Chapter 15
Semantic Services for Business Documents Reconciliation

Michele Missikoff
Consiglio Nazionale delle Ricerche, Italy

Fabrizio Smith
Consiglio Nazionale delle Ricerche, Italy

Francesco Taglino
Consiglio Nazionale delle Ricerche, Italy

ABSTRACT
This chapter intends to report about the ongoing activities in the COIN European project concerning semantic reconciliation of business documents for supporting interoperability of software applications in e-government and e-business scenarios. The approach is based on a reference ontology against which business documents are mapped through semantic annotation and building of reconciliation rules. The work starts from the semantic reconciliation suite developed in the ATHENA European project and intends to improve the suite by providing an automatic support to the definition of mappings. In order to do that, three automatic services have been developed: (1) semantic annotation of business documents; (2) definition of transformation rules; (3) rules optimization and fusion. A running example concerning the exchange of a legal verification document in a scenario of cross border cooperation between European chambers of commerce will guide through the description of the services.

INTRODUCTION
In this chapter, we present a service-oriented architecture for supporting semantic interoperability. The goal of semantic interoperability is to allow the (seamless) cooperation of software applications, which were not initially developed for this purpose, by using ontology-based semantic methods. We will focus on the exchange of documents in an e-government scenario. Such an exchange can take place:
Semantic Services for Business Documents Reconciliation

- between different public / private organizations,
- within the same organization.

The starting point of this work is the semantic suite for business documents reconciliation developed in the ATHENA project. The suite is conceived for semantic annotation of business documents and building of transformation rules aimed at the exchange of documents between heterogeneous software applications. The suite has been now acquired by the COIN project as part of the so called baseline services, and will be at the basis of the development of innovative services.

The objective of this chapter is to describe the ongoing work in the COIN project aimed at improving the ATHENA semantic reconciliation suite, providing an automatic support to the usage of the suite. In fact, some of the steps required by the reconciliation suite, the semantic annotation (declarative mapping) of business documents and the definition of transformation rules (operational mapping) are currently mainly manual activities, being therefore time consuming and error prone. For this reason, a set of innovative services are being developed in the COIN project with the aim of automatically support and optimize the entire reconciliation process.

The rest of the chapter is organized as follows: a section about methodological aspects concerning formalisms for representing domain ontologies, documents and mappings; an overview of the three innovative services: (1) declarative mapping discovery; (2) operational mapping generation; (3) transformation rules fusion; related works on existing standards for documents representation and exchange, platforms for semantic reconciliation of documents and methods for representing and discovering mappings. Finally the Conclusions section ends the chapter.

**METHODOLOGICAL ISSUES**

**Reference Ontology as a Common View of the Business Domain**

Our approach is based on the idea that there is a common view of the world or, more precisely, of the addressed domain (e.g., legal verification, e-procurement). If a common understanding is missing at “human level”, application interoperability will necessarily fail, no matter what technology will be used. In the reconciliation suite, this common view is represented by a reference ontology that aims at representing, in a formal and unambiguous way, a portion of the reality. Then, the data schemata of the software applications involved in the reconciliation process are mapped against the reference ontology in order to solve the interoperability clashes among those schemata.

In the semantic suite the reference ontology can be an OWL-DL ontology or can be built according to the OPAL methodology (D’Antonio, Missikoff, & Taglino, 2007). OPAL (Object, Process and Actor Language) is an ontology building methodology, proposed by IASI-CNR, strongly based on OWL. With respect to OWL, OPAL can be seen as a set of ontological design patterns aimed at providing semantic guidance to the ontology engineers; furthermore, the modeling constructs of OPAL exhibit a domain adequacy for eBusiness that OWL, being a general purpose language, does not have. Based on OPAL, the ATHOS ontology management system (Taglino & Missikoff, ATHENA IP Deliverable A3.2, 2005) has been developed.

**Document Representation**

Concerning the representation of the documents to be mapped against the reference ontology, different representation formalisms have been considered. In particular at the current status of the project we accept as input: