Ontology Alignment
Bridging the Semantic Gap

SEMANTIC WEB AND BEYOND
Computing for Human Experience

Series Editors:

Ramesh Jain
University of California, Irvine
http://ngs.ics.uci.edu/

Amit Sheth
University of Georgia
http://lsdis.cs.uga.edu/~amit

As computing becomes ubiquitous and pervasive, computing is increasingly becoming an extension of human, modifying or enhancing human experience. Today's car reacts to human perception of danger with a series of computers participating in how to handle the vehicle for human command and environmental conditions. Proliferating sensors help with observations, decision making as well as sensory modifications. The emergent semantic web will lead to machine understanding of data and help exploit heterogeneous, multi-source digital media. Emerging applications in situation monitoring and entertainment applications are resulting in development of experiential environments.

SEMANTIC WEB AND BEYOND
Computing for Human Experience
addresses the following goals:

- brings together forward looking research and technology that will shape our world more intimately than ever before as computing becomes an extension of human experience;
- covers all aspects of computing that is very closely tied to human perception, understanding and experience;
- brings together computing that deal with semantics, perception and experience;
- serves as the platform for exchange of both practical technologies and far reaching research.

Additional information about this series can be obtained from
http://www.springer.com

AdditionalTitles in the Series:

Semantic Web Services: Processes and Applications edited by Jorge Cardoso, Amit P. Sheth, ISBN 0-387-30239-5

Canadian Semantic Web edited by Mamadou T. Koné., Daniel Lemire; ISBN 0-387-29815-0

Semantic Management of Middleware by Daniel Oberle; ISBN-10: 0-387-27630-0

Ontology Alignment
Bridging the Semantic Gap

by

Marc Ehrig
University of Karlsruhe
Germany

Springer

Marc Ehrig
University of Karlsruhe
76139 Karlsruhe
Germany

Library of Congress Control Number: 2006928852

Ontology Alignment: Bridging the Semantic Gap
by Marc Ehrig

ISBN-10: 0-387- 32805-X
ISBN-13: 978-0-387-32805-8
e-ISBN-10: 0-387-36501-X
e-ISBN-13: 978-0-387-36501-5

Printed on acid-free paper.

Dissertation, genehmigt von der Fakultät für Wirtschaftswissenschaften der Universität Fridericiana zu Karlsruhe, 2005. Referent: Prof. Dr. Rudi Studer,
Korreferenten: Prof. Dr. Kuno Egle, Dr. (habil) Jérôme Euzenat

9 8 7 6 5 4 3 2 1

springer.com

Contents

Part III Implementation and Application

List of Figures

List of Tables

Preface

In today's knowledge society, a large number of information systems use many different individual schemas to represent data. Ontologies are one promising approach for representing knowledge in a formal way. Many such ontologies have been developed in recent years. Semantically linking these ontologies is a necessary precondition to establish interoperability between agents or services, or simply humans working with them. Consequently, ontology alignment becomes a central issue, when building a world-wide Semantic Web. Integrating data per se is a billion dollar industry. As one can easily imagine, this cannot be done manually beyond a certain complexity, size, or number of, here, ontologies. Automatic or at least semi-automatic techniques have to be developed to reduce the burden of manual creation and maintenance of alignments.

The purpose of this book is to foster understanding in new semantic technologies, data integration, and the interaction between the two fields. In this application-driven work, the reader is presented a methodology and advice for a concrete tool for aligning ontologies. This is going to be done on theoretical and practical level for both research-focused audiences and developers. Goal is not to align ontologies by only integrating the syntax, but actually bringing together entities which have the same meaning, thus bridging the semantic gap.

The book begins with a short motivation, followed by a thorough investigation of the foundations including up-to-date related work on ontology alignment and application scenarios with their respective requirements. The six-step ontology alignment process consists of determining relevant features of individual entities, selection of promising alignment candidates, similarity assessment and aggregation, interpretation of the similarities for alignment, and, if applicable, several iterations thereof. As result one receives those pairs of entities which correspond to each other. Complex similarity considerations are claimed the key for identifying these alignments. The basic approach is extended through novel methods focusing on efficiency, machine learning optimization, active user inclusion, scenario-adaptive alignment, and an in-

tegrated strategy. The implementation and evaluation shows that both the (semi-) automatic ontology alignment process itself and its output improve significantly. Examples of running applications using the new strategies including one commercial product prove the practical value. Further pointers for next steps in ontology alignment are given including a generalization for other structures and schemas before a summary closes this work.

Using semantic features can help to reach levels of alignment which have never been possible before. The exploitation and application of these advantages is just starting. Methods in this work are basic elements of this development and are expected to be continuously enhanced. In this sense, they will lastingly affect future research and implementation. Therefore, the topic of ontology alignment coupled with the application-focused methodology is appropriate to excite interest of a broad readership.

Karlsruhe, April 2006
Marc Ehrig

Acknowledgements

This book would not have been possible without the invaluable support of many persons. Each of them guided me further towards the goal of creating a comprehensive work on ontology alignment, for which I thank every single one very much.

In first place, it was my advisor Prof. Dr. Rudi Studer, who gave me the challenge and the chance to do this research. He granted me the freedom, the trust, and the help I needed.

I also thank the reviewers comprising Prof. Dr. Kuno Egle, Dr. habil. Jérôme Euzenat, Prof. Dr. Hartmut Schmeck, and Prof. Dr. Ute Werner, who through their objective reflection provided valuable feedback. Further, I am grateful to Susan Lagerstrom-Fife and Sharon Palleschi from Springer for the successful and smooth collaboration on publishing this book.

My friends from Karlsruhe, Peter Haase, Jens Hartmann, Prof. Dr. Steffen Staab, Dr. York Sure, and Christoph Tempich were the closest co-workers, but especially main contributors and critics when I was developing new ideas. This list needs to be extended to all the great friends and colleagues at LS3WIM in Karlsruhe. The nice atmosphere was indeed an essential basis resulting in many professional and social events, which make my time in Karlsruhe unforgettable. Numerous people around the world continuously improved this work through helpful and detailed discussions.

I would like to acknowledge that this research would not have been possible without the funding of the European Union through the projects SWAP, SEKT, and Knowledge Web.

Finally, I thank my family: Aline, Helga, and Volker, who since the very beginning have always been reliable partners at my side motivating me to continue on my path.