TPDL 2017

21st INTERNATIONAL CONFERENCE ON THEORY AND PRACTICE OF DIGITAL LIBRARIES

PROGRAM

18-21 SEPTEMBER

THESSALONIKI GREECE GRAND HOTEL PALACE

RESEARCH COMMITTEE
ARISTOTLE UNIVERSITY OF THESSALONIKI

DEMOCRITUS UNIVERSITY OF THRACE

Lecture Notes in Computer Science
LNCS LNAI LNBI
Springer

cni
Coalition for Networked Information

http://www.tpdl.eu/tpdl2017/

INDEX

PREFACE

The 21st International Conference on Theory and Practice of Digital Libraries (TPDL), takes place in Thessaloniki, Greece from September 18 to 21, 2017. The conference is organized by the Aristotle University of Thessaloniki and the Democritus University of Thrace. The 21st version of the International Conference on Theory and Practice of Digital Libraries holds the general theme "Part of the Machine: turning complex into scalable" and wants to create a dialogue that addresses the challenge of creatively transforming these highly synthesized environments into solutions that can scale for the benefit of varied communities.

TPDL2017 received a number of 85 full paper submissions, up from 50 at TPDL2016 and 44 at TPDL2015, making the conference in 2017 very competitive and selective, and requiring the program committee to uphold the highest possible academic standards.

We introduce a two-layered structure for oral presentations, long and short oral, in order to include an adequate number of interesting papers that expand the field of digital libraries on innovative topics and to strengthen the already known. Of the 85 long paper submissions, only 20 (24%) submissions have been accepted with long oral presentation, and an additional 19 (22%) long papers have been accepted with a shorter oral presentation. This makes a grand total of 39 (46%) full papers accepted for the proceedings.

Of the 8 short paper submissions, only 4 (50%) submissions have been accepted, and of the 5 poster/demo submissions, only 2 (40%) submissions have been accepted. Selected full paper submissions were redirected for evaluation as potential short or poster/demo papers, following the recommendations of the reviewers.

Each submission was reviewed by at least three program committee members, and two senior program committee members, and the two chairs oversaw the reviewing and often extensive follow-up discussion. Where the discussion was not sufficient to make a decision, the paper went through an extra review by the program committee. Each paper was discussed individually, based on the reviews, the metareviews, and the discussion at a PC Meeting, where the final decisions were made.

The conference is honored by three very interesting keynote speeches by Paul Groth on "Machines are people too", Elton Barker on "Back to the future: annotating, collaborating and linking in a digital ecosystem" and Dimitrios Tzovaras on "Visualization in the big data era: data mining from networked information". All three cover important areas of the digital library field.

The program of TPDL2017 also includes a Doctoral Consortium track and four tutorials on "Enriching digital collections using tools for text mining, indexing and visualization", "Putting historical data in context: how to use DSpace-GLAM", "Innovation search" and "Enabling precise identification and citability of dynamic data – recommendations of the RDA working group on data citation". Finally, four workshops are organized in conjunction with the main conference, namely the longestablished "European Networked Knowledge Organization Systems (NKOS)" in its 17th version and the newly introduced "(meta)-data quality workshop", "International Workshop on Temporal Dynamics in Digital Libraries" and "Modeling Societal Future (FUTURITY)".

We would like to thank all our colleagues who trusting their papers in the conference, as well as our Program Committee members, both the Senior and the Regular, for the precise and thorough work they put in reviewing the submissions. A word of gratitude must be addressed to our Workshop Chairs Philipp Mayr and Kjetil Nørvåg, our Tutorial Chairs Thomas Risse and Gianmaria Silvello, our Panel Chair Cristina Ribeiro, our Posters/Demo Chairs Vangelis Banos and Annika Hinze and our Doctoral Consortium Chairs Maja Žumer and Heiko Schuldt, for the substantial effort in running effectively their tracks.

Thessaloniki 2017

Jaap Kamps
Giannis Tsakonas
Yannis Manolopoulos
Lazaros Iliadis

ORGANISATION COMMITTEE

GENERAL CHAIRS

Yannis Manolopoulos, *Aristotle University of Thessaloniki, Greece*
Lazaros Iliadis, *Democritus University of Thrace, Greece*

PROGRAM CHAIRS

Jaap Kamps, *University of Amsterdam, The Netherlands*
Giannis Tsakonas, *University of Patras, Greece*

ORGANISING CHAIR

Apostolos Papadopoulos, *Aristotle University of Thessaloniki, Greece*

PUBLICITY/PUBLICATION CHAIR

Ioannis Karydis, *Ionian University, Greece*

WORKSHOP CHAIRS

Philipp Mayr, *GESIS, Germany*
Kjetil Nørvåg, *Norwegian University of Science and Technology, Norway*

TUTORIAL CHAIRS

Thomas Risse, *L3S, Germany*
Gianmaria Silvello, *University of Padua, Italy*

PANEL CHAIR

Cristina Ribeiro, *University of Porto, Portugal*

POSTERS/DEMO CHAIRS

Vangelis Banos, *Aristotle University of Thessaloniki, Greece*
Annika Hinze, *University of Waikato, New Zealand*

DOCTORAL CONSORTIUM CHAIRS

Maja Žumer, *University of Ljubljana, Slovenia*
Heiko Schuldt, *University of Basel, Switzerland*

PROGRAM COMMITTEE

SENIOR PC MEMBERS

Trond Aalberg, *Norwegian University of Science and Technology, Norway*
David Bainbridge, *University of Waikato, New Zealand*
Tobias Blanke, *University of Glasgow, UK*
Jose Borbinha, *IST / INESC-ID, UK*
George Buchanan, *City University London, UK*
Donatella Castelli, *CNR-ISTI, Italy*
Stavros Christodoulakis, *Technical University of Crete, Greece*
Milena Dobreva, *University of Malta, Malta*
Nicola Ferro, *University of Padua, Italy*
Edward Fox, *Virginia Polytechnic Institute and State University, USA*
Ingo Frommholz, *University of Bedfordshire, UK*
Norbert Fuhr, *University of Duisburg-Essen, Germany*
Richard Furuta, *Texas A&M University, USA*
Marcos Goncalves , *Federal University of Minas Gerais , Brazil*
Annika Hinze, *University of Waikato, New Zealand*
Sarantos Kapidakis, *Ionian Uninversity, Greece*
Laszlo Kovacs, *Hungarian Academy of Sciences, Hungary*
Clifford Lynch, *CNI, USA*
Wolfgang Nejdl, *L3S and University of Hannover, Germany*
Michael Nelson, *Old Dominion University, USA*
Erich Neuhold, *University of Vienna, Austria*
Christos Papatheodorou, *Ionian University, Greece*
Edie Rasmussen, *University of British Columbia, Canada*
Andreas Rauber, *Vienna University of Technology, Austria*
Thomas Risse, *L3S Research Center, Germany*
Laurent Romary, *INRIA & HUB-ISDL, France*
Seamus Ross, *University of Toronto, Canada*
Heiko Schuldt, *University of Basel, Switzerland*
Mário J. Silva, *Universidade de Lisboa, INESC-ID, Portugal*
Hussein Suleman, *University of Cape Town, South Africa*
Costantino Thanos, *ISTI-CNR, Italy*
Herbert Van De Sompel, *Los Alamos National Laboratory, USA*

PROGRAM COMMITTEE

PC MEMBERS

Hamed Alhoori, *Northern Illinois University , USA*
Robert Allen, *Yonsei University, South Korea*
Avishek Anand, *L3S Research Center, Germany*
Vangelis Banos, *Aristotle University of Thessaloniki, Greece*
Valentina Bartalesi, *ISTI-CNR, Italy*
Christoph Becker, *University of Toronto, Canada*
Maria Bielikova, *Slovak University of Technology in Bratislava, Slovakia*
Pável Calado, *INESC-ID, Universidade de Lisboa, Portugal*
Vittore Casarosa, *ISTI-CNR, Italy*
Lillian Cassel, *Villanova University, USA*
Panos Constantopoulos, *Athens University of Economics and Business, Greece*
Fabio Crestani, *University of Lugano, Switzerland*
Sally Jo Cunningham, *Waikato University, New Zealand*
Theodore Dalamagas, *IMIS-"Athena" R.C., Greece*
Makx Dekkers, *Independent Consultant*
Giorgio Maria Di Nunzio, *University of Padua, Italy*
Fabien Duchateau, *Université Claude Bernard Lyon 1 – LIRIS, France*
Maria Economou, *University of Glasgow, UK*
Schubert Foo, *Nanyang Technological University, Singapore*
Nuno Freire, *INESC-ID, Portugal*
Dimitris Gavrilis, *Athena Research Centre, Greece*
Manolis Gergatsoulis, *Ionian University, Greece*
C. Lee Giles, *Pennsylvania State University, USA*
Julio Gonzalo, *UNED, Spain*
Paula Goodale, *University of Sheffield, UK*
Sergiu Gordea, *Austrian Institute of Technology, Austria*
Stefan Gradmann, *KU Leuven, Germany*
Jane Greenberg, *Drexel University, USA*
Mark Michael Hall, *Edge Hill University, UK*
Bernhard Haslhofer, *AIT-Austrian Institute of Technology, Austria*
Frank Hopfgartner, *University of Glasgow, UK*
Nikos Houssos, *RedLink, Greece*
Antoine Isaac, *Europeana Foundation, EU*

PROGRAM COMMITTEE

Adam Jatowt, *Kyoto University, Japan*
Nattiya Kanhabua, *Aalborg University, Denmark*
Roman Kern, *Graz University of Technology, Austria*
Claus-Peter Klas, *GESIS – Leibniz Institute for Social Sciences, Germany*
Martin Klein, *Los Alamos National Laboratory, USA*
Petr Knoth, *KMi, The Open University, UK*
Stefanos Kollias, *National Technical University of Athens, Greece*
Ronald Larsen, *University of Pittsburgh, USA*
Séamus Lawless, *Trinity College Dublin, Ireland*
Hyowon Lee, *Singapore University of Technology and Design, Singapore*
Suzanne Little, *Dublin University, Ireland*
Zinaida Manžuch, *Vilnius University, Lithuania*
Bruno Martins, *IST – Instituto Superior Técnico, Portugal*
Philipp Mayr, *GESIS, Germany*
Cezary Mazurek, *Poznań Supercomputing and Networking Center, Poland*
Robert H. Mcdonald, *Indiana University/Data to Insight Center, USA*
Dana Mckay, *Swinburne University of Technology, Australia*
Andras Micsik, *SZTAKI, Hungary*
David Nichols, *University of Waikato, New Zealand*
Jeppe Nicolaisen, *University of Copenhagen , Denmark*
Ragnar Nordlie, *Oslo and Akershus University College, Norway*
Moira Norrie, *ETH Zurich, Switzerland*
Kjetil Nørvåg, *Norwegian University of Science and Technology, Norway*
Raul Palma, *Poznan Supercomputing and Networking Center, Poland*
Nils Pharo, *Oslo and Akershus University College of Applied Sciences, Norway*
Dimitris Plexousakis, *Institute of Computer Science, FORTH, Greece*
Panayiota Polydoratou, *ATEI of Thessaloniki, Greece*
Cristina Ribeiro, *University of Porto, Portugal*
Ian Ruthven, *University of Strathclyde, UK*
J. Alfredo Sánchez, *UDLAP, Mexico*
Michalis Sfakakis, *Ionian University, Greece*
Gianmaria Silvello, *University of Padua, Italy*
Nicolas Spyratos, *Unirsity of Paris South, France*
Shigeo Sugimoto, *University of Tsukuba, Japan*

PROGRAM COMMITTEE

Tamara Sumner, *University of Colorado at Boulder, USA*
Anastasios Tombros, *Queen Mary University of London, UK*
Theodora Tsikrika, *Information Technologies Institute, CERTH, Greece*
Chrisa Tsinaraki, *Joint Research Center (JRC), European Union*
Douglas Tudhope, *University of Glamorgan , IK*
Yannis Tzitzikas, *University of Crete and FORTH-ICS, Greece*
Stefanos Vrochidis, *Information Technologies Institute, CERTH, Greece*
Michele Weigle, *Old Dominion University, USA*
Marcin Werla, *Poznań Supercomputing and Networking Center, Poland*
Iris Xie, *Universty of Wisconsin-Milwaukee, USA*
Maja Žumer, *University of Ljubljana, Slovenia*

MON**18 13:30 Room Patoulidou B**

KEYNOTE

Elton Barker

Elton Barker is Reader in Classical Studies, having joined The Open University as a Lecturer in July 2009. Before then, he had been a Tutor and Lecturer at Christ Church, Oxford (2004-09), and also lectured at Bristol, Nottingham and Reading. He has been a Junior Research Fellowship at Wolfson College, Cambridge (2002-04) and a Visiting Fellow at Venice International University (2003-04). From 2012-2013 he had a Research Fellowship for Experienced Researchers awarded by the Alexander von Humboldt Foundation for research at the Freie Universität Berlin and the University of Leipzig. He has been awarded a Graduate Teaching Award from Pembroke College (Cambridge) and twice won awards from the University of Oxford for an Outstanding Contribution to Teaching.

His research interests cross generic and disciplinary boundaries. Since 2008, he has been leading and co-running a series of collaborative projects, which are using digital resources to rethink spatial understanding of the ancient world. The Hestia project investigates the underlying ways in which Herodotus constructs space in book 5 of his Histories. Meanwhile, the Pelagios project has been establishing the Web infrastructure by which data produced and curated by different content providers – from academic projects like the Perseus Classical Library to cultural heritage institutions like the British Museum – can be linked through their common references to places.

Back to the future: annotating, collaborating and linking in a digital ecosystem

Classical philology has rarely been a self-enclosed discipline: in order to interpret Greek and Latin texts, it is necessary to place them in context—grounding them in the histories of the time and exploring them in and against those cultural horizons. Using the linking potential of the Web, Pelagios Commons (http://commons.pelagios.org/) has been pioneering a means of digital 'mutual contextualization', whereby any online document—be it a text, map, database or image—can be connected to another simply by virtue of having something in common with it, and then draw on this external content to enrich its own, or in turn be drawn upon by and enrich another. In Pelagios this linking is achieved through the method of annotating places. From having originally been seeded in collaboration with partners who already curated data and had the technical know-how to align datasets, Pelagios Commons now offers any researcher, librarian, museum curator, student or member of the public a simple, intuitive means to encode place information in a document of their choosing.

This presentation will set out and explain this annotation process in the Web-based, Open Source platform, Recogito (http://recogito.pelagios.org/) developed

by the Pelagios team. It will go through the steps that the researcher would take in order to geoannotate their material—first identifying the place entity in their document, then resolving that information to a central authority file: i.e. a gazetteer of placenames (e.g. http://pleiades.stoa.org/). It also considers the potential uses of this kind of semantic annotation, outlining the mapping of places in texts, the repurposing of the data in other systems (such as GIS), and the linking to other related resources. Throughout, however, it will be concerned to identify challenges and persistent issues that are not only related to the technical development and use; using Recogito puts a primary demand on defining and conceptualising place. Thus, contrary to much current thinking, this presentation hopes to show how digital tools can enhance the close reading of texts and facilitate a more nuanced understanding of the status and role of places in our historical sources.

Paul Groth

Paul Groth is Disruptive Technology Director at Elsevier Labs. He holds a Ph.D. in Computer Science from the University of Southampton (2007) and has done research at the University of Southern California and the Vrije Universiteit Amsterdam. His research focuses on dealing with large amounts of diverse contextualized knowledge with a particular focus on the web and science applications. This includes research in data provenance, data science, data integration and knowledge sharing. He lead architecture development for the Open PHACTS drug discovery data integration platform. Paul was co-chair of the W3C Provenance Working Group that created a standard for provenance interchange. He is co-author of "Provenance: an Introduction to PROV" and "The Semantic Web Primer: 3rd Edition" as well as numerous academic articles. He blogs at http://thinklinks.wordpress.com. You can find him on twitter: @pgroth.

Machines are People Too

The theory and practice of digital libraries provides a long history of thought around how to manage knowledge ranging from collection development, to cataloging and resource description. These tools were all designed to make knowledge findable and accessible to people. Even technical progress in information retrieval and question answering are all targeted to helping answer a human's information need.

However, increasingly demand is for data. Data that is needed not for people's consumption but to drive machines. As an example of this demand, there has been explosive growth in job openings for Data Engineers – professionals who prepare data for machine consumption. In this talk, I overview the information needs of machine intelligence and ask the question: Are our knowledge management techniques applicable for serving this new consumer?

Dimitrios Tzovaras

Dr. Dimitrios Tzovaras is a Senior Researcher Grade A' (Professor) and Director at CERTH/ITI (the Information Technologies Institute of the Centre for Research and Technology Hellas). He received the Diploma in Electrical Engineering and the Ph.D. in 2D and 3D Image Compression from the Aristotle University of Thessaloniki, Greece in 1992 and 1997, respectively. Prior to his current position, he was a Senior Researcher on the Information Processing Laboratory at the Electrical and Computer Engineering Department of the Aristotle University of Thessaloniki. His main research interests include network and visual analytics for network security, computer security, data fusion, biometric security, virtual reality, machine learning and artificial intelligence. He is author or co-author of over 110 articles in refereed journals and over 300 papers in international conferences.

Since 2004, he has been Associate Editor in the following International journals: Journal of Applied Signal Processing (JASP) and Journal on Advances in Multimedia of EURASIP. Additionally, he is Associate Editor in the IEEE Signal Processing Letters journal (since 2009) and Senior Associate Editor in the IEEE Signal Processing Letters journal (since 2012), while since mid-2012 he has been also Associate Editor in the IEEE Transactions on Image Processing journal. Over the same period, Dr. Tzovaras acted as ad hoc reviewer for a large number of International Journals and Magazines such as IEEE, ACM, Elsevier and EURASIP, as well as International Scientific Conferences (ICIP, EUSIPCO, CVPR, etc.).

Since 1992, Dr. Tzovaras has been involved in more than 100 European projects, funded by the EC and the Greek Ministry of Research and Technology. Within these research projects, he has acted as the Scientific Responsible of the research group of CERTH/ITI, but also as the Coordinator and/or the Technical/Scientific Manager of many of them (coordinator of technical manager in 21 projects – 10 H2020, 1 FP7 ICT IP, 7 FP7 ICT STREP, 3 FP6 IST STREP and 1 Nationally funded project).

Visualization in the Big Data Era: Data Mining from Networked Information

Network graphs have long formed a widely adapted and acknowledged practice for the representation of inter- and intra-dependent information streams. Nowadays, they are largely attracting the interest of the research community mainly due to the vastly growing amount (size & complexity) of semantically dependent data produced world-wide as a result of the rapid expansion of data sources.

In this context, the efficient processing of the big amounts of information, also known as Big Data forms a major challenge for both the research community and a wide variety of industrial sectors, involving security, health and financial applications.

In order to address these needs the current presentation describes a proprietary platform built upon state-of-the-art algorithms that are combined to implement a top-down approach for the facilitation of Data & Graph Mining processes, like behavioral clustering, interactive visualizations, etc.

The applicability of this platform has been validated on α series of distinct real-world use cases that involve large amounts of intra-exchanged information and can be thus help as characteristic examples of modern Big Data problems. In particular, they refer to (i) DoS attacks in a real-world mobile networks and (ii) early event detection in social media communities, (iii) traffic management and (iv) DNA sequences analysis.

In all these cases, the large volumes of data are addressed via a Data Minimization approach that starts with an aggregated overview of network at its whole, and gradually the focus is put on smaller data subsets (i.e. approach upon successive levels of abstraction). In parallel, insights on the network's operations are allowed through the detection of behavioral patterns. Similarly, a dynamic hypothesis formulator and the corresponding backend solver can subsequently be exploited through graph traversing and pattern mining. This way, an analyst is provided with the appropriate equipment to set and verify concrete hypotheses through simulation and extract useful conclusions.

17th European Networked Knowledge Organization Systems (NKOS) Workshop

https://at-web1.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2017

The 17th NKOS workshop at TPDL2017 will explore the potential of Knowledge Organization Systems, such as classification systems, taxonomies, thesauri, ontologies, and lexical databases in the context of current developments and possibilities. These tools help to model the underlying semantic structure of a domain for purposes of information retrieval, knowledge discovery, language engineering, and the semantic web. The workshop provides an opportunity to discuss projects, research and development activities, evaluation approaches, lessons learned, and research findings.

MAIN WORKSHOP THEMES

- KOS Alignment. KOS alignment or terminology mapping plays a vital role in NKOS for many years. This year we want to sort out the needs (use cases) of KOS alignments in the new environment of Linked Open Data. We plan to collect methodologies, best practices, guidelines and tools. This includes manual and automatic alignments.
- KOS Linked Open Data. Recent years have seen an increasing trend to publication of KOS as Linked Data vocabularies. We need discussion of practical initiatives to link between congruent vocabularies and provide effective web services and APIs so that applications can build upon them.
- KOS and Document Retrieval. Documents or parts of documents are nowadays not only accessible via their metadata but their abstracts and in many cases the full texts are electronically available. Thus, these documents also can be found by search engines. Given this possibility of full text search the role of classification and annotation has to be redefined. Questions like the following ones arise: can traditional knowledge organization and document annotation improve full text retrieval? Are classification, categorisation, annotation, tagging, and full text retrieval complementary, or how can they be made complementary? What should be the focus of annotation, if full text retrieval is available?

CHAIRS AND COMMITTEES

Philipp Mayr, *GESIS – Leibniz Institute for the Social Sciences, Germany*
Douglas Tudhope, *University of South Wales, UK*
Koraljka Golub, *Linnaeus University, Sweden*
Christian Wartena, *Abteilung Information und Kommunikation, Germany*
Ernesto William De Luca, *Georg-Eckert-Institut – Leibniz-Institut für internationale Schulbuchforschung, Germany*

(meta)-data quality workshop

http://qualitics.org/mdqual2017

It is well known that we are rapidly moving towards a data driven world where all aspects in our everyday lives are data driven. In all domains from healthcare to retail and finance, data is collected, analysed and used to make decisions, usually utilizing machine learning techniques.

Data Science involves collecting, cleansing and integrating data prior of analysis. The quality of this data is critical and directly affects the outcome of all data science related tasks. Moreover, metadata is used to annotate data and facilitate data organization and retrieval. Metadata quality also directly affects retrieval and other operations (such as data integration) and workflows that are metadata driven.

Although various metrics have been proposed to measure metadata and data quality, in most cases they are highly subjective and/or domain specific. Moreover, they are directly related to the intended use of the data, meaning that a dataset could be of high quality for one use and of low quality for another. In all cases, (meta)data quality has a tremendous impact on data science related tasks and ultimately in everyday life.

The proposed workshop aims at exploring the various quality issues found in people working with both data and metadata across domains. An inter-disciplinary workshop where data scientists across different domains will meet and:

- share and exchange experiences regarding (meta)data quality
- identify patterns in (meta)data quality
- share methodologies and metrics that will help to measure (meta)-data quality
- share / propose tools that can be used effectively in improving (automatically) (meta)-data quality.

This initiative aims at bringing together a community of data scientists that have expertise in a diverse set of domains such as archives and libraries, healthcare, biology, humanities, computer science and engineering, environment, agriculture, economics, etc.

Apart from sharing metrics and methods to identify and resolve quality issues and evaluate datasets, the proposed workshop aims at promoting the use of tools and services for the automatic measurement and improvement of (meta) data quality. Although few such tools are available in the market, a good number of standalone micro-services are available and can be used to automatically improve (meta)data quality.

CHAIRS

Dimitris Gavrilis, *University of Patras, Greece*
Christos Papatheodorou, *Ionian University, Greece*

PROGRAMME COMMITTEE

Trond Aalberg, *NTNU, Norway*
Amir Aryani, *Australian National Data Service, Australia*
Donatella Castelli, *CNR, Italy*
Valentine Charles, *Europeana Foundation, The Netherlands*
Peter Doorn, *Data Archiving and Networked Services, The Netherlands*
Pythagoras Karampiperis, *Agroknow, Greece*
Laurent Romary, *French Institute for Research in Computer Science and Automation, France*
Timos Sellis, *Swinburne University of Technology, Australia*
Giannis Tsakonas, *University of Patras, Greece*
Michalis Sfakakis, *Ionian University, Greece*

1st International Workshop on Temporal Dynamics in Digital Libraries (TDDL-2017)

https://tddl2017.github.io/

In Digital Libraries, which can often span several epochs, time is a critical factor. It is the means by which understanding, searching, and exploring these collections of data. Temporal dynamics, i.e. time-based patterns and trends, underpin language usage, entity references, and cultural and economic trends. Users accessing the information contained in Digital Libraries have to deal with their partial knowledge of these phenomena (word meaning variation, entity temporal ambiguity, specific events and time-related trends), as well as their own temporal evolution, i.e. their change in interests, preferences, and goals over time. Intercepting, representing, and predicting these dynamics is fundamental to the intelligent information access in Digital Libraries. This workshop proposes to bring together researchers and practitioners from different backgrounds in order to identify and discuss research trends, challenges, and new opportunities related to the time-aware intelligent access to Digital Libraries.

TOPICS OF INTEREST

- Diachronic analysis of language
- Time-aware Information Retrieval for Digital Libraries
- Time-aware Recommender Systems for Digital Libraries
- Timeline Summarization
- Time-aware User Modeling for Digital Libraries
- Event detection
- Time-aware entity disambiguation
- Topic detection and tracking
- Temporal clustering
- Timeline interfaces
- Temporal queries
- Historical studies and computational history
- Topic and entity evolution
- Opinion changes over time
- Web archive-related topics

WORKSHOP ORGANIZERS

Annalina Caputo, *ADAPT centre, School of Computer Science and Statistics, Trinity College Dublin, Ireland*

Nattiya Kanhabua, *Database, programming and web technologies (DPW) Group, Department of Computer Science, Aalborg University, Denmark*

Pierpaolo Basile, *SWAP group, Department of Computer Science, University of Bari Aldo Moro, Italy*

Séamus Lawless, *Knowledge and Data Engineering Group, School of Computer Science and Statistics, Trinity College Dublin, Ireland*

PROGRAMME COMMITTEE

Ingo Frommholz, *School of Computer Science and Technology, University of Bedfordshire*

Ted Underwood, *University of Illinois, Urbana-Champaign*

Michele Filannino, *SUNY @ Albany and MIT Clinical Decision Making group*

Adam Jatowt, *Kyoto University*

Wolfgang Nejdl, *L3S Research Center, Hannover*

Kjetil Nørvåg, *Norwegian University of Science and Technology*

Thomas Risse, *L3S Research Center, Hannover*

Rachele Sprugnoli, *Digital Humanities at Fondazione Bruno Kessler (FBK)*

Workshop on Modeling Societal Future (FUTURITY-2017)

https://profs.info.uaic.ro/~futurity/

People seek and share ideas, information, experiences, expertise, opinions, and emotion with both acquaintance and strangers on the Internet, based on the effect of the Wisdom of Crowds. Over the last few years, the use of Social Media has increased tremendously all over the world. The huge popularity of social networks provides an ideal environment for scientists to test and simulate new models, algorithms and methods to process knowledge. Structured social knowledge can be used by different actors (companies, public institutions, researchers and scholars interested in formal and empirical analysis of social trends) to understand the behaviors in users or groups.

As recent advances in information and communication technologies continue to reshape the relationship between governments and citizens, opportunities emerge at both ends. Citizens route their voices through new electronic channels, hoping to have their opinions heard at any time from any place. At the same time, companies are willing to identify user's opinion and perceived contexts about their products.

MAIN WORKSHOP THEMES

- Extracting knowledge from social web,
- Collaborative and interactive search,
- Conversational search interaction,
- Community behavioral analysis,
- Intelligent personal assistants,
- Semantics in digital libraries,
- Extracting and mining forum data,
- Social media and linked data methodologies in real-life scenarios,
- Collaborative tools and services for citizens, organizations, communities,
- Creating and using structured social media-based resources through social web mining,
- Exploring crowdsourcing and user communities,
- Strategic early warning systems and detection of week signals,
- Using the social web to foster innovation,
- Exploring the digital cultural heritage,
- Interaction with the web as a mental, social and physical extension of people.

ORGANISATION COMMITTEE

Daniela Gîfu, *"Alexandru Ioan Cuza" University of Iași, Romania and Romanian Academy - Iași branch*
Diana Trandabăț, *"Alexandru Ioan Cuza" University of Iași, Romania*

PROGRAM COMMITTEE

Nuria Bel, *Universitat Pompeu Fabra, Barcelona, Spain*
Paul Buitelaar, *National University of Ireland, Galway*
Costin Bădică, *University of Craiova, Romania*
Mihaela Colhon, *University of Craiova, Romania*
Dan Cristea, *"Alexandru Ioan Cuza" University of Iași, Romania*
Thierry Declerck, *Universität des Saarlandes, Saarbrücken, Germany*
Daniela Gîfu, *"Alexandru Ioan Cuza" University of Iași, Romania and Romanian Academy - Iași branch*
Radu Ion, *Microsoft Ireland*
Bernardo Magnini
Rada Mihalcea, *University of Michigan, USA*
Vivi Năstase, *Fondazione Bruno Kessler, Trento, Italy*
Octavian Popescu, *IBM Research, USA*
Dan Ștefănescu, *Vantage Labs, USA*
Diana Trandabăț, *"Alexandru Ioan Cuza" University of Iași, Romania*
Dan Tufiș, *Romanian Academy, Research Institute forArtificial Intelligence "Mihai Drăgănescu", Romania*
Piek Vossen, *Vrije Universiteit, Amsterdam, the Netherlands*
Michael Zock, *Aix-Marseille Université, France*

Axel J. Soto & Riza Batista-Navarro

Riza Batista-Navarro (http://personalpages.manchester.ac.uk/staff/riza.batista) is a Lecturer at the School of Computer Science of the University of Manchester. She obtained her bachelors and masters degrees in Computer Science from the University of the Philippines (where she was assistant professor at the Department of Computer Science). She then came to Manchester to join the National Centre for Text Mining (NaCTeM) and to pursue her PhD in Computer Science. Riza has worked in a number of challenging areas such as event extraction, coreference resolution and named entity recognition for the biomedical domain. She has been involved in a number of projects including the transatlantic collaboration Mining Biodiversity, in which she led the work on enhancing the Biodiversity Heritage Library archives with text-mined semantic metadata and search capabilities.

Axel J. Soto (http://personalpages.manchester.ac.uk/staff/Axel.Soto/index.html) is a Research Associate at the University of Manchester and the National Centre for Text Mining (NaCTeM), UK. Before this position, he was a Research Associate with the Faculty of Computer Science and Adjunct Professor at Dalhousie University (Canada). He received his B.Sc. in Computer Systems Engineering and his PhD in Computer Science at Universidad Nacional del Sur (Argentina) in 2005 and 2010, respectively. Most of his research focusses on the area of text mining, machine learning and visual analytics. During his postdoctoral time he has been conducting research in the area of visual text analytics in order to improve analytical processes for mining large collections of documents and how interactive visualisations can help the exploration and extraction of knowledge from text.

Enriching digital collections using tools for text mining, indexing and visualisation

The tutorial will demonstrate a suite of tools for text mining, semantic indexing and visualisation that will facilitate enhanced searching and exploration of digital collections. Specifically, we aim to provide: (1) an introduction to modular text mining and indexing workflows developed using the Argo platform; (2) an overview of the Elasticsearch indexing engine and the Kibana visualisation platform; and (3) the know-how on building and visualising semantic indexes over textual collections without any programming effort.

The tutorial will cover the end-to-end automatic generation, querying and visualisation of a semantically enabled search index over textual collections. By the end of the tutorial, the audience will have gained knowledge on: (1) exemplar digital collections (e.g., the Biodiversity Heritage Library, British Medical Journal) enhanced with text-mined semantic metadata and visualisation tools; (2) information extraction methods for generating semantic metadata over textual collections; (3) generating Elasticsearch indexes to search over digital collections that were semantically enriched by constructing text mining workflows using Argo; and (4) using the Kibana platform to generate dashboards and visually explore digital collections indexed with Elasticsearch.

Requirements: A Web browser, preferably Google Chrome, Mozilla Firefox or Safari, is required.

Andrea Bollini & Claudio Cortese

Andrea Bollini is Chief Technology and Innovation Officer (CTO / CTIO) at 4Science, with the responsibility of ensuring the use of the most efficient technologies to effectively achieve the results of each single project. Andrea is actively involved in various international open source and open standards communities, often with leading roles: DSpace committer, Deputy Leader of the euroCRIS' CERIF Task Group "CERIF" and member of the COAR Working Group "Next Generation Repositories". Chair, speaker and reviewer in important conferences, before joining 4Science Andrea worked for two Italian University Consortia, CILEA and CINECA, where he was responsible for the development, design and management of IT solutions and projects in the field of research, electronic publishing and open access repositories.

Claudio Cortese is Project Manager and Business Analyst at 4Science. He has a PhD in archaeology and is lecturer of "Computer Applications in Archaeology" at the Catholic University of Milan. For several years he has been dealing with modeling, management and analysis of data for cultural heritage using a great variety of methods, standards and technologies: from relational databases to GIS, from Digital Libraries to architectures based on the Semantic Web. Before joining 4Science, he worked for two Italian University Consortia, CILEA and CINECA, mainly dealing with Digital Library Management Systems to preserve, use and distribute digital cultural resources. In the historical/archaeological field Claudio focused on the design and creation of databases for many Archaeological Missions and University Research Units. He constantly gives consultancies, lessons and courses to universities and public/private institutions.

Putting Historical Data in Context: How to use DSpace-GLAM

The proposed tutorial will introduce attendees to DSpace-GLAM (Galleries, Libraries, Archives, Museums), the Digital Library Management System based on DSpace and DSpace-CRIS, developed by 4Science for the management, analysis and preservation of digital cultural heritage, covering its functional and technical aspects.

DSpace-GLAM is an additional open-source configuration for the DSpace platform. It extends the DSpace data model, providing the ability to manage, collect and expose data about every entity important for the cultural heritage domain, such as persons, events, places, concepts and so on.

The extensible data model will be explained in depth, through examples and discussions with participants.

Other main topics will be DSpace-GLAM "components", relationships management and network analysis.

Finally 4Science new add-ons for digital cultural resources fruition and analysis (the IIIF – International Image Interoperability Framework – Image Viewer, the Audio/Video Streaming Module, the OCR Module and the CKAN integration) will be illustrated.

At the end of the tutorial the participants will be able to understand the DSpace-GLAM data model, to adapt it to their needs and to evaluate if DSpace GLAM fits the needs of their institution.

Requirements: The level of this tutorial is introductory. It is addressed to librarians, archivists, historians, archaeologists, researchers and to those who want to build their own digital library but do not want to write their own software nor buy a proprietary solution. No programming ability is required. Basic knowledge of digital libraries and repositories architectures and of the relational model, though not mandatory, can guarantee a better learning experience.

Michail Salampasis

Michail Salampasis is a professor at the department of Informatics at the TEI of Thessaloniki, Greece. He has a BSc in Informatics and a PhD in Computing. His main research interests are in applied studies in information science, innovation search and distributed information retrieval. He has published about 80 papers in refereed journals and international conferences. He was the coordinator of the Cost Action IC1002 on "Multilingual and Multifaceted Interactive Information Access' and a Marie Curie Fellow at the Institute of Software Technology and Interactive Systems, Vienna University of Technology on a program entitled "Pluggable Platform for Personalised Multilingual Patent Search". Salampasis has taught at different levels, from BSc/MSc lectures on information retrieval and patent search to tutorials aimed at PhD students and researchers. Web site: www.it.teithe.gr/~cs1msa.

Innovation Search

Innovation search presents many challenges to the research community and also to professional searchers and search solution providers. Patents are complex technical documents, whose content appears in many languages and contains images, chemical and genomic structures and other forms of data, intermixed and cross-referring with the text material. Further much innovation search involves the search of other forms of technical information such as scientific papers, or the integration of open linked data and so on with the patent data. Finally the realistic presentation of search and analysis results to often non-technical and time-poor audiences for purpose of strategic decision making presents particular challenges.

The course will review the state of the art and point out where the key challenges are, especially for early stage researchers and innovation professionals in patent search and related disciplines. The objectives of the tutorial are:

- Understand the international patent system, patent searching and the relevant state-of-the-art.
- Understand the key limitations and challenges for the research community in the development of patent retrieval and innovation search systems in general.
- Understand how recent developments in information retrieval, multilingual and interactive information access may be applied to patent searching research.

Requirements: We expect the target audience to consist mainly of two groups. First, postgraduate students and post-doc researchers from academia engaging in studies related to information retrieval, professional search systems and natural language processing. Second, researchers from other related disciplines and professionals (e.g. search solutions providers) who will be given the opportunity to enhance their expertise towards the area of innovation search. Since the class will introduce foundations and basic concepts of patents and inno-vation search, it will also be accessible to individuals not familiar with the field of information retrieval. We hence will not rely on any particular prior knowledge.

Andreas Rauber

Andreas Rauber is Associate Professor at the Department of Software Technology and Interactive Systems (ifs) at the Vienna University of Technology (TU-Wien). He furthermore is president of AARIT, the Austrian Association for Research in IT, a Key Researcher at Secure Business Austria (SBA-Research) and Co-Chair of the RDA Working Group on Dynamic Data Citation. He received his MSc and PhD in Computer Science from the Vienna University of Technology in 1997 and 2000, respectively. In 2001 he joined the National Research Council of Italy (CNR) in Pisa as an ERCIM Research Fellow, followed by an ERCIM Research position at the French National Institute for Research in Computer Science and Control (INRIA), at Rocquencourt, France, in 2002. From 2004-2008 he was also head of the iSpaces research group at the eCommerce Competence Center (ec3).

His research interests cover the broad scope of digital libraries and information spaces, including specifically text and music information retrieval and organization, information visualization, as well as data analysis and digital preservation, all of which started to merge recently under the umbrella of reproducible science.

Enabling Precise Identification and Citability of Dynamic Data – Recommendations of the RDA Working Group on Data citation

"Sound, reproducible scholarship rests upon a foundation of robust, accessible data. For this to be so in practice as well as theory, data must be accorded due importance in the practice of scholarship and in the enduring scholarly record. In other words, data should be considered legitimate, citable products of research. Data citation, like the citation of other evidence and sources, is good research practice and is part of the scholarly ecosystem supporting data reuse." (Data Citation principles, [1])

While the importance of these Data Citation Principles is by now widely accepted, several challenges persist when it comes to actually providing the services needed to support precise identification and citation of data, particularly in dynamic environments. In order to repeat an earlier study, to apply data from an earlier study to a new model, we need to be able to precisely identify the very subset of data used. While verbal descriptions of how the subset was created (e.g. by providing selected attribute ranges and time intervals) are hardly precise enough and do not support automated handling, keeping redundant copies of the data in question does not scale up to the big data settings encountered in many disciplines today. Conventional approaches, such as assigning persistent identifiers to entire data sets or individual subsets or data items, are not sufficient to meet these requirements. This problem is further exacerbated if the data itself is dynamic, i.e. if new data keeps being added to a database, if errors are corrected or if data items are being deleted.

Starting from the Data citation Principles we will review the challenges identified above and discuss the solutions and recommendations that have been elaborated within the context of a Working Group of the Research Data Alliance (RDA) on Data Citation: Making Dynamic Data Citeable. These approaches are based on versioned and time-stamped data sources, with persistent identifiers being assigned to the time-stamped queries/expressions that are used for creating the subset of data. We will review examples of how these can be implemented for different types of data, including SQL-style databases, comma-separated value files (CSV) and others, and take a look at operational implementations in a variety of data centers.

[1] Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego, CA: FORCE11; 2014.

BRIEFPROGRAM

TPDL2017

MON18

08:00-09:00	REGISTRATION	
09:00-10:30	**TUTORIALS** *pages 22, 32*	**DOCTORAL CONSORTIUM** *page 32*
10:30-11:00	COFFEE	
11:00-12:30	**TUTORIALS** *pages 22, 32*	**DOCTORAL CONSORTIUM** *page 32*
12:30-13:30	LUNCH	
13:30-14:30	**BARKER** *pages 9,32*	
14:30-16:00	**LINKED DATA** *page 34*	**CORPORA** *page 35*
16:00-16:30	COFFEE	
16:30-17:30	**SHORT PAPERS I** *page 34*	**SHORT PAPERS II** *page 35*
17:30-18:00	MINUTE MADNESS	
18:00-19:00	**POSTER/DEMO SESSION** *page 36*	
19:00	CONFERENCE RECEPTION	

TUE19

08:00-09:30	REGISTRATION	
09:30-10:30	**GROTH** *pages 11, 38*	
10:30-11:00	COFFEE	
11:00-12:30	**DATA IN DL** *page 38*	**QUALITY IN DL** *page 39*
12:30-14:00	LUNCH	
14:00-15:30	**DIGITAL HUMANITIES** *page 40*	**ENTITIES** *page 41*
15:30-16:00	COFFEE	
16:00-17:30	**SCHOLARLY COMMUNICATION** *page 40*	**SENTIMENT ANALYSIS** *page 41*
21:00	CONFERENCE DINNER	

WED20

08:00-09:00	REGISTRATION	
09:00-10:30	**ANN_ENG** *page 42*	**CLAS_PR2** *page 43*
10:30-11:00	COFFEE	
11:00-12:00	**TZOVARAS** *pages 12, 42*	
12:00-12:30	CLOSING SESSION	
12:30-14:00	LUNCH	
14:30-18:30	VISIT TO VERGINA	
19:00	WORKSHOP RECEPTION	

THU21

08:00-09:00	REGISTRATION
09:00-11:00	**WORKSHOPS** *pages 14, 44*
11:00-11:30	COFFEE
11:30-13:00	**WORKSHOPS** *pages 14, 44*
13:00-14:30	LUNCH

KEYNOTES
MAIN SESSIONS
WORKSHOPS
TUTORIALS / DC

08:00 ↓ 09:00

Registration

Room Patoulidou B **TUT**

09:00 ↓ 12:30

Tutorials (in parallel)

Enriching digital collections using tools for text mining, indexing and visualization
Axel J. Soto and Riza Batista-Navarro

Putting Historical Data in Context: How to use DSpace-GLAM
Andrea Bollini and Claudio Cortese

Innovation Search
Michail Salampasis

Enabling Precise Identification and Citability of Dynamic Data – Recommendations of the RDA Working Group on Data citation
Andreas Rauber

10:30 ↓ 11:00

Coffee break

12:30 ↓ 13:30

Lunch

Room Patoulidou B

13:30 ↓ 14:30

Keynote Speech 1

Back to the future: annotating, collaborating and linking in a digital ecosystem
Elton Barker

Chair **Giannis Tsakonas**

DC **Room Olympia D**

09:00 ↓ 12:30

Doctoral Consortium

Chairs **Heiko Schuldt, Maja Žumer**

Research data in Norway: How does expectations, demands and solutions correspond in the knowledge infrastructure for research data?
Live Kvale

The data practices of researchers within a Horizon2020-project - a Scholarly Communication study
Madeleine Dutoit

Top-down and Bottom-up Approaches to Identify the Users, the Services and the Interface of a 2.0 Digital Library
Elina Leblanc

Studying conceptual models for publishing library data to the Semantic Web
Sofia Zapounidou

10:30 ↓ 11:00

Cross-Language Record Linkage across Humanities Collections using Metadata Similarities among Languages
Yuting Song

Machine Learning Architectures for Scalable and Reliable Subject Indexing: Fusion, Knowledge Transfer, and Confidence
Martin Toepfer

Explaining Pairwise Relationships between Documents
Nils Witt

12:30 ↓ 13:30

Lunch

(lo): Long Oral: 30 minutes
(so): Short Oral: 15 minutes
(sp): Short Paper: 10 minutes

14:30 ↓ 16:00

Room Patoulidou B — **LD**

Session 1: **Linked Data**

Chair **Nuno Freire**

Exploiting Interlinked Research Metadata (lo)
Shirin Ameri, Sahar Vahdati and Christoph Lange

Preserving bibliographic relationships in mappings from FRBR to BIBFRAME 2.0 (lo)
Sofia Zapounidou, Michalis Sfakakis and Christos Papatheodorou

Exploring ontology-enhanced bibliography databases using faceted search (so)
Tadeusz Pankowski

What Should I Cite? Cross-Collection Reference Recommendation of Patents and Papers (so)
Julian Risch and Ralf Krestel

16:00 ↓ 16:30

Coffee break

16:30 ↓ 17:30

Room Patoulidou B — **SP**

Short papers

Chair **Christos Papatheodorou**

Towards building knowledge resources from social media using semantic roles
Diana Trandabat

Social Dendro: social network techniques applied to research data description
Nelson Pereira, João Rocha Da Silva and Cristina Ribeiro

Semantic Enrichment of Web Query Interfaces to enable Dynamic Deep Linking to Web Information Portals
Arne Martin Klemenz and Klaus Tochtermann

Analysis of Interactive Multimedia Features in Scientific Publication Platforms
Camila Wohlmuth Da Silva and Nuno Correia

Incidental or influential? - Challenges in automatically detecting citation importance using publication full texts
David Pride and Petr Knoth

14:30 ↓ 16:00

CO — **Room Olympia D**

Session 2: **Corpora**

Chair **Marcin Wela**

Corpus-based concept summaries for document annotation (lo)
Ikechukwu Nkisi-Orji, Nirmalie Wiratunga, Kit-Ying Hui, Rachel Heaven and Stewart Massie

A Comparative Study of Language Modeling to Instance-based Methods and Feature Combinations for Authorship Attribution (lo)
Olga Fourkioti, Symeon Symeonidis and Avi Arampatzis

A Digital Repository for Physical Samples: Concepts, Solutions and Management (so)
Anusuriya Devaraju, Jens Klump, Victor Tey, Ryan Fraser, Simon Cox and Lesley Wyborn

Facet Embeddings for Explorative Analytics in Digital Libraries (so)
Sepideh Mesbah, Kyriakos Frageskos, Christoph Lofi, Alessandro Bozzon and Geert-Jan Houben

16:00 ↓ 16:30

Coffee break

16:30 ↓ 17:30

SP — **Room Olympia D**

Short papers

Chair **Trond Aalberg**

A Complete Year of User Retrieval Sessions in a Social Sciences Academic Search Engine
Philipp Mayr and Ameni Kacem

User Interactions with Bibliographic Information Visualizations
Athena Salaba and Tanja Merčun

Building the Brazilian Academic Genealogy Tree
Wellington Dores, Elias Soares, Fabrício Benevenuto and Alberto Laender

What Should I Cite? Cross-Collection Reference Recommendation of Patents and Papers
Julian Risch and Ralf Krestel

When a Metadata Provider Task is Successful
Sarantos Kapidakis

Extending R2RML with Support for RDF Collections and Containers to Generate MADS-RDF Datasets from Relational Databases
Christophe Debruyne, Lucy McKenna and Declan O'Sullivan

(lo): Long Oral: 30 minutes
(so): Short Oral: 15 minutes
(sp): Short Paper: 10 minutes

17:30 ↓ 18:00

Room Patoulidou B — **MM**

Minute Madness

Chair **Jose Borbinha**

18:00 ↓ 19:00

Room Patoulidou B — **PDS**

Poster/Demo Session

Chair **Jose Borbinha**

Development of an RDF-Enabled Cataloguing Tool
Lucy McKenna, Marta Bustillo, Tim Keefe, Declan O'Sullivan and Christophe Debruyne

Towards Semantic Quality Control of Automatic Subject Indexing
Martin Toepfer and Christin Seifert

Introducing Solon: A system for Modelling, Managing and Mining Legal Sources
Marios Koniaris, George Papastefanatos, Marios Meimaris and Giorgos Alexiou

Towards finding animal replacement methods
Nadine Dulisch and Brigitte Mathiak

Challenges in Scientific Workflow Management
Kyriakos Kritikos and Friederike Klan

Sustaining Cultural Treasures within the Internet of Things: Monitoring Libraries with MonTreAL
Marcel Großmann, Steffen Illig and Cornelius Matějka

Towards a Semantic Search Engine for Scientific Articles
Bastien Latard, Germain Forestier, Jonathan Weber and Michel Hassenforder

18:00 ↓ 19:30

Steering Committee Meeting

Conference Reception

18:00 ↓ 20:00

08:30 ↓ 09:30 — Registration

Room Patoulidou B

09:30 ↓ 10:30 — **Keynote Speech 2**: Machines are People Too — **Paul Groth**

Chair **Jaap Kamps**

10:30 ↓ 11:00 — Coffee break

Room Patoulidou B — **DDL**

11:00 ↓ 12:30 — Session 3: **Data in Digital Libraries**

Chair **Petr Knoth**

Extracting Event-Centric Document Collections from Large Scale Web Archives (lo)
Gerhard Gossen, Elena Demidova and Thomas Risse

Automatic Hierarchical Categorization of Research Expertise Using Minimum Information (lo)
Gustavo Siqueira, Sergio Canuto, Alberto Laender and Marcos André Gonçalves

Challenges of Research Data Management for High Performance Computing (so)
Björn Schembera and Thomas Bönisch

Information Governance Maturity Model - Final Development Iteration (so)
Diogo Proença, Ricardo Vieira and Jose Borbinha

QDL — **Room Olympia D**

11:00 ↓ 12:30 — Session 4: **Quality in Digital Libraries**

Chair **João Rocha da Silva**

Can Plausibility Help to Support High Quality Content in Digital Libraries? (lo)
Jose Maria Gonzalez Pinto and Wolf-Tilo Balke

How Linked Data can aid Machine Learning-based Tasks (lo)
Michalis Mountantonakis and Yannis Tzitzikas

Classifying document types to enhance search and recommendations in digital libraries (so)
Aristotelis Charalampous and Petr Knoth

Understanding the Influence of Hyperparameters on Text Embeddings for Natural Language Processing Tasks (so)
Nils Witt and Christin Seifert

12:30 ↓ 14:00 — Lunch — Lunch

Room Patoulidou B — **DH**

14:00 ↓ 15:30

Session 5: **Digital Humanities**

Chair **Mark Hall**

Metadata aggregation: assessing the application of IIIF and Sitemaps within cultural heritage (lo)
Nuno Freire, Glen Robson, John B. Howard, Hugo Manguinhas and Antoine Isaac

What Information Users Search for Using Europeana and Why (lo)
Paul Clough, Timothy Hill, Monica Lestari Paramita and Paula Goodale

On the uses of word sense change for research in the digital humanities (so)
Nina Tahmasebi and Thomas Risse

A Decade of evaluating Europeana - What have we learned? (so)
Vivien Petras and Juliane Stiller

EN — **Room Olympia D**

14:00 ↓ 15:30

Session 6: **Entities**

Chair **Anastasios Gounaris**

RussianFlu-DE: A German Corpus for a Historical Epidemic with Temporal Annotation (lo)
Van Canh Tran, Katja Markert and Wolfgang Nejdl

Multi-aspect Entity-centric Analysis of Big Social Media Archives (lo)
Pavlos Fafalios, Vasileios Iosifidis, Kostas Stefanidis and Eirini Ntoutsi

Semantic Author Name Disambiguation with Word Embeddings (so)
Mark-Christoph Mueller

What Others Say About This Work? Scalable Extraction of Citation Contexts from Research Papers (so)
Petr Knoth, Phil Gooch and Kris Jack

15:30 ↓ 16:00

Coffee break

Coffee break

15:30 ↓ 16:00

Room Patoulidou B — **SC**

16:00 ↓ 17:30

Session 7: **Scholarly Communication**

Chair **Maja Žumer**

Integration of Scholarly Communication Metadata using Knowledge Graphs (lo)
Afshin Sadeghi, Christoph Lange, Maria Esther Vidal and Sören Auer

Towards a Knowledge Graph Representing Research Findings by Semantifying Survey Articles (lo)
Said Fathalla, Sahar Vahdati, Christoph Lange and Sören Auer

Analysing Scholarly Communication Metadata of Computer Science Events (so)
Said Fathalla, Sahar Vahdati, Christoph Lange and Sören Auer

High-Pass Text Filtering for Citation Matching (so)
Yannis Foufoulas, Lefteris Stamatogiannakis, Harry Dimitropoulos and Yannis Ioannidis

SA — **Room Olympia D**

16:00 ↓ 17:30

Session 8: **Sentiment Analysis**

Chair **Kostas Stefanidis**

Mining Semantic Patterns for Sentiment Analysis of Product Reviews (lo)
Sang-Sang Tan and Jin-Cheon Na

Sentiment Classification over Opinionated Data Streams through Informed Model Adaptation (lo)
Vasileios Iosifidis, Annina Oelschlager and Eirini Ntoutsi

A Comparison of Pre-processing Techniques for Twitter Sentiment Analysis (so)
Dimitrios Effrosynidis, Symeon Symeonidis and Avi Arampatzis

Employing Twitter hashtags and linked data to suggest trending resources in a digital library (so)
Konstantinos Kyprianos, Ioannis Papadakis, Apostolos Karalis and Christos Douligeris

21:00 ↓ 23:00

Conference Dinner
(Mediterranean Palace Hotel)

08:30 ↓ 09:00 — Registration

Room Patoulidou B — **IB**

09:00 ↓ 10:30 — Session 9: **Information Behavior**

Chair **Juliane Stiller**

The Ghost in the Museum Website: Investigating the General Public's Interactions with Museum Websites (lo)
David Walsh, Mark Hall, Paul Clough and Jonathan Foster

Social Tagging: Implications from Studying User Behaviour and Institutional Practice (lo)
Õnne Mets and Jaagup Kippar

Evaluating the Usefulness of Visual Features for Supporting Document Triage (so)
Dagmar Kern, Maria Lusky and Dirk Wacker

Building User Groups based on a Structural Representation of User Search Sessions (so)
Wilko van Hoek and Zeljko Carevic

IR — **Room Olympia D**

09:00 ↓ 10:30 — Session 10: **Information Retrieval**

Chair **Philipp Mayr**

Multiple Random Walks for Personalized Ranking with Trust and Distrust (lo)
Dimitrios Rafailidis and Fabio Crestani

Plagiarism Detection based on Citing Sentences (lo)
Sidik Soleman and Atsushi Fujii

Lexicon Induction from Text for Interpretable Classification (so)
Jeremie Clos and Nirmalie Wiratunga

The Clustering-based Initialization for Non-negative Matrix Factorization in the Feature Transformation of the High-dimensional Text Categorization System: A Viewpoint of Term Vectors (so)
Nguyen Hoai Nam Le and Bao Quoc Ho

10:30 ↓ 11:00 — Coffee break

10:30 ↓ 11:00 — Coffee break

Room Patoulidou B

11:00 ↓ 12:00 — **Keynote Speech 3**

Visualization in the Big Data Era: Data Mining from Networked Information

Dimitrios Tzovaras

Chair **Jose Borbinha**

Room Patoulidou B

12:00 ↓ 12:30 — Closing Ceremony

12:30 ↓ 14:00 — Lunch

14:30 ↓ 18:30 — Visit to Vergina

19:00 ↓ 20:00 — Workshop Reception

09:00 ↓ 13:00

Room Kallipatira

NKOS-2017

Welcome

Classifying Medical Literature Using k-Nearest-Neighbours Algorithm
Andreas Lüschow and Christian Wartena

Wikidata as a linking hub for knowledge organization systems? Experiences from an authority mapping project and and outlook on thesaurus mappings
Joachim Neubert

CCWeb: KOS Mapping Application System
Morsheda Akter and Umamaheswari Balakrishnan

The Publication of a Classification Knowledge Organization System as a Linked Data Vocabulary
Olga Lavrenova

Towards VocBench 3: Pushing Collaborative Development of Thesauri and Ontologies Further Beyond
Armando Stellato, Andrea Turbati, Manuel Fiorelli, Tiziano Lorenzetti, Eugeniu Costetchi, Christine Laaboudi, Willem Van Gemert and Johannes Keizer

Knowledge Organisation for Digital Libraries
Lena-Luise Stahn, Ingetraut Dahlberg and Ernesto William De Luca

Aquacold – a crowdsourced query understanding and query construction tool for the linked data web
Nicholas Collis and Ingo Frommholz

Room Patoulidou B

MDQual-2017

Welcome

Session 1

Refining Duplicate Detection for Improved Data Quality
Yu Huang and Fei Chiang

Compiling Keyphrase Candidates for Scientific Literature Based on Wikipedia
Hung-Hsuan Chen

Metadata in Mexican television news broadcasts on the web
Catalina Naumis and Soto Silvano

Keynote session

A Simple Approach to Assessing the FAIRness of Data in Trusted Digital Repositories
Peter Doorn and Eleftheria Tsoupra

Discussion

Demos & tools session

Repolytics: identifying measurable insights for digital repositories
Vangelis Nomikos

Session 2

Data Quality Assessment in Europeana: Metrics for Multilinguality
Valentine Charles, Juliane Stiller, Péter Király, Wener Bailer and Nuno Freire

Identifying low quality issue patterns: a National Library of Greece Use Case
Matthaios Stratis, Michalis Gerolimos, Filippos Tsimpoglou, Stavros Angelis, Vangelis Nomikos and Dimitris Gavrilis

Discussion & closing

09:00 ↓ 13:00

Room Ilida

TDDL-2017

Welcome

A Half-Life Decaying Model for Recommender Systems with Matrix Factorization
Panagiotis Ardagelou, and Avi Arampatzis

HEI: Hunter Events Interface
Antonio Sorgente, Antonio Calabrese, Gianluca Coda, Paolo Vanacore, and Francesco Mele

Investigating Per-user Time Sensitivity Of Search Topics
Jivashi Nagar, and Hussein Suleman

Room Olympia D

FUTURITY-2017

Welcome

Keynote speakers

Daniela Gîfu
Diana Trandabăț

Users Classification in an Online Community of Romanian Tourists
Mihaela Colhon, Costin Bădică

Developing a Technology Allowing (Semi-) automatic Interpretative Transcription
Daniela Gîfu, Mihaela Onofrei

Neutrosophy, Method of Uncertainties Process Analysis
Florentin Smarandache, Mirela Teodorescu

How can we reconstruct stories based on memories?
Andreea Macovei, Diana Trandabăț

Assistive technology in stock market decision making
Radu A. Ciora, Carmen M. Simion and Marius Cioca

09:00 ↓ 13:00

For all Workshops

Coffee break 11:00 ↓ 11:30

Lunch 13:00 ↓ 14:30

14:30
↓
18:30

Visit to Vergina

Vergina is a small town in northern Greece, located in the regional unit of Imathia, Central Macedonia. Since the 2011 local government reform it is part of the municipality Veroia, of which it is a municipal unit.

The town is better known for its remains of Aigai, the first capital of Macedon. It was here in 336 BC that Philip II was assassinated in the theatre and Alexander the Great was proclaimed king.

Aigai has been awarded UNESCO World Heritage Site status as "an exceptional testimony to a significant development in European civilization, at the transition from classical city-state to the imperial structure of the Hellenistic and Roman periods".

It became internationally famous in 1977, when the Greek archaeologist Manolis Andronikos unearthed the burial site of the kings of Macedon, including the tomb of Philip II, father of Alexander the Great which unlike so many other tombs had not been disturbed or looted. It is also the site of an extensive royal palace and of many rich ancient tombs. A museum now contains Philip's tomb and a new museum is being constructed for the palace and other finds. The objects and paintings found in the tombs at Vergina are also of extraordinarily high quality and historical importance.

The Greek archaeologist Manolis Andronikos became convinced that a hill called the Great Tumulus concealed the tombs of the Macedonian kings. In 1977, Andronikos undertook a six-week dig at the Great Tumulus and found four buried tombs, two of which had never been disturbed. Andronikos claimed that these were the burial sites of the kings of Macedon, including the tomb of Philip II, father of Alexander the Great (Tomb II) and also of Alexander IV of Macedon, son of Alexander the Great and Roxana (Tomb III). This view was challenged by some archaeologists, but in 2010 research based on detailed study of the skeletons, vindicated Andronikos and supports the evidence of facial asymmetry caused by a possible trauma of the cranium of the male, evidence that is consistent with the history of Philip II. More recent research by a team of Greek researchers has confirmed that the bones indeed belong to the Macedonian King Philip II.

From 1987 the burial cluster of the queens was discovered including the tomb of Queen Eurydice. In March 2014, five more royal tombs were discovered in Vergina, possibly belonging to Alexander I of Macedon and his family or to the family of Cassander of Macedon.

Some artifacts excavated at Vergina may be treated as influenced by Asian practices or even imported from Achaemenid Persia in late 6th and early 5th centuries BC, which is during the time Macedon was under the Persian sway.

Conference Venue

Grand Hotel Palace
305, Monastiriou, Thessaloniki, Greece
+30 2310 549 000
http://grandhotelpalace.gr/

Conference Dinner

Mediterranean Palace Hotel
3, Salaminos & Karatassou, Thessaloniki, Greece
+30 2311 240 400
http://mediterranean-palace.gr/

More moving around the city

www.oasth.gr/index_eng.php

City of Thessaloniki official webpage

www.visitgreece.gr/en/main_cities/thessaloniki

Emergency numbers

Police 100
Fire brigade 199
Ambulance 166

Telephone directory enquiries

Local 11888
Theatres & cinemas 1422

Taxi Companies

White Tower (Lefkos Pyrgos) +30 231 024 6104
Makedonia +30 231 055 0500
Taxi Mercedes +30 231 052 4499

Thessaloniki Airport "Macedonia"

Information Centre +30 231 047 3212, 231 047 3312
Flights information +30 231 047 3977, 231 047 3720

Thessaloniki Tourist Office

136, Tsimiski str • 546 21 Thessaloniki
Tel. Center +30 231 025 2170
Tourist Facilities +30 231 025 4834

About Thessaloniki

Thessaloniki (520 km. north of Athens) is the second largest city of Greece and the most important centre of the area. Built at the back of the Thermaïkos Gulf, it is a modern metropolis bearing the marks of its stormy history and its cosmopolitan character. Take a tour in the centre of Thessaloniki and plan to visit its nearby destinations. Also, it is worth going up to Halkidiki.

Visit Thessaloniki's Archaeological sites

- The **ancient forum** (dated to the late 2nd or the early 3rd century AD) with squares, porticoes, additional buildings and odeum (293-395 AD), the palace complex of Galerius Maximianus (4th c. AD), the thermae, the hippodrome, the temples and other monuments and moveable finds (among them mosaics of exquisite art) brought to light in excavations and surveys. In the south square, is the famous Stoa of the Idols, which was two-storeyed and lavishly decorated.
- The **Triumphal Arch of Galerius (Kamara),** built in AD 305 to commemorate his military successes in general in the eastern Roman Empire.
- The **Rotunda** (early 4th cen.) converted into a Christian church.

Visit Thessaloniki's Byzantine monuments

Thessaloniki, with its host of Byzantine monuments (due to it's significance during the Byzantine period), justifiably is considered an open-air museum of Byzantine art. Wandering through the city, it is worthwhile to see:

- The **churches** of Acheiropoietos (5th century) a three-aisled, timber-roofed basilica, the Holy Wisdom of God (Hagia Sophia) (7th century), the Panaghia (Virgin) Chalkeon (1028), Hosios David (12th century), St Panteleemon (late 13th or the early 14th century), is of four-columned cross-in-square type, Ayioi Apostoloi (1310-1314), Taxiarches (14th century), Panagouda a three-aisled basilica with significant icons, Agios Ioannis Prodromos (Nymphaion),Vlatadon monastery a 14th century foundation of which only the katholikon and two cisterns within the precinct survive, Ayios Demetrios a splendid basilica dedicated to the patron saint and protector of the city, etc.
- The **byzantine walls** of the city
- The **archaeological site** in 3 Septemvriou St., with remnants of a cemetery basilica, a martyrion and Early Christian graves.

- The **byzantine bathhouse** (late thirteenth century).
- The **Heptapyrgion castle** was raised in stages, from the early years of the Byzantine Age into the Ottoman period.

Amazing Ottoman monuments

- The **White Tower** (15th century), the hallmark of the city.
- The **Mosques of the Hamza Bey Cami** (15th century), the **Aladja Imaret Cami** (1484) and the **Yeni Cami** (1902).
- Hamams (turkish bathhouses): The **Pazar Hamam** (15th century), the **Pasha Hamam** (15th century), **Bey Hamam** (16th century), **Yeni Hamam** and the **Yahudi Hamam.**
- **Bezesteni**, a rectangular building with lead-covered domes and four entraces was built in the late fifteenth century and operated as a cloth market.

Discover neighbourhoods and focal points in the city

- The **Old City (Ano Polis),** in which many notable examples of Ottoman and traditional Macedonian architecture still stand, alongside humble dwellings put up by the refugees who reached Thessaloniki in droves, after the Greek defeat in Asia Minor, in 1922.
- The historical quarter of the **Ladadika**. In recent years, a series of interventions to rehabilitate the urban fabric have helped to enhance the Ladadika as a quarter for leisure pursuits.
- The traditional markets: the **Modiano,** which is housed in a rectangular building of 1922, with pedimented facade and glass roof; the **Kapani** or **Vlalis market; Athonos Square** and the **'Louloudadika'** (literally flower market).
- **Vasilissis Olgas Avenue,** lined with many representative Neoclassical buildings and examples of late 19thcentury eclectic architecture.
- The central **Aristotelous Square**, surrounded by monumental buildings and open to the waterfront for a width of 100 metres.

Other monuments and buildings in the city

- **Mylos** (literally mill). An old industrial complex, built in 1924, today have been remodelled to house cultural events and leisure activities, as well as the industrial buildings of the old FIX Brewery and the VILKA plant.

- **Lazarist monastery** (1886) by the monastic order of the Brothers of Mercy, and now used for cultural events.
- **Royal Theatre**
- **Thessaloniki Concert Hall.** A newly-built, magnificent yet austere, multipurpose venue for cultural and other events.
- **YMCA Building,** a building of 1924, with a mixture of Neocolonial and Byzantesque architectural elements.

Museums

It's worth seeing the **Archaelogical Museum,** the **Museum of Byzantine Culture,** the **Folk and Ethnographic Museum,** the **State Museum of Contemporary Art,** the **Teloglion Foundation of Art,** the **Thessaloniki Cinema Museum,** the **Thessaloniki Science Center and Technology Museum,** et. al.


CITYMAP
Grand Hotel Palace
Railway Station
Mediterranean Palace Hotel
Ancient forum
Heptapyrgion
Rotonda
Aristotelous Square
Kamara
Aristotle University of Thessaloniki
White Tower of Thessaloniki
Λευκός Πύργος της Θεσσαλονίκης
Archaelogical Museum
University of Macedonia
Museum of Byzantine Culture
Macedonia Airport
Ampelokipoi
Αμπελόκηποι
XIROKRINI
ΞΗΡΟΚΡΗΝΗ
ELPIDA
ΕΛΠΙΔΑ
METEORA
ΜΕΤΕΩΡΑ
Πολίχνη
Neapoli
Νεάπολη
Sykies
Συκιές
KASTRA
ΚΑΣΤΡΑ
ANO POLI
ΑΝΩ ΠΟΛΗ
Agios Pavlos
Αγ. Παύλος
Triandria
Τριανδρία
ANO TOUMP
ΑΝΩ ΤΟΥΜΠ
Thessaloniki
Θεσσαλονίκη
Evosmos
Stavroupoli
Pefka
Πεύκα
Pylaia
Πυλαία
Kalamaria
Καλαμαριά
Nea Krini
Νέα Κρήνη
Peraia
Περαία
Agia Triada
Αγ. Τριάδα
Neo Rysio
Egnatia
Tsimiski
Leof. Nikis
Monastiriou
Stathmou
Karatasou
Agiou Dimitriou
Kassandrou
Olimpiados
Akropoleos
Eptapirgiou
Leof. Andrea Papandreou
El. Venizelou
Giannitson
Geor. Gennimata
A/D PATHE
26is Oktovriou
Sapfous
Fragkon
Filippou
Ermou
Pronias
Paraskeva
Kiprou
Karaiskaki
Elpidos
Digeni
Zakka
Eleftherias
Aggelaki
Leof. Kon/nou K
©2014 Google Maps