

*TOWARDS A SYSTEM OF TERRITORIAL INTELLIGENCE: GEOGRAPHICAL INDICATORS AND
DATA FOR THE EUROPEAN ACTORS*

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 deals with the realization of a territorial information gate for the European actors. It defines the sources of territorial information and the files allowing the cartography of the administrative and statistical cuttings. The acquisition cost of information is also putted forward. It establishes the link between the technical realization of an interactive interface of webmapping and the problem of the data and metadata management.

Résumé : Cet article traite de la réalisation d'un portail d'information territoriale pour les acteurs européens. Il définit les sources de l'information territoriale et des fichiers permettant la cartographie des découpages administratifs et statistiques. Le coût d'acquisition des informations est également mis en avant. Il fait le lien vers la réalisation technique d'une interface interactive de webmapping et la problématique de la gestion des données et des métadonnées.

Keywords: Portal of information, territorial information, indicators, data, metadata.

Mots clés : Portal d'information, information territoriale, indicateurs, données, métadonnées.

Towards a system of territorial intelligence: geographical indicators and data for the European actors

I. Preamble

Within the framework of the CAENTI European project (coordination action of European Network of Territorial Intelligence) of the FP6 (Sixth European framework program), research on the territorial indicators, their treatment and their mode of representation, was undertaken. The objective is to represent, for actors implied in several sectors (in particular social), data which they use for the method CATALYSE.

After multiple research on the indicators, their accessibility and their definition in the European Union of twenty-seven (CHEN *and Al*, 2006; CAENTI 2008), it was necessary to treat the problem of the installation of the territorial portal of information for the actors, available on Internet. Concretely, it is a question of allowing the visualization of data characterizing the socio-economic environment of the populations to increase the comprehension of the territories, their dynamic and their problems, and of allowing a prospective view of the data collected by the actors near the populations at a more appropriate scale.

The cartographic representation was privileged to allow a better legibility of information, the comprehension of the space disparities and the possibility of comparison with other areas. All in all, although the reading of maps is not acquired for all, it is a powerful communications tools.

It was decided that this tool would be available for free on Internet, which generates complex technical and legal constraints, in terms of treatments, of representation and spread of information. Nevertheless, thanks to this free use, this tool will be used by the local actors, often confronted with budgetary and human important restrictions.

Our work was thus focused on the following tasks, in order to emit a list of specification for the creation of a European gate of territorial information, and to improve the current data (offers, distribution):

- obtaining the contextual and geographical data (administrative cuttings and statistics);
- the study and the implementation of the representation modes of cartographic on-line information;
- the study and the implementation of the techniques of information and meta-information storage;

This article will initially treat of territorial information and geographical files (shapefiles¹) collection making it possible together to create cartography of socio-economic and environmental contextual information. This part will take into account the constraints related to the mode of representation selected and will determine the essence of the characterization of this one. The second part will give the broad outlines of the cartographic representations selected for the prototype of the cartographic tool. The third part will treat management of the data and metadata within the system. Then we will conclude.

II. Data-gathering

The data-gathering is a crucial step in the realization of the gate of territorial information: without those, it is obvious that the project does not have substance. In order to carry out cartography of the indicators or data, two types of files are necessary: on the one hand, data to be represented, and on the other hand a file representing spatial cutting in entities (communes, country...). It is within the entities of the latter that the values of the indicators will be represented.

1. Contextual data

The contextual data form an integral part of the analysis by the method CATALYSE (CAENTI, 2006c). In order to limit the data retrievals, indicators were selected on the basis of European guide (CAENTI, 2006a). They are 15, and their availability near the European Office of the Statistics (EUROSTAT) and the national suppliers of statistics was studied (CAENTI, 2008). Nevertheless, the creation of a gate of information is not summarized with the simple visualization, new constraints thus appeared.

a. Constraints

Various constraints quickly emerged: whereas certain sites of statistics provide the data on a tabular format, others allow only one posting and a visualization of information on screen. It is however essential to have a tabular form to allow the introduction of the data into the cartographic software. Certain Internet sites generally force the users to digitalize information, which is a long and expensive operation.

The cost of acquisition of the data remains indeed an important problem since it will generally be paid by the end users, what we want to avoid. We thus supported the data obtaining for free on Internet sites. EUROSTAT allows this, and certain national sites also put a broad offer of free socio-economic data at the users' disposal. Nevertheless, the majority of the consulted statistical sites ask money for the most precise or specific data. They are primarily the data at the communal or infra communal level, like the statistical sectors.

The collection of information is thus subjected to a double constraint, depending on the format and on the acquisition costs of the data. Concretely, that considerably reduces the data availability for their insertion, whereas the preceding statements brought a very promising vision of the accessibility of the data at the European level (CHEN *and al.*, 2006).

Moreover, the problem of the data update remains not exhausted today. Indeed, the comparison of the situation statements on Internet between the years 2006, 2007 and 2008 showed a strong variability of the web-addresses giving access to the information sources. It seems complex to consider an automatic update without preliminary convention with the statistical partners.

b. Results

The difficulty of information collection appeared as of the use of the files resulting from EUROSTAT. As proposed previously (CAENTI, 2008), EUROSTAT mainly collects furnished information with thirds (national institutes...), which generates information gaps within the databases. In this case, it is sometimes necessary to use older years under review, by taking into account the dynamics of constitution of the European Union and thus the disparities of information for the countries integrated more recently. Moreover, certain indicators of the European Office of the Statistics are named in an unsuitable way, or by creating a confusion: indicators which one thinks acquired can appear not very relevant according to their method of calculating.

On the level of the national suppliers, it is the multitude which increases the difficulty of obtaining the data. Acquisition can moreover be expensive for certain very small spatial levels (for example for the communes in France and for the statistical sectors in Belgium). Major indicators (total population, unemployment rate,...) are often free, even for restricted spatial levels. But as soon as the indicators become more complex, they are not free anymore (cost can rapidly increase). It is indeed necessary to take into account the number of countries and the number of entities which gear down the costs of acquisition, without ensuring regular updates for all.

Within four pilot countries, we could collect several simple indicators (total population, population density, unemployment rate, and foreigner rate, income of households...) in a continuous way, from the NUTS 1 until the LAU2/NUTS5. More complex indicators present in a systematic way of the gaps and discontinuities, either verticals (of the space levels low towards highest), or horizontal (between entities of the same administrative level).

2. Geographical data

The geographical data constitute the base of the spatial representation of the indicators. Primarily exchanged and distributed in the form of file vector "shapefiles", we tried to preserve the homogeneity of the format for the whole of the collected data. The national and international distributors of geographical data propose generally various alternative formats, although the shapefiles are dominant.

a. Constraints

The cartography of various countries in the same interface and in a continuous way can be confronted with various problems. They primarily result from the methods and characteristic of the various cartographic

products. Indeed, the creation of the shapefiles in general, and layers of information on administrative cuttings particularly, underlies a certain number of variables which can differ from a creator of data to the other. Among those, we will retain the reference levels of the digitalization, the basic scale of the cartography, the inaccuracy being able to reach several kilometers or employed projections.

Many of these constraints can find an answer by the exploitation of the possibilities of the cartography software, which has function of correction or modification of projections. However, the problems of inaccuracy cannot find an easy solution. The risk of the assembly of shapefiles of different origins for the European countries is thus to create inconsistencies and discontinuities in the cartography.

In addition, the constraints related to the cost of the data also remain for geographical information.

b. National suppliers

Each country has its official supplier of geographical data, even if this one can sub-contract the resale of information as well as their creation. In certain nations, only one institute deals with the digitalization of the administrative limits and their distribution, but this case of figure remains rare.

The data generally have a cost, which varies according to the country, the zone which one wishes to acquire and the administrative level of cuttings. The countries which give the geographical data free are a minority in European space. The majority of them require a price growing according to the precision of the data and in certain cases the format (certain data are free in image format but need to be paid in shapefile format).

In addition, all the countries do not distribute the geographical files. Certain personally contacted national institutes confirmed that they do not place the data of administrative cuttings at the disposal of third, not even to have it in formats usable by the software cartography. These answers are certainly debatable, but it illustrates the difficulty of obtaining such information.

Our work of collection of information was thus made more complex by the multitude of suppliers, the incompatible data formats with the most widespread software of cartography, and the often prohibitory costs of the data moreover subjected to certain legal restrictions in terms of diffusion. In addition, fears concerning the disparities of resolution, reference frame and projection proved founded. The inconsistencies noted at the borders are consistent and make interpretations of the indicators very difficult.

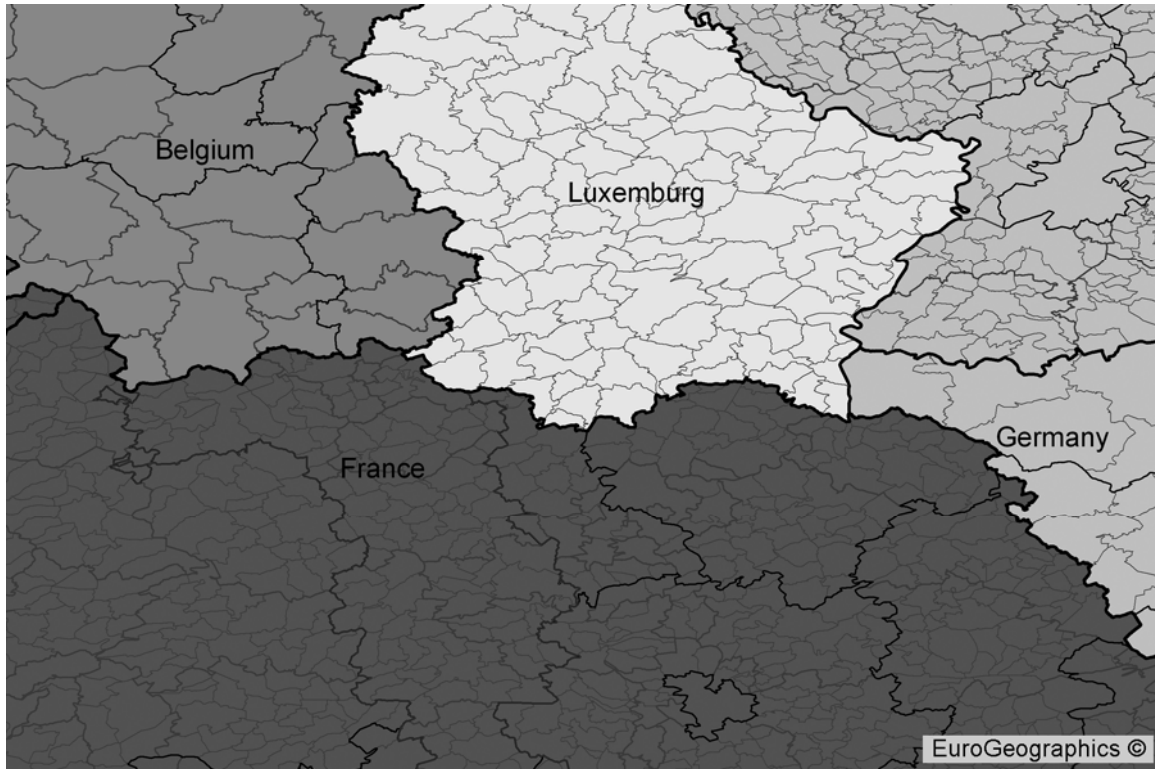
The only official organization providing administrative cuttings for the whole of the European Union at an infra national level is the European Office of the Statistics. Its Internet site places for free at the disposal of all a map of the administrative level divisions NUTS 1, NUTS 2 and NUTS 3. But the precision of these data is poor (scale of reference 1: 1.000.000) and does not resolve the problem of the supra-communal, communal and infra-communal levels, which are the most important levels for the local actors.

c. Single supplier

No traditional public agency provides data from the level of the countries to the communes. This level is the main objective of our collection of information since the “communes²” are a basic unit present in all the European countries. Our research thus led us to seek private or parapublic distributors for these data. The use of products of a single supplier having in addition harmonized the data makes it possible to be freed from a big number of constraints quoted before.

We found only one distributor of data for the whole of the European Union and certain countries partenaires³. This supplier, EuroGeographics, places at the disposal various products at the scale of the Union of which a map of administrative cuttings having for geometrical base the administrative communes or equivalents. Moreover, the base assignee of these files ensures the relation between certain infra-communal entities and the higher entities. The base assignee of this supplier also eludes the question of the diversity of the identifiers (met for national suppliers) since those are harmonized according to European nomenclature NUTS/LAU. Each of the 125.286 entities represented is thus identified by its single code. The data are regularly updated by the company.

In order to constitute its product, EuroGeographics joined the various national distributors. Those thus remain holders of the copyright but EuroGeographics can diffuse the products of the various countries.



The cost of implementation of these data is prohibitory; a license for a single user must be acquired for 15.340 euros for the UE27 and for a 2 years duration. Additional expenditure must be taken into account for diffusion on Internet. Nevertheless, EuroGeographics has agreements with the European Union: the European research projects can have the EuroGeographics data during the duration of the project and for interns use only.

The use of such a product is optimal because it makes it possible to have uniform layers of information whose precision is satisfactory taking into consideration the objective of the tool. Nevertheless its cost in the long term makes its exploitation impossible for a big number of local actors having reduced financial means. It would be interesting to be able to develop in a near future a similar product, of which accessibility would not be subjected to such financial constraints.

III. Data and metadata

Volumes of data treated by the system of representation of territorial information are considerable: it is a question of distributing data of 27 countries for many indicators. Potentially, taking into account the administrative levels, that represents a number of several million. The preceding studies (CAENTI, 2008) showed in addition that the sources of the indicators are numerous, and that the conditions of obtaining are particularly diversified.

In this context, several arguments are in favor of the collection and the aggregation of metadata describing at the same time the data themselves, and the indicators. Firstly, the end users must be able to analyze the data by knowing the methods of calculating of the indicators, the year of the data, their source, etc. In the second place, the metadata are useful within the framework of the updates and the various automatic treatments being able to be carried out on the database.

The research undertaken within the framework of program CAENTI (DAMY *and al.*, 2008) aims to harmonize the metadata and to adapt them to the geographical data and the indicators. The research undertaken in this direction is based on Dublin Core and the ISO standards. Work led to a proposal of grid for the metadata, which was successfully tested on European and national indicators, like on geographical files. The fields will be validated by the comparison with the conclusions of the project INSPIRE⁴ which must end in November 2008.

IV. Cartographic interface

The cartographic interface is the element of the presentation system the most seen by the end users. This one, in its design, must adapt to all data-processing environments in order to have a public potentially more important. Nevertheless, it should be stressed that a good gate of information depends less on its aspect than on what it contains, but its aspect contributes to the visibility and the handiness of the tool.

It is about a tool of *web mapping* making it possible to carry out interactive maps (PEHANI *and al.*, 2008). In addition to mapping the values of the indices, the interface also makes it possible to inform the uses of the metadata and the conditions of obtaining and creation of the visible data on the representation.

The cartographic interface thus necessarily uses the whole of the previously prepared resources. The prototype fulfills the requirements that we had initially fixed ourselves, while being based on *open-source* technologies.

V. Conclusion

The realization of a European gate of indicators as component of the cognitive process of the territories requires a particularly important basic work. After a selection of the indicators represented on the basis of guide adapted to the method CATALYSE, the relative questions to obtaining the data and the metadata emerged. They highlighted the need for collecting both data simultaneously to inform the end user as well as possible.

The collection of information was also particularly instructive to show the availability and the cost of acquisition of geographical information or indicators. The disparities of information spread shown during the research also show the difficulties encountered by the actors to obtain abundant and relevant territorial information. The centralization of the information and its standardization seem to be an essential prerequisite for the processes of the territory knowledge, efficient and accessible for the various actors.

In parallel, the future developments of the cartographic tool will probably make it possible to permit to introduce their own contextual data meeting their needs specifically: in this case, if the data are actually provided by the end user, the shapefiles will have to be provided by the owner of the platform.

The use of an interactive cartographic interface based on technologies of *Web mapping* makes it possible to exploit with relevance the elements implemented before and to widely diffuse near users, whom are sometimes not accustomed with the methods of cartography essential for the comprehension of the territories. The existence of a functional prototype based on real and territorialized data tends to confirm the feasibility of one gate on a European scale, realizing substantial improvements on the level of the accessibility of the data, objective which will probably be achieved thanks to European engagements through the directive INSPIRES. The work undertaken by the caENTI project will make it possible to make recommendations at the European Union to improve the territorial information spreading and to increase the effectiveness of the local actions.

¹ *Shapefiles is a format of file initially implemented by the company ESRI, de facto reference in the software of cartography and geographical information system (GIS). This format is used (and compatible) by a majority of software SIG / cartography. It contains a geographical layer of information being able to be used as support with the cartography of contextual data.*

² *definitions of this level differ however. Nevertheless, the democratic structures are overall similar and correspond to a "local basic level".*

³ *countries partners are, for example, certain applicant countries at the entry in the European Union or of the frontier countries.*

⁴ *European project INSPIRE answers the directive 2007/2/CE of the European Parliament which aims to harmonize and improve geographical information spread in the European Union. A specific shutter is devoted to the geographical metadata.*

Bibliography

CAENTI, 2006a, *List of territorial indicators available on internet for comparison with CATALYSE Guide Data*, Deliverable 53, 28p.

CAENTI, 2006b, *State-of-art about of territorial information available on internet and sources in Europe*, Deliverable 24, 23p.

CAENTI 2006c, *European contents specification for a CATALYSE guide for diagnosis and evaluation*, Deliverable 51, 155p.

CAENTI 2006d, *Guidance notes for the use of CATALYSE information and tools*, Deliverable 56, 192p.

CAENTI, 2008, *Reasoned catalogue of territorial information available on internet and sources in Europe*, Deliverable 29, 83p.

CHARLIER J., DEVILLET G., BRUNETTI E., CUSUMANO C., DELVOYE J.-M., DOYEN P., URBAN A., 2007, *To a TI Community System : shared diagnostic and territorial animation of a Seraing 's area (Belgium)*, in Acts of International Conference of Territorial Intelligence, Huelva 2007 (caENTI)

CHEN D., DEVILLET G., 2006, *Territorial information, themes, indicators and sources*, in Acts of the international Conference of Territorial Intelligence, Vol 2 caENTI, Alba Iulia, October 21st-22nd, 2007

DAMY S., HERRMANN-PHILIPPE B., 2008, *Metadata for caENTI*, in Acts of the International Conference of Territorial Information 2008, UFC : Besançon, à paraître

GIRARDOT J.-J., 2006, *Activities and prospects of research activities concerning tools of territorial intelligence for sustainable developments actors*, in Acts of the international Conference of Territorial Intelligence, Vol 2 caENTI, Ailba Iulia, October 21st-22nd, 2007

LAZZERY Y., 2006, *Les indicateurs territoriaux de développement durable. Questionnement et expériences*, L'Harmattan, 323p.

PEHANI P., OSTIR K., LANDRE M., 2008, *Interactive Map for caENTi – Application of the webmapping technology*, in Acts of the International Conference of Territorial Information 2008, UFC : Besançon, à paraître

SANCHEZ-LOPEZ C., GIRARDOT J.-J., 2006, *Specifications of the contents of the european guide of diagnosis and evaluation*, in Acts of the international Conference of Territorial Intelligence, Vol 2 caENTI, Alba Iulia, October 21st-22nd, 2007

VERMEYLEN P., 2006, *Aménagement et urbanisme, L'Agenda 21 des Communes durables*, Namur : CPDT, 97p.