Expert Systems with Applications 38 (2011) 1203–1212

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

ERP data sharing framework using the Generic Product Model (GPM)

Souleiman Naciri, Naoufel Cheikhrouhou *, Michel Pouly, Jean-Charles Binggeli, Rémy Glardon

Ecole Polytechnique Fédérale de Lausanne (EPFL), Laboratory for Production Management and Processes, Station 9, 1015 Lausanne, Switzerland

ARTICLE INFO

Keywords:
Generic Product Model
Management data
Enterprise Resource Planning
Information sharing

ABSTRACT

Nowadays, all product life cycle processes are investigated deeply in order to get an advantage over competitors. To support these processes, several software applications are available. However, this wide range of heterogeneous applications leads to a large variety of hardware and operating systems, data management software, data models, schemas, and data semantics that hinder the information sharing process. To tackle this problem, Hitachi Company has developed a modeling language called Generic Product Model (GPM) in addition to several translators from native formats to GPM for storing, sharing and visualizing product data in a single data warehouse. In order to broaden the range of application data able to be stored and shared using the GPM data warehouse, this paper presents a methodology and a translator that allow management data to be included in a GPM data warehouse. Given the fact that most management data are stored in Enterprise Resource Planning (ERP) software, and that the latter allows extraction of database contents into Excel files, we propose a translator that handles Enterprise Resource Planning management data that are already extracted into Excel Format.

The proposed framework enables data management contained in flat Excel Files to be translated into structured GPM data. Translation methodologies are given for the translation of mass production and customized product data through two case studies, one in the computer retail sector and the other in the extrusion machine industry.

1. Introduction

Due to financial, technological and environmental constraints, all the steps of the product life cycle management are nowadays deeply analyzed and optimized. Experts can rely on software applications dedicated to each specific field to support this activity. The variety of domain needs and development approaches to tackle a problem have led to the development of a large number of applications and formats which often have limited compatibility between each other. This heterogeneity becomes problematic as soon as people want to share data and information within a company or between different companies. These difficulties hinder collaborative efforts during the design, the management and the end-of-life steps of a product. To improve data sharing, different methods and tools have been developed, such as Electronic Data Interchange (EDI), ETL tools (Vassiliadis, Simitsis, Georgantas, Terrovitis, & Skiadopoulos, 2005; Vassiliadis, Vagena, Skiadopoulos, Karayannidis, & Sellis, 2001) and web-based part libraries (Li, Lu, Liao, & Lin, 2006). However, these solutions have some limitations as outlined in Section 2. To circumvent these limitations, the Japanese Hitachi company has proposed translating data from several CAD/CAM application fields into a single and long lasting format, called the Generic Product Model (GPM), based on STandard for Exchange of Product model (STEP), which can then be visualized in a web browser. This paper proposes a methodology for storing and sharing management data in a structured manner by designing and developing new translators between ERP data and the GPM structure.

The paper is structured as follows. In Section 2, a state of the art of the different technologies and works in data sharing and exchange, with a particular focus on ERP, is presented. According to this, motivations for developing new information sharing techniques are described. Section 3 details the GPM concept and the different features on which the current development relies. Section 4 is dedicated to the presentation of the global platform TECHNO-INFRA that hosts the different data and application layers. The structure of exchanged ERP data as well as the design and development of the information translator are addressed in Sections 5 and 6, respectively. Two implementation cases are presented in Section 7: a fictitious computer manufacturing company and a real Swiss manufacturing company in extrusion machine building. Conclusions are finally provided in Section 8.

2. Motivation and state of the art

The focus of this research work is put on management data that refer to materials-, products- and processes-related data used for

* Corresponding author. Tel.: +41 (0)21 693 4149; fax: +41 (0)21 693 3553.
E-mail addresses: souleiman.naciri@epfl.ch (S. Naciri), naoufel.cheikhrouhou@epfl.ch (N. Cheikhrouhou), michel.pouly@epfl.ch (M. Pouly), jean-charles.binggeli@epfl.ch (J.-C. Binggeli), remy.glardon@epfl.ch (R. Glardon).

0957-4174/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.05.013

production planning and control purposes. They also include information that is *technically* oriented, such as bill of materials (BOM) or operation sequences and others that are *management* oriented such as inventory levels, safety stocks and delivery times. In the early 1970s, these data were stored in MRP (Material Requirement Planning) and MRP II (Manufacturing Resource Planning) software that was used for building production and purchasing plans. Since the 1990s, management data are stored and managed within software called ERP (Enterprise Resource Planning), that host not only management data, but also data related to other fields, such as human resources management and accounting. These commercial software packages are defined by Jacobs and Weston (2007) as frameworks for organizing, defining, and standardizing the business processes necessary to plan and control an organization effectively. Thus, ERP covers a wide range of processes which require the handling of huge amounts of data.

To share such data, several techniques have been developed and used for a while. Indeed, with the advent and democratization of the World Wide Web, techniques such as Electronic Data Interchange (EDI) appeared. EDI is defined by Holland, Lockett, and Blackman (1992) as the "process of computer-to-computer, business-to-business data transfer of repetitive business processes involving direct routing of information from one computer to another without human interference, according to predefined information formats and rules". While this technique is useful for sharing information such as quotation requests, order confirmations and invoicing, it shows some limitations in terms of modularity when the purpose is to share specific data that are not included in existing standardized transactions. With the development of XML (eXtended Markup Language), a common framework was provided to structure data, especially using DTD (Document Type Definition) and XML schemas (Huang & Lin, 2010). Although there is the presence of tags that are supposed to provide information about the contents of XML files, misunderstanding can still happen and hinder the information sharing process. Another approach to sharing information between two or more applications consists of extracting data from the source application(s), transforming it and loading it into the target application or a data warehouse. These powerful tools called ETL (Extract, Transform and Load) are widely used and are able to transfer frequently huge amounts of data. However, these tools often limit their translation capabilities to formats that are based on their own technology. This is the case with SAP Business Warehouse (BW), which is successful for gathering data from other SAP components based on mySAP technology, but cannot handle data coming from other proprietary formats such as Microsoft Access and Excel. When they are not limited to a proprietary format, ETL tools provide capabilities for handling data from heterogeneous applications and translating them into a specific format. However, no recommendations or propositions are made for using a generic model language that may handle data originally stored in heterogeneous applications. Several other approaches focusing on linking and using primary existing ERP databases in different enterprises of the supply chain have been developed (Stefansson, 2002; Themistocleous, Irani, & Love, 2004).

Different frameworks are proposed in order to rely upon ERP databases and functionalities to share and manage data used for management purposes (Nurmilaakso, 2008; Pramatari, 2007). However, none of these technologies claim to overcome all integration problems but rather need to be pieced together to support linking diverse heterogeneous applications that often exist within supply chains. Huang and Lin (2010) propose a solution to share knowledge within a supply chain using a semi-structured knowledge model and ontologies. Unfortunately, this approach requires supply chain partners to use ontologies and additionally requires that a knowledge engineer validates the automatic mapping between similar concepts.

Other techniques for sharing information have been proposed, such as the common data storage method and the use of ontologies developed in (Oh & Yee, 2008) where the objective is to give a formal and agreed description of concepts and their relationships in order to share data in a consistent way (Lee, Chae, Kim, & Kim, 2009; Ziegler & Dittrich, 2004). However, the main challenge remains the data mapping problem. Several methods seek to achieve automatic data mapping, such as the Similarity Flooding algorithm (Melnik, Garcia-Molina, & Rahm, 2002) or transformations heuristic search (Fletcher, 2005) to identify similarities between source and target data sources. However, in the frame of this research, automatic mapping is not the appropriate approach for two main reasons. First of all, the final goal is not to map a whole company' database to another, but rather to share specific sets of data from a company to a shared data warehouse and, thus, avoid information management overload. The second characteristic concerns the nature of the data. Indeed, it is assumed that ERP data have firstly been extracted in a specific file using Excel. This means that the data which are originally structured in the legacy systems are flattened in an Excel file table. Consequently, not only heads of columns may be opaque but also most of the attributes contained in the legacy database are neither correctly nor completely transferred.

As a consequence, neither linguistic nor schema structure recognition can be used to automate the mapping process efficiently.

On the operations management side, the literature has shown the positive impacts of synchronization of supply chain partners by using shared information. Indeed, several research works have outlined the benefits of information sharing between partners throughout the supply chain (Cachon & Fisher, 2000; Lee, So, & Tang, 2000; Lee & Whang, 2000; Shuai, Yi-Fen, & Chyan, 2007). Sharing data and information such as machine loads, points of sales or inventory levels has proven to increase companies and supply chain key performance indicators such as the fulfillment rate and the product cycle time (Chen, Yang, & Yen, 2007) and to decrease order fluctuations (Croson & Donohue, 2003) that characterize the bullwhip effect. Moreover, Kelle and Akbulut (2005) identify the role of ERP tools in enhancing data sharing and propose a quantitative model able to calculate the potential monetary value of policy coordination, to promote cooperation, and minimize the total supply chain system cost. Su and Yang (2010) propose a structured equation model for analyzing the impact of ERP on firm competencies within a supply chain. The positive impact of information sharing on supply chain performance has no longer a need to be proven. However, the tools engaged in the information sharing process still have some limits that are addressed and circumvented in this work.

In this paper, the proposed approach to integrate several applications' data is to translate and load the information contained in heterogeneous applications into a single data warehouse where all the information is translated and stored in a common language: the Generic Product Model (GPM). By doing so, people continue to work on the company side on their usual applications, independently from the standards and tools used by the partner. Then, once engineers want to share information within the company or with the company's partners (suppliers or customers), they put the relevant data in a common data warehouse by using translators that are able to match the structure of each single application with the GPM structure. Then, all the information is gathered, integrated and reachable by all the authorized people. These pieces of information can then either be visualized within the data warehouse using a web-based viewer, or by downloading the relevant data into the corresponding applications using reverse translators. Several translators have already been developed within the IMS-VIP-NET (Intelligent Manufacturing Systems-Virtual Production Enterprise Network). These translators handle either P&ID (Process