# Territorial information, shapefiles and indicators accessibles for actors until commune and infra communal levels (Synthesis).

Christophe Breuer, Guénael Devillet

# *Territorial information, shapefiles and indicators accessibles for actors until commune and infra communal levels (Synthesis)*

**Breuer Christophe**

Licencié en Sciences géographiques
Maître en Urbanisme et Aménagement du Territoire
christophe.breuer@ulg.ac.be, +32 4 366 57 47

**Devillet Guénaël**

Licencié en Sciences géographiques
D.E.S. en Urbanisme et Aménagement du Territoire
G.Devillet@ulg.ac.be, +32 4 366 53 19

**Adresse professionnelle**

Service d'Étude en Géographie Économique Fondamentale et Appliquée (SEGEFA)
Université de Liège
2, Allée du 6 Août, B-4000, Liège, www.segefa.be

**Summary**: This article presents the carried out researches within the framework of the creation of an indicators gate for the local actors. We evaluate in this one the accessible territorial data for the actors, and the cartographic files of the administrative or statistical entities to represent them. We analyze the sources of information, and determine its accessibility in the various countries of the European Union. We moreover deal with problems of the relevant territorial scales and cartographic representation of the data.

**Résumé** : Cet article rend compte des recherches effectuées dans le cadre de la création d'un portail d'indicateurs pour les acteurs locaux. Nous évaluons dans celui-ci les données territoriales accessibles pour les acteurs, et les fichiers cartographiques des entités administratives ou statistiques pour les représenter. Nous analysons les sources de l'information, et déterminons son accessibilité dans les différents pays de l'Union européenne. Nous traitons en outre de la problématique des échelles territoriales pertinentes et de la représentation cartographique des données.

**Keywords**: Indicators, territorial cutting, accessibility of the data, cartographic representation, gate of territorial data.

**Mots clés** : Indicateurs, découpage territorial, accessibilité des données, représentation cartographique, portail de données territoriales.

# Territorial information, shapefiles and indicators accessible for actors until commune and infra communal levels

## 1.    INTRODUCTION

Our team takes charge of the realization of a portal of on-line territorial information. This work brings together three university teams which respectively work on the realization of the data representation interface, on the data-gathering (spatial and contextual data), and on the metadata as well as the management style of the indicators.

The current work aims at working out a portal of information for the attention of the local actors. The concretization of this project requires many theoretical and technical reflections as for the development of the tool, usable on Internet.

The essential phase of work was focused on the technical and operational approach of the online tool. The research undertaken until now aims at answering the following questions:

- how to obtain the contextual data up to the communal and infra-communal level;
- which are the definitions of the statistics and indicators for each country;
- how to obtain the shapefiles to allow a cartographic representation of the data;
- how to create a portal of indicators for local actors.

This paper will firstly deal with the inventory of the socio-economic and environmental data in Europe, and of their accessibility. Secondly, we will tackle the question of the selected indicators according to the guide CATALYSE, method which confronts the populations needs with the services offered by taking account of the environmental context. Then, we will discuss the creation of the contextual information portal, for finally concluding our paper.

## 2.    INVENTORY OF SOCIO-ECONOMIC AND ENVIRONMENTAL INFORMATION AVAILABLE ON INTERNET

### 2.1    Sources and gathering

The information sources available on Internet are very diversified, just like their working. Only a part of the European spatial levels allows an on line consultation of their statistical data, and even rarer are those which allow a direct and free access.

The data-gathering is naturally a crucial step of the creation procedure of the territorial information portal. Research has been leaded, on the one hand, at the level of EUROSTAT, and on the other hand, at the level of the national suppliers (for the LAU 1-2, and the infra-communal data). The first work has been completed and it putted forward the sources of the data and the bonds to reach them. This collection of information had however not taken the type of information, the structure of their working and of the type of file into account.

EUROSTAT makes it possible to have statistical data for all the European Union from the Community or national level, until the equivalent of the NUTS 3 for each country. The data are not necessarily available for the past year, it is generally necessary to take 2 years former periods so that the statistical data are available for all the Union countries. The files obtained can take various forms and are displayable in a spreadsheet or a database manager, which makes it possible to have an easy relationship with the spatial entities at the time of the representation of those.

This side of NUTS 3 level (supra communal), information have to be specifically collected near each national supplier. All the national suppliers of statistical data belong to a ministry or an independent organization of the state taking charge of the data collection and processing. They have each an Internet site at one's disposal on which it is possible to obtain basic information, often at the national scale (economic and social indicators, export figures, etc). The local data are more rarely available, often in direct display or downloading.

Within the framework of the portal of the territorial intelligence network, it has to be noted that the useful data are restricted to the data available for downloading in the spreadsheet or database format. The files in a compressed, only reading, or text format are unusable within the framework of an automated treatment of information (for display).

An inventory of the data availability was carried out during the year 2007, country by country, by taking down in a database the Internet localization of each file, without prejudging its contents. The posterior analysis of this document shows a strong variability of the Web addressing, a regular incompatibility between the database managers and the files suggested by the suppliers. Finally, it has to be stressed that each national institute sales the files in spreadsheet form with the statistical data by spatial entity.

For the data of cuttings LAU 1 and 2, as a precaution, one should treat even the nature of the indicators cautiously so much the definitions can differ, on the one hand between the EUROSTAT data and the data of the national providers, and on the other hand between the national suppliers themselves. It will be necessary to take into account the differences in definition and thus of the statistics calculation, to harmonize (*cfr. infra*) or to inform the user. In the same way, metadata will be essential to determine the various useful data information (date, origin, margin of error in some cases, etc).

The problems are exactly similar for the infra-communal data, which moreover are generally available only for exact dates depending directly on the periods of investigation or census.

The available data near the national providers are, on the one hand, national and regional data (often directly accessible, but maybe making redundant use with EUROSTAT), and on the other hand local or infra-communal data (often one has to pay for them, or only available to display).

The access to the data is heterogeneous, and the data files are seldom downloadable. The definition of the statistics and indicators is variable according to the country, which is seldom explained on the sites of the suppliers. The map of the collected data can moreover easily be related with the chart of the answers by country.

The data are often directly displayed on the screen, or distributed in the form of downloadable pdf files. These documents, although very useful at the local level and for specific projects, cannot allow a systematic automation of the data processing, and are thus useless for us as it is. A solution could be to recode the information, but that would be time consuming and expensive. For example, if we have access to the statistical data for France at the communal level, for 15 variables, that represents more than 551.745 entries to seize, which is unimaginable within the framework of local actors project, of which it is not the objective. Information has a cost, which we will try to specify on the European scale.

## 2.2 Examples of observatories in Europe

### National suppliers of statistical data

Each country comprises its own general supplier of statistical data. These structures can be regarded as observatories because they have in certain cases the same structure, and sometimes the same tasks. As mentioned above, each one of these institutes has its own Internet site generally comprising statistics.

|    | Country | Institution | Website |
|----|---------|-------------|---------|
| 1 | Austria | Statistics Austria | http://www.statistik.at/ |
| 2 | Belgium | FPS Economy - Directorate-general Statistics Belgium | http://www.statbel.fgov.be/ |
| 3 | Bulgaria | National Statistical Institute of Bulgaria | http://www.nsi.bg/ |
| 4 | Cyprus | Statistical Service off Cyprus | http://www.mof.gov.cy/mof/cystat/ |
| 5 | Czech Republic | Czech Statistical Office | http://www.czso.cz/eng/redakce.nsf/i/home |
| 6 | Denmark | Statistics Denmark | http://www.dst.dk/ |
| 7 | Estonia | Statistics Estonia | http://www.stat.ee/ |
| 8 | Finland | Statistics Finland | http://www.stat.fi/ |
| 9 | France | INSEE | http://www.insee.fr |
| 10 | Germany | Statistisches Bundesamt Deutschland | http://www.destatis.de/jetspeed/portal/cms/ |

| 11 | Greece | General Secretariat off National Statistical Service off Greece | http://www.statistics.gr/ |
|----|--------|------------------------------------------------------------------|--------------------------|
| 12 | Hungary | Central Hungarian Statistical Office | http://portal.ksh.hu/portal/ |
| 13 | Ireland | Central Statistics Office Ierland | http://www.cso.ie/ |
| 14 | Italy | Istituto nazionale di statistica | http://www.istat.it/ |
| 15 | Latvia | Exchange Statistical Office off Latvia | http://www.csb.lv/ |
| 16 | Lithuania | Statistics Lithuania | http://www.stat.gov.lt/ |
| 17 | Luxembourg | STATEC | http://www.statec.public.lu/ |
| 18 | Malted | National Statistics Office | http://www.nso.gov.mt/ |
| 19 | Netherlands | Statistics Netherlands | http://www.cbs.nl/ |
| 20 | Poland | Exchange Statistical Office | http://www.stat.gov.pl/ |
| 21 | Portugal | Instituto Nacional de Estatistica | http://www.ine.pt/ |
| 22 | Romania | National Institutul of Statistica | http://www.insse.ro/ |
| 23 | Slovakia | Statistical Office off the SR | http://portal.statistics.sk/ |
| 24 | Slovenia | Statistical Office off the Republic off Slovenia | http://www.stat.si/ |
| 25 | Spain | Insitituto Nacional de Estadistica | http://www.ine.es/ |
| 26 | Sweden | Statistics Sweden | http://www.scb.se/ |
| 27 | UK | Office for National Statistics | http://www.statistics.gov.uk/ |

### EIS Euregional Information Service

The Program EIS depends on European financings. It covers Euregio Meuse-Rhine areas (Netherlands, Belgium, Germany) for which it collects and harmonizes the statistical data. The harmonization must allow interregional comparisons and a putting in prospect for the actors, often involved in international cooperation.

The harmonization of the statistical data and of the indicators does not go without encountering difficulties: the task is long and can go up far in the process of data acquisitions. It is sometimes necessary to take again the base of the statistics itself, which is strongly time consuming and sometimes not very relevant. It is thus appropriate, depending on the time of realization available, to wonder about the possibility of eluding the question of the harmonization.

### Atlantic Institute of Town and Country Planning

Created in 1994, the main aims of the IAAT are the realization of a mutualisation platform of the statistical data and of the territorial tools, in addition to allowing the access to documents and studies useful for regional planning.

The documents are diversified: working papers, maps, statistic tables, instruments, etc. The topics tackled go from the agriculture to the economy and from the demography to the environment.

## 3.    SELECTION OF INDICATORS FOR THE GUIDE CATALYSE

CATALYSE is a method who confront the population needs (collected with questionnaire) with services offered to population (based on a list). This method takes account of the environment with various environmental indicators.

The selected indicators for the guide CATALYSE were highlighted during previous interdisciplinary work. Those sought the available statistical elements which answer the best to questions of the guide (personal questionnaire). All these indicators are not available for each country, moreover, it will certainly be necessary to consider the aggregation of classes in certain cases to form indicators usable by the actors starting from the statistical data (for example for the age groups). Please find bellow the table of correspondence.

| Topic of the guide | Indicator |
|--------------------|-----------|
| Territorial ID and information | Evolution of population |
| Information initial butt reception | Population by gender |

| | Population by old groups |
|---|---|
| | Nationality |
| | Composition of the household |
| | Education Level |
| | Activity status |
| | Tenure status |
| | Health status |
| Personal dated and family links | Marital status |
| | Income of the household |
| Housing | Area of the dwelling |
| Social Individual and autonomy | Means of transport of the household |

## 4. PORTAL OF INDICATORS

### 4.1 Relevance of cartography for the representation of indicators

The cartographic representation of the statistical data is not always the more relevant and the more useful for the actors. When the number of entities represented is too weak, or when the difference between the maximum values is reduced, the reading of the documents will not be easy. For the user, these cases would be better perceived in tables.

In addition, some indicators do not allow a representation in the form of uniform ranges (population pyramid…). In this case, one has to choose between a system of complex representation (symbols) or a set of maps on the same topic: in these two cases, the reading of the information is made more complex for the actors.

However in the majority of the cases, the cartography is very useful to represent indicators. Indeed, it makes it possible to synthesize information and to improve the communication. The cartography is particularly interesting when the choice of the indicators makes it possible to give an account of a relevant situation.

### 4.2 Identification of relevant cuttings

Relevant cuttings are relating to the level of cutting which is the more appropriate for the applications which will be given to the portal of the territorial intelligence. It has to be noted within this framework that the majority of the actors are local and that their working scale can be lower than the commune. The basic level that absolutely has to be reached is thus the commune level, moreover it is the lowest level for which the statistical data are regularly available and can be ordered for all the countries.

The communal level is in addition a level which can be find everywhere in Europe, it is thus a solid working base. We thus note that the relevant level of representation depends firstly on the uses which are made by it, and in addition on the availability of the information. By considering that the data of the higher spatial entities can be found by the aggregation of the lower spatial levels (what is not always checked, in particular because of the calculating methods), we will thus always seek the data at the lowest level as possible

### 4.3 Gathering shapefiles

The computer data relating to the representation of administrative and statistical cuttings are difficult to obtain. On the one hand, the operators of the statistics are seldom the same ones as the suppliers of geographical data, on the other hand these last regularly work with subcontractors. One will also note that for certain spatial levels (especially infra-communal), in certain countries it is the statistical institutes which remain the holders of the numerical maps and sub-contract the execution of it.

Within the framework of the European representation of the spatial data, a major problem comes from the diversity of the reference system and the resolutions of the computer documents. These essential differences interfere at the time of the realization of the visual representation, which can affect the quality of it (distortions, discontinuities of the borders, etc). Finally, some countries do not distribute their *shapefiles,* as for example Luxembourg which has a sufficiently exiguous territory so that the cartographic representation does not use software specific for the mapping (and thus do not require these files).

The separation of the sources, the national traditions and the subcontracting of the realization lead to various problems. Firstly, there exist important differences between the European conventionalism (and the territorial structures which result from this) and the local (national) application of the cuttings. Typology appreciably differs depending on the source, which makes complex the structure and the comprehension of this one. Secondly, the systems of subcontracting by the Member States towards third parties generate a cost, this one can only moderately be reduced by the public characters of the requests (for example the caENTI project).

In order to collect the entirety of administrative cuttings for the whole of the European Union (27), a personalized contact was carried out with each national statistical institute. The requests related at the same time on the numerical maps (shapefiles) of the statistical cuttings, and on the obtaining of the statistics themselves up to the communal or infra-communal level. The procedure envisaged a first e-mail, followed by a first, and then by a second e-mail of recall.

The answers are mainly contrasted: few national institutes could answer our request, some not having the files (subcontracting of the cartography), or others having to pay for them (at a prohibitory cost).

We can note that the countries which have been able to inform us on the obtaining conditions of the shapefiles are very few, and are in Western Europe. The access terms considerably vary between the countries, and the result of the collection is particularly heterogeneous. A big number of countries reorient the requests towards other operators (who do not always answer the information requests).

The data are uniformly available up to level NUTS 3; indeed, EUROSTAT provides shapefiles for levels NUTS 0,1, 2 and 3 on its Internet portal. However the precision of these shapefiles is reduced since the basic scale is the 1: 1.000.000 what is not in adequacy with the local use of the portal of territorial intelligence.

This side of NUTS 3, little country offers the data free. The Netherlands are the exception thanks to the data acquisition which is easy and didactic (possibility of Google Earth compatibility, etc). It should be noted that the all the countries do not have a statistical Internet site with an English translation, which makes the collection of information appreciably more complex.

The collection of information via the national providers remains problematic: little country, on the level of the EU27 envisaged procedures to acquire the shapefiles easily and free. A majority of country does not even sale these data, or at prohibitive prices. In addition, other technical problems reinforce these difficulties of acquisition.

Indeed, the format of the files appreciably differs from one country to another, just like the identifiers which make it possible to establish the link between the spatial entities and the assignee base. The modifications of reference scales at the borders present main issues of representation in these zones.

An important work of standardization of the "graphic data" and assignees tables are thus necessary downstream from the work of data acquisition. Moreover to require an important time of collection, mainly due to the procedures, the subcontracting and the administrative delays of the organizations which have the rights on the files, the passage by each institution separately requires the data processing, which is both time consuming and expensive.

The ideal solution would be to find a single supplier, which would have as a principal advantage to facilitate the relation between the assignees bases (especially European) and the spatial entities by standardizing the identifiers. The others advantages of the single supplier are mainly esthetic and practical: better representation, coherence of the minimal and maximal scales, single intermediary, reduced waiting periods and administration expenses.

The Eurogeographics company offers various products which cover the entirety of the European Union. Its EuroBoundaryMap file proposes various statistical and administrative cuttings at least up to LAU 2. This side of, the situation is variable according to the countries, but a relation is ensured between the statistical sectors and the higher spatial levels. For this purpose, Eurogeographics has rights over the products of the national suppliers. In parallel, this company receives from EUROSTAT a yearly rental fee for the use of the Eurogeographics files by the European projects for free. This formula seems perfectly appropriate for the project of portal of territorial intelligence.

The scale of reference of this product is the 1: 100.000, which could be supplemented by national files to have a scale more precise.

The whole of the infra-communal data can only be obtained by the national suppliers, or by the Regions (for example in Germany). Only two countries provide the infra-communal data in an entirely free way (at the 28/06/2008): Netherlands and Denmark.

**4.4    Additional work**

The data-gathering is part of the creation of the gate of territorial information for the actors. Other works are related to those described above. They mainly deal with the data and metadata management, and the constitution of the tool for cartography based on technology webgis.

Firstly, the gate must manage a great number of data, which is the result of the multiplication of the number of indicators by the number of entities represented. In order to inform the users on visible information, it is necessary to have data describing the data (date of acquisition, supplier, precision…).

Secondly, the visualization of information by the end user depends on the use of a technology of on-line cartography. This technology webgis makes it possible to the local actors to create interactive charts meeting their needs.

These two tasks are carried out by the universities partners of Franche-Comté and Slovenia (ZRC-ZAZU).

**5.    CONCLUSION**

The gate of territorial information for the actors is based on both the technical system of management and representation, and on the data which it contains. The search for these data on a European scale is made more complex by the various data distributers, and the acquisition cost of those.

We highlighted the information sources in Europe, and the gaps of the centralized distributers. Moreover, we treated cartographic representation of the data, and thus of the shapefiles of administrative and statistical cuttings. The effective collection of information showed the difficulty of information concentration, and proposes various tracks to answer it.