



CAENTI
Coordination Action of the European Network of Territorial Intelligence
A project funded under FP6 of the E.U.
<http://www.territorial-intelligence.eu>



CAENTI

Coordination Action of the European Network of Territorial Intelligence

Conceptual and methodological specifications
for a Territorial Information Community System,
including technical specifications for integration of
“PRAGMA” with the software of qualitative data
analysis “ANACONDA” and “NUAGE”.

March 2006, 1st – December 2006, 31st

Deliverable N° 55

Jean-Jacques GIRARDOT
Work package leader
Université de Franche-Comté

Cyril MASSELOT
Coordination group leader
Université de Franche-Comté

December 2006, 31st

Summary

| | |
|---|-----------|
| INTRODUCTION | 3 |
| 1. Integration of data and spatial analysis software | 4 |
| 1.1. Data analysis software | 5 |
| 1.1.1. PRAGMA software | 6 |
| 1.1.2. ANACONDA software | 8 |
| 1.1.3. NUAGE software | 10 |
| 1.2. Spatial analysis software | 11 |
| 1.2.1. SITRA action and software | 11 |
| 1.2.2. ICASIT research project | 14 |
| 1.3. Technical specifications for PRAGMA integration with the software of qualitative data analysis ANACONDA and NUAGE, and with a Territorial Information System. | 14 |
| 2. Integration of analysis functions, editorial workflow and uses in a Territorial Intelligence Community System. | 16 |
| 2.1. Territorial Intelligence Community System | 17 |
| 2.2. Integration of data analysis protocols | 19 |
| 2.2. Integration of the editorial workflow | 19 |
| 2.3. Integration of the community uses | 20 |
| Conclusion | 22 |
| Bibliography | 23 |

INTRODUCTION

Research activities about the integration of the software of quantitative study PRAGMA and the software of qualitative analysis ANACONDA and NUAGE, were quickly enlarged to the design of a “Territorial Intelligence Community System” TICS.

During the evaluation and negotiation period of CAENTI project, we can quickly develop new versions of data analysis software thanks to the help of the French Ministry of Research and Higher Education and the French National Center of Scientific Research, which funded ICASIT project through the Institute of Humanities and Social Sciences (MSH) and the Institute of Information and Communication Technologies. ICASIT action allowed developing new standalone multiplatform version of NUAGE, which represented a good step towards integration.

On September 2005, we showed in an invited communication “Integration of data analysis tools within territorial information systems” during the International Conference on Theory and Applications of Mathematics and Informatics, ICTAMI 2005, in ALBAC (Romania) [GIRARDOT, 2005] a first general schema of integration for data and spatial analysis.

One new idea of this communication was that a chain of data processing functions result from this integration, from data collection to Internet publishing.

During the first month of CAENTI, the matter of the editorial aspect of this chain became clearer. The actors used more websites and they wanted to publish more quickly the results of data analysis as information about their activity. They wanted that tools provide tables, graphics and maps directly publishable on Internet. The coordination group wp6c directed the research activities about the management of documents, as an editorial workflow towards an information system.

So wp6p directed the research activity towards information systems, more precisely towards territorial information systems which can be accessible to actors and facilitate the distributed data gathering and the cooperative analysis, we called territorial intelligence systems, as an information system that did not only allow information functions but also allowed improving the collective knowledge and intelligence of actors to better act together.

On September 2006, during the International Conference of Territorial Intelligence in ALBA IULIA (Romania) we presented a communication about "the editorial function of the territorial intelligence systems."

When we are working with wp6g on the data analysis protocols and the use of CATALYSE tools, we enlarge the topic of the collective use of PRAGMA to the use of territorial intelligence system tools by a community or, more concretely, by a development partnership. In the context of the current evolution of Internet towards social use, we bridged Internet community and community development, as a group of people joined by a common project. It allowed clearly introducing the uses of a community, in the design of the system, in accordance with the wp6g governance that modeled the CATALYSE governance in the deliverable n° 56 “ Guidance notes for the use of CATALYSE information and tools”.

We will successively present:

- The integration of data and spatial analysis software, within the technical specifications for PRAGMA integration with the software of qualitative data analysis ANACONDA and NUAGE.

- The integration of analysis functions, editorial workflow and uses within a Territorial Intelligence Community System.

1. INTEGRATION OF DATA AND SPATIAL ANALYSIS SOFTWARE

Our research activities about data, then spatial, analysis software follow the diffusion of information and communication technologies.

The computer science, information systems, digital networks, and then digital edition via Internet development allowed us designing and progressively making solutions that appeared to be useful for the territorial information systems. Computer science firstly allowed constituting important data sets and analyzing them with quantitative statistical methods. Then it made possible the application of statistical qualitative methods as well as the factorial analysis or classification [BENZECRI, 1982].

The data sets developed, and then it was the case of the information systems. Information that was conserved on different supports, were digitalized so as to be massively stocked on a same support within which they lose their readability. Internet popularized the information sharing, their collaborative processing and the massive publication at the world scale.

Geographic Information Systems became progressively popular in the Humanities and Social Sciences. They offer simple mappings as more complex spatial analysis functions.

Because it is incontrovertible within a territorial approach, they complete data analysis in territorial intelligence.

1.1. Data analysis software

Since the 1970's we have developed convivial data processing solutions in order to diffuse data analysis protocols coming from Matter and Nature Sciences among the researchers in Humanities and Social Sciences.

They are based on a notion that constitutes a base of the editorial function of the TICS: friendliness.

The software ANACONDA – Analyse CONViviale des Données (data friendly analysis) was published in 1981 to diffuse the data qualitative analysis methods (factorial analysis and classification) in Humanities and Social Sciences.

PRAGMA, which is quantitative analysis of questionnaires software, was made in 1991 to facilitate the drafting of the data tables so as to make a data statistical qualitative analysis.

NUAGE completed ANACONDA with the representation of data analysis results.

The concept of friendliness that is fundamental in their conception implied the ideas of economy, simplicity in the use, accessibility, sharing and cooperation. During their development within the research group “Techniques Nouvelles en Sciences de l'Homme” (new techniques in humanities and social sciences), the conception of this software took into account specificities that are proper to:

- Humanities and Social Sciences (HSS);
- Multi-discipline uses;
- Systematic and collective gathering of large data sets;
- Complementarily with spatial analysis.

We will present PRAGMA, ANACONDA and NUAGE in the order of the data processing chain.

1.1.1. PRAGMA software

PRAGMA consists in software of analysis of questionnaires surveys that is directed towards the data quantitative processing. It was initially developed in 1991 [GIRARDOT, 1991] to make the coding operations that are necessary:

1. To make simple and cross quantitative sorting;
2. Then, to make the data table in the prospect of the data qualitative analysis that will be made by ANACONDA.

The specifications of PRAGMA have been defined in the deliverable 54 “Specifications for the software PRAGMA of quantitative collection and treatment software”, so we just remember the basic principles and functions which are important for integration.

The basic principles of its conception are

- **Intuitive use.** PRAGMA aimed being a simple to use, friendly software. The data keying in has become an operation that can be made by any person who know how to use a word processor. The researchers and the professionals could make the coding, recoding, selection, statistic sorting operations, and then complete the exploitation and communication of the results thanks to usual office automation tools.

- **Collective tool.** PRAGMA was designed for a collective use, to mutualize and share the data upstream, and to collectively analyze the results of the data analysis downstream. It allowed gathering data entries that were made by uncollected users without any complication, very closely to the survey places. PRAGMA also emphasizes the collective dimension of friendliness that is based on the concepts of sharing and cooperation. It favors the constitution of a community that is united by a common project, and the transfer of scientific and technical skills about the methods, protocols and tools of data analysis.

- **Robustness and automatic control of the data quality.** So as to favor the access of an increasing number of users to the information quantitative and qualitative analysis techniques PRAGMA integrates control procedures of data quality. These controls essentially concern the questionnaire and the questions form. These formal controls make PRAGMA a robust tool. People without any data processing or statistical qualification could use it in extremely diversified conditions, closest to the observation situations.

According to the first aim of this report, the integration between quantitative and qualitative data analysis function, we are particularly interested in the coding and recoding

functions that provide help to the user who prepares data for quantitative study and qualitative analysis:

- Management of the missing answers, according to a data qualification protocol;
- Coding of the opened answers;
- Recoding of the closed questions;
- Selection of questions and modalities;
- Coding in the prospect to make data analysis.

We define and analyze in the deliverable 54 these functions as the steps of the analysis chain. They presently need more automation, appraisal and assistance to avoid as many errors as possible. Nevertheless, each step must also organize a human decision. The global quantitative balance, updated at the end of each steps, is currently the tool for deciding then controlling the good implementation of the decision.

The missing answers management concerns at the same time individuals and questions. It is the first step of the protocol of data qualification insofar as the control of missing answers leads to the elimination of individuals and questions. It is possible the next steps eliminate data, so this control must be repeated at the end of each step.

The opened questions coding aims at transforming them into closed questions that is to say the answers that are expressed under values form into codes. PRAGMA simply offers to gather values according to a new modality. It is also possible to directly code a value which meaning already corresponds to an existing modality.

The recoding of the closed questions is a similar re-gathering operation that does not concern values any more but answers that already have a codes form. It concerns the modalities which frequency is too low to be considered as representative. PRAGMA allows removing these modalities, or when the modalities meaning allows it, gathering them with other modalities or some of them so as to avoid the information loss.

Recoding is also a mean of questions or modalities selection according to the stage of analysis, exploratory or deepened, or a mean to prepare the results communication.

These coding, recoding and selection functions were particularly developed to select the characters that will constitute the Boolean data table in the prospect of the data qualitative analysis. PRAGMA allows globally selecting the questions to be processed and then to

choose and define the characters to be analyzed from modalities, possibly by gathering several modalities.

The “data table” constitutes an important link for PRAGMA and ANACONDA integration. Conversely, the typology that is determined by ANACONDA can be imported in PRAGMA as a question which classes constitute the modalities, so as to calculate the frequencies of the original modalities in each class, by a cross sorting.

These operations are also useful for the results communication, to associate the partners who participated to the data mutualization to the results interpretation and to diffuse the results on Internet. In both cases, the gross results of the quantitative analysis cannot be published in their state. There should be a selection among them, and specific recoding.

We emphasized the need to make restitution the results to the partners’ network who accept to share their data. With PRAGMA, this restitution is made at an individual level, since the keying in, since each partner can make the quantitative and qualitative analyses he wants from his data, including before putting them in common. But restitution has above all a collective dimension, after the data gathering. The partners can participate to the analysis, direct the processing progression and enrich the results interpretation through the contribution of experience data, information or specific knowledge. This restitution implies an animation that is based on the results communication.

As regards the publication of the results on Internet, it implies a progressive drafting of the documents under a digital form. These documents have many particular characteristics. They are new in information systems that are generally designed for more traditional documents. They are automatically produced, but they are also re-made by human interventions according to technical or communicational imperatives. They are also very varied.

1.1.2. ANACONDA software

It was developed in 1981 [GIRARDOT, 1981]. Its main objective was to provide a cheap and accessible solution to analyze data. They are multi-criteria methods that allow analyzing statistical individuals who are described by a set of multidimensional characters. The factorial analysis determines the structural factors or trends, of an important data set. The classification dissociates its main classes. These methods imply important calculations that impose to use a computer, contrary to the quantitative tables that we can “manually” make [BENZECRI, 1982]. During the 70’s, these calculations were made in calculations centers,

what cost much money for the HSS laboratories. As soon as the characters were more than one hundred or the number of individuals was higher than one thousand, we had to go to PARIS. Indeed, there were frequently data sets composed by more than a thousand of individuals in geography. ANACONDA allowed analyzing data sets until one hundred characters and one thousand individuals on an Apple II (48k). These limits were quickly pushed back.

ANACONDA combines Jean-Pierre BENZECRI's Correspondences Factor Analysis with Michel ROUX's Hierarchic Ascendant Classification¹ [1982]. It is based on a matrix diagonalization algorithm that was designed by Jean-Philippe MASSONIE and initially programmed by Xuan LUONG.

ANACONDA is directed to the HSS researchers; it was designed as free and easy-to-use software for users without computer science qualification. Micro processing allowed simplifying the use. Of course, it was possible to make data analysis on one's personal computer. It presented a very simple interface in comparison with the data processing software. It was interfaced with the usual office automation module. ANACONDA was designed to import the data in the form of a data table², that is constituted by a word processing or a spreadsheet and it is saved in a file that is in format "text"³. It is totally automated for a current use. It applies a principle that was popularized by Jean-Pierre FÉNELON [1981]: "do we need a thesis in thermodynamics to drive a car?" It does not need any other knowledge in mathematics than the ones that are useful to understand the results. The choices are defined by parameters with initial values that correspond to 99% of the uses. It is enough to indicate the name of the file that includes data to analyze so as the calculation of contingences, the factorial analysis and the ascending hierarchic classification are made. Then, the "NUAGE" software simplifies the results analysis and interpretation thanks to a 3D animation.

But, ANACONDA is above all adapted to the main methodological specificity of the HSS where the analysis mainly concerns the individuals in comparison with the characters. The experimental methods that come from the Matter Sciences focus on the relations between the variables, whilst neutralizing the individuals. More than in the Nature Sciences that are

¹ It is based on the distance of the Chi2 and on the aggregation by the focused moments of order two.

² The data table presents the individuals on lines and their characters on columns. The first line is reserved for the characters identification codes and the first column is reserved of the individuals' indicators.

³ Suffix txt. Each line of the table is separated from the continuation by the "return" character (ASCII 13). In each line, each column is separated from the following one by the tabulation character (ASCII 11).

characterized by the variety, the HSS are singularized by the individuals' diversity and by individuality. The main interest of the qualitative analysis consists in its ability to study this diversity. We chose the correspondences factorial analysis that allows analyzing not only the relations between the variables but also the relations between the individuals, and the relations between the individuals and the variables. The classification allows studying the individuals' diversity by determining the main classes of individuals and their characters' profiles. ANACONDA allows associating both approaches by the representation of the classification in the factorial space.

Thus, the ANACONDA friendliness is based on:

- Automation of functions which knowledge is not indispensable to the results understanding;
- Choice parameters definition with initial values that cover the usual situations;
- An interface with the most popular office automation standards;
- An adaptation to the specificities of Humanities and Social Sciences.

ANACONDA was used in diversified fields: in the scientific research on the landscapes analysis in geography and on the structure and dynamic of the archaeological establishments. It is an generic toll used in diversified research activities fields of applied research and research-action: studies of the structure of the museums collections and of the audiences of cultural manifestations, diagnosis of the variety and of the complexity of the needs of underprivileged people and evaluation of the actions that aim at improving these persons' situation, etc.

1.1.3. NUAGE software

NUAGE is software of representation in three dimensions of the results of the Correspondences Factorial Analysis of and of the Hierarchic Ascendant Classification. It was developed in 2003 so as to replace the American software MACSPIN which development had been stopped since 1991. The software of data graphic analysis allowed making a three-dimension representation of the ANACONDA results. This representation was much simpler to use than the representations through maps on sheets of paper. It was above all more intuitive and allowed the collective interpretation of the results of the qualitative analyses in small groups. But MACSPIN could also be used on Apple computer and it became almost impossible to use it after 2002, because of the generalization of the exploitation system "X" of APPLE.

NUAGE that was developed in JAVA by Hervé FEHNER is multi-platform software. It makes the restitution of all the graphic functions of MACSPIN, but its interface is specifically adapted to the double representation of the factorial analysis and of the classification and to the interpretation of the data qualitative analysis. Indeed, it automatically integrates a series of operations that had to be manually made before. It also memorizes the results of the analyses and representations. Despite an important work on the interface, it essentially produces work documents. The thinking should presently direct to the production of documents that can be published.

1.2. Spatial analysis software

Thanks to collaboration with computer science researchers, the SITRA, Territorial Information Systems for Actors Network, research action allowed completing the data statistical analysis by tools of spatial analysis and cartography.

ICASIT, Integration of Analytical Chain of the Territorial Information Systems, started updating and making the software of data statistical and spatial analysis that were previously developed evolve.

1.2.1. SITRA action and software

The research action « Territorial Information Systems for Actor Network » was led from 2002 to 2004, so as to model convivial territorial information systems that are accessible online to networks of local actors of the sustainable territorial development so as to develop their ability of mobilization, decision, action and evaluation. The objective was especially to define a TIS architecture that is adapted to the work in network and to complete the data statistic analysis with convivial cartographic solutions. The SITRA solutions, often called SITRA software, were applied to draft the ADRI (Agency for the Development of the Intercultural Relations) resources management system, the resources management system of the Inter-ministries Delegation of Town (DIV) and to make the territorial “CATALYSE” observatories evolve.

The word « network » referred to the reality of actors who cooperate on a common project, but working in different organizations at the institutional level, without any hierarchic link and being geographically far away the ones from the other ones. It prefigured the concept of partnership that constitutes a fundamental economical principle of sustainable development. The latter one advocates the use of local resources in a spirit of cooperation rather than the creation of new structures. It implies a multidiscipline and multisector approach. Partnership

is presently a current word at the level of territorial action. The European programme of social-economical action “EQUAL” funds for example actions that are led by consortiums that are called “development partnerships”. Because of the development of cooperation at the level of the data analysis thanks to PRAGMA friendliness, the SITRA stakes always concerned the information sharing and its cooperative exploitation. From now on, we had to mobilize in coherent way extremely varied information that is referenced in space and time. They should be statistically analyzed or be used for something else. The SITRA action gathered researchers in Human and Social Sciences and Computer Science, whereas we used to make software of data analysis inside the organizations. We wanted to use the Internet potential and it implied wider knowledge and skills in computer networks, modeling of information systems and databases.

SITRA allowed:

- Starting the integration of functions of statistical and spatial analysis by designing friendly cartographic tools.
- Defining a procedure of global management of all the documents, despite their variety.
- Designing a system of rights that would be appropriated to a partnership management.

The word « Territorial Information Systems » (TIS) refers to a project that would be at the same time more ambitious and simpler than the one of « Geographical Information System » (GIS). The GIS allow geographically referencing the information, making spatial analyses, analyzing them and mapping the information and results. From this point of view, a GIS is the TIS basis module. Nevertheless, the experience of the “ARCHAEOMEDES” project and of the “CATALYSE” observatories shown the data bases of the GIS could not manage in a coherent way neither the very varied involved documents, nor the actors diversity nor the documentary flow. Besides, their use remains expensive and complex and the performances poor as regards digital edition on Internet. If this cost and this complexity could be justified for specialized spatial analyses, it was not the case to make simples spatial analyses and to edit maps, functions for which there was a free and more friendly offer. It was necessary that the TIS architecture be inspired by a GIS structure to integrate these functions. ARCHAEOMEDES uses a GIS to make spatial analyses, whereas the bad called SIGVILLE draws maps according to the demand as regards public social-economical indicators that are

appropriated to the territories observation and to the evaluation of devices on all the territories of the town policy (around 700 infra and inter communes zones) thanks to a cartographic online tool (ALOVMAP). The manipulation is extremely simple as it is enough to select the zone and the data to map. This friendly solution is also used for the cartography of the diagnoses and of the territorial indicators on the “CATALYSE” observatories, by using in particular files in format text to map the results that come from PRAGMA and ANACONDA. Indeed, SITRA kept making the integration of the statistic software with the new cartographic functions. It started the execution of a free and multi-platform version of ANACONDA in Java, and a free online prototype of PRAGMA, epragma, in PHP/MySQL.

A TIS manages the access online of a great variety of documents in Extranet secured network or on Internet. SIGVILLE also manages the documentation that is linked to the actions of the town policy (contracts, maps, reports). It is an element of the website of online resources of the town policy: books, magazines, circulars, reports, maps, conventions, numbers tables, graphs, experiences cards, links to partners websites... This variety is also enriched by the multi-disciplinary and multi-sector approach that characterizes territorial action. The website of the ADRI (that has become the “Cité Nationale de l'Histoire de l'Immigration”, CNHI, French national city of the immigration history) offers a simplified system of metadata so as to improve the accessibility of the data through Internet. Whilst it offered several researches according to the kind of document, SITRA modeled for the Resources of the Town Policy a TIS that references and that manages in the same way referenced and digital documents. Thus, we have access with only a key word to the whole documents that are linked to a theme, whatever their original support, their digital format, their kind or their function. A TIS is characterized by the production of new documents from other documents. The documents management also concerns the referencing of the new documents and the management of the analytical flow.

SITRA also started the modeling of a security system that is adapted to the partnership networks, by offering a procedure that has generalized from this moment. The administrator offers a login and a password that can be modified by the user. SITRA has the same organization of individuals in groups that PRAGMA. This notion offered the possibility to gather the persons who are informed by a same structure or who live in a same territorial zone... SITRA introduced the differencing of the accesses in function of the groups, what allowed, by defining the groups according to an organization for example, limiting to the users of this organization the right to have access to the persons who belong to this group. The

works of secure of the TIS led to deepen the link between metadata, document and information, according to the information updating. We can illustrate this thinking as follows. We can constitute for the cartographic representation of indicators a table, which columns are the indicators and the lines the geographic units to be mapped. We can consider that this table is a document and attach to it metadata, as the updating date, and rights. It will be insufficient if different users can independently update each indicator or each zone. In this case, we should attach the metadata and the rights to the column or the line that corresponds in the table. If we are susceptible to independently modify each piece of information, we should attach metadata to each piece of information of the table.

Thus, SITRA allowed making a first list of specifications to make TIS and experimenting them with important executions as SIGVILLE.

1.2.2. ICASIT research project

The ICASIT action “Integration of the Analytical Chain of the Territorial Information Systems” followed SITRA from 2004 to 2006 by emphasizing the notion of analytical chain that will eventually lead to the one of editorial chain. SITRA put the bases of an analytical chain that integrates the software functions within the TIS, from the data gathering with PRAGMA until the online publication of the results, mainly maps. ICASIT followed the integration effort of the software of analysis with two practical objectives:

1. Making integrated, free and multi-platform versions of the TIS software functions.
2. Putting these functions online.

NUAGE was developed in multi-platform version JAVA to represent in 3D the ANACONDA results. The specifications of a multi-platform version of PRAGMA were established. The cartographic functions were extended with the flows representation.

A beta version of online PRAGMA, epragma, was experimented in the framework of the ACCEM migrations observatories.

1.3. Technical specifications for PRAGMA integration with the software of qualitative data analysis ANACONDA and NUAGE, and with a Territorial Information System.

The integration of the data analysis software will follow two steps:

- Integration of ANACONDA and NUAGE, in a new ANACONDA 2.0 version;

- Development of a multiplatform version of PRAGMA in JAVA and integration with ANACONDA 2.0 in a future ANACONDA 3.0 version.

The next schema (diagram 1) shows the complete integration of data analysis and spatial analysis functions.

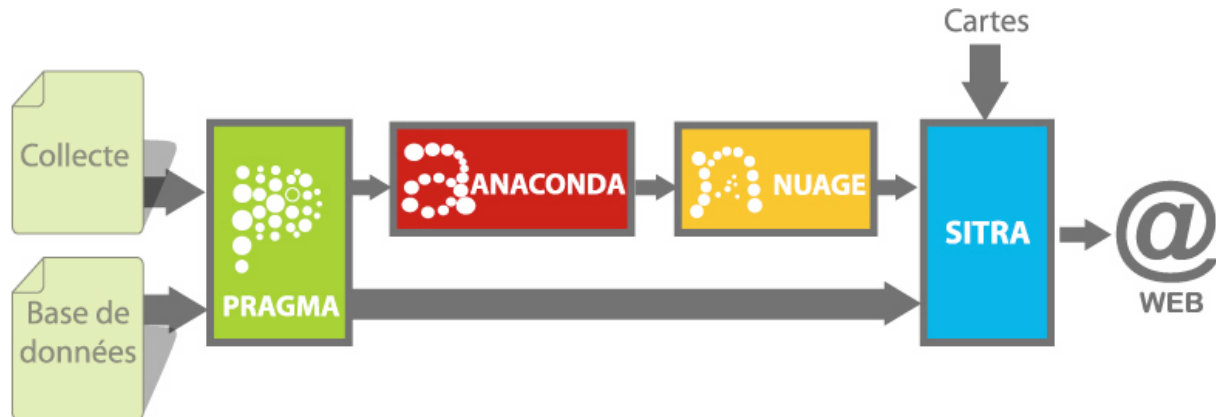


Diagram 1: Schema of data analysis and spatial analysis software
(Collecte : data collection, Base données : database, Cartes : maps, SITRA : TIS)

The integration of ANACONDA and NUAGE, is simple to do because:

- We achieved a standalone multiplatform and multilanguage of ANACONDA (version 1.9.2.) in JAVA, as previously for NUAGE (version 0.9.1.)
- They are presently linked by only one intermediate file.

So this integration can be the first step in the integration between quantitative and qualitative analysis software. We propose that the new software be simply called ANACONDA 2.0. As actually it will provide the following results files, all in a “text“ format:

- A technical file (VEC) with the mathematical results of matrix diagonalization;
- An historical results file (AD) that allows printing all the results as during the good old times (we need to control the 3D representation function that allows consulting all the results)
- The "contingence table" (CTG) that can be edited as a data sheet.
- The intermediary file (NGE) that memorizes the ANACONDA results and the work with the 3D representation software, NUAGE.

Then, the integration of PRAGMA and ANACONDA 2.0, assumes the following steps:

- Improving standalone multifunction and multilingual versions of PRAGMA, a “collect” version and a “complete” version, as recommended in the PRAGMA specifications (deliverable n° 54).

- Integrating it in an ANACONDA 3.0 version.

The intermediary files from PRAGMA to ANACONDA (see diagram 2):

- Data table (TAB);

- Codes meaning list (COD) will stay useful (but all information must be accessible on screen, without paper or edited file)

The integration with PRAGMA must not only be designed from a technical point of view. The concept of data analysis, based on characters, is useful to solve many problems of quantitative study, based on questions and modalities, as missing answers, multiple answers and questionnaire evolution. So we recommend to use the data collecting mnemonic codes, as data analysis, instead of digital ones.

ANACONDA also provides a better solution with contingences tables instead of PRAGMA cross-sorting.

We presently need to study the retroactivity between quantitative and qualitative analysis. For example we reintroduce the result of the classification made by ANACONDA as a question in PRAGMA to calculate the characters frequencies within classes. We can automate this link and develop others. It can also be directly done with a contingency table. The integration between PRAGMA and ANACONDA will allow such links.

In this schema SITRA assigns a Territorial Information System of which the deliverable 56 contains (in chapter four) all the technical specifications.

2. INTEGRATION OF ANALYSIS FUNCTIONS, EDITORIAL WORKFLOW AND USES IN A TERRITORIAL INTELLIGENCE COMMUNITY SYSTEM.

This second part of report summarizes general considerations. They are less results than research orientations.

We will present:

- The concept of Territorial Intelligence Community System

- The present realizations and prospects orientations of the coordination research.

2.1. Territorial Intelligence Community System

Here we want to distinguish territorial information system and territorial intelligence system, these concepts introduce the concern for the community interest.

An **information system** is a structured whole of data bases that gather a defined set of information and functions that allow establishing relations between these information, defining the direction of these relations or, better, measuring them.

In a **territorial information system (TIS)**, the information concerns a territory, that is to say a geographical space and its community. They are referenced in space and time. Functions of statistical and spatial analysis aim at helping the territorial actors producing knowledge on this territory.

Thus, a TIS allows:

- Gathering and regrouping information;
- Analyzing them according to scientific protocols, especially in a spatial prospect;
- Interpreting the results of these analysis;
- Representing the information and the results;
- Drafting decision-making scenarios;
- Managing and valuating the policies, programs, devices, projects and actions that result from the decisions.

A **territorial intelligence system** is a Territorial Information System that is at the service of a partnership of territorial actors that want to develop democratic governance at the service of sustainable development.

- It favors the information sharing within a territorial development partnership;
- It instruments the data cooperative analysis and the results participative interpretation;
- It introduces the citizens' participation in the process of decision-making;
- It provides the actors with the useful information to draft projects, and then to manage them and value them.

It exploits the potential that computer science offers to gather information, share this information, favor the actors' partnership from the information cooperative analysis, and to

increase the citizens' participation by improving their information and their access to information. It respects the constraints that sustainable development, particularly participation and partnership, imposes to information processing and publication protocols.

Our research activities converged, since the development of ANACONDA, from the friendliness notion toward the editorial chain one. This evolution is linked to the information technologies diffusion but also to the permanent concern of making them accessible to a great number of people without any prejudice for the information quality.

Working on analytical chain, we became aware of the matter of its editorial function that is directed towards online edition of the data analysis results. In the context of the territorial intelligence systems, this editorial chain concerns production of digital documents that should be shared within an actors' partnership. It is also a work flow that can not be organized only according to the technical protocols of data analysis, but also so as to allow participation within the partnership, but above all beyond it, within the territorial community. That is why we call them Territorial Intelligence Community Systems (TICS).

The word « territorial community » does not refer to the biological, historical or religious dimension. It assigns a set of people that are united by a common project. It is also the meaning that is used in computer science to name a users community. This meaning covering will allow studying the conditions in which the territorial community can constitute the TICS users community. We should also take into account the fact that if the information analysis can constitute a project for the TICS: it is an intermediary step at the service of an objective of sustainable development for the territorial community, or more simply for the local actors who want to plan, lead and evaluate actions of local development.

Deepening and instrumenting the TICS editorial function is the research prospect of CAENTI after the CATALYSE Toolkit. It corresponds to our initial plan to work on an Inclusion Itinerary Accompaniment File in 2007, then on a Internet Territorial Information System in 2008.

In addition to the analysis function integration, we engaged the following research orientations:

- Integration of data analysis protocols.
- Integration of the editorial workflow.
- Integration of community uses.

2.2. Integration of data analysis protocols

The effort of evolution and integration will continue with the constant concern to improve the accessibility of the software data analysis tools following data analysis protocols. The definition of a multi-sector guide of diagnosis and evaluation offered a selection of questions and for each question a precise data analysis protocol according to the expertise and experience of the CAENTI partners (chapter 5 of deliverable 56).

- By using and making free software that are widely opened to the developers' communities.

- By making multiplatform versions, which work on all the computers, on all the exploitation systems and independent from their evolution.

- By making multilingual software that are widely accessible to the users' communities.

- By automating, as it was initiated with ANACONDA, all the functions of which neither the expertise nor the technical control are indispensable to understand the results. The operational and decisional steps remain; they need an intervention of the users via the man-machine interface. The pursuing of the processing implies new information that is generally transmitted through documents or decisions that are transmitted under the form of parameters that condition the treatment.

- By integrating the expertise. The experts usually attribute initial values to the parameters. The user has a proposal that corresponds either to the usual uses or at expert advice, and that they can modify. For example, in SIGVILLE, the user can request the representation of a map by indicating the zone and the indicator. Experts define the choices of the classes and of the colors representing each indicator as parameters. They integrate an expertise on the one hand that is adapted to each indicator and on the other hand to the cartographic representation.

- By memorizing the users' decisions. The evaluation and observation steps often imply the repetition of decisions from a period to another one. The parameters are used to memorize the decisions.

2.2. Integration of the editorial workflow

The research activities that are led within SITRA on the integration of the functions of information analysis emphasized the documents flow that is gathered and produced within the

TIS, from the gathering to the data publication. As we previously underlined, a TIS references and digitalizes a great variety of documents, among which some remain few usual (statistical tables and graphs, maps...) that is widens by the production of the production of digital documents. ICASIT put the bases of the modeling of the documentary flow, by starting the inventory of the used (input) and produced documents (output) by the analytical chain. The objective is to define a metadata system that describes the information source and the made processing, if possible in an automatic way. In the continuity of the effort of simplification that was initiated by SITRA, these research activities aim at integrating the “DUBLIN CORE” standard, especially the norms that are presently developed for the geographical information. It is also a condition for the metadata to constitute vectors of communication with other systems.

We can indicate some specifications for meta data that are an important key for the editorial chain integration:

- All the information (that are numeric, imported and produced) should be documented.

- The meta-data should be linked to the information source and to the data processing that are made in the analytical and editorial chain.

- As it is an important job, the meta-data should be able to automatically generated.

- As the territorial information systems mix extremely different types of information, among which some remain few used (statistical tables and graphics, maps...) and meta-data are communication vectors with other systems and portals, the meta-data of the SIT should respect the standard (Dublin Core), even the standards that are currently developing (for example for geographical information).

2.3. Integration of the community uses

ICASIT emphasized the TIS editorial function and their community dimension.

The integration of the analytical process on Internet is conditioned by the online publication of the results. The knowledge production and the information documentation are inscribed in a “workflow” of digital scientific and technical edition that organizes the continuity of the technical functions from the information gathering to the results publication, via the data sharing, the results analysis and interpretation, the documents drafting and documentation and their validation.

Thanks to the digital edition on Internet, the TIS offer a huge potential for the sustainable development. The cooperation online tools are indispensable to the partnership. Digital edition allows informing the territorial community, guaranteeing the visibility of its identity and developing participation by the means of interactive services. This potential makes the TIS systems of territorial intelligence that offer concrete solutions of instrumentation of partnership and participation. Nevertheless, the experts initially control this potential. Territorial information accessibility and conviviality of processing are the conditions to release this potential at the service of the collective intelligence of the territorial community.

This orientation implies the categorization of the TICS users – *a priori*; experts, actors, partners and community, and the study of their uses. Which functions need an expertise? Which functions are transferable to the users? According to which pedagogy and with which accompaniment? How to organize the sharing of the territorial information, and then the cooperation at the level of their analysis and their edition? Which are the actors, partnership, and territorial community functions? How to associate the community by the means of the participative interactive systems? This modeling implies an adaptation of the steps of the data analysis and of the results interpretation, as well as of the editorial flow, to the communication, animation objectives of the partnership, the participation and of governance of the territorial community.

The following schema the four levels of modeling and research:

- functions of data analysis;
- data analysis protocols, according to deliverable 56 (chapter 5), to be completed;
- documents chain and editorial workflow;
- community uses, according to CATALYSE governance in development partnership established by coordination group wp6g (chapter 2 of deliverable 56).

- The modeling of their architecture and the definition of their functional specifications according to logics that are proper to analysis, documentation, edition and communication in the partnership and participation framework;
- Deepen the description of the processing protocols;
- The modeling of the contents, documents and editorial workflow by integrating the human individual and collective interventions;
- The metadata specifications;
- The modeling of the decisional, governance, partnership and participation process within the territorial community.
- The uses evaluation;

BIBLIOGRAPHY

- BENZECRI J.-P., 1982: "Histoire et préhistoire de l'analyse des données.", Dunod, PARIS.
- BENZECRI J.-P. et F., 1976: "Pratique de l'analyse des données.", 2 tomes, Dunod, PARIS.
- FENELON J.P., 1981: "Qu'est-ce que l'analyse des données ?", Lefonen, PARIS.
- GHIGLIONE R., MATALON B., 1978: "Les enquêtes sociologiques." Théorie et pratique, Colin, PARIS.
- GIRARDOT, J.-J., 2006: "Activities and prospects of CAENTI". in: *Acts of International Conference of Territorial Intelligence*, ALBA IULIA (Romania), September 20th-22nd, 2006, 9 p. <URL: <http://www.territorial-intelligence.eu/telechargement/albaiulia2006/Alba06-Girardot1.pdf>>
- GIRARDOT, J.-J., 2006: "Activities and prospects of research activities concerning tools of territorial intelligence for sustainable development actors. Work Package 6 "Tools for Actors" of CAENTI", in: *Acts of International Conference of Territorial Intelligence*, ALBA IULIA (Romania), September 20th-22nd, 2006, 8 p. <URL: <http://www.territorial-intelligence.eu/telechargement/albaiulia2006/Alba06-Girardot3.pdf>>
- GIRARDOT, J. -J., 2006: "The editorial function of the territorial intelligence systems". in: *Acts of International Conference of Territorial Intelligence*, ALBA IULIA (Romania), September 20th-22nd, 2006, 6 p. <URL: <http://www.territorial-intelligence.eu/telechargement/albaiulia2006/Alba06-Girardot4.pdf>>
- GIRARDOT, J.-J., 2005: "Concepts, principes et outils de la méthode CATALYSE", in : *3e colloque international du Réseau Européen d'Intelligence Territoriale*, LIÈGE (Belgique), 20 et 21 octobre 2005, 5 p. <URL:<http://mti.univ-fcomte.fr/reit>>
- GIRARDOT, J.-J., 2005: "Integration of data analysis tools within territorial information systems.", International Conference on Theory and Applications of Mathematics and Informatics, ALBAC (Romania), September, 15-17.
- VAN DER LEEUW S., FAVORY F., GIRARDOT J.-J., 2004: "The archaeological study of environmental degradation. An example from Southern France.", in CH. L. Redman & al., "*The Archaeology of global change*", Smithsonian Institution, WASHINGTON, 112-129.
- GIRARDOT, J.-J., 2000: "Principes, Méthodes et Outils d'Intelligence Territoriale. Évaluation participative et Observation coopérative", in : *Conhecer melhor para agir melhor, Actes du séminaire*

européen de la Direction Générale de l'Action Sociale du Portugal, EVORA (Portugal), 3-5 mai 2000, DGAS, LISBONNE, décembre 2000, 7-17.

GIRARDOT, J.-J., 1987: "Expertise du logiciel MACSPIN d'analyse graphique de données.", Sciences et Vie Micro, n°37, 100-103, PARIS (France).

GIRARDOT, J.-J., 1982: "ANACONDA, système conversationnel d'analyse des données.", Cahiers du SURF, 1, nouvelle série, 37-74, Université de Franche-Comté, BESANÇON (France).

MASSELOT C., 2006: "Systèmes d'information territoriaux et politiques sociales : quand l'observation territoriale s'empare du net", in : 15ème Colloque de la Société Française des Sciences de l'Information et de la Communication, BORDEAUX (France), mai 2006.

MASSONIE, J.-P., 1990: "Pratique de l'analyse des correspondances.", Collection "Techniques Nouvelles en Sciences de l'Homme", Annales Littéraires, BESANÇON.

ROUX, M., 1985: "Algorithmes de classification.", Masson, PARIS.

SANDU D., 2005: "Dezvoltare comunitară. Cercetare, practică, ideologie.", Iami, Editura Polirom.

Sélection de liens :

Portail Internet du centre de ressources de la délégation interministérielle à la ville (DIV, PARIS, France) <http://i.ville.gouv.fr>

Système d'information géographique de la politique de la ville (PARIS, France)

<http://i.ville.gouv.fr/sigville>

Observatoire des migrations ODINA en Asturies (Espagne) <http://www.odina.info>

Observatoire des migrations OPASI de SIGÜENZA (Espagne) <http://mti.univ-fcomte.fr/siguenza>

Observatoire communautaire OPTIMA de SERAING (Belgique) <http://www.optima-obs.org>

Observatoire Intercommunal INTEGRA (Belgique) <http://www.integraplus.be>

Agence pour le Développement des Relations Interculturelles ADRI (PARIS, France)

<http://www.adri.fr>

<http://www.alterites.com>,

<http://www.histoire-immigration.fr>