

1 – Semantic expertise



From LUZ 2001 to LUZ 2004: spatial and temporal variability

2001: Functional  Administrative

NAME	Clear functional definition	Probable functional definition*	Administrative definition, except for Capital city	Administrative definition
Belgium				
France				
Sweden				
Finland				
Germany				
Ireland				
Luxemburg				
Netherlands				
United Kingdom				
Austria				
Denmark				
Greece				
Italy				
Portugal				
Spain				

2004: Functional  Administrative

NAME	Clear functional definition	Probable functional definition*	Administrative definition, except for Capital city	Administrative definition
Belgium				
Croatia				
France				
Netherlands				
Portugal				
Spain				
Sweden				
Finland				
Germany				
Norway				
United Kingdom				
Austria				
Cyprus				
Denmark				
Slovakia				
Slovenia				

 Type of definition
 Change between two dates

*: The methodology used to build LUZ is described very shortly in National Final Reports, Methodological Handbooks or other official sources and need to be clarified by furthered research

2 – The UMZ database as a benchmark for databases comparison

- ❑ Basically, **interesting properties**:
 - quite automatic construction process
 - objective and reproducible

 - relatively clear specifications
 - traceable

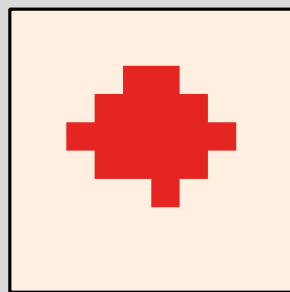
 - same statistical and geometrical sources over all Europe
 - internationally coherent, completeness
 - geometrically compatible

- ❑ Not so used until now, **need to be enriched** by other attributes :
 - no name : how to give a name to each object?
 - in which country? In which region?
 - updating population (last version of grid pop.)
 - giving a geometrical center (centroid)

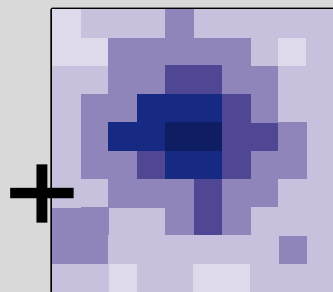
2 – The UMZ database as a benchmark for databases comparison

Giving a name to UMZ:

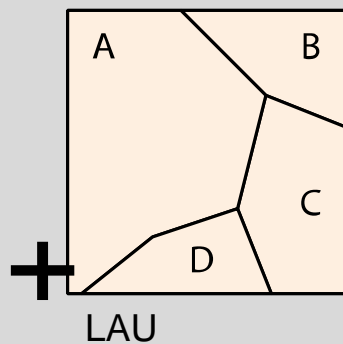
Attributing the **LAU name** (one or several) corresponding to the core of the agglomeration



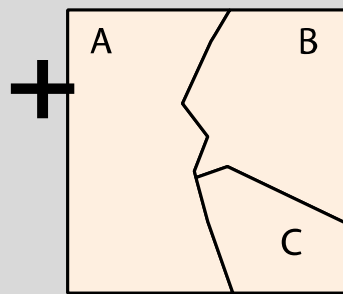
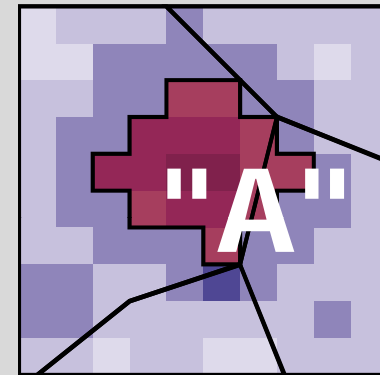
UMZ



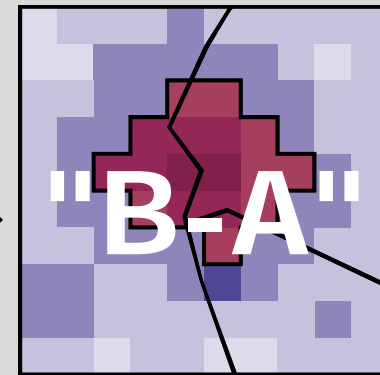
Pop. density disaggregated with Corine land cover 2000, **version 4.1** (last version)



LAU



LAU



1 - clear morphological core within one LAU (>50% pop UMZ) → **1 name**

2 - otherwise → **2 or more names** (pop LAU ∩ UMZ >50% pop main LAU ∩ UMZ)

3 – Cross expertise for data base comparisons (comparison protocol)

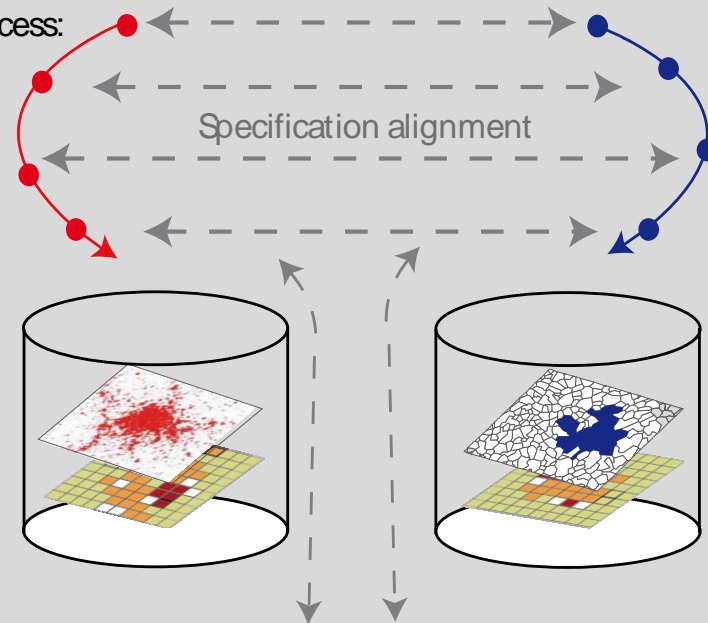
Step 1: Extracting the rules used to build urban objects: spatial relations, thresholds...

Step 2: Aligning specifications

Step 3: Qualitative evaluation of the quantitative differences between data bases

Specifications of the construction process:

- spatial and temporal resolution
- constraints
- definition of the core class and parameters of the resolution
- aggregation methods

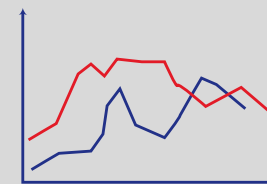


First Results:
MUAs and UMZs

Comparison and identification
of the different sources of differences:

- at macro level
- term to term

Qualitative and quantitative evaluation



Challenge 7 : Extra-ESPON data exchange

Objectives :

- Exchange of data and expertise with main data providers
- Identify long-run synergies
- Promotion of ESPON DB

Situation:

- Ongoing contacts with Eurostat (update NUTS2&3) and EEA (grids)

Strategy :

- Set up of regular meetings, strategic and technical, with institutions and CU
- 1 DB team <-> 1 external institution

Coord : University of Luxembourg



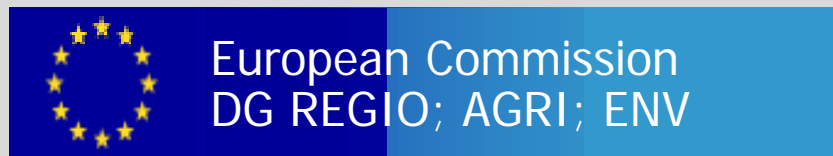
7.1. Activities and progresses



European Environment Agency



ORGANISATION
FOR ECONOMIC
CO-OPERATION
AND DEVELOPMENT



Others (local): National Statistics Offices,
Regional observatories, etc...

Developing
2-way FLOWS of

DATA
NUTS
WUTS
GRIDS

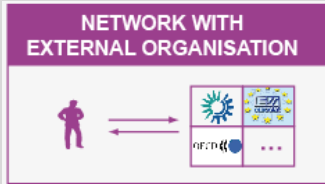
EXPERTISE &
KNOWLEDGE
in
Spatial Data

Analysis,
Integration,
Management (meta's)
Publication (web)



ESPON
DB2013

7.1. Activities and progresses

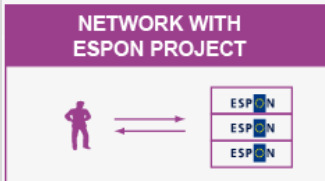


Networking activities

Follow-up on EUROSTAT-ESPON Action Plan

Information exchange with Commission Services

Wider networking (e.g. ESPON workshop on 6 May, Brussels)



Thematic access to the ESPON Database

The capacities of using a thesaurus for information retrieval

Literature review of standardised relationships for thesauri

Examples of online thesauri developed by international organisations

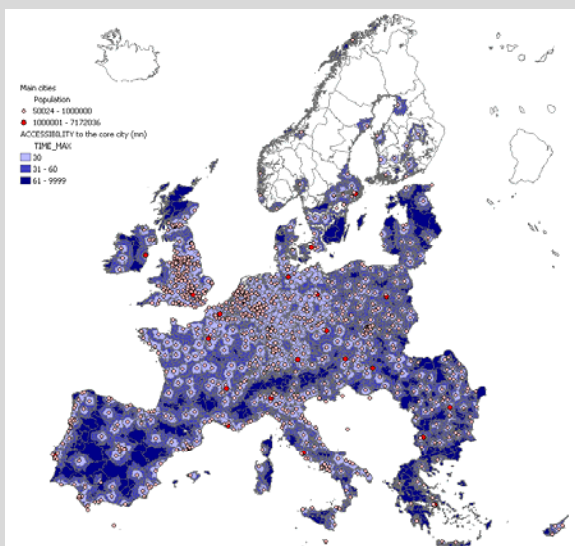
The added-value of the experience acquired by the ESPON 2006 database

Explore qualitative text analysis tools for thematic structuring

7.2. Networking activities: External data flow

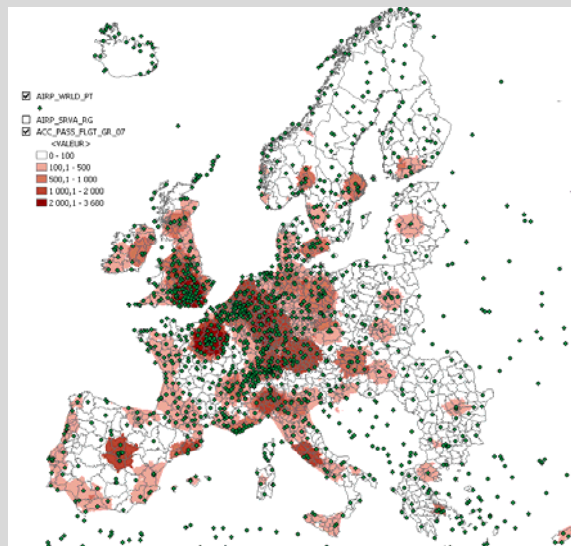
DG REGIO: Geodatabases describing various topics

Accessibility to cities



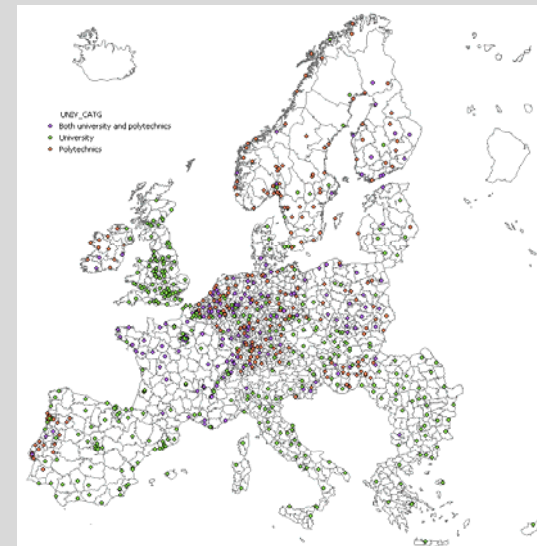
This polygon feature class represents accessibility zones to cities having more than 50 000 inhabitants. Accessibility zones have been divided into four time breaks: 0-30 min, 30-45 min, 45-60 min and more than 60 min.

Accessibility to passenger flights



This raster gives the cumulated daily number of passenger flights available within 90 minutes of travel by road for each square kilometer cell in Europe in 2007.

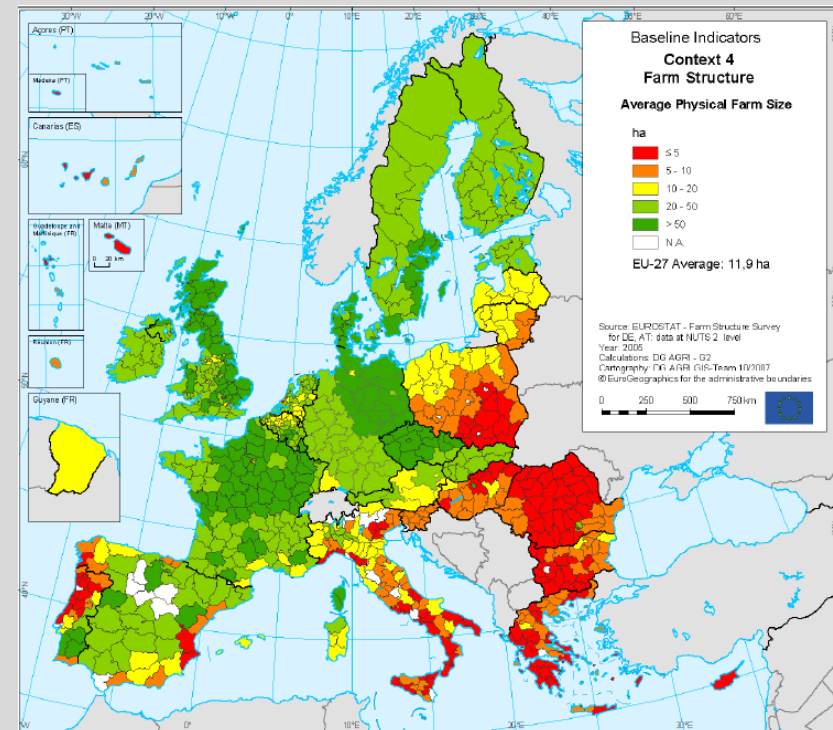
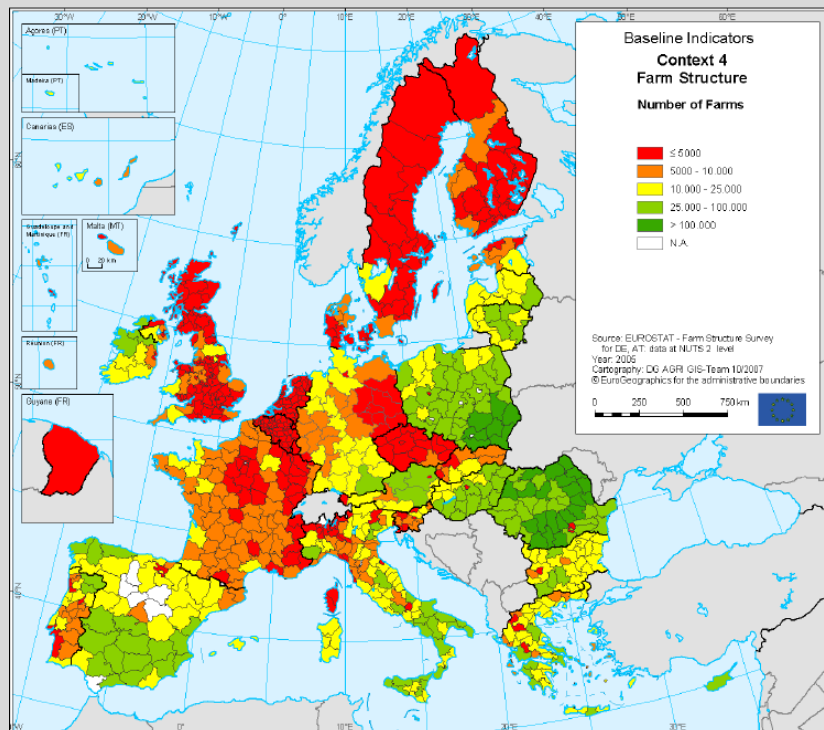
Location and characteristics of universities



Point locations of universities in the EU (number of students by location), Based on the point layer from the Study on Mountain Areas in Europa (2004) (DG REGIO / Nordregio)

7.2. Networking activities: External data flow

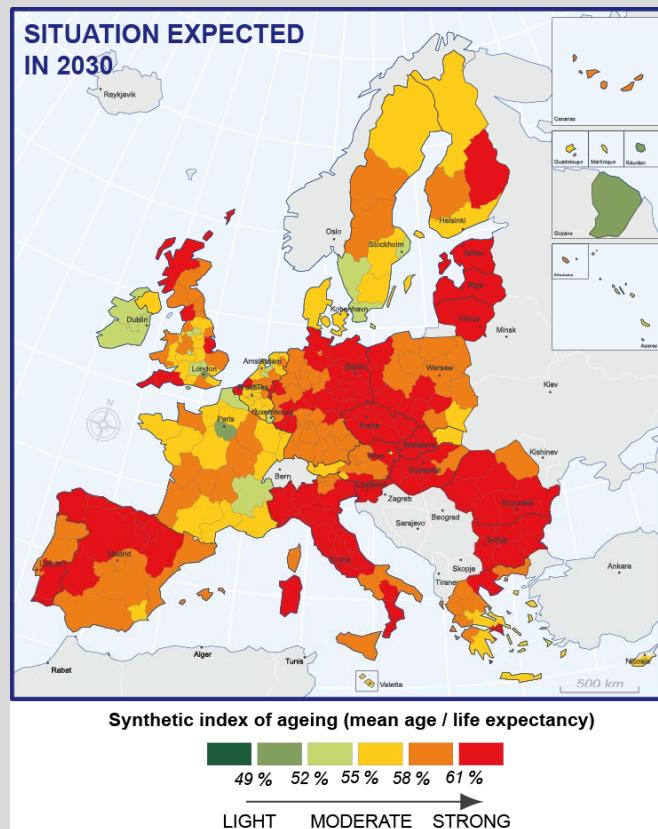
DG AGRI: Lot of data at NUTS2 and NUTS3 level, in particular concerning agriculture settlements and employment



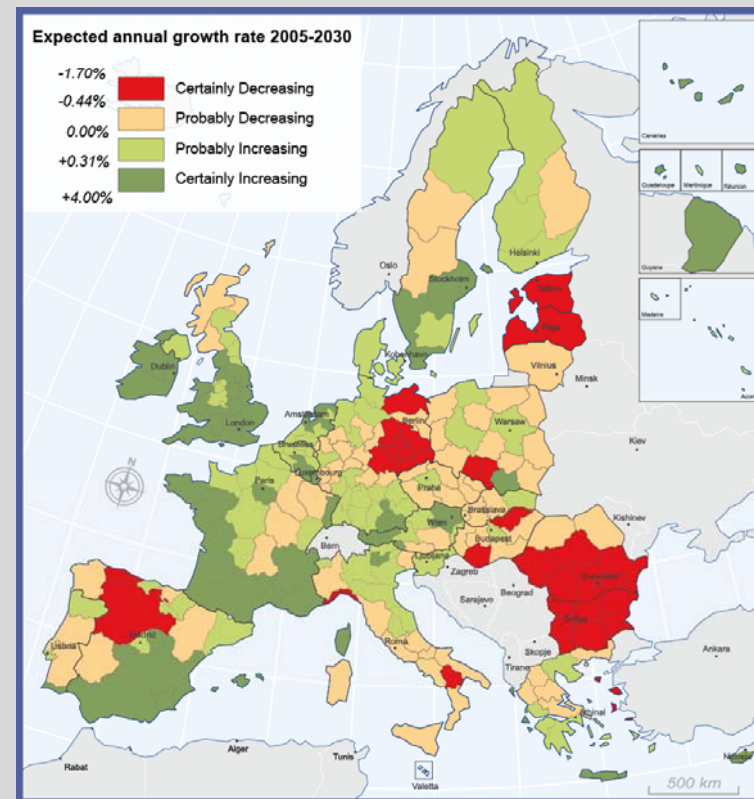
7.2. Networking activities: External data flow

European Parliament: Results of studies

Synthetic index of ageing



Population projections



7.2. Networking activities: External information flow

JRC: Inspire metadata profile – ISO19115 Norm

The screenshot shows the 'INSPIRE Metadata Editor' web application. The browser address bar displays 'http://www.inspire-geoportal.eu/InspireEditor/'. The page header includes the INSPIRE Geoportal logo and the JRC (Joint Research Centre) logo, with the text 'the EU portal for Geographic Information' and 'EUROPEAN COMMISSION'. Below the header, there are navigation links: 'Exit', 'Save', 'Validation (ON)', and 'User Guide'. A tabbed interface is visible with the following tabs: 'Metadata', 'Identification', 'Classification', 'Keyword', 'Geographic', 'Temporal', 'Quality&Validity', 'Conformity', 'Constraints', and 'Organization'. The 'Metadata' tab is active, showing a section titled 'Metadata on metadata'. This section contains a 'Metadata point of contact' form with the following fields and controls:

- E-mail Address:** A text input field with an 'Add' button to its right.
- Remove Selected:** A button located below the E-mail Address field.
- Organisation Name:** A text input field with an 'Add' button to its right.
- Remove Selected:** A button located below the Organisation Name field.
- Metadata date:** A date input field showing '2009-01-31' with a calendar icon.
- Metadata language:** A dropdown menu currently set to 'English'.

At the bottom of the browser window, the status bar shows 'Terminé' on the left and the 'zotero' logo on the right.

7.2. Networking activities: External information flow

European Environmental Agency: EEA profile for geographical data

Metadata Editor - EEA-MSGI V1.1d - Urban Morphological Zones 2000 F1v0 (vector)

Metadata On Metadata | Dataset Identification | Distribution Information | Other Information | Validate and Publish | Help

* Point of contact

- * Organisation name: European Environment Agency
- * Individual name: Andres Bastholm
- Position name:
- Address: Delivery point: Kongens Nytorv 6
- Address: City: Copenhagen
- Address: State, Province: K
- Address: Postal Code: 1050
- Address: Country: Denmark
- Address: E-mail: eea@eea.europa.eu

Stored contacts


←

Add contact

Last modified: 20090420 (YYYYMMDD)

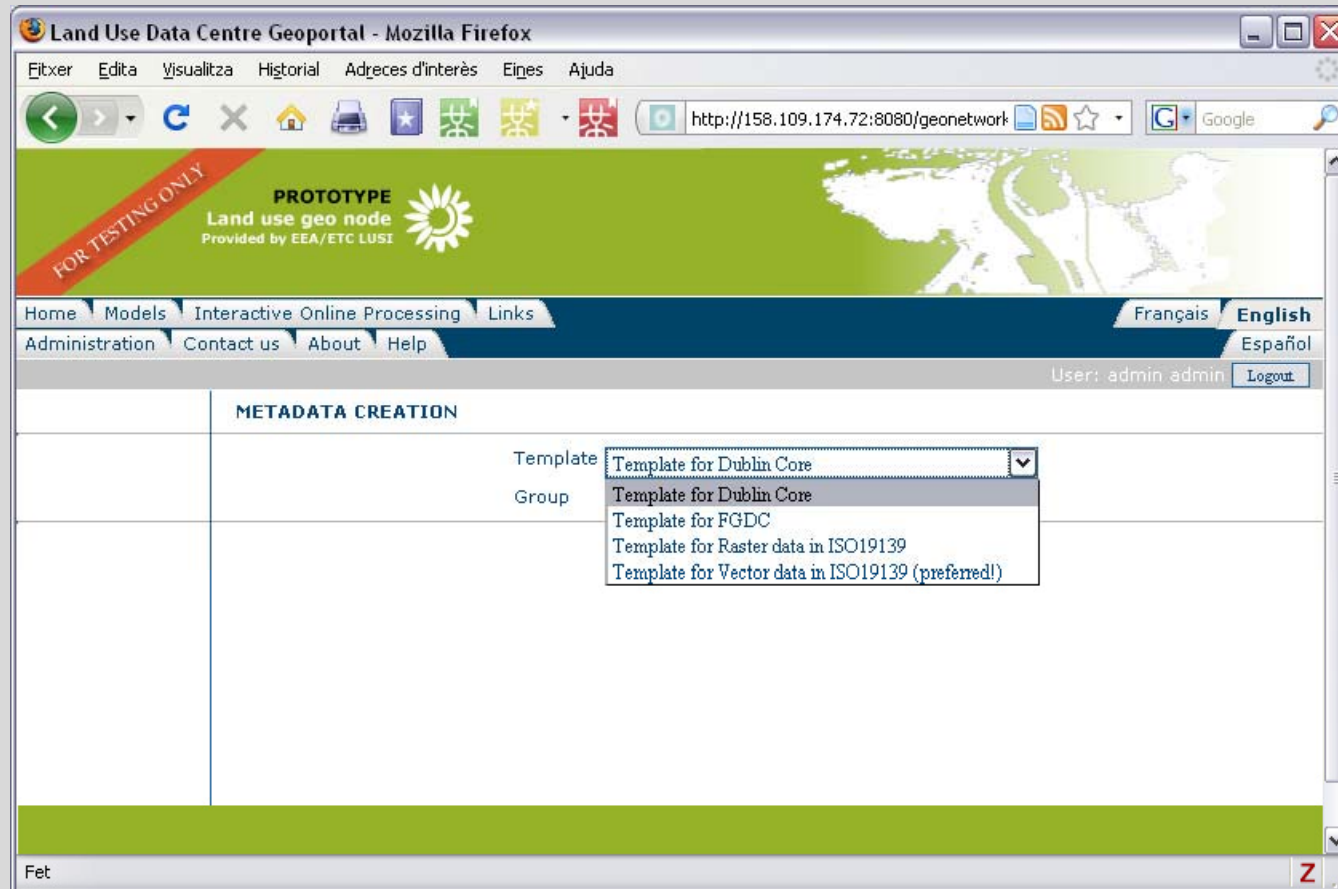
Name of standard: EEA-MSGI/ISO19115 (First Edition)

Version of standard: EEA-MSGI 1.1d

European Environment Agency 

7.2. Networking activities: External information flow

Geonetwork: An open source solution for metadata



7.3. Toward an ESPON Thesaurus ?

Some considerations in the creation of an ESPON thesaurus:

- Preserve a **neutral-technical** approach
- Explore **quantitative text analysis** tools for defining appropriate vocabulary and support an efficient indexing
- Improve the ability in defining standards and further advance on the **harmonisation and coherence** of concepts
- Ensure the overall meaning of terms for **information retrieval** by final users

Next steps towards the thematic structuring of the ESPON Database:

- Deepen the **literature review** on controlled vocabulary (i.e. thesauri) developed by other international organisations (e.g. OECD, ILO, UNESCO)
- Perform **qualitative and quantitative text analysis** on various scientific reports (ESPON, DGs, CoR, EESC, European Parliament)
- Determine links between different terms, namely by **word co-occurrence analysis** and **types of relationships**
- Link those terms with data for information retrieval (indicators, typologies) and **structure the ESPON DB 2013**

8.3 Toward an ESPON Theasaurus ?

01 Spatial Typologies	10 Household-Oriented Infrastructure
02 Population	102 <i>Education</i>
021 <i>Population structure</i>	11 Land use
022 <i>Population movement</i>	111 <i>Natural resources</i>
03 Employment and Labour Market	112 <i>Land use</i>
031 <i>Employment and sector structure</i>	12 Environment
032 <i>Structure of persons employed</i>	124 <i>Natural hazards</i>
033 <i>Unemployment</i>	13 Agriculture
04 Wealth and Production	131 <i>Land use</i>
042 Income and consumption	132 <i>Farmer structure</i>
06 Transport	133 <i>Employment</i>
061 <i>Transport infrastructure</i>	134 <i>Livestock</i>
062 <i>Passengers and goods transport</i>	135 <i>Production</i>
066 <i>Accessibility</i>	14 Social Situation
067 <i>Impacts of transport policies</i>	141 <i>Poverty</i>
07 Research and Development	17 Tourism
071 <i>Invention and innovation</i>	171 <i>Arrival and stays</i>
072 <i>Facilities and employment</i>	18 Public Sector
073 <i>Finance and expenditures</i>	183 <i>Regional policy</i>
09 Communication Technology	19 Other Data
091 <i>Infrastructure</i>	191 <i>Area</i>

Source: ESPON Project 3.4.1

7.3 Toward and ESPON Thesaurus ?

archipelago borders centre-periphery connections cooperation cities countries

database demographic development **eastern**

economic enlargement **Europe evidence** flows

functional labour **integration** Mediterranean **metropolitan**

migration monitoring neighbourhood

networks **political population** relations regional **regions** strategy

southern spatial **territorial** trade **visions** western **world**

Table 1 – Most commonly used words in the ‘Europe in the World’ report.

Note: Ranking frequency is defined by letter size.

Challenge 8 : Intra-ESPON data exchange

Objectives :

- 1) **Provide** an efficient interface between ESPON projects (Priority 1 and Priority 2) and statistical or cartographic institutions (Eurostat, EEA, Eurogeographics...)
- 2) **Stimulate** the launching of future ESPON Projects by incorporating indicators from neighbouring countries (Balkans, Ukraine, Belarus, Moldova, Maghreb, Turkey...) or elaborating long term (1960-1980-2000) series for a selected number of indicators
- 3) **Ensuring** the quality of the data flow with ESPON

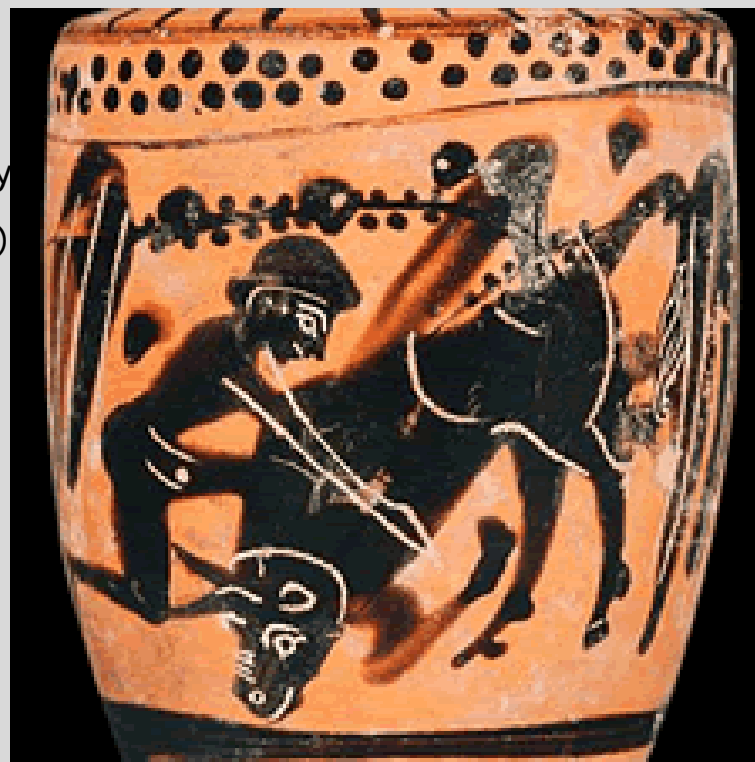
Situation :

- 1) Regular contacts with some TPG's are effective
- 3) A metadata model ISO 19115 compliant has been established. This template will allow to integrate in a short term indicators from ESPON TPG's into the database

Next steps :

- 2) Technical report on data availability in neighbouring countries (in relation with challenge 10)
- 3) Creation of a metadata editor into the ESPON DB

Coord : RIATE & LIG



8.1 The metadata model

An operational short-term solution: Excel implementation

Dataset level

Metadata information		
point of contact	email	manager@espondb.eu
	organization	"ESPON 2013 Database project"
	last update date	28/05/2009
data filename	ESPON_basic_indicators_2003.xls	
copyright	ESPON 2013 Program	
use rights	free	free text
access rights	free	free text
metadata read right	yes	boolean: indicates whether metadata can be read
Maintenance information		
productionRegular	Yes	if some new acquisition are planned
productionFrequency	yearly	indicates to the user if can be used
Spatial representation Information		
nomenclature	NUTS	or WUTS, or UMZ, etc...
version	2003	free text

Indicator level

Indicator 1 = sheet 1						
1	Identification	area				
2	code	area				
3	name	area				
4	units	km ²				
5	abstract	Area of the region (land use total)				
6	Methodology					
7	language	english				
8	Classification	theme*	General			
9	keywords*	Supercity, Geometry				
10	Temporal extent					
11						
Indicator 2 = sheet 2						
1	Identification	pop_1				
2	code	pop_1				
3	name	Population				
4	units	Thousands inhabitants				
5	abstract	Annual average population (both sex)				
6	Methodology					
7	language	english				
8	Classification	theme*	Demography			
9	keywords*	Population				
10	Temporal extent					
11						
12	start	2003				
13	end	2003				
Indicator 3 = sheet 3						
1	Identification	gdp_eur				
2	code	gdp_eur				
3	name	GDP in Euro				
4	units	Million Euros				
5	Methodology					
6	abstract	Gross Domestic Product (Euros)				
7	language	english				
8	Classification					
9	theme*	Economy				
10	keywords*	GDP, Euro				
11	Temporal extent					
12	start	2003				
13	end	2003				
Indicator 4 = sheet 4						
1	Identification	Indicator_Metaddata_0				
2	code	Indicator_Metaddata_0				
3	name	Indicator_Metaddata_0				
4	units					
5	Methodology					
6	abstract					
7	language	english				
8	Classification					
9	theme*					
10	keywords*					
11	Temporal extent					
12	start	2003				
13	end	2003				

Record level

Lineage			
14	scope		
15	label	temporal	1
16	temporal	start	2003
17		end	2003
18	lineage	provider	EUROSTAT
19		date	07/2007
20		URL	http://epp.eurostat.ec.europa.eu/
21		methodology	
22	Reliability	Official estimation	Yes
23		Quality	High
24		estimation	yes/no
25		Quality	low - medium
26	scope		
27	label	temporal	2
28	temporal	start	2003
29		end	2003
30	lineage	provider	EUROSTAT
31		date	07/2008
32		URL	http://epp.eurostat.ec.europa.eu/
33		methodology	
34	Reliability	Official estimation	Yes
35		Quality	High
36		estimation	yes/no
37		Quality	low - medium
38	scope		
39	label	temporal	2'
40	temporal	start	2003
41		end	2003
42	lineage	provider	EUROSTAT
43		date	07/2008
44		URL	http://epp.eurostat.ec.europa.eu/
45		methodology	Aggregation from NUTS3 level/NE
46	Reliability	Official estimation	No
47		Quality	Low
48		estimation	yes/no
49		Quality	low - medium
50	scope		
51	label	temporal	2'
52	temporal	start	2003
53		end	2003
54	lineage	provider	EUROSTAT
55		date	07/2008
56		URL	http://epp.eurostat.ec.europa.eu/
57		methodology	Aggregation from NUTS3 level/NE
58	Reliability	Official estimation	No
59		Quality	Low
60		estimation	yes/no
61		Quality	low - medium

- Based on the INSPIRE initiative
- Following the ISO and SDMX standards
- This solution allows a three tiers characterisation: the whole dataset level, the variable level and the record level

8.2 Networking with TPG's

ESPON 2013 DATABASE – Metadata Guide-lines for territorial units



ESPON 2013 DATABASE

GUIDE-LINES FOR METADATA ADAPTED TO TERRITORIAL UNITS WITHIN ESPON 2013 PROJECTS.

Provisional document

ESPON 2013 DATABASE – Metadata Guide-lines for territorial units

The metadata challenge

Scarcity of data documentation within the previous ESPON program has been seen as an important impediment to the building and use of the ESPON database. Difficulties stemmed from uncertainties about legal constraints, sources, units, etc...

In the ESPON database 2013 project, the ESPON database will be further enriched and expanded in the time, spatial and thematic dimensions (see the First Interim Report of ESPON Database project). Information about the data made available is thus even more crucial. Building a rich database would be useless without a strong effort to describe precisely the data that have been gathered and integrated within the database. This information about data is known as *metadata*. **Creating and organising metadata is therefore an additional, important, and transversal challenge for the ESPON community.**

To be useful for ESPON projects and other end-users, data should always be accompanied by metadata, including information about their quality, their sources, and their lineage. It is also particularly important that metadata are created in a manner that is consistent with international standards (INSPIRE, SDMX) so as to ensure the use of the database in the longer-run and to insure compatibility with other national and international database initiatives.

The **short-term strategy** is aimed at proposing a metadata solution for ongoing ESPON projects (Priority 1 and 2) when exchanging data with the ESPON Database. This provisional solution is needed as long as our project is developing the tools that are necessary in order to ease the edition of good metadata. Defining a complete and non-redundant metadata model is a research in itself and should be made carefully because it conditions future data quality and access. The short-term metadata solution is based on a simple spreadsheet format, implementing the ESPON database metadata model, which is to accompany data provided to the database by the different projects. The spreadsheet has the same conceptual structure as the one of the mid term solution. This metadata profile is adapted for statistical data gathered on geographical objects defined by territorial footprints (like NUTS territorial units or cities). Of course, the conceptual model defined at this stage is likely to evolve and be extended if the need to accommodate different types of geographical objects or indicators requires it.

The **mid-term solution** will be based on a **web-editor** and linked to the Database Import pool. We have explored a few metadata editors in order to find out the most suitable one for the ESPON community (Inspire Metadata Editor, CadMEdit, Geonetwork,...). We found that a web metadata editor (rather than a desktop one) should be preferred in the longer-run because it requires no install (and updates) from the users but also because it can more easily provide help to the users when filling in the different necessary fields.

The aim of this paper is to show what is the conceptual statistical metadata model that has been defined by the ESPON Database 2013 project and, with the help of some concrete examples (datasets with various indicators, datasets dealing with time-series, contingency tables like the age pyramid), to illustrate how to produce data and metadata compliant with the ESPON Database project.

It is important to say that the metadata model presented by the ESPON Database project is aimed at having some statistical information as detailed as possible (importance of the lineage of the information), while other information (that can be extracted from the data themselves, like completeness, spatial extent, etc.) are left

Challenge 9 : Data model and integration

Objectives : To build an operational software architecture that ensures the importation, the integration and the exportation of sets of indicators (data)

Situation :

Strategy :

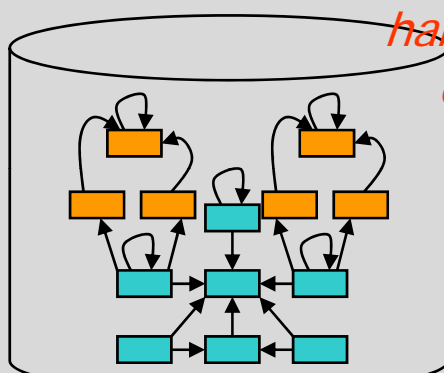
Coord : LIG



Operational implementation of the ESPON 2013 Database application: inclusion of metadata

id	level	pop.J	inv.pop.J	inv.pop.J_200
1	TEMPORAL_START	2000	2001	200
2	TEMPORAL_END	2000	2001	200
3	TEMPORAL_START	2001	2002	30
4	#111 NUTS3	381.1	377.1	37
5	#112 NUTS3	137.7	140.8	141
6	#113 NUTS3	100.3	87.9	87
7	#121 NUTS3	234.6	237.7	237
8	#122 NUTS3	248.5	245.9	246
9	#123 NUTS3	140.3	142.2	142
10	#124 NUTS3	235.9	223.9	223
11	#125 NUTS3	129.1	123.3	123
12	#126 NUTS3	256.5	276.8	276
13	#127 NUTS3	292.3	292.8	294
14	#130 NUTS3	1551.2	1550.3	1573
15	#211 NUTS3	287.1	269.4	25
16	#212 NUTS3	131.7	131.8	131
17	#213 NUTS3	161.4	159.1	156
18	#214 NUTS3	366.1	360	365
19	#222 NUTS3	81.2	82.2	82
20	#223 NUTS3	188.8	176	174
21	#224 NUTS3	298	268.2	268
22	#225 NUTS3	187	190.5	190
23	#226 NUTS3	111.3	109.4	106

ESPON MEGABASE



*harmonization
enriching*

Data and metadata files

Importation
checking

Transfer

*dynamic
generation*

Exportation

id	level	pop.J	inv.pop.J	inv.pop.J_200
1	TEMPORAL_START	2000	2001	200
2	TEMPORAL_END	2000	2001	200
3	TEMPORAL_START	2001	2002	30
4	#111 NUTS3	381.1	377.1	37
5	#112 NUTS3	137.7	140.8	141
6	#113 NUTS3	100.3	87.9	87
7	#121 NUTS3	234.6	237.7	237
8	#122 NUTS3	248.5	245.9	246
9	#123 NUTS3	140.3	142.2	142
10	#124 NUTS3	235.9	223.9	223
11	#125 NUTS3	129.1	123.3	123
12	#126 NUTS3	256.5	276.8	276
13	#127 NUTS3	292.3	292.8	294
14	#130 NUTS3	1551.2	1550.3	1573
15	#211 NUTS3	287.1	269.4	25
16	#212 NUTS3	131.7	131.8	131
17	#213 NUTS3	161.4	159.1	156
18	#214 NUTS3	366.1	360	365
19	#222 NUTS3	81.2	82.2	82
20	#223 NUTS3	188.8	176	174
21	#224 NUTS3	298	268.2	268
22	#225 NUTS3	187	190.5	190
23	#226 NUTS3	111.3	109.4	106

#111	NUTS3	815.525661	845.709883	857.788843
#112	NUTS3	2768.837348	2260.319326	3070.418371
#113	NUTS3	1678.383794	1460.872282	1662.723874
#121	NUTS3	4704.933194	4981.938054	5111.868838
#122	NUTS3	4651.288077	4495.524466	5136.988212
#123	NUTS3	3699.53075	3622.286282	3954.918701
#124	NUTS3	4132.457209	4218.886783	4431.813555
#125	NUTS3	1787.462222	1743.627916	1823.239887
#126	NUTS3	4303.037746	6278.889562	6357.478887
#127	NUTS3	5383.258354	9800.452365	10319.828111
#130	NUTS3	80814.22287	81542.81847	83021.868229
#211	NUTS3	7284.128787	7350.352236	7725.023766

ESPON DATABASE

id	level	pop.J	inv.pop.J	inv.pop.J_200
1	TEMPORAL_START	2000	2001	200
2	TEMPORAL_END	2000	2001	200
3	TEMPORAL_START	2001	2002	30
4	#111 NUTS3	381.1	377.1	37
5	#112 NUTS3	137.7	140.8	141
6	#113 NUTS3	100.3	87.9	87
7	#121 NUTS3	234.6	237.7	237
8	#122 NUTS3	248.5	245.9	246
9	#123 NUTS3	140.3	142.2	142
10	#124 NUTS3	235.9	223.9	223
11	#125 NUTS3	129.1	123.3	123
12	#126 NUTS3	256.5	276.8	276
13	#127 NUTS3	292.3	292.8	294
14	#130 NUTS3	1551.2	1550.3	1573
15	#211 NUTS3	287.1	269.4	25
16	#212 NUTS3	131.7	131.8	131
17	#213 NUTS3	161.4	159.1	156
18	#214 NUTS3	366.1	360	365
19	#222 NUTS3	81.2	82.2	82
20	#223 NUTS3	188.8	176	174
21	#224 NUTS3	298	268.2	268
22	#225 NUTS3	187	190.5	190
23	#226 NUTS3	111.3	109.4	106

Data and metadata files

Statistical metadata profile for ESPON 2013 Database

Existing metadata standards (INSPIRE, ISO 19115) are not precise enough for use with statistical data : **without good metadata, data devalue on the long term.**

What's new ?

- **A complete and non-redundant profile for statistical metadata**
 - Complete: allow traceability of source and quality for long term storage
 - Non-redundant: any metadata that can be automatically derived is not required from the projects
 - **May seem cumbersome** to fill, **but** it's at least equally benefic for data searching (in catalogues) and for understanding data
- This **profile is in the form of an excel file and has been distributed to ESPON projects** accompanied by a set of **explanatory guidelines and examples**
- A first prototype of metadata Web editor based on the Geonetwork Metadata Editor
 - More user friendly, universally accessible via a Web browser

Excel metadata template

	A	B	C	D	E	F
3		email	manager@espondb.eu	an email		
4		organization	"ESPON 2013 Database project"	ESPON internal project name ou "ESPON DB 2013"		
5		last update date	31/05/2009	a date : DD/MM/YYYY		
6		data filename	Delivery_june_2009.xls			
7		Constraint information This rubrique is optional : fill if the data comes with a particular right for use and diffusion				
8		copyright	ESPON 2013 Program	free text : the copyright attached to the provided data		
9		use rights	free	free text : usage is authorized		
10		access rights	free	free text : further diffuion of this dataset is authorized		
11		metadata read right	yes	boolean : indicates whether metadata could be exposed on ESPON DB web site		
12		Maintenance information				
13		productionRegular	Yes	if some new acquisition are planned out for some sub-set of units, precise this by free text		
14		productionFrequency	yearly	indicates to the user if can can expect further acquisition of those data, and in which time.		
15		Spatial representation Information				
16		nomenclature	NUTS	or WUTS, or UMZ, etc..		
17		version	2006	free text		
18						
19						
20						
27						
28						
29						
30						
31						
32						
33						
34						
35						
36						
37						
38						

Dataset metadata

Excel metadata template

The screenshot shows an Excel spreadsheet with a metadata template. The template is organized into sections: Identification, Methodology, Classification, Temporal extent, Lineage, and Reliability. The 'Indicator metadata' section (rows 1-13) includes fields for code, name, units, abstract, language, theme, keywords, start, and end. The 'Value metadata' section (rows 15-43) includes fields for label, temporal, lineage, provider, date, URL, methodology, reliability, official estimation, and quality. The spreadsheet is titled 'Metadata_delivery_june_2009.xls' and shows a grid with columns A through Q and rows 1 through 51.

Section	Field	Value	Notes
Identification	code	pop_t	
	name	Total population	free text
	units	Thousands inhabitants	free text or choose in a list
	abstract	Annual average population (both sex)	free text
	Methodology		free text, Methodology used to calculate the indicator (in the case of complex indicators, typologies, ratios etc)
Methodology	language	english	
Classification	theme	Demography	use a thesaurus (GEMET, SDMX or other)
	keywords	Population	use a thesaurus (GEMET, SDMX or other)
Temporal extent	start	2000	
	end	2006	
Lineage	scope	1	
	temporal		
	start	2000	
	end	2006	
	lineage		
	provider	EUROSTAT	
	date	02/2009	
	URL	http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1090,30070682,1090_33078576&_dad=portal&_schema=PORTAL	
	methodology		
	Reliability		
Official estimation	Yes	yes no	
Quality	No	yes no	
	High	low - medium - high	
Value metadata	scope	2	
	temporal		
	start	2000	
	end	2006	
	lineage		
	provider	ESPON Database 2013 Project	
	date	02/2009	
	URL		
	methodology	Sum of NUTS2	
	Reliability		
Official estimation	No	yes no	
Quality	Yes	yes no	
	High	low - medium - high	
Value metadata	scope	3	
	temporal		
	start	2000	
	end	2006	
	lineage		
provider	ESPON Database 2013 Project		

Extension of the ESPON databases schemas

The ESPON databases:

- « **ESPON MEGABASE** » : a complex structure for internal use, for long term storage of data, but also for data harmonization, metadata enriching
- « **ESPON DATABASE** » : It is a simpler structure for projects and external users, meant as a repository for harmonized data exported from the megabase; it is easier to use

What's new ?

- Extension of the ESPON DATABASE in order to hold enriched metadata exported from the ESPON MEGABASE
- Extension of the ESPON MEGABASE in order to include :
 - Ontological information (general metadata e.g. measure unit for an indicator)
 - Detailed metadata information (specific metadata e.g. the contact for the person filling the metadata)

Why all this for? Improve the export tools!

Metadata creation requires considerable effort to be created and maintained... Yes

BUT... Data are much more usable (they can be searched in data catalogues and can be understood easier)

Present situation: simple queries

- The ESPON Database Web application now integrates metadata support
 - Can be queried by selecting some relevant parameters
 - The result is exported as a standard Excel file

Future (expected) situation: complex queries

- Giving users the possibility to browse, search and query all metadata in order to find the most useful data for their purpose
- Formulating (with the help of a wizard) a query like:
 - « *Show me all available data on demography of the ESPON space between 1991-1995 in NUTS 2010 nomenclature. I'm interested only in data for NUTS3 units, with a least 90% territorial coverage, sorted by completeness. Prioritize official information* »

The ESPON Database Web Application



Welcome to the ESPON 2013 Database

Please select at least one value in each list of parameters before exporting data as an Excel file

Study area

- ▶ Study areas
- ▼ Countries
 - Austria
 - Belgium
 - Bulgaria
 - Cyprus
 - Czech Republic
 - Denmark

NUTS revision

- NUTS_2003
- NUTS_2006

NUTS levels

- 0
- 1
- 2
- 3

Publication date

- 2009-01-01

Reset selections

Export to excel

Indicators

- Activity rate
- Gross domestic product in euro
- GDP in euro per inhabitant
- GDP in Purchasing Power Parities
- GDP in Purchasing Power Parities per inhabitant
- Total population
- Productivity in euro
- Productivity in Purchasing Power Parities
- Unemployment
- Unemployment rate

Production dates

- 2000-01-01
- 2001-01-01
- 2002-01-01
- 2003-01-01
- 2004-01-01
- 2005-01-01
- 2006-01-01

Data sheet

Microsoft Excel - EsponDBQuery[1].xls

File Edit View Insert Format Tools Data Window Help

Type a question for help

100% Arial 10 B I U

H22 Eurostat, 2009

	A	B	C	D	E	F	G	H	I	J	K	L
1	Id	Official name	activ	Origin	activ	Origin	activ	Origin	activ	Origin	activ	Orig
2	NUTS_2006		2000		2001		2002		2003		2004	
3	at111	Mittelburgenland										
4	at112	Nordburgenland										
5	at113	Südburgenland	44	Eurostat, 2009	43.5	Eurostat, 2009	43.8	Eurostat, 2009	45.3	Eurostat, 2009	44.3	Euro
6	at121	Mostviertel-Eisenw	108.8	Eurostat, 2009	109.1	Eurostat, 2009	112.2	Eurostat, 2009	113.2	Eurostat, 2009	115.2	Euro
7	at122	Niederösterreich-S	114.7	Eurostat, 2009	115	Eurostat, 2009	118.2	Eurostat, 2009	119.2	Eurostat, 2009	119.5	Euro
8	at123	Sankt Pölten	67.5	Eurostat, 2009	67.7	Eurostat, 2009	69.6	Eurostat, 2009	70.2	Eurostat, 2009	70.4	Euro
9	at124	Waldviertel	102.1	Eurostat, 2009	102.4	Eurostat, 2009	105.2	Eurostat, 2009	106.2	Eurostat, 2009	105.5	Euro
10	at125	Weinviertel	55.4	Eurostat, 2009	55.5	Eurostat, 2009	57.1	Eurostat, 2009	57.6	Eurostat, 2009	57.6	Euro
11	at126	Wiener Umland/No	132.4	Eurostat, 2009	132.7	Eurostat, 2009	136.4	Eurostat, 2009	137.6	Eurostat, 2009	139.2	Euro
12	at127	Wiener Umland/Sü	141.9	Eurostat, 2009	142.3	Eurostat, 2009	146.2	Eurostat, 2009	147.5	Eurostat, 2009	147.8	Euro
13	at130	Wien	801.6	Eurostat, 2009	807.7	Eurostat, 2009	766	Eurostat, 2009	794.5	Eurostat, 2009	781.8	Euro
14	at211	Klagenfurt-Villach	123.5	Eurostat, 2009	123.9	Eurostat, 2009	126.6	Eurostat, 2009	126.4	Eurostat, 2009	128.2	Euro
15	at212	Oberkärnten	56.7	Eurostat, 2009	56.9	Eurostat, 2009	58.1	Eurostat, 2009	58	Eurostat, 2009	58.2	Euro
16	at213	Unterkärnten	68.7	Eurostat, 2009	69	Eurostat, 2009	70.5	Eurostat, 2009	70.4	Eurostat, 2009	70.4	Euro
17	at221	Graz	173.4	Eurostat, 2009	175.6	Eurostat, 2009	174.1	Eurostat, 2009	178.9	Eurostat, 2009	178.5	Euro
18	at222	Liezen	37.3	Eurostat, 2009	37.8	Eurostat, 2009	37.4	Eurostat, 2009	38.5	Eurostat, 2009	37.5	Euro
19	at223	Östliche Obersteie	76.5	Eurostat, 2009	77.5	Eurostat, 2009	76.8	Eurostat, 2009	79	Eurostat, 2009	76.6	Euro
20	at224	Oststeiermark	127.4	Eurostat, 2009	129.1	Eurostat, 2009	127.9	Eurostat, 2009	131.5	Eurostat, 2009	129.3	Euro
21	at225	West- und Südstei	89.1	Eurostat, 2009	90.2	Eurostat, 2009	89.4	Eurostat, 2009	91.9	Eurostat, 2009	90.4	Euro
22	at226	Westliche Oberste	47.9	Eurostat, 2009	48.5	Eurostat, 2009	48.1	Eurostat, 2009	49.4	Eurostat, 2009	48.2	Euro
23	at311	Innviertel	129.5	Eurostat, 2009	130.1	Eurostat, 2009	130.8	Eurostat, 2009	135.5	Eurostat, 2009	132.4	Euro
24	at312	Linz-Wels	258.6	Eurostat, 2009	259.9	Eurostat, 2009	261.2	Eurostat, 2009	270.5	Eurostat, 2009	265.2	Euro
25	at313	Mühlviertel	95.4	Eurostat, 2009	95.8	Eurostat, 2009	96.3	Eurostat, 2009	99.8	Eurostat, 2009	98.1	Euro
26	at314	Steyr-Kirchdorf	71.1	Eurostat, 2009	71.5	Eurostat, 2009	71.8	Eurostat, 2009	74.4	Eurostat, 2009	72.6	Euro
27	at315	Traunviertel	105	Eurostat, 2009	105.5	Eurostat, 2009	106	Eurostat, 2009	109.8	Eurostat, 2009	107.4	Euro
28	at321	Lungau	9.7	Eurostat, 2009	9.7	Eurostat, 2009	9.8	Eurostat, 2009	9.9	Eurostat, 2009	9.8	Euro
29	at322	Pinzgau-Pongau	78.5	Eurostat, 2009	78.7	Eurostat, 2009	79.1	Eurostat, 2009	80.3	Eurostat, 2009	79.7	Euro
30	at323	Salzburg und Umg	168.2	Eurostat, 2009	168.7	Eurostat, 2009	169.6	Eurostat, 2009	172.1	Eurostat, 2009	171.3	Euro
31	at331	Außerfern	14.7	Eurostat, 2009	14.8	Eurostat, 2009	15.2	Eurostat, 2009	15.7	Eurostat, 2009	15.9	Euro
32	at332	Innsbruck	126.9	Eurostat, 2009	128.2	Eurostat, 2009	131	Eurostat, 2009	135.5	Eurostat, 2009	136.1	Euro

Metadata / NUTS0 / NUTS1 / NUTS2 / NUTS3

Ready NUM

Metadata sheet

Microsoft Excel - EsponDBQuery[1].xls

Type a question for help

File Edit View Insert Format Tools Data Window Help

100% Arial 10 B I U

Reply with Changes... End Review...

	A	B	C	D	E	F
1	Exported from the ESPON Database on Mon Jun 01 02:05:14 CEST 2009					
2						
3	Dataset metadata					
4						
5		email	manager@espondb.eu			
6		organization	"ESPON 2013 Database project"			
7		last update date	01/01/2009			
8	data filename	ESPONDBQuery.xls				
9						
10	Constraint information					
11	copyright	ESPON 2013 Program				
12	use rights	free				
13	access rights	free				
14	metadata read right	yes				
15	Maintenance information					
16	productionRegular	Yes				
17	productionFrequency	yearly				
18	Spatial representation Information					
19	nomenclature	NUTS				
20	version	2006				
21						
22	Indicator metadata					
23						
24	code	pop_t				
25	name	Total population				
26	units	Thousands inhabitants				
27	abstract	Annual average population (both sexes)				
28	Methodology					
29	language	english				
30	Classification					
31		theme*	Demography			
32		keywords*	Population			

Metadata / NUTS0 / NUTS1 / NUTS2 / NUTS3 /

Ready NUM

Challenge 10 : Spatial analysis for quality control

Objectives : Develop spatial analysis and datamining methods in order to identify exceptional values.

Situation : Different methods exist to find outliers

Strategy : 3 dimensions to explore:

- 1) Attribute-space
- 2) Geographical-space
- 3) Temporal-space

Coord : NCG & LIG



An index for the measure of discontinuities in time series

Extraction of suspect data in the dataset of population 2000-2006

NUTS2_NAME	2000	2001	2002	2003	2004	2005	2006
bg31 Severozapaden	1074.2	1032.1	1016	999.5	982.9	966.3	950.8
bg32 Severen tsentralen	1043.9	991.4	980.7	971.8	962.9	954.1	945.3
bg34 Yugoiztochen	1206.5	1166	1158.8	1150.8	1143.3	1137.3	1132.3
bg41 Yugozapaden	2142.9	2097	2103.1	2107.1	2112.4	2116.8	2117.8
bg42 Yuzhen tsentralen	1680.7	1605.1	1595	1585.3	1575.8	1566.1	1557.6
ch01 Région lémanique	1295.3	1310.3	1325.5	1339.6	1339.6	1369.5	1383.1
ch04 Zürich	1201.2	1218.3	1235.6	1245.6	1245.6	1267.2	1278.3
ch06 Zentralschweiz	676.7	684.4	692.2	696.7	696.7	706.1	711.2
ch07 Ticino	310.2	311.7	313.2	316.3	316.3	321.1	323.6
cz01 Praha	1183.9	1164.7	1158.8	1163.8	1168.1	1176.1	1184.9
dee0 Sachsen-Anhalt	2633	2598.4	2564.8	2535.9	2508.7	2482.1	2427.1
fr83 Corse	264	266.9	269.9	272.9	275.7	278	296.3
fr91 Guadeloupe (FR)	427.1	430.8	434.5	438.2	442	445.5	438
fr93 Guyane (FR)	166.5	173.9	181.7	189.9	196	200	209.7
is00 Iceland	281	285	288	290	292	296.7	303.8
li00 Liechtenstein	33	33	34	34	34	34.8	35
nl23 Flevoland	323.1	335.3	346.7	355.8	362.9	368.3	372.5
nl31 Utrecht	1112.9	1123.7	1146.1	1157.2	1166.8	1175.7	1185.3
ro11 Nord-Vest	2847.2	2839.1	2756.5	2746.8	2743	2735.9	2729.2
ro12 Centru	2643.3	2640.2	2549.8	2545.9	2538.5	2533.9	2529.3
ro21 Nord-Est	3825.7	3835.5	3745.1	3744.6	3739.2	3735.2	3731.4
ro22 Sud-Est	2935.6	2935.2	2867.7	2859.2	2852.5	2846.8	2839
ro31 Sud - Muntenia	3469.3	3462.6	3376.1	3359.4	3344.2	3329.8	3313.1
ro32 Bucuresti - Ilfov	2279.3	2268.9	2211	2208.2	2209	2212.7	2223.9
ro41 Sud-Vest Oltenia	2401.5	2397.2	2342.2	2330.5	2319.5	2307.9	2293.8
ro42 Vest	2041.2	2032.2	1954.8	1947.3	1939.1	1932.1	1927.9
sk01 Bratislavský kraj	617.2	599.1	599	599.8	600.5	602.4	605.2
uki1 Inner London	2722.4	2853.9	2886.9	2908	2921.7	2942.1	2972.6
uki2 Outer London	4382	4459.3	4476.3	4486	4493.9	4511.3	4536.8

Challenge 11 : Enlargement to neighbourhood

Objectives : To collect data at regional and local level for neighbouring countries of ESPON territories

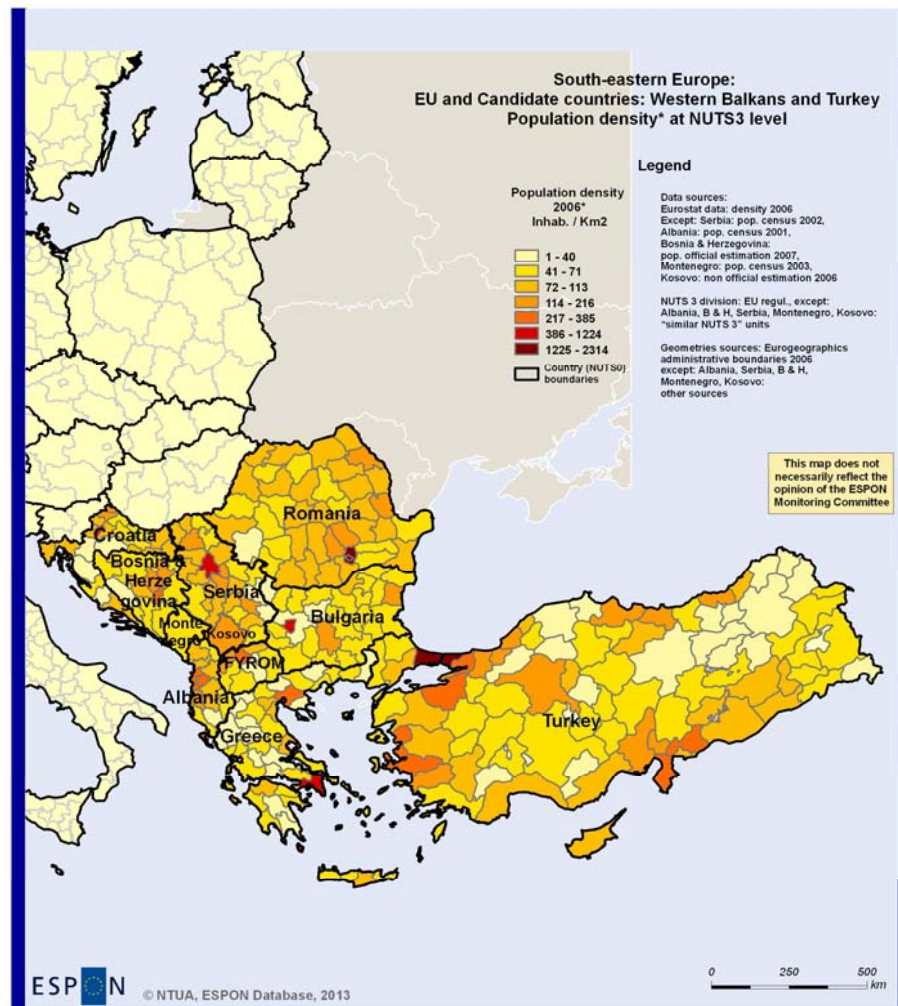
Strategy :

- 1) Evaluation of the situation of data availability in EU candidate countries
- 2) Establishment of contacts with CC statistical offices

Coord : RIATE & NTUA



Regional population density in South-Eastern Europe, 2006



Challenge 12 : Individual data and surveys

Objectives : Examine how to integrate individual data based on census or surveys in the ESPON DB

Situation : In Nordic countries, register data of the population has been made available for research purpose

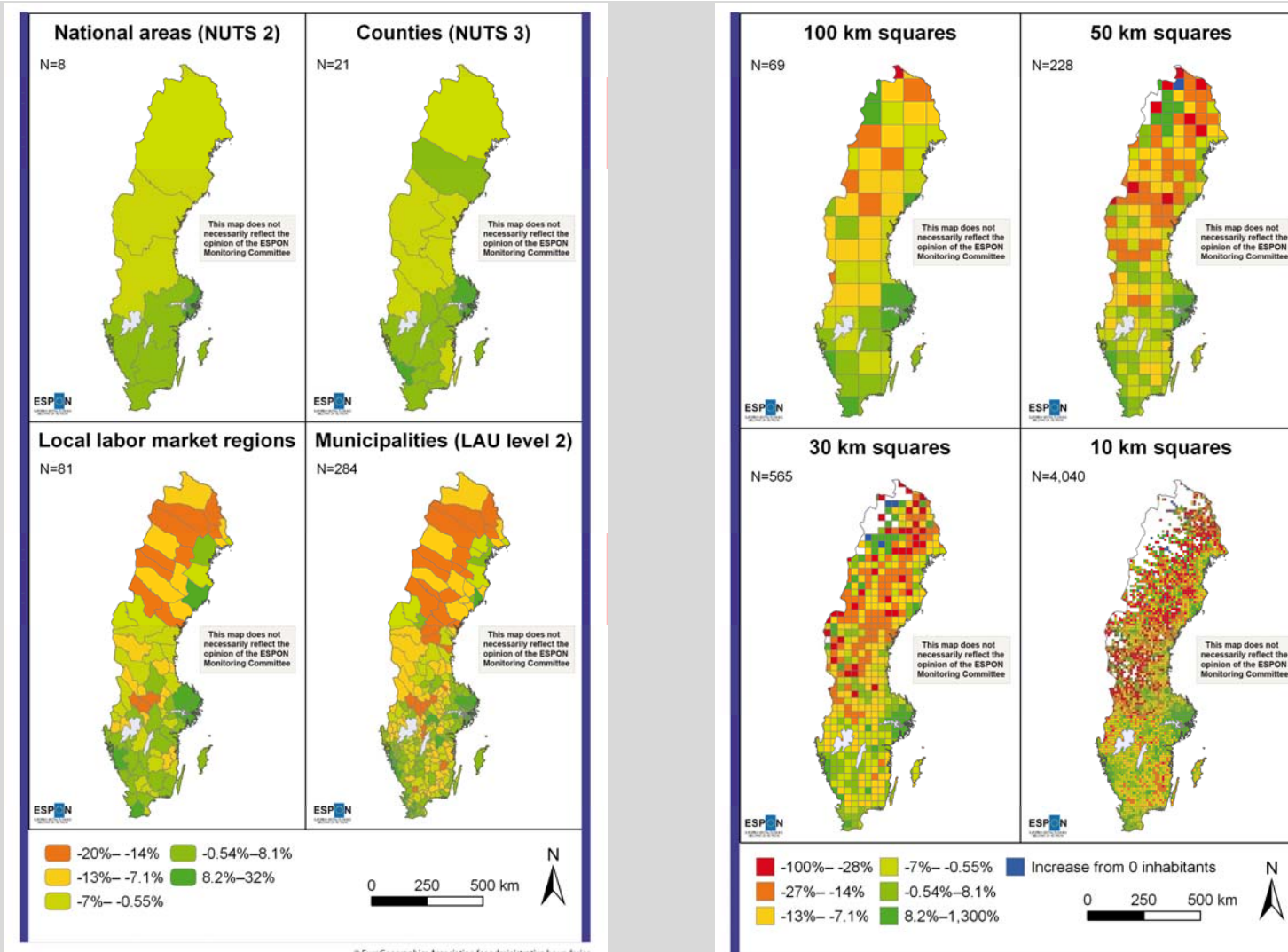
Strategy :

- 1) Availability and usefulness of survey data
- 2) MAUP
- 3) Comparison Sweden-Europe wide data
- 4) Integration of survey data in the ESPON DB

Coord : RIATE & UMEA



Population change, 1990-2005





THANK YOU
FOR
YOUR
ATTENTION !