

## **Contents**

Editorial Preface- Youakim Badr, Richard Chbeir	i
<b>Research</b>	
Context-informed Knowledge Extraction from Document Collections to Support User Navigation- Mario Cataldi, Claudio Schifanella, K. Selcuk Candan	74
Conference knowledge modelling for conference-video-recordings querying and visualization- Maria Sokhn, Elena Mugellini, Omar Abou Khaled, Ahmed Serhrouchni	95
Enhancing Multimedia Data Fragmentation- Richard Chbeir, Dominique Laurent	112
Semantic Classification, Keyword Mining and Search Space Optimization for digital ecosystems- Nikunj Yadav, Yanu Gupta, Manish Kumar, Ratna Sanyal	131

## **Funding, Research Fellowship and Post Doc Positions**

## Editorial Preface

**Youakim Badr<sup>1</sup>, Richard Chbeir<sup>2</sup>**

<sup>1</sup> INSA de Lyon, France

[youakim.badr@insa-lyon.fr](mailto:youakim.badr@insa-lyon.fr)

<sup>2</sup> Bourgogne University, France

[richard.chbeir@u-bourgogne.fr](mailto:richard.chbeir@u-bourgogne.fr)

This special issue includes extended and revised papers from the ACM MEDES'09 Conference series which aims to develop and bring together a diverse community from academia, research laboratories and industry interested in exploring the manifold challenges and issues related to resource management of Digital Ecosystems and how current approaches and technologies can be evolved and adapted to this end. In essence, in the world of the Internet, the rapid growth and exponential use of digital medias leads to the emergence of virtual environments namely digital ecosystems composed of multiple and independent entities such as individuals, organizations, services, software and applications sharing one or several missions and focusing on the interactions and inter-relationships among them. The digital ecosystem exhibits self-organizing environments, thanks to the recombination and evolution of its “digital components”, in which resources provided by each entity are properly conserved, managed and used. The underlying resources mainly comprehend data management, innovative services, computational intelligence and self-organizing platforms.

Due to the multi-disciplinary nature of digital ecosystems and their characteristics, they are highly complex to study and design. This also leads to a poor understanding as to how managing resources will empower digital ecosystems to be innovative and value-creating. The application of Information Technologies has the potential to enable the understanding of how entities request resources and ultimately interact to create benefits and added-values, impacting business practices and knowledge. These technologies can be improved through novel techniques, models and methodologies for fields such as data management, web technologies, networking, security, human-computer interactions, artificial intelligence, e-services and self-organizing systems to support the establishment of digital ecosystems and manage their resources.

After a very tight review process, four original research papers were only accepted for inclusion in this special issue out of twenty-two candidate papers initially submitted. The acceptance rate was thus around 18% and the selected works reflect the high standards for excellence used by the many esteemed members of the reviewing board who contributed to this special issue.

The first paper is entitled “Context-informed Knowledge Extraction from Document Collections to Support User Navigation” and authored by Mario Cataldi, Claudio Schifanella, K. Selçuk Candan, Maria Luisa Sapino, and Luigi Di Caro. In this paper, the authors propose an innovative approach to exploring text collections using a novel keywords-by-concepts (KbC) graph, which supports navigation using domain-specific concepts as well as keywords that are characterizing the text corpus. The KbC graph is a weighted graph, created by tightly integrating keywords extracted from documents and concepts obtained from domain taxonomies. Documents in the corpus are associated to the nodes of the graph based on evidence supporting contextual relevance; thus, the KbC graph supports contextually informed access to these documents. The construction of the KbC graph relies on a spreading-activation like technique which mimics the way the brain links and constructs knowledge. In this paper, the authors also present CoSeNa (Context-based Search and Navigation) system that leverages the KbC model as the basis for document exploration as well as contextually-informed media integration.

In the second paper, Maria Sokhn, Elena Mugellini, Omar Abou Khaled, and Ahmed Serhrouchni present their work titled “Conference knowledge modelling for conference-video-recordings querying

and visualization”. In essence, the evolution of the web in the last decades has created the need for new requirements towards intelligent information retrieval capabilities and advanced user interfaces. Nowadays, effective retrieval and usage of multimedia resources have to deal with the issues of creating efficient indexes, developing retrieval tools and improving user oriented visualization interfaces. To that end, the authors put forward an integrated framework named CALIMERA. The framework is based on a High-level model for cOnference (HELO) and aims at enhancing the information management, retrieval and visualization of recorded talks of scientific conferences. This paper presents the conference model and its uses within the framework: performing high level annotation of scientific talk recordings, offering granular search facilities and complex queries, and enhancing the knowledge visualization of the recordings. As a proof-of-concept, they present the prototypes that have been implemented.

Authored by Richard Chbeir and Dominique Laurent, the third paper is entitled “Enhancing Multimedia Data Fragmentation”. It focuses on the problem of data(base) fragmentation, initially consisting of reducing irrelevant data accesses by grouping data frequently accessed together in dedicated segments, in a multimedia context. Here, the authors mainly address the issue of query and predicate implication required in current fragmentation algorithms, and provide a formal approach to identify such implications, in order to partition multimedia data efficiently. It is worthy to note that the provided approach is capable of considering multimedia-based as well as semantic comparisons, based on a generalized notion of functional dependencies, which are called multimedia functional dependencies.

The last paper of this special issue is written by Nikunj Yadav, Yanu Gupta, Manish Kumar, and Ratna Sanyal, and titled “Semantic Classification, Keyword Mining and Search Space Optimization for digital ecosystems”. Here, the authors present a novel approach to classify the documents in a digital repository and find the semantically significant keywords related to those documents to make the organization and the retrieval of the documents faster and more efficient. They approach this problem using Probabilistic Latent Semantic Analysis with incomplete training data to organize them and mark the relevant keywords. This approach makes the classification faster and instead of the unlabeled clustering gives classification with well defined topics relating to human logic.

We hope this special issue motivates researchers to take the next step beyond building models to implementing, evaluating, comparing, and extend proposed approaches. Many people helped us that this issue becomes a reality. We would first like to gratefully acknowledge and sincerely thank all the reviewers for their timely and insightful valuable comments and criticism of the manuscripts that greatly improved the quality of the final versions. Of course, thanks are due to the authors, who provided excellent articles and timely revisions.