

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
2nd AMICUS Workshop

To learn more about I-DAS, including the Book Series, please visit the webpage
<http://www.i-das.org/>

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 (<http://creativecommons.org/licenses/by-nc/3.0/>), 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	6
Prof. Christos Skourlas	
Towards the Preservation and Availability of Historical Books and Manuscripts: A Case Study	8
Eleni Galiotou	
An Extensive Experimental Study on the Cluster-based Reference set Reduction for Speeding-up the k-nn Classifier	12
Stefanos Ougiaroglou, Georgios Evangelidis and Dimitris A. Dervos	
Exploiting the Search Culture Modulated by the Documentation Retrieval Applications	16
Nikitas N. Karanikolas and Christos Skourlas	
Information and Knowledge Organization: The Case of the TEI of Athens	22
Anastasios Tsolakidis, Manolis Chalaris and Ioannis Chalaris	
Providing Access to Students with Disabilities and Learning Difficulties in Higher Education through a Secure Wireless framework	26
Catherine Marinagi and Christos Skourlas	
Improving Query Efficiency in High Dimensional Point Indexes	30
Evangelos Outsios and Georgios Evangelidis	
Text Segmentation Using Named Entity Recognition and co-Reference Resolution in Greek Texts	34
Pavlina Fragkou	
KINISIS, a Graphical XQuery Language	42
Euclid Keramopoulos, Achilleas Pliakas, Konstantinos Tsekos and Ignatios Deligiannis	
Dimensionality Curse, Concentration Phenomenon and the KDB-tree	46
Nikolaos Kouiroukidis and Georgios Evangelidi	

Applying Balanced Scorecard Strategic Management in Higher Education	50
Manolis Chalaris, Anastasios Tsolakidis and Ioannis Chalaris	
A Web Portal Model for NGOs' Knowledge Management	54
Zuhal Tanrikulu	
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	66
Lizette de Jager and Albertus AK Buitendag and Potjie (JS) van der Walt	
4TH SYMPOSIUM ON BUSINESS AND MANAGEMENT AND DYNAMIC SIMULATION MODELS SUPPORTING MANAGEMENT STRATEGIES	71
Dr. Damianos Sakas	
New Political Communication Practices: No Budget Events Management. The New Challenge	73
Evangelia N. Markaki, Damianios P. Sakas and Theodore Chadjipantelis	
Free Software – Open Source Software. A Powerful Tool for Developing Creativity in the Hands of the Student	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	85
Anastasia J. Chournazidis	
Managing Scientific Journals: A Cultural Viewpoint	87
Marina C. Terzi, Damianos P. Sakas, and Ioannis Seimenis	
A Conceptual Framework for Analyzing Knowledge-based Entrepreneurship	92
Nikos S. Kanellos	

SESSION ON INFORMATION HISTORY: PERSPECTIVES, METHODS AND CURRENT TOPICS	96
Prof. Laszlo Karvalics	
Emerging Research Fields in Information History	98
Laszlo Z. Karvalics	
Information Management through Elementary Data Clusters: New Observations on Pridianum-Type Roman Statistical Documents	102
Gergő Gellérfi	
Information and Secrecy on the Silk Road. Methods of Encryption of Legal Documents in Inner Asia (3th-4th century)	106
Szabolcs Felföldi	
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	114
Zsolt Hunyadi	
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	123
Marton Holczer	
SYMPOSIUM ON INTEGRATED INFORMATION: THEORY, POLICIES, TOOLS	126
Prof. Georgios Giannakopoulos	
Approaching Information as an Integrated Field: Educating Information Professionals	128
Georgios Giannakopoulos, Daphne Kyriaki Manesi and Strydon Zervos	
Special Libraries as Knowledge Management Centers	132
Eva Semertzaki	
Digital Libraries' Developers and their Suitability: A Case Study	136
Maria Monopoli	

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	150
Aikaterini Yiannoukakou	
School Archives and their Potentials in Teaching: Aspects of Greek Reality	156
Sonia Geladaki and Panagiota Papadimitriou	
Research on School Libraries in Greece and Suggestions on its Further Development	160
Georgios D. Bikos	
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	172
Keita Tsuji, Erika Kuroo, Sho Sato, Ui Ikeuchi, Atsushi Ikeuchi, Fuyuki Yoshikane and Hiroshi Itsumura	
Developing a National Database on Librarianship and Information Science. The Case of E-VIVA, the Hellenic Fulltext Database	176
Filippos Ch. Tsimpoglou, Vasiliki V. Koukounidou and Eleni K. Sakka	
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	181
Ms. Mitra Zarei and Ms. Maliheh Farrokhnia	
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 194

Hara Brindesi and Sarantos Kapidakis

Aggregating Metadata for Europeana: The Greek Paradigm 198

Alexandros Koulouris, Vangelis Banos and Emmanouel Garoufallou

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 209

Michail Agathos and Sarantos Kapidakis

Integration of Metadata in BWMETA-2.0.0 Format 213

Katarzyna Zamlynska, Jakub Jurkiewicz and Lukasz Bolikowski

SESSION ON EVIDENCE-BASED INFORMATION IN CLINICAL PRACTICE 216

Dr. Evangelia Lappa

Applicability of Data Mining Algorithms on Clinical Datasets 218

Wilfred, Bonney

Changing Roles of Health Librarians with Open Access Repositories 221

Christine Urquhar and Assimina Vlachaki

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

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

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)	314
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	343
Panagiotis Reklitis and Anestis Fotiadis	
Strategic Alignment of ERP, CRM and E-business: A Value Creation	347
Catherine C. Marinagi and Christos K. Akrivos	
The Impact of Occupational Stress on Performance in Health Care	351
Panagiotis Trivellas Panagiotis Reklitis and Charalambos Platis	

The Impact of Emotional Intelligence on Job Outcomes and Turnover Intention in Health Care	356
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	365
Anastasia Argyropoulou, Panagiota Dionysopoulou, Georgios Miaoulis	
G.N.T.O. (Greek National Tourism Organization) Communication Strategy in Advertising Campaigns 1991-2006	370
George Stafylakis and Panagiota Dionysopoulou	
GENERAL PAPERS	375
The role of Environmental Education within the Framework of the Environmental Policy of a Regional Municipality	376
Vassiliki Delitheou and Dimitra Thanasia	
Issues of Social Cohesion: A case study from the Greek Urban Scenery	380
Evgenia Tousi	
Merging Activity and Employee Performance: The Greek Banking System	384
Panagiotis Liargovas and Spyridon Repousis	
Sustainable Development and Corporate Social Responsibility in Higher Education: Some Evidence from Greece	387
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 Repositories: 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. Dendrinis, Technological Institute of Athens
Milena Dobрева, 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].

Aggregating Metadata for Europeana: The Greek paradigm

Alexandros Koulouris[†], Vangelis Banos[‡] and Emmanouel Garoufallou^{‡†}

[†] *Technological Educational Institute of Athens. Department of Library Science and Information Systems, 12210, Egaleo, Greece.
akoul (at) teiath.gr*

[‡] *Aristotle University of Thessaloniki. Department of Informatics, 54124, Thessaloniki, Greece.
vbanos (at) gmail.com*

^{‡†} *Technological Educational Institute of Thessaloniki, Department of Library Science and Information Systems, P.O. BOX 141, 57400, Thessaloniki, Greece.
mgarou (at) libd.teithe.gr*

Abstract: *Europeana is an evolving service that tries to be a single access point for Europe's cultural heritage. The metadata aggregation for Europeana is a procedure that needs interoperability and software tools for transforming metadata to Europeana Semantic Elements (ESE). This aggregation imposes requirements difficult to be implemented in many small institutions. To accommodate these needs and help the Greek institutions contribute their content to Europeana, we developed a series of tools. The Hellenic Aggregator that has established a single communication point with Europeana. The Open Archives Engine (OAE) software for creating digital library aggregators. The oaipmh.com tool for validating OAI-PMH enabled repositories. Other wrappers that enable repeatable generation and harvesting of ESE-compatible metadata via OAI-PMH. We present these tools and the way that helped the Greek cultural institutions in contributing content to Europeana.*

Keywords: *Europeana, EuropeanaLocal, OAI-PMH, Metadata harvesting, Europeana semantic elements.*

I. INTRODUCTION

Repositories and digital libraries are distributed among European Countries. Various formats, different content types, multiple metadata schemes are used. This knowledge either in cultural or in science sector should be accessible to Europeana citizens for awareness and dissemination. From this need various aggregation schemes have derived. *Europeana* (the European Digital Library) is an evolving service that tries to be a single access point for Europe's cultural heritage. According to recent researches, Europeana seems to be service of a vital importance for European cultural awareness. This may be certified by recent surveys (IRN Research, 2011). In this context and in order these content to be ingested various tools have been developed.

This paper analyses the Greek paradigm and the toolset that has been developed for data providers.

II. EUROPEANA AND EUROPEANALOCAL: WHERE WE STAND TODAY

Europeana service (Koninklijke Bibliotheek, 2009) is designed to increase access to digital content across Europe's cultural organizations (i.e. libraries, museums, archives and audio/visual archives). In order to achieve these goals *European Union* (EU) launched various projects. One of the most fruitful was EuropeanaLocal, which ran from 1 June 2008 to 31 May 2011. It had 32 partners from 27 countries, 1031 plus person months and €4.3 million budget. Up to 21 July 2011, EuropeanaLocal partners made available to Europeana live service 4.984.952 items, and 160.000 items from Region Marche ingested via *CulturaItalia* (2006), make o total of 5.144.952 items. More than 20 million items, held across 27 countries. Finally, another 1.075.791 items are on the way for harvesting.

In conclusion, EuropeanaLocal project, till the 4th of July 2011, contributed 26% of total current Europeana content. It had a great impact on Europeana strategy and awareness, documentation and guidelines, workflows and on tools and support. Technical and interoperability challenges were solved and the European aggregations infrastructure was enhanced. However, long term systemic problems (e.g. finance, staff availability that is qualified), remained. Finally, more work has to be done in ingesting more local archival content from municipal/regional archives, church councils, local history associations, etc (Rowlatt *et al.*, 2011).

III. THE GREEK CASE: SOLVING TECHNICAL ISSUES AND DEVELOPING SOFTWARE TOOLS

Greece is participating in EuropeanaLocal with content providers and the Hellenic Aggregator created and supported by the Veria Central Public Library (VCPL). One of the most important aspects in the process of creating a Europeana Compliant digital repository is the support for ESE. Existing digital repository software in general does not support ESE by default. The DSpace plugin for Europeana Semantic Elements webpage (Banos, 2010), developed by the VCPL and the Hellenic National Documentation Centre (EKT), provides specific information about the process (Houssos *et al.*, 2011).

DEiXto, Hellenic Aggregator, Open Archives Engine and oaimh.com are tools created to help organizations participate in aggregation schemes such as Europeana, even if they do not use open digital repository platforms, but closed source technologies or legacy software, which do not support OAI-PMH or any other form of automatic metadata exchange.

IV. THE HELLENIC AGGREGATOR FOR EUROPEANA

Until August 2011, thirteen Greek cultural organizations are already participating in EuropeanaLocal (Koulouris *et al.*, 2010) and are making their metadata available to the Europeana service. These organizations are:

#	Organization	Records
1	Pandektis - National Documentation Center of Greece (http://pandektis.ekt.gr/)	38.880
2	Medusa - Veria Central Public Library (http://medusa.libver.gr/)	1.547
3	The Historical Archives of the American Farm School of Thessaloniki (http://ouranos.afs.edu.gr/dspace)	712
4	Technical Chamber of Greece Regional Department of Corfu (http://lib.teekerk.gr/)	141
5	Central Library of NTUA (http://dspace.lib.ntua.gr/)	3.103
6	Music Library - Lilian Voudouri (http://digma.mmb.org.gr/)	2.796
7	Corgialenios Digital Library (http://www.corgialenios.gr/library/)	7.053
8	University of Athens – Pergamos digital library (http://pergamos.lib.uoa.gr/)	74.494
9	Hellenic Ministry of Education – Educational Television (http://www.edutv.gr/)	661
10	Anatolia College - Digital Archives & Special Collections (http://www.anatolia.edu.gr/digitalarchives)	447
11	Technical Chamber of Greece Library (http://library.tee.gr)	5.783
12	Serres Central Public Library (http://ebooks.serrelib.gr)	464
13	Levadia Central Public Library (http://ebooks.liblivadia.gr)	142
	Total	136.223

Table 1. Greek cultural organizations that participate in Europeana.

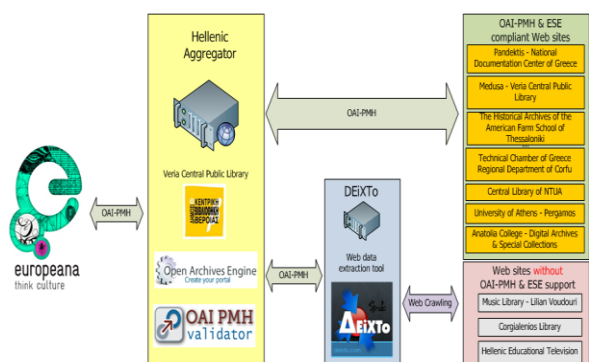


Figure 1. The Hellenic Aggregator Architecture.

Essentially, becoming part of Europeana means that Europeana is able to retrieve specific information from a digital library. One way of doing such a thing would be to connect with each digital library using the appropriate network communication protocol and retrieve the following data:

- Metadata (descriptive, administrative) describing a digital object. The metadata must be mapped to the ESE v3.4 (Europeana v1.0, 2011),
- A preview or thumbnail of the described object,
- Persistent identifiers - active and stable links to the described digital object on the provider's site or the portal's site.

Obviously, the amount and type of content, the technical infrastructure, the output formats and the documentation available can vary significantly among all these content providers. It is, therefore, practically impossible for Europeana to work individually with every content provider due to the enormous amount of work that the harmonization and normalization of metadata would require. As a result, Europeana works with an intermediate layer of content providers, the aggregators.

The Hellenic Aggregator (Veria Central Public Library, 2010) functions as an intermediary on a national level, retrieving data from each participating organization and creating a single communication point with Europeana. What is more, the Hellenic Aggregator's role within Europeana is not confined to submitting metadata. Aggregators also play a key role in other fields:

- Disseminating the vision and objectives of Europeana to their network of institutions in order to increase support for and involvement with Europeana.
- Providing valuable feedback about the issues and discussions from their field.
- Promoting and implementing standards further along the content provision chain.
- Providing domain specific expertise and skills to institutions and Europeana.

A. Registering the Hellenic Aggregator

The process of registering a new digital library to the Hellenic Aggregator is described below:

1. Initially, the digital library web site is examined by an expert who concludes whether it contains content suitable for Europeana.
2. If the digital library supports OAI-PMH, an expert from the EuropeanaLocal Group conducts metadata tests, identifies potential problems and suggests possible solutions.
3. If the digital library does not support OAI-PMH, DEiXto software is used to harvest the required metadata from the target HTML pages.

4. As soon as the digital library's metadata comply with the Europeana standards, it is registered in the Hellenic Aggregator.
5. Content Provider Agreement is signed by the digital library director.
6. The digital library content is published in Europeana.

B. Software Tools

The software infrastructure of the Hellenic Aggregator consists of three different tools which function complementarily in order to implement the full lifecycle of digital library validation, data extraction, storage and communication with Europeana. The core of metadata harvesting, storage and communication with Europeana is implemented by Open Archives Engine while specific data extraction tasks are handled by DEiXTo software. Last but not least, OAI-PMH protocol support and standards compliance for all partners is evaluated using OAIPMH.com.

V. OPEN ARCHIVES ENGINE

Open Archives Engine (OAE) (Banos, 2009) software can be used to create a metadata aggregator and search portal using OAI-PMH enabled, web accessible digital repositories. OAE utilizes the OAI-PMH protocol in order to retrieve metadata from multiple digital libraries and create an index, which then can be used not only to search and filter information but also to export information in a variety of formats such as OAI-PMH Dublin Core (DC) and ESE. Additionally, OAE leverages the technology of DEiXTo in order to extract metadata from legacy digital libraries.

The main components of OAE are the metadata harvester and the web interface.

A. OAE Metadata Harvester

The OAE Metadata Harvester is responsible for connecting with digital libraries and extracting their metadata. After retrieving the metadata from a content provider the one way or the other, the OAE software applies filtering and normalization techniques in order to prevent errors and increase the quality of the metadata.

- XML document encoding and structure is checked using the HTML Tidy (Ragget, 2008) library and a number of errors such as adding missing tags or removing and resolving inappropriate XML characters are resolved.
- Validation against ESE & Dublin Core XML Schemas (DCMI, 2011) is performed.
- Validation for invalid metadata values such as invalid URLs, dates or missing fields is performed.
- Special library-specific bug fixes are applied.

Finally, system indexes are updated and the Hellenic aggregator is ready to publish the aggregated data.

B. OAE Web Service

OAE web interface is responsible not only for creating a web portal from which users are able to search, browse and view metadata records and navigate to the original archives, but also for outputting metadata in a variety of formats such as Dublin Core and ESE, as well as OpenSearch (A9.com, Inc, 2011) and JSON (1999). Currently, Europeana communicates with the Hellenic Aggregator using the OAI-PMH interface.

VI. DEIXTO

Digital libraries developed in the past did not support up to date technologies such as OAI-PMH or any kind of metadata extraction using web services. As a result, the inclusion of such libraries in the Hellenic Aggregator and the Europeana would be impossible without the use of advanced data extraction tools such as DEiXTo.

DEiXTo (Donas, 2010) is a powerful web data extraction tool that is based on the W3C Document Object Model (DOM) (W3C, 2005). It allows users to create highly accurate "extraction rules" (wrappers) that scribe what pieces of data to scrape from a website. DEiXTo can contend with a wide range of websites with high precision and recall. It provides the user with an arsenal of features aiming at the construction of well-engineered extraction rules. Wrappers built with GUI DEiXTo can be scheduled to run automatically providing automated access to resources of interest and saving users a lot of time, energy and repetitive effort.

DEiXTo extracts data and stores it in various formats. In order to facilitate data extraction for the Hellenic Aggregator, a special output plugin has been developed for DEiXTo in order to generate ESE XML files.

VII. OAIPMH.COM

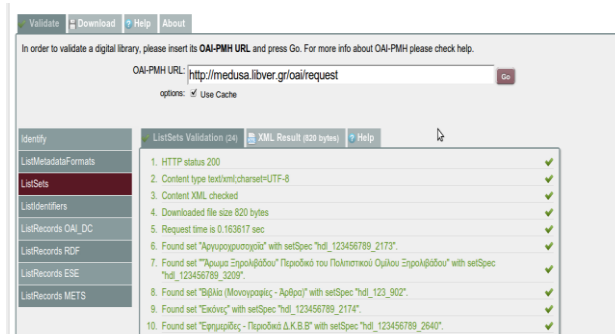
The process of validating an OAI-PMH enabled digital library is quite complex and may become tedious when dealing with a large number of digital libraries. In order to alleviate the task of validating a large number of libraries contributing content to Europeana, a novel online tool has been developed. OAIPMH.com is a web application capable of performing all the necessary checks required to ensure that an OAI-PMH enabled digital library is ready to be part of Europeana.

OAIPMH.com (Banos, 2011) consists of a server-side application running in the background and a modern web interface running on the web. Users can either validate a digital library or download all the records from one or more digital libraries in their computers.

- **Validation:** The validation of an OAI-PMH enabled digital library requires only the submission of the OAI-PMH web service URL. The application issues all OAI-PMH commands to the digital library and evaluates the output according to XML Schemas of DC and ESE as well as the Europeana Guidelines.

- **Metadata extraction:** Users can input a list of OAI-PMH URLs and retrieve all the metadata records which are available from them in parallel. Using this feature, users can retrieve a large number of metadata records from multiple libraries rapidly and easily, thus enabling them to inspect them and evaluate them.

OAI-PMH.com has improved the process of validating new and existing OAI-PMH enabled libraries. Anyone concerned can evaluate digital libraries using a quick and intuitive tool.



VIII. CONCLUSIONS AND FUTURE DEVELOPMENTS

The Greek Cultural Heritage Institutions implemented interoperable digital libraries, digitized valuable historic content, populate, and preserve it through Europeana service. They enriched their metadata quality, familiarized their staff with digital skills and developed new services for their users. On the other hand, the EuropeanaLocal Greek team developed software tools, which facilitated the Geek CHI to implement interoperable digital collections by reducing the cost and the human effort. Automated harvesting procedures were established.

The Geek Cultural Heritage Institutions have to take advantage of the knowledge that achieved during EuropeanaLocal, to implement repositories and to contribute more content to Europeana service. These synergetic schemes, especially in this economic crisis, enhance the viability of the Greek Cultural Heritage Institutions and content. In this context, European Commission should support similar projects, because this is of great importance for the European cultural heritage.

Finally, as the most important future obstacle and because VCPL has the aggregation obligation only through EuropeanaLocal, a national Hellenic aggregator that will continuously support the content providers, share knowledge and expertise between partners should be established.

REFERENCES

A9.com, Inc., *OpenSearch protocol*, <http://www.opensearch.org/> (2011).

Banos, V., *DSpace plugin for Europeana Semantic Elements (ESE)*, <http://vbanos.gr?p=189> (2010).

Banos, V., *Open archives engine software*, <http://openarchivesengine.com> (2009).

Banos, V., *Open archives initiative protocol for metadata harvesting validation & data extraction tool*, <http://oai-pmh.com> (2011).

CulturaItalia, *CulturaItalia*, <http://www.culturaitalia.it/Language/LanguageGateway?lang=en&T=1312795936520> (2006).

DCMI, *XML schemas to support the guidelines for implementing Dublin Core in XML*, <http://www.dublincore.org/schemas/xmls/> (2011).

Donas, K., *DEiXTo*, <http://www.deixto.com> (2010).

Europeana v1.0, *Europeana semantic elements specifications. Version 3.4, 11/08/2011*, http://www.version1.europeana.eu/c/document_library/get_file?uuid=77376831-67cf-4eff-a7a2-7718388eecl&groupId=10128 (2011).

Europeana, *Europeana content checker user guide*, http://version1.europeana.eu/c/document_library/get_file?uuid=d9ca0106-affb-4a38-83e9-e886289dd0d9&groupId=10602 (2010).

EuropeanaLocal, *EuropeanaLocal*, <http://www.europeanalocal.eu/> (2008).

Houssos, N., Stamatis, K., Banos, V., Kapidakis, S., Garoufallou, E. and Koulouris, A., "Implementing enhanced OAI-PMH requirements for Europeana", *Proc. International Conference on Theory and Practice of Digital Libraries (TPDL 2011)*, Berlin, Germany, September 25-29, 2011, and *Lectures Notes in Computer Science (LNCS)* (2011).

IRN Research, *Europeana – online visitor survey: research report*, version 3, 23rd June 2011, https://version1.europeana.eu/c/document_library/get_file?uuid=334beac7-7fc2-4a4e-ba23-4dcc1450382d&groupId=10602 (2011).

Javascript Object Notation (JSON), <http://www.json.org/> (1999).

Koninklijke Bibliotheek, *Europeana*, <http://www.europeana.eu> (2009).

Koulouris, A., Garoufallou, E. and Banos, E., "Automated metadata harvesting among Greek repositories in the framework of EuropeanaLocal: dealing with interoperability", *Proc. 2nd Qualitative and Quantitative Methods in Libraries and International Conference (QOML2010)*, Chania, Greece, 331-338 (2010).

Ragget, D., *HTML tidy library project*, <http://tidy.sourceforge.net> (2008).

Rowlatt, M., Davies, R. and Komen, L., *EuropeanaLocal: it's objectives, activities and impact. Project presentation: results D1.11*, <http://www.europeanalocal.eu/eng/Document-Library/Project-Deliverables> (2011).

Veria Central Public Library, *Europeana Local Aggregator*, <http://aggregator.libver.gr> (2010).

W3C, *Document Object Model (DOM)*, <http://www.w3.org/DOM/> (2005).

Integrating a Repository with Research Output and Publications: The case of the National Technical University of Athens

Dionysis Kokkinos

National Technical University of Athens. Central Library, 9 Heroon Polytechniou str., Zografou Campus, 15780, Athens, Greece.

dennis(at)central.ntua.gr

Abstract: *The paper describes the progress of the main Institutional Repository (IR) of the National Technical University of Athens (NTUA) and its current development. The new challenges presented are the enhancement of the repository with new services and collections. One of the major collections which is under development is the University's research output. We have already started populating the collection with papers and presentations published by the members of NTUA's academic community. Our first step was to establish new self-archiving routines for the academic staff and faculty members. This included new forms and directions, to support the deposit of their publications. The essential goal of the IR integration is to gather all of NTUA's scientific and technical content and simultaneously to provide new, personalized and advanced services.*

Keywords: *Institutional repositories, Integration, Scholarly publishing, Open access, Digital content.*

I. INTRODUCTION

The National Technical University of Athens (NTUA) is the oldest and most prestigious educational institution of Greece in the field of technology, and has contributed unceasingly to the country's scientific, technical and economic development since its foundation in 1836. Recently (scoring results 2010), NTUA achieved to rank 17th among the European Research Institutions according to the European Research Ranking (2011). The NTUA personnel include more than 700 professors and scientific assistants and there are many researchers as well. The total number of undergraduate students is about 13000 and the post-graduate students are 4000. Professors, researchers, students and staff consists a vital and productive community participating in numerous valuable projects and publications. A recent survey on the number of research projects sponsored by the European Commission classifies NTUA in the 2nd place among all European universities and in the 4th place among all European organizations (NTUA, 2011). The amount and the high standards of this research are proved by the numerous publications in scientific journals and proceedings of international conferences. All these publications need to be collected and disseminated to broader scientific communities. The concept is to exploit all the publications and the other research output. The ideal way to exploit these publications is through the IR service. The whole idea is described in detail in a related Conference paper (Kouis

and Kokkinos, 2009). More specifically, the suggested model of a digital library with technical digital content is illustrated with regard to aggregating and displaying the research papers or data produced at NTUA.

The Central Library of NTUA started operating in 1914. It serves the NTUA teaching staff, assistants, researchers, administrative staff and students, as well as scientists and students from other universities and institutions. It offers a variety of services either locally, or remotely. There are 15.000 registered Library users (students, staff and faculty). The Central Library has established a new Open Access Institutional Repository (IR) policy on July 2007. The IR contains *Electronic Theses and Dissertations* (ETDs) combined with a submission tool/service and a growing collection of scientific publications (NTUA Central Library, 2011). It is based on a DSpace installation (DSpace, 2011), using *Dublin Core* (DC) descriptive metadata and *Open Access Initiative Protocol for Metadata Harvesting* (OAI-PMH) for interoperability and inter-connectivity reasons. At the time of writing this paper (September 2011), there are approximately 3000 ETDs stored in the repository, where the majority are diploma theses, providing worldwide access to full text in PDF format. The NTUA IR, as part of the NTUA Digital Library, is participating in the *European Digital Library* (Europeana, 2011) contributing cultural digital content, that gave the opportunity to make local digital content available to all the citizens in Europe. Despite the fact that the majority of our digital items are in Greek, the cultural and scientific digital content is now accessible to a wide European community. It is remarkable that Europeana's collection reaches 15 million items and 1500 institutions have contributed to Europeana.

The NTUA IR is indexed in the *Directory of Open Access Repositories* (DOAR) and in the *Registry of Open Access Repositories* (ROAR); additionally the NTUA Digital Library, including the IR, is indexed in the *Greek Digital Resources Index*. It is useful to point out that, through the NTUA's Digital Library is also provided access to unique digitized collections. There are two separate collections, the first with digitized rare books (Greek Scientific and Technical Bibliography: 1830-1940) in jpeg format (one image per page) and the second one with digitized rare journals in PDF format (1 PDF file per article). The digitized material (monographs and journals) belongs to the Historical Library of NTUA and their subjects cover, among others, the fields of technology, history of sciences and