database) which allows access to and querying of data concerning NUTS territorial units (see Figure 4).

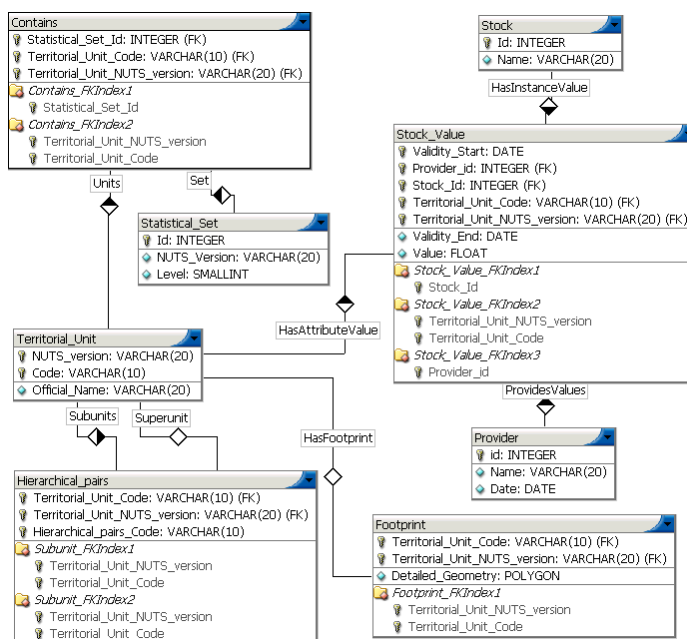


Figure 3. Implementation schema of the external database

The schema of the external database contains the following classes:

1. *Territorial\_Unit* describes NUTS territorial units, from states to local units. A territorial unit is identified by a nomenclature version and a code in this nomenclature. It also contains an official name.
2. *Footprint* describes the geometrical contour of the territorial units and it is meant for storing the detailed geometries of territorial units required for GIS manipulation.
3. *Statistical\_Set* describes custom defined study areas (like EU27, EU 15, etc.). They are defined as sets of territorial units.
4. The *Hierarchical\_Pairs* relation contains all the administrative inclusion relations between territorial units belonging to consecutive NUTS levels.
5. *Stock* describes the names of the indicators contained in the database.
6. *Provider* describes the different data sources for the data contained in the database (e.g. Eurostat, DG Agri Database, etc.)
7. *Stock\_Value* is the relation that stores the indicator values for the territorial units. There may be one value indicator per NUTS revision, TU, indicator, production date and provider.

The first version of the external database has been implemented using PostgreSQL 8.3 and the data acquisition process has started. The data acquisition is evolving in parallel with the development of the import tools. An SQL file which allows restoring the database on any PostgreSQL 8.3 server is provided in annex to this document. For the moment, the database contains only the NUTS 2003 and 2006 territorial units (from NUTS0 to NUTS3), some providers and a basic set of indicators. A demo will be made at the ESPON Seminar in Bordeaux, running on a basic set of indicator values.

In the next step (February 2009 FIR) the external database will be made available through a Java Web server, while its query interface will be improved, taking into account the remarks and comments made by first users.

In more advanced stages of the project, the database will be filled with all the data provided by the lead partner RIATE and, later on, with harmonized data extracted from the internal ESPON 2013 database. It will allow the exportation of both the data (under various formats: Excel, text, GIS, etc.) and the metadata (enriched INSPIRE compliant XML file).

## **II. FLOWS FROM ESPON DB TO ESPON PROJECTS**

In this section, we describe what has been delivered by ESPON DB project the 28<sup>th</sup> November 2008 as first version of ESPON Database.

### **II.1 Delivery of data**

The ESPON database is organised in four directories that defines the process of data elaboration and the type of data that can be delivered at the different stages of elaboration.

- **The import pool** is that part of the database where new data are initially received from various sources in raw format. This can be, for example, a statistical table download from the Eurostat website, a data file transmitted by an ESPON project, a database received from a non ESPON project (e.g. study from European Parliament on Shrinking Regions) or from different services of the EC (e.g. a database received from

DG AGRI). It can also be national sources collected for non EU countries like a database received from the Ministry of Employment of Switzerland or a publication named « Lichtenstein in Figures »... The common point of the data that are available in the import pool is the fact that (1) they have not been subject to any transformation by the ESPON DB project, (2) they have not been validated by a quality control, (3) they are not necessarily compatible with the standard of ESPON DB and, in particular with the cartographic standards. It is therefore a « buffer area » for data that are subject to be eventually introduced in the ESPON DB but also an « Archive » making possible to go back to initial data collection if doubts appear in the future concerning more elaborated data.

- **The integration pool (or internal database)** is that part of the database where harmonised data (from ESPON point of view) are elaborated combining various sources and where quality control and estimation of missing values can be introduced. It is the « heart » of the ESPON database i.e. the place where a real added value is obtained in terms of creation of new information. To be clearer, the data that are introduced in the integration pool are organised in order to fulfil specific objectives or challenges. The aim is not to collect the maximum amount of data but to be able to fulfil specific aims through the choice of relevant metadata, ontologies, etc. For example, if the first challenge is to produce complete regional data sets at NUTS2 or NUTS3 level (not necessary derived from Eurostat), we are obliged to create a complete dictionary of spatial units belonging to the NUTS system, not only for actual revision (NUTS version 2006) but also for the older divisions (2003, 1999, 1995). And we are also obliged to imagine what are the potential non EU countries to be integrated into the database, etc... This is only one example and other parts of the integration pool can be further elaborated to the integration of grid data, of urban units, of networks ... It is important to note that the integration pool is a mixture of expert knowledge and computer application, with an attempt to translate as much as possible the first into the second. The basic idea is that first integration of data is made by experts who produce a technical report on how they have solved some difficulties. This technical report can be transmitted to future experts (who can use and improve it) but we hope also that, in many case, this expert knowledge can be partially or fully automatized through computer applications. As an example, the production of data at level NUTS2 through aggregation of data at level NUTS3 is typically a procedure that can be introduced in the data model (even if it is not so obvious at it looks ...).
- **The export pool (or external database)** can be defined as that part of the database where fully elaborated data are delivered either to the ESPON community but also to the external world. In the first case (internal delivery of data) it is typically a web application which users can access with restriction (ESPO members only) and download data files that have been validated and elaborated and have, in a sense received an official stamp « ESPON Database 2013 ». The choice of the best interface is not a question that will be solved immediately but a process of continuous improvement. As the aim of this export pool is to fulfil the objectives of the end users (i.e. the ESPON Projects, the ESPON CU, the ESPON MA), each new version will be presented to the community and improved according to feedback received. The first step from this point of view will be the presentation of a prototype of the export pool at the ESPON Seminar in Bordeaux in December 2008. We consider that ESPON Projects teams should be able to select and export the parts of the database that are useful for their projects. Because this data is an output of the database project, the

accompanying metadata will be consistent with the whole ESPON database and will be INSPIRE compliant. In the short run, since the export tool is in the process of being created, we propose that the data needed by the projects and available within the database will be required and then provided only through the canal of the ESPON DB representative within each project. This will ensure that the most updated information is used. A data request questionnaire will be designed and used for that purpose.

- **The Diffusion pool (publication) is the** normal form of delivery of data from ESPON to external world. The « publication » does not necessarily mean the use of a paper copy (it can be just a spreadsheet file to be downloaded from the ESPON 2013 website) but it is really equivalent to the printing of a book in the sense that data are published at a precise date where the data are considered as frozen. Any modifications of external data on the ESPON website should be considered as a « new publication », even if the change is minor.

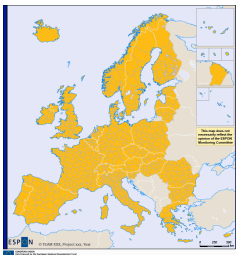
At each territorial level for each country, the completeness of data indicators is easily noticeable in a synthetic table.

COUNTRY	LEVEL	area_2003	pop_t_2003	gdp_eur_2003	gdp_pps_2003	actv_2003	unemp_2003	CLC03NATU_2000	Pop_density	GDP_per_inh	GDP_pps_per_inh	Productivity	Productivity_pps	Activity_rate	Unemp_rate	CLC03_per_inh	Completeness of data			
EU27+4	NUTS0	100	100	100	100	97	97	55	100	100	100	97	97	97	97	55				
	NUTS1	100	100	100	100	99	99	84	100	100	100	99	99	99	99	84				
	NUTS2	100	100	100	100	100	100	85	100	100	100	100	100	100	100	85				
	NUTS3	100	100	100	100	100	72	91	100	100	100	100	100	100	72	91				
Austria	NUTS0	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100				
	NUTS1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100				

### II.1 Delivery of tools for mapping (Mapkit tool)

The supplied new map kit tool for ESPON 2013 data base named Version 0.1 (MAP\_KIT\_v0.1) allows partners to reach three types folders:

- 📁 **A Eurogeographics map kit:** Available for the NUTS2003 and NUTS2006 version, this map kit is based on Eurogeographics geometries (20M).



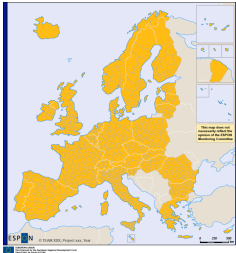
Data type: geometries

Format: shp + mxd

Projection: Lambert azimuthal equal area. 15°E50°N

Resolution: 1/20 000 000

📁 A generalized map kit: Available for the NUTS2003 and NUTS2006 version, this map kit is based on a more generalized geometry of territorial units (20M).



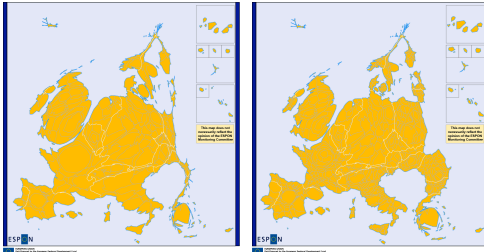
Data type: geometries

Format: shp + mxd

Projection: Lambert azimuthal equal area. 15°E50°N

Resolution: *adapted for thematic representation*

📁 An anamorphosis map kit: Available only for NUTS2003 version, this map kit proposes two new layers based on GDP and population for all nuts levels.



Data type: geometries

Format: shp + mxd

Projection: Lambert azimuthal equal area. 15°E50°N

Resolution: *adapted for thematic representation*

📁 A template folder: this folder contains an overview of maps for each map kit. It contains also different elements that the partners have always to put on the map (logo, disclaimer).



This map does not necessarily reflect the opinion of the ESPON Monitoring Committee



Data type: geometries

Format: eps (vector) + png (raster)



### **III. FLOWS FROM ESPON PROJECTS TO ESPON DB**

In this section, we describe the proposals for delivery of materials from ESPON Projects to ESPON DB 2013, in terms of data and maps.

#### **III.1 Lessons learned from first experiences of delivery of data**

Actually, only two projects has delivered data to be included in the new ESPON database. In both cases, problems has been encountered for an immediate inclusion of this data, because the rules are not actually précised on how this delivery should be made.

- *Update of demographic typology* : this data has been received from ESPON coordination unit the 21 November 2008 in the form of 4 excel files containing either raw data (directly download from Eurostat) either harmonised data where the target indicators are completed (i.e. population variation, natural variation, migratory variation, resulting typology) for a level that is a mixture of NUTS2 and NUTS3 units, different from the official mixture defined in the ESPON DB project. These files has been introduced in the "Import pool" but has not been transferred to the integration pool or the export pool as many questions has to be solved before to proceed to this integration. One of the most important is the fact that the typology is based on the use of basic Eurostat data (population, birth, death) that should be normally directly collected by ESPON DB project. In this situation, the real added value is the fact that the expert propose estimation of missing values (e.g. Denmark, Switzerland, ...) and the fact that he proposes formula for the estimation of target indicators. But this method should be normally available at all NUTS levels (NUTS0, NUTS1, NUTS2, NUTS23, NUTS3) in order to comply with the structure of the integration pool. We propose therefore to complete the work done by the expert before to realise the transfer toward the integration pool. And we suggest to store the methodology in order to be able to reproduce it in the future.
- *First delivery of data by ReRisk Project* : we have received the 28<sup>th</sup> November 2008 an excel file from Daniela Velte from ReRisk project containing one single indicator related to 2006 estimates of regional energy consumption of households in toe per inhabitants. The data are elaborated in NUTS2006 divisions and covers EU27 at NUTS2 level with some exceptions (Denmark is only available at NUTS0, Ultra peripheral regions of France are not available FR91 to FR 94). This data was apparently elaborated with help of DG Regio but they are no metadata at all in the Excel file and the origin is therefore unclear. Another problem is the fact that we have received the ratio (energy consumption. / inhabitant) but not the raw count variables that are behind. Finally, the question is open on how it could be possible to complete this file for non EU countries members of ESPON (CH, NO, IS, LI).

This two examples indicates clearly that there is an urgent need to propose some rules for delivery of data to ESPON Database Project, but at the same time, it is important to be aware that rules should be elaborated in common between, ESPON DB project and other ESPON projects. We have therefore suggested to CU to organize a specific meeting on this subject at the next ESPON seminar in Bordeaux. The rules that are proposed below should be therefore considered as a discussion paper.

### III.2 Rules for delivery of data to ESPON DB (to be discussed in Bordeaux)

We should first remind that for each project, one member of the database project is responsible for linking with the project. This person/team will ensure smooth information flows about data and updates in both directions. This role is particularly important in the short-run since the data flow and access to the database is not fully automatic.

As explained before, any data that is introduced in the ESPON database is firstly stored in the import pool « as such » without any modification. The *import pool* will be managed by the database project. Access is not directly given to each project since data needs to be checked for duplication, quality, and metadata compliance beforehand. It is only when data has been controlled that they can be transmitted to the integration pool or the export pool in order to be made available for ESPON Projects.

We can further identify two kinds of IN-FLOWS:

- EXTERNAL DATA, i.e. data gathered from a particular project directly from an external source (National offices, research institutions, other research projects, etc...)
- ESPON-MADE DATA, i.e. data created within the ESPON project.

This distinction is not always quite clear. For example, in the case of the datafile received from ReRisk project, it seems that the data was elaborated by DG Regio and it is therefore an External Data. But if ReRisk decide to complete the dataset with other sources and/or proceed to estimation of missing values, it is the second situation of an Espon-made data ... The same is true with the case of the update of demographic typology where basic datasets are clearly external data (Eurostat, National statistical offices) but where the resulting indicators and typology are a genuine ESPON data.

Data input from ESPON projects is composed of two types of files (preferably with the same name):

1. A data file containing the actual supplied dataset. A sample Excel file is presented in annex to this report as an example of normalized excel data file. The structure of the file is as follows:
  - a. A first data sheet (named "dataset") contains three columns with data describing the whole dataset: "NUTS\_version" describes the NUTS revision to which belong the territorial units described by the file (e.g. "NUTS\_2003"), "Date" describes the year or the date for the production of the indicator values (e.g. "2001" for the total population of EU27 countries in 2001) and "Indicator", describing the standard name of the indicator (e.g. "Total population").
  - b. The second data sheet (named "values") contains two columns: "Code" contains the NUTS code of all described territorial units, while the column "Value" contains the corresponding indicator value for that territorial unit.
2. An INSPIRE compliant XML metadata file describing the dataset. Metadata are necessary in order to keep track of the source, quality and methods used and will be augmented within the ESPON Database project with data completeness statistics. The metadata file are required to be filled using the INSPIRE metadata editor accessible online on the INSPIRE site at <http://www.inspire-geoportal.eu/InspireEditor/>. In more

advanced stages of the ESPON 2013 Database Project the INSPIRE editor will be replaced with an ESPON version, enriched with data dictionaries allowing a more precise description of the datasets. The tool will be implemented by extending one of the available open source INSPIRE compliant metadata editors (most probably CatMDEdit, accessible at <http://catmdedit.sourceforge.net/>).

### **III.3 Rules for delivery of maps and figures to ESPON DB (to be discussed in Bordeaux)**

Looking forward to the discussions, concerning the opportunity to choose a more generalized template as official one, during the meeting ESPON in Bordeaux, the partners can send realized maps as well according to template1 (Eurogeographics) as template2 (Generalized). A special 'send form' will be given to partners in order to help everybody to send elements in their appropriate shape.

**In any case, it is important to keep in mind that this rules are provisional and that a general "ESPON design" is currently elaborated by ESPON CU with help of expert advice.**

In the meantime, we can just suggest to ESPON project a minimum set of rules to be further discussed and elaborated at the ESPON Seminar in Bordeaux

**Rule 1** : Maps format must be vector shape file (ai, svg, emf, eps, ....).

**Rule 2** : Logos and disclaimer must be correctly included on maps (do not change the template).

**Rule 3** : The sources must be well identified and complete.

**Rule 4** : Data used to realize maps (if they are constructed or if they came from another source than EspoN DB 2013) must be given with their meta data.

**Rule 5** : Every map must be accompanied with a comment from 3 to 4 lines, with the possibility of an additional longer comment of 1500 signs.<sup>1</sup>

**Rule 6** : Design of borders and some countries must follow precise rules for political reason. In general, ESPON follows the rules established by European Commission. When these rules does not exists at EU level (for example because of lack of consensus) the rules of UN are used as reference. Actually, the situation has been clarified for the following situations. It is also important to consider that, out of political consideration, each state contributing to the ESPON Program has specific requests concerning the visualisation of his countries (remote territories, islands, ...).

- Rule 6.1 for Cyprus:

---

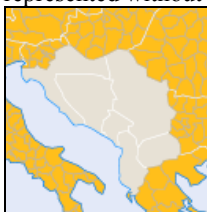
<sup>1</sup> This longer comment could be very usefull for external dissémination of ESPON results by the coordination Unit. For example, web publication of « Maps of the month » or printing of brochure like « ESPON briefing » or « ESPON in Progress ».

Cyprus is represented in two different colours. The two parts are separated with a NUTS line (without border line). The area not controlled by the government appears as “No data available”.



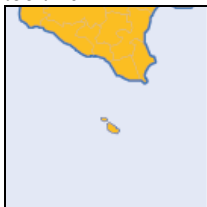
- Rule 6.2 for Kosovo:

According to the European Union, Kosovo is a province of Serbia. So, it has to be represented without border line.



- Rule 6.3 for Malta

If you use the Eurogeographics template, take care the lines of Malta's coats are not too thick



**(YES)**



**(NO)**

If you use the more generalized template, Malta is clearly apparent in any case.



- Rule 6.4 for Remote territories

Don't forget that the remote territories of France (Martinique, Guadeloupe, Guyane française, Réunion), Spain (Ceuta, Mellila, Canarias) and Portugal (Açores, Madeira) are territories members of the European Union that should be represented on maps. Therefore, you should never remove this régions from the Template (squares on the top right), even if data are not available for this regions. In this case, you let the area in blank with a legend « No data ».

## *Addendum : Catalogue of data currently available in the ESPON database*

### **(1) Data available at NUTS 2 and NUTS3 according to 2006 delimitation**

<b>Variable name</b>	<b>Definition</b>	<b>Unit</b>	<b>NUTS version</b>	<b>NUTS level</b>	<b>Years</b>	<b>% Missing values</b>
Area	Area of the regions (land use total)	Square kilometer (km2)	2006	3	1990-2007	27
Area	Area of the regions (land use total)	Square kilometer (km2)	2006	2	1990-2007	22
Population (males)	Annual average population (males)	1000 inh.	2006	3	1990-2006	35
Population (males)	Annual average population (males)	1000 inh.	2006	2	1990-2006	25
Population (females)	Annual average population (females)	1000 inh.	2006	3	1990-2006	35
Population (females)	Annual average population (females)	1000 inh.	2006	2	1990-2006	25
Total population	Annual average population (both sex)	1000 inh.	2006	3	1990-2006	12
Total population	Annual average population (both sex)	1000 inh.	2006	2	1990-2006	10
GDP	Gross domestic product (GDP) at current market price	Millions of euro (from 1.1.1999)/Millions of ECU (up to 31.12.1998)	2006	3	1995-2005	8
GDP	Gross domestic product (GDP) at current market prices	Millions of euro (from 1.1.1999)/Millions of ECU (up to 31.12.1998)	2006	2	1995-2005	3
GDP (PPS)	Gross domestic product (GDP) at current market prices	Millions of Purchasing Power Parities	2006	3	1995-2005	8
GDP (PPS)	Gross domestic product (GDP) at current market prices	Millions of Purchasing Power Parities	2006	2	1995-2005	3
Active population	Economically active population (both sex - 15 years and over)	1000 inh.	2006	3	1999-2007	20
Active population	Economically active population (both sex - 15 years and over)	1000 inh.	2006	2	1999-2007	8
Unemployment	Unemployment (both sex - 15 years and over)	1000 inh.	2006	3	1999-2007	37
Unemployment	Unemployment (both sex - 15 years and over)	1000 inh.	2006	2	1999-2007	8

## 2. Data available at NUTS 2 and NUTS3 according to 2003 delimitation

Variable name	Definition	Unit	NUTS version	NUTS level	Years	% Missing values
Area	Area of the regions (land use total)	Square kilometer (km <sup>2</sup> )	2003	3	1990-2005	27
Area	Area of the regions (land use total)	Square kilometer (km <sup>2</sup> )	2003	2	1990-2005	22
Population (males)	Annual average population (males)	1000 inh.	2003	3	1990-2004	33
Population (males)	Annual average population (males)	1000 inh.	2003	2	1990-2004	23
Population (females)	Annual average population (females)	1000 inh.	2003	3	1990-2004	33
Population (females)	Annual average population (females)	1000 inh.	2003	2	1990-2004	23
Total population	Annual average population (both sex)	1000 inh.	2003	3	1990-2004	6
Total population	Annual average population (both sex)	1000 inh.	2003	2	1990-2004	6
GDP	Gross domestic product (GDP) at current market prices	Millions of euro (from 1.1.1999)/Millions of ECU (up to 31.12.1998)	2003	3	1995-2004	2
GDP	Gross domestic product (GDP) at current market prices	Millions of euro (from 1.1.1999)/Millions of ECU (up to 31.12.1998)	2003	2	1995-2004	2
GDP (PPS)	Gross domestic product (GDP) at current market prices	Millions of Purchasing Power Parities	2003	3	1995-2004	2
GDP (PPS)	Gross domestic product (GDP) at current market prices	Millions of Purchasing Power Parities	2003	2	1995-2004	3
Active population	Economically active population (both sex – 15 years and over)	1000 inh.	2003	3	1999-2005	14
Active population	Economically active population (both sex – 15 years and over)	1000 inh.	2003	2	1999-2005	5
Unemployment	Unemployment (both sex – 15 years and over)	1000 inh.	2003	3	1999-2005	30
Unemployment	Unemployment (both sex – 15 years and over)	1000 inh.	2003	2	1999-2005	6
Land use	Land use by CORINE category	ha	2003	3	2000	9
Land use	Land use by CORINE category	ha	2003	2	2000	15
Land use	Land use by CORINE category	ha	2003	3	1990	9
Land use	Land use by CORINE category	ha	2003	2	1990	15