International Journal of Information Management 31 (2011) 366–373

Contents lists available at ScienceDirect

International Journal of Information Management

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

Collaborative management of business metadata

Kai M. Hüner*, Boris Otto, Hubert Österle

Institute of Information Management, University of St. Gallen, Müller-Friedberg-Strasse 8, 9000 St. Gallen, Switzerland

ARTICLE INFO

Article history:
Available online 17 January 2011

Keywords:
Collaboration
Data quality management
Design science research
Metadata management
Semantic wiki

ABSTRACT

Legal provisions, cross-company data exchange and intra-company reporting or planning procedures require comprehensively, timely, unambiguously and understandably specified business objects (e.g. materials, customers, and suppliers). On the one hand, this business metadata has to cover miscellaneous regional peculiarities in order to enable business activities anywhere in the world. On the other hand, data structures need to be standardized throughout the entire company in order to be able to perform global spend analysis, for example. In addition, business objects should adapt to new market conditions or regulatory requirements as quickly and consistently as possible. Centrally organized corporate metadata managers (e.g. within a central IT department) are hardly able to meet all these demands. They should be supported by key users from several business divisions and regions, who contribute expert knowledge. However, despite the advantages regarding high metadata quality on a corporate level, a collaborative metadata management approach of this kind has to ensure low effort for knowledge contributors as in most cases these regional or divisional experts do not benefit from metadata quality themselves. Therefore, the paper at hand identifies requirements to be met by a business metadata repository, which is a tool that can effectively support collaborative management of business metadata. In addition, the paper presents the results of an evaluation of these requirements with business experts from various companies and of scenario tests with a wiki-based prototype at the company Bayer CropScience AG. The evaluation shows two things: First, collaboration is a success factor when it comes to establishing effective business metadata management and integrating metadata with enterprise systems, and second, semantic wikis are well suited to realizing business metadata repositories.

1. Introduction

1.1. Motivation and problem statement

A clear and unambiguous understanding of data is indispensable for the effective management of multidivisional companies. At the same time, clear and unambiguous definitions of business objects (e.g. materials, customers, suppliers) are imperative for implementing and benefiting from integrated, automated business processes (Kagermann, Österle, & Jordan, 2010). Fundamentally, smooth and efficient business process execution is ensured by metadata describing the data structure of business objects (i.e. their attributes and relations) and providing information for correct usage of business objects in business processes. High-quality metadata (i.e. metadata that is up to date, accurate and complete) helps to create and maintain a common understanding of business objects and business processes (Schmidt & Otto, 2008). It minimizes the occurrence of errors in automated activities and reduces waiting times between activities. The following three examples illustrate how poor metadata quality may negatively affect companies' business process performance.

DB Netz AG (40,300 employees, EUR 4.1 billion revenue in 2009), a subsidiary of Deutsche Bahn AG, is responsible for Germany's railway infrastructure. The company has to make sure that its entire infrastructure is accurately kept in an inventory database in order to be able to report to national authorities on an annual basis for the purposes of obtaining public funding. This task is impeded by the fact that there is no company-wide definition of basic business objects (i.e. tracks, switches, and tunnels). For example, while a tunnel in one geographical region is considered as an underground track segment, with the tunnel length being the distance between tunnel entrance and tunnel exit, in another region "tunnel length" refers to the distance between two underground stations. Moreover, crucial knowledge (e.g. on the objects' geographical position, maintenance, and monetary value) is distributed across the entire company. As a result of these problems, any consolidation of infrastructure data needed for inventorization is virtually impossible.

SBB Cargo AG (3700 employees, CHF 915 million revenue in 2009), a subsidiary of Schweizerische Bundesbahnen, is responsi-

* Corresponding author. Tel.: +41 71 224 2420; fax: +41 71 224 2777.
E-mail addresses: kai.huener@unisg.ch (K.M. Hüner), boris.otto@unisg.ch (B. Otto), hubert.oesterle@unisg.ch (H. Österle).

0268-4012/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijinfomgt.2010.12.002

ble for Switzerland's railway freight transportation. The trains of SBB Cargo run on tracks that are used by a certain train type within a given period of time. To book a track, SBB Cargo only needs to provide the train type used and the period of time. No further attributes are necessary to specify the business object from a data exchange perspective. For internal purposes (e.g. specifying transportation orders, identifying dangerous goods to be transported, putting in place work schedules for the staff), additional attributes are relevant, the values of which are set by various information systems at various points in time. In order to be able to achieve flexibility in planning and short response times (e.g. in case of disturbance or malfunction), an unambiguous model of the business objects (e.g. track, train, train conductor) is needed that is understood in every department of the company. Another important requirement relates to business object definitions of business partners, since during any one trip the trains of SBB Cargo may run on tracks of various providers in various countries, making it necessary to take into account different business object definitions when a track is booked.

Bayer CropScience AG (18,700 employees, EUR 6.5 billion revenue in 2009), a market leader in the field of agrochemical products (e.g. for plant protection, insect control, seed growing), is a business division of Bayer AG. Bayer CropScience is conducting a worldwide business process harmonization project, in the course of which master data is to be consolidated. Currently, the structure, meaning and usage of the company's master data vary from region to region. For example, materials are counted by weight in Europe, and by item (e.g. bottle, box) in other regions. In order to enable corporate analyses (e.g. spend analysis, demand planning), data structures and business metadata have to be harmonized as well.

1.2. Research objective and paper structure

A common feature of the cases described in Section 1.1 is that the knowledge needed to create consistent business metadata is spread across various departments, divisions or lines of business. One possible solution to this problem could be to collaboratively identify, record, discuss and consolidate knowledge on various and differing aspects of business objects (e.g. technical features, accounting attributes, regional characteristics in the case of DB Netz's "tunnel") with the help of business experts. To do so, the use of a collaboration tool following the concept of Wikipedia[1] (i.e. a software tool supporting an iterative and collaborative approach for building a repository that any user can edit) could facilitate the process of collaboratively creating and maintaining consistent business metadata.

A comparison of the cases described in Section 1.1 with Wikipedia indicates the potential to use Wikipedia as a model for solving the problems described in the cases. Wikipedia even describes concepts like "tunnel", "train" or "train conductor", and provides collaboratively maintained and commonly accepted definitions. If the maintenance of Wikipedia was restricted to employees of DB Netz for example, the provided description of "tunnel" could be used as an accepted, company-wide definition. The question is, however, whether collaboration effects brought about by software similar to the Wikipedia platform can also be accomplished in a business context. Therefore, the research objectives are (a) to identify business requirements to be met by a repository which supports the collaborative management of business metadata, (b) to implement a wiki-based business metadata repository, and (c) to evaluate its use in a business context.

[1] Wikipedia's English website: http://en.wikipedia.org.

The next section of the paper defines business metadata according to literature on the subject. Furthermore, advantages and disadvantages of wikis to be used for creating business metadata repositories are discussed, and some scientific findings on the use of wikis in business contexts are presented. Section 3 specifies the research approach on which the paper is based and the techniques used for identifying and evaluating requirements to be met by a business metadata repository. Section 4 deals with the design of a business metadata repository, starting with the identification of business requirements, followed by the derivation of generic usage processes from these requirements, and finally leading to the implementation of a wiki-based prototype that meets the requirements and supports the usage processes. Section 5 presents the results of a requirements evaluation with business experts from various companies and users at Bayer CropScience using a wiki-based prototype. Finally, Section 6 presents a summary of the paper and gives recommendations for future research.

2. State of the art

2.1. Business metadata management

Metadata describes other data, i.e. it gives this data a certain meaning in a certain context. For example, a stick has a length (a physical property) of two meters (a measurement unit). Here, for the description of an object (the stick) a data item (the number "2") and two metadata items ("length" and "meters") are given (Sen, 2004, p. 151). Like the stick in this example, a business object is a real or virtual element in a company's process of providing services or products (e.g. a purchase or production order, a contract, a business partner, a certain material) that is characterized by a number of properties represented by data elements (Martin, 1975). Metadata referring to a business object describe its characteristics, providing both technical aspects (e.g. data structure, data distribution) (Tozer, 1999, pp. 184–194) and business-related aspects like information regarding correct use of the business object in certain business processes (Burnett, Ng, & Park, 1999, p. 1213; Marco, 2000, pp. 49–52). Table 1 gives an overview of the business objects, data attributes, and metadata mentioned in the examples given in Section 1.1, e.g. length of a tunnel in the case of DB Netz. In this context of metadata describing business objects, a business metadata repository is an information system that supports the collaborative management of business metadata and that aims at describing business objects comprehensively, unambiguously and understandably (Schmidt & Otto, 2008, p. 215; Schreiber et al., 2000, pp. 14–19).

2.2. Systems for collaborative metadata management

Due to the success of Wikipedia, companies increasingly express their desire to use wikis for business purposes (Bughin, Chui, & Miller, 2008, p. 2, 2009, p. 14). As research has shown, the successful use of wikis in companies is basically possible (Majchrzak, Wagner, & Yates, 2006), mainly as effective tools for specification processes (Wagner & Majchrzak, 2007). However, the development of Wikipedia has shown that unstructured metadata, as created and used in a classical wiki, may lead to redundancies and inefficient usage of wiki content (Krötzsch, Vrandečić, Völkel, Haller, & Studer, 2007).

In general, a wiki is a website providing content which can easily be edited by any (registered or anonymous) user (Cunningham, 1995). Depending on the software used, a wiki may come with different functionality. There are two functions, however, that can be found in practically any wiki (Kane & Fichman, 2009, p. 2):