



Dr. **JAVIER NOGUERAS ISO** is an assistant professor of Computer Science at the University of Zaragoza and a member of the Advanced Information Systems Laboratory. Holding MS and PhD degrees in Computer Science from the University of Zaragoza, he has co-authored numerous publications and has collaborated on several R+D projects in the SDI context.



Dr. **F. JAVIER ZARAZAGA-SORIA** is an associate professor of Computer Science at the University of Zaragoza and a member of the Advanced Information Systems Laboratory. Holding a MS degree in Computer Science from the Technical University of Valencia and a PhD degree in Computer Science from the University of Zaragoza, he has co-authored numerous publications and has collaborated on several R+D projects in the SDI context.



Dr. **PEDRO R. MURO-MEDRANO** is an associate professor of Computer Science at the University of Zaragoza and the head of the Advanced Information Systems Laboratory. He has been visiting researcher in Carnegie Mellon University, the University of Maryland and the US National Institute of Health. Holding MS and PhD degrees in Industrial Engineering from the University of Zaragoza, he has co-authored numerous publications and has been main researcher in numerous R+D projects in the SDI context.

Nogueras-Iso  
Zarazaga-Soria  
Muro-Medrano



Geographic Information Metadata  
for Spatial Data Infrastructures

Nogueras-Iso · Zarazaga-Soria · Muro-Medrano  
Geographic Information Metadata  
for Spatial Data Infrastructures

Metadata play a fundamental role in both DLs and SDIs. Commonly defined as “structured data about data” or “data which describes attributes of a resource” or, more simply, “information about data”, it is an essential requirement for locating and evaluating available data. Therefore, this book focuses on the study of different metadata aspects, which contribute to a more efficient use of DLs and SDIs. The three main issues addressed are: the management of nested collections of resources, the interoperability between metadata schemas, and the integration of information retrieval techniques to the discovery services of geographic data catalogs (contributing in this way to avoid metadata content heterogeneity).

ISBN 3-540-24464-6



springeronline.com

Javier Nogueras-Iso  
F. Javier Zarazaga-Soria  
Pedro R. Muro-Medrano

# Geographic Information Metadata for Spatial Data Infrastructures

Resources, Interoperability  
and Information Retrieval

 Springer

Gestalter  
**ERICH KIRCHNER**  
Heidelberg

Druckfarben  
HKS 82 braun  
HKS 42 blau

Dieser Farblaser-Ausdruck  
dient nur als Anhaltspunkt  
für die farbliche Wiedergabe  
und ist nur bedingt  
farbverbindlich.  
Bitte benutzen Sie Ihre  
Pantone- oder HKS-Fächer  
um die Farbkombinationen  
zu überprüfen.



Javier Nogueras-Iso  
F. Javier Zarazaga-Soria  
Pedro R. Muro-Medrano

# Geographic Information Metadata for Spatial Data Infrastructures

Resources, Interoperability  
and Information Retrieval

With 83 Figures

 Springer

**DR. JAVIER NOGUERAS-ISO**  
**DR. F. JAVIER ZARAZAGA-SORIA**  
**DR. PEDRO R. MURO-MEDRANO**  
UNIVERSITY OF ZARAGOZA  
MARÍA DE LUNA, 1  
50018 ZARAGOZA  
SPAIN

*E-mail: jnog@unizar.es*  
*javy@unizar.es*  
*prmuro@unizar.es*

ISBN 10 3-540-24464-6 **Springer Berlin Heidelberg New York**  
ISBN 13 978-3-540-24464-6 **Springer Berlin Heidelberg New York**

Library of Congress Control Number: 2005920698

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, reuse of illustrations, recitation, broadcasting, reproduction on microfilm 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-Verlag. Violations are liable to prosecution under the German Copyright Law.

**Springer is a part of Springer Science+Business Media**  
springeronline.com  
© Springer-Verlag Berlin Heidelberg 2005  
Printed in The Netherlands

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: E. Kirchner, Heidelberg  
Production: A. Oelschläger  
Typesetting: Camera-ready by the Authors  
Printing: Krips, Meppel  
Binding: Litges+Dopf, Heppenheim

Printed on acid-free paper 30/2132/AO 5 4 3 2 1 0

---

## Acknowledgements

There are many people to whom we are grateful for their support during the evolution of this book, many more than the ones we have space to acknowledge here.

First of all, we would like to thank to the members and friends of the Advanced Information Systems Laboratory (IAAA) of the University of Zaragoza during these years of hard work, making a especial mention to Pedro Álvarez, Rubén Béjar, Miguel A. Latre, Joaquín Ezpeleta, Silvia Laiglesia, David Portolés, Rodolfo Rioja, Javier Lacasta, David Anaya, David Gayán, Oscar Cantán, Rafael Tolosana, Covadonga Fernández, Pablo Gallardo, Rafael Martínez, Rubén Moreno, Jesús Barrera, Raquel Miguel and Javier López. All of them have contributed in some way (ideas, developments, tools, discussions, ...) to the successful end of this book. Additionally, we are also indebted to the other members of the TeIDE Consortium from the University Jaume I and the Polytechnic University of Madrid because of their help and enthusiasm to boost the interest in Spatial Data Infrastructures in Spain.

We also want to remember here the people from the Geomatics & Remote Sensing General Office of the National Geographic Institute of Spain and especially to Sebastián Más and Antonio Rodríguez because of their interest and support.

We are very grateful to Michael Gould, Lars Bernard and José Angel Bañares, reviewers of some sections and chapters, and especially to Eva Méndez who made an extensive, deep, interdisciplinary and optimistic analysis. They found time in their busy schedules to read the text and provide valuable suggestions for its improvement.

Finally, we are absolutely grateful to our families and friends for all their patience, support and love. Much time have been stolen from our personal lives for the creation of this text. Undoubtedly, without their generous understanding this work would never have come into existence.

---

## Contents

<b>1</b>	<b>Spatial Data Infrastructures and related concepts</b>	<b>1</b>
1.1	Introduction	1
1.2	Components of a Spatial Data Infrastructure	3
1.3	Integrating Digital Libraries concepts within Spatial Data Infrastructures	5
1.3.1	Digital Libraries and Geolibraries	5
1.3.2	Digital Libraries versus Spatial Data Infrastructures	8
1.4	Metadata types and standardization initiatives	9
1.4.1	Introduction	9
1.4.2	General purpose metadata standards	10
1.4.3	Metadata schemas for geographic resources	13
1.4.4	Metadata schemas for service description	17
1.5	Technical components of Spatial Data Infrastructures and the role of metadata	18
1.6	Ontologies and Knowledge Representation in the context of Spatial Data Infrastructures	23
1.7	Conclusions	29
<b>2</b>	<b>A metadata infrastructure for the management of nested collections</b>	<b>31</b>
2.1	Introduction	31
2.2	Related work	35
2.2.1	Addressing collections and relations in metadata standards	35
2.2.2	Collections in Digital libraries and Geolibraries	43
2.2.3	Addressing relations in knowledge representation	47
2.3	Defining the desired functionality of a collection enabled catalog system	49
2.4	The Metadata Knowledge Base	55
2.4.1	Building the catalog services over a metadata knowledge base	55

2.4.2	The knowledge base model .....	57
2.4.3	Automatic generation of metadata .....	65
2.4.4	Intelligent query answering .....	72
2.5	Building aggregation relations .....	82
2.6	Conclusions and future work .....	85
<b>3</b>	<b>Interoperability between metadata standards .....</b>	<b>89</b>
3.1	Introduction .....	89
3.2	Related work .....	93
3.2.1	Ontology based semantic interoperability .....	94
3.2.2	Crosswalk based semantic interoperability .....	96
3.3	Construction of crosswalks between metadata standards .....	99
3.3.1	Harmonization .....	100
3.3.2	Semantic mapping .....	102
3.3.3	Additional rules for metadata conversion .....	103
3.3.4	Implementation of crosswalks: the use of style sheets ...	106
3.4	Putting the method to work .....	115
3.4.1	Transformation between CSDGM and ISO 19115 .....	116
3.4.2	Transformation between ISO 19115 and Dublin Core ...	120
3.5	Conclusions and future work .....	125
<b>4</b>	<b>The use of disambiguated thesauri to improve information retrieval .....</b>	<b>129</b>
4.1	Introduction .....	129
4.2	Basic concepts about thesaurus and WordNet .....	133
4.2.1	Thesaurus .....	133
4.2.2	WordNet .....	135
4.3	The Semantic Disambiguation of Thesauri .....	136
4.3.1	State of the art in Semantic Disambiguation .....	136
4.3.2	Description of the semantic disambiguation method ...	142
4.3.3	Testing the method .....	148
4.4	The information retrieval model .....	152
4.4.1	State of the art in sense based information retrieval ...	152
4.4.2	Introduction to the vector-space retrieval model .....	154
4.4.3	The indexing of metadata records .....	155
4.4.4	The indexing of queries .....	157
4.4.5	Testing the retrieval model .....	158
4.5	Conclusions and future work .....	166
<b>5</b>	<b>Integrating the concepts within the components of a Spatial Data Infrastructure .....</b>	<b>169</b>
5.1	Introduction .....	169
5.2	The catalog services component .....	171
5.2.1	Introduction .....	171

5.2.2	Integrating the interoperability between metadata standards .....	173
5.2.3	Integrating the concept based information retrieval ....	176
5.3	A metadata editor .....	178
5.3.1	Introduction .....	178
5.3.2	Import/Export of metadata .....	181
5.3.3	Collection Metadata Edition .....	181
5.3.4	Thesaurus Management .....	185
5.4	The Web Portal of a Spatial Data Infrastructure .....	192
5.5	Conclusions and future work .....	194
<b>6</b>	<b>Conclusions and future work</b> .....	197
<b>A</b>	<b>Collections</b> .....	207
A.1	Consistency of the metadata model .....	207
A.2	Metadata Inference .....	217
A.2.1	Generation of complete values .....	217
A.2.2	Update of whole-part hierarchy .....	219
A.2.3	Example of a <i>wholeInferredValues</i> specification .....	221
<b>B</b>	<b>Crosswalks</b> .....	225
B.1	CSDGM→ISO 19115 stylesheet .....	225
B.2	ISO 19115→DC stylesheet .....	227
B.3	DC→ISO 19115 stylesheet .....	228
<b>C</b>	<b>Applications</b> .....	231
C.1	Revision of geographic metadata editors .....	231
C.2	Revision of thesaurus tools .....	236
	<b>References</b> .....	241
	<b>Index</b> .....	259