

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

University of Dortmund, Germany

Madhu Sudan

Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos

New York University, NY, USA

Doug Tygar

University of California, Berkeley, CA, USA

Moshe Y. Vardi

Rice University, Houston, TX, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

Stefano Spaccapietra (Ed.)

Journal on Data Semantics IV

Volume Editor

Stefano Spaccapietra
EPFL-IC-IIF-LBD, INJ 236 (Bâtiment INJ), Station 14
1015 Lausanne, Switzerland
E-mail: stefano.spaccapietra@epfl.ch

Library of Congress Control Number: 2005937594

CR Subject Classification (1998): H.2, H.3, I.2, H.4, C.2

ISSN 0302-9743
ISBN-10 3-540-31001-0 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-31001-3 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2005
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11603412 06/3142 5 4 3 2 1 0

The LNCS Journal on Data Semantics

Computerized information handling has changed its focus from centralized data management systems to decentralized data exchange facilities. Modern distribution channels, such as high-speed Internet networks and wireless communication infrastructure, provide reliable technical support for data distribution and data access, materializing the new, popular idea that data may be available to anybody, anywhere, anytime. However, providing huge amounts of data on request often turns into a counterproductive service, making the data useless because of poor relevance or inappropriate level of detail. Semantic knowledge is the essential missing piece that allows the delivery of information that matches user requirements. Semantic agreement, in particular, is essential to meaningful data exchange.

Semantic issues have long been open issues in data and knowledge management. However, the boom in semantically poor technologies, such as the Web and XML, has boosted renewed interest in semantics. Conferences on the Semantic Web, for instance, attract crowds of participants, while ontologies on their own have become a hot and popular topic in the database and artificial intelligence communities.

Springer's *LNCS Journal on Data Semantics* aims at providing a highly visible dissemination channel for the most remarkable work that in one way or another addresses research and development on issues related to the semantics of data. The target domain ranges from theories supporting the formal definition of semantic content to innovative domain-specific application of semantic knowledge. This publication channel should be of the highest interest to researchers and advanced practitioners working on the Semantic Web, interoperability, mobile information services, data warehousing, knowledge representation and reasoning, conceptual database modeling, ontologies, and artificial intelligence.

Topics of relevance to this journal include:

- Semantic interoperability, semantic mediators
- Ontologies
- Ontology, schema and data integration, reconciliation and alignment
- Multiple representations, alternative representations
- Knowledge representation and reasoning
- Conceptualization and representation
- Multimodel and multiparadigm approaches
- Mappings, transformations, reverse engineering
- Metadata
- Conceptual data modeling
- Integrity description and handling
- Evolution and change
- Web semantics and semi-structured data
- Semantic caching
- Data warehousing and semantic data mining
- Spatial, temporal, multimedia and multimodal semantics

- Semantics in data visualization
- Semantic services for mobile users
- Supporting tools
- Applications of semantic-driven approaches

These topics are to be understood as specifically related to semantic issues. Contributions submitted to the journal and dealing with semantics of data will be considered even if they are not within the topics in the list.

While the physical appearance of the journal issues is like the books from the well-known Springer LNCS series, the mode of operation is that of a journal. Contributions can be freely submitted by authors and are reviewed by the Editorial Board. Contributions may also be invited, and nevertheless carefully reviewed, as in the case for issues that contain extended versions of the best papers from major conferences addressing data semantics issues. Special issues, focusing on a specific topic, are coordinated by guest editors once the proposal for a special issue is accepted by the Editorial Board. Finally, it is also possible that a journal issue be devoted to a single text.

The journal published its first volume in 2003 (LNCS 2800), its second volume at the beginning of 2005 (LNCS 3360), and its third volume in Summer 2005 (LNCS 3534). The first two volumes are special issues composed of selected extended versions of the best conference papers. The third volume is a special issue on “Semantic-Based Geographical Information Systems”, coordinated by guest editor Prof. Esteban Zimányi. This fourth volume is the first “normal” volume, consisting of spontaneous submissions on any of the topics of interest to the journal. Currently planned volumes include a special issue on emergent semantics.

The Editorial Board comprises one Editor-in-Chief (with overall responsibility) and several members. The Editor-in-Chief has a four-year mandate to run the journal. Members of the board have a three-year mandate. Mandates are renewable. More members may be added to the board as appropriate.

We are happy to welcome both readers and authors, and hope we will share this privileged contact for a long time.

Stefano Spaccapietra
Editor-in-Chief
<http://lbdwww.epfl.ch/e/Springer/>

JoDS Volume IV - Preface

This fourth JoDS volume is the outcome of the selection of papers spontaneously submitted to the journal, in particular in response to a Call for Papers issued on July 17, 2004. The call invited submissions on any topic that falls within the scope of the journal.

Altogether, 38 submissions were reviewed. After the first round of reviews, 24 submissions were asked to perform a major revision and resubmit. Most of these were actually resubmitted and went through a second round of reviews, with the same reviewers as allocated for the first round. Eventually, 10 papers were accepted for publication, after some last modifications suggested by the reviewers.

Accepted papers cover a wide range of topics, from traditional data semantics (information modeling, data model transformation, knowledge representation, data and schema integration) to the newest trends (multimedia, Semantic Web annotation, information extraction, and knowledge discovery).

A new Call for Papers is open at the moment for a volume to appear in 2006. We are looking forward to your contributions.

Stefano Spaccapietra
Editor-in-Chief

In Memoriam

Prof. Maurizio Panti

On July 3, 2005, at 9:45AM, Prof. Maurizio Panti, Head of the Department of Computer, Management and Automation Engineering at the Polytechnic University of Marche, passed away.

Approximately seven months earlier, he had been diagnosed with a late-stage aggressive cancer. In this brief period of illness, he never stopped his activities, showing devotion to his work and students and a strong will power that will be an example for all of us. He spent his final hours surrounded by family and friends.

Prof. Maurizio Panti was full professor of Information Systems and Data Bases. He promoted the growth of informatics both in academia, contributing to the foundation of the Informatics Institute, which then developed into the present department, and in the regional economic arena, also serving as a member in the Scientific and Technical Board for the Regional Information System.

His loss will be greatly felt by all those who knew and worked with him.

Prof. Hongjun Lu

Prof. Hongjun Lu was a world-renowned researcher who served the database community with dedication and distinction in various capacities over the years. He was a trustee of the VLDB Endowment, a member of the ACM SIGMOD Advisory Board (1998–2002), an associate editor of IEEE Transactions on Knowledge and Data Engineering (TKDE), Chair of the Steering Committee of the International Conference on Web-Age Information Management (WAIM), and Co-chair (1998-2001) and Chair (2001-2003) of the Steering Committee of Pacific-Asia Conference of Knowledge Discovery and Data Mining (PAKDD). In December 2004, the China Computer Federation Database Society honored him with a Contribution Award and, just recently, he was honored with this year's inaugural PAKDD Distinguished Contribution Award.

His passing is a great loss not only to the China Database Society but also to the International Database Society. He has made sustained and outstanding contributions to the international database research community as well as to database research.

Organization

External Reviewers

All Editorial Board members contributed reviews for the selection of the submissions. In addition, a number of colleagues helped us in this reviewing task. We would like to express here our gratitude for their cooperation and our warmest thanks for the job they did.

Bettina Berendt

Sonia Bergamaschi

Claudio Bettini

Mokrane Bouzeghoub

Anita Coleman

Jos De Bruin

Flavio De Paoli

Giovanni Denaro

Ying Ding

Gillian Dobbie

Anastasios D. Doulamis

Oleksandr Drutskyy

Birte Glimm

Sven Hartmann

Georges Hébrail

Martin Hepp

Jane Hunter

Mustafa Jarrar

Roland Kaschek

Evgeny Zolin

Yehuda Koren

Mong Li Lee

Lei Li

Xiaonan Lu

R. Manmatha

Andrea Maurino

Rym Mili

Moira Norrie

Axel Polleres

Helen Purchase

Brian Roark

Ingo Schmitt

Richard Smiraglia

Peter Spyns

Paolo Terenziani

Christelle Vangenot

Michalis Vazirgiannis

James Z. Wang

X. Sean Wang

JoDS Editorial Board

Carlo Batini, Università di Milano Bicocca, Italy

Tiziana Catarci, Università di Roma La Sapienza, Italy

Lois Delcambre, Portland State University, USA

David W. Embley, Brigham Young University, USA

Jerome Euzenat, INRIA Rhône-Alpes, France

Dieter Fensel, University of Innsbruck, Austria

Nicola Guarino, National Research Council, Italy

Jean-Luc Hainaut, FUNDP Namur, Belgium

Ian Horrocks, University of Manchester, UK

Larry Kerschberg, George Mason University, USA

Maurizio Lenzerini, Università di Roma La Sapienza, Italy

Tok Wang Ling, National University of Singapore, Singapore

Salvatore T. March, Vanderbilt University, USA

Robert Meersman, Vrije Universiteit Brussel (VUB), Belgium

John Mylopoulos, University of Toronto, Canada

Shamkant B. Navathe, Georgia Institute of Technology, USA

Antoni Olivé, Universitat Politècnica de Catalunya, Spain

José Palazzo M. de Oliveira, Universidade Federal do Rio Grande do Sul, Brazil

Christine Parent, Université de Lausanne, Switzerland

John Roddick, Flinders University, Australia

Klaus-Dieter Schewe, Massey University, New Zealand

Bernhard Thalheim, Brandenburg University of Technology, Germany

Yair Wand, University of British Columbia, Canada

Esteban Zimányi, Université Libre de Bruxelles (ULB), Belgium

Table of Contents

Generic Relationships in Information Modeling <i>Mohamed Dahchour, Alain Pirotte, Esteban Zimányi</i>	1
EMMA – A Formal Basis for Querying Enhanced Multimedia Meta Objects <i>Sonja Zillner, Werner Winiwarter</i>	35
Comparing and Transforming Between Data Models Via an Intermediate Hypergraph Data Model <i>Michael Boyd, Peter McBrien</i>	69
iASA: Learning to Annotate the Semantic Web <i>Jie Tang, Juanzi Li, Hongjun Lu, Bangyong Liang, Xiaotong Huang, Kehong Wang</i>	110
A Survey of Schema-Based Matching Approaches <i>Pavel Shvaiko, Jérôme Euzenat</i>	146
An Overview and Classification of Adaptive Approaches to Information Extraction <i>Christian Siefkes, Peter Siniakov</i>	172
View Integration and Cooperation in Databases, Data Warehouses and Web Information Systems <i>Hui Ma, Klaus-Dieter Schewe, Bernhard Thalheim, Jane Zhao</i>	213
Semantic Integration of Tree-Structured Data Using Dimension Graphs <i>Theodore Dalamagas, Dimitri Theodoratos, Antonis Koufopoulos, I-Ting Liu</i>	250
KDD Support Services Based on Data Semantics <i>Claudia Diamantini, Domenico Potena, Maurizio Panti</i>	280
Integrating the Two Main Inference Modes of NKRL, Transformations and Hypotheses <i>Gian Piero Zarri</i>	304
Author Index	341