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

DOI: 10.4018/jdst.2012070104

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 Busineness 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.

Efforts have been made by groups such as the *Linking Open Data* movement (W3C, 2011) to give more structure to the data cloud through Linked Data. The basic idea is to make the large amount of publicly available information more useful by exposing it through a computer interpretable structure, based on web technologies. Arbitrary information entities can be identified by URIs (Uniform Resource Identifiers) so that people and computers can refer to them (Berners-Lee, 2006; Sauermann, Cyganiak, Ayers, & Völkel, 2008). Independent information spaces are interconnected by means of ontologies, mitigating the problem of ambiguities. While some issues still remain such as correctly detecting and fixing inconsistencies (Haslhofer & Popitsch, 2009), several commercial off-the-shelf and Open Source products are already available to retrieve information from large-scale distributed information spaces.

However, none of these efforts provide a solution that is tailored for virtual enterprises and is focused on industry needs. Despite of the huge amount of available content in the form of corporate knowledge, the complexity of search and missing end user applications still hinder broad usage in commercial domains (Latif, Höfler, Stocker, Saeed, & Wagner, 2009).

PROPOSED APPROACH

Forming a Cross-Enterprise Information Space using Linked Data

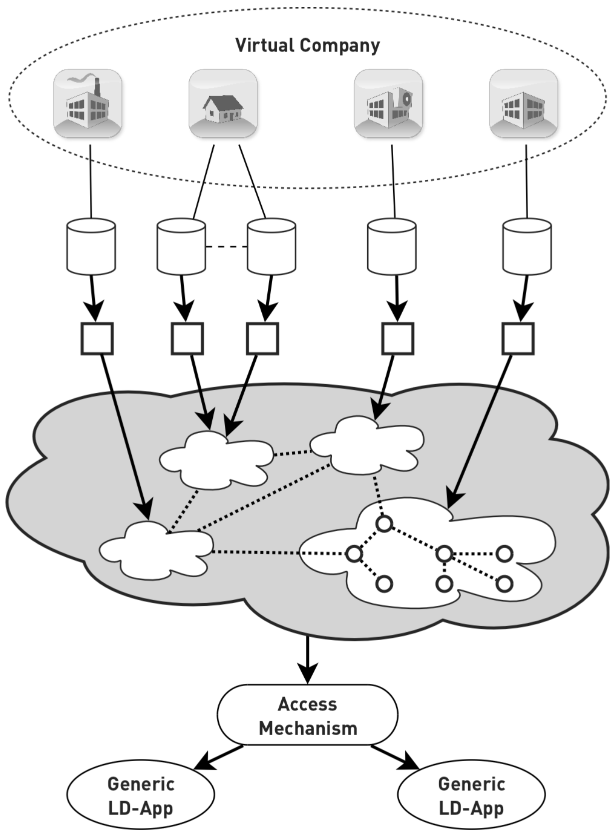
We propose the use of Linked Data to provide a single unified access point to all kinds of information in enterprises. Each data source with its concepts and domain specific semantics can be parsed into a semantic network with acceptable effort. All these semantic networks can be merged into a single Linked Data Cloud using appropriate adapters and converters. Data can be stored on several distributed servers building a decentralised network. The Linked Data Cloud has a controlled vocabulary and can be manipulated by means of graph theory to provide custom views of the data for different client applications. Clients can access the data via unified technologies with generic access concepts depending on the desired interaction pattern. The generic access concepts are able to operate on the same semantic network in parallel from any client. Security policies may be established to provide appropriate access control. The proposed methodology can be seen in Figure 1.

Linked Data and its Concepts

Linked Data is a promising technology for storing and providing structured data. It uses the principles and technologies of the semantic web (Bizer, Heath, & Berners-Lee, 2009) to publish, maintain and interlink data. Thus, all entities are referenced by URIs using the standard web protocol HTTP. Each entity can reference arbitrary other entities using arbitrary relations. This way, existing knowledge can be reutilised by linking to it which can substantially increase the value of modelled data. This makes it possible to connect various data sources and to build a decentralised, dynamic and extensible collaborative information space. Furthermore, the use of standardised languages enables computers to automatically read and interpret information so that applications can gather the desired information from different sources in a generic way.

Linked Data consists of a collection of triples expressed in the Resource Description Framework (RDF) which was developed by the W3C (2004). A triple is a simple semantic statement with a subject, predicate and object (Figure 2). Subjects and predicates have to be URIs whereas the object can either be a URI or a literal. One subject can have multiple predicates (an example is shown in Figure 2). Literals can, among others, be strings in different languages, numbers or dates. RDF itself is an abstract model and therefore has to be serialised for transferring and reading the data. The most common serialisation format is the machine-readable RDF/XML or XHTML with embedded RDFa. RDF can also be serialised in the RDF/N3 or Turtle format which are designed to be more

Figure 1. Domains use different concepts that can be harmonised by Linked Data technology

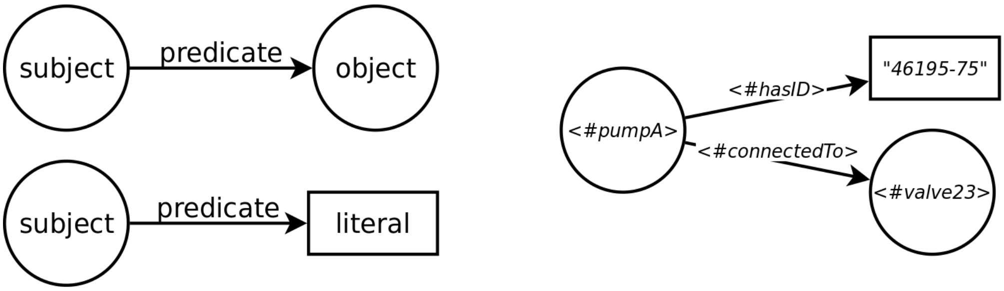


human-readable. RDF data is usually stored in special databases called triple stores which are designed to manage this data structure with high efficiency.

Extensibility is one of the main concepts of Linked Data. It is compatible to the open-world assumption that everything not explicitly stated is undetermined (Bizer, Carroll, Hayes, & Stickler, 2005). This concept eases integration for industrial purposes because subsequent additions of data from other domains into the information cloud impose no substantial barrier. In contrast, under the closed-world assumption something that is not explicitly stated doesn't exist. Thus, the whole cosmos of the system has to be modelled within itself, which widely constricts extensibility.

In Linked Data information is modelled on the basis of concepts and relations expressed in ontologies. These semantic vocabularies are also modelled in RDF themselves. Many ontologies have been developed which can be used as a base for building new and adapted ontologies. The essential ontology languages for building new ontologies in Linked Data are RDF Schema (RDFS) and – as a further development – the Web Ontology Language (OWL/OWL2) which extends RDFS with new constructs for better expressiveness (Yu, 2010). With these ontologies complex taxonomies and compound logics can be developed. However, the development and use of heavy-weight ontologies exploiting all the power of rules, restrictions and logics is error-prone and expensive. For a fast and seamless migration to Linked Data, we recom-

Figure 2. Linked Data triple consisting of subject, predicate and object (left), example for Linked Data with a URI and a literal as objects for two predicates of the same subject (right)



mend the use of light-weight ontologies without heavy reasoning.

Linked Data can reveal its full strength when connecting several information spaces from different domains with their own specific vocabulary. In opposition to that, traditional methods of data storage can only be bound together with rather high effort. Two basic methods are available to fulfill this task in Linked Data. One method is to harmonise data from different semantic domains by explicitly transferring the concept of one domain to another. This can be achieved by extending one ontology with further properties that link to the other ontology, for example via *sameAs* or *seeAlso* links. The other method is to develop a meta-ontology that summarises all the underlying concepts into a single harmonised one, so that the separate underlying ontologies don't have to know from each other. Even the meta-ontology itself can consist of various reusable ontologies which cover specific parts of the semantic.

Various meta-ontologies have been presented that promise such meta-level connections between different ontologies. Usable approaches are, e.g., Simple Knowledge Organisation System (SKOS) (Yu, 2010), Vocabulary of Interlinked Dataset (VOID) (Hyland, 2010), Friend of a Friend (FOAF) (Yu, 2010), Semantically-Interlinked Online Communities (SIOC) (Breslin, Harth, Bojārs, & Decker, 2004) and DublinCore (International Organization for Standardization, 2009).

Linking Industrial Data

In industrial information systems, data is usually integrated on the application level. Thus, every application has to solve the same problems linking different data sources. In contrast, Linked Data integrates directly at the data level creating a new RDF layer on the legacy data. However, Linked Data for industrial applications is still a matter of ongoing research - especially the possibilities of using Linked Data in the field have not yet been thoroughly investigated. Simple generic methods rely on either browsing data or querying data. There is almost no experience about how to apply these methods in a proper and usable way. Evolved data structures which were not designed to meet the requirements of Linked Data are obstructing the use of Linked Data in industrial applications. However, there are two strategies to generate Linked Data from existing industrial data. The first strategy utilises Linked Data to provide a standardised access mechanism for data stored in proprietary databases and Content Management Systems (CMS). An appropriate Linked Data adapter (Bizer & Cyganiak, 2006; Auer, Dietzold, Lehmann, Hellmann, & Aumueller, 2009) is required for this purpose. The second strategy is the complete migration of the databases to native Linked Data stores. The drawback of a migration is that all tools working with the data have to be migrated first.

Some pieces that are needed for industrial applications are still missing. For example, standard data formats and ontologies adapted to industrial use have not yet been established to a sufficient degree. A few domain specific ontologies exist such as OntoCape (Morbach, Yang, & Marquardt, 2007), but further investigation and development is necessary.

While the Linked Data movement has its origin in the W3C initiative (W3C, 2009) for the semantic web with its open and public data, industrial applications make further demands. Trust, data integrity and access control are important issues. Confidential parts of a company's knowledge must be protected. Technologies required for ensuring this remain to be evaluated in the relevant context.

A schematic illustration of connected data for industrial applications is given in Figure 3. It shows the domains of a Distributed Control System (DCS), the Computer Aided Engineering (CAE) system and an external link to a chemical knowledge ontology. A *tag* in the DCS world has connections to data sources for further information retrieval. It also has a link indicating that it implements a *measurement and control function* in the CAE system. Additional information regarding the function is provided by the *electrical engineering cloud* which could be hosted externally. The ontology representing the *physical property* that is measured by the *measurement and control function* and that belongs to a *material flow* is shown in detail. Each *physical property* has a *unit* and a *value*. Links to one or more *constraints* such as minimal or maximal values may also exist. The connection to the chemical domain is established via the *material flow* which is also connected to the *measurement and control function*. In this case it is a *mixture* which consists of several *pure materials*. The chemical properties of this material can be retrieved from public or proprietary databases like the REACH database (Lahl & Hawxwell, 2006). This example shows that a comprehensive view of industrial processes and entities requires information from various domains which can be brought together by Linked Data.

Using Graph Theory for Linked Data

The simple triple pattern used for expressing Linked Data information can also be interpreted as a directed graph. The predicate serves as the link, and subject and object comprise the nodes. In that sense, a semantic network is nothing else than a big directed graph. With this knowledge in mind, one can use the methods and tools from graph theory to solve most issues stated in the previous section.

Graph theory has a well established theoretical and mathematical foundation with a comprehensive ecosystem of tools (Bizer, Carroll, Hayes, & Stickler, 2005). Techniques such as transformation, routing, matching and visualisation are investigated (Marcus, 2008) and can easily be applied to semantic networks. An example for restructuring a graph through filtering can be seen in Figure 4. The upper half of the image shows a strongly interconnected network with many linkages. For a specific use case, only some of the nodes and their direct connections might be of interest. To get a simple network as illustrated in the lower half of the figure, several connections between the nodes need to be condensed. Using an appropriate graph transformation, the server can provide a pre-processed network that is already adapted to the specific task. Consequently, clients of the Linked Data Cloud do not have to perform costly computations to produce useful results. This is important to ensure scalability as the performance of a server can be increased much more easily than the performance of clients, especially mobile ones. Furthermore, graph theory can be used to realise validation and consistency checks.

TECHNICAL FRAMEWORK

Filling the Linked Data Cloud

All of the information that is necessary to generate a Linked Data Cloud for industrial ap-

Figure 3. Heterogeneity of industrial data from different domains can be overcome by the use of Linked Data (Urbas, Ziegler, Doherr, & Schmidt, 2010)

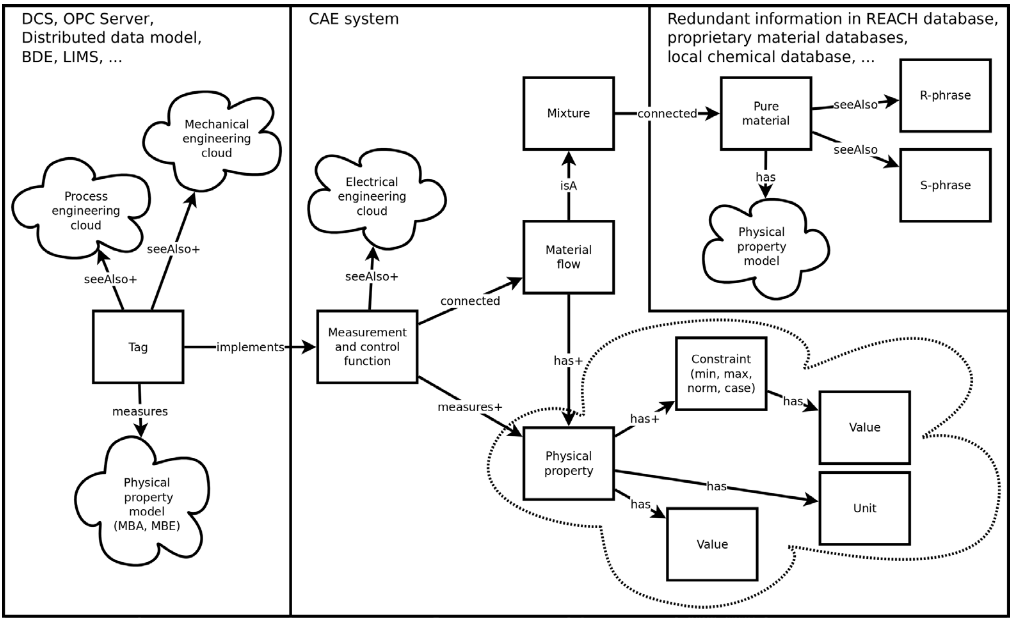
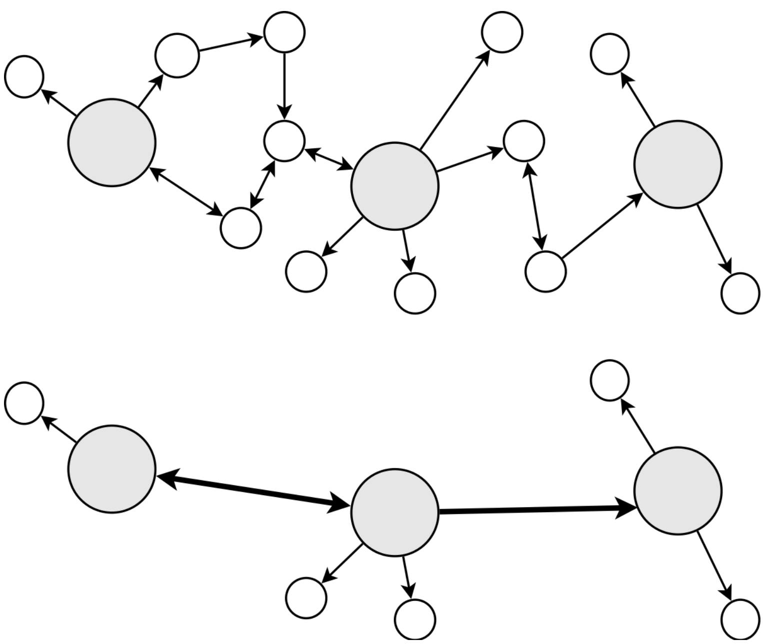


Figure 4. Graph theory allows filtering of Linked Data networks (Graube, Pfeffer, Ziegler, & Urbas, 2011)



plications is usually available within computer-aided engineering systems, public and local databases and other enterprise Information and Communication Technologies (ICT). Information on devices, equipment, pipes, materials, processes, staff, as well as their relationships can be retrieved from digital sources. In industrial environments almost all data exists in a structured or at least semi-structured format but without explicit semantics. It is usually stored in relational databases, spreadsheets, XML files or other structured formats. These formats are based on an implicit or explicit data model including concepts and relations, and are therefore easy to handle for automatic processing. Little information is stored in unstructured formats like plain text files. It requires data mining, text analytics and text analysis techniques, or manual tagging to make this data machine-readable. Therefore, the processing of unstructured data is very expensive. Since it only provides little additional information in this context, it is not considered in this first approach. Consequently, appropriate adapters have to be created to provide Linked Data from proprietary data sources. Linked Data applications can then use the information as a semantic graph while conventional access to the original data is preserved. The creation of adapters further allows for an iterative migration process. Because of the open-world assumption, adapters for subsets of data can be developed step-by-step. The experiences gained in the first iterations can then support the migration of the remaining data. There are several tools for a conversion to Linked Data (e.g., Corcho, Fernández-López, & Gómez-Pérez, 2006). The migration of machine-readable formats such as XML into ontologies requires only a couple of simple transformations. These can be executed by generic tools like XSD2OWL and XML2RDF (García & Gil, 2010). Better results can be achieved by taking the specific vocabulary into account and transforming those XML files with adapted XSL transformations. Even spreadsheets (xls, ods, csv) can be converted to RDF via XLWrap without much effort (Langegger & Wöß, 2009). Furthermore, there are many libraries for operating on RDF as well as on most structured and proprietary data formats which allow easy transformation with the programming language of choice. A good example for an adapter between conventional data and Linked Data is the RDF-View in Virtuoso which provides an adapter for SQL data (Schaffert, Pellegrini, Tochtermann, & Auer, 2009). In other cases the development of novel adapters is necessary, for example when providing live data from a CAE system.

Graphs from various domains are combined and merged through the use of meta-ontologies. This allows further extensibility by linking new information clouds to existing ones without the need to adapt their ontologies. Only the meta-ontology has to be modified so that the concepts of the new data structures are modelled accordingly. This allows companies to keep their own vocabulary. Thus, there is no need to define a unified world model that tries to cover each and every domain.

Accessing the Linked Data Cloud

After transformation into Linked Data, the new information space can be made accessible by means of RDF files on common web servers. However, these files have a static nature and do not allow easy modification. Alternatively, the data can be stored in triple stores like Jena (Yu, 2010) or quad stores like Virtuoso (Schaffert, Pellegrini, Tochtermann, & Auer, 2009). These special databases are optimised to store triples and to guarantee high performance even for large graphs. They enable changes and modifications of the Linked Data from the client side. Some of these triple stores also provide adapters between relational databases and Linked Data.

Another important access mechanism is provided by SPARQL endpoints. They allow graph based queries in the *SPARQL Protocol and RDF Query Language* to select parts of the semantic network. Triple patterns, conjunctions, disjunctions and optional patterns can be used to form SPARQL queries.

DISCUSSION

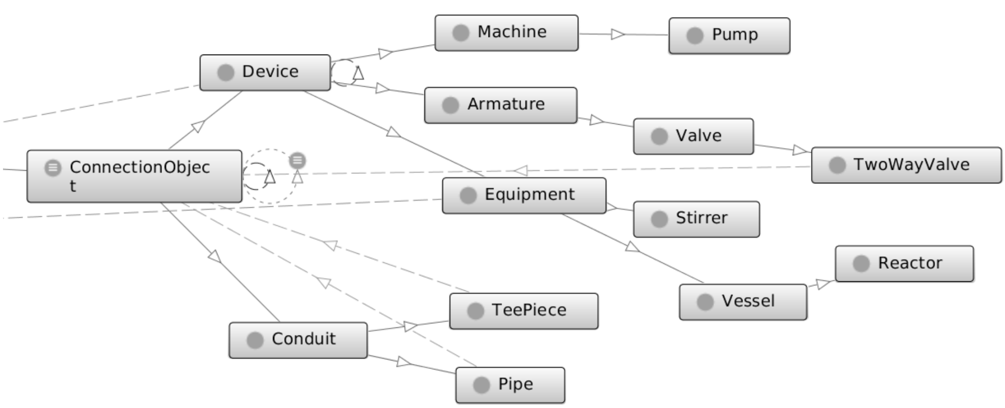
The development of a fully functional Linked Data environment containing access components, allowing graph manipulation and providing data from different sources is a task that requires considerable effort and time. The development has only begun recently, but prototypes for the components have already been created to show the feasibility of the proposed approach.

For this purpose we chose the use case of industrial maintenance. There is data of various plants available from previous projects. The first step of the migration to Linked Data was providing information that is useful for maintenance work in the field. The data of the chosen small-sized laboratory plant is stored in a Comos CAE system (Siemans, 2012). Data about the structure of the plant was exported to an XML file using a Comos *XML-Connector*. The transformation of this information into RDF was executed by a python script using the *rdflib* and *minidom* libraries. The used vocabulary is an ontology which was created with Protègè (Gennari et al., 2003). The *mso* (maintenance structure ontology) is a lightweight ontology describing the structure of the plant and the flow of matter between different equipments. It makes use of the internal Comos model and the IEC 61512-1 to model the structure of the plant with unit, process cells and equipments. An excerpt of the *mso* can be seen in Figure 5. The transformed data as well as the ontology were stored on a Virtuoso server (OpenLink Software, 2011) in different graphs. HTML representations of the RDF data and a SPARQL endpoint were set up at the same server.

First prototypes for generic visualisation of the information stored in the Linked Data cloud have also been implemented. The chosen use case requires an application for use in mobile contexts which provides useful information to the worker during maintenance in the field. An example of such an application has been implemented on the android platform (Urbas, Pfeffer, & Ziegler, 2011). It allows browsing through devices in the plant by following the up or down stream of device identified by its RFID tag.

Although the feasibility of the approach could be shown, there are still numerous challenges on the road for Linked Data to become a valuable part of the industrial ICT ecosystem. The most urgent issue is the formal description and validation of Linked Data Clouds. Linked Data is a generic format and does not provide a well-defined, consistent information meta model. Hence, it will be necessary to explore extensible vocabularies and requirements for

Figure 5. Excerpt from the mso ontology showing structure of equipment



industrial Linked Data Clouds in order to assure data integrity and operability. This also implies that it is presently not possible to validate such information spaces based on existing technologies. Yet, since Linked Data Clouds represent a giant directed graph, it seems promising to leverage tools from the domain of graph theory to analyse and validate Linked Data Clouds. Validation tools for Linked Data information spaces might be based on these theories and configured by domain- and database-specific requirements to assure that Linked Data Clouds are well-formed, consistent and appropriate in their context of use. Another important issue is the lack of end-user frontends. Established interaction metaphors for semantic networks such as search or browse cannot exploit the true potential of Linked Data. Again, graph-based visualisation and navigation algorithms may be an appropriate solution and lead to advanced interaction metaphors. Based on graph abstractions and transformations it might be possible to reduce complexity and diversity of Linked Data Clouds to a reasonable level.

CONCLUSION

Using Linked Data for data integration allows quick dissemination of new concepts. The necessary tools are widely used within the web ecosystem and have their roots in genuine internet technology. Existing enterprise ICT can easily be enhanced with Linked Data technology at minimal expense. It will be possible to set up and maintain shared information spaces across enterprise boundaries. Flexibility, interoperability and time-to-market are key factors for economic success. Linked Data as integrating technology for industrial data addresses exactly these factors.

As promising as the technology may be, it needs further investigation. Future work will focus on joining all components and on creating a meta-ontology to link different ontologies of a specific domain.

ACKNOWLEDGMENT

The research leading to these results is receiving funding from the European Community's Seventh Framework Programme under grant agreement no. FP7-284928 ComVantage.

REFERENCES

Allemang, D. (2010). Semantic Web and the linked data enterprise. In Wood, D. (Ed.), *Linking enterprise data* (pp. 3–23). New York, NY: Springer. doi:10.1007/978-1-4419-7665-9_1

AT&T Knowledge Ventures. (2008). *Collaboration across borders*. Retrieved from http://www.corp.att.com/emea/docs/s5_collaboration_eng.pdf

Auer, S., Dietzold, S., Lehmann, J., Hellmann, S., & Aumueller, D. (2009). Triplify - light-weight linked data publication from relational databases. In *Proceedings of the 18th International World Wide Web Conference* (p. 621).

Berners-Lee, T. (2006). *Design Issues - Linked Data*. Retrieved from http://www.w3.org/DesignIssues/LinkedData.html

Bizer, C., Carroll, J. J., Hayes, P., & Stickler, P. (2005). Named graphs, provenance and trust. In *Proceedings of the 14th International Conference on World Wide Web* (pp. 613-622).

Bizer, C., & Cyganiak, R. (2006). D2R Server – Publishing relational databases on the semantic Web. In *Proceedings of the 5th International Semantic Web Conference* (p. 26).

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data -The story so far. *International Journal on Semantic Web and Information Systems*, *5*(3), 1–22. doi:10.4018/jswis.2009081901

Breslin, J. G., Harth, A., Bojārs, U., & Decker, S. (2004). Towards semantically-interlinked online communities. In A. Gómez-Pérez & J. Euzenat (Eds.), *Proceedings of the 2nd European Semantic Web Conference*, Heraklion, Greece (LNCS 3532, pp. 500-514).

Broser, C., Fritsch, C., Gmelch, O., Pernul, G., & Schillinger, R. (2009). SPIKE – A collaboration platform for short-term virtual business alliances. In *Proceedings of the Workshop on Technological Trends in Enterprise Systems for SMEs and Large Enterprises: Heading towards the Future Internet*, London, UK.

Corcho, O., Fernández-López, M., & Gómez-Pérez, A. (2006). Ontological engineering: Principles, methods, tools and languages. In Calero, C., Ruiz, F., & Piattini, M. (Eds.), *Ontologies for software engineering and software technology* (pp. 1–48). New York, NY: Springer. doi:10.1007/3-540-34518-3_1

García, R., & Gil, R. (2010). Linking XBRL financial data. In Wood, D. (Ed.), *Linking enterprise data* (pp. 103–125). New York, NY: Springer. doi:10.1007/978-1-4419-7665-9_6

Gennari, J. H., Musen, M. A., Fergerson, R. W., Grosso, W. E., Crubézy, M., & Eriksson, H. (2003). The evolution of Protégé: an environment for knowledge-based systems development. *International Journal of Human-Computer Studies*, *58*(1), 89–123. doi:10.1016/S1071-5819(02)00127-1

Graube, M., Pfeffer, J., Ziegler, J., & Urbas, L. (2011). Linked data as integrating technology for industrial data. In *Proceedings of the 14th International Conference on Network-Based Information Systems* (pp. 162-167).

Haslhofer, B., & Popitsch, N. (2009). DSNotify - Detecting and fixing broken links in linked data sets. In *Proceedings of the 20th International Workshop on Database and Expert Systems Application* (pp. 89-93).

Hepp, M., Leymann, F., Domingue, J., Wahler, A., & Fensel, D. (2005). Semantic business process management: a vision towards using semantic Web services for business process management. In *Proceedings of the IEEE International Conference on e-Business Engineering* (pp. 535-540).

Hyland, B. (2010). Preparing for a linked data enterprise. In Wood, D. (Ed.), *Linking enterprise data* (pp. 51–64). New York, NY: Springer. doi:10.1007/978-1-4419-7665-9_3

International Electrotechnical Commission. (2008). *IEC 62424: Representation of process control engineering - Requests in P&I diagrams and data exchange between P&ID tools and PCE-CAE tools*.

International Organization for Standardization. (1998). *ISO 10303-22: Industrial automation systems and integration - Product data representation and exchange - Part 22: Implementation methods: Standard data access interface*. Geneva, Switzerland: Author.

International Organization for Standardization. (2007). *ISO 15926: Industrial automation systems and integration: Integration of life-cycle data for process plants including oil and gas production facilities*. Geneva, Switzerland: Author.

International Organization for Standardization. (2009). *ISO 15836: Information and documentation - The Dublin Core metadata element set*. Geneva, Switzerland: Author.

Lahl, U., & Hawxwell, K. A. (2006). REACH - The New European chemicals law. *Environmental Science & Technology*, *40*(23), 7115–7121. doi:10.1021/es062984j

Langegger, A., & Wöß, W. (2009). XLWrap - Querying and integrating arbitrary spreadsheets with SPARQL. In A. Bernstein, D. R. Karger, T. Heath, L. Feigenbaum, D. Maynard, E. Motta, & K. Thirunarayan (Eds.), *Proceedings of the 8th International Semantic Web Conference*, Washington, DC (LNCS 5823, pp. 359-374).

Latif, A., Höfler, P., Stocker, A., Saeed, A., & Wagner, C. (2009). The linked data value chain: A lightweight model for business engineers. In *Proceedings of the 5th International Conference on Semantic Systems*.

Marcus, D. A. (2008). *Graph theory: a problem oriented approach*. Washington, DC: Mathematical Association of America Textbooks.

Marquardt, W., Morbach, J., Wiesner, A., & Yang, A. (2010). Related work on ontologies for engineering applications. In Marquardt, W., Morbach, J., Wiesner, A., & Yang, A. (Eds.), *OntoCAPE: A reusable ontology for chemical process engineering* (pp. 369–390). Berlin, Germany: Springer-Verlag. doi:10.1007/978-3-642-04655-1_11

Morbach, J., Yang, A., & Marquardt, W. (2007). OntoCAPE – A large-scale ontology for chemical process engineering. *Engineering Applications of Artificial Intelligence*, *20*(2), 147–161. doi:10.1016/j.engappai.2006.06.010

OpenLink Software. (2011). *Virtuoso universal server*. Retrieved from http://virtuoso.openlinksw.com/

Prud'hommeaux, E., & Seaborne, A. (Eds.). (2008). *SPARQL Query Language for RDF.* Retrieved from http://www.w3.org/TR/rdf-sparql-query/

Sauermann, L., Cyganiak, R., Ayers, D., & Völkel, M. (2008). *Cool URIs for the Semantic Web*. Retrieved from http://www.w3.org/TR/cooluris/

Schaffert, S., Pellegrini, T., Tochtermann, K., & Auer, S. (2009). *Networked knowledge - networked media: Integrating knowledge management, new media technologies and semantic systems (Studies in Computational Intelligence)*. New York, NY: Springer.

Siemans. (2012). *COMOS Platform: The software basis for object-oriented, worldwide consistent data management*. Retrieved from http://www.automation.siemens.com/mcms/plant-engineering-software/en/comos-platform/Pages/Default.aspx

Thompson, K. (2008). *The networked enterprise: Competing for the future through virtual enterprise networks*. Tampa, FL: Meghan Kiffer.

Urbas, L., Pfeffer, J., & Ziegler, J. (2011). iLD-Apps: Usable mobile access to linked data clouds at the shop floor. In *Proceedings of the Workshop Visual Interfaces to the Social and the Semantic Web.*

Urbas, L., Ziegler, J., Doherr, F., & Schmidt, T. (2010, October). NetGen:X – automated planning and configuration of communication structures. In *Proceedings of the Workshop on Middleware in Industrial Automation*, Ladenburg, Germany.

Usman, Z., Young, R. I., Case, K., & Harding, J. A. (2010). A manufacturing foundation ontology for product life cycle interoperability. In Popplewell, K., Harding, J., Poler, R., & Chalmeta, R. (Eds.), *Enterprise Interoperability IV: Making the Internet of the future for the future of enterprise* (pp. 147–155). Berlin, Germany: Springer-Verlag. doi:10.1007/978-1-84996-257-5_14

van Heck, E., & Vervest, P. (2007). Smart business networks: how the network wins. *Communications of the ACM*, *50*, 28–37.

W3C. (2004). *Resource Description Framework (RDF)*. Retrieved from http://www.w3.org/RDF/

W3C. (2009). *About W3C*. Retrieved from http://www.w3.org/Consortium/

W3C. (2011). *Linking open data*. Retrieved from http://esw.w3.org/SweoIG/TaskForces/CommunityProjects/LinkingOpenData

Yu, L. (2011). *A developer's guide to the semantic Web*. New York, NY: Springer. doi:10.1007/978-3-642-15970-1

Ziegler, J., & Urbas, L. (2009). Realization of tangible mobile human machine interfaces using wireless personal area networks. In *Proceedings of the International Conference on Intelligent Interactive Assistance and Mobile Multimedia Computing* (pp. 37-48).

Markus Graube is a research scientist and PhD-student at Faculty of Electrical and Computer Engineering at TU Dresden. In 2010, he obtained a Master degree in Mechatronics. He conducts active research in various information technology related areas including usability, semantic technologies and information modeling. This includes specific problems regarding the integration of data and task models in the fields of process industries.

Johannes Pfeffer is a research scientist and PhD-student at Faculty of Electrical and Computer Engineering at TU Dresden. In 2010, he obtained a Master degree in Mechatronics. His research focuses on model driven software engineering, usability on mobile devices, component orchestration of mobile solutions for supervisory control and maintenance applications.

Jens Ziegler is a research scientist and PhD-student at Faculty of Electrical and Computer Engineering at TU Dresden. In 2008, he obtained a diploma in electrical engineering. His research includes task and context modeling, usability engineering and human-machine interaction design for mobile solutions for supervisory control and maintenance applications. Currently, he is the head of the Mobile Information Systems research group at TU Dresden and an active member of the GI Working Group "Tangible Interaction in Mixed Realities".

Leon Urbas (VDI, DECHEMA, VDE, GMA, GI) is a full professor at TU Dresden. His research includes applied cognitive user modeling, usability engineering, human-centered process automation and interactive support system design. After his PhD on internet technologies and operator training simulation he gained three years professional experience as project manager in large scale automation and process information systems at a leading fine chemicals company. In 2000 he went back to university to set up a cognitive human performance modeling group at TU Berlin. Since 2006 he is head of the Chair for Process Control Systems Engineering at TU Dresden. His group works on integrated information models and workflows for control systems engineering and innovative mobile information systems. Prof. Urbas is member of the board of processNet PAT, spokesman of the GMA Task Force 5.16 "Middleware in Industrial Automation", and Editor-in-Chief of the scientific journal ATP Edition.