

A MULTI-DRIVEN APPROACH TO REQUIREMENTS ANALYSIS OF DATA WAREHOUSE MODEL: A CASE STUDY

Jorge Oliveira e Sá, *Centro Algoritmi, University of Minho, Campus de Azurém, Guimarães, Portugal*

Claus Kaldeich, *Lecturer & Head of the Center for Strategic Intelligence, Innovation, R&D, Technical Trainers College, Riyadh, Kingdom of Saudi Arabia.*
Technical and Vocational Training Corporation (TVTC), Saudi Arabia and Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, Germany

João Álvaro Carvalho, *Centro Algoritmi, University of Minho, Campus de Azurém, Guimarães, Portugal*

ABSTRACT

In this paper, a multi-driven approach to data modeling in data warehousing will be presented, which integrates three existing approaches normally used separately: goal-driven, user-driven and data-driven; and two approaches usually not used in data warehousing field: process-driven and technology-driven. Goal-driven approach produces subjects and KPI's (Key Performance Indicators) of main business fields. User-driven approach produces analytical requirements represented by measures and dimensions of each subject. Process-driven approach propose improvements in business processes (by using and creating subject oriented enterprise data model) to satisfy the KPI's, measures and dimensions identified in the previous approaches. Technology-driven approach is an enabler or an obstacle to be considered in a data warehouse model. Data-driven approach is a combination of the results of previous approaches and results in a data warehouse model. By using a multi-driven approach with five stages, a layered data warehouse model more aligned with business and individual needs can be obtained. This will be illustrated by using examples of a case study.

KEYWORDS

Requirement Analysis, Data Warehouse Model, Case Study.

A MULTI-DRIVEN APPROACH TO REQUIREMENTS ANALYSIS OF DATA WAREHOUSE MODEL: A CASE STUDY

1. INTRODUCTION

Data warehousing enables organizations to have access to management information that is crucial to obtain significant gains, in particular, increased sales, reduced costs, and by offer new services and/or products. Most large companies have established data warehousing systems as a component of information systems landscape (Herrmann and Melchert, 2004). However, the implementation of these systems in organizations suffers from a high failure rate (Chenoweth et al, 2006). Several initiatives to implement data warehousing started with significant expectations, but quickly found that was not guaranteed to satisfy them.

Data warehouse departments are usually confronted with a large amount of requirements from several different data warehouse stakeholders, like, for instance, end users (top managers, middle managers and others), data warehouse sponsors, and operational IT departments. The amount of requirements regularly exceeds the available resources for their implementation (Herrmann and Melchert, 2004).

Nowadays, researchers, practitioners, consultants, among others, have not yet arrived at a consensus in determining the right methodology, method, approach and architecture to implement data warehousing. However, one of the most important issues in data warehousing is how to develop an appropriate data warehouse model to support analysis, like, querying, exploring, and reporting. Consequently, it is still important to research data warehouse modeling in data warehousing field.

Existing data warehouse modeling approaches can fall within three basic groups: goal-driven, data-driven and user-driven (List et al, 2002). Each of the three approaches is usually used separately:

1. Goal-driven approaches consists to analyze the business to obtain a set of visions by interviewing the top and middle management, these set of visions must converge in order to obtain a set of quantifiable Key Performance Indicators (KPI's) (Malinowski and Zimányi, 2009). Goal-driven approaches have a disadvantage to depend on the availability of the top and middle management to be involved.
2. Data-driven approaches consist to study the data sources to determine the data warehouse model (Malinowski and Zimányi, 2009; and Inmon, 2005). Some authors (Song et al, 2008; and Moody and Kortink, 2000) proposed an automatic or semi-automatic methods to create data warehouse models from Entity Relationship Models (ERM) of data sources. Data-driven approaches have a disadvantage to create a set of information structures for storing information in the data warehouse and this information may never be used and failing to adequately involve the future data warehousing users.
3. User-driven approaches consists to determine information requirements of several business users, these set of needs must be integrated to obtain a multidimensional model (Winter and Strauch, 2003; and Kimball et al, 2008). User-driven approaches have a disadvantage to mapping the information requirements with the available data sources only *a posteriori*, sometimes this mapping fails and may cause users' disappointment.

Data warehouse model obtained from a single principle are usually incomplete, which cannot obtain satisfaction and trust of organizations and individuals simultaneously (Guo et al, 2006). Some authors propose a combination of user-driven and data-driven approaches (Winter and Strauch, 2003). Others combines a goal-driven, data-driven and user-driven

approaches and comprises four stages: goal-driven stage produces subjects and KPI's of main business fields; data-driven stage produces subject oriented enterprise data model; user-driven stage yields analytical requirements represented by measures and dimensions of each subject; and a combination stage combines the triple-driven results, this propose claims to get a more complete, more structured and more layered data model of a data warehouse (Guo et al, 2006).

In this paper a multi-driven approach will be described, this approach is a combination of business process management approach (process-driven approach) with goal-driven, user-driven, technology-driven and data-driven approaches, and comprises five stages: KPI's stage, by identifying user requirements and organizational goals (goal-driven and user-driven approaches); technological stage, by identifying the existing technological capability in the organization to implementing a data warehousing; process stage, by modeling organizational processes (as-is); process improvement stage, by modeling organizational processes and its data, based on the requirements and goals identified at stage 1 and its implementation by ensuring that they are aligned with business goals (to-be); and mapping stage; by combining the results of the integrated process-driven, the data warehouse model can be obtained from a data-driven approach taking into account the requirements, technological capability and goals identified.

A case study was conducted in small medium-sized printing/graphic arts industrial company located in Oporto, Portugal. This company does not have any experience in data warehousing, as does not have a department of information systems, depending on IT vendors to improve and develop their information system. This case study aimed to make an implementation of a data warehousing solution using a multi-driven approach with five stages to achieve the requirements analysis to obtain the conceptual model of the data warehouse. The combination of different approaches allows obtain the “right” data warehouse conceptual model, i.e., more adjusted to the organization and individuals (users) needs. In this paper, the emphasis will be given to the process-driven approach (stage 3 and 4). To do a process-driven approach, started with a requirement analysis of the current information system (IS) (as-is) and according to organizational goals and users' needs, some improvements to the IS (to-be) will be proposed.

Through the presentation of the multi-driven approach, the paper aims to answer this research question: How a multi-driven approach with five approaches/stages helps to obtain a better data warehouse model?

The remaining sections are organized as follows: Section 2 presents the related work. Section 3 describes the as-is model. Section 4 the to-be model. Section 5 describes some insights about process-driven approach. Section 6 presents the multi-driven approach step by step. Finally, section 7 points out conclusions of this study and future work.

2. IMPLEMENTATION STRATEGY

This case study will focus not on the organization itself, but on its information system (IS). So when this case study started in the company there were three Enterprise Resources Planning (ERP) systems not integrated, namely: an ERP called Keren covering the production management of the company (however this ERP does not cover the complete production cycle of the company); a second one to cover the accounting and finance; and a third one to manage human resources.

A MULTI-DRIVEN APPROACH TO REQUIREMENTS ANALYSIS OF DATA WAREHOUSE MODEL: A CASE STUDY

This case study started by modeling the as-is IS, to do this a tool called ARIS 6.2 and the techniques proposed by this tool were used. These techniques allow modeling the relationship between the existing ERP and the business processes. The tool ARIS 6.2 is recommended by the ERP SAP R/3 to define organizational functions or processes as well the workflows to which these functions or processes are connected to facilitate the implementation of ERP (Scheer, 2001). ARIS model describe four integrated views on organizations: (1) Organization, (2) Data, (3) Control, and (4) Function. One advantage is that a model of well-structured business processes can reduce the time and cost of implementing an ERP, the steps are: the model as-is - the model of business processes as they are implemented in the company; and the model to-be - the model of business processes as they shall be implemented in the company. At this point, the involvement of company management is essential, because now is the time to reengineer organizational processes. These modeling and reengineering are needed for the parameterization of the ERP based on the characteristics and objectives of the company. From ARIS 6.2 two techniques are used: EPC (Event driven-Process-Chain); and ERM (Entity Relationship Model).

A technique called ADAPTTM was used to model the KPI's identified in the to-be model.

2.1 EPC

The EPC is composed by events - the events can trigger functions/processes, for example: "end of the month" triggers the "payroll"; functions - are the basic unit of an EPC, for example, "IRS calculation"; sub-processes - are a set of functions, for example, "calculate the salary of an employee"; processes - are a set of sub- processes, for example, "process the salaries of the company"; flow controls - are symbols whose semantics is the formal logic and determining the flow of tasks / sub-processes within a process; and data clusters - contains the set of data that a function needs, this dataset can be one or more tables of Database. The EPC diagram is very interesting because it integrates organizational units, their functions and the data necessary for the operation of the function. It allows modeling the dynamic characteristics (workflow) and static (cluster data) in a single diagram.

2.2 ERM

The existent clusters in the EPC's, as already mentioned, may have data from one or more entities in the database. Thus, the ERM model describes the data model, in terms to know the entities defined there. The clusters, in turn, assume the task of making the link between the EPC's (i.e. the models of organizational processes) with the database through the ERM.

2.3 ADAPTTM

With ADAPTTM (Bulos, 2000) the focus was on information about a business perspective, and create a logical structure to transform data into information. In our point of view, this diagram represents the path and the transformation process between the initial data (which may be operational data stored in the ERP system, or another KPI) and the final KPI.

The example obtained from the case study, see figure 1, shows how to calculate the KPI - Value Budget, based on the cube Amount Budgeted (yellow color) which represents data from an operational ERM, and the cube Production Value (white color) represents a computed

value, i.e., a KPI. This example shows the possibility to calculate KPI's from other KPI's. In figure 2, the dimensions time, customer and product are presented.

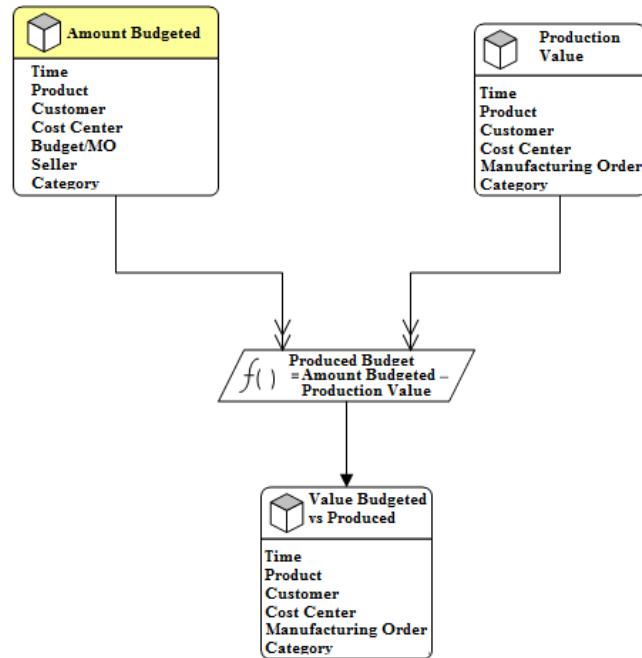


Figure 1. ADAPT™ KPI example.

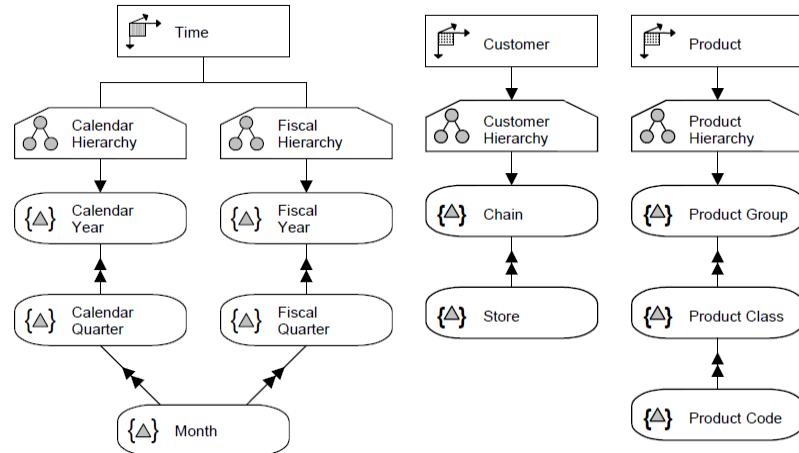


Figure 2. ADAPT™ Dimensions example.

A MULTI-DRIVEN APPROACH TO REQUIREMENTS ANALYSIS OF DATA WAREHOUSE MODEL: A CASE STUDY

3. AS-IS MODEL

The case study started with a survey of existing business processes. Several EPC diagrams were designed, but it is not intention of this study to show the EPC diagrams, but rather to describe the results obtained:

1. Firstly the identification of the ERPs in the organization and the level of integration were made. There are three ERPs: manufacturing, administrative/financial and human-resources, and the three ERPs are not completely integrated. One of these ERPs - Keren, because of its importance, was the subject of further detail. Keren can be characterized as an ERP for manufacturing/production management for the printing/graphics arts industry. Keren is not composed of modules and follows an architecture positioned between the client-server and web-based architecture. An installation of a web-browser and a small component of reporting are enough to work with Keren, Keren can be installed in one or more servers, depending on the desired configuration. In the case study, it appears that there was only a single server that also contained the data base management system (DBMS) in this case the choice was based on a database Oracle 10g, but it could be another DBMS. Keren, includes the following features: budgeting for the production of type: offset rotary, flexography and packaging; orders and customer management; work order (interface for production); stocks management/logistics; concept of perpetual and intermittent stocks; supply chain management: integration of orders/receipts/returns; vendor management/quality; invoices management; production data collection - data collection by multi-terminal through sensors installed on production machines; integration with production equipment using, for example, the CIP4 protocol; and shipping and freight.
2. The case study identify a lack of information in various levels of management, despite the Keren produce a wide range of reports, however, they all had a very large operational perspective, i.e., not allow analysis over time, for example: the comparison of the monthly production of a machine over the past five years. Moreover, there was no integration of information that allows a preparation of analysis of information from the three ERPs. It was found that despite collecting some information to spreadsheets a lot of information was not properly stored. Another area with a large gap was the business analysis, i.e., On-Line Analysis Processing (OLAP). There were no defined dashboards; as well and there were no defined business indicators – KPI's.

One of the characteristics that facilitated all this case study was Keren can be "seen from within", i.e., it is possible to have access to Keren "code" and to physical ERM. This facilitates the extension of the existing features in the Keren.

In a parallel job, a market survey to identify other ERP solutions for this business area was made and allow identify several ERP solutions, including: Agfa Delano, Kodak Epicor Vantage, CRC Information Systems, Prism WIN MIS, among others. This case study not pretend to make a comparative analysis of these systems, but only check if the ERP Keren features are in line with market needs and the conclusion was the ERP Keren complies with the essential features needed to manage a production of a printing/graphic arts industry.

4. TO-BE MODEL

As mentioned above, the involvement of company management is essential. Thus a set of meetings were conducted with the top and second management levels of the company. As a result of these meetings, a clearly identification of a "business pain" and a set of KPI's (business goals) were emerged.

A "business pain" is something that needs to be improved within the company and may justify this intervention. In this sense, the company needs to better understand their customers in order to be able to respond to their needs in a most appropriate way. By not knowing their customers, the company cannot determine, upon arrival from customers a request for quote or order the priority of this task. The need arises to classify and rank the customers to conclude who are the most important and to determine their potential. Therefore, the company needs, because the costs are very high, to know how many customer quotes or contacts were made until being materialized in an order.

The highest levels of management of the company identified a large number of KPI's, but it was found that these KPI's reflect very tactical concerns rather than strategic, for instance: number of printed sheets in a particular printer; number of minutes spent on mobile phone by vendors; etc.

After these steps, a set of personal interviews with middle management to identify their information needs are conducted. As a result, a transformation of these information needs into KPI's and an identification and validation of these KPI's are new or have already been identified are obtained. There are also some KPI's related with the "business pain" described above. Thus a new KPI's list is obtained.

The next step was to model the KPI's in ADAPTTM (see figures 1 and 2 above). In this step, various analytical models (cubes and reports) necessary to meet the management information needs are modeled. At this point, it is possible to realize what are the operational data needed to satisfy the KPI's and to realize that the Keren cannot provide all information (through the existing operational data) to meet the needs of the KPI's, especially those who were related with "business pain".

To satisfy the "business pain" and the KPI's, it was considered relevant develop and implement a system to manage the relationship with customers fully integrated with Keren, however it was found that was not really urgent to implement a "complete" Customers Relationship Management System (CRM), but only a subset of the features of a CRM with the goal to know better the customers. A data warehouse emerges as the best solution to integrate data from CRM system and from ERP Keren and also to integrate data from the others two existing ERP systems. This data warehouse can provide the KPI's required to satisfy the top and middle management needs.

4.1 CRM System

The company intends to implement a system that allows register for a specific request of a customer (or potential customer), how many quotes were elaborated, the rejection reasons, and, if accepted, the reason for the approval and processing into an order. This system must be integrated with Keren, see figure 3.

A MULTI-DRIVEN APPROACH TO REQUIREMENTS ANALYSIS OF DATA WAREHOUSE MODEL: A CASE STUDY

This system has the characteristic of registering the entire workflow of the customer relationship, i.e., receive customer request, validate the request, elaborate the quote, accept the quote and transform the quote in an order. This system allows the top and middle management to have more knowledge about their customers in order to classify and rank them to determine who the best customers are and to know the potential of each customer. Now they know how many quotes are requested by customers until be transformed into an order, or know the customers who withdraws the quote and the reason to do it, or even the quality demanded by customers (note that often the customer in the quoting process diminish the quality of the product to obtain a lower unit price to pay), see figure 4 and 5.

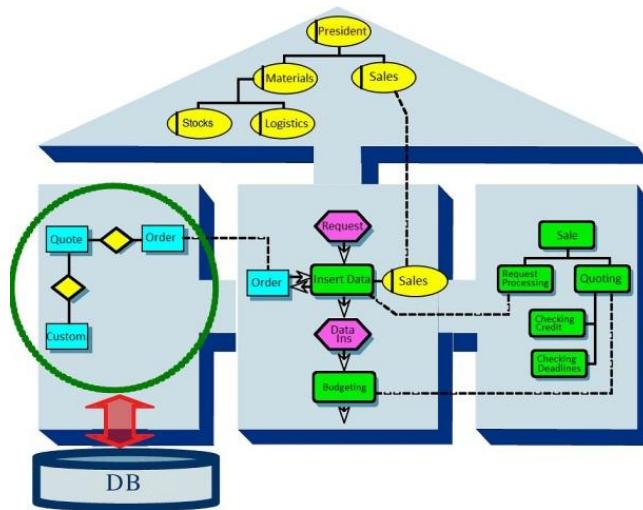


Figure 3. Quoting Process.

Based on this, a new quoting system has been developed. Beyond the technical issues of developing a quoting request process (see examples in figures 6 and 7), in terms of: the definition of the workflow; a choice of materials; and the imposition plans; it also records, for a customer, each quotation requested, allowing to know how many requests are made and what changed between each one (through a version based system). If the customer accepts the quotation value, the production orders are automatically created and if the customer refuses the quotation value the CRM allow registering the reject reasons.

As described above, the CRM allows storing several indicators, for instance: the number of quotes to obtain an order and number of quotes refused. In addition, the CRM stores data from various version budgets, this let knowing the changes that customer made in the quoting process to obtain the right product and unit price. All this data is stored in an operational DBMS system.

4.2 Data Warehouse Conceptual Model

The data warehouse conceptual model is obtained by combining several ADAPTTM cubes. The ADAPTTM model was crucial to obtain the data warehouse model, i.e., the fact tables that with their indicators and the dimensions tables with their attributes hierarchies.

The Extract, Transform and Load (ETL) process can be also obtained by ADAPTTM, i.e., what data will be extracted from the operating system and what is the process of transformation this data suffer until they are loaded (and stored) in the data warehouse. If the existing business processes do not store these data in the company operational system, then the business processes must be studied and modified. This happened in this case study when a new CRM system was developed and implemented to satisfy the information needs.

The data warehouse model follows the recommendations proposed by Kimball, i.e., several data marts are built and integrated by a "Data Warehouse Bus Architecture", this architecture is primarily an implementation of "the bus" - a collection of conformed dimensions and conformed facts; and the dimensions are shared (in a specific way) between facts in two or more Data Marts (Kimball et al, 2008). This approach ensures that the data warehouse model meets all the informational needs by the company and "heals the business pain" through the provision of the identified KPI's.

Related to data warehouse's data, analytical exploration (OLAP) was performed. Many cubes were created so that users could "navigate" in the data warehouse information. Experiments of data analysis through techniques of data mining were conducted, to do this a tool SAS® Enterprise Miner was chosen, where, for example, a correlation analysis were conducted between the months of the year and sales of a certain value of turnover, to analyze by month the sales values and identify sales clusters.

5. SOME INSIGHTS INTO THE PROCESS-DRIVEN APPROACH ("INTEGRATED-PROCESS-DRIVEN")

Let us focus on the last process showed in figure 4 – "EPC (Quoting Process)". The top manager asked for another KPI – "costs variation of quotes versioning".

The background: the quotation process in the printing/graphic industry is significant cost factor, because the sales-man, who calculates and issues the quotations, should be an expert (with a high salary) and the quotation process are quite time consuming. The sales-man has to know about the different production processes, the different raw materials, the different printing machines using different sets of inks and so on. On the top of all, he has to advise the customer about some technical aspects, costs and quality of the printing job.

In order to reduce the unit-price costs, the customer changes some requirements (for instance, changing the set of colors of the printing job, which, for instance, will change some production processes, the production plan or the assignment to a specific printing machine), asking for a 2nd version of the original quotation, and so on.

The managers want to know the costs variation of all those quotations (original, 2nd version, 3rd version and so on...), until the customer approve the last quotation and authorize the printing job. The final quotation amount of some "smaller" printing jobs is sometimes so high, that the profit on a printing job is close to null. This is one of the factors, which increases the "business pain".

A MULTI-DRIVEN APPROACH TO REQUIREMENTS ANALYSIS OF DATA WAREHOUSE
MODEL: A CASE STUDY

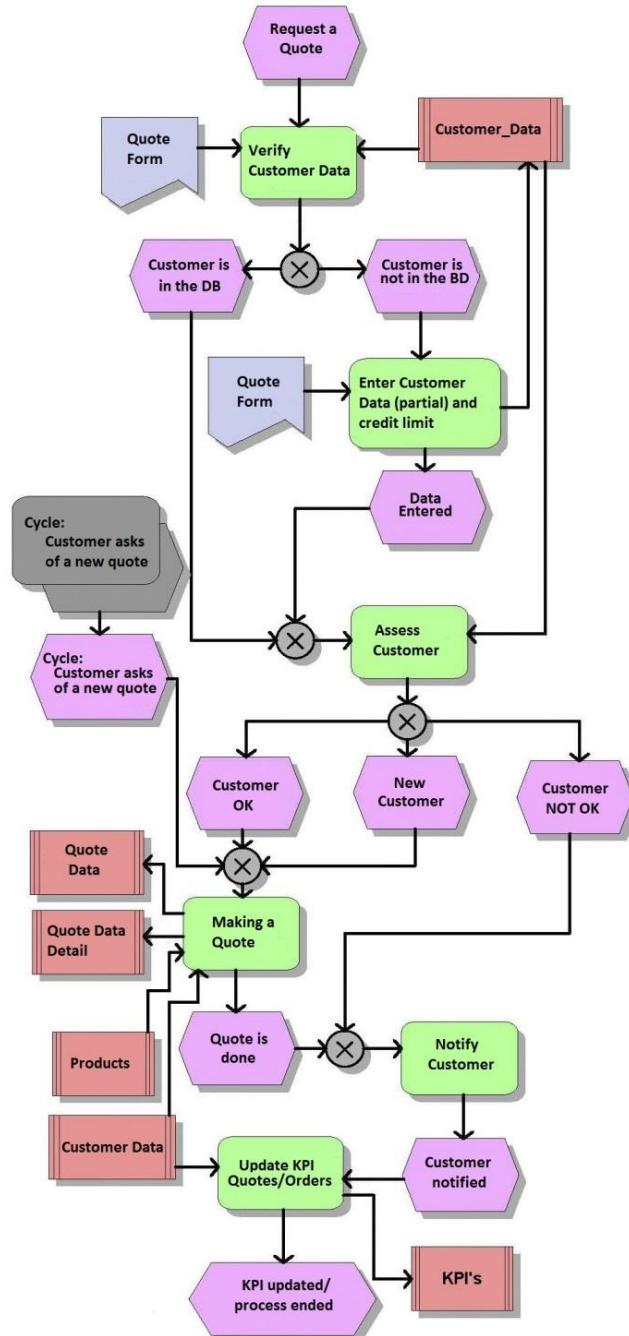


Figure 4. EPC (Quoting Process).

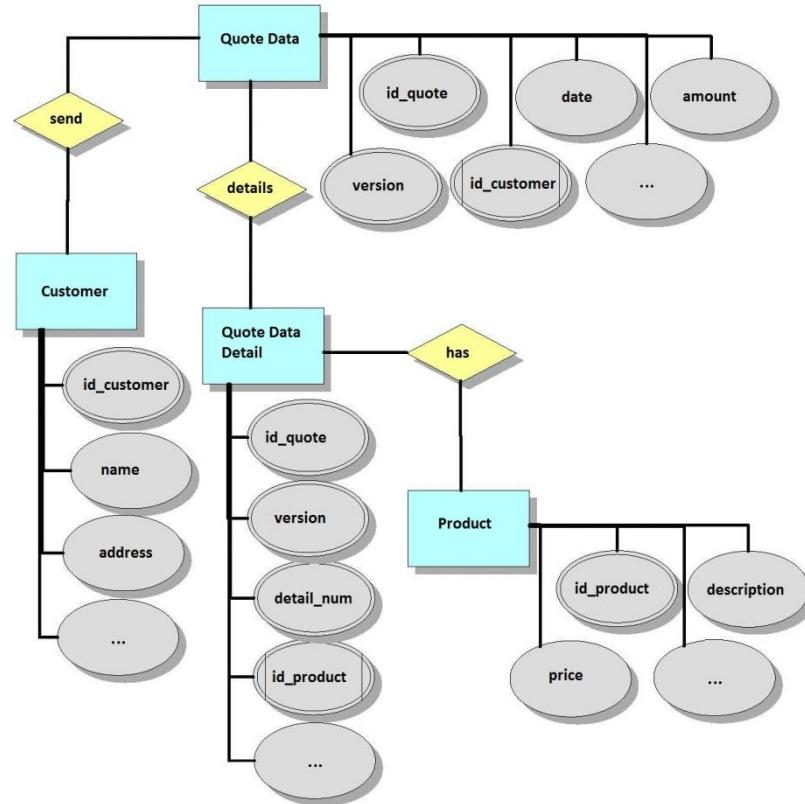


Figure 5. ERM from the EPC (Quoting Process).

Out of this example, is easy to realize, that a significative number of business processes will be involved, in order to gather the right data, on the right time to implement that KPI (see fig. 9). The example below involves also the definiton of "Congruence Relation in Relational Models" and related issues (Kaldeich, 1995; Kaldeich and Sá, 2004).

The example below (see figures 8, 9 and 10) will be presented with no details for obvious reasons, but it will be conclusive enough to allow the understanding of this methodology and its goals.

This example shows:

- i) the connections between two different EPCs, showing schematically how those processes are working and where data is coming from and going to;
- ii) the improvement of the data model by a *congruence relation* (see figure 10), in order to allow the implementation of the "new" KPI. In this case, the data warehouse ETL process will include the congruence relation, in order to get data of the "new" KPI.

The goal of the congruence relation is to define new relationships and gather data from different sources ("congruent data", data related to each other "in a certain (formal) manner"), in order to allow an improvement of the data model.

The EPCs shed light the dependencies of the involved business processes and the data flow.

A MULTI-DRIVEN APPROACH TO REQUIREMENTS ANALYSIS OF DATA WAREHOUSE MODEL: A CASE STUDY

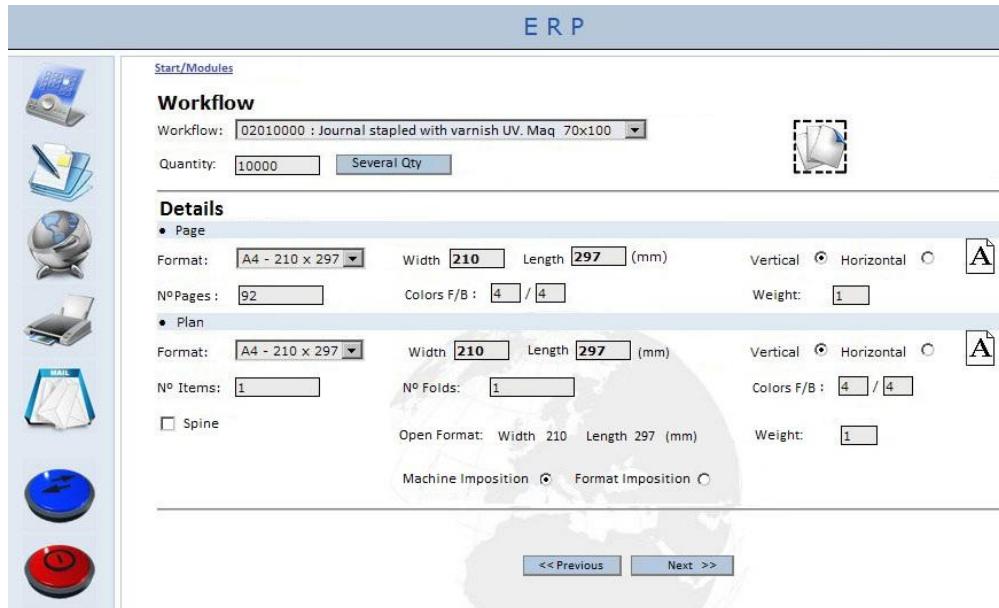


Figure 6. Keren (Quoting Process).

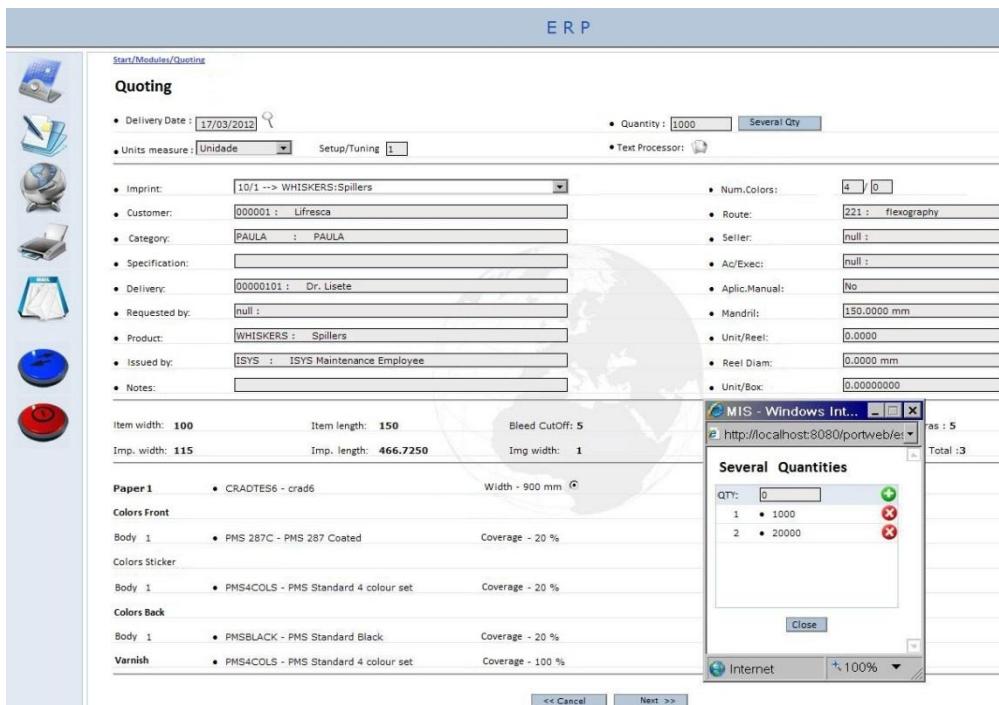


Figure 7. Keren (Workflow Production Definition).

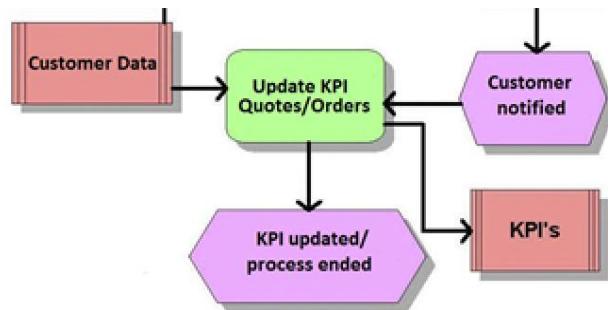


Figure 8. EPC (Update KPI).

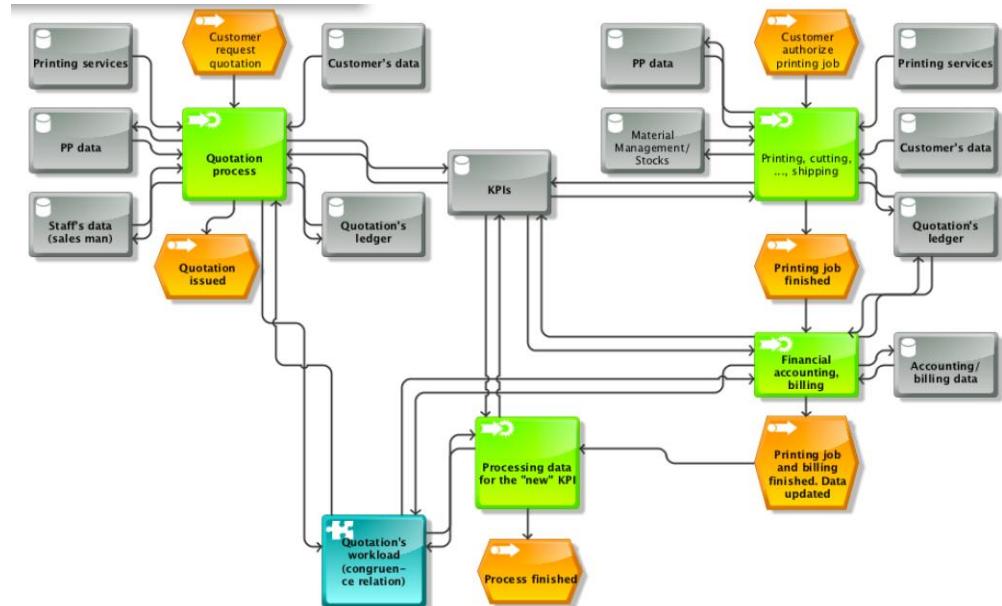


Figure 9. EPC (The "new" KPI and the *congruence relation*)

6. MULTI-DRIVEN APPROACH

A multi-driven approach was followed with five stages:

1. KPI's stage - started by an identification of the company issue (or issues), this is called "business pain"; followed by the identification of the business areas and related topics which will be covered by data warehousing, for instance: sales, procurement, production, human resources. This is very important and should be given greater attention, the issues identified allow to set priorities among the various subjects or topics and to make an agenda, which implies setting the data warehousing, in terms of its content and the numbers of iterations to implement - ideally each subject will be an

A MULTI-DRIVEN APPROACH TO REQUIREMENTS ANALYSIS OF DATA WAREHOUSE MODEL: A CASE STUDY

development iteration in the data warehousing; the identification of KPI's – these KPI's can be related with one or many subjects or topics; and key users identification, which will be involved in the gathering requirements process. Goal-driven results consist in two documents: one with the business pain description and a list of the subjects / topics and key users; and other with a list of KPI's – this is a goal-driven approach. By interviewing key users (identified in the goal-driven approach) with the aim of gathering users' requirements (user-driven approach) and adjusting the KPI's obtained at goal-driven approach. As a result one new document and a update of an existing document: one with a list of users requirements; and an update of a list of KPI's obtained in the goal-driven approach;

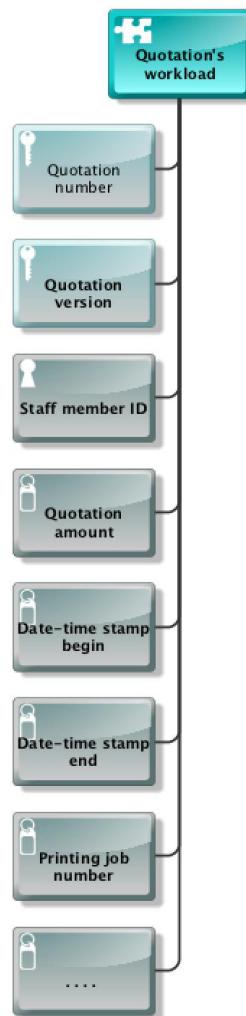


Figure 10. The congruence relation.

2. technological stage - identify the technological capability installed in the company, namely, computer networks, existing servers hardware characteristics, operating systems, computer applications, database management systems, etc. This result in a diagnosis (snapshot) of the current state of the technology in the company and the level to which organizational processes (essentially those related with data warehousing) and the type and size of data warehouse can be implemented and grow with the technological capability identified. There may be recommendations for investment in hardware and / or software which will increase this capability, but, in this particular case study, was not necessary;
3. process stage – taking into account the business areas and requirements identified in the KPI's stage, the organizational systems (technologies) identified in technology stage, the goal of this stage was modeling organizational processes by using EPC notation;
4. process improvement stage – based on processes identified in the previous stage, this stage pretends to optimize organizational processes using EPC notation, the EPCs enable an identification of the operational ERM, which, in turn, includes new information required in the optimization of business processes; and
5. mapping stage – this is the most important point of the multi-driven approach, because the data warehouse model is obtained. By the prioritized list of subjects / topics covered by data warehouse a list of selected KPI's relating with the subject / topic are modeled in a top-down fashion with ADAPT™. As result an analytical model was obtained, the ETL process and operational data required. If the operational data model do not include the required data, then the requirements list must be updated as equal the process or processes must be modified (this implies to go to the process improvement stage again). It may also happen that is not even possible to obtain such operational data and, if so, the KPI's stage must be repeated in order to adjust the KPI's, or, decide to close the project.

A multi-driven approach embraces goal-driven, user-driven, technological-driven, process-driven and data-driven approaches and there is used at all stages described above. In the stage number 1, the approaches goal-driven and user-driven are used. Technological-driven approach is used in the stage number 2. Process-driven and data-driven approaches are used at stage number 3 and 4. Finally, in the stage number 5 a data-driven approach was used.

The multi-driven approach includes two approaches usually not used in data warehousing methodology: technological-driven and process-driven approaches. The technological-driven approach is very important because the technological capability installed in the company can limit the data model of the data warehouse. The process-driven approach allows align business process with the information (KPI's) need by top and middle management. The others approaches, goal-driven, user-driven and data-driven are usually used to obtain data warehouse model, but, normally not combined together.

Now it is possible to answer the research question – How a multi-driven approach with five approaches/stages helps to obtain a better data warehouse model? The case study shows that the data warehouse model obtained from a combination of several approaches can improve the satisfaction and trust of organizations and individuals simultaneously, because by following a multi-driven approach that covers, goals and user requirements, and better processes (with right data) and, finally, appropriate technology.

A MULTI-DRIVEN APPROACH TO REQUIREMENTS ANALYSIS OF DATA WAREHOUSE MODEL: A CASE STUDY

7. CONCLUSION

This was a small and medium-sized industrial company, without having any previous experience of data warehousing, and without an IT department, depending on IT vendors to improve and develop their information system. In this way, the case study allowed to, internally, discuss about processes and technologies, resulting in the emergence of an environment of reflection among the various collaborators, this was achieved because this study involve several collaborators and they had moments away from daily routines, making them able to discuss and find solutions appropriate to the reality of the company.

The company has undergone several changes, note the company at the end of the case study is distinct from that initiated, because have emerged needs of CRM solutions (the need for customer information), new applications of quoting process, production planning, ERP integration, etc. This led their managers to ask KPI's completely different from those that were identified at the beginning, i.e., from simple indicators such counting indicators, as number of printed sheets, to indicators of productivity and customer classification, among others.

The need experienced by the company managers in terms of data analysis, namely, based on business KPI's can really understand what is happening in the company. This is enabled by data analysis through reaching data details which allow understanding the semantic meaning of the value of the KPI's obtained. To facilitate this, an integration of data from multiple repositories (ERPs, management applications and even spreadsheets) was made, which justify the development of a data warehousing, thus allowing "heal the existing business pain".

The improvement of organizational processes has led to obtain further informational records for operational databases in order to ensure conformity with the organizational processes model (to-be). This can be considered a determining success factor for data warehousing in organizations. Ensuring that, business processes are aligned with organizational strategy and identify management indicators (KPI's) that will measure these business processes in order to feed the organizational goals, is critical to the organizations success and information technologies that support organizations.

The combination of several approaches is becoming a concern of several authors, in order to be able to properly define the data warehouse model and this is one of the challenges facing to those who want implementing a data warehousing. It appears that all these approaches were important for defining the data warehouse model, however, user-driven approach, despite its importance and, in this case study, helped essentially to validate and identify KPI's, suffers from the difficulty that users have to think beyond the limits of existing technology or processes. Of course this can be minimized if potential data warehousing users were more experienced in using information analytical tools, even if they are basic, such as pivot tables from spreadsheets. In this case study, due to the inexperience of users, they did not realize the capabilities of the system until they have been able to try it, failing to clearly state their informational requirements (business indicators, levels of detail of information and reporting).

In conclusion, the multi-driven approach with the combination of five stages allow bridging inexperience and difficulties that often occur in data warehousing projects and this guarantee that the data warehouse model is aligned with business strategies.

As future work, a recommendation to analyze each approach by identifying if it has positive or negative impact in the data model of the data warehouse, i.e., if it is positive, increments (more KPI's and information) the data warehouse model, but if it is negative, contract the (less KPI's and information) data warehouse model. For instance, technological-driven approach has negative impact in the data model of the data warehouse, but is important to measure this impact.

REFERENCES

- Bulos, D., 2000, *Getting started with ADAPTTM*, Retrieved May 14 2003, from http://www.symcorp.com/tech_expertise_design.html.
- Chenoweth, T. et al, 2006, Seven Key Interventions for Data Warehousing Success. *Communications of ACM* 49(1), pp. 115-119.
- Guo, Y., S., et al, 2006, *Triple-Driven Data Modeling Methodology in Data Warehousing: A Case Study*, In Proceedings of the International Workshop on Data warehousing and Online Analytical Processing (DOLAP), Arlington, VA, USA.
- Herrmann, C. and Melchert, F., 2004, *Sponsorship models for data warehousing: two case studies*. In Proceedings of the AMCIS. New York City, USA.
- Inmon, W. H., 2005, *Building the Data Warehouse*, 4th Edition, Wiley Publishing, Inc, the University of Michigan.
- Kaldeich, C. 1995, *Congruence relations in relational databases: incomplete information*; The 2nd Workshop on Non-Standard Logic and Logical Aspects of Computer Science - NSL'95, Irkutsk, Russia.
- Kaldeich, C., Sá, J. O., 2004, *Data Warehouse Methodology: A Process Driven Approach*. In Proceedings of the 16th International Conference on Advanced Information Systems Engineering (CAiSE 2004), Riga, Letonia.
- Kimball, R., et al., 2008, *The Data Warehouse lifecycle toolkit: expert's methods for designing, developing, and deploying data warehouses*, 2nd edition, John Wiley & Sons, New York, USA.
- List, B. et al., 2002, *A Comparison of Data Warehouse Development Methodologies Case Study of the Process Warehouse*. In Proceedings of the 13th International Conference on Database and Expert Systems Applications, Aix en Provence, France.
- Malinowski, E. and Zimányi, E., 2009, *Advanced Data Warehouse Design From Conventional to Spatial and Temporal Applications*, Springer-Verlag - DCSA, Heidelberg, Germany.
- Moody, D. L. and Kortink M. A. R., 2000, *From Enterprise Models to Dimensional Models: A Methodology for Data Warehouse and Data Mart Design*, In Proceedings of the International Workshop on Design and Management of Data Warehouses (DMDW), Stockholm, Sweden.
- Scheer, A.-W., 2001, *Modellierungsmethoden, Metamodelle, Anwendungen*, Springer-Verlag; 4. Aufl. Edition, Heidelberg, Germany.
- Song, I-Y., et al., 2008, *SAMSTAR: An Automatic Tool for Generating Star Schemas from an Entity-Relationship Diagram*, In Proceedings of the 27th International Conference on Conceptual Modeling, Barcelona, Spain.
- Winter, R. and Strauch B., 2003, *A method for demand-driven information requirements analysis in data warehousing projects*, In Proceedings of the 36th Hawaii International Conference on System Sciences, HICSS-03, Hawaii, USA.