Linked Data as Integrating Technology for Industrial Data

Markus Graube, Technische Universität Dresden, Germany
Johannes Pfeffer, Technische Universität Dresden, Germany
Jens Ziegler, Technische Universität Dresden, Germany
Leon Urbas, Technische Universität Dresden, Germany

ABSTRACT

In a globalised world the process industry faces challenges regarding data management. Rising demands for agility and rapid shortening of innovation cycles have lead to project-based collaborations. Highly specialised small and medium enterprises are forming “virtual companies” for their mutual benefit. However, today’s industrial data structures are very heterogeneous, complicating collaborative work and hindering the flow of data between stakeholders from different domains. Existing solutions are too rigid and potentially cumbersome. A broad gap still exists between the need of virtual companies to share data from mixed sources in a controlled way and the technologies available. The authors’ approach uses semantic web technologies to represent industrial data in a generic way. Major advantages in comparison to traditional approaches arise from the inherent merging abilities and the extensibility of Linked Data. Distributed information spaces from different domains can be condensed into an interlinked cloud. Existing data can be integrated either on-the-fly using appropriate adapters or by complete migration. Furthermore, operations from graph theory can be performed on the Linked Data networks to generate aggregated views. This article discusses a set of proven web technologies for cloud-driven industrial data sharing in virtual companies and presents first results.

Keywords: Graph Theory, Information Science, Knowledge Management, Linked Data, Meta-Ontologies, Semantic Technology, Virtual Companies

INTRODUCTION

Due to ever shorter release-cycles and growing pressure to innovate, independent small and medium enterprises are entering into short-term, project-based collaborations to develop and manufacture competitive products (AT&T Knowledge Ventures, 2008). In particular in fast moving markets, the concept of virtual companies (a loosely coordinated set of complementing companies that temporarily work together) promises an advantage over monolithic enterprises (van Heck & Vervest, 2007): “Agility is the name of the game in modern business” (Allemang, 2010).

For trustful collaboration and agility data must be shared quickly and easily. Consequently, corporate knowledge from all stakeholders and from various heterogeneous sources has
to be integrated. Usable access has to be provided to different types of databases but also to documentation, intelligent equipment and many other sources of data. At the same time, while virtual companies have a need to share some of their knowledge, there may be confidential information that needs to be protected.

We outline a novel approach that tackles the above challenges by leveraging semantic and web technologies beyond the concept of Hepp, Leymann, Domingue, Wahler, and Fensel (2005) which focuses only on Semantic Business Process Management. Extending our main idea and vision (Graube, Pfeffer, Ziegler, & Urbas, 2011) we present a highly integrated framework making it possible to access industrial engineering data through an interlinked semantic network. Our approach allows the use of SPARQL (Prud’hommeaux & Seaborne, 2008) for querying intelligent graph patterns. In addition, we propose universal endpoints for generic data views instead of propagating the creation of a collaborative portal like SPIKE from Broser, Fritsch, Gmelch, Pernul, and Schillinger (2009).

The remainder of this article is structured as follows. The next section introduces the relevant context of use and lists the prevalent ways of dealing with heterogeneous data in industrial applications. Afterwards, an approach is presented to overcome the stated challenges using Linked Data as the core technology. In the following section the technical structure of the proposed framework is explained. Finally, we present first results and discuss the proposed approach.

STATE-OF-THE-ART

Context of Use

In industrial process plants and on production sites, electrical, mechanical and hydraulic equipment is connected in a multitude of ways to form a productive system. Professionals from many different trades (among others mechanical, electrical & computer engineers) work together and share information. All these professionals may potentially work for different stakeholders – especially when virtual companies are considered (Thompson, 2008). At any time during the phases of planning, manufacturing and productive use of a plant, they may need collaborative access to different types of data, both sequentially and in parallel.

Industrial corporate environments have higher requirements for trusted and secure information exchange than the open web. Besides that, however, the need to coordinate large distributed information islands is quite similar to that of the public semantic web. Both have to make unstructured or semi-structured data available for their target users (Allemang, 2010).

Conventional Dealing with Heterogeneity

Conventional paths to integrate Enterprise Resource Planning systems (ERP), for instance after a merger, are usually very time-consuming. While several approaches to meet this challenge have been attempted in the past, there is still no commonly accepted standard for controlling and exchanging data over the companies’ boundaries of trust.

In the process and manufacturing industries, several approaches to a single unified world-model have been developed, for instance STEP which defines numerous domain-specific APIs (International Organization for Standardization, 1998, 2007) or more recently the Manufacturing Foundation Ontology from Usman, Young, Case, and Harding (2010). Yet, all these world models have turned out to be too complex to be easily applicable and too unspecific to meet the stakeholders’ individual requirements without significant customisation (Marquardt, Morbach, Wiesner, & Yang, 2010). So far, the most promising approaches seem to be meta-models, e.g., CAEX – Computer Aided Engineering Exchange (International Electrotechnical Commission, 2008). These provide sufficient formalism to allow transformations between particular implementations of different partners while maintaining flexibility during the life cycle of a product or a production plant.
Location Update Improvement Using Fuzzy Logic Optimization in Location Based Routing Protocols in MANET

Toward a Proof of Concept Implementation of a Cloud Infrastructure on the Blue Gene/Q