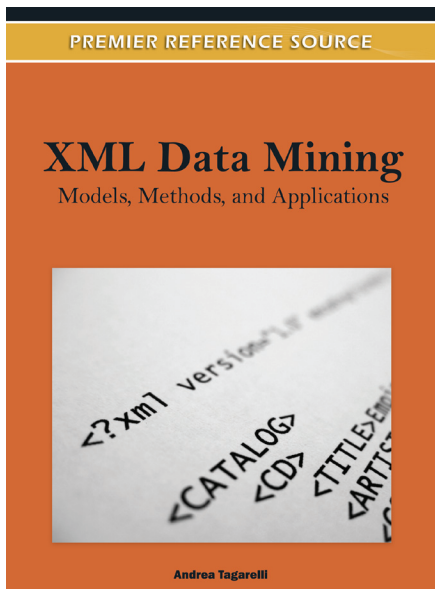


# XML Data Mining: Models, Methods, and Applications

**Edited By:** Andrea Tagarelli (University of Calabria, Italy)



The widespread use of XML in business and scientific databases has prompted the development of methodologies, techniques, and systems for effectively managing and analyzing XML data. This has increasingly attracted the attention of different research communities, including database, information retrieval, pattern recognition, and machine learning, from which several proposals have been offered to address problems in XML data management and knowledge discovery.

**XML Data Mining: Models, Methods, and Applications** aims to collect knowledge from experts of database, information retrieval, machine learning, and knowledge management communities in developing models, methods, and systems for XML data mining. This book addresses key issues and challenges in XML data mining, offering insights into the various existing solutions and best practices for modeling, processing, analyzing XML data, and for evaluating performance of XML data mining algorithms and systems.

**ISBN:** 978-1-61350-356-0; © 2012; 538 pp.

**Hard Cover:** US \$195

**Online Perpetual Access:** US \$295

**Print + Online Perpetual Access:** US \$390

## Topics Covered:

- XML Models for Data Mining
- XML Similarity Search and Detection
- Approximate Matching of XML Documents and Schemas
- Clustering of XML Data
- Classification of XML Data
- Frequent Pattern Discovery of XML Data
- Association Rule Mining of XML Data
- Mining of Uncertain XML Data
- Mining of Evolving XML Data Streams
- XML Mining for Semantic Web
- Semantics-aware Mining of XML Data
- Domain-specific XML Mining Applications: Credit Risk Assessment, Social Network User Modeling of Geographical Maps, P2P systems

**Pre-pub Price:\***

**Hard Cover: US \$185; Online Perpetual Access: US \$280**

\* Pre-pub price is good through one month after publication date.

## Market:

This premier publication is essential for all academic and research library reference collections. It is a crucial tool for academicians, researchers, and practitioners and is ideal for classroom use.

"I am confident that this timely volume will be the go-to reference for finding the latest methods and developments in XML data mining. It will be an invaluable resource for students, researchers, and practitioners in the field."

Mohammed J. Zaki, Rensselaer Polytechnic Institute, USA

**Excellent addition to your library! Recommend to your acquisitions librarian.**

[www.igi-global.com](http://www.igi-global.com)

### Section 1: Models and Measures

#### Chapter 1

*A Study of XML Models for Data Mining: Representations, Methods, and Issues*  
Sangeetha Kutty, Queensland University of Technology, Australia  
Richi Nayak, Queensland University of Technology, Australia  
Tien Tran, Queensland University of Technology, Australia

#### Chapter 2

*Modeling, Querying, and Mining Uncertain XML Data*  
Evgeny Kharlamov, Free University of Bozen-Bolzano, Italy & INRIA Saclay, France  
Pierre Senellart, Télécom ParisTech, France

#### Chapter 3

*XML Similarity Detection and Measures*  
Sanjay Kumar Madria, Missouri University of Science and Technology, USA  
Waroporn Viyanon, Missouri University of Science and Technology, USA

#### Chapter 4

*Efficient Identification of Similar XML Fragments Based on Tree Edit Distance*  
Hongzhi Wang, Harbin Institute of Technology, China  
Jianzhong Li, Harbin Institute of Technology, China  
Fei Li, Harbin Institute of Technology, China

### Section 2: Clustering and Classification

#### Chapter 5

*Approximate Matching Between XML Documents and Schemas with Applications in XML Classification and Clustering*  
Guangming Xing, Western Kentucky University, USA

#### Chapter 6

*The Role of Schema and Document Matchings in XML Source Clustering*  
Pasquale De Meo, University of Messina, Italy  
Giacomo Fiumara, University of Messina, Italy  
Antonino Nocera, University Mediterranea of Reggio Calabria, Italy  
Domenico Ursino, University Mediterranea of Reggio Calabria, Italy

#### Chapter 7

*XML Document Clustering: An Algorithmic Perspective*  
Panagiotis Antonellis, University of Patras, Greece

#### Chapter 8

*Fuzzy Approaches to Clustering XML Structures*  
Michał Koziński, Silesian University of Technology, Poland

#### Chapter 9

*XML Tree Classification on Evolving Data Streams*  
Albert Bifet, University of Waikato, New Zealand  
Ricard Gavaldà, UPC Barcelona Tech, Spain

#### Chapter 10

*Data Driven Encoding of Structures and Link Predictions in Large XML Document Collections*  
Markus Hagenbuchner, University of Wollongong, Australia  
Chung Tsoi, Macau University of Science and Technology, China  
Shu Jia Zhang, University of Wollongong, Australia  
Milly Kc, University of Wollongong, Australia

### Section 3: Association Mining

#### Chapter 11

*Frequent Pattern Discovery and Association Rule Mining of XML Data*  
Qin Ding, East Carolina University, USA  
Gnanasekaran Sundarraj, Pennsylvania State University, USA

#### Chapter 12

*A Framework for Mining and Querying Summarized XML Data through Tree-Based Association Rules*  
Mirjana Mazuran, Politecnico di Milano, Italy  
Elisa Quintarelli, Politecnico di Milano, Italy  
Angelo Rauseo, Politecnico di Milano, Italy  
Letizia Tanca, Politecnico di Milano, Italy

#### Chapter 13

*Discovering Higher Level Correlations from XML Data*  
Luca Cagliero, Politecnico di Torino, Italy  
Tania Cerquitelli, Politecnico di Torino, Italy  
Paolo Garza, Politecnico di Milano, Italy

### Section 4: Semantics-Aware Mining

#### Chapter 14

*XML Mining for Semantic Web*  
Rafael Berlanga, Universitat Jaume I, Spain  
Victoria Nebot, Universitat Jaume I, Spain

#### Chapter 15

*A Component-Based Framework for the Integration and Exploration of XML Sources*  
Pasquale De Meo, University of Messina, Italy  
Antonino Nocera, University Mediterranea of Reggio Calabria, Italy  
Domenico Ursino, University Mediterranea of Reggio Calabria, Italy

#### Chapter 16

*Matching XML Documents at Structural and Conceptual Level using Subtree Patterns*  
Qi Hua Pan, Curtin University, Australia  
Fedja Hadzic, Curtin University, Australia  
Tharam S. Dillon, Curtin University, Australia

### Section 5: Applications

#### Chapter 17

*Geographical Map Annotation with Significant Tags Available from Social Networks*  
Elena Roglia, University of Turin, Italy  
Rosa Meo, University of Turin, Italy  
Enrico Ponassi, University of Turin, Italy

#### Chapter 18

*Organizing XML Documents on a Peer-to-Peer Network by Collaborative Clustering*  
Francesco Gullo, University of Calabria, Italy  
Giovanni Ponti, ENEA, Italy  
Sergio Greco, University of Calabria, Italy

#### Chapter 19

*Incorporating Qualitative Information for Credit Risk Assessment through Frequent Subtree Mining for XML*  
Novita Iksari, Curtin University, Australia & University of Indonesia, Indonesia  
Fedja Hadzic, Curtin University, Australia  
Tharam S. Dillon, Curtin University, Australia

## Andrea Tagarelli

Andrea Tagarelli is an Assistant Professor of Computer Science with the Department of Electronics, Computer and Systems Sciences, University of Calabria, Italy. He graduated in Computer Engineering, in 2001 and obtained his Ph.D. in Computer and Systems Engineering, in 2006. He was visiting researcher at the Department of Computer Science & Engineering, University of Minnesota at Minneapolis, USA. His research interests include topics in knowledge discovery and text/data mining, information extraction, Web and semistructured data management, spatio-temporal databases and bioinformatics. On these topics, he has coauthored journal articles, conference papers and book chapters and developed practical software tools. He has served as a reviewer as well as a member of program committee for leading journals and conferences in the fields of databases and data mining, information systems, knowledge and data management and artificial intelligence. He has been a SIAM member since 2008 and an ACM member since 2009.

**Excellent addition to your library! Recommend to your acquisitions librarian.**