Pervasive Business Intelligence Platform to Improve the Quality of Decision Process in Primary and Secondary Education – A Portuguese Case Study

Andreia Ferreira, Filipe Portela, Manuel Filipe Santos

Algoritmi Research Centre, University of Minho, Portugal

a58354@alunos.uminho.pt; {cfp, mfs}@dsi.uminho.pt;

Abstract. Business Intelligence (BI) can be seen as a method that gathers information and data from information systems in order to help companies to be more accurate in their decision-making process. Traditionally BI systems were associated with the use of Data Warehouses (DW). The prime purpose of DW is to serve as a repository that stores all the relevant information required for making the correct decision. The necessity to integrate streaming data became crucial with the need to improve the efficiency and effectiveness of the decision process. In primary and secondary education, there is a lack of BI solutions. Due to the schools reality the main purpose of this study is to provide a Pervasive BI solution able to monitoring the schools and student data anywhere and anytime in real-time as well as disseminating the information through ubiquitous devices. The first task consisted in gathering data regarding the different choices made by the student since his enrolment in a certain school year until the end of it. Thereafter a dimensional model was developed in order to be possible building a BI platform. This paper presents the dimensional model, a set of pre-defined indicators, the Pervasive Business Intelligence characteristics and the prototype designed. The main contribution of this study was to offer to the schools a tool that could help them to make accurate decisions in real-time. Data dissemination was achieved through a localized application that can be accessed anywhere and anytime.

Keywords: Business Intelligence, Innovation, Information System Architecture, Education, Pervasiveness.

1. Introduction

Business Intelligence (BI) can be understood as a method aiming to help the companies to take intelligent decisions, by means of data and information collected from diverse information systems. BI systems use the data available in organizations to provide relevant information for decision process. BI systems are traditionally associate to the use of Date Warehouses (DW). They are considered as being a warehouse of data, more concretely an integrated repository that allows the storage of excellent information for the decision-making. With the need to optimize the decision-making process, it is necessary to integrate real-time data. In order to support the existing school software it

was proposed to hold a BI platform with pervasive characteristics (PBI), applied to the area of Education, more specifically, school clusters of elementary and secondary education. The work has directly contributed to solving the initial problem: How does the education system can benefit from the introduction of Business Intelligence techniques? With this solution it is possible to check the information quality, improving business processes management, as well as improving organizations (schools) performance, more specifically, in its decision making. Now and using the PBI, they are able to have inside information at the decision time. With this solution the schools can take decisions and provide more accurate indicators in real-time taking advantages of a continuous data streaming. These information will be always updated (autonomous tasks) and will be in constant change, depending on stakeholders needs.

This article is divided into seven chapters. In chapter two is a framework of Education in Portugal, the BI and Data Warehouse relationship and finally the Pervasive BI. In chapter three is presented the methodology and the tools used to carry out the prototype. In chapter four it is presented the architecture. In chapter five the data model is presented and chapter six presents the indicators adopted and the results obtained. Finally the conclusions are presented, as well as future work.

2.   Background

2.1.   Education in Portugal

The educational in Portugal at a non-graduated level has two levels, Basic and Secondary. Basic is to students with at least of 9 years of education and secondary from 9 to 12 years studying. This work was carried out in a real context using data from a leading company in educational management (they provide the systems present in the schools clusters). With the need to optimize the process of decision and support the existing school software provided by the company it was proposed to hold a BI platform, applied to the Education area. For the initial goals, it was necessary collect data related to various schools in that encompassed from the registration process to its completion within that school year. To do this it was necessary to survey all the requirements, as well as study all the reports and Key Performance Indicators (KPIs) already requested by the government. From there it was prepared a dimensional model able to support the BI platform and answer to the government requirements.

2.2.   Business Intelligence and Data Warehouse

BI can be defined as a set of mathematical models and analysis methodologies that exploit the available data to generate useful information and knowledge for making complex decisions [1]. The main objective of BI is to provide useful knowledge, using tools and methodologies able to enabling them to make timely and effective decisions.

A Data Warehouse (DW) is simply a consolidation of information from a variety of data sources. It is designed to support strategic and tactical decisions. Its main purpose is to provide a coherent picture of the business at a point in time [1]. After a market

analysis, it can be seen that there are some BI solutions in the education area, however without pervasiveness capabilities:

a) QuantumCube

The CIL QuantumCube Education [2] solution is based on the provision of education services and information management defined by a Cloud architecture. This architecture comes to provide users a new technological environment, flexible, scalable and rapid implementation. The Education Portal software solution allows countries, States, counties, Schools and schools group easily to create, implement and link their education portals. This solution is Portuguese and it is applied in the universe including Portugal. This solution provided, also, content management, communication functionality and analysis of the state for teachers, students, parents and education service providers. The BI education, QuantumCube, is a solution designed to support decision-makers from education in their decision-making, planning, analysis and monitoring of projects, based on indicators and reporting, in the context of simple, effective access information and structured.

b) Logi Analytics - Logi Info

In order to streamline processes, Logi Analytics is a computer applications company. Over the years this was the best on BI applications to facilitate implementation, decision and management of processes (Jackson, 2013). This solution came to provide a better overview of all the results, as well as all the decisions taken. This application is not Portuguese so even operates in the US and eventually are more addressed to universities, healthcare and industrial area [3]. It is an extremely agile and flexible technology and currently have clients from various sectors, such as: Financial; Health; Education; Technology Companies; Industries and Government.

2.3. Pervasive Business Intelligence

The Pervasive Business Intelligence Solution (PBI) is a BI solution with pervasiveness characteristics. Some of its features are [4]: real-time, policy, interoperability, scalability, ubiquity, data security and privacy and context awareness. Based on a literature review [5-9] on pervasive computing and business Intelligence, a PBI must include a set of requirements, being the most important:
- Automate the process of knowledge discovery – real time data acquisition and processing.
- Have the information always updated when necessary;
- There is a need to adapt the system to a changing environment and the own characteristics of the user;
- Disclose information indicators dashboards through electronic platforms situated and ubiquitous devices;
- Help make the best decision at the right time, anywhere and at anytime;
- Create an intelligent alert system able to notify administrators about some major changes in the data dashboards.

3. Material and Methods

For the development of this work it was followed Kimball methodology to design the Data Warehouse and to build the PBI prototype.

3.1. Kimball Methodology

Kimball [10] defends a Bottom-Up approach. He defends that, after the most important business indicators, Data Marts are created first. They will provide a more refined view of organizational data. Kimball [11] defines the DW as "a copy of transaction data specifically structured for query and analysis." This organizational modelling focuses on the user's ease of access and provides a high level of performance for the DW. Kimball methodology is known for Dimensional Lifecycle Business and analysed following three forms. One of that defines the project itself: determines the shape of project management and explains which steps the tasks need to follow in order to be able to properly implement a data warehouse (figure 1).
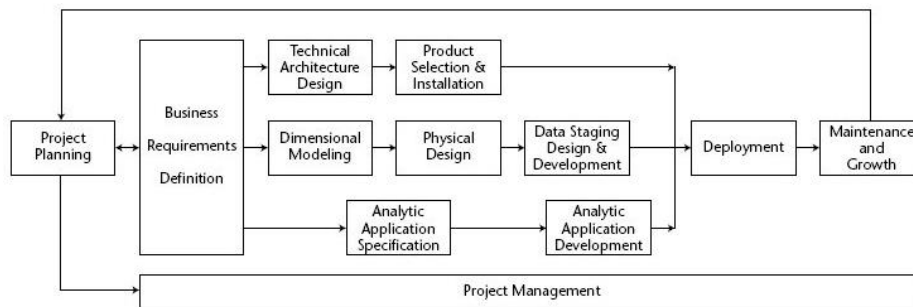

Fig. 1 – Kimball Methodology[11]

As can be seen, the Kimball methodology starts with the Project Planning. For this, it was proposed the Readiness Litmus Tes that fits this project and focuses on the following points: 1) Manager with integrated executive powers in the company that will define the projects objectives, as well as ensure all necessary resources for their development; 2) Business motivation; 3) The partnership between the University of Minho and the company management; 4) The analytical culture of the organization; 5) Projects technical feasibility.

Despite the five points, there are two more important for the success of the project: the analytical culture of the organization against the decision taken and the projects technical feasibility. Regarding the former, if the company's decisions are taken based on the instinct of decision-makers and not on the facts that there is, must be pointed out that a DW will not bring them significant benefits. Otherwise, if these decisions are based on facts presented on the company's activity, then a DW will have a significant weight in the decision-makers choices [11].

3.2. Tools selection

The tools used to implement the PBI were: SQL Server and QlikView. SQL Server was used to make the whole process of extraction, transformation and loading data (ETL) [12]. QlikView [13] was the tool chosen to make the whole BI process. Despite being a tool with particular characteristics, capable of supporting invasive solutions, performs well in terms of real-time dashboards, a high level of security and have a high degree of data security. As a first step it was made the ETL. Then the data was prepared in order to corresponding to indicators made. Finally it was developed the dashboards.

4. Platform Architecture

The main objective of this prototype is to help schools in their decision process, making them easier and containing information in real-time. The architecture (Fig. 2) is divided in five subsystems [14]. Data Acquisition refers to the collection of data by the entities (data source). These data can be in two formats universal (XML) or database (SQL). The ETL process is performed in the data processing subsystem. The tasks associated to data validation and creation of new variables are executed in this subsystem. The Business Intelligence subsystem is responsible to the development of the Data Warehouse and induce data mining models always some new future predictions is requested. This subsystem interacts with the Inference subsystems, where it is all the data mining scenarios and business indicators. Finally it is Interface, which is responsible for the system interaction with the user. This architecture is composed by several intelligent agents which are responsible to automatize the tasks of each subsystem [14]. These agents are responsible for: Automate data collection; Data processing; Data Mining models Inductions; Information retrieval; Ensure the data quality of the PBI system; Present the results in real time.
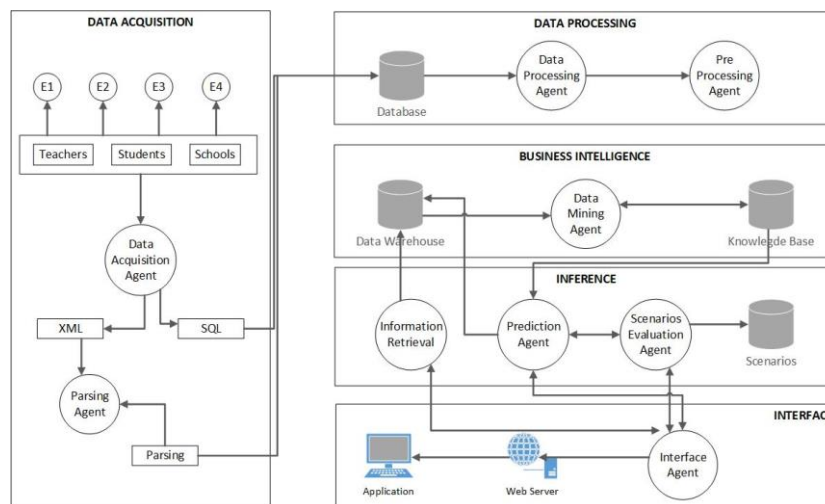

Fig. 2 – Architecture retrieved from [14]

5.   Data Warehouse

It was developed a data model (Fig. 3) allowing the organization to run the entire business process dynamically. The data model consists in a set of tables which support all the school management process. In this model is stored all the information about the students enrolment, as well as their notes. It is also stored all the information on the school management, from the disciplines, courses, student, parents, professors, human resources and others.

This model comprised three facts tables. These tables contains all the KPIs requested by the government that has available information to be produced. In this model is kept all the relevant data for future analysis. There is a need to have three fact tables due the fact of some dimensions are not interchangeable. This happens because there are different records with different information and there are a certain number of records that change throughout the year.

The Registration Evaluation table records the data about the student evaluations and information about the conclusion year. In the Registration table is stored all the information held in the beginning of the year. Teacher Fact table refers to all information associated to teachers.

This DW has pre-defined ETL procedures for an automatic data handling and data loading into the tables. The ETL system extracts data from the source systems, enforcing data quality, standards consistency and combine separate data sources. If appears new data with different patterns it will be necessary optimizing the ETL system by adding or modifying the existing tasks.
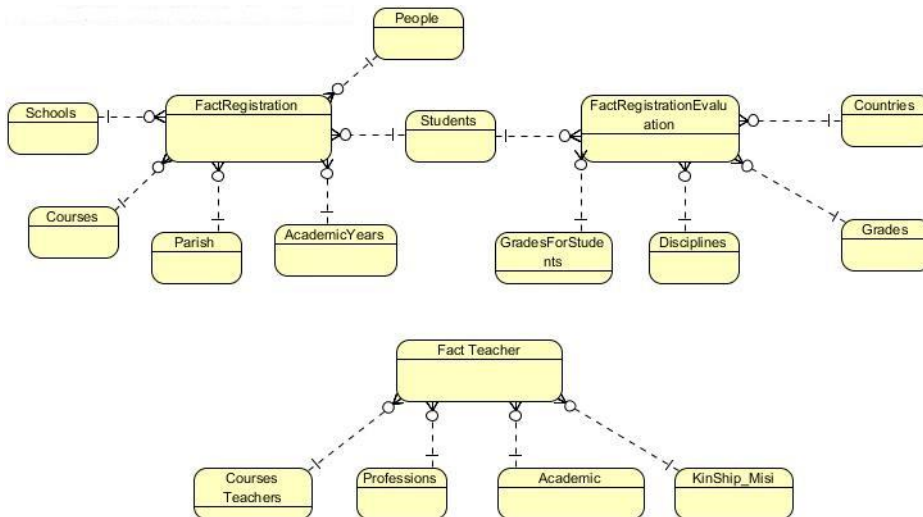

Fig. 3 – Dimensional Model

6. Pervasive Business Intelligence Solution

6.1. Solution Features

The features of the platform took into account the whole process done and englobes a set of pervasive features. Therefore, this platform meets some functional characteristics:

- **Adaptability** - The platform has the ability to automatically optimize models with new data when necessary. This information is obtained from the results;
- **Real-time** - by obtaining the data, this platform is able to present real-time information (e.g. transitions, information concerning the registration, class, household, etc).
- **Security** - Data is stored and securely in the database, so that they can be accessed by whomever they have access privileges.
- **Optimization** - The models are updated over time, so that the platform is always more precise and concise in presenting their results.
- **Accuracy** - It is necessary to have available accurate and reliable data. For this it is essential than the schools be able to providing such data validated.
- **Interoperability** – The platform can integrate data from several data sources and in different formats.
- **Context Awareness** – The system was designed to adapt to a new environment (i.e. new school cluster) and also has user access politics according to the context.
- **Autonomous ETL** – All the ETL process is already defined and it is automatically executed by the agents when a new set of data arrives.

6.2. Indicators

The KPIs are quantifiable measures used to understand whether the objectives are being achieved or not. Consequently, these indicators determine whether to take different attitudes to improve results. Therefore, defined a set of KPIs to make more concise. The indicators reflect more the student's level, that is, everything that encompasses from the stage of registration to the exit of ratings. The DW is prepared to supply the PBI in order to create dashboards and indicators in the following domains:

Domain 1 - Results:
- Academic Success;
- Participation School Life;
- Behaviour and Discipline

Domain 2 - Educational Services provides:
- Teaching practice of monitoring in the classroom.

Domain 3 - Organization and School Management:

- Design, Planning and Development of the activity;
- Human Resource Management:
    - The Financial and Material Resources Management;
    - Participation of parents and other members of the school community;

Domain 4 - Leadership:
- Vision and Strategy;
- Motivation and commitment

Most concretely and taking as example, in the first domain it is possible create indicators to analyse:
- Transitions for School Year;
- Number of students for year, class and subject;
- Students information (behaviour, discipline, results, other);
- Grades by student, subject, school, year;
- Average grades by subject;
- Number of working student (%) and their relation with the final grades;
- Number of absences (students, professors);
- Transitions per school and year;
- Alerts about student performance;

In the Fig. 4 it is possible observe a snapshot of the platform, containing some of the indicators. All these data can be analysed by student, year and school.
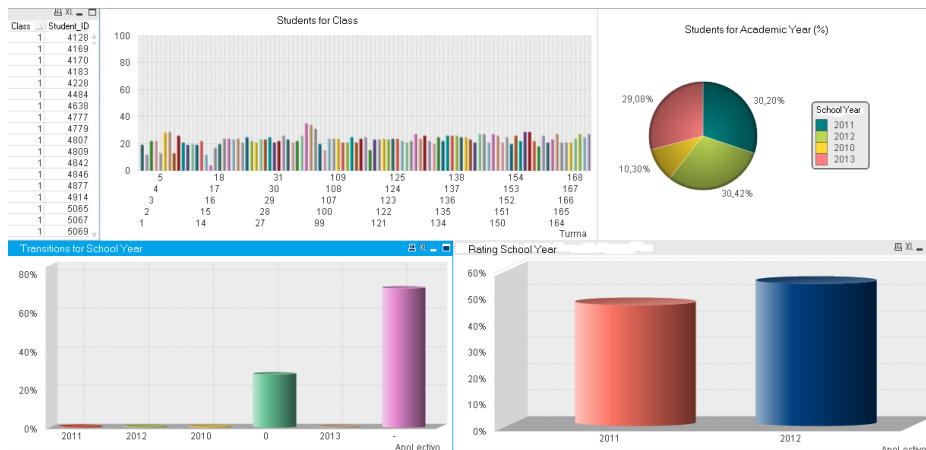

Fig. 4 – Student data dashboard

7. Conclusion & Future Work

The work developed has directly contributed to solving the initial problem. With this work it was possible conclude that the primary and secondary clusters schools can

benefit from the introduction of a PBI. With the PBI the schools can included the majority of their stakeholders (professors, managers, directors) in their business intelligence solution. This prototype can take a variety of dashboards using real-time information about the school performance. In order to evaluate to PBI success should be done analysis of the influence and effectiveness of the solution at driving schools performance. This solution is able to:

- Create added value to the schools from real-time outstanding insights, accessing to the data, anywhere, anytime.
- Provides significant performance results where the dashboards can be used by anyone who wants to increase the indicators of each domain presented (1-4).
- Improve operational efficiency by uncovering new best practises and driving those practises from all the schools clusters. With the PBI all schools use the same BI platform being the information grouped by each one.
- Empower all the stakeholders with dashboards to measure the progress against strategic corporate goals and KPIs of school cluster (it can be done in real-time and the schools automatically monitored).
- Send notifications pre-defined by the user to for example alert for the number of absences by students and professors, a decreasing in the student grades, an increasing of school quality in terms of performance results and other changes in dashboards / indicators.

It is important to note that the solution presented can be used by other countries being only necessary to make some adaptations in the relational model and if necessary create more indicators having all the characteristic mentioned in the section 2.3.

This project opens new doors at a research level and can be seen as a starting guideline to implement a PBI in the primary and secondary schools. In the future this solution should be monitored and the data in failure should be added in order to improve and complete the PBI solution.

Acknowledgments

References

1. Vercellis, S.: Data mining and optimization for decision making. John Wiley & Sons Ltd (2009)
2. Cube, Q.: Quantum Cube. Available at http://www.q-cube.net/pdf/Flyer_QCube.pdf. Accessed on December, 8, 2014 (2014)
3. Logianalytics.com: Dashboard   Software   -   Logi   Analytics.   Available: http://www.logianalytics.com/. Accessed on December, 8, 2014 Accessed on December, 8, 2014 (2014)
4. Portela, C.F., Santos, M.F., Silva, A., Machado, J., Abelha, A.: Enabling a Pervasive Approach for Intelligent Decision Support in Critical Health Care. In: CruzCunha, M.M.,

Varajao, J., Powell, P., Martinho, R. (eds.) Enterprise Information Systems, Pt 3, vol. 221, pp. 233-243 (2011)

5. Satyanarayanan, M.: Pervasive computing: vision and challenges. Personal Communications, IEEE 8, 10-17 (2002)

6. Henricksen, K., Indulska, J., Rakotonirainy, A.: Infrastructure for pervasive computing: Challenges. pp. 214-222. (Year)

7. Turban, E., Sharda, R., Delen, D.: Decision Support and Business Intelligence Systems. Prentice Hall (2010)

8. Michalewicz, Z., Schmidt, M., Michalewicz, M., Chiriac, C.: Adaptive Business Intelligence. Springer (2007)

9. Rahman, N.: Refreshing data warehouses with near real-time updates. Journal of Computer Information Systems 47, 71-80 (2007)

10. Kimball, R.: The data warehouse lifecycle toolkit: expert methods for designing, developing, and deploying data warehouses. John Wiley & Sons (1998)

11. Kimball, R., Ross, M.: The data warehouse toolkit: the complete guide to dimensional modeling. John Wiley & Sons (2011)

12. Microsoft.com: Explore o SQL Server 2012-2014 | Microsoft. http://www.microsoft.com/pt-br/server-cloud/products/sql-server/. Accessed on December, 8, 2014 (2012)

13. Qlik: What Makes QlikView Unique. Whitepaper (2011)

14. Ferreira, A., Castro, A., Soares, D.d.S., Portela, F., Santos, M.: Towards of an information service to educational leaderships: business intelligence as analytical engine of service. (2014)