Krassimira Ivanova, Milena Dobreva, Peter Stanchev, George Totkov
(editors)

Innovative Applications of Automated Metadata Generation

Plovdiv University Publishing House "Paisii Hilendarski"
2012, Plovdiv, Bulgaria

Access to Digital Cultural Heritage:
Innovative Applications of Automated Metadata Generation

Edited by:
Krassimira Ivanova, Milena Dobreva, Peter Stanchev, George Totkov

Authors (in order of appearance):
Krassimira Ivanova, Peter Stanchev, George Totkov, Kalina Sotirova, Juliana Peneva, Stanislav Ivanov, Rositza Doneva, Emil Hadjikolev, George Vragov, Elena Somova, Evgenia Velikova, Iliya Mitov, Koen Vanhoof, Benoit Depaire, Dimitar Blagoev

Reviewer: Prof., Dr. Avram Eskenazi

Published by: Plovdiv University Publishing House "Paisii Hilendarski"
2012, Plovdiv, Bulgaria
First Edition

The main purpose of this book is to provide an overview of the current trends in the field of digitization of cultural heritage as well as to present recent research done within the framework of the project D002-308 funded by Bulgarian National Science Fund. The main contributions of the work presented are in organizing digital content, metadata generation, and methods for enhancing resource discovery.

Printed in Bulgaria by Plovdiv University
24, Tsar Assen, Str., Plovdiv-4000, Bulgaria

ISBN: 978-954-423-722-6
Plovdiv, 2012

Acknowledgements

This book summarises the outcomes of several recent research projects. It is quite a complex project in terms of scope and number of contributors. We would like to particularly thank all authors for their enthusiasm and commitment.

We would also like to thank all our colleagues who were supportive of our work and thus helped to develop our competences on the topic.

The projects which helped to develop our ideas and test them in real life were supported by:

- the Bulgarian National Science Fund, namely the Project D002-308 "MetaSpeed: Automated Metadata Generating for e-Documents Specifications and Standards";
- the Framework Programme 7 (FP7) of the European Commission, namely the project RI-246686 "OpenAIRE: Open Access Infrastructure for Research in Europe";
- the Hasselt University in Belgium, namely the Projects R-1875 "Search in Art Image Collections Based on Colour Semantics" and R-1876 "Intelligent systems' memory structuring using multidimensional numbered information spaces".

The involvement of Bulgarian researchers in these projects in particular was made possible through the cooperation of several institutions: we would like to express our gratitude to the *Institute of Mathematics and Informatics – Bulgarian Academy of Sciences*, *Plovdiv University "Paisii Hilendarski"*, *New Bulgarian University*, and *Hasselt University in Belgium* for providing excellent conditions for collaboration.

A number of international as well as national events gave us possibilities to present our visions and discuss with other colleagues from different professional communities; these discussions were also of great contribution to our work. Of particular help were the international events organised by the Member of the European Parliament Emil Stoyanov; he also welcomed the idea of creating this collection.

We also want to thank our reviewer Professor Avram Eskenazi for his helpful remarks during the process of prepariation of the content of this publication.

Last but not least, we would like to thank Emilia Todorova from Glasgow Caledonian University for the help with language revision and to Viktoria Naoumova from the Institute of Mathematics and Informatics for technical assistance.

Table of Contents

List of Abbreviations

5M	Multicultural, Multilingual, Multimodal, Multivariate, Modelling
5S	Streams, Structures, Spaces, Scenarios, and Societies
AAT	Art and Architecture Thesaurus
ACRI	Associative Classifier with Reoccuring Items
ACTA	Anti-Counterfeiting Trade Agreement
AIP	Archival Information Package
APICAS	Art Painting Image Colour Aesthetics and Semantics
ARC-AC	Association Rule-based Categorizer for All Categories
ARC-BC	Association Rule-based Categorizer By Category
ARM	Association Rule Mining
ArM	Archive Manager
ARUBAS	Association RUle BAsed Similarity framework
BIDL	Bulgarian Iconographical Digital Library
CAD	Computer-aided Design
CAR	Class-Association Rules
CATCH	Continuous Access to Cultural Heritage
CBA	Classification Based on Associations
CBIR	Content-Based Image Retrieval
CCA	Curvilinear Component Analysis
CCSDS	Consultative Committee for Space Data Systems
CDWA	Categories for the Description of Works of Art
CH	Cultural Heritage
CHO	Cultural and Historical Objects
CIDOC CRM	International Committee for Documentation – Conceptual Reference Model
CL	Colour Layout
CMAR	Classification based on Multiple Association Rules
CMY	Cyan-Magenta-Yellow
CONA	Cultural Objects Name Authority
CorClass	Correlated Association Rule Mining for Classification
CORDIS	Community Research & Development Information Service
CPAR	Classification based on Predictive Association Rules
CS	Colour Structure

CSDGM	Content Standard for Digital Geospatial Metadata
DACS	Describing Archives: a Content Standard
DC	Dominant Colour
DC	Dublin Core
DCP	Data Coverage Pruning
DELOS	Network of Excellence on Digital Libraries
DHO	Digital Humanities Observatory
DIP	Dissemination Information Package
DL	Digital Library
DLRM	Digital Libraries Reference Model
DOI	Digital Object Identifier
DWT	Discrete Wavelet Transform
EAD	Encoded Archival Description
EC	European Commission
ECDL	European Conference on Digital Libraries
EDL	European Digital Library
EDM	Europeana Data Model
EH	Edge Histogram
EMD	Earth Mover's Distance
EOF	Empirical Orthogonal Function
Fedora	Flexible Extensible Digital Object Repository Architecture
FOIL	First Order Inductive Learner
FP7	Seventh Framework Programme
FRBR	Functional Requirements for Bibliographic Records
FRBROO	FRBR – Object Oriented
GIS	Geographic Information System
GLAM	Galleries, Libraries, Archives, Museums
GLOH	Gradient Location and Orientation Histogram
GPS	Global Positioning System
HARMONY	Highest confidence clAssification Rule Mining fOr iNstance-centric classifYing
HSIS	Humanities Serving Irish Society
HSL	Hue-Saturation-Luminance
HSV	Hue-Saturation-Value
HT	Homogeneous Texture
HTML	Hyper-Text Markup Language
ICCROM	International Centre for the Study of the Preservation and Restoration of Cultural Property
ICOM	International Council of Museums
ICT-CIP	Information and Communication Technologies – Competitiveness and Innovation Framework Programme
IDABC	Interoperable Delivery of European eGovernment Services to public Administrations, Businesses and Citizens
IFLA	International Federation of Library Associations

IMI-BAS	Institute of Mathematics and Informatics – Bulgarian Academy of Sciences
IRI	International Resource Identifier
ISAAR(CPF)	International Standard Archival Authority Record for Corporate Bodies, Persons and Families
ISAD(G)	General International Standard Archival Description
ISOC	Internet Society
IT	Information Technology
JISC	Joint Information Systems Committee
LESH	Local Energy based Shape Histogram
LIDAR	Light Detection And Ranging
LIDO	Light Information Describing Objects
LLE	Locally Linear Embedding
MARC	MAchine-Readable Cataloging
MBR	Minimum Bounding Rectangle
MDS	Multi Dimensional Scaling
MET	Metropoliten Museum of Art
METS	Metadata Encoding & Transmission Standard
MINERVA	MInisterial NEtwoRk for Valorising Activities in digitisation
MODS	Metadata Object Description Schema
MPEG	Moving Picture Experts Group
NSDL	National Science Digital Library
NURBS	Non Uniform Rational BSpline
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OAIS	Open Archival Information System
OpenAIRE	Open Access Infrastructure Research for Europe
ORE	Ontology Rule Editor
OWL	Web Ontology Language
PCA	Principal Component Analysis
PGN	Pyramidal Growing Network
PP	Projection Pursuit
R&D	Research and Development
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
REGATTA	REGional Aggregator of heTerogeneous culTural Artefacts
RGB	Red-Green-Blue
RYB	Red-Yellow-Blue
SAIL	Semi-Automated Interactive Learning systems
SC	Scalable Colour
SDA	Symbolic Data Analysis
SGML	Standard Generalized Markup Language
SIFT	Scale-Invariant Feature Transform
SIP	Submission Information Package
SRES	Self-supervised web relation Extraction System
SRSWOR	Simple Random Sample WithOut Replacement

SRSWR	Simple Random Sample With Replacement
SURF	Speeded Up Robust Feature
SVD	Singular Value Decomposition
SVM	Support Vector Machines
TEL	The European Library
TFPC	Total From Partial Classification
TGN	Thesaurus of Geographic Names
TPDL	Theory and Practice of Digital Libraries
ULAN	Union List of Artist Names
UNESCO	United Nations Educational, Scientific and Cultural Organization
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
VQ	Vector Quantization
VRA	Visual Resources Association
W3C	World Wide Web Consortium
WDL	World Digital Library
WIPO	World Intellectual Property Organization
XML	eXtensible Markup Language