Ontology-based Interoperability Services for Semantic Collaboration in Open Networked Systems [*]

Silvana Castano, Alfio Ferrara, Stefano Montanelli
{castano,ferrara,montanelli}@dico.unimi.it

Università degli Studi di Milano
DICo - Via Comelico, 39, 20135 Milano - Italy

Abstract. In this paper, we describe an ontology-based collaboration model for supporting semantic interoperability in open networked systems. We characterize discovery and matchmaking *semantic interoperability services* for retrieving information resources semantically related to a target request, to enable a coordinated and virtualized access to distributed heterogeneous information resources.

Key Words: *Semantic Interoperability, Ontologies, Networked Organizations*

1 Introduction

Semantic interoperability is a crucial problem in open networked systems where many different and independent enterprise parties need to cooperate and share heterogeneous information resources often in response to opportunities or challenges that cannot be anticipated in advance and require a rapid response. Accessing heterogeneous and distributed informational resources in a coordinated and virtual way requires appropriate semantic interoperability techniques to enable a seamless access and retrieval of the right information resources, while preserving the information representation and management requirements of each single party involved in the coalition [1, 15]. In addition, a further requirement for effective semantic interoperability techniques regards the availability of semantically rich descriptions of information resources in use by an enterprise party [12]. Ontologies are generally recognized as an essential tool for allowing communication and knowledge sharing among distributed users and applications, by providing a semantically rich description and a common understanding of a domain of interest. Due to the Semantic Web, a large body of research has been recently devoted to ontologies and ontology language standard proposals [2]. In this context, a crucial role is related to the availability of matchmaking techniques based on Semantic Web technologies (e.g., OWL [16]) in order to discover information

resources based on available ontology descriptions. Work related to this topic has been addressed in [3, 8], where intelligent techniques based on a Description Logics approach are described, which compare the knowledge contained in different concept ontologies, by looking for semantic mappings denoting similar concepts. Recent research in P2P systems focuses on providing techniques for evolving from basic P2P networks supporting only file exchanges using simple filenames as metadata, to more complex systems like schema-based P2P networks, capable of supporting the exchange of structured contents (e.g., documents, relational data) by exploiting explicit schemas to describe knowledge, usually using RDF and thematic ontologies as metadata [10, 11].

In this paper, we focus on the information resource discovery problem, and we propose a semantic collaboration model for open networked systems, where autonomous enterprise parties require a coordinated access to heterogeneous and distributed information resources. We rely on ontologies for representing the structure and the semantics, including interdependencies and relationships, of the information resources required and in use by a given enterprise. We characterize semantic interoperability services which exploit ontology descriptions for realizing a semantic collaboration model for networked organization contexts [7, 9, 10]. In particular, we characterize three ontology-based *interoperability services*, namely, the *matching service* for performing semantic affinity evaluations on ontology elements, the *discovery service* for query composition, propagation, and processing, and the *acquisition service* for information resource access.

The paper is organized as follows. In Section 2, we discuss semantic interoperability requirements and a collaboration model for open networked systems. In Section 3, we present ontology-based semantic interoperability services. In Section 4, we describe an application example of the proposed collaboration model. Finally, concluding remarks are presented in Section 5.

2 Semantic interoperability requirements in open networked systems

The following features affect collaboration in open networked systems, and need to be addressed by appropriate techniques: (i) *dynamism of the system*, regards the fact that enterprise parties are allowed to join and leave the networked organization at any moment; (ii) *autonomy of enterprises*, in that each enterprise is responsible for its own information resource management and representation; (iii) *absence of a-priori agreement*, about ontology specification vocabulary and language to be used for knowledge specification; (iv) *equality of responsibilities*, no centralized entities with coordinating tasks are recognized and each party enforces interaction facilities with other parties for resource sharing and collaboration.

In an open networked system, each involved enterprise party is autonomous and acts as a node (peer) like in typical open distributed systems (e.g., P2P, Grid). In particular, each enterprise party takes part to the networked system by ex-

posing the information resources to be shared and by providing an ontological representation for them.

2.1 Ontology-based information resource description

In order to support the discovery of relevant information with respect to a target request, information resources need to be described in a way that is understandable and usable by the networked organization. To this end, each enterprise party provides an ontological description of its information resources, according to a Semantic Web-compatible language for its specification (e.g., OWL [16]). In order to describe the interoperability services in an ontology language-independent way, we refer to a reference ontology model, in terms of concepts, properties and semantic relations between concepts. A concept is characterized by a name and a set of properties that represents its features. We distinguish between *strong* and *weak* properties, to denote mandatory properties (i.e., properties with minimal cardinality $\geq$ 1), and optional properties (i.e., properties with minimal cardinality $=$ 0), respectively. Each property is associated with a name and a value, which can be a datatype or a reference to another concept. Semantic relations are defined between concepts, to express the most appropriate relations existing between them. In particular, semantic relations that are specified include the typical relations provided by the Semantic Web languages (e.g., equivalentClass, subClassOf in the OWL language). For example, a detailed description of how to map OWL on the reference model is provided in [5].

As an example, we consider the enterprise party EP_1 which provides information resources related to the travel domain. As shown in Figure 1(a), the EP_1 ontology contains the Accommodation, Reservation, and Travel_document concepts; the Reservation concept is a specialization (i.e., subClassOf) of the Travel_document concept. Furthermore, the value of the property Accommodation of the Reservation concept refers to the Accommodation concept. In Figure 1(b), the fragment of OWL code specifying the Reservation concept is reported. In particular, strong property constraints are defined with a property restriction by setting the minCardinality clause to the value 1 (e.g., reservation_code).

2.2 The peer-based collaboration model

The peer-based collaboration model enforces resource sharing and discovery in open networked systems by exploiting a number of internetworked ontologies. The effectiveness of the knowledge discovery process depends on the capability of retrieving the information related to a target resource request, by exploiting semantic features of enterprise ontology descriptions.

Each enterprise party in the system acts both as a client and as a server in the networked organization interacting with other parties directly, by submitting queries containing a request for one or more resources. The architecture of the peer-based collaboration model is shown in Figure 2. Two different query types are supported in the peer-based collaboration model, namely the *probe query*
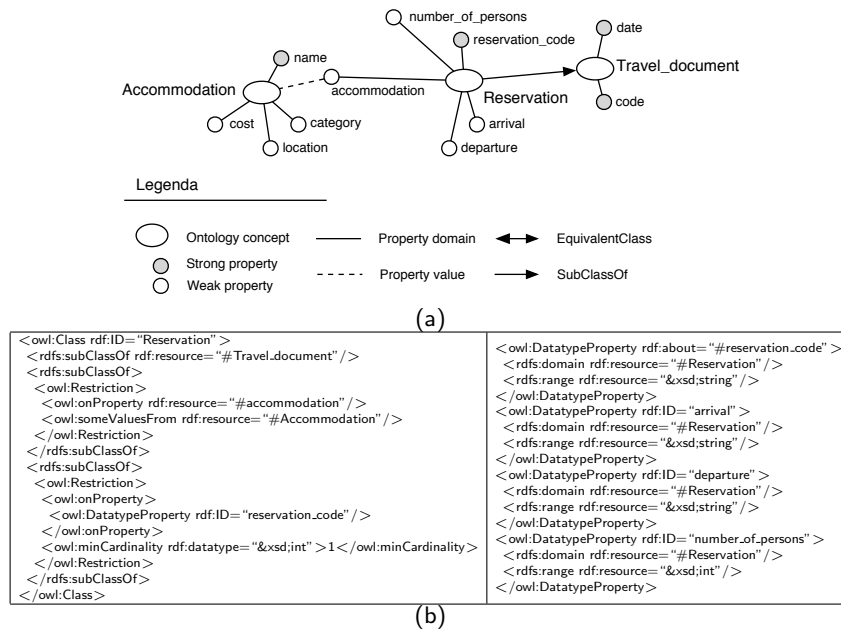
Fig. 1. (a) An example of enterprise party ontology; (b) a fragment of OWL code

and the *search query*. A probe query is used for discovering potential collaborating enterprises, and contains specifications of target concepts describing the resources of interest. A search query is used in order to acquire resource data related to one or more target resources, once a collaborating enterprise has been identified. In order to support ontology-based resource sharing and discovery, each enterprise party realizes three *semantic interoperability services*:

- The *discovery service*: it is responsible for performing the activities related to probe query composition and propagation, and of invocation of probe query processing. The discovery service is invoked by an enterprise party in order to discover which collaborating parties within the networked organization can provide relevant information resources with respect to a target request, based on ontology descriptions.
- The *matching service*: it performs probe query processing against the ontology of an enterprise party. The matching service is based on a set of flexible matchmaking techniques which enable each enterprise party to compare the incoming requests against its ontology in order to identify whether there are information resources (i.e., concepts) matching the target.
- The *acquisition service*: it is related to the accomplishment of the collaboration between two enterprise parties by realizing the acquisition of resource data. The acquisition services of two collaborating parties are responsible for
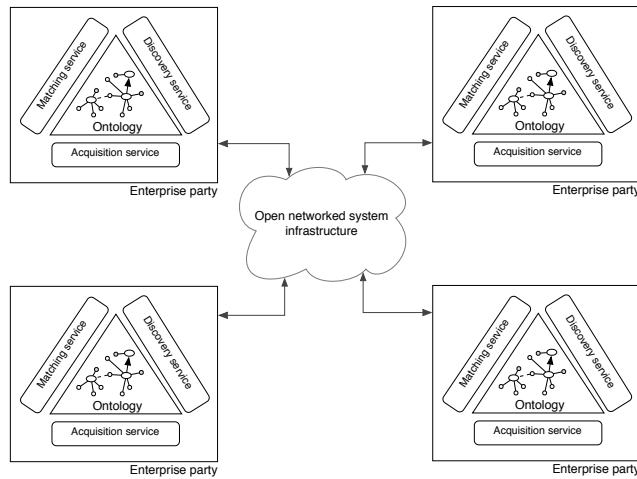
Fig. 2. Architecture of a open networked system for peer-based collaboration

managing search queries and for retrieving resource data by exploiting the
internal services (e.g., Web services) for data access.

In the remainder of the paper, we describe these interoperability services in more
detail and we provide an example of their application for semantic collaboration.

3   Semantic interoperability services

In this section, we describe the interoperability services by showing how they
exploit ontology knowledge for supporting inter-enterprise semantic collabora-
tion.

3.1   The matching service

The matching service is responsible for performing comparison of ontology con-
cept descriptions for probe query processing. According to the ontology model
presented in Section 2, the meaning of ontology concepts depends basically on
the names chosen for their definition and on their contexts, that is, the proper-
ties and the relations they have with other concepts in the ontology. Different
ontologies can describe the same resources using different modeling choices. For
this reason, the matching service is based on a set of ontology matchmaking
techniques capable of coping with different levels of detail in modeling the re-
sources of interest, by considering various ontology elements separately or in
combination. With the goal of providing a wide spectrum of metrics suited for
dealing with many different matching scenarios, ontology matchmaking tech-
niques support four different matching models: surface, shallow, deep, and intensive

matching models. Each model calculates a semantic affinity value $SA_{c,c'}$ of two concepts c and c' which expresses their level of matching. $SA_{c,c'}$ and is produced by considering linguistic and contextual features of concept descriptions [5].

- *Linguistic features.* Linguistic features refer to names of ontology elements and their meaning. To capture the meaning of names in an ontology, the matching service refers to a thesaurus Th of terms and terminological relationships among them. Th is automatically derived from the lexical system WordNet [13]. In order to express the implication of terminological relationships for semantic affinity, a weight $W_{tr} \in [0,1]$ is associated with each terminological relationship tr derived from WordNet.
- *Contextual features.* Contextual features refer to properties and concepts directly related to a given concept (i.e., adjacents) in an ontology. A weight $W_{sr} \in [0,1]$ is associated with each semantic relation sr to denote the strength of the connection expressed by the relation on the involved concepts for semantic affinity evaluation purposes. The greater the weight associated with a semantic relation, the higher the strength of the semantic connection between concepts. Furthermore, a weight W_{sp} is associated with strong properties sp, and a weight W_{wp} with weak properties wp, respectively, with $W_{sp} \geq W_{wp}$ to capture the importance of the property in characterizing the concept for matching.

With respect to the matching models available in the matching service, the surface matching considers only concept names. The shallow matching considers concept names, concept properties, and information about the presence of cardinality constraints on properties. The deep matching model is defined to consider concept names and the whole context of concepts, in terms of properties and semantic relations. In addition to the previous model, the intensive matching includes also property values in the matching process. In each matching model, the semantic affinity evaluation is calculated as a weighted sum of linguistic and contextual features, whose relevance in the semantic affinity evaluation process can be properly established, by setting the weight of the linguistic affinity $W_{la} \in [0,1]$.

We have implemented and tested such matchmaking techniques in the framework of HELIOS, a peer-based system where knowledge sharing and evolution is based on interactions among peer and on peer ontologies for resource description. A more detailed description of these matchmaking techniques is given in [4, 6].

3.2 The discovery service

The discovery service performs all the activities related to composition, propagation, and processing of probe queries.

The discovery service is based on a query/answer paradigm in which there is not a centralized authority managing the collaborations, and the involved parties are dynamically selected based on the semantic affinity of the information resources with the given target. Given a target request, the discovery service is invoked in

order to query a set of prospective collaborating parties and evaluate the affinity with respect to the target using ontology descriptions. The receiving parties, compare the request against their ontologies and reply whether they can provide relevant information resources (i.e., resources matching the target). In particular, query management requires an expressive representation capable to support the description of target resources in terms of ontology concepts searched (target concept(s)), with possible properties and semantic relations. To this end, a reference query template for information resource discovery is provided in Figure 3 and it is composed of the following clauses:

- *Find:* list of target concept(s) names.
- *With:* (optional) list of properties of the target concept(s).
- *Where:* (optional) list of conditions to be verified by the property values, and/or (optional) list of concepts related to the target by a semantic relation.
- *Matching model:* (optional) specification of the matching model asked by the requesting peer to process the query.
- *Matching threshold:* (optional) specification of the threshold value t, with $t \in (0, 1]$ to be used for the selection of matching concepts based on the semantic affinity value determined by the matching process. If a matching threshold is not specified in the query, the answering peer adopts its own default threshold.

Probe query template

Find	Target concept name [, ...]
[*With*	⟨Property name⟩ [, ...]]
[*Where*	Condition, ⟨related concept, semantic relation name⟩ [, ...]]
[*Matching model*	Matching model to be used]
[*Matching threshold*	$t \in (0, 1]$]

Fig. 3. The reference query template

At the same, the discovery service has to provide an expressive representation of query answers. A query answer can be considered as a list of matching concepts. As described in Figure 4, the structure of the answer template contains the following clauses:

- *Concept:* name of the matching concept.
- *Properties:* (optional) list of properties of the matching concept.
- *Adjacents:* (optional) list of concepts related to the matching concept by a semantic relation.
- *Matching:* set of pairs ⟨target concept, affinity value⟩, specifying the target concept with which the matching concept matches, together with the corresponding affinity value.

– *Matching model:* specification of the matching model applied to process the query.
– *Matching threshold:* (optional) specification of the threshold value t, with $t \in (0, 1]$ used for the selection of matching concepts based on the semantic affinity value determined by the matching process.

Probe answer template

{*Concept*	Concept name
[*Properties*	⟨Property name⟩ [, ...]]
[*Adjacents*	⟨related concept, semantic relation name⟩ [, ...]]
Matching	⟨Target concept, affinity value⟩[, ...]
Matching model	Matching model name
[*Matching threshold*	$t \in (0, 1]$]]}

Fig. 4. The reference answer template

3.3 The acquisition service

The acquisition service is responsible of the collaboration establishment between two enterprise parties. When a requesting enterprise has identified a relevant partner for the collaboration, it sends a search query in order to access and acquire data about shared information resources. Search queries contain the *Find, With* and *Where* clauses of the probe query template. An enterprise party provides appropriate techniques to access its repositories containing resource data, in order to support search query processing.

– The *Web Service-based* approach. Each enterprise party provides a standard access to its shared information resources by means of a Web Service. Standard protocols (e.g., SOAP, WSDL) can be adopted by the acquisition service in order to interact with the Web Service and provide a seamless access to the underlying information resources. The acquisition service interacts with the Web Service by means of the SOAP protocol which supports well-defined XML-based message communications. The WSDL document provides the specification of the set of methods which can be invoked by the acquisition service as well as the structure of the returned data extracted from the information resources.
– The *Wrapper-based* approach. The access to the shared information resources is provided by means of a wrapper service. The acquisition service interacts with the wrapper service by submitting target queries for information resource access. The wrapper service manages mapping rules defining the concepts of the enterprise ontology and the underlying information resources in order to reformulate the target query in terms of queries over specific

structures of the repository where the information resources are stored (e.g., relational database structure). Finally, the answer to a search query sent back to the requesting party is an XML document containing the information about the shared information resources.

4 Exploiting interoperability services for semantic collaboration

In this section, we describe the use of interoperability services in the peer-based collaboration model wuth an application example.

4.1 The peer-based semantic collaboration model

In Figure 5, we outline the main interactions, together with service invocations, between two collaborating parties in the peer-based model. An enterprise party
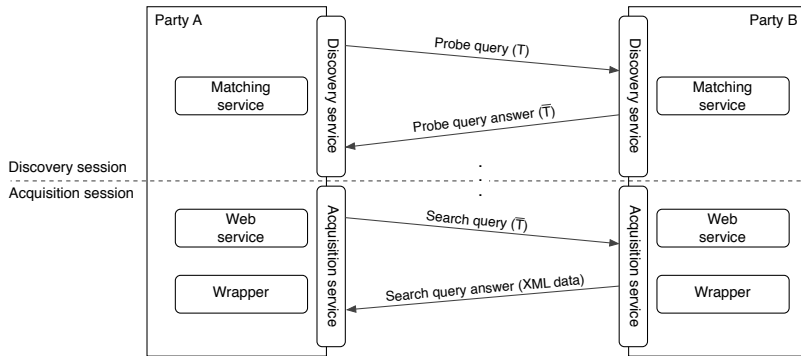

Fig. 5. Interactions and services in the peer-based collaboration model

invokes its discovery service when it intends to find potential collaborating enterprises of the networked organization capable to provide information semantically related to one or more target concepts of interests. According to the query template of Figure 3, a probe query is composed by exploiting the discovery service containing the specified target concepts. Such a request is submitted to the other enterprise parties (i.e., B in Figure 5). In the context of a peer-based collaboration, many enterprise parties can be involved in the organization, and if each probe query is propagated to all the nodes of the network, performance and efficiency of the overall system can drop dramatically. For this reason, the discovery service of the requesting enterprise party defines semantic routing rules in order to restrict the query propagation to a subset of nodes following semantic criteria: the probe query is sent to the parties whose ontologies are expected to contain

relevant concepts with respect to the target query. Preliminary results on this topic are described in [14]

Receiving a probe query, the discovery service of a receiving enterprise party interacts with the matching service in order to identify if there are concepts matching the target request. In particular, the discovery service provides to the matching service an ontological description of the target concept(s) (extracting such information from the Find, With, and Where clauses), as well as the matching model and the threshold to apply (derived from the Matching model and the Threshold clauses, respectively). As a result, the matching service returns a (possibly empty) ranked list of concepts semantically related to the target, and, for each entry, the corresponding semantic affinity value in the range $(0, 1]$, computed by the ontology matchmaking techniques. Finally, the results of the matching service are organized according to the answer template in Figure 4 by the discovery service, and such an answer is replied to the requesting enterprise party (i.e. the answer $\overline{T}$ in Figure 5).

Collecting query replies from answering parties, the requesting enterprise evaluates the results and decides whether to establish a collaboration with the enterprise parties found to be relevant by the discovery service (collaborating enterprise(s)). To this end, the acquisition service is invoked to send an appropriate search query formulated over the matching concepts of each collaborating enterprise (i.e., the concepts provided in the probe query answer during the discovery session). The acquisition service retrieves data for a search query according to the information resource access approach supported by the answering party, and sends back resource data in an XML format (i.e., search query answer in Figure 5).

4.2 Application example

As an example of discovery in a peer-based collaboration, we consider networked organization composed of four enterprise parties, namely EP_1, EP_2, EP_3, and EP_4. As shown in Figure 6, the ontology of each enterprise party contains concepts related to the travel domain. We suppose that EP_1 intends to discover whether any enterprise party in the networked organization can provide relevant resources with respect to the Accommodation concept. To this end, EP_1 invoke the discovery service and submits to the system the query in Figure 7. Exploiting its semantic routing rules, EP_1 sends the query to EP_2 and EP_3. The discovery service of EP_2 and EP_3 catches the incoming query and invokes the matching service which performs ontology matching using the deep model to evaluate the semantic affinity between the incoming query and the concepts contained in each enterprise ontology. According to the matching service results, the discovery service of EP_2 replies to EP_1 with the answer shown in Figure 8. Following the same procedure, EP_3 does not identify matching concepts over the specified threshold in its ontology, since the matching results between the concepts flight and car and the target concept accommodation are 0 and 0.125, respectively. Nevertheless, by exploiting its semantic routing rules, EP_3 forwards the query to EP_4 which is expected to provide semantically related concepts. According to
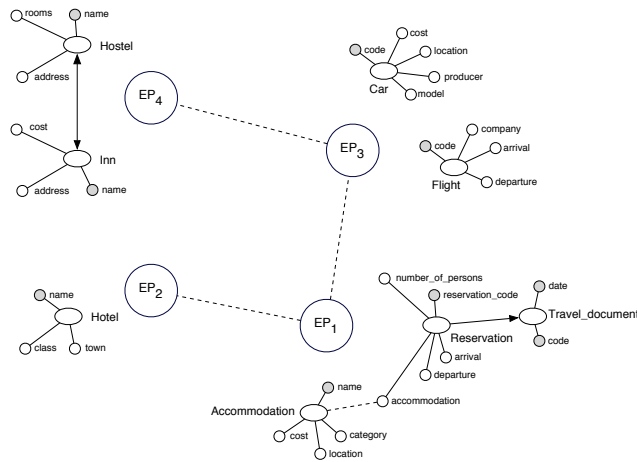
Fig. 6. An example of networked enterprise ontologies

FIND	accommodation
WITH	name, category, location
Matching model	deep
Matching threshold	0.5

Fig. 7. The probe query submitted by EP_1 for resource discovery

its matching service results, the discovery service of EP_4 replies directly to EP_1 with the answer shown in Figure 9.

5 Concluding remarks

In this paper, we have presented an ontology-based collaboration model and semantic interoperability services for supporting discovery and sharing of heterogeneous information resources. Future research issues will regard the implementation of the semantic interoperability services. For what concern the matching service, we have developed H-MATCH, an algorithm for ontology matching and we are now testing performance and effectiveness of such an algorithm [5]. In the context of the discovery service, we are working on the development of a semantic routing protocol, in order to enhance the discovery functionalities by adopting advanced network protocols, capable of taking into account information about semantic neighborhood among nodes of the organization for semantic query forwarding [14]. Finally, we are working on the definition of flexible techniques for

```
CONCEPT                hotel
PROPERTIES             name, class, town
MATCHING               ⟨accommodation, 0.75⟩
MATCHING MODEL         deep
MATCHING THRESHOLD     0.5
```

Fig. 8. The answer provided by EP_2

```
CONCEPT                hostel
PROPERTIES             name, address, rooms
ADJACENTS              ⟨inn, equivalentClass⟩
MATCHING               ⟨accommodation, 0.625⟩
MATCHING MODEL         deep
MATCHING THRESHOLD     0.5
CONCEPT                inn
PROPERTIES             name, address, cost
ADJACENTS              ⟨hostel, equivalentClass⟩
MATCHING               ⟨accommodation, 0.625⟩
MATCHING MODEL         deep
MATCHING THRESHOLD     0.5
```

Fig. 9. The answer provided by EP_4

the acquisition service capable to deal with different scenarios in information resource access.

References

1. H. Afsarmanesh, C. Garita, and L.O. Hertzberger. Virtual Enterprises and Federated Information Sharing. In *Proc. of the 9th Int. Conference on Database and Expert Systems Applications (DEXA 1998)*, pages 374–383, Vienna, Austria, August 1998. LNCS.
2. T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web. *Scientific American*, May 2001.
3. P. Bouquet, B. Magnini, L. Serafini, and S. Zanobini. A SAT-based Algorithm for Context Matching. In *Proc. of the 4th Int. and Interdisciplinary Conference on Modeling and Using Context (CONTEXT 2003)*, pages 66–79, Stanford, CA, USA, June 2003. Springer Verlag.
4. S. Castano, A. Ferrara, and S. Montanelli. H-MATCH: an Algorithm for Dynamically Matching Ontologies in Peer-based Systems. In *Proc. of the 1st VLDB Int. Workshop on Semantic Web and Databases (SWDB 2003)*, Berlin, Germany, September 2003.
5. S. Castano, A. Ferrara, S. Montanelli, and G. Racca. From Surface to Intensive Matching of Semantic Web Ontologies. In *Proc. of the 3rd DEXA Int. Workshop on Web Semantics (WEBS 2004)*, Zaragoza, Spain, September 2004. IEEE Computer Society.

6. S. Castano, A. Ferrara, S. Montanelli, and G. Racca. Matching Techniques for Resource Discovery in Distributed Systems Using Heterogeneous Ontology Descriptions. In *Proc. of the Int. Conference on Coding and Computing (ITCC 2004)*, Las Vegas, Nevada, USA, April 2004. IEEE Computer Society.

7. S. Castano, A. Ferrara, S. Montanelli, and D. Zucchelli. HELIOS: a General Framework for Ontology-based Knowledge Sharing and Evolution in P2P Systems. In *Proc. of the 2nd DEXA Int. Workshop on Web Semantics (WEBS 2003)*, Prague, Czech Republic, September 2003. IEEE Computer Society.

8. A. Doan, J. Madhavan, P. Domingos, and A. Halevy. Learning to Map between Ontologies on the Semantic Web. In *Proc. of the 11th Int. World Wide Web Conference (WWW 2002)*, pages 662–673, Honolulu, Hawaii, USA, May 2002.

9. J. Broekstra et al. A Metadata Model for Semantics-Based Peer-to-Peer Systems. In *Proc. of the 1st WWW Int. Workshop on Semantics in Peer-to-Peer and Grid Computing (SemPGRID 2003)*, Budapest, Hungary, May 2003.

10. W. Nejdl et al. EDUTELLA: a P2P Networking Infrastructure Based on RDF. In *Proc. of the 11th Int. World Wide Web Conference (WWW 2002)*, Honolulu, Hawaii, USA, May 2002.

11. A. Kementsietsidis, M. Arenas, and R. J Miller. Mapping Data in Peer-to-Peer Systems: Semantics and Algorithmic Issues. In *Proc. of ACM SIGMOD Int. Conference on Management of Data*, San Diego, California, USA, June 2003.

12. L. Li, B. Wu, and Y. Yang. An Ontology-Oriented Approach for Virtual Enterprises. In *Proc. of the 6th Asia-Pacific Web Conference (APWeb 2004)*, pages 834–843, Hangzhou, China, April 2004. LNCS.

13. G. A. Miller. WordNet: A Lexical Database for English. *Communications of the ACM (CACM)*, 38(11):39–41, 1995.

14. S. Castano and A. Ferrara and S. Montanelli and E. Pagani and G. P. Rossi and S. Tebaldi. On Combining a Semantic Engine and Flexible Network Policies for P2P Knowledge Sharing Networks. In *Proc of the 1st DEXA Workshop on Grid and Peer-to-Peer Computing Impacts on Large Scale Heterogeneous Distributed Database Systems (GLOBE 2004)*, Zaragoza, Spain, September 2004. IEEE Computer Society. To Appear.

15. N. Silva, J. Rocha, and J. Cardoso. E-Business Interoperability Through Ontology Semantic Mapping. In *Proc. of the Processes and Foundations for Virtual Organizations, IFIP TC5/WG5.5 4th Working Conference on Virtual Enterprises (PRO-VE 2003)*, pages 315–322, Lugano, Switzerland, October 2003. Kluwer.

16. Michael K. Smith, Chris Welty, and Deborah L. McGuinness (Eds.). OWL Web Ontology Language Guide, 2004.