

Iterative Effort Reduction in B2B Schema Integration via a Canonical Data Model

Michael Dietrich, SAP AG, Walldorf, Germany

Jens Lemcke, SAP AG, Walldorf, Germany

Gunther Stuhec, SAP AG, Walldorf, Germany

ABSTRACT

Nowadays, B2B integration still remains a big cost driver for companies. On the one hand, standardization efforts were able to reduce the mapping effort between e-Business schemas. However, the effort for creating customized messages from the huge and underspecified standard templates increased. Due to the myriad of different requirements by different companies, a great variety of standards coexist. Instead of forcing companies to adopt huge standards, this article propagates an iteratively improving schema and mapping derivation system in the cloud. Thus, we provide flexibility, but streamline companies' integration efforts based on an evolving canonical data model. This approach reduces the need for explicit standardization to a minimum. Our simulation based on real schemas shows a potential to reduce guide creation effort by 50% and mapping effort from 6% to almost 100%.

Keywords: B2B Integration, Canonical Data Model, Cloud, E-Business Schemas, Guide Creation Effort, Mapping Effort

INTRODUCTION

In current B2B integration scenarios, companies have to face the great overhead in B2B message templates. In order to exchange messages, they need to customize these standard message templates published by standardization organizations like the UN/CEFACT (UN/CEFACT, 2003). These templates contain thousands of fields in order to cover all the business needs of one industry domain or even across industries.

For the creation of a new message (also called message guide) only a small amount of fields from the standard template is used. Creating such a message guide means customizing the standard template: redundant fields have to be removed and missing fields have to be added. This manual process is time consuming and error prone. Another big issue is the different semantic understanding or the miss-use of a field: business partners might use the same syntactical field for different purposes. Ad-

DOI: 10.4018/ijst.2013100102

ditionally, market leaders might ignore given template structures and force partners to adapt to their interpretation. Over the last decades, a great number of various EDI standards emerged. Standardization organizations tried to cover industry domain-specific requirements on the one hand, as well as industry-independent demands on the other hand. Furthermore, if companies need to support various standards, integration cost increase even more. For each standard, software has to be adapted or additional modules need to be bought. A step towards the reduction of the standard heterogeneity is the creation of subsets for different industries within one standard. However, in this case the same compatibility problems mentioned above occur as well: business partners might not use the identical subset or misuse fields.

In case two business partners want to exchange messages using different standards or subsets of standards, they require a mapping between these messages. The mapping assigns each field of the source message to a corresponding field of the target message guide. The creation of a message mapping between two standards is also a time consuming process. Usually, consultants are hired by the companies in order to create these mappings as expert knowledge of the involved standards is required.

In total, integration cost still make up around 40% of companies' IT budgets (Kastner and Saia, 2006). In terms of guide creation and mapping creation, there is a lot of potential for effort reduction.

As a pre-study, we examined concrete standard business document templates of 7 different message types from 15 different e-business standards of different industry domains, and their interpretations from 50 different companies. Our analysis revealed that on average more than 60% of the structure and the elements of each schema are semantically comparable. However, only 5% are syntactically similar. Consequently, a precise and commonly understandable lingua franca with consistent and semantically unambiguous meaning of structure and elements is feasible and the key solution. Our approach

incorporates a canonical data model (CDM) as the single view of data for multi-enterprises, enterprises, divisions, or processes and can be independently used by any system or business partner. In this article, we enrich the approach presented in (Dietrich et al., 2013) by refining the notion of the knowledge base to the concept of the Unified Data Model (UDM).

There are several related research approaches, like the MOMIS tool (Beneventano et al., 2001), the Porsche approach (Saleem et al., 2008), the Xyleme project (Delobel et al., 2003), and the BInXS approach (Santos Mello & Heuser, 2005). Besides the existing research, several solutions have found their way into the industry, for example the Contivo solution from Liasion (2012) or Boomi AtomSphere (Boomi, 2012). However, all related works only partly tackle the problems presented in this article. The central data model is mostly manually modeled and a common structure is never computed. This means that especially cross-standard communication remains a big issue for companies. In addition to our previous research, a comprehensive evaluation of the existing work regarding the individual features of our approach is provided.

Our approach covers cross-domain as well as cross-standard communication. It is different from existing endeavors, as these rather focus on the mapping task. However, the core problem is the heterogeneity of the schemas, which leads to high mapping costs. Thus we automatically compute an evolving canonical data model which consolidates the current knowledge of the correspondences of existing schemas. As can be seen in Figure 1, our approach aims to increase homogeneity by applying the knowledge in the CDM. With that, we instantly reduce the mapping effort.

Participation can only be stimulated if constraints are minimal and benefits are substantial and well distributed. The effort reduction in our approach is produced by the following features, which become possible through the CDM, and which each address a key challenge in contemporary B2B integration:

23 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage:

www.igi-global.com/article/iterative-effort-reduction-in-b2b-schema-integration-via-a-canonical-data-model/103865?camid=4v1

This title is available in InfoSci-Journals, InfoSci-Journal Disciplines Library Science, Information Studies, and Education. Recommend this product to your librarian:

www.igi-global.com/e-resources/library-recommendation/?id=2

Related Content

Comparison of Baccalaureate Nursing Students' Experience of Video-Assisted Debriefing versus Oral Debriefing following High-Fidelity Human Simulation

Colleen Royle and Kathleen Hargiss (2015). *International Journal of Strategic Information Technology and Applications* (pp. 40-49).

www.igi-global.com/article/comparison-of-baccalaureate-nursing-students-experience-of-video-assisted-debriefing-versus-oral-debriefing-following-high-fidelity-human-simulation/136267?camid=4v1a

A Relative Comparison of Leading Supply Chain Management Software Packages

Zhongxian Wang, Ruiliang Yan, Kimberly Hollister and Ruben Xing (2010). *Strategic Information Systems: Concepts, Methodologies, Tools, and Applications* (pp. 220-235).

www.igi-global.com/chapter/relative-comparison-leading-supply-chain/36692?camid=4v1a

Strategic Alignment for Electronic Commerce

Christian Bauer (2001). *Strategic Information Technology: Opportunities for Competitive Advantage* (pp. 258-272).

www.igi-global.com/chapter/strategic-alignment-electronic-commerce/29771?camid=4v1a

Cooperative Decision Support Systems

Pascale Zaraté (2010). *Strategic Information Systems: Concepts, Methodologies, Tools, and Applications* (pp. 1551-1558).

www.igi-global.com/chapter/cooperative-decision-support-systems/36774?camid=4v1a