Informatics in Education, 2016, Vol. 15, No. 1, 63–84

DOI: 10.15388/infedu.2016.04

MEMORE: an Environment for Data Collection and Analysis on the Use of Computers in Education

Ronaldo GOLDSCHMIDT[1], Isabel FERNANDES de SOUZA[2], Monica NORRIS[3], Claudio PASSOS[4], Claudia FERLIN[5], Maria Claudia CAVALCANTI[1], Jorge SOARES[6]

[1] *Military Institute of Engineering (IME) – Brazil*
[2] *Rural Federal University of Rio de Janeiro (UFRRJ) – Brazil*
[3] *Education Office of Piraí – Brazil*
[4] *Pedro II School – Brazil*
[5] *Faculty of Technological Education of Rio de Janeiro – Brazil*
[6] *University of the State of Rio de Janeiro – Brazil*
e-mail: {ronaldo.rgold, isabelfs.ufrrj, mncnorris, cpassos.cp2, claudia.ferlin, maryoko, jasoares}@gmail.com

Received: September 2015

Abstract. The use of computers as teaching and learning tools plays a particularly important role in modern society. Within this scenario, Brazil launched its own version of the 'One Laptop per Child' (OLPC) program, and this initiative, termed PROUCA, has already distributed hundreds of low-cost laptops for educational purposes in many Brazilian schools. However in spite of the numerous studies conducted in the country since PROUCA was launched, Brazil shows a lack of proficiency in basic information crucial for managing and improving any OLPC initiative (e.g., number of effectively used laptops, use time and distribution per subject, use location and school performance of users, and others). Therefore, the focus of this article is to introduce MEMORE, a computational environment for longitudinal on-line data collection, integration and an analysis of how PROUCA laptops are used by schools. Technical details about MEMORE's architecture, database and functional models are supplied and the results from real data collected from Brazilian public schools are presented and analyzed. They elucidate how MEMORE can be a valuable management tool in OLPC contexts.

Keywords: elementary education, architectures for educational technology system, country-specific developments, media in education.

1. Introduction

The rapid innovations and advances in Information and Communication Technologies (ICT) have provided different computational resources with practical and strong applications in several segments of society (Castells, 2009; Laurindo, 2008). Computers used as educational tools are particularly relevant and current themes in multidisciplinary research, especially in the areas of education and computer science (Tajra, 2008).

1.1. *Background of the Study*

Inspired by the "One Laptop per Child" (OLPC) initiative of MIT's Media Lab (Massachusetts Institute of Technology), several countries have implemented programs geared towards distributing low-cost computers in schools. By and large, most programs are motivated by the possibility that the students' use of ICT will lead to learning improvements and academic productivity (Hansen *et al.*, 2012).

It was within this context that the Brazilian federal government initially launched the "One Computer per Student" project – *Um Computador por Aluno* (UCA) – which was followed by 'One Computer per Student Program' – *Programa Um Computador por Aluno* (PROUCA) – both inspired by the OLPC initiative (CNPq, 2010). While the UCA project directly distributed laptops to 300 schools throughout the country, PROUCA has provided special financing conditions for purchasing laptops to the local governments interested in participating in the OLPC proposal and implementing the use of laptops in their schools. For the sake of simplification, except when both Brazilian OLPC modalities should be differentiated, in this paper they will be interchangeably referred to as PROUCA.

Ever since the launching of PROUCA in 2007, several projects, research, experiences and experiments have been conducted by universities, schools, enterprises and government bodies in Brazil (Coelho *et al.*, 2010; Schneider *et al.*, 2011; Fonseca, 2011; Meneses, 2011; Siqueira and Costa, 2012; Sampaio and Miranda, 2012). Substantial information about the program and its experiences has been and continues to be produced in several parts of the country (Sampaio and Elia, 2012).

Some examples typifying the outcomes from initiatives linked to PROUCA include: statements, reports, summaries, articles, interviews, books, theses and dissertations, lectures and banners (MEC, 2013). The formats/media wherein these products are presented vary as well, such as texts, videos, images, presentations, spread sheets, charts, tables and diagrams. Moreover, the communication channels to disseminate PROUCA production are also diversified, namely blogs, forums, chats, wikis, social networks, scientific databases and websites, and others.

The decision-making process in terms of the large amounts of data (mainly in textual format) is not an easy task for the administrators (Weiss *et al.*, 2005), as it requires identifying valid, relevant, recurrent and potentially useful information in the context under study (Feldman and Sanger, 2007; Jurafsky and Martin, 2007). In the case of the Brazilian OLPC, the complexity of the process is even greater due to the fact that the existing information is spread over different formats and storage sites.

The Brazilian government lacks consolidated information concerning common pedagogical practices that take into account the use of laptops in several schools (CNPq, 2010). However in spite of the numerous studies conducted in the country since PROUCA was launched, Brazil lacks some basic information, such as how many computers are being effectively used in each school, the average use time laptops are used by the students in each school, how time is distributed per subject, where the computers are being used and the school performance of the students benefited from the program, among additional important basic information necessary for the administration and improvement of any OLPC initiative.

Nugroho and Longsdale (2010) carried out a comprehensive literature review to assess OLPC programs throughout the world. The outcome observed that several of the programs implemented in other countries face problems similar to those of PROUCA, namely the same lack of basic data regarding their development. Most OLPC initiatives, including PROUCA, have concentrated the evaluation of their programs on specific case studies (Nugroho and Longsdale, 2010). Questionnaires are commonly applied as data collection instruments for cases under study. The questionnaires applied consist of open and/or close-ended questions. The answers to open-ended questions require more processing and analysis effort as they entail interpreting the written texts in natural language.

Some types of questions are more recurrent, as for example: the frequency level used for certain software applications, opinions on how much the student has learned about a particular theme, how appealing are the classes, and others (Kozma *et al.*, 2004). Nevertheless, the contents of most of the questions vary according to the purpose of the evaluation (Hansen *et al.*, 2012). In several situations the data collected were limited to textual reports about the experiences undergone by the beneficiaries of the programs (Nugroho and Longsdale, 2010).

Another characteristic that most OLPC programs seem to share regards the time the questionnaires are applied. The data collection is generally conducted at some specific moment, usually at a time unrelated to daily computer tasks (Nugroho and Longsdale, 2010). The student has to provide answers about something he accomplished in the past. In these situations, the accuracy of the information collected depends on the student's memory, among others aspects. And the longer the time span between using the laptop and filling out of the questionnaire, the more the information accuracy will depend on the interviewee's memory capacity.

Selecting the study groups that will fill out the evaluation questionnaires is customarily performed in a timely and separate manner by many OLPC programs (Penuel, 2006), thus rendering unfeasible a longitudinal follow-up of most of the cases studied.

Except for the electronic spread sheets and web surveys, the existence of other computational tools to support the data collection process were not identified in any of the programs studied by Nugroho and Longsdale (2010). One drawback in the use of tools such as spread sheets and surveys is that they are restricted to portraying the data collected at a specific moment. Integrating these data with the data collected at other moments, indispensable for a viable longitudinal follow-up, is not a simple task and requires technical computing knowledge.

Another aspect discussed in Hansen *et al.* (2012) refers to the difficulty to compare studies that assess how OLPC programs influence the school context. The different investigation factors, objectives, the accuracy and the methodological approach used in each study have contributed to render a comparative analysis unfeasible. It should also be emphasized that, to the extent of our observation, there are few works that make available the original data used in the evaluation studies of OLPC programs. This information gap also helps to make difficult any intention for setting a parallel between the research conducted in the area.

1.2. *MEMORE Environment – Overview*

The objective of this paper is to present a computational environment called MEMORE and the results of its implementation. This environment was created to provide teachers and educational supervisors updated, integrated and continuous information on the use of computers with regards to PROUCA. Accordingly, this environment collects, integrates, consolidates and relays information on the use of laptops by the program's beneficiaries. The information provided by MEMORE includes the time frame the laptops are used per subject, per software, per location and type of activity developed. This and other types of information can be displayed separate or integrated, taking into consideration different geographic contexts (schools, cities, states and country) and time frames.

MEMORE operates in six distributed key modules that complement each other:

(a) Data Collection Module: based on the concept of purely reactive software agents (Russel and Norvig, 2010) which captures real-time information on how each laptop is used by the user. It autonomously identifies the applications accessed, use time and duration, and sites visited; it interacts in real time with the laptop user to obtain more details about the time period used, such as the location and the subject studied; it reduces the user's responsibility to remember and provide precise information regarding the previous times it was used. Once installed in a laptop, the data collection agent follows the user continuously in the successive computer use sessions.

(b) Data Transfer Module: transports the collected data in each laptop of each school to a single repository called 'analysis center'. By integrating and centralizing the information in a single database, this module feeds an appropriate structure for subsequent longitudinal studies in different geographic perspectives.

(c) Data Analysis Module: based on well-known data mining techniques (Witten and Frank, 2005), it enables to consult different studies on the information collected and stored in the analysis center and provides resources to process structured data and data in textual format.

(d) Configuration Module: stands out by providing educational administrators flexibility to formulate and apply the questionnaires. The objective of the questionnaire, the content and type of each question (open/closed) as well as the target audience and the study's period of application should be defined by the administrator during the construction of the tool. Examples of questionnaires already incorporated into MEMORE include reviewing the socioeconomic profile of teachers and students and the schools' operational infrastructure report. This module also enables to capture information on school performance, student absentee rates, among others.

(e) Visualization of Results Module: enables teachers, students, educational leaders and authorized visitors to visualize statistics of the collected data and access the results generated by the data analysis module.

(f) Beneficiary Support Module: enables teachers, students and directors to answer previously formulated questionnaires in the system by educational leaders. It also enables the beneficiaries of the program to continuously inform their textual reports using a logbook individualized for each user of the environment.

As conceived, MEMORE enables an ongoing follow-up of the entire school community benefited from PROUCA laptops, therefore exempts population sample procedures.

However, two equally relevant points should be emphasized, both related to MEMORE database project. The data model conceived during the construction of the environment allows the gradual inclusion of schools linked to PROUCA, for those interested in knowing how their computational resources linked to the program are being used. On the other hand, this model can store detailed information on the program. Such information is available so that educational leaders, researchers, sponsors and other interested parties are free to analyze the program data according to several interests and levels of details over time. Although ethic issues restrain the overall available data, we believe this can be a path for constructing a common database to enable future comparisons among the studies performed.

2. Method

This section describes the methodology used to develop MEMORE, the system with its main functional requirements, the conceptual data model and the details about the prototype built.

2.1. *Project Development*

MEMORE environment, stemming from a collective construction experience that includes universities, government bodies and public schools, emerged from the Brazilian federal government's interest in knowing the particulars of the pedagogical practices undertaken with regards to the use of laptops in several PROUCA schools (CNPq, 2010). The project proposal was initially conceived in partnership with a research group comprised of researchers from five Brazilian universities.

The starting point of an information system requires an initial overview of the pre-requisites (Paula Filho, 2005). Based on the general requirements determined by CNPq (2010), it was necessary to look for schools that could support the modeling requirement process, as well as be used as pilot in the tests and validation of MEMORE. The most experienced Brazilian city in terms of PROUCA was Piraí, Rio de Janeiro, which embraced the proposal of the project. The Education Office of Piraí selected two schools to participate in the project. The schools selected had to comply with the two Brazilian OLPC modalities, affording a greater representativeness of the study. Thus, one of the schools selected is linked to the UCA project and the other benefits from the PROUCA initiative. The Education Office of Piraí selected four pilot basic education classes for the project – two first grade classes and two second grade classes.

Four types of meetings to promote the project were held in each school with: (a) the administrative staff; (b) the teaching staff; (c) the students of each pilot class and (d) those responsible for the students in the pilot classes. All meetings included the presentation of the project's goals and the proposed methodology. Specific materials were created with adequate language to each target audience. Moreover, meetings were held with those responsible for the students, requiring their signing a formal term authorizing the students' participation in the project.

As for the technical aspects, the software process used to design MEMORE was Práxis (Paula Filho, 2005). The modeling language used was UML (Bezerra, 2014).

Several meetings to assess and analyze the requirement were held with the Municipal Education Office and the pilot schools. Over one hundred and twenty work meetings were held to discuss the project, which included the use of Brainstorming and Participatory Design techniques (Paula Filho, 2005) during the creation of the system. The agents' data collection interface project with PROUCA beneficiaries was built with the support of Human Machine Interface modeling techniques (Rocha and Baranauskas, 2003; Pressman, 2006). Observation sessions were conducted of classes where the students used the laptops.

The Verification and Validation (V&V) process of MEMORE (Delamaro *et al.*, 2007) was in effect throughout the development cycle of the environment. Software artifact inspection techniques (Pressman, 2006) were continuously employed in the requisites assessment, modeling and system project phases. Several test cases were prepared and implemented in the programming and validation phases of MEMORE. The approval records of the requisites and models of the system were documented in the proceedings of specific meetings for the execution of the V&V process. Moreover, there were meetings to homologate the software and the provision of training courses for the users to use the environment.

The development of the proposed environment was based on the basic principles of object orientation (Shlaer-Mellor Lee, 1996).

MEMORE data modeling was inspired by the attribute identification approach based on cardinal points. This approach is commonly used in the construction of multidimensional systems in decision support processes (Inmon, 2005), which consists of identifying dimensions (attributes) from key questions such as "who", "when", "where" and "when". In the MEMORE context, such questions were adapted to collect information such as laptop use sessions: "who used it", "when was it used", "where was it used" and "what was used".

It should also be highlighted that, due to the multidisciplinary nature of the MEMORE project, the development and implementation of the environment enabled integrating teams from different areas, namely collaborators with computing experience such as researchers, systems analysts, programmers and undergraduate students, and also teachers, administrators and technicians with proven track records in the area of basic education. Those involved benefited from the many opportunities to exchange knowledge and experiences.

The development process of MEMORE resulted in collaboration in the dynamics of the classes involved. Therefore, the work surely exceeded Práxis and during the process the exercise of action-research was observed (Thiollent, 1998).

2.2. *The Modeling Developed*

MEMORE is organized in six functional modules. Fig. 1 depicts a partial view of MEMORE´s conceptual data model with emphasis on the most important data structures of the environment.

MEMORE was designed to meet the following classes of users: (a) beneficiaries of the program, students or teachers with laptops granted by PROUCA; (b) school directors and other educational leaders involved in the program; (c) data analysts professionals qualified to apply techniques and mining algorithms to the collected data and, (c) visitors not officially linked to PROUCA but interested in the practices and outcomes created by the program.

The next section provides details of the functional modules and comments about the conceptual data model.

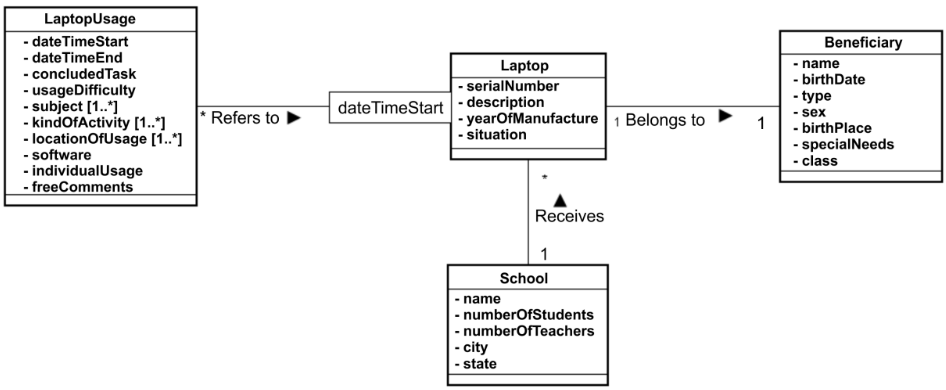


Fig. 1. MEMORE´s conceptual data model: a simplified partial view.

2.2.1. *Data Collection Module*

The data collection module is a purely reactive software agent (Russell and Norvig, 2004) installed in each PROUCA laptop. This agent will obtain and register information in real time from each session the laptop is used by its beneficiary. The use session begins when the laptop is turned on and ends when the computer is turned off.

The data collection agent gathers information during each use session. This agent perceives when each software is turned on, recording the date and hour of this event. Additionally, this agent is also responsible for recording the date and hour when each software is off. To perform its function, the data collection agent needs to know the software list selected by the school that has to be monitored and also be able to follow the events of the operational system in order to detect the initialization and closing situations of the list of selected software.

When the user opens the web browser, the data collection agent identifies and registers all the electronic addresses visited during the user's navigation.

At the end of each session, the data collection agent interacts with the user to enhance information on the resource used. At this time, when it was used for academic purposes, the agent interacts with the beneficiary and asks, for example, about the subjects involved, the type of activity accomplished, the type of work (individual or in group) and the laptop's use location. The user's answers are stored in a log file with the initialization and closing times of each software accessed during the session. The log file is identified by the laptop's serial number.

The interaction content between the agent and the user varies according to the type of user. For example, the set of activities performed by teachers is different from the set of activities performed by the students. As PROUCA provides assistance to all levels of basic education, there are also students in the literacy stage, therefore the interaction between the data collection agent and this set of students needs to be differentiated. To meet this specific group, the data collection agent was designed to use images and sounds as complementary support resources to the text-formulated questions.

Therefore, three data collection agent versions were created. Each version is aimed at a specific audience: literate students, literacy stage students and teachers. Before handing out the laptop to the beneficiary in question, the data collection agent version compatible with the type of receptor is then installed.

2.2.2. *Data Transfer Module*

The data transfer module also consists of a purely reactive software agent installed in each PROUCA laptop. As the name itself suggests, the data transfer agent transports the collected data to a single database center (called 'analysis center').

At the beginning of each use session, the data transfer agent checks if there is internet access availability via laptop. If so, the agent then transports to the analysis center the available data in the log file generated by the collection agent. The transference process occurs unnoticeably by the laptop user.

The amount of data in the analysis center is processed in a chronicled structure to accumulate the information collected during the use sessions. Fig. 1 shows that the data of the use sessions are linked to the respective laptops and can answer questions such as "who used it", "where was it used", "what was used" and "when was it used". The laptop's serial number available in the log file enables to associate the correct information, enabling to associate the use sessions and the respective users.

As the use session chronicles are filled out by MEMORE, the longitudinal studies on the use of laptops become feasible.

2.2.3. *Data Analysis Module*

This module provides well-known pre-processing and data mining algorithms (Witten and Frank, 2005; Goldschmidt *et al.*, 2015). The analyst has autonomy to select and apply as many algorithms as necessary. The application scope available is the MEMORE database of the analysis center.

One of the algorithms implemented in this module is *Apriori* (Goldschmidt *et al.*, 2015), a classical example to perform the association rule mining task. This task finds association rules from the sets of information that appear simultaneously and frequently in a database.

Formally, an association rule is an implication of the form: (R_i) X ==> Y, where: (a) X and Y are non-empty sets of items so that $X \cap Y = \varnothing$; (b) the support of the rule, calculated by $sup(R_i) = |X \text{ and } Y| / |D|$ (the frequency rate that the items of X and Y occur simultaneously in relation to the total available cases in the given set), should be bigger than or equal to a minimal support established by the user (MinSup); (c) the confidence of the rule, expressed by $conf(R_i) = |X \text{ and } Y| / |X|$ (the frequency rate that the items of X and Y occur together in relation to the total of cases in which the items of the antecedent of R_i occur in the data set), should be bigger than or equal to a minimal confidence established by the user (MinConf).

R is an example of an association rule: (R) activities developed individually and using Google Chromium ==> Concluded all the activities and had no difficulty in resources use (sup = 6.9%; conf = 95%). X has two items: students developed activities individually and students used Google Chromium. And so does Y: students concluded all their activities and they had no difficulty in resources use. The support measure indicates that in 6.9% of database records, items in X and Y occur simultaneously. On the other hand, confidence measure states that in 95% of database records items in X occur, items in Y also occur.

2.2.4. *Beneficiary Support Module*

By means of this module, teachers, students and school directors can answer questionnaires.

With the configuration module MEMORE enables the educational administrators to elaborate questionnaires aimed to specific target audiences. By accessing the support module, the beneficiary has access to all the questionnaires that should be answered.

Table 1 shows some of the questionnaires that should be answered by the respective groups of beneficiaries.

Inspired by Fonseca (2011), the report of the school's operational situation enables the school director to describe, by means of a specific questionnaire, the institution's operational infrastructure conditions in relation to PROUCA. Table 2 shows some sample questions to be answered by the director in this report.

Table1
Examples of questionnaires and respective target audiences

Questionnaire	Target audience
Report of the school's operational situation	School Directors
Teacher's socioeconomic profile	Educators / School Directors
Student's socioeconomic profile	Students
PROUCA evaluation by teachers	Educators / School Directors
POUCA evaluation by students	Students

Table 2
Examples of questions of the report on the school operational situation

Subject	Question proposition	Possible Answers
Safety	Does the school have appropriate window and door grids for laptop safety?	Yes, it does / Partially, it does/ No, it does not.
Physical Infrastructure	Does the internet access speed meet the school demand?	Fully / Partially / Poorly / No, it does not.
Training	Did teachers, school directors and technicians get initial training to use the didactic devices received?	Yes / They are under training / No, they did not.
Technical Support	Does the school provide operational technical computing support for the pedagogic activities related to PROUCA (classes, meetings, works, etc.)?	Yes, fully satisfactory / Yes, partially satisfactory / No, it does not.

The questionnaires to describe the socioeconomic profile include questions about the beneficiary's family income, age, type of residence, among others. The questionnaires were elaborated taking into consideration the questions asked in the Brazilian school census, an instrument used by the schools at a national level (INEP, 2012).

The questionnaires to evaluate PROUCA (teacher and student versions) were elaborated based on previous program evaluation studies (Coelho *et al.*, 2010; Fonseca, 2011) and are composed of closed questions and one open question. For the closed questions, the beneficiary has to inform, for example, how he/she evaluates the program with regards to the continued education offered, the laptops' operational conditions and the school's network and internet. The open question allows the beneficiary to freely express him/herself by giving his opinion on the program's strong and weak points.

This module also capacitates teachers, students and administrators to relate everyday experiences in terms of PROUCA. Therefore, they should candidly fill out a logbook describing facts textually and/or expressing opinions about the program and their experiences.

Completing the questionnaires and updating the logbooks help to enrich the database of MEMORE in the analysis center, as this information is then integrated to the data records obtained by the data collection module, and can subsequently be taken into consideration in the analysis and data mining processes.

2.2.5. *Visualization of Results Module*

This module enables the educational leaders, school directors and teachers to access the collected and analyzed data with the support of MEMORE.

The results are presented in reports organized in the following groups:

(a) Data Analysis Results – this group contains the reports generated from the main and most relevant conclusions obtained with the support of the data analysis module.

(b) Managerial consultations – this group contains consolidated statistic data about the main subjects available in MEMORE, such as use of laptops, schools' operational situation in terms of PROUCA; class performance; IDEB (development index of basic education[1]) and socioeconomic profile of teachers and students.

All reports of the managerial consultation group are preceded by filters that allow users to select the desired data subset about the subject in question. According to MEMORE data modeling and the design of multidimensional systems, all the filters were created to enable building search conditions from the cardinal points. For instance, by consulting the use session chronicles, the user can apply filters which combine restrictions of the attributes "who used it", "when was it used", "where was it used" and "what was used". Table 3 broadens this example by illustrating each of the dimensions with some attributes.

Another concern regarded in the elaboration of the reports was to enable a cross analysis between the information, such as a report presentation that includes the laptops' average use time distribution in relation to the subject and the software used. Thus, the operations to group the attributes are built dynamically from the users' choices.

Table 3

Examples of dimensions and attributes used to compose query filters

Dimension	Attribute	Comment
Who used it?	School School Serie Class	Seeks to delimit the use sessions per group(s) of users responsible for this execution.
When was it used?	Initial date Final date	Aims to reduce the use sessions that took place during the informed period.
Where was it used?	Use location	Seeks to limit consultation in the use sessions that occurred in a specific location.
What was used?	Type of software Software	Enables to filter the use sessions by the type of software or the software used during the session.
Used for what?	Subject Type of activity	Aims to select only the use sessions of the subject and the type of activity indicated.

[1] This index measures, in a scale from zero to ten, the quality of each school and each Brazilian education system. It is calculated based on the student's performance in official evaluations carried out nationwide, his/her school attendance and his/her schooling development.

2.2.6. *Configuration Module*

This module is crucial for MEMORE, as through it all the basic information is fed, necessary for the system's operationalization. This includes, for example, registering users and their access rights, the inclusion of schools and the laptops granted to the beneficiaries. It is also responsible for providing the software list to be monitored by the data collection agent in each school. It is the school director's responsibility to keep this list updated, by means of this module.

The configuration module also enables the administrators to formulate questionnaires aimed to specific public audiences, consisting of closed and/or open question. In addition, the design of each questionnaire also requires determining the period in which the tool should be applied. Table 1 contains questionnaire samples built with this module.

The importation of each school's IDEB value is also a functionality of the module to configurate the environment. Also by means of this module, teachers can inform the student's academic performance progress in each subject during the school year, as well as the absentee rate.

2.3. *Implementation*

This subsection gives details about the architecture and technologies used in the implementation of the aforementioned modeling.

2.3.1. *Architecture*

The Brazilian government selected *Classmate* as the laptop model in PROUCA. With a simplified configuration of 512 MB of RAM and 4 GB of HD, the laptops' operational system is *Linux* adapted to *Megoo* system.

The data collection agents and data transference were developed so they could operate with the configuration described. The programming languages C (Schildt, 1996), Python (Summerfield, 2007) and Bash (Costa, 2010) were employed.

The beneficiary support, configuration and results visualization modules were developed in PHP (Dall'Oglio, 2007) to operate in the web and to facilitate the educational administrators' data access. Two servers were configured, one of them for application storage and the other for database storage. Both are under the protection of PRODERJ, a public agency responsible for the ICT in the state of Rio de Janeiro. PostgreSQL was the database management system chosen (Gonzaga, 2007).

Given the high computational cost typical of data mining applications, the data analysis module should operate in a local environment. Its implementation was developed in C language with *PostgredSQL* data and the minimum requirements recommended for the performance platform of this implementation are 2GB of RAM and 20 MB disk.

2.3.2. *Examples of Interfaces*

Fig. 2 and Fig. 3 display partial views of interaction interfaces of the data collection agent with literate students and students in the literacy stage, respectively.

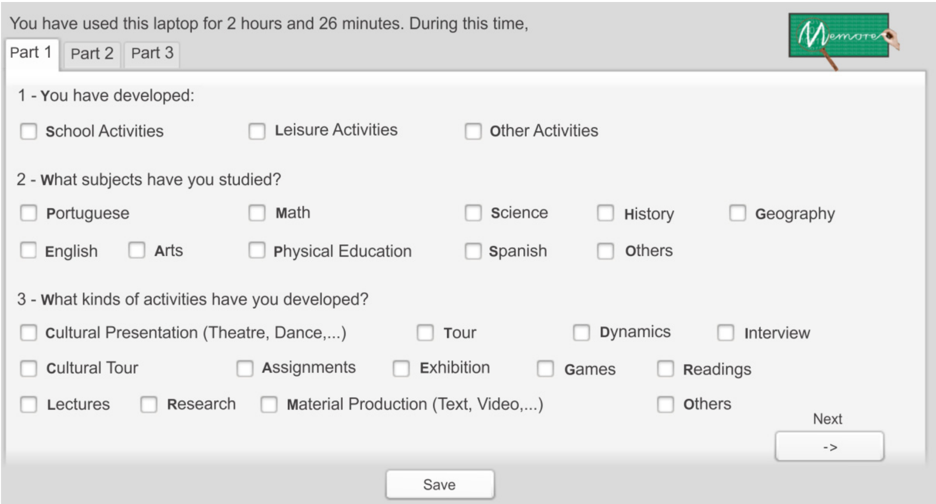


Fig. 2. Interface between data collection agent and literate student.

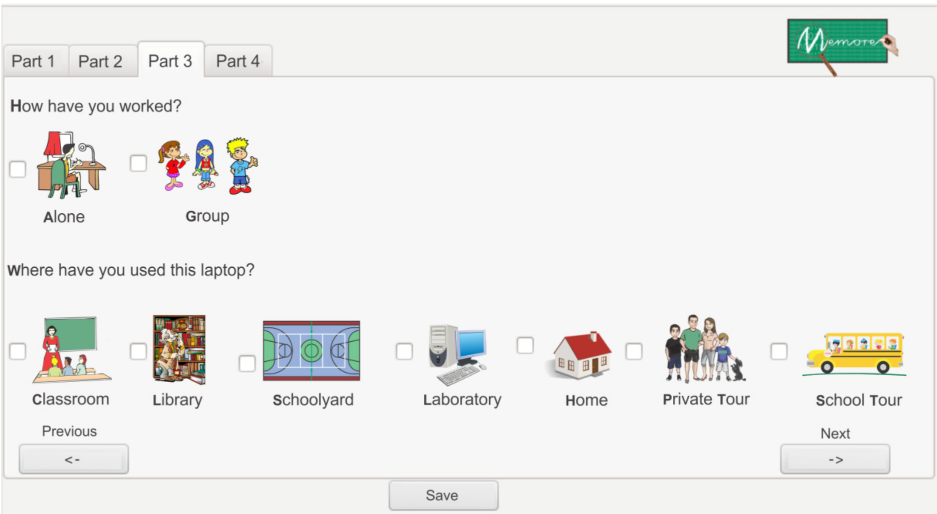


Fig. 3. Interface between data collection agent and student in the literacy stage.

3. Results and Discussion

This section presents and discusses some results produced by MEMORE which depicts the environment's potential to provide the lack of information of Brazilian school leaders on PROUCA development. These results are organized according to aspects that show the main contributions of the system as a management support tool to the OLPC program.

3.1. *General Remarks*

For the experiments described in this section, the data collection and transfer agents were installed in sixty-seven laptops distributed in four pilot classes. All students of the four classes took part in the experiment. The data collection occurred between 1 April and 30 June 2015.

From the beginning of the collection period until the end of May 2015, the computers were kept at school, according to the general rule for all municipal schools. After this date, all the students of the municipality received permission by the Municipal Educational Office to take the laptops home every day. At the end of the school year, the computers were returned to the school for maintenance.

It was seen that during the collection period some laptops in the pilot classes required maintenance, and as the schools had a technical support staff, the maintenance time never exceeded more than two regular school days.

3.2. *Data Integration and Consolidation*

The collected data was transferred, integrated and consolidated in a single database (analysis center). We counted a total of 1,549 hours, 6 minutes and 9 seconds of use in 1,413 use sessions. Based on these data, we estimated an average use of almost two hours per week, per laptop (estimated memory calculation: 1,549/1,413 = 1 hour per session; 1,549 hours/67 = 23 hours per laptop; 23/3 months =7 hours per month = 1.8 hours per week).

Table 4 shows the distribution of laptops, the number of use sessions and the total use time by the four pilot classes. Older students used the laptops longer. The use of laptops by the younger students, mainly in class 201 (average 6 years of age and in the literacy stage), was more restricted to periods when they were helped by the teacher in the classroom.

Interestingly, only 3% of the average laptop use time during the period studied was devoted to group activities. This number rises to 18% when we limited the overall observation to the use sessions performed in the classroom. This variation shows a higher demand for group activities in classroom tasks than in extra-curricular tasks.

Table 4
Distribution of laptops and use time per pilot classes

Class	No. of Laptops/Students	Average Age	No. of Sessions	Total Time (Hours)
201	13	6 years old	120	136:03
402	18	9 years old	201	317:31
701	19	12 years old	634	702:01
801	17	14years old	431	393:13

The software with higher average use time in the period under analysis included Chromium (15 min and 42 sec), Tux Paint (14 min and 41 sec), LibreOffice Writer (12 min and 40 sec) and Gcompris (08 min and 09 sec), respectively. Due to the high internet demand, the most accessed software application was, as expected, the browser Chromium. Also, not surprisingly, the recurrent use for drawing software (Tux Paint) and writing software (LibreOffice Writer) as they are applications frequently associated with academic activities in basic education. Likewise, the popularity of Gcompris also explains the use of the software since it is an educational software package containing highly attractive games for children and young teens.

For 4.4% of the laptop use sessions, the students reported having many doubts on how to manipulate the computational resources. For 16.6% of the cases, they declared having few doubts regarding how to use the computer. No doubts were reported for the other use sessions (79% of the cases). This data indicate the students' good assimilation level with regards to the technology provided by the laptops.

Fig. 4 and Fig. 5 show the distribution of the laptops average use time per subject and location in the period. The high laptop use frequency in school tours reflects the several activities performed in the pilot classes with the local community, nearby the school.

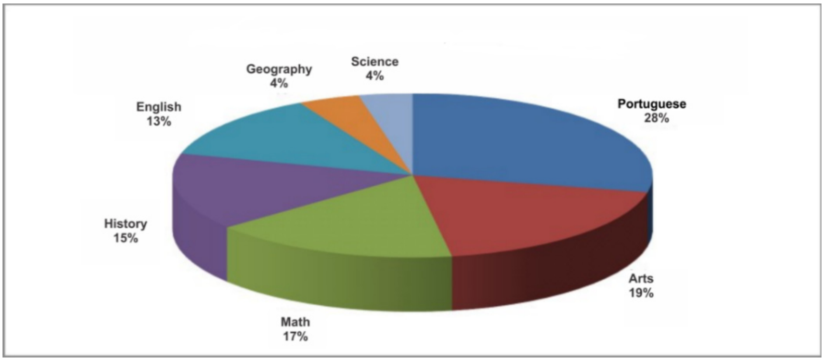


Fig. 4. Average laptop use time: distribution per subject.

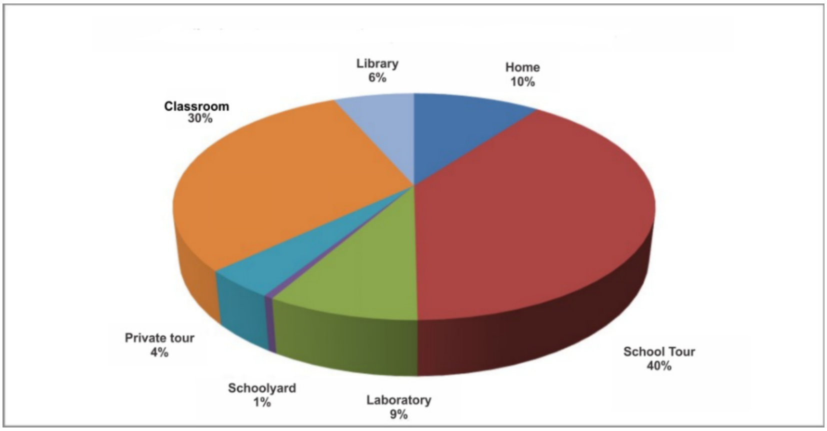


Fig. 5. Average laptop use time: distribution per location.

3.3. *Longitudinal Follow-up*

Fig. 6 illustrates the progress of the students' average performance in the pilot classes, which were evaluated every two months. Although the evaluation interval exceeded the data collection period of the project, the graph represents a longitudinal follow-up instrument for the classes. The averages shown in the figure were taken from the students in Mathematics class.

3.4. *Crossing Information*

The data model of MEMORE enables crossing information through the different dimensions of the problem. As stated earlier, the environment enables the educational leaders to interactively and dynamically select and combine in any amount and order, the attributes of the problem, in order to perform the desired distributions.

Table 5 depicts the laptops' average use time distribution combining two dimensions – "use location" and "type of activity" – developed by the pilot classes of the experiment. Other distribution samples were omitted due to space restrictions.

3.5. *Integrated Formulation of Questionnaires*

The following examples illustrate two reports prepared with the data collected from questionnaires formulated with the support of MEMORE. The environment provides flexibility to construct such type of instruments. The data from the questionnaires are integrated to all background information collected by the system.

The administrators of both pilot schools completed the report of their schools' operational conditions. This was done at the end of the first semester of 2015. Table 6 shows a summary with the main operational aspects identified in the schools studied. The school that required technical computing support and internal pedagogic coordina-

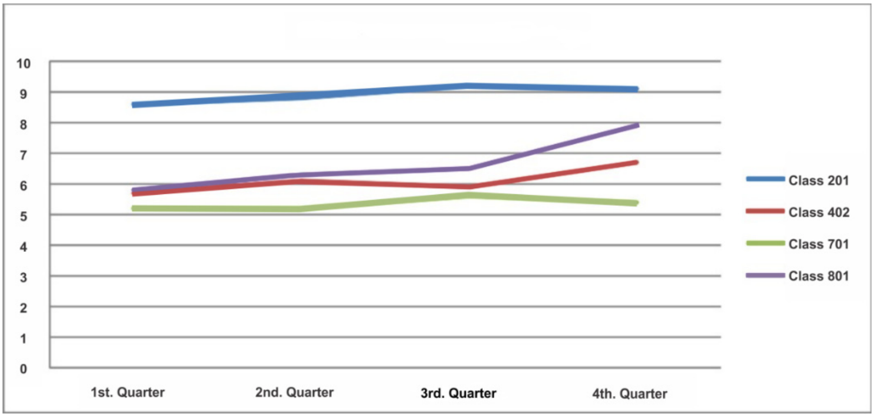


Fig. 6. Progress of average academic performance of the pilot classes in Math.

Table 5

Distribution of average use time regarding location and activity

Use Location	Type of Activity	Average Time
Library	Assignments	00:16:07
Library	Reading	00:32:51
Library	Research	01:18:09
Home	Dynamics	00:33:08
Home	Assignments	07:30:05
Home	Games	05:04:29
Home	Reading	09:39:02
Home	Research	07:58:52
Home	Material Production (Text, Video, etc.)	00:49:42
Laboratory	Assignments	01:17:01
Laboratory	Games	08:43:00
Laboratory	Reading	01:21:19
Laboratory	Assignments	01:55:13
Laboratory	Material Production (Text, Video, etc)	00:56:09
School excursion	Dynamics	00:23:44
School excursion	Assignments	01:37:00
School excursion	Assignments	05:07:15
School excursion	Exhibition	04:16:35
School excursion	Games	02:46:21
School excursion	Reading	03:51:04
School excursion	Research	05:55:12
School excursion	Material Production (Text, Video, etc)	00:56:09
School excursion	Dynamics	00:10:05
School excursion	Games	00:14:41
Classroom	Interview	03:02:46
Classroom	Cultural Tour	01:07:25
Classroom	Assignments	02:00:17
Classroom	Exhibition	01:12:47
Classroom	Games	02:25:05
Classroom	Reading	03:11:46
Classroom	Research	03:07:44
Classroom	Material Production (Text, Video, etc.)	00:48:11

Table 6

Key operational aspects collected in the pilot schools in April 2015

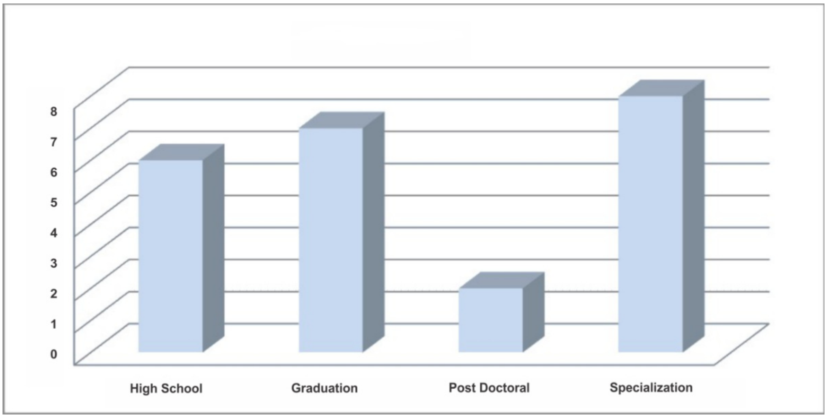
Both schools regarded the safety of the laptops as partially satisfactory.
The rooms to recharge the laptops are appropriate and meet the requirements.
Both schools regarded the internet as stable and with satisfactory speed.
One of the schools needs an internal pedagogical coordinator responsible for PROUCA.
One of the schools needs technical computing support for PROUCA pedagogical activities.
In both schools, the educators, directors and technicians get continuous laptop use training.

tion of the OLPC program belongs to the UCA project. The other school, which belongs to PROUCA, had all its needs provided for.

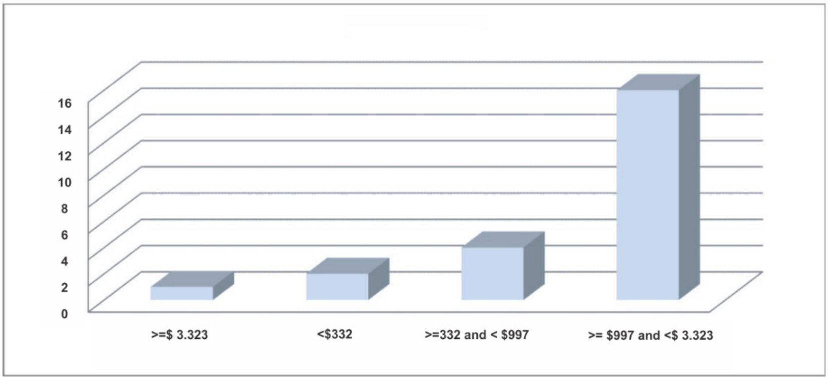
In May 2015, twenty-three teachers of the pilot schools (including the teachers of the pilot classes) answered the MEMORE socioeconomic profile questionnaire. Fig. 7 shows some of the outcomes.

3.3. *Mining Association Rules*

Table 7 displays some examples of association rules created from the application of the *Apriori* algorithm over the data gouped in the analysis center. The 3% and 7% values were used as minimum support and confidence, respectively. Rule R_1, for example, indicates that whenever the students worked alone and used *Chromium* as one of their tools, they completed all of their activities. Rule R_2 shows that all leisure activities during the observed period were developed individually by the students. Rule R_5 shows that 88% of



(a)



(b)

Fig. 7. (a) Characteristics of educators' socioeconomic profiles: Highest Degree.
(b) Characteristics of educators' socioeconomic profiles: Household Income.

Table 7

Association rules generated by *Apriori* (SupMin=3%;ConfMin=70%)

Rule	Sup.	Conf.
(R_1) Worked individually and used Google Chromium ==> Concluded all activities undertaken.	7.3	100
(R_2) Developed leisure activity ==>Performed individually.	6.9	100
(R_3) Developed activities individually and used Google Chromium ==> Concluded all activities and had no difficulty in using the resources.	6.9	95
(R_4) Developed leisure activity ==> No difficulty to use the equipment.	6.5	94
(R_5) Developed Mathematics activity ==>Was able to conclude activity.	3.1	88
(R_6) Developed Science activity ==> Was able to conclude activity.	2.7	86
(R_7) Developed Portuguese activity ==> Was able to conclude activity.	3.4	78
(R_8) Group activity, not used for leisure and developed school activities ==> Concluded all the activities undertaken.	3.4	82
(R_9) Worked at home ==>Concluded all the activities developed.	3.3	96
(R_{10}) Worked at home ==> Developed school activities.	3.5	43
(R_{11}) Worked at home ==> Developed leisure activities.	3.6	46
(R_{12}) Worked in group==> Developed school activities.	3.3	100

the times the students used the laptops to conduct Mathematics-related activities, these activities were concluded. Rules R_{10} and R_{11} suggest a surprising balance indicative between the types of activities performed at home: 46% were leisure activities and 43% were school activities. They contradict the expectations of the educational administrators in the pilot municipality that the laptops would be used at home only for leisure activities.

4. Final Remarks

Along with several developing countries, the Brazilian federal government launched the "One Computer per Student Program" – *Programa Um Computador por Aluno* (PROUCA) – focused on promoting the implementation of laptops for educational purposes in schools as an incentive to the digital inclusion of students, educators and directors (Meneses, 2011). Notwithstanding, the PROUCA initiative lacks the instruments needed to enable the educational leaders to be informed, follow and evaluate the pedagogic actions associated with the use of laptops inside and outside classrooms (CNPq, 2010).

Given this scenario, the objective of this paper was to introduce MEMORE, a computational environment which resulted from the integrated endeavors of the federal government, universities and public schools in pursuit of constructing tools to assist the educational leaders to follow and evaluate PROUCA. This paper describes the development process of MEMORE and some of the environment´s main functionalities. The preliminary experiments are reported as well as their respective results, thus illustrating MEMORE's potential as a management information system in order to provide support to the evaluation and the improvement of educational policies and pedagogic practices within the PROUCA context.

Although initially created as a technological instrument, MEMORE has opened opportunities for developing several scientific research activities in terms of computer science applied to education. Therefore, in addition to MEMORE's natural evolution as a management information system, some potential future works could include: searching for metrics to evaluate the effects of PROUCA considering the guiding principles of the program such as mobility, connectivity and immersion; a comparative analysis of the learning projects planned by educators and the practical effects from the execution of such projects; the construction of a social network to share MEMORE database with schools and with the scientific community as a whole; an investigation to understand the relationship between the use of laptops and the students' academic performance; the application of text mining techniques using the free textual reports collected in MEMORE questionnaires by the beneficiaries of the program; and also a detailed study on the ethical questions related to the use of the data collected with the system.

References

Bezerra, E. (2014). *Princípios de Análise e Projeto de Sistemas com UML*. 3ª ed., Rio de Janeiro: Campus.

Castells, M. (2009). *A Era da Informação: Economia, Sociedade e Cultura.* Coleção A Sociedade em Rede. 12ª. Reimpressão. São Paulo: Paz e Terra.

CNPq/CAPES/SEED-MEC (2010). Edital no. 76/2010. PROUCA Fase 2. Retrieved from. http://memoria.cnpq.br/editais/ct/2010/docs/076.pdf

Costa, D.G. (2010). *Administração de Redes com Scripts: BASH SCRIPT, PYTHON E VBSCRIPT*. Rio de Janeiro: Brasport.

Dall'Oglio, P. (2007). *PHP: Programando com Orientação a Objetos*. São Paulo: Novatec.

Delamaro, M.E., Maldonado J.C., Jino, M. (2007). *Introdução ao Teste de Software*. Rio de Janeiro: Elsevier.

Coelho, F., Jardim, M.H. (2010). Relatório de sistematização III. Guia de implementação, monitoramento e avaliação. Programa UCA. Retrieved from. http://www.uca.gov.br/institucional/downloads/experimentos/DFguiaImplementacao.pdf

Feldman, R., and Sanger, J. (2007). *The Text Mining Handbook.* Cambridge.

Fonseca, A.L.B. (2011). *Projeto UCA – Um Computador por Aluno: Analisando as Condições da Implantação em uma Escola da Rede Pública do Distrito Federal.* Trabalho Final de Curso. Faculdade de Educação/UnB.

Goldschmidt, R., Bezerra, E., Passos, E. (2015). *Data Mining: Conceitos, Técnicas, Algoritmos, Orientações e Aplicações*. Rio de Janeiro: Elsevier.

Gonzaga, J.L. (2007). *Dominando o PostgreSQL*. Rio de Janeiro: CM.

Hansen, N. *et al.* (2012). Laptop use affects abstract reasoning of children in the developing world. *Computers and Education* 59, 989–1000.

INEP (2012). Censo Escolar da Educação Básica Brasileira. Retrieved from. http://portal.inep.gov.br/basica-censo

Inmon, W. (2005). *Building the Data Warehouse.* Indianapolis, Indiana: Wiley Publishing.

Jurafsky, D., Martin, J. (2007). *Speech and language Processing: an Introduction to Natural Language Processing, Computacional Linguistics and Speech Recognition*. 2nd edition. New Jersey: Prentice-Hall.

Kozma, R. *et al.* (2004). Closing the digital divide: evaluation of the world links program. *International Journal of Educational Development* 24, 361–381.

Laurindo, F.J.B. (2008). *Tecnologia da Informação: Planejamento e Gestão de Estratégias.* São Paulo: Atlas.

MEC – Ministério da Educação e Cultura (2013). *Portal do Projeto Um Computador por Aluno*. Retrieved from. http://www.uca.gov.br/institucional/index.jsp

Meneses, S.C.P. (2011). UCA – Um Computador por Aluno: era da inclusão digital. In *Anais do XXII SBIE – XVII WIE*. Aracaju: SBC. Retrieved from. http://www.br ie.org/sbie-wie2011/SBIE-Trilha7/92966_1.pdf

Nugroho, D., Longsdale, M. (2010). Evaluation of OLPC programs globally: a literature review (Version 4). *Australian Council for Education Research*. Retrieved from `http://wiki.laptop.org/images/a/a5/olpc_Lit_Review_v4_Aug2010.pdf`

Paula Filho, W.P. (2005). *Engenharia de Software*. Rio de Janeiro: LTC.

Penuel, W.R. (2006). Implementation and effects of one-to-one computing initiatives. *Journal of Research on Technology in Education* 38, 329–348.

Pressman, R.S. (2006). *Engenharia de Software*. São Paulo: McGraw-Hill.

Rocha, H.V., Baranauskas, M.C. (2003). *Design e Avaliação de Interfaces Humano-Computador*. Campinas: NIED/UNICAMPI.

Russell, S., Norvig, P. (2004). *Inteligência Artificial*. Rio de Janeiro: Campus.

Sampaio, F., Elia, M. (orgs) (2012). *Projeto um Computador Por Aluno: Pesquisas e Perspectivas*. Rio de Janeiro: NCE/UFRJ.

Sampaio, F., and Miranda, L.C. (orgs). (2012). Anais do 18º Workshop de Informática na Escola. *Congresso Brasileiro de Informática na Educação*. Rio de Janeiro: SBC. Retrieved from. `http://www.br-ie.org/pub/index.php/wie/index`

Schildt, H. (1996). *C Completo e Total*. São Paulo: Pearson do Brasil.

Schneider, H.N., Nunes, M.A.S.N., Silveira, I.F., Bittencourt, I.I., Machado, G.J.C. (orgs). (2011). In: *Anais do 22º Simpósio Brasileiro de Informática na Educação*. *Congresso Brasileiro de Informática na Educação*. Rio de Janeiro: SBC. Retrieved from. `http://www.br-ie.org/sbie-wie2011/#`

Shlaer-Mellor Lee, M.M. (1996). *Object Oriented Development: A Manager's Practical Guide*. Retrieved from. `http://www.projtech.com`

Siqueira, S.W.M., and Costa, R.M.E.M. (orgs). (2012). Anais do 23º Simpósio Brasileiro de Informática na Educação. *Congresso Brasileiro de Informática na Educação*. Rio de Janeiro: SBC. Retrieved from. `http://www.br-ie.org/pub/index.php/sbie/index`

Summerfield, M. (2007). *Rapid GUI programming with Python and Qt: the definitive guide to PyQt programming*. New Jersey: Prentice-Hall.

Tajra, S.F. (2008). *Informática na educação*. São Paulo: Érica.

Thiollent, M.J.M. (1998). Metodologia da pesquisa-ação. São Paulo: Cortez.

Trucano, M. (2005). Knowledge Maps: ICTs in education. Washington, DC: infoDev/WorldBank. Retrieved from. `http://www.infodev.org/en/Publication.154.html`

Weiss, S., Indurkhya, N., Zhang, T., and Damerau, F. (2005). *Text mining: predictive methods for analyzing unstructured information*. New York: Springer.

Witten, I., and Frank, E. (2005). *Data mining: practical machine learning tools and techniques*. San Francisco: Morgan Kaufmann.

R. Goldschmidt is an associate professor at the Computer Engineering Department of the Military Institute of Engineering (IME), Rio de Janeiro, Brazil. He received the D.Sc. degree in electrical engineering from the Pontificia Catholic University of Rio de Janeiro (PUC-Rio), in 2004. Currently his main research interests are in artificial intelligence, data mining, decision support systems and formal languages.

I. Fernandes, Ph.D., is currently working as a professor at engineering department in a Community College located in South of Brazil. She teaches technologies in education. Her research experiences are in the areas of computer and technologies in education.

M.N. Ribeiro – Graduate at Lingua e Literatura Portuguesa e Inglesa from Fundação Educacional Rosemar Pimentel (1989) and Master's at Educacional Technology from Federal University of Rio de Janeiro (2008). Has experience in language, acting on the following subjects: digital inclusion, science education and science and technology literacy. Works at Education Office of Piraí and University of Volta Redonda, Rio de Janeiro, Brazil.

C. Passos received his master's degree in systems and computing from Instituto Militar de Engenharia (Military Institute of Engineering, IME) (1996) and graduated in mathematics from the college Celso Lisboa (1990). He is currently professor of the school Pedro II and PhD student at UFRJ in complex algorithms. He has experience in computer science, with emphasis on Programming Languages, acting on the following topics: artificial intelligence, computational intelligence, intelligent agents, collaborative systems, virtual reality, image processing and pattern recognition.

C. Ferlin received the D.Sc. degree in computer engineering from the Federal University of Rio de Janeiro, Rio de Janeiro-RJ, Brazil, in 2008. She is currently professor of PUC-Rio (Pontificia Catholic University of Rio de Janeiro) and FAETERJ-Rio (Faculty of Technological Education of Rio de Janeiro). Her current research interests include artificial intelligence, data mining and computer aided education.

M.C. Cavalcanti received the D.Sc. degree in computer engineering from the Federal University of Rio de Janeiro, Rio de Janeiro-RJ, Brazil, in 2003. She has been active in the database research community for more than ten years. Since 2004, she has been a Faculty Member with the Computer Engineering Department, Military Institute of Engineering (IME), Rio de Janeiro, Brazil. In addition, she has been collaborating with the Computational Biology and Systems Doctoral Program of FIOCRUZ, Brazil, since 2008. Her current research interests include scientific workflow and data management on distributed architectures as well as ontology interoperability.

J. Soares received the D.Sc. degree in computer engineering from the Federal University of Rio de Janeiro, Brazil, in 2007. He is currently professor of University of the State of Rio de Janeiro (UERJ), Brazil. His current research interests include data mining and decision support systems.