Preface

The Handbook on Ontologies in Information Systems provides a comprehensive overview of the current status and future prospectives of the field of ontologies.

In the early nineties, ontologies have been a research topic being addressed in a rather small research community, the Knowledge Acquisition Community. There, ontologies came into play by the conceptual shift from the 'knowledge transfer'-view on knowledge acquisition to the 'knowledge modeling'- view, most prominently reflected in the CommonKADS methodology for knowledge engingeering and management [10]. Being applied in the context of developing knowledge-based systems, ontologies were classified into domain ontologies, method ontologies, task ontologies, and top-level ontologies that cover the different aspects that are relevant when modeling knowledge-based systems [8, 11].

This rather limited role and impact of ontologies changed drastically in the late nineties by the insight that a conceptual, yet executable model of an application domain provides a significant added value for all kinds of application scenarios like knowledge management or eCommerce, to mention just a few of them. Obviously, the main push for ontologies was given by the vision of the Semantic Web as coined by Tim Berners-Lee [1, 3]. In the Semantic Web ontologies provide the conceptual underpinning for making the semantics of metadata machine interpretable. Being nowadays an important research and application topic in many subject areas, ontologies constitute a field of activities that evolves very fast, both in research and industry.

Being used in such diverse application contexts, it is not easy to agree on a common definition for an *ontology*. However, in the informatics community there has been gained some agreement on using the following definition, based on [6]: "An ontology is a formal explicit specification of a shared conceptualization for a domain of interest." What we can see from this definition on the one hand is the fact that an ontology has to be specified in a language that comes with a formal semantics. Only by using such a formal approach ontologies provide the machine interpretable meaning of concepts and relations that is expected when using an ontology-based approach. On the other hand, ontologies rely on a social process that heads for an agreement among a group of people with respect to the concepts and relations that are part of an

ontology. As a consequence, domain ontologies will always be constrained to a limited domain and a limited group of people. Otherwise, such an agreement will not be feasible. That is clearly one lesson learned from the failure of specifying enterprise-wide data models in the Eighties.

The Handbook on Ontologies in Information Systems covers the fascinating developments in the field of ontologies and thus serves as a reference point for the fast growing community of people being involved in ontology research and applications.

In order to reflect the various aspects that are relevant when dealing with ontologies, the Handbook is organized into several parts.

Part A:

Ontology Representation and Reasoning addresses the main (paradigms of) languages that are currently used for formally specifying ontologies. Description Logics (Baader, Horrocks, and Sattler; Chapter 1) on the one hand and Frame Logic (Angele and Lausen; Chapter 2) on the other hand cover the two main knowledge representation paradigms that have been and are used for describing ontologies. The development of RDF(S) may be considered as one of the cornerstones in establishing basic languages for the Semantic Web (McBride; Chapter 3). The recent development of the Web ontology language OWL reflects the additional requirements that have to be met for deploying powerful ontologies in the (Semantic) Web (Antoniou and van Harmelen; Chapter 4). The last chapter in this first part of the Handbook develops the notion of an ontology algebra (Mitra and Wiederhold; Chapter 5) aiming at a well-defined set of ontology manipulation operators that should be comparable to the well-known relational algebra for relational databases.

Part B:

Ontology Engineering is devoted to various methodological aspects of developing ontologies. The first chapter describes a methodology for introducing ontology-based knowledge management solutions into enterprises (Sure, Staab, and Studer; Chapter 6). In spite of all the tool environments that are nowadays available, there is still a lack of large domain ontologies. A large scale case study in developing a medical ontology is described in the next chapter (Hahn and Schulz; Chapter 7). When dealing with such large resources there comes the urgent need to clean up the top level of the ontology in order to provide some first principles on which to build and extend a given ontology — such as defined by the OntoClean methodology (Guarino and Welty; Chapter 8). The overhead involved in building up an ontology, however, remains a major obstacle in applying ontologies. Therefore, an important research topic is ontology learning that aims at providing (semi-) automatic support for building up ontologies. This aspect is addressed in Chapter 9 by Mädche and Staab. Knowlede patterns as outlined in Chapter 10 cover inherent structures that may be found when investigating axioms in detail. Such patterns enable reuse as known from the field of Software Engineering (Clark, Thompson, and Porter; Chapter 10). The tight interaction between ontologies on the one hand and lexicons as handled in linguistics is discussed in Chapter 11 by Hirst. In the majority of applications one has to deal

with multiple ontologies that reflect the different views people have on a subject field. Therefore, the need arises for being able to reconcile these views and thus to relate the corresponding ontologies to each other (Hameed, Preece, and Sleeman; Chapter 12). Goméz-Pérez introduces a field of research that emerged very recently: ontology evaluation (Chapter 13). Obviously, ontology evaluation gains more and more importance when considering the fast growing number of deployed ontologies. The last chapter in this part surveys state-of-the-art ontology engineering environments (Mizoguchi; Chapter 14) — touching on some of the aspects of the following part.

Part C:

Ontology Infrastructure covers various important aspects of managing and using ontologies. This part starts with a description of a flexible and extensible server environment for managing ontologies (Oberle, Volz, Staab, and Motik; Chapter 15). Ontologies and problem-solving methods are tightly related to each other with respect to providing a conceptualization of the static and dynamic aspects of knowledge-based systems. These relationships and the roles ontologies have for describing problem-solving methods are discussed in Chapter 16 by Crubézy and Musen. Multi-Agent systems heavily rely on the communication and cooperation between the involved agents. Therefore, ontologies play an important role in providing a semantic basis for these interactions (Sycara and Paolucci; Chapter 17). The next chapter provides the description of a tool environment that supports users in merging ontologies and define mappings between different ontologies (Noy; Chapter 18) . The topic of mapping ontologies to each other is further developed in the next chapter in which a learning approach is introduced for learning mappings between ontologies (Doan, Madhavan, Domingos, and Halevy; Chapter 19). An enduser point of view on handling ontologies is described in the last two chapters of this part. First, an exploring and browsing environment is described that supports the graphical browsing of ontologies, illustrated for RDF-based ontologies (Eklund, Cole, and Roberts; Chapter 20). The last chapter provides an insight into the visualization techniques that can be exploited for supporting the life-cycle of ontologies (Fluit, Sabou, and van Harmelen; Chapter 21).

Part D:

Finally, *Ontology Applications* provides a broad overview of the various application areas that nowadays exploit ontologies. The first chapters discuss knowledge management applications that address various information technology aspects of knowledge management. The general role that ontologies might play for knowledge management applications is discussed in the beginning (Abecker and van Elst; Chapter 22). An approach for ontology-based content management taylored to the needs of virtual organizations is outlined in Chapter 23 (Mika, Iosif, Sure, and Akkermans). Chapter 24 discusses flexible recommender systems that rely on user profiles which are based on an ontology (Middleton, de Roure, and Shadbolt). Ontologies also provide means for building up semantic portals that are able to integrate knowledge

from heterogeneous sources. In Chapter 25 the OntoWeb portal is described as a concrete application based on such a semantic portal framework (Oberle and Spyns). The presentation of portal content is closely related to the question of how to create (non-linear) hypertext. The ability to model a domain of discourse with an ontology holds out the promise of computationally reasoned and reasonable linking services. Chapter 26 describes some attempts to augment hypermedia systems with ontologies in order to provide, or improve, hypertext. Chapter 27 (Domingue, Dzbor, and Motta) introduces a system environment in which Web resources are embedded into a semantic context that is based on ontologies. Thus, user-specific viewpoints may be offered in a flexible way. eLearning and knowledge management are fields that get more and more integrated. In Chapter 28 metadata- and ontology-based approaches for eLearning are discussed (Brase and Nejdl).

Besides knowledge management, issues of interoperability and integration of enterprise applications are a second major application field for ontologies. In Chapter 29 (Grüninger) a process language is defined that is based on ontologies and provides means for facilitating the exchange of information between manufacturing systems. Chapter 30 addresses issues of information integration in eCommerce scenarios (Ying, Fensel, Klein, Omelayenko, and Schulten). Semantic interoperability issues are further discussed in Chapter 31. They are illustrated by application scenarios from the domain of tourism (Missikoff and Taglino). Finally, ontologies may be used for capturing the semantics of vast amount of data that are handled e.g. in molecular biology (Stevens, Wroe, Lord, and Goble; Chapter 32).

As can be seen from the large collection of different application scenarios that are discussed in *Part D: Ontology Applications* ontologies become a major conceptual backbone for a broad spectrum of applications (cf. also [2, 4]). Integration of web resources, intranet documents, and databases as well as cooperation of web services and enterprise application, all these scenarios require a semantic characterization of the meaning of their contents and/or their functionalities. Therefore, ontologies gain a strongly increasing importance in these kinds of applications. Technologies for managing ontologies in the context of these applications have reached a maturity level that enables their usage in real life applications.

Nevertheless, a considerable amount of research problems have to be solved in the future in order to meet the increasing challenges. Some recent developments address the evolution of ontologies in distributed environments [9], creation of metadata conforming to ontologies [7], the integration of description logics and rules [5], and performance evaluation [12, 13]. Furthermore, there remain open research problems, like ontology mediation on-the-fly, efficient ontology querying, or a rule language that fits to the proposed ontology languages and meets the requirements of the application developers.

The Handbook on Ontologies in Information Systems provides a detailed overview of the current achievements and challenges in the ontology area. With its coverage of research and applications it provides valuable insights, both for researchers and practicioners. The lot of developments that are currently under way may result in a second volume that might appear in some years from now.

Karlsruhe, July 2003

Steffen Staab
Rudi Studer

Acknowledgements

We gratefully acknowledge efforts by all the authors who also acted as peer reviewers and helped to further improve the quality of the papers. We thank Juliane Willsenach for heroic efforts typesetting Word documents.

References

1. T. Berners-Lee, J. Hendler, and O. Lassila (2001). The Semantic Web. *Scientific American*, May 2001.
2. J. Davies, D. Fensel, and F. van Harmelen, eds. *Towards the Semantic Web: Ontology-driven Knowledge Management.* John Wiley & Sons, 2003.
3. D. Fensel and M.A. Musen, eds. The Semantic Web: A Brain for Humankind. *IEEE Intelligent Systems, Special Issue Cooking up the Semantic Web*, 16, 2 (March/April 2001), pp. 24-25.
4. D. Fensel, W. Wahlster, H. Lieberman, and J. Hendler, eds. *Spinning the Semantic Web*. MIT Press, 2003.
5. B. Grosof, I. Horrocks, R. Volz, S. Decker. Description Logic Programs: Combining Logic Programs with Description Logic. *Proceedings of WWW-2003*, ACM Press, 2003.
6. T. Gruber. A Translation Approach to Portable Ontology Specifications, *Knowledge Acquisition*, Vol.5, 1993, pp.199-220.
7. S. Handschuh and S. Staab (eds.). *Annotation for the Semantic Web*. IOS Press, 2003.
8. G. van Heijst, G. Schreiber, and B.J. Wielinga. Using Explicit Ontologies in KBS Development. *Intl. Journal on Human-Computer Studies*, 46 (2/3) (1997), pp. 183-292.
9. A. Maedche, B. Motik, L. Stojanovic, R. Studer, and R. Volz. An Infrastructure for Searching, Reusing and Evolving Distributed Ontologies. *Proceedings of WWW-2003*, ACM Press, 2003.
10. G. Schreiber, H. Akkermans, A. Anjewierden, R. de Hoog, N. Shadbolt, W. van de Velde, and B. Wielinga. *Knowledge Engineering and Management — The CommonKADS Methodology*. The MIT Press, Cambridge, Massachusetts; London, England, 1999.
11. R. Studer, V.R. Benjamins, and D. Fensel. Knowledge Engineering: Principles and Methods. *Data & Knowledge Engineering*, 25 (1998), pp. 161-197.
12. Y. Sure, J. Angele, and S. Staab. OntoEdit: Guiding Ontology Development by Methodology and Inferencing. *1st International Conference on Ontologies, Databases and Applications of Semantics for Large Scale Information Systems - ODBASE 2002*. October 29 - November 1, Irvine, California. LNCS, Springer, 2002.
13. R. Volz, B. Motik, I. Horrocks and B. Grosof. Description Logics Programs: An Evaluation and Extended Translation. Technical Report, 2003.

Contents

Author Index