# Real-Time Business Intelligence Platform to Maternity Care

Andreia Brandão, Eliana Pereira, Filipe Portela, Manuel Santos, António Abelha and José Machado

*Abstract—* **The motivation for the implementation of a decision support system in maternity care came from the fact that people are constantly making quick decisions based on incomplete information. There is a significant impact on the patient health, as well as in increasing medical errors. To implement this system, it was resorted to the technologies of Business Intelligence, which involved the construction of two data warehouses with a dimensional structure in a star shape, for two distinct modules, in Gynecology and Obstetrics cares. The feasibility of an evidence-based practice and medical decision making in real time with universal and interoperable features are some of the benefits resulting from the implementation of decision support system in maternity care. In this paper we present the architecture of BI solution, some clinical outcomes and some benefits of the BI solution in a real world context.**

## I. INTRODUCTION

The Information and Communication technologies have a high potential to facilitate information sharing, communication and collaboration between health professionals, increasing the quality and efficiency of the health system as well as the use of Electronic Health Records (EHR) [1], [2],[3]. In recent times, the technologies of Business Intelligence (BI) have been the target of interest to health professionals and to the professionals of Information Technology (IT), due to its applicability in EHR [4]. BI is a process that encompasses several methodologies, applications and technologies for collecting, storing, manipulating, analyzing and providing access to data to help enterprise users make better and faster business decisions.

Thus BI has the ability to operationalize the repository content of EHR in supporting evidence-based practice and improving the quality of healthcare delivery [3], [4],[5]. In the case of healthcare organizations, the majority of clinical data that document their daily activities are stored in Relational Database Management System (RDBMS).

Because of the extensive amount of information, this information is stored in different ways and therefore highly heterogeneous with each other. On the other hand, a decision-making process, where it is necessary to integrate multiple data provided by clinical, medical, financial and administrative systems, where the sources are quite heterogeneous, large and complex becomes extremely important to meet the data quality that directly interferes in the success of the Knowledge Discovery Database (KDD) process [5][6].

So, with this increasing amount of information there is also a corresponding need to apply Data Mining (DM) technologies to extract data quality from information stored in databases to provide real-time decisions [4]. Most clinical data are not structured and the techniques of DM work well with structured data. It is inferred another advantage to using BI as a decision support technology since it allows the combination of structured and unstructured data [4].

In the new Centro Materno Infantil do Norte (CMIN), formerly known as Maternidade Júlio Dinis (MJD), the EHR platform is supported by the Agency for Integration, Archive and Diffusion of Medical Information (AIDA) [7]. This platform is considered a good source for the application of BI technologies, since it is stored many records of patients [3], [4]. However, the large amount of data stored hinders the health professional to have a sense of what type of patients are received and their characteristics. Thus, the implementation of a BI platform in CMIN aims to help and support the decision making process, providing access to key perform indicators (KPIs) for the care of patients in gynecology and obstetrics.

As referred, this article is focused on presenting the architecture of a global BI platform that can be used in Maternity Care Hospitals, taking as a case of study the CMIN. It was planned apply BI and Data Mining (DM) techniques in order to develop a platform where it is possible generate useful KPIs (clinical and management) for healthcare professionals in the context of the Obstetric and Gynecological (GO) module and in the Voluntary Interruption of Pregnancy (VIP) module. For the construction of this platform, two Data Warehouse (DW) were built using the star schema, one for each module of CMIN (GO and VIP). For better accuracy of the solution and validity of information generated throughout the work has been pleading with clinicians and specialists. The solution in development is a pervasive real-time web-based BI application and due to their characteristics allows it to be accessed anywhere and anytime.

Beyond the introduction, the article includes six sections. The first one is related to the background and relative work done in the CMIN. Topics such as interoperability, the work done on the GO module and work done on VIP module are

addressed. The second section discusses in general terms the BI process used in this project. The third section discusses the process of DM. The fourth section outlines the BI system that is being implemented in CMIN and some of the indicators to get at the end. Finally, some discussions and conclusions are made and future work is suggested.

## II. BACKGROUND AND RELATED WORK

### A. Gynecology and Obstetrics specialties in Maternity Care

The Centro Hospitalar do Porto (CHP) results from the union of three hospitals, namely, the Hospital de Santo António (HSA), Hospital Joaquim Urbano (HJU) and Centro Materno Infantil do Norte (CMIN). Each one of these units were separate entities until the creation of the CHP in 2007. With the creation of CHP, women who need urgent care in the specialties of Gynecological and Obstetrics are directed to MJD, now CMIN, which was not previously checked [7].

The CMIN is a health institution, which one of the specialties is GO. In this unit, the support is provided to the pregnant throughout the gestation period, from the beginning of the pregnancy until giving birth, and the neonatal period. Thus, doctors and nurses specialize in GO analyze these types of patients, to meet their needs and to support them in the first cares. Patients who use this skill in CMIN are usually pregnant women, no pregnant women, postpartum women, women who plan VIP or women who wish to take the examination of Cardiotocography (CTG) [7].

The HSA has a general emergency department with a Manchester Triage System (MTS), implemented since 2000. In order to provide better care to women in emergency situations, they are directed to the emergency department of CMIN. But a system of triage as MTS was not efficient for the specific cases. Therefore, a specific pre-triage system has been used with a specific questionnaire for these cases, allowing identifying more precisely and accurately the severity of the patient. This system classifies patients as emergent if it is a very serious case, and the pregnant is answered immediately. Urgently, being the pregnant forwarded to the emergency service where she has to wait a few seconds/minutes to be attended or so if it is a less urgent the patient is forwarded for outpatient service. This system is not yet a system of priorities as the MTS implemented in the general emergency department. This system was implemented just to distinguish between urgent and less urgent cases [7]. However the triage system for this environment using a priority model is under developing.

In VIP module of CMIN, clinically non-surgical methods are used for this purpose. More specifically, World Health Organization (WHO) recommends a drug method that has been proven safe and is used quite effectively. This method consists of the combination of medication (mifepristone and misoprostol). The process of VIP consists of several stages. Initially, prior to performing the abortion, there is a mandatory meeting with a physician, a reflection period of three days and still optional meetings with a psychologist or a social worker. Later, after the patient is aware of her decision, there are three moments of monitoring with the nursing team. The first concerns to a nursing procedure, where it is applied the first dose of medication and the triage is performed to verify if the patient will need to be

followed in the administration of the second dose of medication or if she will be able to self-administer at home. Finally, there is the collaboration of nursing, after the interruption of pregnancy, and a medical appointment of control and family planning. The idea of implementing a BI platform is a follow up to facilitate the work of health professionals, who were responsible for obtaining reports, which should contain indicators related to the information recorded during the proceedings. These reports are held annually and are used for evaluation of CMIN as is the example of VIP module [8].

### B. Interoperability System

At Centro Hospitalar do Porto (CHP), the interoperability between information systems is ensured by AIDA. This system of intelligent agents allows communication of the different CHP systems, as well as standardization of clinical systems, and overcoming the medical and administrative complexity of the different sources of information.

Thus, AIDA ensures the sending and receiving of information from the hospital, managing, storing and responding to requests for information from third parties [9],[10]. All information systems of CHP are connected through AIDA, for example, the SAPE (Nursing Support System) and SONHO (Clinical Information System) [9], [10].

## III. BUSINESS INTELLIGENCE

Business Intelligence (BI) is defined as the transformation of information into knowledge and has the ability to provide the right information to the right user at the right time to support the decision-making process [4],[11]. BI integrates a set of tools, technologies and programming products that are used to collect, integrate, analyse and make data available [11].

BI is a popular theme and can be combined with Data Mining (DM) and Knowledge Discovery in Database (KDD). This technology is important because it provides applicable technologies to transform operational DW, providing strategic and/or business logic information to support the decision making process of any organization. BI has become an important product in the IT industry, due to the lack of efficient tools to support the DM of unstructured data sets. With BI that problem does not exist because it works very well with sets of structured and unstructured data[12].

DW is a basis for business intelligence. The principal architectures of development of DW are the Top-Down Approach and the Bottom-Up approach. In this project, it was used the Bottom-Up approach, because it is faster, being implemented in several steps. In this architecture, data flows from source into data marts, then into the DW [6][13][14]. An example is the Kimball's approach, which is the most used methodology and it was the approach followed in the development of this BI platform.

In figure 1, it is identified the sequencing of tasks and it is highlighted the activities that must happen simultaneously, requiring the following steps:

- **Program/Project planning**- In this stage it was performed tasks, such as survey of KPIs (clinical

and management) that would be needed in CMIN, was made an initial planning of all of the Kimball approach stages and were defined some deadlines for various tasks.

- **Management Program / Project** - This phase follow the entire project from start to finish and it is responsible for tasks such as strengthening the project plan initially developed.

- **Definition of business requirements** - At this stage it was made a planning / survey of all necessary KPIs explored in two modules. Both in the case of GO module as in the VIP module was discussed the possible final KPIs, together with healthcare professionals and IT professionals involved in CMIN, in order to get a more precise idea of the final requirements of the BI platform.

In the diagram of figure 1 it can be seen that three tracks project occur simultaneously. These are focused on technology, data and BI applications and divided into several steps. These steps are:

- **Technological task includes:**
  1. Technical architecture design - A technical drawing of the overall architecture was done.
  2. Product Selection and Installation - At this stage should be done the selection and installation of all the most suitable BI tools to build the platform. In this case, was chosen Oracle Database, Oracle SQL Developer and Pentaho BI Open Source as OLAP tool.

- **Data Tasks includes:**
  1. Dimensional modeling – It is the DW design phase that intended to offer a facilitated support to end users queries. To model this project, at this stage was chosen the star schema. The star schema integrates a fact table (center of the star) and dimensions in multiple tables [15].
  2. Physical Design - In this stage it is defined the physical structure that has involved tasks as: setting up the database environment and setting up appropriate security.
  3. ETL design and development - This phase includes the extraction, transformation and loading of data. In extraction, initially data from databases used in CMIN were extracted and subsequently these were extracted again, after the transformation and cleaning processes. In the transformation phase, were performed cleaning, standardization and integration processes, so as to obtain accurate, consistent and unambiguous data. Finally, the load, the data resulting from the previous steps are inserted into a multidimensional structure for particular end users and application systems [16].

- **The Business Intelligence application includes**:
  1. BI Application Design - The goal is to provide resources that are accepted by CHP / CMIN to increase their support and decision making, so it can be identified the candidate to the BI applications and appropriate navigation interfaces to address the CMIN's needs.
  2. BI Application Development - In this phase, it was performed the construction and validation of the specified operational BI and analytic applications.

- **Deployment** - Begins the actual development of the BI application, so it is essential to do a proper planning to make sure that the results of technology, data and BI application tracks are tested and fit properly.

- **Maintenance** - Occurs when the system is in production. Includes technical operational tasks that are necessary to keep the system performing optimally.

- **Growth** - DW systems tend to expand. In this case, the cycle should be started again and develop the foundations already laid, with a focus on new requirements. In case of future, this BI platform can be expanded with new modules of interest in CMIN for the process of decision-making.
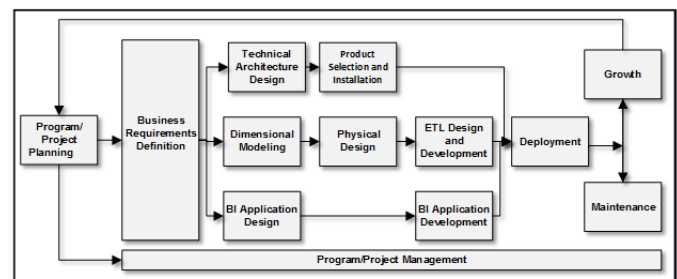


Figure 1. The Kimball Lifecycle Diagram.

IV.  DATA KNOWLEDGE DISCOVERY IN DATABASES

The Knowledge Discovery in Databases (KDD) process includes five stages and aims at extracting knowledge in order to support the process of decision-making. The steps are as follows:

- **Selection** - At this stage occurs the selection of the data set that will be needed to perform the DM. For this case were used data from AIDA, SAPE and SONHO;

- **Pre-processing** - This step includes cleaning and processing of data in order to make them consistent. When the data is selected from the DW, these are already treated, since it already underwent an ETL

process. Here are validated through clinical principles in which there is for example the elimination of null and noise values;

- **Transformation** - In this phase the data is worked out according to the target. For example has been done a reduction of the search space, where were eliminated some of the rows and/or columns to analyze, as well as changes in attributes with continuous values for attributes with discrete values;

- **Data Mining** - consists of finding patterns or relationships that may exist in the data stored in repositories. In this phase, the type of task being performed (classification, segmentation, summarization, dependency modeling) is defined and identified the technique to be used (decision trees, association rules, linear regression, neural networks, among others). In this particular case, classification tasks on problems of DM were used. An example of the application of the classification task is the DM problem that pretends to predict if the VIP patient performed the second part of the process at home or in ambulatory at CMIN. In this issue, were applied the DM techniques of Support Vector Machine, Decision Tree, Generalized Linear Model and Naive Bayes

- **Interpretation / Evaluation** - Consists in the interpretation and evaluation of the patterns obtained. The validity of the results is verified by applying patterns found in new data sets [17].

## V. BUSINESS INTELLIGENCE PROJECT

Following the methodology of Kimball addressed in section II and III, a BI project for Gynecology and Obstetrics unit was developed. This project fits in consideration of the case study: CMIN.

### A. The Main Purpose of BI Platform and Benefits

All the records of patients who use the CMIN are recorded and stored in AIDA, which supports the EHR. However, there is no system for decision support that gives utility to the stored data, transform them into useful knowledge and allow a clinical medical evidence-based practice [7]. The data recorded in the EHR, as well as the information stored in the databases are not very useful for the extraction of knowledge, both in the GO module as the VIP module. When there is such data collection, are used time consuming and expensive techniques for extracting knowledge as rudimentary techniques such as spread sheets to obtain reports or charts, as has been used in recent years in the VIP module. In this case, the idea of implementing a BI platform in the VIP module emerged as a way to facilitate the work of health professionals responsible for obtaining the reports with the information collected during the procedures [8].

Thus, a major objective for implementing this BI platform in CMIN is to give usability to the large amount of data recorded, enabling the creation of new knowledge and allow the optimization and automation of the entire process of obtaining indicators, so that these could be available at the time of decision. In the GO module there is no process of knowledge extraction implemented. Thus, this BI platform emerges as a novelty in this module, enabling physicians perform a safer clinical practice, reducing wasted time and decreasing the number of medical errors. In addition to these specific situations, the BI platform addresses the following benefits for CMIN:

- Organization of dispersed information, extracting only those that may be relevant to the maternity, and build scenarios to help health professionals in decision-making;

- Greater autonomy and flexibility of users when it comes to creating reports;

- Faster and simpler analyses of clinical data;

- Better support for decision making;

- From graphs, for example, the processes and work flows can be monitored, the difficulties identified and improved processes;

- Compare outcomes of decisions taken in the past with present results, in order to record the evolution.

- Better protection of patient data is also possible, providing access to data to only those with appropriate access levels.

### B. BI Project Planning

After the definition of the problem, the following step is the planning process, according to the Kimball's methodology. Thus, the process begins with the collection and selection of information that is required to generate the indicators, which in this case was held in EHR and SAPE. In this project, the data were stored in separate databases interconnected by AIDA. After collecting the data, they are evaluated and processed through ETL procedures defined. Subsequently, these data are loaded into two different DWs, whose multidimensional model was previously defined using the star schema [6], [13], [14].

For each module of the CMIN (VIP and GO) was built a different DW with distinct dimensions, so that each module could generate their own indicators. This is due to the fact that, despite the type of patients be the same (pregnant women), processes and the data collected are different, hence the need to create two DW, which will feed into the same BI system. However, if there is information common to both, it is possible crossing information. Thus, the data related to each of the modules were loaded into the respective DWs.

In the ETL procedures, they were used to eliminate null and repeated data and to transform information. In the loading phase it was used procedures to load / update information for the DW. These procedures are predefined and scheduled tasks in order to automate the process. After loading the DW the next step was the development of On-Line Analytical Process (OLAP) Cube.

The OLAP cube consists of the methodology to provide quick answers to ad-hoc queries on dimensions and metrics stored in the DW. Therefore OLAP cubes structures are highly optimized to perform ad-hoc queries by storing pre-computed results for different combinations of dimensions and metrics. Thus, it was selected the Pentaho Open Source BI as a tool for building OLAP cubes and the visualization of the indicators requested by CMIN through reports, charts, dashboards, among others. Figure 2 represents the BI architecture in terms of the process previously presented.
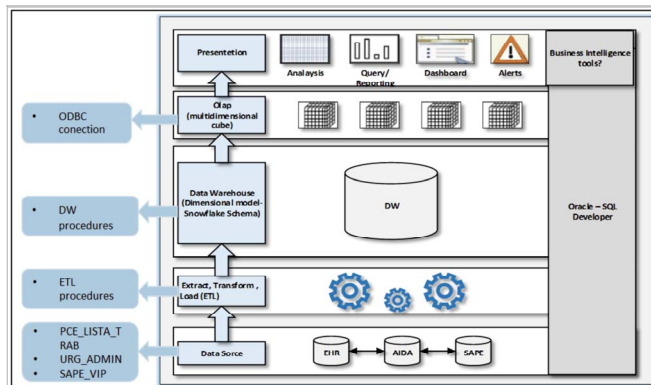


Figure 2. BI architecture.

### C. Key Perform Indicators

#### 1) Obstetric and Gynecological Module

In this module the indicators that are under development are concerned with activity of physicians between a certain time periods; daily activity; professional high times; doctors high clinical; number of urgent episodes and outpatient services; evolution of the high hourly in clinical obstetric module; flow of patients arriving in the CMIN, number of births per year; number of obstructed births; number of births by skilled nurses; number of live and stillborn births per year, among others.

#### 2) Voluntary Interruption of Pregnancy Module

The implementation of a BI platform to the VIP module has the major advantage of extracting knowledge from information recorded in the SAPE. Thus, the indicators generated for some of the activities undertaken and the results obtained give an overview of the operation of this module and consequently implement improvement measures. Some of the indicators that will have a positive contribution to the CMIN are characterization of the user group by age; characterization of the user group for the number of pregnancies; characterization of the user group for the place where the previous VIP has been done; characterization of the user group for the number of miscarriages; contraceptive methods used at the end of the VIP process; number of VIP failures according to the number of weeks of pregnancy; among others.

### D. Predict Clinical Events

The BI platform also intends to incorporate predictive models of DM. In the VIP module, models were developed to predict whether women need up monitoring of nursing

staff in the administration of the second dose of medication or if they are able to do so at home, taking into consideration some variables. Another model of DM that also intends to be implemented in the BI platform is providing which is the risk group of patients who resort to CMIN to hold a VIP, i.e. the patients who not attend to the last review appointment, where the final outcome of the procedure is verified. In the GO module, has been worked on models to predict the waiting time for patient is answered after triage.

### E. Business Intelligence Application

The BI application is a web platform that can be accessed anywhere, anytime and by different people with different access privileges [18]. The web application that integrates indicators of GO module and VIP module has been developed and is already implemented, having been created for the extraction of knowledge related with other units of the CHP. Thus, two new sections in this application will be presented where indicators are represented in graphs, tables or diagrams and further indicators from manipulation of OLAP cubes are created. In figure 3, there is shown a display in a bar graph, which shows the characteristics of the group of patients who have made a VIP for number of pregnancies. In this graph, the x-axis corresponds to the number of pregnancies, the y-axis corresponds to the number of associated patients and the different colors are associated with different time periods, recorded as 'year-month'.

Regarding the creation of OLAP cubes, these allow users to perform ad-hoc analysis of data considering multiple dimensions, thus providing the necessary information for better decision-making. Thus, this technology provides performing various operations such as roll-up, drill-down, slice and dice and pivot [6], [15].
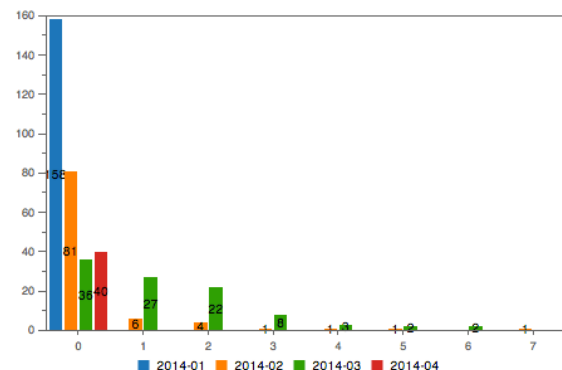


Figure 3 - Group's characterization by number of pregnancies per year-month represented in a bar graph.

## VI. DISCUSSION

The implementation of a BI platform in CMIN can bring many benefits to this institution, namely, increased autonomy and flexibility of the users with regard to creating reports, faster and easier analysis of the clinical data, support clinical decision-making, among others.

Scientifically, this project will also have a significant contribution in that it is a pioneering work in the BI area applied to maternity care. From the knowledge gained through the indicators generated, this can be considered a

starting point for the implementation of improvement measures in the GO and VIP modules. Furthermore, it was still possible knowledge extraction through different DM models, which can be considered good models, in that the sensitivity values of around 90% were obtained.

This project can still be useful for implementing BI platforms similar to the GO and VIP specialties in other health institutions, as well as other specialties.

Thus, this BI platform can allow having an overall overview of CMIN operation, giving usability to the large amount of recorded clinical data and extracting from them the inherent knowledge.

## VII. Conclusion and Future Work

With the achievement of this project it was possible to explore the importance of integrating BI technology in CMIN. The implementation of this technology guarantees the quality of the extracted data analysis, although most of the clinical data are unstructured.

The establishment of this BI platform involves the construction of two DW, one for each data of GO and VIP module, respectively, necessary for obtaining the indicators stored. Thus, from this platform, it can be generated indicators to the GO module, for example, a medical clinical activity over a given period of time, and for the VIP module, such as the characterization of patients by age. The indicators mentioned in this article are just some of the most important, however, more indicators that are the target of interest can be added. In addition, this platform also integrates predictive DM models in order to extract knowledge from data obtained from electronic records and to provide real-time support in the clinic decision-making process to doctors and health professionals.

The implementation of a BI platform in CMIN, in GO and VIP services, is quite innovative, in that it is the first time that BI technologies are applied in these areas.

As already mentioned this platform is undergoing tests. After the implementation of the platform's prototype, it is intended to perform an analysis based on TAM3 to evaluate the user acceptance. Moreover, after getting the results is intended to do all the improvements to undergo a process of making decisions with greater confidence. In future, it is pretended to make a personalized patient care through the creation of new indicators to better exploit the capabilities of DM under the development of useful predictions in CMIN, constantly improving the platform by meeting the needs of users (doctors, nurses, managers, patients) of CMIN. This allows for evidence-based clinical practice increasingly. Also it is intended to extend the platform to other areas of interest for CMIN.

## References

[1] A. Abelha, C. Analide, J. Machado, J. Neves, and P. Novais, Ambient Intelligence and simulation in health care virtual scenarios, in Establishing the Foundation of Collaborative Networks, Camarinha-Matos L. Afsarmanesh H., Novais P., Analide C., (Eds), Springer-Verlag, 2007.

[2] S. Khodambashi, "Business Process Re-engineering Application in Healthcare in a Relation to Health Information Systems," *Procedia Technol.*, vol. 9, no. 2212, pp. 949–957, Jan. 2013.

[3] F. Portela, M. Vilas-Boas, M. F. Santos, A. Abelha, J. Machado, A. Cabral, and I. Aragao, "Electronic Health Records in the Emergency Room," *Computer and Information Science (ICIS), 2010 IEEE/ACIS 9th International Conference on* Computer and Information Science,. pp. 195–200, 2010.

[4] W. Bonney, "Applicability of Business Intelligence in Electronic Health Record," *Procedia - Soc. Behav. Sci.*, vol. 73, pp. 257–262, Feb. 2013.

[5] T. Mettler and V. Vimarlund, "Understanding Business Intelligence in the Context of Healthcare," *Health Informatics J.*, vol. 3, no. 15, pp. 254–264, 2009.

[6] M. Viana, O. Oliveira, A. Abelha and J. Machado, "Step towards Simulation and Monitoring of Hospital Waiting Lists", in Proceedings of the ESM 2012, Eurosis, Essen. Germany, 2012.

[7] F. Portela, A. Cabral, A. Abelha, M. Salazar, C. Quintas, J. Machado, J. Neves and M. Santos, "Knowledge Acquisition Process for Intelligent Decision Support in Critical Health Care," in *IGI Global Book*, 2013.

[8] C. Valente, T. Cristina, F. Rosário, and B. Alcina, "Centro Hospitalar do Porto - Unidade da Mulher e da Criança Acompanhamento de enfermagem na interrupção da gravidez por opção da mulher ( I . G . O .) Índice," 2012.

[9] J. Machado, V. Alves, A. Abelha and J. Neves; "Ambient Intelligence via Multiagent Systems in Medical arena", International Journal of Engineering Intelligent Systems, Special issue on Decision Support Systems; vol. 15, n.3, pp. 167-173, 2007

[10] H. Peixoto, M. Santos, A. Abelha, and J. Machado, "Intelligence in Interoperability with AIDA," in *Foundations of Intelligent Systems SE - 31*, vol. 7661, L. Chen, A. Felfernig, J. Liu, and Z. Raś, Eds. LNCS, Springer, 2012, pp. 264–273.

[11] Reinschmidt, J., and Francoise, A., "*Business Intelligence Certification Guide*", IBM, International Technical Support Organization, 2002.

[12] H. Baars and H.-G. Kemper, "Management Support with Structured and Unstructured Data-An Integrated Business Intelligence Framework," *Inf. Sys. Manag.*, vol. 25, no. 2, pp. 132–148, 2008.

[13] A.Alacantara, "Business Intelligence (BI) como auxílio à Gestão do Negócio," Faculdade de Tecnologia da Zona Leste., Brazil, pp. 1–83, 2010.

[14] S. Chaudhuri, U. Dayal, and V. Narasayya, "An Overview of Business Intelligence Technology," *Commun. ACM*, vol. 54, no. 8, pp. 88–98, 2011.

[15] A. Ribeiro, "Implementação de um Sistema de Business Intelligence para a análise da Doença Pulmonar Obstrutiva Crónica," Universidade do Minho, 2011.

[16] S. El-Sappagh, A. Hendawi, and A. Bastawissy, "A proposed model for data warehouse ETL processes," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 23, no. 2, pp. 91–104, Jul. 2011.

[17] A. Azevedo, "Data mining languages for business intelligence," , University of Minho, 2011.

[18] U. Varshney, "Pervasive Healthcare: Applications, Challenges and Wireless Solutions," *Communications of the Association for Information Systems*: Vol. 16, Article 3, 2005.