GUIDELINES ON GEOGRAPHIC ONTOLOGY INTEGRATION

Margarita Kokla

School of Rural and Surveying Engineering, National Technical University of Athens, 9 H. Polytechniou Str., 157 80 Zografos Campus, Athens, Greece - mkokla@survey.ntua.gr

Technical Commission II, WG II/6

KEY WORDS: GIS, Geography, Interoperability, Comparison, Extraction, Integration

ABSTRACT:

Semantic integration is a complex process influenced by different factors or characteristics. The problem is further complicated by the confusion concerning many issues entailed in the integration process. The present paper (a) analyzes and compares existing integration approaches through the adoption of an ontological framework and (b) describes the principal directions to perform semantic integration of geographic ontologies. In order to make the above analysis more comprehensive, a scenario is demonstrated, which involves the semantic integration of two terminological geographic ontologies. The aim of the paper is to assist the user in following the proper procedure and selecting the appropriate approach in an integration endeavour. The focus is put on semantic integration of geographic ontologies.

1. INTRODUCTION

Interoperability is usually defined as the ability of heterogeneous information systems to communicate, process, and interpret the information exchanged. In literature, interoperability is associated with a number of related but not quite equivalent terms such as coordination, alignment, mapping, merging, partial compatibility, unification, etc., (Sowa, 2000; Klein, 2001). Semantic interoperability aims at the comparison and association of different concepts of interoperating systems on the basis of their semantics, in order to match them and possibly create integrated knowledge bases. Therefore, integration is a core issue in semantic interoperability, especially in a heterogeneous setting, such as the World Wide Web, where different ontologies are used. Semantic integration inevitably leads to inter-ontology mapping, also called *ontology integration*.

Ontology integration is concerned with existing ontologies. It is a crucial issue for the reuse and exchange of knowledge among different domains and also for creating knowledge pools. Furthermore, since knowledge evolves and new knowledge is created, so do the associated ontologies. It is therefore of utmost importance to provide tools for associating present and future ontologies when evolution occurs. This is a common case in the spatiotemporal geographic domain, for instance when National Statistical Agencies change/expand their classification nomenclatures from one census to the next.

There are various approaches of matching, detecting and resolving conflicts, and eventually associating heterogeneous information. This problem is poorly understood (Partridge, 2002; Calvanese et al., 2001; Uschold and Gruninger, 2002) and there is also confusion in the terms found in literature. This is expected for there is not a general agreement on what semantics is or what integration is, what is more, what semantic integration denotes.

The present paper attempts to clarify issues relating to the semantic integration of geographic ontologies and set the principal directions for performing different integration tasks. Existing integration approaches are evaluated on the basis of the ontological framework introduced by Kavouras (2005). In order to make the analysis more comprehensive, an integration scenario is demonstrated.

2. GUIDELINES ON ONTOLOGY INTEGRATION

Integration approaches may and usually do differ considerably due to a number of reasons.

The first reason is the different intended use and objective an approach is designed for, e.g., content explication, data integration, query support, and ontology evolution. The intended use imposes constraints that integration has to comply with. An important issue to deal with is how the resource ontologies are integrated: they may be directly associated, or through a top-level ontology.

The second reason is that approaches may rely on different semantic elements, either because these are simply the only ones available or because these are considered as semantically more important or reliable than others. Most approaches to ontology integration (Calvanese et al., 2001) just rely on term similarity to express mappings between concepts. This proves to be a simple but not always an effectual mechanism; other approaches incorporate additional descriptive information, such as attributes and parts or subsumption relations among concepts in existing schemata or hierarchies. Definitions are also considered as a rich source of semantic information and are used as a comparison basis. Finally, some approaches may resort to an extesional mapping of concepts based on their corresponding instances when of course available. Attributes and instances are not considered semantically very rich and are usually employed by database-oriented approaches.

The third reason is the context of comparison. Ontologies and their concepts have been developed according to different thematic domains. Therefore, it is expected that they present various conflicts. Concepts can be compared in different ways and the result can be diverse. Therefore, one has to decide in advance the parameters/dimensions to be considered as the base/reference of comparison. If too many parameters intrude in the comparison, it is less likely to determine any similarity.

The fourth reason of difference is the role experts or users play during integration. Differences may be reconciled automatically (without user involvement), semi-automatically (resolving some heterogeneities automatically and presenting possible/available choices to select from for the rest), or completely manually (based on intuition, experience, or agreement).

In order to understand and evaluate ontology integration approaches from a huge literature, a framework is usually employed (Tama and Visser, 1998; Pinto et al., 1999; Wache et al., 2001; Calvanese et al., 2001; Klein, 2001; Ding and Foo, 2002; Uschold and Gruninger, 2002; Kavouras, 2005). These frameworks analyze the integration process from a specific perspective and do not attempt to exhaust every possible aspect of the problem. More specifically, Uschold and Gruninger (2002) present an analysis of semantic integration on the web, in which various architectures are evaluated. Klein (2001) focuses on issues applying to ontology combination. The examination of these frameworks shows that there is no general agreement on terminology but more importantly on the issues and dimensions of ontology integration. Furthermore, the integration of geographic ontologies entails additional issues. These are:

- the way concepts are semantically defined and how such semantic information can be derived from existing sources,
- the reconciliation of existing semantic heterogeneities,
- the selection of the most appropriate type of ontology integration approach according to the available semantic information, the objectives set and the constrains of the integration endeavour.

In order to incorporate issues relating to the semantic integration of geographic ontologies, Kavouras introduced an ontological framework, in which the semantic integration process involves three sub-processes: (a) semantic information extraction, (b) concept / ontology comparison, and (c) integration. The first sub-process aims at extracting semantic information from existing sources, such as text, data dictionaries, database schemata, ontologies, etc. This can be empirically implemented by using expert knowledge to establish the basic concepts in a domain. However, natural language processing (NLP) techniques can also be applied in order to automatically or semi-automatically extract important information from available data sources. This information is used as a basis for identifying and resolving semantic heterogeneities between concepts in the second sub-process. In order to compare similar concepts between different ontologies, similarity computation methods are usually employed. However, this process also deals with the resolution of semantic heterogeneities. The third sub-process is the integration of the original concepts. In the literature, several terms are used to denote different integration types: alignment, partial combatibility, unification, true integration (Sowa, 2000). Integration approaches vary according to the following three dimensions (Kavouras, 2005): (1) the possible distortion of the original ontologies, (2) the number of ontologies resulting from the integration process, and (3) the use of a target ontology as a basis for the integration.

In this paper, the framework introduced by Kavouras (2005) is adopted in order to analyze integration approaches, evaluate them and propose their utilization according to different scenarios. The most important issue in selecting or developing an integration approach is the principle or core idea behind it. In literature, there is a lot of confusion about the directions one may possibly take in an integration task. A fundamental objective of all approaches, no matter what methodology or architecture they subsequently employ, is more or less to compare the semantics of the given ontologies and determine the following:

- Whether the given ontologies are to some degree similar, related, or disjoint.
- How to compare concepts in overlapping or related ontologies, in order to identify equivalent, similar (overlapping), related or disjoint concepts.
- How to associate the ontologies on the basis of these findings and the possible architectures.

While the objectives seem clear, the context, perspective, and way the above issues are tackled differ a lot. Herein, some alternatives are presented which are commonly pursued in existing approaches. The objective is to clarify the principal directions and not to exhaust all variants -known or possible.

1. Conforming to a single global ontology. Approaches following this principle attempt to establish a single global ontology that all users employ. This limited principle follows the old standardization paradigm – a way to enforce semantic interoperability, in which the need for mappings is entirely eliminated. Such approaches are not successful only suffice for very narrow applications and small community needs, and for a limited time since they do not handle ontology evolution.

2. Manual ad-hoc mappings. This is a simple and commonly used approach, which lets the user/expert define arbitrary mappings between concepts of the two ontologies. The vast majority of mappings are still established manually. KRAFT (Visser et al., 1999; Preece et al., 2000) is a characteristic example of such an approach. Its major advantage is simplicity and user-controlled result. Since however it is an entirely subjective process many inconsistencies shall arise while it is not certain that semantics are preserved (Wache et al., 2001). Being also laborious and error prone make it highly inefficient to deal with many, large complicated ontologies with many overlapping concepts.

3. Intuitive mappings based on "light" lexical information. More refined (and of course complex) approaches exploit basic (light) lexical information, such as terms (concept names) and their synonyms, to enable a more intuitive mapping between concepts. OBSERVER (Mena et al., 1998) is a typical such case. The advantage of approaches following this direction is that they are less subjective than the first one. As a result, some parts of the process can be semi-automated in form of alternatives suggested to the user. The disadvantage is that mere term similarity (even with the use of vocabularies) is not sufficient to encapsulate the semantics of concepts.

4. Intuitive mappings based on explication characteristics. Some integration approaches, despite being called "semantic", attempt to solve explication problems resulting to a distortion of ontologies in order for example to make them computationally equivalent. These approaches resemble those from the field of database integration, where concepts (entities in this case) are compared/matched with respect to their syntactic similarity on explication characteristics (such as names, data types, and structures) of representation elements (attributes, relations, constraints, and instances). These approaches are very useful to integration at the explication level. Such syntactic information however is either inappropriate (leads to wrong conclusions) or insufficient to reveal semantic similarity, relation, or difference.

5. Intuitive mappings based on structural similarity. Some approaches originating from research on schema integration (Rahm and Bernstein, 2001), are developed on the following assumption: similar ontologies also exhibit some structural (schematic) similarity. Along the same lines, if concepts in different ontologies match, then it is explored if the associated super/sub/side concepts also match. Such approaches no matter how logical, present various problems because ontology hierarchies are based on different contexts/domains and it is not necessary that two equivalent concepts from two ontologies have also equivalent super or sub concepts. The validity of the assumption that schematic similarity is under conditions positively related to semantic similarity has been identified and investigated (Sheth and Kashyap, 1992; Kashyap and Sheth, 1996).

6. Relating (grounding) to a single shared or top-level ontology. Another family of approaches avoids the determination of direct correspondences between concepts from different ontologies by using a single common top-level ontology. Each resource ontology only inherits superconcepts from the top-level ontology. This approach, known as *top-level grounding* (Wache et al., 2001), has some practical advantages, the most important being that the semantics of resource ontologies remain unaltered. The fact however that only indirect correspondences are supported via more general superconcepts may create problems when exact correspondences are needed (ibid.).

7. Direct mappings based on "deep" semantics. Similarly to the semantic correspondences by Wache et al. (2001), the objective of this direction is to avoid (a) indirect mappings via a top-level ontology (direction 5), and (b) subjective direct mappings based on "shallow" semantic information (direction 1). In this family of approaches, in order to support direct mappings among concepts based on "deeper" semantics (e.g., semantic relations), it is essential that such information is derived from the available sources. *Linguistic techniques* such as NLP are usually applied on *unstructured* data, while *constraint-based approaches* are more suitable in the case of *semi-structured* data (Sheth et al., 2005). Common vocabularies may be used to establish semantic correspondences between concepts from different ontologies. Wache et al. (2001) make use of semantic labels to compute correspondences between database fields. Stuckenschmidt et al. (2000) propose a description logic model of terms from different information sources, while relations between different terminologies can be established using subsumption reasoning. Formal Concept Analysis based on semantically "deep" properties, also establishes direct mappings between the concepts involved, in a concept lattice.

8. Integration by view-based query processing. This is a family of approaches in which querying mechanisms play a dominant role (Pottinger and Halevy, 2001; Capezzera, 2003). Integration in this context is expressed by mappings between a global and the local ontologies. Such a service is usually provided by mediators, which usually offer abstract (non-materialized) integrated views over heterogeneous data sources.

9. Compound similarity measures. Concepts might be compared and matched on the basis of the available semantic components assumed in the previous paragraph; i.e., term comparison, relation/property/attribute comparison, or instance comparison. This may conclude as to whether two concepts are equivalent, different, related, etc. Another way of dealing with semantic correspondences and concept matching is by establishing compound similarity measures among the compared concepts from different ontologies (Maedche and Staab, 2002; Maedche and Zacharias, 2002; Rodriguez and Egenhofer, 2003). The result of such approaches is usually a similarity (or dissimilarity) matrix, which may additionally have other uses. From the practical point of view however, compound similarity values do not enlighten us about (a) how heterogeneities must be reconciled and (b) how to create an integrated ontology.

10. Extensional mappings based on common reference. Many of the above families of approaches associate concepts relying on their intensional information. There are however approaches which associate concepts on the basis of their extensional information, when of course available. The simple assumption made here is that concepts having the same instances are likely equivalent. An advantage of these approaches is that there exists an extensional base for comparison and reconciliation. There are however several disadvantages: (i) extesional information may be unavailable, unknown, incomplete or circumstantial; (ii) instances do not necessarily (or fully) describe the semantic domain of a concept; and (iii) the degree to which extensional resemblance is directly and positively related to concept resemblance is not known or justified.

3. AN INTEGRATION SCENARIO

This section demonstrates a semantic integration scenario of two terminological ontologies. A terminological ontology consists of a hierarchy of concepts defined by natural language definitions. It is the most commonly encountered type of existing geographic metadata source. The purpose of the semantic integration scenario is to analyze geographic category definitions in order to extract immanent semantic information, which will be subsequently used to identify similarities and resolve heterogeneities between original categories. The example is small due to space limitation, but includes complex kinds of semantic conflicts that may occur between two concepts (e.g., concepts with overlapping definitions). A basic requirement is to resolve these kinds of conflicts without altering the original concepts. Another requirement is to perform true integration between the original ontologies, i.e., to generate a single integrated ontology, without altering the original ones, nor imposing a target ontology. A further limitation is to perform the process with minimum human interaction in order to ensure maximum objectivity. Relatively to the principal directions introduced in the previous section, the present scenario focuses on the most challenging direction, i.e., the establishment of direct mappings between concepts based on "deep" semantics.

3.1 Semantic Information Extraction

For the first process, a methodology for analyzing definitions and extracting immanent semantic information in the form of semantic elements (e.g., LOCATION, PURPOSE, IS-PART-OF, etc.) is adopted. The methodology was introduced by Jensen and Binot (1987), and further pursued by Vanderwende (1995) and Barriere (1997). This approach is based on:

- parsing (syntactic analysis) of definitions, and
- application of rules that locate certain syntactic and lexical patterns (or defining formulas) in definitions

Parsing determines the structure of a definition, i.e., the form, function, and syntactical relationships of each part of speech. An appropriate tool called parser performs syntactic analysis. The result is usually presented as a parse tree. For the present research, parsing was performed by DIMAP-4 (CL Research, 2001), a program for creating and maintaining dictionaries for use in natural language and language technology applications.

The parsing result is subsequently used by a set of heuristic rules (Dolan et al., 1993). These rules examine the existence of syntactic and lexical patterns, i.e., words and phrases in definitions systematically used to express specific semantic information. For example, the PURPOSE semantic property is determined by specific phrases containing the preposition "for" (e.g., for (the) purpose(s) of, for, used for, intended for) followed by a noun phrase, present participle, or infinitival clause. The rule for extracting this semantic property from definitions is the following (Vanderwende, 1995):

If the verb used (intended, etc.) is post-modified by a prepositional phrase with the preposition "for", then there is a PURPOSE semantic property with the head(s) of that prepositional phrase as the value.

The HAS-PART semantic relation is determined by phrases such as "consist of", "comprise of", "composed of", and "made of". The rule to extract this semantic relation is formulated as following:

If the verb consist (comprise, compose, etc.) is post-modified by a prepositional phrase with the preposition "of", then there is a HAS-PART semantic relation with the head(s) of that prepositional phrase as the value.

Based on the above methodology, geographic category definitions are analyzed and formalized according to their semantic elements. Table 1 shows a list of geographic category terms and their definitions related to hydrography. The definitions are compatible with the structure: "term: genus + differentiae".

Ontology A
Stream: natural flowing body of fresh water
River: natural stream of water, normally of a large volume
Lake: body of water surrounded by land
Canal: artificial waterway created to be paths for boats, or for irrigation
Ontology B
Stream: natural flowing watercourse
Canal: man-made or improved natural waterway used for transportation
Lake: body of water surrounded by land

Table 1. Geographic category terms and definitions related to hydrography

Table 2 shows the set of semantic elements and values of the above geographic categories. Thus, each geographic category definition is replaced by a set of semantic elements and their values.

Category comparison is based on terms, semantic elements and corresponding values. However, in order to perform this process, it is necessary to find synonyms and hypernyms for category terms and values. Reference ontologies, dictionaries or thesauri may provide this information, however human intervention may also be necessary at this phase.

		SEMANTIC ELEMENTS						
		IS-A	**COVER**	**PURPOSE**	**NATURE**	**SIZE**	**SURROUNDNESS**	**FLOW**
ONTOLOGY A	**stream**	body	fresh water		natural			flowing
	river	stream	water		natural	large volume		
	lake	body	water				land	
	canal	waterway		boats or irrigation	artificial			
ONTOLOGY B	**stream**	watercourse			natural			flowing
	canal	waterway		transportation	man-made or improved natural			
	lake	body	water				land	

Table 2. Example of semantic elements and values for geographic categories

3.2 Category Comparison

Category comparison consists in the identification of similarities and heterogeneities between similar categories. This process relies on available elements, which describe categories' semantics, such as terms and definitions. According to the previous section, definitions can be further analyzed into semantic elements and values. Therefore, if we assume that a category definition is analyzed into a set of semantic elements and their corresponding values, then a category C_i is represented by the triple $<T_{Ci}, E_{Ci}, V_{eiCi}>$, where T_{Ci} is the term, E_{Ci} the set of semantic elements and V_{eiCi} the set of corresponding values, i.e.,:

$$E_{C_i} = \{e_{1C_i}, e_{2C_i}, \ldots, e_{nC_i}\} \quad (1)$$

$$V_{e_iC_i} = \{v_{e_1C_i}, v_{e_2C_i}, \ldots, v_{e_nC_i}\} \quad (2)$$

Different combinations of T_{Ci}, E_{Ci} and V_{eiCi} lead to four possible comparison cases (expressing degree of equivalence) between two categories:

- equivalence, when the categories are identical in meaning
- difference (non-equivalence), when the categories have different meanings
- subsumption (partial equivalence), when one category has broader meaning than the other
- overlap (inexact equivalence), when categories have similar, but not precisely identical meanings.

Although many combinations between terms, semantic elements and corresponding values may technically occur, in practice comparison is meaningful mainly for semantically similar categories, i.e., categories with the same or synonymous terms and categories with common semantic elements and values.

Table 3 includes indicative, meaningful combinations between T_{Ci}, E_{Ci}, V_{eiCi}, the comparison result and the action required to resolve the case. This approach can also prove to be useful in cases where terms are neither equal nor synonymous, but appear to present some similarity in certain semantic elements and their corresponding values. Some of these cases are straightforward and can be easily resolved. For example, categories with $T_{C1} \neq T_{C2}$, $E_{C1} = E_{C2}$ result in:

- an overlap $C_1 \mathbf{o} C_2$ when $V_{eC1} \mathbf{o} V_{eC2}$,
- a subsumption $C_1 \supset\subset C_2$ when $V_{eC1} \supset\subset V_{eC2}$

Some other cases however, which involve different terms, overlapping sets of semantic elements and overlapping values are complicated and require more detailed analysis and possibly expert's involvement.

Each comparison case is dealt with differently. The first three (equivalence, difference, and subsumption) are easily resolved. In case of equivalence between two categories, a direct correspondence (equality) between them is specified and they appear as one category in the integrated ontology. In the opposite case, i.e., when the categories are different, no correspondence is specified and the integrated ontology includes both categories. In case of a category being more general than another, a subsumption (IS-A) relation is defined in the integrated ontology. The fourth case is the most difficult to resolve. In this case, it is necessary to split the common from the different parts of overlapping categories.

Reconciliation is implemented by a conceptual analysis procedure known as semantic factoring. Semantic factoring decomposes original categories into a set of non-redundant, non-overlapping conceptual building blocks (Sowa, 2000). These building blocks constitute categories themselves and are called semantic factors. The procedure is based on the comparison results of the previous process.

Semantic factoring (Kokla and Kavouras, 2005; Kokla, 2005) proceeds bottom-up from specific to general categories. At this point, it is necessary to rely on a general reference ontology, which will provide the most specific categories to initiate the comparison. For the purpose of the running example, WordNet is used as reference ontology.

The third process consists in building the integrated ontology from the semantic factors. The proposed methodology (Kokla, 2005) is based on *Formal Concept Analysis* (Wille 1992; Ganter and Wille, 1999), a theory for the formal representation of conceptual knowledge. It models a specific context, namely formal context, as consisting of sets of objects and attributes and the binary relation between them. In our case, the formal context is given by the sets of semantic factors and their corresponding semantic properties and relations.

The result of the third process is the list of final concepts and their subsumption relations, which form a concept lattice (Figure 3). This concept lattice represents the final integrated ontology.

4. CONCLUSIONS

The present paper focuses on the semantic integration of geographic ontologies. It aims to clarify the reasons for the differentiation of integration approaches and the confusion existing in the literature around this complicated issue. A classification of existing integration approaches is presented with an emphasis on the semantics of the compared ontologies. The final objective is to describe principal directions and alternatives that can be pursued for performing an integration task. Furthermore, the paper presents and evaluates integration methods and proposes their utilization according to different integration tasks.

The above are demonstrated with a complex integration scenario, which includes three integration sub-processes. The purpose of this scenario is to analyze geographic category definitions in order to extract immanent semantic information, which will be subsequently used to identify similarities and resolve heterogeneities between original categories. The basic requirements are (a) to perform true integration between the original ontologies, and to implement the process with minimum human interaction in order to ensure maximum objectivity.

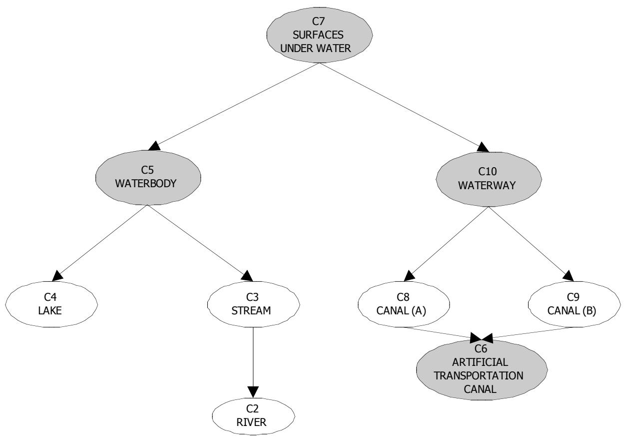

Figure 3. Concept lattice of the final integrated ontology

REFERENCES

Barriere, C., 1997. *From a Children's First Dictionary to a Lexical Knowledge Base of Conceptual Graphs.* PhD Thesis, School of Computing Science, Simon Fraser University.

Calvanese, D., De Giacomo, G., and Lenzerini, M., 2001. A framework for ontology integration. In *Proc. 1st Semantic Web Working Symp. (SWWS 2001)*, 303.

Capezzera, R., 2003. Query processing by semantic reformulation, thesis, Faculty of Engineering at Modena, *Universitŭ di Modena e Reggio Emilia*, Italy, http://www.dbgroup.unimo.it/tesi/capezzera.pdf

CL Research, 2001. DIMAP-4, Dictionary Maintenance Programs, http://www.clres.com

Ding, Y. and Foo, S., 2002. Ontology research and development, Part 2 – A review of ontology mapping and evolving. *Journal of Information Science,* 28 (5), pp. 375-388.

Dolan, W.B., Vanderwende L. and Richardson S.D., 1993. Automatically Deriving Structured Knowledge Base from On-line Dictionaries. In: *Proceedings of the Pacific Association for Computational Linguistics*, Vancouver, British Columbia.

Ganter, B., and Wille, R., 1999. *Formal Concept Analysis, Mathematical Foundations*. Springer-Verlag, Berlin.

Jensen K. and Binot J.L., 1987. Disambiguating prepositional phrase attachments by using on-line dictionary definitions. *Computational Linguistics*, 13 (3-4), pp. 251-60.

Kashyap, V. and Sheth, A., 1996. Schematic and semantic similarities between database objects: A context-based approach, *Int. J. on Very Large Data Bases*, 5 (4), pp. 276-304.

Kavouras, M., 2005. A unified ontological framework for semantic integration. In: *Next Generation Geospatial Information*, Agouris, P., (Ed.), A.A. Balkema Publishers - Taylor & Francis, The Netherlands, pp. 147-156.

Klein, M., 2001. Combining and relating ontologies: An analysis of problems and solutions. In: *Proc. of IJCAI'01-Workshop on Ontologies and Information Sharing,* A. Gomez-Perez, A. et al., Eds., Seattle, USA, Aug. 4-5. http://citeseer.ist.psu.edu/klein01combining.html

Kokla, M. and Kavouras, M., 2005. Semantic information in geo-ontologies: extraction, comparison, and reconciliation". *Journal on Data Semantics*, Stefano Spaccapietra, Esteban Zimányi (Eds.), Lecture Notes in Computer Science 3534 Springer, pp. 125-142.

Kokla, M., 2005. *Semantic Interoperability in Geographic Information Science*. PhD Thesis, School of Rural & Surveying Engineering, National Technical University of Athens, Athens, Greece.

Maedche, A. and Staab, S., 2002. Measuring similarity between ontologies. In: *Proc. of the European Conference on Knowledge Acquisition and Management - EKAW-2002, LNCS/LNAI 2473,* Springer-Verlag, Berlin, 2002, pp. 251-263.

Maedche, A. and Zacharias, V., 2002. Clustering ontology-based metadata in the semantic web. In: *Principles of Data Mining and Knowledge, Proc. 6th European Conference, PKDD 2002, LNCS 2431,* Springer-Verlag, Berlin, 2002, pp. 348-360.

Mena, E., Kashyap, V., Illarramendi, A., Sheth A., 1998. Domain specific ontologies for semantic information brokering on the global information infrastructure. In *Formal Ontology in Information Systems*, Guarino, N., Ed., IOS Press, Amsterdam, pp. 269-283.

Partridge, C., 2002. The role of ontology in semantic integration. In: Second International Workshop on Semantics of Enterprise Integration at OOPSLA, Seattle.

Pinto, H.S., Gomez-Perez, A., and Martins, J.P., 1999. Some issues on ontology integration. In: *Proc. of AI99 Workshop on Ontologies and Problem Solving Methods: Lessons Learned and Future Trends*, Benjamins, V.R. et al., Eds., CEUR Publications, Amsterdam, 18, pp. 7.1-7.12.

Pottinger, R. and Halevy, A., 2001. MiniCon: A scalable algorithm for answering queries using views. *VLDB Journal*, 10 (2-3), pp. 182-198.

Preece, A.D., Hui K., Gray, W.A., Marti, P., Bench-Capon T.J.M., Jones, D.M., Cui, Z., 2000. The KRAFT architecture for knowledge fusion and transformation. *Knowledge Based Systems*, 13(2-3), pp. 113-120.

Rahm, E. and Bernstein, P.A., 2001. On matching schemas automatically. Technical Report MSR-TR-2001-17, Microsoft Research, Microsoft Corporation, One Microsoft Way, Redmond, WA 98052-6399.

Rodríguez, A. and Egenhofer, M., 2003. Determining Semantic Similarity Among Entity Classes from Different Ontologies. In: *IEEE Transactions on Knowledge and Data Engineering*, 15 (2), pp. 442-456.

Sheth, A. and Kashyap, V., 1992. So far (schematically), yet so near (semantically). In: *Proc. IFIPTC2/WG2.6 Conf. on Semantics of Interoperable Database Systems, DS-5*, IFIP Transactions A-25, North Holland, pp. 283-312.

Sheth, A., Ramakrishnan, C., and Thomas, C., 2005. Semantics for the Semantic Web: The implicit, the formal and the powerful. In: *Int. Journal on Semantic Web & Information Systems*, 1(1), pp. 1-18.

Sowa, J.F., 2000. Knowledge Representation: Logical, Philosophical, and Computational Foundations, Brooks/Cole, Pacific Grove.

Stuckenschmidt, H., Wache, H., Vögele, T., Visser, U., 2000. Enabling technologies for interoperability. In: *Workshop on the 14th International Symposium of Computer Science for Environmental Protection*, Visser, U. and Pundt, H., Eds., TZI, Univ. of Bremen, Bonn, Germany, pp. 35-46.

Tamma, V.A.M and Visser, P.R.S., 1998. Integration of heterogeneous sources: Towards a framework for comparing techniques. In: *Proc. of the "6 Convegno dell'Associazione Italiana per l'Intelligenza Artificiale (AI*IA), Workshop su Strumenti di organizzazione ed accesso intelligente per informazioni eterogenee"*, Padua, 89.

Uschold, M. and Gruninger, M., 2002. Creating semantically integrated communities on the World Wide Web. Invited talk in Workshop on the Semantic Web, Co-located with WWW 2002, Honolulu, HI.

Vanderwende, L., 1995. *The Analysis of Noun Sequences using Semantic Information Extracted from On-Line Dictionaries*. Ph.D. thesis, Faculty of the Graduate School of Arts and Sciences, Georgetown University, Washington, D.C.

Visser, P.R.S., Beer, M.D., Bench-Capon, T.J.M., Diaz, B.M., Shave, M.J.R., 1999. Resolving ontological heterogeneity in the KRAFT project. In *10th International Conference and Workshop on Database and Expert Systems Applications DEXA* '99. University of Florence, Italy, August, pp. 668-677.

Wache, H., Vögele, T., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H., Hübner, S., 2001. Ontology-based integration of information - A survey of existing approaches. In: *Proc. of IJCAI-01 Workshop: Ontologies and Information Sharing*, Stuckenschmidt, H., Ed., Seattle WA, pp. 108-117.

Wille, R., 1992. Concept Lattices and Conceptual Knowledge Systems. *Computers and Mathematics with Applications*, 23(6-9), pp. 493-515.