

INTEGRATION OF THE E&T DATABASE IN THE E-PRAGMA SYSTEM

Catherine Caille-Cattin

Assistant Professor of Geography
THEMA-30, rue Mégevand
25030 Besançon cedex

Ahmed Hammad

Assistant-Professor of Science Computer
LIFC-16 Route de Gray
25030 Besançon cedex

Jean Louis Poirey

Assistant Professor of Geography
THEMA-30, rue Mégevand
25030 Besançon cedex

Remi Thomas

Phd Student Teritorial-Intelligence
THEMA-30, rue Mégevand
25030 Besançon cedex

Summary :

Using of Epragma system in the E&T project. The digitization of the different filled questionnaires from young students in the East and South-east of France and methods.

Résumé :

Utilisation du système Epragma dans le cadre du projet E&T. Numérisation des différents questionnaires saisis des écoliers de l'est et sud-est de la France et méthodes.

Keywords :

Epragma, E&T, academic success, Youth, methodology, databases

Mots clés :

Epragma, E&T, succès académique, jeunesse, méthodologie, bases de données



1. Introduction

The project of the Observatory of Education and Territories (E&T) has started in 1999, it focused on better knowledge of rural students in their specific geographical, social, family and school. The research issues are in link with the issues of academic success, project life and employability of young people in environments that are often characterized by cultural isolation. The limited range of environmental and socio small school infrastructure. It represents the impact that the territory of life can have on the future of students.

In despite of numerous regulatory actions carried out under policies of land use since several decades in France, the geographical area of education and the school is far from homogeneous.

There are significant differences inter-academic in all areas of school life, both in terms of infrastructure (network facilities, equipment ...) that educational support (by staffing level, average age of teachers ...) to student success rate (repetition, pass exams ...), offer training (specialties of technical and professional ...) or simply the desire of families (enrollment in the various sectors or university ...).

These differences inter-academic or inter-regional, often much stronger than might assume, is added gap between urban and rural areas. Places and school environments are not neutral, because life class, working students and teachers who are in unequal conditions, fall in economic, social and cultural rights.

Issues arising from these observations have led researchers to conduct a followed cohort of students belonging to various types of rural and mountain of CM2¹ in the final. The data collected through surveys have been grouped in a large database.

2. The monitoring of the cohort of students attending colleges and rural mountain

Monitoring this cohort of 2400 students at the Second CM2 belonging to various rurality of six departments of the East and South-east France (Ain, Alpes de Haute Provence, Ardèche, Drôme, Haute-Loire and Haute-Saone) allows you to examine how the project is built professional college, how it focuses on the family and the local economy, how it affects the investment school students. This questioning is particularly in the opening strategies implemented by families, schools, local governments or associations to address situations of isolation or to enhance the image of local environments.:

- Regular use of TIC/TICE²,
- Working in network
- and building partnerships with outside organizations to ministry of national education.

The panel includes an equal number of students from three rural typologies defined by INSEE / INRA 1998:

- rural centers, offering jobs 2000 to 5000 and whose number of jobs offered is greater or equal than the number of active residents and their peripheries of which over 20% or more works in the rural center
- rural, low urban influence which over 20% or more of active residents working in an urban area,
- isolated rural areas where less than 20% of active residents working in a center or a rural urban center

It also includes an equal split between "mountain" and "out of area mountain" as defined in 1985, completed in 1997 on the altitude and slope.

2.1. The criteria taken into account

To learn more about the journey of each student, the reasons for its choices and its investment school, it was collecting information on :

- the student himself,
- his family environment,
- operation of his school.

Information on human factors are quantitative and qualitative. It was important to know a number of "representations" that students and their families are doing their territory, school and professional integration.

The body contains twelve groups of information, information relating to students:

- gender
- the municipality of residence and (the) town (s) of Education
- the length of school transport

¹ Middle Course second year

² Information Technology and Communication for Education



- opinion about the school and the environment where they live,
- opinion on school subjects
- Professional projects
- draft curriculum
- knowledge of trades
- the process of determining policy choices
- recreation
- the use of TIC
- test scores and reviews (Tests assessment Sixth Patent and Colleges), the guidelines (Repetition, fate after the Third Class)

He has eight sets of information to know the family environment:

- occupation of parents
- the educational level of parents
- Family Size
- opinion on the type of school attended by children
- opinion on the level of school children
- wishes to guidance for the child
- the process of determining policy choices
- desired occupations for children

Finally, the body of information about each student includes information on schools:

- the size of the structure
- strategies to break isolation
- the level of availability of TICE for students
- information measures in place for orientation

2.2. Scientific methodology

A database containing all this information has been formed, enriched over the years. To do this, the methodology and tools used: CATALYSIS were used with software Pragma, Anaconda, Nuage, to form the database (approximately 20 000 questionnaires before), make quantitative statistical treatment (flat sort, crossed sort, balance sheet, extractions of sub-bases ...) and qualitative (CFA ... CAH). The information system territorial Sitra, designed in partnership with the Laboratory for Computer Science at the University of Franche-Comte is a potential that has not yet been used ...

The information comes from five sources:

- questionnaires to students,
- questionnaires sent to their parents,
- questionnaires to schools,
- information requested in Rectorats (notes for tests and evaluation results of the patent colleges, where enrollment of students "lost")
- few interviews.

Some of these information are confidential, it was necessary to obtain authorization from the CNIL to establish these bases and monitoring.

3. Data Bases

Data collection was therefore spread over several years over the raising, as follows:

- 1999 (year coded "n"): students in CM2 (the last year in elementary school)
- 2000 (year "n+ 1") collection of scores of assessment tests in the 6th,
- 2001 (year "n+ 2"): students who follow a curriculum in the 5th,
- 2003 (year "n+ 4"): students who follow a curriculum in the 3rd,
- 2004 (year "n+5"): students who follow a curriculum in Second,
- 2006 (year "n+7"): students who follow a curriculum are final.



The latest arrest has been delayed because of the difficulty of matching some departments to access files of their rectors and their directions of agricultural education and the respective institutions. It is nearing completed. In total, since the beginning, more than 20 000 questionnaires were filled.

During each campaign to collect data, a series of additional information had completed those already collected for each student. There is therefore a personalized for each student.

But during the time of collect, the number of students has decreased, in some important, due to several factors:

Established change without the Principal of the College retains the new assignment (usually start to a private school)

- relocation of the family and change of establishment,
- leaving the school system during cycle (learning, FSC) or at the end of the cycle,
- failure of the monitoring team of students,
- unwillingness of heads of institutions ...

This decrease was expected. Checks indicate maintains the balance between boys and girls between the types of territories. It was feared that subjectively "evaporation" particularly affects students experiencing academic difficulties. But a study on "lost sight of students" just to prove that they do not affect the overall results, because people still remaining representative of the initial contingent.

Furthermore, the current collection made from various statistics Rectorats, the Chamber of Agriculture for agricultural and private sector, will significantly raise the number of individuals of the "n+ 7".

The state of digital lines is as follows:

- Initial Base: 2394 students (50 descriptors)
- Base "n+1": 2241 students (54 descriptors)
- Base 'n +2 »: 1828 students (106 descriptors)
- Base 'n+ 4: 1365 students (164 descriptors)
- Base 'n+ 5: 1084 students (203 descriptors)
- Base "n+ 7" being formed (204 descriptors)

Thus, the study of each base is relevant. Over the years, the mass of information about each student is enriched with its progress in the school curriculum and its maturation over its projects, but the number of students decreases.

3.1. Methodological limitations and e-Pragma solution

The data is made through resources that allow Pragma flat sort, crossed sort, the extraction of "values", ie words or numerical values introduced by the open questions, the establishment of sub-bases ...

The chain of treatment made by Anaconda and Nuage allows qualitative analysis (CFA ... CAH). The set is built according to the methodology CATALYSIS.

3.2. Difficulties in managing the geographical distance.

The methodology involves CATALYSE diversity of users of the database (Pragma) working on networking sites. Team members E&T are spread over five departments that stretch of the "Haute-Saone" to the "Alpes of Haute-Provence". Within each department are also raises the issue of communication between researchers and actors (teachers) and the shared use of tools.

Seizures require the establishment of sub-bases for the teams that can work in conjunction and consultation in real time is impossible. The transmission of files is sometimes random because of their weight: some refuse messaging transmission. The same is true of course, the final base obtained after aggregation of departmental bases...

In short, teamwork is hampered by distance. The same is true for even consulting the database by the players, or by other researchers who want to use data from the E&T.

4. E-Pragma

The development of Electronic Pragma (e-Pragma) began in 2004 with a project SITRA (Territorial Information System Network and Actors). It integrates the logic of operation of Pragma classic and adds new features and design more innovative. It is used by several major projects like the ACCEM, Cocagne, etc ... It has moved in version RC (Release Candidate) from the project SGAR (representing Europe in Franche-Comte) of 2007.

Developed by the Territorial Intelligence Team (ThéMA-CNRS), available from any computer connected to the



Internet, ePragma allowed in a first time to carry out the fill of networked questionnaires. Now it is able to perform basic operations: scanning the questionnaire, data entry, flat sorts, exports and imports.

The success of this experiment led the team to develop a more efficient and faster which is the E&T

4.1. The new technical solutions

The system e-pragma currently being developed for the E&T uses PHP: Hypertext Preprocessor and MySQL for its database system. The interest of this tool is is enabling the team to function network, centralizes and provides access from any location.

The encoding data of individuals and items using standard questionnaires caENTI:

- 2 characters for descriptors,
- 4 characters for the procedures,
- 12 characters for individuals: the code of each individual has 3 characters for the group, 7 for the main number and 2 for the version number..

4.2. Features

Based on the classic Pragma, the online version e-Pragma brings other advantages as the centralization of information in only one space (database) for network-based users. New features were added.

The project management system. Any investigation is the result of a project that combines a number of partners at different levels of responsibility and separates tasks. The system of project management e-Pragma is essential for a structure of a network team whose members are working on different geographic locations. The hierarchy of permissions and therefore the roles of each partner involved in the project allows for more sophisticated management.

The different categories of users have a specific function, greater or lesser extent depending on their precise role in the project: consultation, seizure, creation and modification of the questionnaire, and overall management of each questionnaire of the E&T. This prioritization is not fixed and can be adjusted according to the amendment of the team. The hierarchy of responsibilities and tasks includes 5 levels that open up "rights":

- The super administrator who is responsible for managing the overall system and which allocates permissions. He is the master of all projects, partnerships, questionnaires, the public investigated. He is unique in the system.
- The project manager manages the only system on the project under his care.
- The administrator questionnaire: Such a user runs a questionnaire with questions and answers arrangements that will provide comments to the database.
- The member functions such user is limited to the seizure of individuals who responded to a questionnaire given.
- Guest: it has the ability to view data forming the base.

The search engine. The databases of E&T include a number of important issues and individuals that evolve over the



campaigns to collect information. The initial base formed when students were in elementary school end of 50 descriptors and 2394 individuals. The following bases include new information, but lose people. The base "n +5" formed when students have achieved a course without delay in the second class of 203 students descriptors and 1084. In this archive, the searches are difficult manually: a search engine is implemented for this purpose to facilitate this.

The visualization of data. One way of viewing data without opportunities to change, is planned. This mode uses the same interface for entering or modifying data, but is secure. This gives us an overview of the responses of an individual.

4.3. Ergonomics

To facilitate the use of the system by non-specialist users, unfamiliar to computers or used to the classic version of Pragma, the system retains the same logic interface with a framed by original version.

While retaining the same logic, the new version improves the clarity of screens or consultation, and the navigation system. The technical vocabulary remains the same as not to confuse users who worked on the classic version of Pragma.

Data migration Data from the E&T are contained in classical bases, five bases and established a base being seized. It is therefore necessary to organize the migration of these data in ePragma. To do this, and before the transfer, it was re CAENTI according to the standard all questionnaires (2 characters for descriptors, 4 characters for procedures) and individuals (12 characters). The recoding of data is done manually and individuals is automated.

The import in e-Pragma has been updated to import data from Pragma classic, while recode individuals.

Figure 1: The top screen displays a list of issues

5. Futur works : Towards a European school

The E&T will use the technology usefully ePragma for:

- allow players to make the recent seizures online (creation of the Base "n +7")
- enable researchers to question this in real time to achieve the initial analysis. In addition, they can also access previous bases to carry out the necessary checks and setbacks. (see description of the 6 bases in paragraph 3)
- access to all data to the writing of Volume 5 of the series "The education and rural montagnard" (University Press franc-comtoises (see presentation in paragraph 4)

The Observatory European project of the school (WP6E) led by Yves Alpe will also use technology ePragma:

- to access the foundations of the E&T in the educational objective to familiarize the European partners of the project with the methodology CATALYSE
- to build their own databases

6. Conclusion

The use of ePragma by the team of the E&T constitutes an important step whose main beneficiaries are:

- Actors: the attachment allows online comfort and time saving, increased security and increased quality scientific data, it also facilitates their consultation, harnessed to improve their professional practice.
- Researchers: data are accessible in real time; the collection have increased security and the rigor and the scientific reliability of these data is being increased

In addition, E&T has the opportunity to initiate the consultation to outside teams interested in exploiting the data capitalized during these seven years.

Finally this tool will be useful for the implementation of the European school (WP6E)

7. References

ALPE (Y.), CHAMPOLLION (P.), FROMAJOUX (R.-C.), POIREY (J-L.), (coord.), (2001), L'enseignement scolaire en milieu rural et montagnard. Tome 1, Espaces ruraux et réussite scolaire, Besançon : Presses Universitaires Franc-comtoises.

ALPE (Y.), POIREY (J-L.), (coord.), (2002), L'Enseignement scolaire en milieu rural et montagnard. Tome 2, Au Seuil du collège, Besançon : Presses universitaires franc-comtoises.



ALPE (Y.), CHAMPOLLION (P.), POIREY(J-L.), (coord.), (2005), L'enseignement scolaire en milieu rural et montagnard. Tome 3, Collégiens à mi-parcours, Besançon : Presses Universitaires Franc-comtoises.

BERTACCHINI Y., GIRARDOT J-J., GRAMACCIA G., 2006. De l'intelligence territoriale : Théorie, Posture, Hypothèses, Définitions. In : TIC et Territoire : quels développements ?. Actes du 5ème colloque, Besançon, 9-10 Juin 2006. Besançon : ISDM, n° 26 [en ligne]. Juin 2006 [réf. du 30 mars 2007], 7 p. Disponible sur : http://isdm.univ-tln.fr/articles/num_archives.htm#isdm26

GIRARDOT, J.-J., 2006. Integration of data analysis tools within territorial intelligence systems. Communication invitée présentée dans le cadre de la conférence de Soochow University, Taipei (Taiwan), 1er Decembre 2006.

GIRARDOT J.-J., 2005. Concepts, principes et outils de la méthode CATALYSE. 3ème colloque international du Réseau Européen d'Intelligence Territoriale, Liège, 20- 22 Septembre 2005.

GIRARDOT J.-J., 2004. Intelligence territoriale et participation [en ligne]. In : Tic et Territoire. Actes du 3ème colloque, Lille, mai 2004. Lille : ISDM, n° 16 [en ligne]. Mai 2004 [réf. du 30 mars 2007], 13 p. Disponible sur: http://isdm.univ-tln.fr/articles/num_archives.htm#isdm16

MySQL 5.0 Reference Manual

