A business-oriented approach to data warehouse development

Enfoque orientado al negocio para el desarrollo del almacén de datos

A. Cravero Leal¹, J. N. Mazón² and J. Trujillo³

ABSTRACT
Several surveys have indicated that many data warehouses fail to meet business objectives or are outright failures. One reason for this is that requirement engineering is typically overlooked in real projects. This paper addresses data warehouse design from a business perspective by highlighting business strategy analysis, alignment between data warehouse objectives and a firm’s strategy, goal-oriented information requirements’ modelling and how an underlying multidimensional data warehouse model may be derived. A set of guidelines is provided allowing developers to design a data warehouse aligned with a prevailing business strategy. A classic case study is presented.

Keywords: Data warehouse, business-oriented, alignment.

RESUMEN
Varios trabajos indican que un porcentaje significativo de los almacenes de datos no cumplen con los objetivos del negocio o que son fracasos rotundos. Una de las razones consiste en que la ingeniería de requerimientos es pasada por alto en proyectos reales. En este trabajo el diseño del almacén de datos es abordado desde un punto de vista empresarial, en el que se da relevancia al análisis de la estrategia del negocio, la alineación entre los objetivos del almacén de datos y la estrategia de la empresa, un modelo de objetivos que permite obtener los requisitos de información y la derivación del modelo multidimensional del almacén. Proveemos un conjunto de directrices que permite a los desarrolladores diseñar un almacén de datos alineado a la estrategia del negocio. Se presenta un caso de estudio clásico.

Palabras clave: almacén de datos, orientado al negocio, alineamiento.

Received: March 20th 2012
Accepted: March 12th 2013

Introduction
For years now, IT executives have considered information system (IS) alignment as a top priority for an enterprise’s strategy (Thevenet et al., 2006). Regarding strategic alignment, Bleistein (2006) has stated that those defining an enterprise’s strategy (i.e. enterprise executives) do not speak the same language as IS engineers and thus they do not have the same vision of such enterprise, thereby hindering good communication between both “worlds.” Consequently, an IS does not provide the expected value for a particular enterprise (Ellis-Braithwaite et al., 2012).

The situation is no different for data warehouse (DW) projects because they are used for supporting decision-making concerned

with analysing an enterprise’s state of development (Gupta et al., 2011).

It follows that analytical business-oriented requirement models must be used for attaining a particular system’s goals from a user viewpoint to focus an organisation’s strategy in a structured way (Simonin et al., 2012).

Our previous work has used goal-oriented models based on the i* framework (Yu, 1995) to define a requirement analysis stage for DWs. The new approach considered business strategy using vision, mission, objectives, strategies, tactics (VMOST) requirements analysis (Bleistein et al., 2006), a business motivation model (BMM) (Object Management Group-OMG, 2010) for aligning DW objectives and organisational strategy, modelling such objectives with i* (Mate et al., 2011) and deriving the underlying multidimensional model of the DW by means of a unified modelling language (UML) profile (Luján-Mora S., 2002).

A business-oriented approach to data warehouse development

Our business-oriented approach to DW development was focused on defining the goals which an organisation must achieve and the relationships amongst stakeholders which would be needed for them to become fulfilled (see Fig. 1). These goals were the main objectives an organisation wanted to achieve by

using a DW and could be classified into strategic, decision-making and information goals. Information requirements were derived from these business goals as information provided by the DW for achieving business goals. Such business goals and information requirements should be represented in an organisational model together with DW users and the necessary relationships between the DW and its users needed to achieve business goals.

Requirements analysis for DW

An IS' requirement analysis phase must deal with analysing, understanding and modelling the business context in which it works. Regarding DW, this phase is aimed at representing DW users, the business goals of the organisation in which a DW is integrated and the relationships between the DW and its users for achieving business goals. This phase is crucial for DW development, as stakeholders usually do not know how to describe information requirements and the DW may not reflect business needs (Golfarelli, 2010). This phase thus allows developers to situate the DW within its business context and relate it to business goals. Designers can then more easily obtain what the DW users need to do with the DW system to achieve business goals (i.e. information requirements) (Sarkar, 2012).

Approach for modeling the DW

The proposal for DW design using business strategy complements Cravero’s work (2010) which included business strategy softgoals (see Figure 1); compared to hard goals, softgoals may be highly subjective and strictly related to a particular context as they enable analysts to highlight quality issues from the outset, making explicit the semantics assigned to them by the stakeholders (Donzelli, 2004).

Business strategy analysis

The first step in strategic alignment of an enterprise is to establish its business strategy (Singh and C.Woo, 2009); the objective of a strategy (the "what") is to define the business’ goals and objectives.

VMOST analysis (Sondhi, 1999) provides a means of deconstructing business strategy into the core components of vision, mission, goals, strategies, objectives and tactics. It aids to understand how these components relate to, align with and provide support for each other by an analyst’s response to a number of key questions presented in Bleistein (2006).

Alignment using BMM

The BMM is the OMG’s specification for modelling business plan elements. These motivate and provide business governance and guidance (Andersson et al., 2007). The BMM model helps in modelling organisation systems so they align IS with the business and provide support for an organisation’s business motivation (Andersson et al., 2007; Berkem, 2006). Based on the same strategic planning components as VMOST, the BMM model describes a framework in which organisational “means” achieve organisational “ends”.

BMM provides business vocabulary which can be understood by business stakeholders for developing, communicating and managing business plans in an organised manner (OMG, 2010). Figure 2 illustrates BMM elements and their relationships to support a business with means to achieve its ends. Means include the mission, course of action (i.e. strategic and tactical, deriving the goal, tasks and resources) and directives (i.e. business policy and business rules, deriving the softgoal) governing the course of action and supporting the achievement of ends. Ends include the vision and desired result (i.e. goal and objective).

Objetives model for DW

The i* modelling language represents the actors, dependencies and relationships with the objectives to be achieved according to a business strategy (Yu, 1995).

This technique consists of two models: strategic dependency (SD) to describe the dependency relationships between different actors within their organisational context and the strategic rationale (SR) model which is used to describe each actor’s interests and concerns and how they might be addressed (Yu, 1995). The central concepts in i* models are the actors’ intentions and
their dependency relationships for achieving their goals. In a DW model, users are agents of the enterprise depending on a DW system’s tasks or information provided by this (resource) to achieve their own goals (Mazón et al., 2007b).

SD models express dependency using four main types of dependency links. For goal dependencies, the depender depends upon the dependee to bring about a certain state in the world. A goal represents a condition or state of the world which might be achieved. For task dependencies, the depender depends upon the dependee to carry out a task. A task represents a particular way of attaining a goal. It can be considered as a detailed description of how to accomplish a goal. For resource dependencies, the depender depends upon a dependee for the availability of an entity. Resources may be physical or informational and considered as the availability of some entity or even the finished product of some action, process or task. For softgoal dependencies, the depender depends upon the dependee to perform some task that meets the softgoal or non-functional requirement, or to perform the task in a particular way (Yu, 1995).

We have also used i* SR models for modelling the intentionality associated with each actor and to decompose and make testable actor-dependent softgoals modelled in the SD model.

Multidimensional conceptual schema

Each task and resource reflected in the SD model for the actor DW may be treated in a multidimensional model MD (Mazón et al., 2007b).

Figure 1 uses a UML profile proposed by Luján-Mora (2002) for conceptual DW design. The most important feature of the multidimensional paradigm is to divide the data into facts (composite of measures) and dimensions to provide data on the required level of granularity. The UML profile is defined by a set of stereotypes and tagged values and elegantly represented by these major properties at the multidimensional conceptual level by a class diagram (see Table 1). Due to lack of space we refer to Mazón (2005) for a more detailed explanation.

Table 1. The main UML profile stereotypes (Mazón, 2005)

<table>
<thead>
<tr>
<th>Stereotype</th>
<th>Description</th>
<th>Icon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fact class</td>
<td>Represents Facuss consisting of measures</td>
<td><img src="image" alt="Fact Stereotype Icon" /></td>
</tr>
<tr>
<td>Dimension class</td>
<td>Represent dimensions consisting of dimension attributes and hierarchy levels</td>
<td><img src="image" alt="Dimension Stereotype Icon" /></td>
</tr>
<tr>
<td>Base class</td>
<td>Represent dimension hierarchy levels</td>
<td><img src="image" alt="Base Stereotype Icon" /></td>
</tr>
</tbody>
</table>

Transformation begins with VMOST analysis. Once the strategy components (from the future DW users) have been obtained, alignment is checked by using BMM rules. Modelling continues by constructing model goals based on i* (SD and SR models). Each task and resource embedded in an SD model for an actor DW should be incorporated in a model MD.

How to build SD and SR models for an MD model of DW has been proposed in Franch et al.’s prescriptive SD guide (2007). This approach covers generic steps from actor discovering to the inclusion of software actors (as a DW) to specify a proposed socio-technical system. This approach was adopted because it leads to well-formed i* models.
A case study

Our case study consisted of a retail company (Kimball and Ross, 2002) which included both retail sales and inventory. The retail business had several grocery stores spread over several geographical regions. Profit came from attracting as many customers as possible in a highly price-competitive environment. This resulted in a policy of “attracting as many customers as possible”, so one way of measuring success would be to determine the number of new customers for each campaign producing more than $1,000 profit.

This case study dealt with various types of inventory models of a shop. Interest was placed on the right on time inventory which measured each day's inventory level, records being kept separately in the database. The main goal of management lies in making decisions to optimise inventory levels to decrease costs. These decisions are meant to ensure that the right product is in store just at the right time to minimise out-of-stock items (if the products are not available on the platform to be sold) and thus reduce the overall cost of transport for the inventory. The daily levels of inventory at hand (by product and by store) must thus be analysed for good inventory management. The inventory manager is also responsible for measuring the speed of inventory movement (how a product moves through the store) to ascertain sales profits. The manager needs to know the gross margin on inventory (GMROI).

Business strategy analysis guidelines

Guideline 1: Business actors had to be discovered. These actors were decision-makers (e.g. managers and senior executives). The DW being constructed was also considered as an actor. Possible actors were “marketing manager”, “inventory manager” and “DW”.

Guideline 2: VMOST analysis provided the mission, vision, strategies and goals. The answers to questions 1 through 6 were:

For example, Question 1: What is the overall, ideal, end-state towards which an organisation strives (vision)? In response to question 1, an executive described his vision of creating a chain of grocery stores where the cost of their inventories would be reduced to meet customer needs on request and increase the number of customers by improving revenue.

Question 6: What goals does an organisation set to determine whether it is competing successfully? Increasing the number of customers represents a marketing manager’s strategic objective and reducing inventory costs is an inventory manager’s strategic objective. Each strategic objective will be represented as a dependency of the actor DW.

Guideline 3: Intentional elements, such as objectives and tasks, should be obtained for each DW actor (decision-maker). A specific actor’s new business goal could mean discovering new business dependencies and hence new strategic or tactical actors. Several guidelines are given for this process:

Guideline 3.1: Questions 7 and 8 from VMOST analysis were used to identify tactics and decision objectives to support decision-making. Relationships between them were represented as means-ends links because such links were used to describe how objectives were achieved.

Question 7: Which activities does an organisation perform to achieve the goals set out in response to Question 6? Some key questions relevant to such objectives would be:

- How can the number of customers be increased? Through determining effective promotion allowing customers to sample new products.
- How can inventory costs be reduced? By taking steps to optimise inventory management.

A decision goal (determine effective promotion) was thus discovered for the following strategic goal: to increase the number of customers. A target decision was found (optimise inventory management) from the following strategic goal: to reduce inventory costs.

Guideline 3.2: VMOST analysis (question 9) was used again for information objectives. Each object was represented as means-ends links.

Question 9: What are the measurable objectives indicating that the goals identified in Question 6 have been achieved? What activities does the organisation perform to achieve such objectives? Once again, further questions had to be asked to obtain such objectives:

- How can effective promotion be determined? By analysing previous promotions.
- How can inventory management be optimised? Required inventory levels and inventory flows need to be analysed.

These were represented as objects linked to decision objectives in the SR model.

Guideline 3.3: Each objective of previously obtained information was related to the analysis of some measurements used to achieve these goals. The analysis that described a reporting requirement was represented as a task. Decision-makers carried out these tasks to get information from the DW since it was necessary to achieve information goals. VMOST question 10 had to be asked.

Question 10: How do the objectives identified in Question 9 support the goals identified in Question 6? Analysing a promotion meant determining the amount of products sold in stores, on which days and in what promotional conditions. An important measurement was the amount of products sold, thereby representing a task: to analyse the quantity sold. The retailer needed to analyse daily levels of product available in the store (the amount at hand) to examine the levels of inventory to ensure that the product was in the store just at the right time. The amount of goods in the store had to be measured for certain periods of time. A task called “analyse the quantity on hand” had to be represented in the DW. The GMROI had to be analysed for “the study of inventory movements” (as described above), so a task called “analyse GMROI” was created. The linking of means and ends was represented in the SR model between each task and goal decision (as shown in Figure 5).

Guideline 4: Business policy and rules had to be ascertained. The business determined the policy of “keeping good customers as much as possible.” This might have resulted in policies like
classifying customers according to the money they spent, resulting in special promotions for better clients (i.e. those who spent most).

**Guidelines for alignment using BMM**

**Guideline 5:** The alignment of goals obtained through BMM rules was checked. There were several guidelines for such alignment:

**Guideline 5.1:** The alignment of the components related to the media was checked. The strategies defined in step 2 had to be components of the mission. The business rules should have produced an effect on tactics. Increasing the number of customers is a marketing manager’s strategic goal and reducing inventory costs is an inventory manager’s strategic goal. They were also aligned with the mission and would be implemented through a strategic IS.

**Guideline 5.2:** The alignment of the goals so obtained was checked (purposes). The vision was amplified and it had to be made sure that informational goals were measurable and delivered information about the strategic goals. The strategic goals were increasing the number of customers and reducing inventory costs, aimed at improving inventory management business, making every convenience store business. This also happened with decisional goals like sales information and inventory information. Information goals were indeed measurable targets since they had sufficient information to analyse historical developments, review inventory levels and study the flow of inventories.

**Guideline 5.3:** Verifying the mission operationalised the vision, the strategic objectives drove effective business strategy and the decisional objectives formulated decision-making tactics. By providing IT strategic information for handling business, the mission would allow stores to keep an inventory and should minimise their costs. Decision-making goals such as to determine effective promotion allowed creating medium term tactics.

**DW objective model guidelines**

**Guideline 6:** The SD model was obtained; the information required by decision-makers was represented as resource dependency between each actor and the DW, since this information was provided by the same DW (Mazón, 2005).

In this case study information regarding sales and inventory information was drawn in the SD model as actor DW resource dependency (see Figure 3).

**Related work**

This section gives a brief description of the most relevant approaches for requirement analysis in DW. Prakash (2004) has described DW elicitation and proposed requirement grouping requirements at several levels of abstraction. They identified...
information for supporting decision-making via information scenarios; an organisation’s goals had to be determined. Decision-making needs were specified and the information needed to cover such decisions was identified (a goal decision information (GDI) diagram was used).

Figure 6: Class diagram created from the multidimensional goal model

Paim (2003) described a data warehouse requirement definition (DWARF) technique. They adapted traditional requirement engineering, even capturing a non-functional requirement such as performance or accessibility, using the NFR framework. They proposed a methodological approach for requirements definition and DW management; however, they focused on technical issues (e.g. how to access data) but did not explicitly obtain business goals.

Giorgini (2008) presented a GRAnD approach for deriving a conceptual multidimensional schema. They proposed gathering multidimensional requirements and later mapped them onto the data sources in a conciliation process. The authors introduced an agent-oriented methodology based on the i*. They argued that it was important to model the organisational setting in which a DW would operate (organisation modelling) and capture DW functional and non-functional requirements (decisional modelling).

Mette (2012) presented an extended version of the critical success factors method for establishing a first version of a demand-driven requirements specification for a DW. The purpose was to help managers and consultants develop DW that would really satisfy managers’ needs for data to improve decision-making.

Sarkar (2012) presented a framework starting by identifying the analytical requirements. The framework consisted of three phases: early requirements analysis, detailed requirements analysis and mapping. DW requirement descriptions were mapped using high level design components from a conceptual object-oriented multidimensional data model.

Cravero and Sepúlveda (2012) made a chronological study of the different approaches to DW design.

The aforementioned approaches paid little attention to providing a mechanism for understanding the business context in which a DW would be deployed. They did not provide a clear set of guidelines for DW requirement analysis and they were not based on a global methodology from which the corresponding conceptual multidimensional schema could be directly obtained for achieving business goals. They also did not incorporate DW goal phase alignment with business strategy.

Conclusions

Requirements analysis plays a key role in DW system development for reducing the risk of failure. A good DW design method should be preceded by requirements elicitation and analysis considering user requirements; a business-oriented approach for DW development was thus devised.

The approach consisted of four phases: VMOST-based business strategy analysis, analysing the elements obtained using BMM to align business strategy with DW, a DW conceptual model using i* and a multidimensional model using a UML profile.

The research so far has made a number of contributions in the field of requirements engineering for DW. It has provided an approach to requirements engineering-oriented business that includes an analysis of business strategy for designing a conceptual multidimensional model of DW in accordance with an organisation’s objectives. While some research has been based on a model of objectives ours is the first investigation of requirements which has used strategic business activities incorporating a technique known strategic business analysis. A validation guide to DW requirements with business strategy through a set of traceable links has also been provided. BMM can be used for a better understanding of who, what, why and how regarding business requirements, thereby enabling DW designers to provide business executives with answers to their DW-based questions. A case study has illustrated business alignment.

Adopting the above approach should help designers to reduce the risk of project failure by ensuring that early requirements are truly needed by a particular business, ensuring a good design and, at the same time, ensuring that the resulting DW schema answers strategic decision-makers’ questions.

Future work will deal with developing a prototype tool for supporting the proposed DW requirements analysis framework. Quality evaluation of the proposed requirements analysis framework will also form a prime objective for such future work.

References


