Computers & Industrial Engineering 83 (2015) 39–60

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

A framework to canonicalize manufacturing service capability models

Boonserm Kulvatunyou [a], Yunsu Lee [a,b,*], Nenad Ivezic [a], Yun Peng [b]

[a] Systems Integration Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA
[b] Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD 21250, USA

ARTICLE INFO

Article history:
Received 6 September 2013
Received in revised form 14 January 2015
Accepted 29 January 2015
Available online 7 February 2015

Keywords:
Manufacturing service capability
Ontology design pattern
Pattern-based ontology transformation
Canonicalization
OWL
Semantic mediation

ABSTRACT

The capability to share precisely defined information models, which reveal a supplier's *manufacturing service capability* (MSC) with anyone who needs it, is key to the creation of more agile supply chains. Today, unfortunately, this capability does not exist. Why? Because most suppliers use proprietary information models to represent and share their MSC information! This limits both the semantic precision in the models, which is needed for interoperability, and the level of agility in the supply chains. The availability of a semantically precise and rich reference MSC ontology could address both of these limitations. Based on our prior research, the development of such an ontology will require a semantic mediation process between the proprietary MSC models and the reference MSC ontology. At the heart of every known, semantic-mediation process is a mapping between a proprietary MSC model and the reference MSC ontology. Such a mapping must deal with the structural and semantic conflicts between the two. In this paper, we propose a new approach, which we call *canonicalization* to address the structural conflicts. The semantic conflicts are addressed using logical mapping. The canonicalization pre-processes the structural representations of the proprietary models and then aligns them using ontology design patterns which are also used in the reference ontology. This simplifies both the mapping problems themselves and the resulting mapping statements considerably. In the paper, we also demonstrate our approach and its benefits in the context of a description-logic-based semantic mediation using the Ontology Web Language (OWL).

1. Introduction

The Smart Manufacturing Leadership Coalition stated in its 2011 workshop report that the capability to share manufacturing service capability (MSC) information was the key to the creation of more agile and better optimized supply chains (SMLC, 2011). Min and Zhou (2002) also showed that this same capability could enable successful supply chain integration. Ameri and Dutta (2006) showed that integration was only possible when that MSC information is semantically precise, complete, and interoperable. Currently, this is not the case, because manufacturing companies provide their MSC information in proprietary MSC data models. Examples of these *proprietary MSC data models* can be found in every online marketplace dedicated to finding manufacturing suppliers for OEMs. These proprietary MSC data models are heterogeneous in their structures and representations, which make it hard for the OEMs to understand those models and find the best supplier that fits their needs. In situations like this, it is clear that information sharing, which is critical to the success of both the OEM and the supplier, is extremely difficult and costly.

Researchers have shown that a reference ontology can enhance the access to and precision of information models. In particular, Ye et al. (2007), Lu et al. (2013), Wang et al. (2013), and Zheng and Terpenny (2013) use Web Ontology Language (OWL) (W3C, 2009a) coupled with Semantic Web Rule Language (SWRL) (W3C, 2004a) to link local and reference ontologies. Kulvatunyou et al. (2013) and Tsinaraki et al. (2004) achieve similar linkages using only OWL axioms.

The *OWL-based semantic mediation* approach in Kulvatunyou et al. (2013) uses an OWL reasoner and OWL mapping axioms to inherit semantics from a semantically rich *reference MSC ontology*.[1] This approach enhances semantic precision and coverage and

* Corresponding author at: 100 Bureau Drive, Stop 8260, Gaithersburg, MD 20899-8260, USA. Tel.: +1 3019758798.
E-mail address: yslee@nist.gov (Y. Lee).

[1] In this paper the term "manufacturing service capability (MSC) model" or data model generally includes both schema and instance data. However, the reference ontology generally does not have instance data. A simple example of an instance data is 'Company A has drilling process capability with 0.025 mm precision'. We use the term 'MSC data model' in a very general sense to refer to any structured or semi-structured MSC information source; while the term 'MSC model' refers to formally encoded information specifically in OWL.

http://dx.doi.org/10.1016/j.cie.2015.01.027
0360-8352/

resolves semantic conflicts across proprietary MSC data models. The approach worked because the reference MSC ontology provided a common domain model and terminology. It has three major steps. First, information in the proprietary MSC data models is transformed into the common RDF syntax (W3C, 2004b) using the OWL semantics. Second, the resulting *OWL-encoded proprietary MSC model* is mapped to the reference ontology using OWL axioms. Finally, the description logic inference, over the OWL-encoded proprietary MSC models, the reference ontology, and the mapping axioms, results in improved MSC information sharing.

In this paper, we focus on transformation and the axioms. There are generally two ways to transform proprietary MSC data models into OWL: purely syntactic or with semantic interpretation. In the purely syntactic way, generic transformation rules are applied to the data source. Those rules are based on the underlying schema language. For example, in the case of relational databases, tables are transformed into classes and columns are transformed into properties. In the semantic-interpretation way, humans write rules that are specific to the data source schema and they use them for the transformation. Such rules are typically developed from the sole viewpoint of the data source owner.

In either case, the resulting models, called *arbitrary OWL-encoded proprietary MSC model*,[2] are not aligned structurally with the target reference ontology (see the top of Fig. 1). Such an arbitrary OWL-encoded proprietary MSC model can render the OWL mapping axioms, which are required by the approach in Kulvatunyou et al. (2013), exceedingly complex, if it is at all possible. Technologies such as the D2RQ (D2RQ, 2014) and the W3C's R2RML (Relational Database to RDF Mapping Language) (Das et al., 2012) support both the pure syntactical as well as the semantic-interpretation transformation practices.

In this paper, we propose a methodology, called a canonicalization approach, to streamline the OWL-based semantic mediation process by simplifying the OWL mapping axioms and the actual mapping itself. First, we transform the proprietary MSC data model automatically by using a common, syntactic, rule set that is independent of its source data schema. Second, a human applies a canonicalization by transforming the data again using a set of design patterns. Third, the human writes the OWL mapping axioms against the reference ontology. Since the design patterns used in the canonicalization are also used in the reference ontology, the resulting *canonicalized OWL-encoded proprietary MSC model* is more structurally aligned and, therefore, simpler to map to the reference ontology. This proposed methodology is illustrated at the bottom of Fig. 1.

In this paper, we also validate our approach by providing quantitative and qualitative analysis for a manufacturing semantic mediation example. The qualitative analysis will show that canonicalization can (1) amend a model not originally suited for semantic mediation via OWL DL, (2) simplify the mapping by avoiding the need for complex, OWL-class expressions in the mapping axioms, and (3) simplify the mapping maintenance by reducing the number of, and complexity of, mapping axioms. The quantitative analysis will show that computational time grows cubically when a certain, yet common, type of structural conflicts is resolved without canonicalization, as opposed to linearly when using canonicalization.

The rest of the paper is structured as follows. In the next section, we provide a literature review. In Section 3, we characterize canonicalization by the types of semantic conflicts it can address. Section 4 introduces the proposed canonicalization framework. It is followed with Section 5, which validates the applicability and usefulness of the framework with a running example. Section 6 presents the qualitative and quantitative analyses. Finally, we provide a conclusion and remarks on the current work and our future plans in Section 7.

[2] By arbitrary, we mean that the MSC model inconsistently and sub-optimally uses one or more approaches to express manufacturing information using the OWL language, whether it involves class-based, property-based, or some general axiomatic representation that is specific to proprietary view of the data.

2. Literature review

The importance of a reference model in semantic mediation has been emphasized in recent research. Bloomfield et al. (2012) proposed a core, manufacturing-simulation reference model to improve the data exchange between manufacturing simulations throughout the product life cycle. Wang et al. (2013) provided a shared-negotiation ontology to address communicative interoperability problems in supply chain negotiation. Zheng and Terpenny (2013) enhanced the semantics of legacy information by combining a global ontology with the legacy domain knowledge. The global ontology served as the reference model to provide additional semantics to the legacy domain knowledge. As noted above, Kulvatunyou et al. (2013) provided details of semantic mediation using the Web Ontology Language (OWL) (W3C, 2009a). In that work, the inference over the mapping between the proprietary and the semantically rich reference OWL model results in semantic enhancement to the proprietary OWL model.

All of the previously mentioned approaches require a mapping between the proprietary model and the reference model. According to Shvaiko and Euzenat (2011), they note, however, that developing such a mapping is one of the most difficult tasks in the semantic mediation, especially if there are structural conflicts. It is not surprising then that none of the previously mentioned semantic mediation approaches provides any methods or tools to assist in the mapping task. In addition, our evaluation of the ontology-mapping approaches described in Noy and Musen (2003) and McGuinness et al. (2000) found that they also do not perform well in the face of structural conflicts. Our hypothesis is that the mapping task could be simplified if the proprietary model is encoded with the same Ontology Design Patterns (ODPs) as the reference ontology. The reason is that concepts in the proprietary model would be represented with the same types of entities and with the same relationship structures used in the reference ontology. We call this, ontology canonicalization.

The approach described in Svab-Zamazal et al. (2009) and Svab-Zamazal and Svatek (2011) includes workable methods and tools for the ODP-based ontology transformations. Together, these methods and tools are called PATOMAT. PATOMAT produces a well-defined XML schema for both pattern definitions and transformation rules. In addition, the authors developed the functionality and software to generate a SPARQL query from the pattern definitions. That software uses an OPPL application interface (OPPL, 2012) for pattern transformation and a GUI editor for capturing both the source and target ontology patterns and the associated transformation rules.

PATOMAT provides a good foundation for our ontology canonicalization approach. However, several enhancements are needed. First, PATOMAT does not deal with the representative artifacts that represent the varying parts of the source ontology pattern. This means that whenever there are multiple pattern instances that use the same source ontology pattern, PATOMAT does not generate the correct, recursive, transformation rules. Second, PATOMAT does not provide any method either to generate a source ontology pattern or to retrieve a reusable target ontology pattern. This means that all the patterns must be defined manually. Lastly, PATOMAT has no facilities to deal with literal value pattern detections and transformations at present. In this paper, we describe a framework that fills these gaps in PATOMAT so that it can be used to fully canonicalize a proprietary ontological model.