

## REVIEW

### **Librarianship and Digital Serendipity: TPDL and CIKM 2011 Conferences**

Adam Sofronijevic<sup>1</sup>  
University of Belgrade,  
University Library "Svetozar Markovic"

Conferences are for the networking, we often hear these days. In spite of this, and one might say in addition to, conferences are still very important venue for knowledge dissemination, even in the era of virtual. Idea exchanges might often be followed by the small or big moments of inspiration. ICT related conferences are especially prone to serendipity since they are often attended by a mix of ICT professionals and professionals of a specific field. Since librarianship

is an information creation industry and therefore is all about the real world it is in a constant need of ideas that might bring its practice closer to practices of the environment. Because of this, two ICT related conferences attended recently by the author are presented in the following text focusing on unexpected possibilities they provided for librarians.

TPDL - Theory and Practice of Digital Libraries is a librarianship conference with a heavy ICT influence. It aims at bringing together

---

<sup>1</sup> [sofronijevic@unilib.bg.ac.rs](mailto:sofronijevic@unilib.bg.ac.rs)

researchers, developers, content providers and users in the field of digital libraries. The 2011 TPDL was held in Berlin, Germany from September 25th till September 29th and was organized by the Humboldt University (Berlin School of Library and Information Science, the Computer and Media Services and the Department of Computer Science). It was the 15th time this conference was held, but for the first time it was called TPDL. Before this year the conference name was ECDL - The European Conference on Research and Advanced Technology for Digital Libraries. The ECDL Conference Series started in 1997 and it has been the leading European scientific forum on digital libraries for 14 years. The proceedings of all the ECDL conferences have been published by Springer Verlag in the Lecture Notes in Computer Science (LNCS) series. These traditions continue with TPDL.

TPDL 2011 Organizing Committee General Chair was Professor Stefan Gradmann of the Humboldt University, Berlin School of Library and Information Science. This leading institution was founded in 1928 and it is the only library school at a German research university and also the only library and information science program in Germany with the right to conduct doctoral degree studies. Current research in this institution blends traditional library science with social science, computer science and linguistics. Digital libraries are important part of most studies at this institution with human-computer interaction playing an increasing role. Ongoing research projects are mostly on long-term digital preservation, electronic publishing,

copyright and open access, information retrieval and semantic network development. One great focus of activities of The Berlin School of Library and Information Science at the Humboldt University is the building of Europeana.

Conference program included various forms of scientific dissemination: paper presentations, workshops, tutorials, demos, panels, poster session and a doctoral consortium. The main reason the author attended TPDL2011 was to present his dissertation at the doctoral consortium. This specific format of dissemination of scientific results was envisioned as a general rehearsal of one's doctoral dissertation defense. Therefore only dissertations at a very late stage of preparation were eligible for presentation. Doctoral Consortium Chair Professor Milena Dobрева of University of Strathclyde, UK managed to gather a very impressive line up of experts in the role of the thesis committee: Dr. Stefan Gradmann of the Humboldt University, Germany, Dr. Giuseppina Vullo of University of Glasgow, HATII, UK and Dr. José Luis Borbinha of IST/INESC-ID, Portugal. The author's dissertation entitled "New Paradigm of Library Collaboration" was presented along with six other dissertations from various European countries. Some very productive discussions followed each presentation and experts in the role of the dissertation committee tried very hard to add constructive critique for every dissertation presented in order to further its final stage quality. Heterogeneity of themes the dissertations presented is the best example of how diversified become the digital library subject.

Experience of presenting a dissertation in such a format is something that may lead to improving quality of doctoral dissertations for any PhDc.

TPDL 2011 keynote talks by Dr. Thomas Hoffman of Google Inc., Dr. Moira Norrie of ETH Zurich and Dr. Clifford Lynch of Coalition for Networked Information were relevant and inspiring. Being a Director of Engineering at Google since 2006 Dr. Hoffman provided an insight into the ongoing research and early phase projects of the company. Machine translation was the first area of interest covered and it should be indicative for all experts designing the future translation and foreign language courses in higher education. The need for translators seems to be decreasing in many areas of communications, but the need for experts in foreign culture, of which a language is an important part, seems to be on the rise. Another issue of high importance for Google Inc. is time as a conceptual problem in page ranking. Users should be provided with the time adjusted content; they might be demanding fresh or historic content and the distinction is not always easy to be determined by looking at search keywords. Dr. Hoffman used the old saying that yesterday's news is no news and clearly stressed the importance of time – the most valuable resource we have. The next on the agenda of this multi-thematic presentation was the issue of better use of existing content and Google Inc. experience with the service Squared. Semantic web turned twice as an issue for Dr. Hoffman: first as the issue of connecting entities and later as a path from search to interface. Freebase project

started by Metaweb Company, now owned by Google Inc. was discussed as an example of the successful collaborative effort of many people working for free, following their pragmatic interest. User generated content was the final issue in this presentation. Possibilities for analysis and reuse of this kind of content seems to be indefinite and coupled with software engines provided by Google Inc. that enable pattern recognition and pattern matching paint a bleak picture of a traditional librarian's future. On the other hand they are prophetic in the sense of librarianship profession paradigm change and educational shift in library and information science curriculums around the world. The final slide of Dr. Hoffman's presentation depicted Bart Simpson writing down for hundreds of times as a punishment the same sentence on a blackboard: "I will use Google before asking dumb questions." Ms. Norrie presented an almost sci-fi technology of a digital paper and pen that has been around for years now and discussed the reasons for the lack of wider adoption and the future of this technology. Dr. Lynch discussed the foundational ideas of a digital library by looking at the contemporary projects such as Europeana and Digital Public Library of America.

In an introductory speech at the beginning of the conference Dr. Steve Griffin, the Visiting Scientist at Library of Congress and Program Director at National Science Foundation, depicted his vision of the new role of libraries. He stressed the importance of data oriented research support. Social network data analysis and availability of digital humanities data were singled out as areas of

high importance for libraries. The distinction between final document, that still needed human recovery and mass data that allows for automated recovery was clearly presented. Dr. Griffin concluded his speech by an appeal to librarians to lead the way in creating new scholarly document model.

Presentations of the day one of TPD2011 were given by some of the top members of international scientific community. Social archives, Microsoft academic research and ghost tags were some of the keywords. Most of the presentations exposed tools that might be quite useful for librarians. Walden's path project presentation described the second phase of this most useful project for anyone interested in organizing online knowledge sources for educational purposes. Reference managers were also discussed in details along with the presentation of research data on their usage in academic environment.

Poster session of TPD2011 presented many interesting projects in domains immersed in or touching on digital libraries. Encoding Diachrony: "Digital Editions of Serbian 18th-Century Texts" was the project from Serbia presented at the poster session of TPD2011 by its authors Mr. Toma Taso- vac and Ms. Natalia Ermolaev.

Days two and three of TPD2011 fol- lowed packed with presentations most inter- esting for both librarians and more ICT fo- cused audience. Invisque – Interactive Visual Search and Query Environment developed at Middlesex University, UK was presented. As a result of networking at this conference the

author of the paper provided the video link presentation of this environment for Serbian audience in November 2011 and a perspec- tive collaboration that may lead to usage of Invisque cutting edge technology in Ser- bia. Detailed discussion of annotation with presentation of research results by the team from Dalhousie University, Canada included also a theoretical framework and was inspir- ing for everyone contemplating the reading process and especially reading for learning purposes. Analysis of user research habits in a digital library environment was presented by the team from the library of Swinburne University of Technology, Australia. Data analyzed was collected in live environment and results that may lead to very pragmatic solutions in libraries have been presented. Important and inspiring results of experi- ments conducted in live library environment were presented by Dr. Elke Greifeneder of the conference hosting Berlin School of Li- brary an Information Science. These results may further the discussion, but also lead to a more constructive one in the area of research of library user's experience.

The experience of attending TPD2011 was complemented by a rich social program that offered participants the experience of contemporary Berlin and Germany.

CIKM – Conference on Information and Knowledge Management has been organized since 1992 by ACM – Association for Com- puting Machinery, which is the world's larg- est educational and scientific computing so- ciety. CIKM 2011 was held at Glasgow, UK from October 24th till October 28th 2011.

This conference aims to bring together leading researchers from database, information retrieval and knowledge management communities. A CIKM conference have always been a venue for the promotion of the collaboration among different, multiple areas of scientific research. In this regards a Best Interdisciplinary Paper award has been and still is a part of this conference series. Conference program was shaped in different formats: workshops, tutorials, paper presentations, poster session, demos, panel and the industry session. The author's reason for attending this conference has been presenting a paper at BooksOnline'11 workshop. BooksOnline workshop series has been held annually since 2008 in conjunction with CIKM conferences, except in 2009 when it was a part of ECDL (TPDL) conference. Participants of this workshop series have been some of the top leading scientists as Dr. Ray Siemens of University of Victoria, Canada and some of the top world business leaders as Mr. James Crawford, the Google Books director. Therefore an opportunity to present a part of his doctoral dissertation regarding Reading 2.0 concept has been an exceptional privilege for the author. This year's workshop BooksOnline organizers Dr. Gabriella Kazai of Microsoft Research, Dr. Carsten Eickhoff of Delft University of Technology, The Netherlands and Dr. Peter Brusilovsky of University of Pittsburgh, USA managed to involve highly inspiring keynote speakers: Dr. Adam Farquhar of The British Library and Mr. Ville Miettinen, entrepreneur and CEO of Microtask. As different as they could be both keynote speakers provided an insight into the different aspects of contemporary

landscape of reading. Other papers presented at the workshop and the subsequent discussion and breakout session proved the huge fragmentation of ideas, interests, theories and practices regarding the reading process and the text formatting.

CIKM 2011 conference gathered together researchers from some of the leading companies in ICT: Google, Microsoft, SAP, HP, Yahoo, Thompson Reuters, AT&T and some of the leading universities and institutes: MIT, Stanford, Kyoto University, University of California Irvine, University of Illinois at Urbana-Champaign, Peking University, Karlsruhe Institute of Technology, Georgia Tech and different Max-Planck-Institutes. From the variety of presented researches and topics some lend themselves as relevant for librarians. Behavioral targeting research presented by Yahoo labs team would be a great addition to any digital library catalogue if ample resources were available. Social search based on social network data analysis presented by group of researchers from Karlsruhe Institute of Technology and Mobile Advertising GmbH constitute a very important data source for users and therefore should be of utmost interest for librarians. Sessions on Classification and Evaluation, Information Filtering, Query Answering and Social Search and Text Mining might have been a specialized, just a bit more tech oriented part of any librarianship conference. A special event that has to be mentioned when presenting CIKM 2011 is Industry Event that took place on October 27th. It presented the state-of-the-art in information retrieval, knowledge management, databases, and

data mining, delivered as keynote talks by influential technical leaders who work in industry. Microsoft researcher Dr. Stephen Robertson, leading world's researcher in the area of information retrieval, opened the event with a talk on importance of the recall in a search. His thesis is that a tradeoff between the recall and precision in the search may not be necessary, but that those two may well go hand in hand. The second speaker was Mr. John Giannandrea of Google Inc. He is the leader of the Freebase project, which is an open database of knowledge which anyone can contribute to. Freebase was developed by Metaweb Technologies of which John was founder and that was acquired by Google Inc. in 2010. Other speakers at this event that run in parallel with other tracks of the conference were from Cloudera, LinkedIn, Thompson Routers, SAP, Yahoo and Yandex. The main contribution of the event was connecting the dots from the theories and the research phase to the industrial practice and users. A thing most needed in most of the contemporary libraries.

Very rich social program of CIKM 2011 included the Civic Reception at Glasgow City Chambers, where the Mayor hosted a reception for delegates. This was a very needed complement to long hours of the working part of the conference.

Participating at two major conferences in just two months and presenting papers at both was challenging, but experience gained and the insight into the contemporary scientific community was worthwhile. One may conclude that the conferences still have their

place in the mosaic of scientific dissemination methods and may be a useful shortcut in the labyrinth of ties connecting international scientific community.

Details of TPDL 2011 conference can be found at [www.tpd12011.org](http://www.tpd12011.org)

Details of CIKM 2011 conference can be found at [www.cikm2011.org](http://www.cikm2011.org)