Ontology-based Interoperability for Interorganizational Applications ⋆

D. Bianchini, V. De Antonellis, M. Melchiori

Università di Brescia
Dip. Elettronica per l'Automazione
Via Branze, 38
25123 Brescia - Italy
bianchin|deantone|melchior@ing.unibs.it

Abstract. In this paper we present an ontology-based approach to support interoperability in interorganizational applications. The ontology contains knowledge, coming from several organizations, structured in different layers to support its effective use and communication. The proposed approach has been experimented in the framework of the Italian VISPO (Virtual district Internet-based Service PlatfOrm) project.

1 Introduction

In recent years, organizations are increasingly looking for opportunities to exploit innovative technologies that use Internet, mobile and wireless devices to enhance communication and cooperation in providing information and services. Modern inter- and intra-organizational applications specifically need to support understanding of shared knowledge.

In literature, different approaches address the problem of interoperability in distributed dinamically evolving environments and propose ontologies as a suitable solution [2, 11]. Several proposals involve the design of enterprise ontologies and many research groups are studying and developing methods and tools for the definition and use of ontologies (*http://ontoweb.aifb.uni-karlsruhe.de/*). The OntoWeb network [11] main goal is to bypass communication bottlenecks among various and heterogeneous research groups and organizations. Ontologies are considered essential part in supporting information exchange processes and business transactions, providing on-line unified access to large volumes of information and knowledge based on machine-processable semantics of data. Tool-supported methodologies for ontology design and several infrastructures for search and reuse of distributed ontologies are proposed in [10, 14, 12]. In [9, 3] the focus is on ontology-based integration of datasources. An open issue regards effective users support to identify knowledge assets. Many research activities are devoted to study the problems of providing different visions of the same realm by different organizations, taking into consideration semantical aspects of the involved datasources.

In this paper, we propose an ontology-based approach to support interoperability in interorganizational applications according to a three-layer ontology

⋆ This work has been partially supported by the Italian MIUR VISPO (Virtual district Internet-based Service PlatfOrm [15]) project and the European EU NoE INTEROP.

architecture. The ontology acts as an informative infrastructure to aggregate information exchanged among organizations that cooperate for business purposes giving rise to a virtual district. The ontology provides an access point to information at local level and supports service aggregation and distribution. In particular, we discuss the use of a three-layer ontology in the virtual district scenario by presenting some results of our research and development activity in the Italian VISPO (Virtual district Internet-based PlatfOrm) project [7, 15]. Starting from heterogeneous XML-based and XML-compliant datasources, a domain ontology is designed in a semi-automated way with the support of the ARTEMIS tool environment [1]. Knowledge in the ontology is organized intro three layers by means of: clusters of similar concepts coming from different sources (*semantic mapping layer*); unified global concepts and semantical relationships between them (*mediation layer*); subject categories derived from available standard taxonomies (*categorization layer*). In particular, we focus on ontology use and deployment.

This paper is organized as follows: Section 2 presents the considered context and provides motivations for the proposed approach; Section 3 presents the ontology design approach; Section 4 shows different modalities to exploit the interorganizational knowledge represented in the domain ontology; Section 5 presents concluding remarks.

2 Ontology support to interoperability and communication intra- and inter-organizations

The availability of methodologies to construct ontologies is a critical issue in supporting the communication and cooperation among different organizations. In fact, the explicit and shared representation of interorganizational knowledge is a main prerequisite for exchanging information and services. According to the experience gained in the VISPO Project [15], that studied the definition of methods and architectures to support the activities of a virtual district, we illustrate two application profiles for ontologies that match some relevant requirements on knowledge representation that we collected in the two considered virtual districts. In the following sections, we describe the ontology architecture, the related design methodology and the deployment primitives that support these application profiles. The requirement analysis developed in the VISPO Project confirmed that various services and tools that support the business activities of the district are to be based on shared knowledge representations that can be usefully organized as ontologies. In particular, this need arises in the application profiles we present: (i) service for the analysis of purchase requests, and (ii) e-procurement service with aggregation of purchase requests. In the first case, we describe the knowledge requirements for the design of services that support the analysis of purchase requests generated into a single company. In the other case, we analyze the requirements of interorganizational knowledge for designing an e-procurement service devoted to companies that would like to aggregate their purchase orders regarding the same products for the purpose of obtaining a commercial advantage and therefore lower prices from suppliers. The context of developing a support for the analysis of purchase requests leads to identify an intra-organizational use of our methodology for ontology construction, while the context of e-procurement services design leads to a inter-organizational use of the methodology.

2.1 An intra-organizational use of three-layer ontology in a Virtual District

In this considered context, we address the use of ontology in the analysis of internal purchase orders, that is, a set of auditing activities to analyze and control the sources, the flows and the volumes of the orders made from the departments of a company. In this case, the employees of each department usually formulates purchase requests according to a terminology and descriptions that are not completely normalized, but vary from department to department and often from employee to employee. The requests are directed and processed from the Purchase Department of the company that is responsible to collect the purchase requests and to send the orders to the suppliers. From the perspective of the analysis of internal purchase orders, the problem is to identify all the requests referred to the same product or to the same product category, since they can be possibly described differently in the various requests and purchase transactions. The main requirement, in this case, is the need for an internal catalog that provides standardized descriptions for the purchase requests. In this way, in fact, it is possible to evaluate the ordered amounts of a given product that results uniquely identified. The internal catalog can be constructed by integrating the terminology and the descriptions of products contained in the purchase requests. In the experimentation of our methodology in the context of the VISPO Project, an internal catalog has been constructed according to the three-layer ontology, presented in Section 3, by integrating the description of purchase requests extracted from the ERP of a selected company and the product descriptions obtained from two public industrial catalogs. The application of our methodology and, in particular, the integration with descriptions from industrial catalogs provide the following benefits: (i) the descriptions and the terms identifying products are made homogeneous and are better standardized, (ii) the descriptions of products are enriched on the basis of the industrial catalog descriptions. The internal catalog can be exploited to formulate a new purchase request in terms of the global concepts the catalog contains (for instance, an `hexagonal-head screwdriver - 4mm diameter`) and, in this case, the tree-layer architecture is followed in a top down way, that is the ontology is browsed from the categorization layer down to the mediation layer until the desired product is identified. If a specific product is required (for instance, an `exagonal-head screwdriver` of given brand and model) the architecture is used bottom-up to classify the product in the associated global concept. The Figure 1 shows schematically the process of standardization of the internal catalog.

2.2 An inter-organizational use of three-layer ontology in a Virtual District

The relevant issue in developing the e-procurement service is that the different purchase requests have to be aggregated into a single order from the e-procurement service to the supplier that provides the better conditions for the ordered items. It is supposed that the e-procurement service relies on different suppliers, each of ones has its own catalog, so from the e-procurement perspective the problem is: (i) to provide a unified supplier catalog from which the client companies can decide their orders and (ii) to map the product description of the unified catalog onto the suppliers' ones. This scenario corresponds

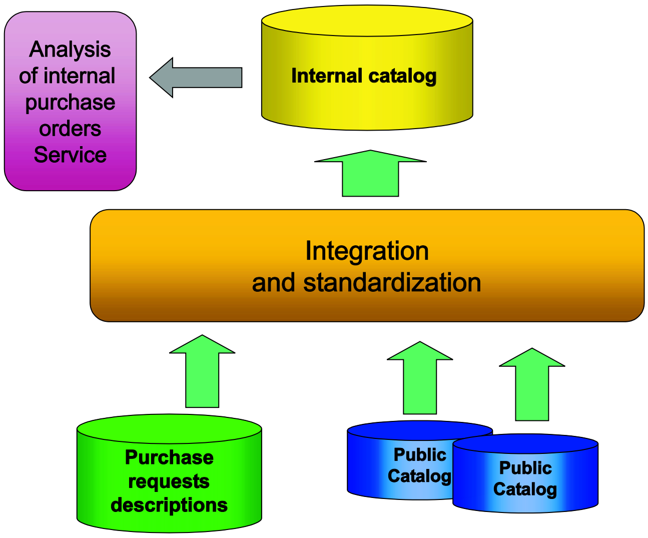

Fig. 1. The internal catalog standardization process

to the well known problem of catalog integration that we analyze with respect to an inter-organizational use of our ontology construction methodology. In general, to enable the integration of more catalogs that list products in overlapping domains, we need to identify correspondences among similar concepts and to construct unified representations of them that maintain references to the original catalogs. This allows one to search for a given product in the unified catalog and, after the product is found, accessing to the associated products listed in the supplier catalogs. The identification of correspondences is required at different levels:

– different catalogs use different product taxonomies as tables of contents; we need to establish mappings among the taxonomies since, in general, they use different terminology and have different organizations (see Fig. 2);
– we need to mediate and to establish correspondences among different representations of similar products adopted from different catalogs and possibly to map them into standardized representations.

A unified catalog represented as three-layer ontology that satisfies the presented requirements can be constructed as illustrated in the next sections, where in particular the products are represented as global concepts in the ontology. On the basis of this ontology, an e-procurement service can be designed to support the aggregation of purchase requests from the companies of the virtual district. In fact, if the requests are referred to products chosen in the unified catalog, the e-procurement service can classify the ordered products of different purchase requests, aggregate the requests related to homogeneous products into single purchasing transactions that will regard bigger quantities, so obtaining greater discounts from suppliers. In this scenario, our thee-layer ontology can be used according to a top-down perspective, where the e-service browses the ontology

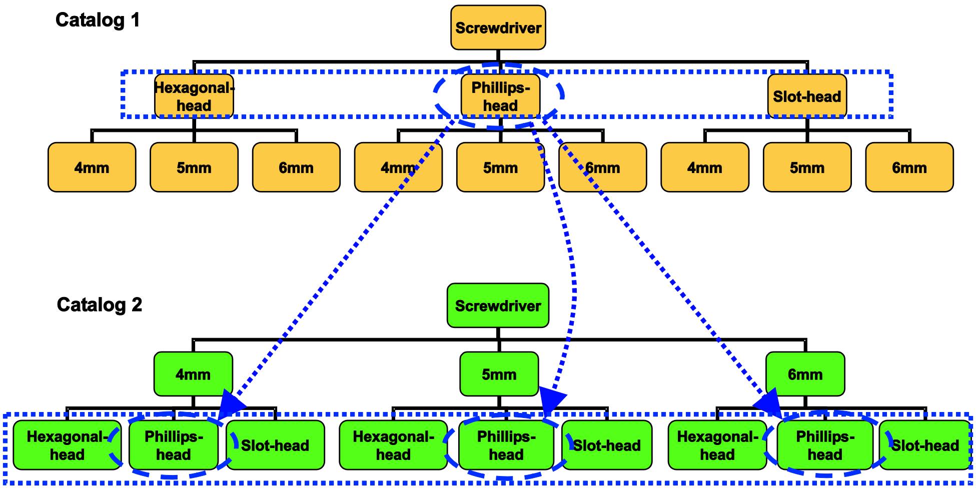

Fig. 2. Heterogeneities in taxonomies of product catalogs

starting from the global concepts and reformulates a purchase request made in terms of global concepts into one referring products in a supplier catalog. The application of ontology to this context is schematically shown in the Figure 3.

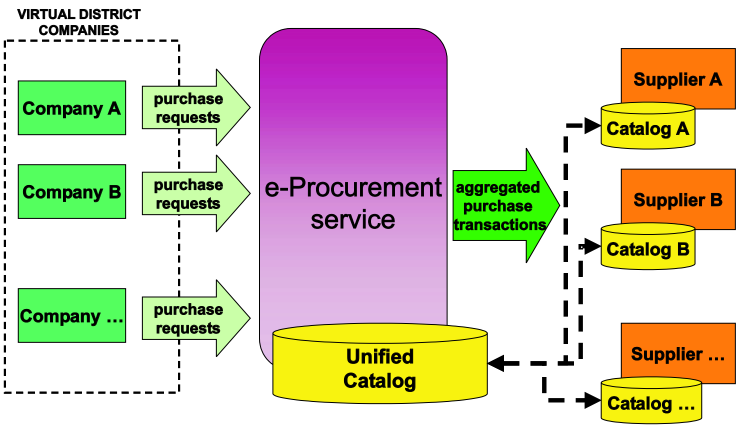

Fig. 3. Tree-layer ontology use in a context of e-procurement

3 Ontology design

To support interoperability and communication in a virtual district scenario, we proposed a three-layer ontology architecture which provides a unified semantic representation of interorganizational knowledge in the considered domain by means of ontological concepts and semantic relationships between them [6]. The ontology is composed by a set of concepts and relationships (inter-layer and intra-layer links) between them.

Ontological concepts. The three layers of the ontology architecture organize interorganizational knowledge through three main kinds of concepts:

- XClasses, that are conceptual elements of the original datasources expressed using a common formalism, the X-Formalism, presented in [4] (*semantic mapping layer*);
- global concepts, that are global XClasses obtained by unification of similar XClasses in different datasources (*mediation layer*);
- subject categories, belonging to available standard taxonomies (*categorization layer*).

These different kinds of concepts are specified as follows. We assume that information in the considered datasources is expressed using XML-based or XML-compliant schema languages. An XClass is thus described by a name, a set of properties or attributes (with simple or built-in data types, such as `string`, `NMTOKEN` or `integer`, and some cardinality constraints) and a set of references to other XClasses (with cardinality constraints). A global XClass is associated to a cluster of XClasses, that are grouped on the basis of their semantic similarity [5, 9]. A global XClass results from the unification of similar XClasses by means of rules for the reconciliation of names, types and cardinality constraints of properties and referenced XClasses. Each global XClass defines a global concept. To provide topic-based view of underlying layers, global concepts are related to subject categories relevant to the domain of interest, as provided in several standard taxonomies. There are several proposals for standard classifications in literature. In particular, in our work we considered the UNSPSC taxonomy [8].

Figure 4 shows a portion of three-layer ontology built in the VISPO project to support e-procurement services in a virtual district operating in the industrial supply market. We have considered two on-line industrial catalogs, the Beta[1] and Usag[2] catalogs, containing detailed descriptions of their products, and a third catalog provided by a company of the virtual district with less detailed product descriptions. From each catalog, we extracted the product descriptions, represented through XClasses, and we grouped them into clusters on the basis of their similarity in the semantic mapping layer (for example, the XClasses `Toggle joint shears for sheet-steel` and `Shears for sheet-steel`). Then, we unified descriptions belonging to the same clusters into global concepts in the mediation layer (for example, the concept `Shears`), finding semantic relationships between them (e.g., `generalization` relationships between `Shears` and `Blades`). Finally, global concepts have been related to subject categories of the UNSPSC taxonomy in the categorization layer.

Intra-layer and inter-layer links. XClasses, global concepts and subject categories at different levels of abstraction are related by means of inter-layer

[1] "http://www.beta-tools.com/"
[2] "http://www.usag.it/"

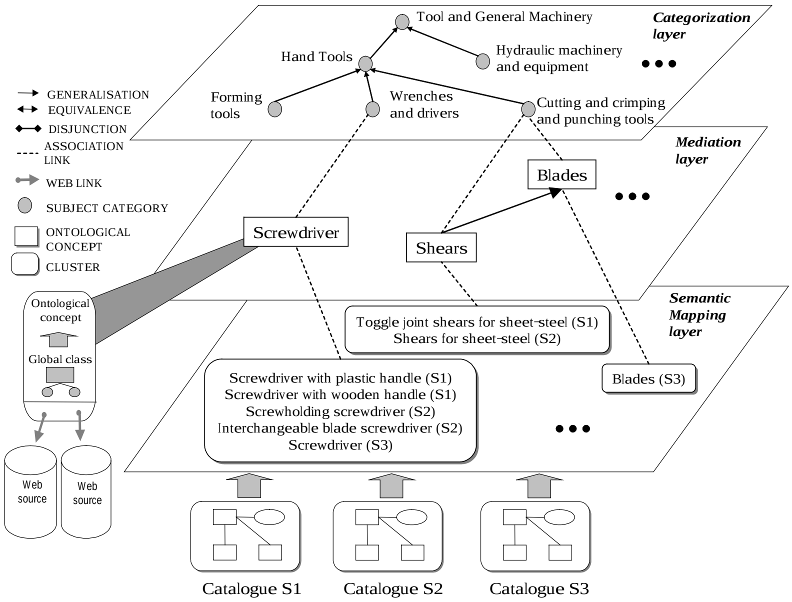

Fig. 4. A portion of three-layer ontology.

and intra-layer links that can be used to browse the three-layer ontology as explained in Section 4. In the semantic mapping layer, local XClasses belonging to the same cluster are related each other by means of `similar-to` relationships, obtained through the evaluation of their structural and name affinity [5]. Clusters of XClasses are connected to corresponding global XClasses (global concepts) in the mediation layer through `association` links.

Global concepts in the mediation layer are organized by means of semantic relationships. We consider three kinds of relationships: (i) *generalization*, a concept α generalizes another concept β if the set of instances of α includes the set of instances of β; (ii) *disjunction*, two concepts α and β are disjoint if the sets of their instances are disjoint; (iii) *equivalence*, two concepts α and β are equivalent if the sets of their instances coincide.

Finally, *association* links are maintained between global concepts and subject categories in the categorization layer. Subject categories are organized in a generalization taxonomy.

Ontological elements and intra/inter-layer links are represented in a common manner by means of the frame structure shown in Table 1, that can be easily exploited by a software agent for the ontology deployment (see Section 4), apart from the logical language used to implement the ontology. Note that not all the fields of the frame structure are always mandatory for every ontological element. Source links for a global concept are obtained by means of the union of source links for local XClasses in the associated cluster; in the same manner, source links for subject categories are the union of source links for associated global concepts (in the case of leaf subject categories) or of source links for the specialized subject categories (otherwise in the taxonomy).

Property	**Description**
Name	Name of the ontological element
Type	XClass, global concept or subject category
Property	List of properties of the global concept or XClass, with associated types and cardinality constraints; empty for subject categories
Kind-of	Names of elements that generalize the current one (`generalization` relationship in the mediation and categorization layer)
Equivalent-to	Names of ontological elements that are related to the current one through an `equivalence` relationship between global concepts in the mediation layer and a `similar-to` relationship between local XClasses in the semantic mapping layer
Disjunction	Names of elements that are disjoint from the current one, in the case of global concepts in the mediation layer
Association	Names of elements that are related to the current one by means of an `association` link (inter-layer links)
SourceLinks	Links to the datasources to which the current element is related

Table 1. The common frame structure for ontological elements.

In [6] we presented a methodology for the construction of the three-layer ontology, articulated into four main steps:

1. *data analysis and conceptualization*, to extract XClasses from datasources and to cluster similar XClasses;
2. *integration*, to unify similar XClasses into the global XClasses;
3. *synthesis and categorization*, (i) to define global concepts and semantic relationships between them starting from global XClasses; (ii) to relate global concepts to subject categories;
4. *implementation*, to formally represent the ontology and to iteratively refine and test the ontology concepts.

The construction of the three-layer ontology is supported by ARTEMIS [1], a semi-automated tool environment which supports the domain expert in extracting information from datasources, integrating global concepts, identifying the semantic relationships and querying the ontology contents. Figure 5 shows the ARTEMIS architecture.

4 Ontology deployment

Domain knowledge ontology is a very useful tool to provide an access point to the interorganizational information within the virtual district and support discovery of information for business purposes. Several searching modalities can be exploited in the three-layer ontology, taking into consideration all different levels of abstraction, experience of user in the considered domain and the kind of the

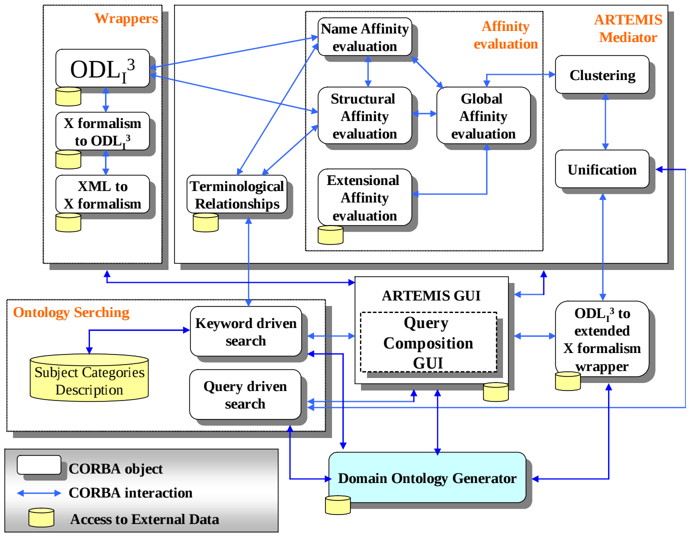

Fig. 5. The ARTEMIS tool environment architecture.

user: human or software agent. The different searching modalities rely on a set of primitives that permit to software agents or applications to exploit ontology searching capabilities and that are listed in Table 2.

Category-driven navigation. The taxonomy of subject categories can be browsed to find the desired one and visit datasources directly connected to it; it is also possible to reach ontological concepts associated to that category and use other searching modalities. According to the portion of ontology shown in Figure 4, the user can browse the UNSPSC taxonomy from the `Tool and General Machinery` category to the `Wrenches and drivers` category and visualize all the datasources associated to the second one; the underlying tool environment supports the user executing the primitives (1) and (2) to browse the taxonomy and the primitive (7) to visualize datasources.

Concept-driven navigation. Global concepts of the mediation layer are exploited as starting points to build queries on the global concepts in the considered domain, propagating these queries towards each involved datasources; in this manner, only one query is requested on the global concepts describing the domain, instead of multiple queries submitted to all datasources; moreover, it is not necessary to know location, terminology and content for each datasource; results of the queries are obtained exploiting association links between global concepts and local XClasses and are combined to obtain the query answer to be presented to the user; we consider a query language proposed by ODMG-93 standard [13], OQL (Object Query Language), a superset of SQL'92 query language; in particular, a subset of OQL-like queries is used:

Primitive	Description
(1) $\{Category\}\ generalizationOf\ (sc)$	Returns all super-categories of a category sc
(2) $\{Category\}\ specializationOf\ (sc)$	Returns all sub-categories of a category sc
(3) $\{OntologicalElement\}\ elementFrom\ (t)$	Returns all ontological elements whose names contain the term t
(4) $``NULL'' \cup relType\ getRelationBetween\ (c_i, c_j)$	Returns the relation type, if exists, between global concepts c_i and c_j
(5) $\{ < attrName, attrType > \}\ attributeOf\ (c_i)$	Returns all the attribute names and types of the global concept c_i
(6) $\{ < Concept, relType > \}\ relationsOf\ (c_i)$	Returns all the concepts that are related to the concept c_i with the corresponding relation type
(7) $\{\ Sources\ \}\ getSourcesOf\ (e_j)$	Returns all the sources related to the ontological element e_i

Table 2. Primitives for ontology deployment.

```
query ::= select *
     from concept 1{, concept 2,...}
     [where condition]
```

This query contains a set of global concepts, a boolean condition on the attributes of them and returns the instances of concepts involved in the query that satisfy the condition. To perform query reformulation and merging of the answers obtained from the datasources, we exploit the mapping rules defined in the mediation scheme, where correspondences between global and local features are represented. In the considered example, suppose that the user retrieves the ontological concept `Blades` through one of the different searching modalities; at this point, he searches for other ontological concepts that are related to `Blades` by means of semantic relationships (the system uses the primitives (4) and (7)) and chooses the concepts related to the current one by the `generalization` relationship (in the example, only the concept `Shears`); finally, he requires datasources related to all visited ontological concepts by means of the primitive (7).

Keyword-driven search. A traditional keyword-driven search can be used to find desired ontological concepts (both at global and at local level of abstraction); users can specify one or more keywords and these terms are matched against the ontology, comparing them with names of ontological elements: (i) in case of names of global concepts, sources related to the matching concept and to its specializations are considered and the list of URL of these sources is presented to the user; (ii) in case of names of subject categories, sources related to the matching category and to its subcategories are considered and the list of URL of these sources is presented to the user. When more than one keyword is specified, the search is repeated for each keyword; in particular, if the keywords are concatenated by the AND operator, we consider the intersection of results obtained for single keywords, while if the keywords are concatenated by the OR operator, we consider the union of the results. As an example, the user specifies the keyword `Screwdriver`; the system visualizes all ontological elements that contain the word `screwdriver` or its synonyms using the primitive (3); the result of the search is the ontological concept `Screwdriver` and local XClasses belonging to the associated cluster; the system visualizes the corresponding datasources using the primitive (7).

Experience of the user about the considered domain and the ontology structure influences the choice of the searching modality: users with less experience generally prefer the category-driven or the keyword-driven search, while the concept-driven search requires more knowledge about query formulation (even if graphical interfaces are used to support this task) and about use of SQL-like query languages. An ontology-based search engine (for which we propose an architecture, described in Figure 6) provides users with an overview of categories, concepts and relevant relationships in the considered domain, supporting them in determining the information that better matches their needs. This kind of search engine can be considered to be complementary to traditional Web search engines which are based only on keyword occurences.

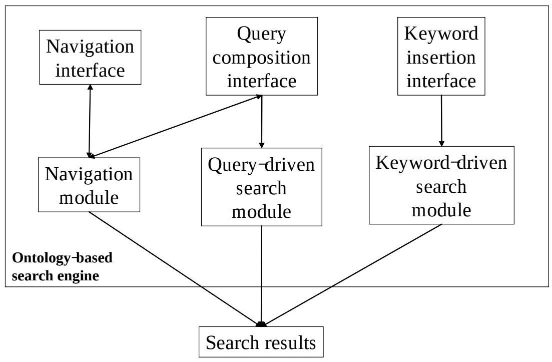

Fig. 6. Architecture of an ontology-based search engine.

5 Conclusions and future work

In this paper we have proposed an ontology-based approach to support interoperability and communication in interorganizational applications, where different organizations cooperate and share information to expand their market possibilities. Our approach organizes ontologies in three abstraction levels to represent the interorganizational knowledge, providing a global view on heterogeneous datasources, where the same information can be represented in different ways. Several searching modalities to exploit the proposed ontological representation of knowledge are provided. A semi-automated tool environment both for ontology design and for searching support is being completed.

Future activities are devoted to the refinement of the representation language to express ontological concepts and semantical relationships at each level of the ontology and to the study of methods and technologies for domain ontology test, reuse and consistency checking, also in the framework of the European INTEROP network of excellence.

References

1. The ARTEMIS Project Home Page. `http://www.ing.unibs.it/~deantone/interdata_tema3/Artemis/artemis.html`.
2. D. Bianchini, V. De Antonellis and M. Melchiori. Domain ontologies for virtual knowledge sharing and service composition in virtual district. In *Proc. of Int. Workshop on Web Semantics (WebS2003)*, Prague, Czech Republic, September 1-5 2003.
3. A. Calì, D. Calvanese, G. De Giacomo, M. Lenzerini, P. Naggar and F. Vernacotola. IBIS: Semantic Data Integration at Work. In *Proc. of the 15th Int. Conference on Advanced Information Systems Engineering (CAiSE 2003)*, Klagenfurt, Austria, June 16th-18th 2003.
4. S. Castano, V. De Antonellis, S. De Capitani and M. Melchiori. Semi-automated Extraction of Ontological Knowledge from XML Datasources. In *Proc. IEEE DEXA 2002 of Int. Workshop on Electronic Business Hubs (WEBH2002), Aix-en-Provence, France*, pages 852–860, 2002.
5. S. Castano, V. De Antonellis and S. De Capitani di Vimercati. Global Viewing of Heterogeneous Data Sources. *IEEE Transactions on Knowledge and Data Engineering*, 13(2), 2001.
6. S. Castano, V. De Antonellis, S. De Capitani di Vimercati and M. Melchiori. Designing a Three-Layer Ontology in a Web-based Interconnection Scenario. In *Proc. IEEE of Int. Workshop WEBH2001*, Munich, Germany, 2001.
7. E. Colombo, V. De Antonellis, C. Francalanci, M. Mecella, M. Melchiori, B. Pernici and P. Plebani. Cooperative Information Systems in Virtual Districts: the VISPO Approach. *IEEE Database Engineering*, 25(4), 2002.
8. ECCMA. UNiversal Standard Products and Services Classification (UNSPSC). `http://www.eccma.org/`.
9. J. Madhavan, P. A. Bernstein and E. Rahm. Generic schema matching with Cupid. In *Proc. of the Int. Conference on Very Large Data Bases (VLDB2001)*, pages 49–58, Rome, Italy, September 2001.
10. A. Maedche, B. Motik, L. Stojanovic, R. Studer and R. Volz. An Infrastructure for Searching, Reusing and Evolving Distributed Ontologies. In *Proc. of the 12th Int. Conf. on World Wide Web (WWW2003)*, Budapest, Hungary, 2003.
11. A. Maedche, B. Motik, L. Stojanovic, R. Studer and R. Volz. Ontologies for Enterprise Knowledge Management. *IEEE Intelligent Systems*, 2003.
12. M. Missikoff and F. Taglino. SymOntoX: a Web-Ontology Tool for eBusiness Domain. In *Proc. of the 4th Int. Conf. on Web Information Systems Engineering (WISE2003)*, Rome, Italy, December 10-12 2003.
13. ODMG Home Page. `http://www.odmg.org/`.
14. Y. Sure. A Tool-supported Methodology for Ontology-based Knowledge Management. *The Ontology and Modelling of Real Estate Transactions. Edited by H. Stuckenschmidt, E. Stubkjaer and C. Schlieder (eds), Ashgate*, 2003.
15. The VISPO Project Home Page. `http://cube-si.elet.polimi.it/vispo/index.htm`.