Chapter 11
Semantic Technologies and E-Business

Ivan Bedini
Orange Labs, France

Georges Gardarin
University of Versailles, France

Benjamin Nguyen
University of Versailles, France

ABSTRACT
This chapter studies what semantic technologies can bring to the e-business domain and how they can be applied to it. After an overview of the goals to be achieved by e-business applications a large panel of existing e-business standards is detailed, with a specific focus on B2B (Business to Business) and their current modus operandi. Furthermore, some of the most relevant e-business ontologies are also presented. Next, the chapter argues that the use of semantic technologies will simplify the automatic management of many e-business partnerships. However the construction of ontologies brings a new level of complexity that might be facilitated by automating the great part of the generation process. For this purpose, the Janus system, which is a prototype to help with the automatic derivation of ontologies from XML Schemas, the de-facto format adopted in e-business standard applications was developed. Differently from existing systems, it permits to retrieve automatically conceptual knowledge from large XML corpus sources and is based on the use of the Semantic Data Model for Ontology (SDMO) whose advantages are presented in this chapter.

INTRODUCTION
Computer mediated networks play a central role in the evolution of Information Systems. For example the sales application must interface with the inventory application or the inventory application must connect to the supplier’s application, or the simple mobile calendar must synchronize with the professional calendar; all the time, applications require efficient and effortless integration with others. Nevertheless the integration of enterprises applications still remains harder than it really
should be. Enterprises are typically composed of several applications that are custom built, acquired from third parties or a combination of both. Moreover it is not uncommon to find an enterprise whose information is segmented between different instances of enterprise software and countless departmental solutions. In consequence, the integration of these application systems becomes a real challenge that requires considerable human effort, especially if the final goal is to connect applications belonging to different enterprises. This last use case refers to what is also called Business to Business (or simply B2B).

Communication between applications is mainly governed by standard protocols and standardized content, as shown in the European e-business report (E-Business W@tch, 2007) among different solutions applicable to e-business, at least three enterprises out of four that implement business exchanges with partners, declare implementing applications standards solutions based on these two technologies (in Europe). The advent of XML along with Web Services, and more generically with the Service Oriented Architecture (SOA), has contributed greatly to the development of such standards-based integration solutions. But the large adoption of these technologies entails a new fragmentation in applications development. As a result standardisation addresses only parts of the integration challenge. The frequent claim that XML is the lingua franca for system integration is somewhat misleading; indeed this statement does not imply common semantics and its adoption has led to the creation of countless dialects and languages which cannot be understood and integrated directly by machines. This problem is reflected in the many existing B2B standards that we present in this Chapter. The analysis we provide is based on the observation of more than 40 of them.

Following this approach, professional exchange integration scenarios are based on a complete transformation of business messages at design time. Although this model works and businesses are able to exchange messages electronically, the effort to produce these standards appears too high. Moreover, it would be impossible to write a standard specification for every possible business communication. Especially for (smaller) firms who are unable to contribute to standardization. For this reason Semantic Web-related technologies are well suited to integrate the e-business architecture in order to fulfil the standardization approach and achieve the needed flexibility.

Another aspect that we tackle in the Chapter is the automatic construction of top-level domain ontologies. As asserted by Euzenat and Shvaiko (Euzenat et al., 2007), the importance of the generation of such kind of knowledge is fundamental for the improvement of the alignment and thus integration problem. However most solutions implicitly assume that reference knowledge exists in compatible format and semantics, but actually it is often inadequate for the application domain or difficult to find, if it even exists at all.

To give a point of comparison, we also present the most adopted approach to e-business data integration. Through this analysis we point out the current architecture limitations and explain why ontologies are a better approach which leads to a gain in flexibility and dynamicity. In this sense we provide an overview of schema matching and ontology alignment solutions and we point out one of the current limitations to their broad adoption and provide a system that facilitates, by automation, the transformation from the current model to the “next one”: from XML to OWL.

The overall outline of the Chapter is as follows: the first section introduces current e-business approaches to data integration and we follow with the presentation of more than 40 existing standards for the B2B and B2C domains. Following this introduction we focus on Semantic Web related technologies applied to the e-business domain. In the survey we detail some of the most relevant works related to product classification and we continue with a section focusing on schema