# ADVANCES ON INFORMATION PROCESSING AND MANAGEMENT

#### **CONFERENCE ORGANIZERS INSTITUTES**

The International Conference on Integrated Information is supported by the following Institutes:

Emerald Group Publishing Limited
Technological educational Institute of Athens, Greece
University of Peloponnese, Greece
National And Kapodistrian University of Athens, Greece
Mednet Hellas, The Greek Medical Network
2<sup>nd</sup> AMICUS Workshop

# INTEGRATED INFORMATION

International Conference on Integrated Information

Kos, Greece September, 29 – October, 3 2011

## **EDITORS**

Georgios A. Giannakopoulos *Technological Educational Institute of Athens, Greece* 

Damianos P. Sakas *University of Peloponnese, Greece* 

# All papers have been peer-reviewed



Piraeus, Greece, 2011

#### **Editors**

Georgios A. Giannakopoulos

Technological Educational Institute of Athens Faculty of Management and Economics Department of Library Science and Information Systems Address: Aghiou Spyridonos Street, 12210, Egaleo

E-mail: gian@teiath.gr

Damianos P. Sakas

University of Peloponnese Faculty of Science and Technology Department of Computer Science and Technology Address: End of Karaiskaki St., 22100, Tripolis, Greece

E-mail: D.Sakas@uop.gr

The copyrights will be owned by the authors under the Creative Commons Attribution-Non Commercial license (<a href="http://creativecommons.org/licenses/by-nc/3.0/">http://creativecommons.org/licenses/by-nc/3.0/</a>), which permits unrestricted use, distribution, and reproduction in any non commercial medium, provided the original work is properly cited.

ISSN:

Printed in the Greece, EU

# **CONTENTS**

PREFACE: Proceedings of the International Conference on Integrated Information (IC-INFO 2011)	1
Georgios A. Giannakopoulos, Damianos P. Sakas	
Conference Details	3
Keynote Speaker	5
SYMPOSIUM ON INFORMATION AND KNOWLEDGE MANAGEMENT Prof. Christos Skourlas	6
Towards the Preservation and Availability of Historical Books and Manuscripts: A Case Study  Eleni Galiotou	8
An Extensive Experimental Study on the Cluster-based Reference set Reduction for Speeding-up the k-nn Classifier  Stefanos Ougiaroglou, Georgios Evangelidis and Dimitris A. Dervos	12
Exploiting the Search Culture Modulated by the Documentation Retrieval Applications Nikitas N. Karanikolas and Christos Skourlas	16
Information and Knowledge Organization: The Case of the TEI of Athens Anastasios Tsolakidis, Manolis Chalaris and Ioannis Chalaris	22
Providing Access to Students with Disabilities and Learning Difficulties in Higher Education through a Secure Wireless framework  Catherine Marinagi and Christos Skourlas	26
Improving Query Efficiency in High Dimensional Point Indexes Evangelos Outsios and Georgios Evangelidis	30
Text Segmentation Using Named Entity Recognition and co-Reference Resolution in Greek Texts	34
Pavlina Fragkou	
KINISIS, a Graphical XQuery Language Euclid Keramopoulos, Achilleas Pliakas, Konstantinos Tsekos and Ignatios Deligiannis	42
Dimensionality Curse, Concentration Phenomenon and the KDB-tree Nikolaos Kouiroukidis and Georgios Evangelidi	46

Applying Balanced Scorecard Strategic Management in Higher Education Manolis Chalaris, Anastasios Tsolakidis and Ioannis Chalaris	50
A Web Portal Model for NGOs' Knowledge Management Zuhal Tanrikulu	54
The Digital Archives System and Application Optimized for the Tradition Knowledge Archives	58
Jeon Hong. Chan, In Deok. Hwang, Jae Hak. Park, Hyeok. Sim, U won. Gwon and Soon Cheol. Park	
A Semi-automatic Emerging Technology Trend Classifier Using SCOPUS and PATSTAT	62
Seonho Kim, Woondong Yeo, Byong-Youl Coh, Waqas Rasheed, Jaewoo Kang	
Presenting a Framework for Knowledge Management within a Web Enabled Living Lab Lizette de Jager and Albertus AK Buitendag and Potjie (JS) van der Walt	66
4TH SYMPOSIUM ON BUSINESS AND MANAGEMENT AND DYNAMIC SIMULATION MODELS SUPPORTING MANAGEMENT STRATEGIES Dr. Damianos Sakas	71
New Political Communication Practices: No Budget Events Management. The New Challenge  Evangelia N. Markaki, Damianios P. Sakas and Theodore Chadjipantelis	73
Free Software – Open Source Software. A Powerful Tool for Developing Creativity in the Hands of the Student  Nacionavlas K. Dimitrias Demicros P. Sakas Konstantinos Massalas	78
Nasiopoulos K. Dimitrios, Damianos P. Sakas, Konstantinos Masselos	
Open Source Web Applications. How it Spread Through the Internet and their Contribution to Education.	82
Nasiopoulos K. Dimitrios, Damianos P. Sakas, Konstantinos Masselos	
Culture in Modern Times in the Frame of Luhmann's System Theory Anastasia J. Chournazidis	85
Managing Scientific Journals: A Cultural Viewpoint Marina C. Terzi, Damianos P. Sakas, and Ioannis Seimenis	87
A Conceptual Framework for Analyzing Knowledge-based Entrepreneurship Nikos S. Kanellos	92

SESSION ON INFORMATION HISTORY: PERSPECTIVES, METHODS AND CURRENT TOPICS	96
Prof. Laszlo Karvalics	
Emerging Research Fields in Information History Laszlo Z. Karvalics	98
Information Management through Elementary Data Clusters: New Observations on Pridianum-Type Roman Statistical Documents  Gergő Gellérfi	102
Information and Secrecy on the Silk Road. Methods of Encryption of Legal Documents in Inner Asia (3th-4th century)  Szabolcs Felföldi	106
The Role of Information and Disinformation in the Establishment of the Mongolian Empire: A Re-examination of the 13th century Mongolian History from the Viewpoint of Information History	110
Márton Gergő Vér	
Early Warning Systems and the Hospitallers in the Eastern Mediterranean Zsolt Hunyadi	114
Information Management as Establishment Dutch Navigational Knowledge on Japan, 1608-1641	118
Gabor Szommer	
Files Everywhere - Register and Training of Men for Military and Civil Purpose in Prussia in the early 18th century  Marton Holczer	123
SYMPOSIUM ON INTEGRATED INFORMATION: THEORY, POLICIES, TOOLS Prof. Georgios Giannakopoulos	126
Approaching Information as an Integrated Field: Educating Information Professionals Georgios Giannakopoulos, Daphne Kyriaki Manesi and Sryridon Zervos	128
Special Libraries as Knowledge Management Centers Eva Semertzaki	132
Digital Libraries' Developers and their Suitability: A Case Study Maria Monopoli	136

A Preliminary Study for the Creation of a Greek Citation index in the Humanities and the Social Sciences (GCI – H&SS)	140
Daphne Kyriaki-Manessi and Evi Sachini	
Archiving as an Information Science. Evidence from a Survey Carried out on a Sample of Greek Students	144
Georgios Giannakopoulos and Ioannis Koumantakis	
Transition Process of E-records Management and Archiving System in Universities: Ankara University	147
S. Özlem Bayram and Fahrettin Ozdemirci	
Government Information: Access and Greece's Efforts for Access Aikaterini Yiannoukakou	150
School Archives and their Potentials in Teaching: Aspects of Greek Reality Sonia Geladaki and Panagiota Papadimitriou	156
Research on School Libraries in Greece and Suggestions on its Further Development Georgios D. Bikos	160
Building Digital Collections for Archeological Sites: Metadata Requirements and CIDOC CRM Extension	164
Georgios S. Gkrous and Mara Nikolaidou	
Museological Claims to Autonomous Knowledge: Rethinking the Conceptual Mode of Display and its Claims to Knowledge	169
Assimina Kaniari and Georgios Giannakopoulos	
Use of Library Loan Records for Book Recommendation Keita Tsuji, Erika Kuroo, Sho Sato, Ui Ikeuchi, Atsushi Ikeuchi, Fuyuki Yoshikane and Hiroshi Itsumura	172
Developing a National Database on Librarianship and Information Science. The Case of E-VIVA, the Hellenic Fulltext Database Filippos Ch. Tsimpoglou, Vasiliki V. Koukounidou and Eleni K. Sakka	176
Integrated Access to Cultural Heritage Information Pieces in Iran Astan-Quds Razavi's Organization of Libraries, Museums and Documents Center: A Theory of Unionization Disparate Information Assets over Imam Reza's Zarih  Ms. Mitra Zarei and Ms. Maliheh Farrokhnia	181
Attitudes of University Librarians and Information Scientists towards the Draft Code of	185

# Library Ethics to Present a Model for Final Library Ethical Codes

Mahsoomeh Latifi, Fatemeh Zandian and Hasan Siamian

SESSION ON OPEN ACCESS REPOSITORIES: SELF-ARCHIVING, METADATA, CONTENT POLICIES, USAGE	188
Dr. Alexandros Koulouris	
Geographical Collections in Greek Academic Libraries: Current Situation and Perspectives	189
Ifigenia Vardakosta and Sarantos Kapidakis	
Information Seeking Behavior: Factors that Affect the Behavior of Greek Astronomers Hara Brindesi and Sarantos Kapidakis	194
Aggregating Metadata for Europeana: The Greek Paradigm Alexandros Koulouris, Vangelis Banos and Emmanouel Garoufallou	198
Integrating a Repository with Research Output and Publications: The Case of the National Technical University of Athens	202
Dionysis Kokkinos	
Implementation of Workflows as Finite State Machines in a National Doctoral Dissertations Archive	205
Nikos Houssos, Dimitris Zavaliadis, Kostas Stamatis and Panagiotis Stathopoulos	
Practices of "Local" Repositories of Legally Protected Immovable Monuments. A Global Scheme for 'Designation – Significance' Information  Michail Agathos and Sarantos Kapidakis	209
Sur sa a man a rea of	
Integration of Metadata in BWMETA-2.0.0 Format  Katarzyna Zamlynska, Jakub Jurkiewicz and Lukasz Bolikowski	213
SESSION ON EVIDENCE-BASED INFORMATION IN CLINICAL PRACTICE Dr. Evangelia Lappa	216
Applicability of Data Mining Algorithms on Clinical Datasets Wilfred, Bonney	218
Changing Roles of Health Librarians with Open Access Repositories Christine Urquhar and Assimina Vlachaki	221
From Medical Records to Health Knowledge Management Systems: The Coding to Health Sector	225
Evangelia C. Lappa and Georgios A. Giannakopoulos	

The Survey of Skill, Attitude and Use of Computer and Internet among Faculty Members Hasan Siamian, Azita Bala Ghafari, Kobra Aligolbandi, Mohammad Vahedi and Gholam Ali Golafshani Jooybari	229
Trends in Scholarly Communication among Biomedical Scientists in Greece Assimina Vlachaki and Christine Urquhart	232
SESSION ON ELECTRONIC PUBLISHING: A DEVELOPING LANDSCAPE Dr. Dimitris Kouis	236
E-Journal and Open Access Journal Publishing in the Humanities: Preliminary Results from a Survey among Byzantine Studies Scholars Victoria Tsoukala and Evi Sachini	238
Preliminary Results on a Printed VS Electronic Text Books Assessment Through Questionnaire  Dimitrios A. Kouis and Kanella Pouli	242
An Interpretation of Aristotelian Logic According to George Boole Markos N. Dendrinos	246
SESSION ON INFORMATION CONTENT PRESERVATION AS OUTCOME OF CONSERVATION OF CULTURAL HERITAGE: ETHICS, METHODOLOGY AND TOOLS  Prof. George Panagiaris and Dr. Spiros Zervos	251
Intrinsic Data Obfuscation as the Result of Book and Paper Conservation Interventions Spiros Zervos, Alexandros Koulouris and Georgios Giannakopoulos	254
Mass Deacidification: Preserving More than Written Information Michael Ramin, Evelyn Eisenhauer and Markus Reist	258
Information Literacy of Library Users: A Case Study of Mazandaran Public Library Users, Iran Hussein Mahdizadeh and Hasan Siamian	260
The Narratives of Paper in The Archives of the New Independent Greek State (Mid 19th c.)  Ourania Kanakari and Maria Giannikou	264
From Macro to Micro and from Micro to Nano: The Evolution of the Information Content Preservation of Biological Wet Specimen Collections Nikolaos Maniatis and Georgios Panagiaris	268

Digital images: A valuable scholar's tool or misleading material?  Patricia Engel	272
Attitudes of University Librarians and Information Scientists Towards the Draft Code of Library Ethics to Present a Model for Final Library Ethical Codes  Mahsoomeh Latifi, Fatemeh Zandianand and Hasan Siamian	277
Investigation of the Degradation Mechanisms of Organic Materials: From Accelerated Ageing to Chemometric Studies	280
Ekaterini Malea, Effie Papageorgiou and Georgios Panagiaris	
SESSION ON DIVERGENCE AND CONVERGENCE: INFORMATION WORK IN DIGITAL CULTURAL MEMORY INSTITUTIONS  Dr. Susan Myburgh	285
Extending Convergence and Divergence in Cultural Memory Institutions: The Old Slave Lodge in the New South Africa  Archie L Dick	287
The Transfer of Knowledge from Large Organizations to Small: Experiences from a Research Project on Digitization in Wales  Clare Wood-Fisher, Richard Gough, Sarah Higgins, Menna Morgan, Amy Staniforth and Lucy Tedd	289
The Usage of Reference Management Software (Rms) in an Academic Environment : A Survey at Tallinn University  Enrico Francese	293
Varialog : How to Locate Words in a French Renaissance Virtual Library Marie-Hélène Lay	297
The Urge to Merge: A Theoretical Approach Susan Myburgh	301
SYMPOSIUM ON ADVANCES INFORMATION FOR STRATEGIC MANAGEMENT Professor Nikolaos Konstantopoulos	304
Empowerment in the Tax Office of Greece Antonios E. Giokas and Nikolaos P. Antonakas	306
Building Absorptive Capacity Through Internal Corporate Venturing Ioannis M. Sotiriou and Alexandros I. Alexandrakis	310

The Monitoring Information System (M.I.S.) - An information and Management System for Projects Co-financed Under the National Strategic Reference Framework (NSRF) and the Community support framework (CSF)	
Catherina G. Siampou, Eleni G. Fassou and Athanassios P. Panagiotopoulos	
Corruption in Tax Administration: The Entrepreneurs View Point	318
Nikolaos P. Antonakas, Antonios E. Giokas and Nikolaos Konstantopoulos	
Conflicts between the IT Manager and the Software House after the Strategic Choice of Outsourcing of the Information Processes in Maritime Companies.	322
Anthi Z. Vaxevanou, Nikolaos Konstantopoulos, Damianos P. Sakas	
Contemporary Forms of Ordering Between the Supply Department and Ship Chandler Companies in the Shipping Industry	325
Anthi Z. Vaxevanou, Nikolaos Konstantopoulos, Damianos P. Sakas	
Strategies Implemented and Sources Used for the Acquisition of Information on Foreign Markets	329
Myropi Garri, Nikolaos Konstantopoulos and Michail G. Bekiaris	
The Effect of High Performance Working Systems on Informative Technology in Enterprises after Organisation Changes such as Mergers & Acquisitions	333
Nikolaos Konstantopoulos and Yiannis Triantafyllopoulos	
Personnel's Absorptive Capacity as a Guiding Concept for Effective Performance in Informative Technology	337
Nikolaos Konstantopoulos and Yiannis Triantafyllopoulos	
SESSION ON CONTEMPORARY ISSUES IN MANAGEMENT: ORGANISATIONAL BEHAVIOUR, INFORMATION TECHNOLOG, EDUCATION & HOSPITAL LEADERSHIP	341
Dr. Panagiotis Trivellas	
Investigating the Importance of Sustainable Development for Hotel SMES Panagiotis Reklitis and Anestis Fotiadis	343
Strategic Alignment of ERP, CRM and E-business: A Value Creation Catherine C. Marinagi and Christos K. Akrivos	347
The Impact of Occupational Stress on Performance in Health Care Panagiotis Trivellas Panagiotis Reklitis and Charalambos Platis	351

The Impact of Emotional Intelligence on Job Outcomes and Turnover Intention in Health Care	
Panagiotis Trivellas Vassilis Gerogiannis and Sofia Svarna	
SYMPOSIUM ON BUSINESS MANAGEMENT AND COMMUNICATION STRATEGIES SUPPORTING DECISION MAKING PROCESS IN TOURISM SECTOR	360
Dr. Panagiota Dionysopoulou	
The Human Factor as a Mediator to the Total Quality in the Tourism Companies. The impact of Employees' Motivation to Quality Improvements	362
Christos K. Akrivos and Panagiotis Reklitis	
Tourist Destination Marketing and Management Using Advanced ICTS Technologies Anastasia Argyropoulou, Panagiota Dionyssopoulou, Georgios Miaoulis	365
G.N.T.O. (Greek National Tourism Organization) Communication Strategy in Advertising Campaigns 1991-2006	370
George Stafylakis and Panagiota Dionyssopoulou	
GENERAL PAPERS	375
The role of Environmental Education within the Framework of the Environmental	376
Policy of a Regional Municipality	
Vassiliki Delitheou and Dimitra Thanasia	
Issues of Social Cohesion: A case study from the Greek Urban Scenery Evgenia Tousi	380
Merging Activity and Employee Performance: The Greek Banking System Panagiotis Liargovas and Spyridon Repousis	384
Sustainable Development and Corporate Social Responsibility in Higher Education: Some	387
Evidence from Greece	
Anastasios Sepetis and Fotios Rizos	
Exploring the Effects of Organizational Culture on Collaborative vs. Competitive	
Knowledge Sharing Behaviors	395
Hanan Abdulla Mohammed Al Mehairi and Norhayati Zakaria	

# Preface: Proceedings of the International Conference on Integrated Information (IC-ININFO 2011)

#### GEORGIOS A. GIANNAKOPOULOS

Department of Library Science and Information Systems, Technological Educational Institute of Athens, Aghiou Spyridonos, Egaleo, 12210, Greece

#### DAMIANOS P. SAKAS

Department of Computer and Technology Science, University of Peloponnese, Praxitelous 89-91, Piraeus, 18532, Greece

## Aims and Scope of the Conference

The International Conference on Integrated Information 2011 took place in Kos Island, Greece, between September, 29 and October, 3, 2011. IC-ININFO is an international interdisciplinary conference covering research and development in the field of information management and integration.

The conference aims at creating a forum for further discussion for an Integrated Information Field incorporating a series of issues and/or related organizations that manage information in their everyday operations. Therefore, the call for papers is addressed to scholars and/or professionals of the fields of Library and Archives Science (including digital libraries and electronic archives), Museum and Gallery Studies, Information Science, Documentation, Information Management, Records Management, Knowledge Management, Data management and Copyright experts the latter with an emphasis on Electronic Publications. Furthermore, papers focusing on issues of Cultural Heritage Management and Conservation Management are also be welcomed along with papers regarding the Management of Nonprofit Organizations such as libraries, archives and museums.

One of the primary objectives of the IC-ININFO will be the investigation of information-based managerial change in organizations. Driven by the fast-paced advances in the Information field, this change is characterized in terms of its impact on organizations that manage information in their everyday operations.

Grouping emerging technologies in the Information field together in a close examination of practices, problems and trends, IC-ININFO and its emphases on integration and management will present the state of the art in the field. Addressed jointly to the academic and practitioner, it will provide a forum for a number of perspectives based on either theoretical analyses or empirical case studies that will foster dialogue and exchange of ideas.

# **Topics of general Interest**

Library Science, Archives Science, Museum and Gallery Studies, Information Science, Documentation, Digital Libraries, Electronic Archives, Information Management, Records / Document Management, Knowledge Management, Data Management, Copyright, Electronic Publications, Cultural Heritage Management, Conservation Management, Management of Nonprofit Organizations, History of Information, History of Collections, Health Information

# Symposia

The Conference offered a number of sessions under its patronage, providing a concise overview of the most current issues and hands-on experience in information-related fields.

- Symposium on Integrated information: Theory, Policies, Tools
- 4th Symposium on Business and Management and Dynamic Simulation Models supporting management strategies

- Session on Open Access Rrepositories: Self-archiving, Metadata, Content policies, Usage
- Session on Evidence-Based Information in Clinical Practice
- Session on Business Management and Communication Strategies supporting Decision Making Process in Tourism Sector
- Session on Electronic Publishing: A Developing Landscape
- Session on Information and Knowledge Management
- Session on Information Content Preservation as Outcome of Conservation of Cultural Heritage: Ethics, Methodology and Tools
- Session on Advances Information for Strategic Management
- Session on Information History: Perspectives, Methods and Current Topics
- Session on Divergence and Convergence: Information Work in Digital Cultural Memory Institutions
- Session on Contemporary issues in Management: Organisational Behaviour, Information Technology, Education & Hospital leadership.

The wide range of aspects that the sessions covered, highlighted future trends in the Information Science.

## **Paper Peer Review**

More than 300 papers had been submitted for consideration in IC-ININFO 2011. From them, 91 were selected for presentation, after peer review in a double blind review process. The accepted papers were presented at IC-ININFO 2011.

#### **Thanks**

We would like to thank all members that participated in any way in the IC-ININFO 2011 Conference and especially:

- The famous publishing house Emerald for its communication sponsorship.
- The co-organizing Universities and Institutes for their support and development of a high-quality Conference scientific level and profile.
- The members of the Scientific Committee that honored the Conference with their presence and provided a significant contribution to the review of papers as well as for their indications for the improvement of the Conference.
- All members of the Organizing Committee for their help, support and spirit participation before, during and after the Conference.
- The Session Organizers for their willing to organize sessions of high importance and for their editorial work, contributing in the development of valued services to the Conference.
- PhDc Marina Terzi for her excellent editorial work, contributing in the production of the Conference proceedings.

#### **CONFERENCE DETAILS**

#### Chairs

Georgios A. Giannakopoulos, Technological Educational Institute of Athens, Greece Damianos P. Sakas, University of Peloponnese, Greece

#### **Co-Chairs**

Daphne Kyriaki – Manesi, Technological Educational Institute of Athens, Greece Dimitrios Vlachos, University of Peloponnese, Greece

#### **Scientific Committee**

Amanda Spink, Queensland University of Technology

Andreas Bagias, European Court

Andreas Rauber, Vienna University of Technology

Astrid van Wesenbeeck, SPARC Europe

Christine Urquhart, Aberystwyth University

Christos Schizas, University of Cyprus

Christos Skourlas, Technological Educational Institute of Athens

Claire Farago, University of Colorado at Boulder

Claus-Peter Klas, FernUniversität in Hagen

Costas Vassilakis, University of Peloponnese,

Dimitris Dervos, Technological Educational Institute of Thessaloniki

Eelco Ferwerda, OAPEN

Elena Garcia Barriocanal, University of Alcalá

Emmanouel Garoufallou, Technological Educational Institute of Thessaloniki

Filippos Tsimpoglou, University of Cyprus

Fillia Makedon, University of Texas at Arlington

George Korres, University of Newcastle

Georgios Evangelidis, University of Macedonia

Georgios Panagiaris, Technological Educational Institute of Athens

Johan Oomen, Netherlands Institute for Sound and Vision

José Aldana, University of Malaga

Konstantinos Masselos, University of Peloponnese

Luciana Duranti, The University of British Columbia

Markos N. Dendrinos, Technological Institute of Athens

Milena Dobreva, University of Strathclyde

Prodromos Tsiavos, London School of Economics and Political Science

Sándor Darányi, University of Borås

Sarantos Kapidakis, Ionian University

Sirje Virkus, Tallinn University

Spiros Zervos, Technological Educational Institute of Athens

Susan Myburgh, University of South Australia

Theodoros Pitsios, University of Athens, Faculty of Medicine

#### **Organizing Committee**

Alexandros Koulouris (Chair), Technological Educational Institute of Athens

Christos Christopoulos, SCEV Scientific Events Ltd

Marina Terzi, University of the Aegean, Greece

Evangelia Markaki, Aristotle University of Thessaloniki

Assimina Kaniari, Athens School of Fine Arts Evangelia Lappa, General Hospital Attikis K.A.T. Dimitris Kouis, Greek Ministry of Education, Lifelong Learning and Religious Affairs Dionysis Kokkinos, National Technical University of Athens

#### KEYNOTE SPEAKER



Professor Amanda Spink

Professor Amanda Spink has published over 340 scholarly journal articles, refereed conference papers and book chapters, and 6 books. Many of her journal articles are published in the Journal of the American Society for Information Science and Technology, Information Processing and Management, and the Journal of Documentation. She is Editor of the Emerald journal Aslib Proceedings. Amanda's research has been published at many conferences including ASIST, IEEE ITCC, CAIS, Internet Computing, ACM SIGIR, and ISIC Conferences. Her recent books include Information Behavior: An Evolutionary Instinct and Web Search: Multidisciplinary Perspectives, both published by Springer. Amanda's research focuses on theoretical and empirical studies of information behavior, including the evolutionary and developmental foundations. The National Science Foundation, the American Library Association, Andrew R. Mellon Foundation, Amazon.com, Vivisimo. Com, Infospace.com, NEC, IBM, Excite.com, AlltheWeb.com, AltaVista.com, FAST, and Lockheed Martin have sponsored her research. In 2008 Professor Spink had the second highest H-index citation score in her field from 1998 to 2008 [Norris, M. (2008)]. Ranking Fellow Scholars and their H-Index: Preliminary Survey Results. Loughborough University, Dept of Information Science Report].

# **Integration of Metadata in BWmeta-2.0.0 Format**

Katarzyna Zamlynska<sup>†</sup>, Jakub Jurkiewicz<sup>‡</sup> and Lukasz Bolikowski

University of Warsaw, Interdisciplinary Centre for Mathematical and Computational Modelling, Warsaw, Poland

† kzaml(at)icm.edu.pl

† kura(at)icm.edu.pl

**Abstract:** This article describes the BWmeta-2.0.0 format in terms of its use in an integrated digital library.

**Keywords:** Mathematical digital libraries, Metadata indexing format, YADDA

#### I. INTRODUCTION

YADDA [1] is a document repository framework for digital publication of scientific works. It is built on an open architecture and therefore is flexible and easily adapts to the particular needs of publishers. It can aggregate contents from a diverse range of sources of varying data type, not only publications, but also programs, experimental results, movies etc. At present, in YADDA repository [2], 5 different collections are available. Those collections cover a wide range of disciplines from humanities to science to mathematics, containing both books and journals, with different access regimes. The variety of content to be held in the YADDA repository is on the one hand very attractive for users, while on the other hand challenging for programmers [3]. Proper description of the content, clear and detailed, is not an easy task. To obtain the required level of flexibility the new metadata format, called BWmeta, was implemented in the YADDA platform. The latest release of BWmeta is BWmeta-2.0.0, which has a range of new features, which further improve our digital library.

#### II. HIERARCHY

The Polish Digital Mathematical Library (PL DML) [4] contains both books and journals. In one of the most popular metadata format for scientific publications tag libraries for books and journals are separated. In BWmeta, description of an object does not depend on its type. Chapters of books and journal articles are all described with the same set of elements. The hierarchy of documents is not strictly defined. It can be:

- publisher book series volume book chapter
- publisher journal volume issue article
  - series conference article

but also any other suitable to the situation and particular needs of a contributor, like: institution – Ph.D. thesis. Another flexibility is the ability to add detailed metadata to every level of hierarchy. It is very useful in journals like Banach Center Publications from the PL DML

collection, in which every volume has its own title and editors.

<structure hierarchy="BWmeta1.hierarchyclass.hierarchy\_Journal"> <level="BWmetal.level.hierarchy\_Journal\_Article"</pre> position="169-178"/> <ancestor level="BWmetal.level.hierarchy\_Journal\_Journal"> <name lang="eng">Banach Center Publication</name> <contributor institution="false" role="editor"> <name>Andrzej Hulanicki</name> </contributor> </ancestor> <ancestor level="BWmeta1.level.hierarchy\_Journal\_Volume"> <name>40</name> <name lang="eng" type="subtitle">Mathematics of *Gravitation*</name> </nacestor> <ancestor level="BWmeta1.level.hierarchy\_Journal\_Issue"> <name>1</name> <name lang="eng" type="subtitle">Lorentzian Geometry and Einstein Equations</name <contributor institution="false" role="editor"> <name>Piotr Chru'sciel</name> </contributor> </structure> ... <name lang="eng">A stable class of spacetimes with naked singularities</name> <contributor institution="false" role="author"> <name type="surname">Kriele</name> </contributor>

This feature is also helpful in multi-volume books like Ouevres of Stefan Banach, in which every volume has its own editors, title and every chapter can have different authors and date of publications.

Each object can belong to more than one hierarchy class. To describe the place in hierarchy one can use the reference to higher object. References can also be done in other elements like contributor or affiliation. It is enough to add the attribute id, which allows the element content to be referenced. For example:

<affiliation id="ICM\_UW"> <text>Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw</text> </affiliation>

The reference to the "ICM UW" can be done later by the element:

<affiliation-ref ref="ICM\_UW"/>

In BWmeta-2.0.0 an identifier is not an XML reference. The definition of the reference element does not

have to be declared in the same XML file. Moreover, during the import to the system the identifiers don't have to be available in the import package, like it was required in BWmeta-1.2.0.

#### III. DIVERSITY OF METADATA

Once the metadata scheme is defined, adding new tags to it creates problems. Instead of trying to predict all possible types of data, we decided to make the tag system as easy as possible. The detailed description of elements can be done by adding parameters. In the PL DML library, few types of description exist. All of them are marked by <description>. To distinguish between an abstract written by author or a comment add by the editor, it is enough to add the attribute type:

```
<description lang="eng" type="abstract">
```

where type can be: abstract, summary, comment, note. If it is needed a new type can be added, for example the table of content in a book, without redefining metadata or changing software presentation. In BWmeta-2.0.0 elements are not strictly defined, they can describe different types of data. For example: the original title and it's translation are described by the same element <name>, and distinctions between them can be made by attribute type:

```
<name lang="eng">Quantum Mechanics</name>
<name lang="pl" type="translated">Mechanika
Kwantowa</name>
```

The attribute type doesn't have to be included. Most of the XML elements can be expanded with information which doesn't suit the schema. It can be done with the help of the element attribute. For example with citations, instead of creating a number of new tags, adding to the element <attribute> parameter key, we are able to keep in the same place the unstructured an structured citation forms.

```
<relation type="reference">
<attribute key="text">
<value>1. Winfree, A. T. The Geometry of Biological
(Springer, New York, 1980)."</value>
</attribute>
<attribute key="parsed">
<attribute
key="type"><value>book</value></attribute>
key="position"><value>1</value></attribute>
<attribute key="authors">
<attribute key="author">
<value>A. T. Winfree</value>
< attribute key = "forenames" > < value > A.
T.</value></attribute>
key="surname"><value>Winfree</value></attribute>
</attribute>
```

```
</attribute>
<attribute key="title">
<value>The Geometry of Biological Time</value>
</attribute> <attribute
key="publisher"><value>Springer</value></attribute>
>
<attribute key="city"><value>New
York</value></attribute>
<attribute
key="year"><value>1980</value></attribute>
</attribute>
</attribute>
</attribute> </attribute>
</attribute> </attribute> </attribute>
```

#### IV. EMEBED TEXT FORMULA

In BWmeta-2.0.0 we decided that we don't want to limit text fields to any specific XML. We would like to allow adding to text: XHTML, MathML and probably some formats we don't know about. We use for text fields XSD type: xsd:any with namespace="##other". This allows us to easily add both XHTML and MathML. Below is example of HTML:

```
<name lang="eng">
Bcl-2-dependent modulation
of Ca<sup
xmlns="http://www.w3.org/1999/xhtml">2+</sup>
homeostasis and store-operated channels in
prostate cancer cells </name>
```

Thanks to this solution we get a nice view of calcium ion: Ca<sup>2+</sup>. Using same construction we input MathML:

What is important here is that we don't put MathML to CDATA section. This means that during the validation process of BWmeta XML documents, MathML, which is inside, is also validated.

#### V. CONCLUSIONS

The flexibility of BWmeta-2.0.0 makes it a more appropriate format for integrated libraries. It allows for easy addition of new attributes, descriptions, and names of elements, without the need to change the format of the metadata definition, presentation software and storage software. This flexibility is extremely important

during migration processes from one library system to another without loosing the elements which are not present in the new system.

#### **ACKNOWLEDGEMENTS**

This work is partially supported by the National Centre for Research and Development (NCBiR) under Grant No. SP/I/1/77065/10 by the Strategic scientific research and experimental development program: SYNAT - "Interdisciplinary System for Interactive Scientific and Scientific-Technical Information"

#### **REFERENCES**

#### YADDA

http://yaddainfo.icm.edu.pl/?page%20id=46

The Virtual Library of Science http://yadda.icm.edu.pl
Katarzyna Zam lynska, Lukasz Bolikowski, Tomasz
Rosiek Migration of the Mathematical Collection
of Polish Virtual Library of Science to the
YADDA platform. DML 2008 Towards Digital
Mathematics Library, Birmingham, UK, July 27th,
2008 Proceedings, Masaryk University, r. 2008

The Polish Digital Mathematical Library

http://pldml.icm.edu.pl/mathbwn/search/article.action