



**4th International Workshop on
Semantic Digital Archives**



**September 12, 2014
London, UK**

Proceedings of the 4th International Workshop on
Semantic Digital Archives
(SDA 2014)

held as part of the
International Digital Libraries Conference (DL2014)
September 12, 2014 in London, UK

<http://sda2014.dke-research.de>

Edited by:

Thomas Risse, L3S Research Center, Hannover, Germany

Livia Predoiu, University of Oxford, United Kingdom

Andreas Nürnberger, Otto-von-Guericke University of Magdeburg, Germany

Seamus Ross, University of Toronto, Canada



Vol-1306
urn:nbn:de:0074-1306-1

Copyright © 2014 for the individual papers by the papers' authors. Copying permitted only for private and academic purposes. This volume is published and copyrighted by its editors.

Preface

The 4th Workshop on Semantic Digital Archives (SDA 2014) has built upon the success of the previous editions in 2011 to 2013 and has been held as part of the International Digital Libraries Conference (DL2014) on September 12, 2014 in London, UK. Organized as full-day workshop, SDA 2014 has aimed to promote and discuss sophisticated knowledge representation and knowledge management solutions specifically designed for improving Archival Information Systems.

Archival Information Systems (AIS) are becoming increasingly important. For decades, the amount of content created digitally is growing and its complete life cycle nowadays tends to remain digital. A selection of this content is expected to be of value for the future and can thus be considered being part of our cultural heritage. However, digital content poses many challenges for long-term or indefinite preservation, e.g. digital publications become increasingly complex by the embedding of different kinds of multimedia, data in arbitrary formats and software. As soon as these digital publications become obsolete, but are still deemed to be of value in the future, they have to be transferred smoothly into appropriate AIS where they need to be kept accessible even through changing technologies.

The successful previous SDA workshops showed: Both, the library and the archiving community have made valuable contributions to the management of huge amounts of knowledge and data. However, both are approaching this topic from different views which shall be brought together to cross-fertilize each other. There are promising combinations of pertinence and provenance models since those are traditionally the prevailing knowledge organization principles of the library and archiving community, respectively. Another scientific discipline providing promising technical solutions for knowledge representation and knowledge management is semantic technologies, which is supported by appropriate W3C recommendations and a large user community. At the forefront of making the semantic web a mature and applicable reality is the linked data initiative, which already has started to be adopted by the library community. It can be expected that using semantic (web) technologies in general and linked data in particular can mature the area of digital archiving as well as technologically tighten the natural bond between digital libraries and digital archives. Semantic representations of contextual knowledge about cultural heritage objects will enhance organization and access of data and knowledge. In order to achieve a comprehensive investigation, the information seeking and document triage behaviors of users (an area also classified under the field of Human Computer Interaction) needs also to be included in the research.

One of the major challenges of digital archiving is how to deal with changing technologies and changing user communities. On the one hand software, hardware and (multimedia) data formats that become obsolete and are not supported anymore still need to be kept accessible. On the other hand changing user communities necessitate technical means to formalize, detect and measure knowledge evolution. Furthermore, digital archival records are usually not deleted from the AIS and therefore, the amount of digitally archived (multimedia) content can be expected to grow rapidly. Therefore, efficient storage management solutions geared to the fact that cultural heritage is not as frequently accessed like up-to-date content residing in a digital library are required. Software and hardware needs to be tightly connected based on sophisticated knowledge representation and management models in

order to face that challenge.

In line with the above, we invited contributions to the workshop that focus on:

- Architectures and Frameworks for semantic AIS and Archival Information Infrastructures (AII)
- Semantic (Web) services implementing AIS & AII
- Contextualization of digital archives, museums and digital libraries
- Linked data for AIS, AII, museums and digital libraries
- Ontologies for AIS, AII, museums and digital libraries
- Semantics of complex content (e.g. Social Media, Multimedia)
- Information integration/semantic ingest (e.g. from digital libraries)
- Semantic search & information retrieval in digital archives, digital museums and digital libraries
- User interfaces for (semantic) AIS, AII, digital museums & semantic digital libraries
- Semantics for Preservation Processes and Protocols
- Preservation of work flow processes
- (Semantic) provenance models
- Semantics for the appraisal and selection of content
- Evolving semantics in long-term archives
- Trust for ingest & data security/integrity check for long-term storage of archival records
- User studies focusing on end-user needs and information seeking behavior of end-users
- Implementations & evaluations of (semantic) AIS, AII, semantic digital museums & semantic digital libraries
- Semantic long-term storage & hardware organization for AIS & AII & digital libraries

We received submissions covering a broad range of relevant topics in the area of semantic digital archives. With the help of our program committee all articles were peer-reviewed. These proceedings comprise all accepted submissions which have been carefully revised and enhanced by the authors according to the reviewers' comments.

We sincerely thank all members of the program committee for supporting us in the reviewing process. Altogether, the diversity of the papers in these proceedings represent a multitude of interesting facets about the exciting and promising research field of semantic digital archives and semantic digital archiving infrastructures. During the workshop itself we had many fruitful and inspiring discussions which would not have been possible without the well done presentations and the interested audience. Many thanks to all workshop attendants for a great workshop!

We would also like to thank Sun SITE Central Europe for hosting these proceedings on <http://ceur-ws.org>.

December 2014

T. Risse, L. Predoiu, A. Nürnberger, and S. Ross

Organizing Committee

- Thomas Risse, L3S Research Center, Hannover, Germany
- Livia Predoiu, University of Oxford, Oxford, UK
- Andreas Nürnberger, Otto-von-Guericke University, Magdeburg, Germany
- Seamus Ross, University of Toronto, Toronto, Canada

Program Committee

- Vassilis Christophides, Foundation of Research & Technology, Hellas, Greece
- Kai Eckert, University Library of Mannheim, Germany
- Alejandra Gonzalez-Beltran, University of Oxford, UK
- Marco Grassi, Università Politecnica delle Marche, Italy
- Tudor Groza, University of Queensland, Australia
- Armin Haller, CSIRO, Australia
- Andreas Harth, KIT, Karlsruhe, Germany
- Steffen Hennicke, Humboldt-Universität zu Berlin, Germany
- Stijn Heymans, SRI International, USA
- Pascal Hitzler, Wright State University, USA
- Christian Keitel, State Archive of Baden-Württemberg, Germany
- Claus-Peter Klas, FernUniversität in Hagen, Germany
- Birger Larsen, Royal School of Library and Information Science, Denmark
- Thomas Lukasiewicz, University of Oxford, UK
- Mathias Lux, Klagenfurt University, Austria
- Maria Vanina Martinez, University of Oxford, UK
- Annett Mitschick, TU Dresden, Germany
- Knud Möller, Talis, Birmingham, UK
- Kai Naumann, State Archive of Baden-Württemberg, Germany
- Gillian Oliver, Victoria University of Wellington, New Zealand
- Jacco van Ossenbruggen, VU University Amsterdam, Netherlands
- Andreas Rauber, Vienna University of Technology, Austria
- Sebastian Rudolph, Karlsruher Institut für Technologie, Germany
- Heiko Schuldt, Universität Basel, Switzerland
- Mike Salampasis, Alexander Techn. Educational Inst. (ATEI) of Thessaloniki, Greece
- Kunal Sengupta, Wright State University, USA
- Gerardo Simari, University of Oxford, UK

- Herbert van de Sompel, Los Alamos National Laboratory Research Library, USA
- Marc Spaniol, Max-Planck-Institut Saarbrücken, Germany
- Manfred Thaller, University of Cologne, Germany

Table of contents

Invited Talk

Understanding Web Archives	1
<i>Helen Hockx-Yu</i>	

Semantics for Interoperability and Identification

Self-contained Information Retention Format for Future Semantic Interoperability	4
<i>Simona Rabinovici-Cohen, Roger Cummings and Sam Fineberg</i>	
Identification Semantics for an Organization Establishing a Digital Library System.....	16
<i>Angela Di Iorio and Marco Schaerf</i>	
Persistent E-mail Identification is Viable!	28
<i>Stefan Haun and Andreas Nürnberger</i>	

Culture meets Semantics

The Digital Online Museum: A new Approach to Experience Virtual Heritage.....	38
<i>Tilman Deuschel, Timm Heuss and Bernhard Humm</i>	
Relations for Reusing (R4R) in a Shared Context: An Exploration on Research Publications and Cultural Objects	49
<i>Andrea Wei-Ching Huang and Tyng-Ruey Chuang</i>	

