Semantic Web based Container Monitoring System for the Transportation Industry

Pinar Gocebe (1), Oguz Dikenelli (1), Umut Kose (2), Juan F. Sequeda (3)
(1) Ege University
(2) Bimar Information Technology Services
(3) Capsenta

Abstract: Goods are transported around the world in containers. Monitoring containers is a complex task. In this presentation, we will present a Container Monitoring System based on Semantic Web technologies. This system is currently being developed by Ege University, Bimar Information Technology Services and Capsenta for ARKAS Holding, one of Turkey's leading logistics and transportation companies. Our presentation consists of 1) introducing the challenges of monitoring containers in the transportation industry, 2) how existing technologies and solutions do not satisfy the needs, 3) why Semantic Web technologies can address the needs, 4) how we are using Semantic Web technologies including architectural design decisions and finally 5) describe lessons learned.

Problem: Monitoring Containers in the Transportation Industry

Logistics and Transportation Industry works as a complex system where different databases interact with each other dynamically. ARKAS Holding is one of Turkey's leading logistics and transportation companies. It operates in different fields such as sea, land, rail, air transportation, ship operations and port operations. The objective is to transport a container from a start location to an end location. One of the most important problems is monitoring containers. Each step of the container transportation process may be performed by a different company. Therefore, monitoring the container's lifecycle in real time is a challenging engineering task because these processes runs in parallel on different software systems. The end goal is to have managers and customers be able to track the container transportation process.

Why Semantic Web technologies?

In the last decade, Electronic Data Interchange (EDI) based standards (EDIFACT, RosettaNet, STEP, AnsiX12), XML standards and Service Oriented Architecture (SOA) approaches are used for solving the integration problems of logistics and transportation industry [1]. The standards provide common syntax for data representation and SOA provides an application integration infrastructure. However, these technologies are not sufficient to ultimately solve the integration challenges in large enterprise for the following reasons:

- In EDI-based standards, EDI messages are pushed among the organizations on a predefined time and these standards are not suitable for real-time applications [2].
- In the SOA approaches, the most important problem is interoperability[3]. Unique identifiers in a database, are understood inside of a system but they may lose their meaning in another system. Finding the operation of a web service that will be called by an identifier is fixed in the software application logic. The resulting application logic is elaborate when considering the complexity of the logistics and transportation industry.

In the recent years, Semantic Web standards and infrastructure are prevalently used to integrate enterprise information and business processes [4]. Semantic Web provides an integration environment which is more flexible, extensible and open to the exterior when necessary. The same identifiers (URIs) can be used across different data sources creating then a huge knowledge base. Software systems can use this knowledge base independently from each other. Semantic Web technologies decreases the dependence on the middle-tier technologies whose management is hard; so maintenance and management processes are become easier. For all these reasons, Semantic Web technologies comprises a new solution to the dynamic, distributed and complex nature of the logistics and transportation industry.

In ARKAS, there are approximately 200 active integration projects being carried out domestically and internationally. Development cost of a new integration between operational systems to supply tracking data each other is ~25-30% and maintenance cost is ~10-15% of the total project cost. The integration of a new database is performed in approximately one month because of the data format identification and transformation process between the different technologies that are being used.

The Project

Ege University, Bimar Information Technology Services and Capsenta are working together to develop a container monitoring system based on Semantic Web technologies[1]. The goal of using Semantic Web technologies is to decrease the cost of integration and dependencies between software systems. The first phase consists of integrating four internal relational databases of ARKAS on ports, agencies, land transport and warehouses. The second phase consists of integrating external relational databases from third-party companies such as a warehouse and land transportation companies. We are currently near the end of the first phase.

How we use Semantic Web technologies

In order to address the problem, we use a hybrid architecture consisting of a combination of Extract-Transform-Load (ETL), Wrapper, Warehouse and Federation. We have created OWL ontologies that describe ports, agencies and warehouses and R2RML mappings between the relational databases and the ontologies. ARKAS' internal databases are ETLed to RDF using Capsenta's Ultrawrap and warehoused in an RDF triplestore. The goal of having a centralized RDF triplestore is to have full control of the data in a single repository and perform analytics over the integrated data. External data sources are integrated into the system in a distributed manner. External relational databases are wrapped with Ultrawrap in order to provide a virtual RDF view. A query federator is used to integrate the external sources with the internal sources. A wrapper is ideal for the external sources because third-party companies are not willing to give up their data. In order to keep updates to the underlying relational databases consistent with the RDF data in the triplestore, we use data capture systems for relational databases.

Current Lessons Learned

- Given the complexity of the transportation domain, creating ontologies for this domain is not straightforward. Existing ontologies do not satisfy our use case.
- Creation of R2RML mappings involves a domain expert and an ontology engineer. For example, it took 15 days to create the mappings for the port database.
- Deciding on the appropriate architecture according to the requirements is a complex process.
- Creating a simple core ontology and mapping it with suitable R2RML patterns are important tasks to provide a scalable architecture.
- Ongoing work is testing the scalability of the systems as a result of integrating new databases.

References

[1] Nurmilaakso, J.-M.: Adoption of e-business functions and migration from EDI-based to XML- based e-business frameworks in supply chain integration. International Journal of Production Economics 113(2), 721-733 (2008)

[2] Harleman, R. 2012. Improving the Logistic Sectors Efficiency using Service Oriented Architectures (SOA). In 17th Twente Student Conference on IT (2012)

[3] The European Interoperability Framework, http://ec.europa.eu/isa/documents/isa_annex_ii_eif_en.pdf (2010)

[4] Frischmuth, P., Klímek, J., Auer, S., Tramp, S., Unbehauen, J., Holzweißig, K., Marquardt, C.-M.: Linked Data in Enterprise Information Integration. Semantic Web – Interoperability, Usability, Applicability. IOS Press Journal (2012)

[1] This project is funded by the Republic of Turkey Ministry of Science, Industry and Technology.