

Knowledge Maps and Information Retrieval (KMIR)

Peter Mutschke¹, Andrea Scharnhorst², Philipp Mayr¹, Aida Slavic³, Preben Hansen⁴

¹ GESIS – Leibniz Institute for the Social Sciences, Cologne, Germany

² Royal Netherlands Academy of Arts and Sciences (KNAW) / Data Archiving and Networked Services (DANS), Amsterdam / The Hague, The Netherlands

³ UDC Consortium, The Hague, The Netherlands

⁴ University of Stockholm, Department of Computer and Systems Sciences, Sweden

Contact: peter.mutschke@gesis.org

Abstract: Information systems usually show as a particular point of failure the vagueness between user search terms and the knowledge orders of the information space in question. Some kind of guided searching therefore becomes more and more important in order to more precisely discover information without knowing the right search terms. Knowledge maps are promising tools for visualizing the structure of large-scale information spaces. However, there is no continuous knowledge exchange between the “map makers” on the one hand and the Information Retrieval (IR) specialists on the other hand. Thus, knowledge maps are still far away from being applicable for searching a digital library, due to a lack of models, explorations and user studies that properly combine insights of the two strands. The workshop aims at bringing together these two communities: experts in IR reflecting on visual enhanced search interfaces and experts in knowledge mapping reflecting on interactive visualizations of the content of a digital library collection. The focus of the workshop is to discuss the potential of interactive knowledge maps for information exploration.

Keywords: interactive information retrieval, information seeking behavior, knowledge mapping, information visualization.

Motivation, Goals, Objectives, Outcome

The success of an information system depends mainly on its ability to properly support interaction between users and information. Current information systems, however, show as a particular point of failure the vagueness between user search terms and the knowledge orders of the information space in question (Mayr et al. 2008, Mutschke et al. 2011). Studies in interactive information seeking behavior have confirmed that the ability to browse an information space and observe similarities and dissimilarities between information objects is crucial for accidental encountering and the creative use of information (Nicholas et al. 2004, Westerman et al. 2005). This is in particular true for heterogeneous information spaces within the open Web. Some kind of guided searching therefore becomes more and more important in order to precisely discover information without knowing the right search terms. Yet, this seems to remain the weakest point of interactive information systems (Ford 2000, Foster 2004, Tang 2007).

Knowledge mapping encompasses all attempts to use visualizations to gain insights into the structure and dynamics of large-scale information spaces. Knowledge maps can take the form of network visualizations, treemaps or specific, geographic map like arrangements of knowledge structures (cf. Börner et al. 2003, Shiffrin/Börner 2004, Börner 2010, Klavans/Boyack 2010, Skupin et al. 2013, Sahal et al. 2013, Boyack/Klavans 2013). As an activity performed in very different disciplines – and often independent from each other – it stands in line with the dominance of *the visual* in our culture (Manovich 2009). Knowledge maps of digital library collections are promising navigation tools through knowledge spaces, but – to the best of our knowledge – still far away from being applicable for searching digital libraries. Most maps are made for special purposes, are static, and usually not

interactive (Akdag Salah et al. 2012). In interactive information systems the use of visual elements to enhance information seeking and discovery is a recurring research issue. However, not much of the experiences made in knowledge mapping have ever led to novel interfaces to digital libraries and collections (Börner/Chen 2001), nor is there a stable and continuous knowledge exchange between the “map makers” on the one hand and the Information Retrieval (IR) specialists on the other hand. Thus, there is also a lack of models that properly combine insights of the two strands, which are driven by quite different epistemic perspectives.

Our workshop aims at bringing together these two communities: experts in IR reflecting on visual enhanced search interfaces and experts in knowledge mapping reflecting on visualizations of the content of a collection that might also present – visually – a context for a search term. The focus of the workshop is to discuss the potential of interactive knowledge maps for information exploration and to create a common ground for experiments aiming at the incorporation of knowledge maps into IR models at the level of the user interface. Our interests include interactive IR, information seeking behavior, knowledge mapping, science modelling, information visualization, and digital libraries. The workshop is closely related to the COST action KNOWeSCAPE¹ (Analyzing the dynamics of information and knowledge landscapes) which aims at implementing new navigation and search strategies based on insights of the complex nature of knowledge spaces as well as visualization principles for knowledge maps.

The proposed workshop can tie up to the success of the first international workshop on “Knowledge Maps and Information Retrieval (KMIR)”, held as part of the DL 2014 in London (see <http://www.gesis.org/en/events/events-archive/conferences/kmir2014/>).

The long-term research goal is to develop and evaluate new approaches for combining knowledge mapping and IR. More specifically, we address questions such as:

- What are appropriate interactive knowledge maps for IR systems?
- How can knowledge maps be utilized for information seeking purposes?
- How to locate an information need on a knowledge map?
- How can interaction with knowledge maps be transformed into IR tasks?
- Can knowledge maps improve searching in large, in particular heterogeneous, cross-language, cross-domain information spaces?
- And the other way around: Can insights from IR also improve knowledge mapping itself?

The availability of new IR test collections that contain citation and bibliographic information like the iSearch collection (see Lykke et al. 2010) or the ACL collection (Ritchie et al. 2006) could deliver an interesting playground for developing or evaluating combined models of IR and knowledge mapping for scholarly searching.

Topical Outline

To support the previously described goals the workshop topics include (but are not limited to) the following:

- Knowledge maps for digital libraries
- Knowledge orders of information spaces
- Information seeking behavior
- Information discovery
- Interactive IR systems
- Human-computer IR
- Information visualization in IR
- Visual interfaces to information systems

¹ <http://knowescape.org/>

- Browsing and navigating information spaces
- Task based user modelling, interaction and personalization
- Evaluation of interactive IR systems.

Format and Structure

We propose a half-day workshop that brings together researchers in IR and knowledge mapping. The workshop will start with an inspirational keynote given by a major researcher in knowledge mapping in order to kick-start thinking and discussion on the workshop topic. This will be followed by a number of paper presentations in the following format: 20 minutes presentations for full papers, 10 minutes presentations for short papers. The workshop will conclude with a round-robin discussion of how to progress in enhancing IR with knowledge mapping approaches.

Audience

The workshop addresses researchers from IR as well as knowledge mapping having an interest in visually enhanced approaches for searching and information discovery in large-scale information spaces. Thus, the intended audience for the workshop includes researchers interested in all aspects of knowledge mapping and search interfaces to digital libraries such as information visualization, in particular visualization of knowledge structures in large-scale information spaces, interactive information retrieval, information seeking behavior, search interface design, human-computer interaction, interfaces to information for novices and experts, retrieval and browsing, studies of human factors in networked information, theoretical models of information interaction and organization, user behavior and modelling, as well as usability evaluation.

Given the experience from the KMIR workshop at DL 2014 we expect 10-20 submissions and about 20-30 attendees. The communication with the authors will be managed via EasyChair and email. Attendees will be identified via email registration. Notifications to attendees will be done via email and the workshop page.

Types of Submissions

Full Papers (6 to 8 pages): Full papers, describing advanced or completed work

Short Papers (4 pages): Position papers or work in progress

Poster and Demonstrations (2 pages): Poster and presentation of systems or prototypes

Submissions have to follow the Springer LNCS Author Guidelines² and should be submitted as PDF files to EasyChair. All submissions will be reviewed by at least two independent reviewers. At least one author per paper needs to register for the workshop and attend the workshop to present the work. In case of no-show the paper (even if accepted) will be deleted from the proceedings AND from the program.

Output

Workshop proceedings will be deposited online in the CEUR workshop proceedings publication service (ISSN 1613-0073) . This way the proceedings will be permanently available and citable (digital persistent identifiers and long term preservation). All accepted workshop papers will be published in the workshop proceedings.

² <http://www.springer.com/computer/lncs?SGWID=0-164-6-793341-0>

Previous workshop

The proposed workshop can tie up to the success of the first international workshop on “Knowledge Maps and Information Retrieval KMIR”, held as part of the International Conference on Digital Libraries 2014 - ACM/IEEE Joint Conference on Digital Libraries (JCDL 2014) and International Conference on Theory and Practice of Digital Libraries (TPDL 2014), London, 11th September 2014 (<http://www.dl2014.org/>). The KMIR workshop at DL 2014 was organized by the authors of this workshop proposal, received 12 submissions and attracted 26 registered participants. See workshop page at <http://www.gesis.org/en/events/events-archive/conferences/kmir2014/>.

The workshop gave floor to one keynote by André Skupin (San Diego State University, USA) and ten paper presentations spanning a wide range of questions around knowledge maps to be implemented as search interfaces for digital libraries. Skupin’s keynote on “Managing Domain Knowledge: Ontology, Visualization, and Beyond” focused on high-dimensional models of so-called Bodies of Knowledge (BoKs) that can be leveraged within interactive visualization. What followed was a search for best ways to create knowledge maps from different kind of data by different kind of visualization methods and metaphors (such as graph visualization of relational data, radial histograms of multi-faceted data, or visualization of knowledge structures by geographic metaphors) as well as perspectives to use knowledge maps as search interfaces. One thread through the presentations was to provide the user with an overview. Another shared topic concerned issues of design, such as fonts, colors, or use of the space. See workshop proceedings: <http://ceur-ws.org/Vol-1311/>.

During these previous workshops it became obvious that there is a growing awareness that interactive knowledge maps as search interfaces to digital libraries could be very beneficial for information seeking purposes, and provide a novel research area for both communities, interactive IR as well as knowledge mapping.

Proposed Program Committee (to be confirmed)

Alkim Almila Akdag Salah, University of Amsterdam (The Netherlands)

Nicholas J. Belkin, Rutgers University (USA)

Katy Börner, Indiana University (USA)

Kevin W. Boyack, SciTech Strategies, Inc. (USA)

Edward A. Fox, Virginia Polytechnic Institute and State University (USA)

Norbert Fuhr, University of Duisburg-Essen (Germany)

Daniel Hienert, GESIS – Leibniz Institute for the Social Sciences (Germany)

Peter Ingwersen, Royal School of Library and Information Science (Denmark)

Claus-Peter Klas, University of Hagen (Germany)

Birger Larsen, Royal School of Library and Information Science (Denmark)

Vivien Petras, Humboldt-Universität zu Berlin (Germany)

André Skupin, San Diego State University (USA)

Catherine L. Smith, Kent State University (USA)

Howard D. White, Drexel University (USA)

References

Akdag Salah, A.A.; Scharnhorst, A.; Ten Bosch, O.; Doorn, P.; Manovich, L.; Salah, A.A.; Chow, J. (2012): Significance of Visual Interfaces in Institutional and User-Generated Databases with Category

Structures. Proceedings of the second international ACM workshop on "Personalized access to cultural heritage (PATCH)". ACM Multimedia Conference, Nara, Japan.

Börner, K.; Chen, C. (2001): Visual interfaces to digital libraries - its past, present, and future. JCDL'01 Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries.

Börner, K.; Chen, C.; Boyack, K. W. (2003): Visualizing knowledge domains. *Annual Review of Information Science and Technology*, 37, 179-255.

Börner, K. (2010): *The Atlas of Science*. MIT Press.

Boyack, K. W.; Klavans, R. (2013, forthcoming): Creation of a highly detailed, dynamic, global model and map of science. *Journal of the American Society for Information Science and Technology*.

Ford, N. (2000): Cognitive Styles and Virtual Environments. *Journal of the American Society for Information Science*, 51(6), 543-557.

Foster, A.E. (2004): A Nonlinear Model of Information Seeking Behaviour. *Journal of the American Society for Information Science and Technology*, 55(3), 228-237.

Klavans, R.; Boyack, K. W. (2010): Toward an objective, reliable, and accurate method for measuring research leadership. *Scientometrics*, 82(3), 539-553.

Lykke, M.; Larsen, B.; Lund, H.; Ingwersen, P. (2010): "Developing a Test Collection for the Evaluation of Integrated Search." *Advances in Information Retrieval*, edited by Gurrin et al., 5993:627-630. *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer.

Manovich, L. (2009): *How to Follow Global Digital Cultures, or Cultural Analytics for Beginners*". *Deep Search: The Politics of Search Beyond Google*, edited by Felix Stalder and Konrad Becker, 198-211 (Wien: Studien Verlag / Transaction Publishers, 2009).

Mayr, P.; Mutschke, P.; Petras, V. (2008): Reducing semantic complexity in distributed digital libraries: Treatment of term vagueness and document re-ranking. *Library Review*, 57(3), 213-224.

Mutschke, P.; Mayr, P.; Schaer, P.; Sure, Y. (2011): Science models as value-added services for scholarly information systems. In: *Scientometrics*, 89 (1), 349-364.

Nicholas, D. et.al. (2004): Reappraising Information Seeking Behaviour in a Digital Environment: bouncers, checkers, returners, and the like, *Journal of Documentation*, 60(1), 24-43.

Ritchie, A.; Teufel, S.; Robertson, S. (2006): Creating a Test Collection for Citation-based IR Experiments. Proceedings of the Main Conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics, 391-398.

Sahal, A.A.; Wyatt, S.; Passi, S.; Scharnhorst, A. (2013): Mapping EINS - An exercise in mapping the Network of Excellence in Internet Science. Conference Proceedings of the First International Conference on Internet Science, Brussels, 75-78.

Skupin, A.; Biberstine, J.R.; Börner, K. (2013): Visualizing the Topical Structure of the Medical Sciences: A Self-Organizing Map Approach. *PLoS ONE* 8 (3): e58779.

Shiffrin, R.; Börner, Katy (2004): Mapping Knowledge Domains. Proceedings of the National Academy of Sciences of the United States of America, Vol. 101(Suppl. 1).

Tang, M. (2007): Browsing and Searching in a Faceted Information Space: a naturalistic study of PubMed users' interaction with a display tool, *Journal of the American Society for Information Science and Technology*, 58(13), 1998-2006.

Westerman, S.J.; Collins, J.; Cribbin, T. (2005): Browsing a Document Collection Represented in Two- and Three-dimensional Virtual Information Space, *International Journal of Human-Computer Studies*, 62(6), 713-736.

Short bios of the proposers

Peter Mutschke: Peter Mutschke is senior researcher at GESIS – Leibniz Institute for the Social Sciences (Cologne) and acting head of the GESIS department “Knowledge Technologies for the Social Sciences”. His research focuses on information retrieval, network analysis and Social Web. He worked in a number of national and international research projects such as DAFFODIL (Distributed Agents for User-Friendly Access of Digital Libraries), INFOCONNEX (interdisciplinary information network for Social Sciences, Education Science and Psychology), IRM (value-added search services), the DELOS/NSF Working Group on reference models for digital libraries, and the EU-funded projects WeGov (Where eGovernment meets the eSociety) and SENSE4US (Data insights for policy makers & citizens). Currently, he is involved in major national and European research networks such as the COST Action KNOWeSCAPE (“Analyzing the dynamics of information and knowledge landscapes”) and the research alliance “Science 2.0” of the German Leibniz Association. For both research networks Peter is member of the management committee. Peter is author of a number of research articles, member of a number of international program committees, was co-organizer of the workshop “Combining Bibliometrics and Information Retrieval” at ISSI 2013 and “Bibliometric-enhanced Information Retrieval (BIR)” at ECIR 2014 and main organizer of the “Knowledge Maps and Information Retrieval (KMIR)” workshop at DL 2014.

Andrea Scharnhorst: Dr. Andrea Scharnhorst is head of e-Research at the Data Archiving and Networked Services (DANS) institution in the Netherlands, a large digital archive for research data primarily from the social sciences and humanities. She also coordinates the computational humanities program at the e-humanities group of the Royal Netherlands Academy of Arts and Sciences (KNAW) in Amsterdam. Starting in physics (Diploma in Statistical Physics) she received her PhD in philosophy of science. She co-edited books in the Springer Series of Understanding Synergetics on Innovation Networks (with A. Pyka) and recently on Models of Science Dynamics (with K. Boerner and P. van den Besselaar). Her current work in the information sciences is devoted to the development of *knowledge maps* for library collections, research data bases and online knowledge spaces such as Wikipedia. Andrea was co-organizer of the workshop “Combining Bibliometrics and Information Retrieval” at ISSI 2013 and co-organizer organizer of the “Knowledge Maps and Information Retrieval (KMIR)” workshop at DL 2014.

Philipp Mayr: Philipp Mayr is a postdoctoral researcher and team leader at the GESIS – Leibniz Institute for the Social Sciences department “Knowledge Technologies for the Social Sciences”. Philipp is a graduate of the Berlin School of Library and Information Science at the Humboldt University Berlin where he finished his doctoral research in 2009. Philipp is a member of the European NKOS network and published widely in the areas Informetrics, Information Retrieval and Digital Libraries. He is member of the editorial board of the journals *Scientometrics* and *Information - Wissenschaft & Praxis*. His research interests include non-textual ranking in digital libraries, bibliometric methods, evaluation of information systems and knowledge organising systems, as well as applied informetrics. Philipp was the main organizer of the workshop “Combining Bibliometrics and Information Retrieval” at ISSI 2013 and the workshop “Bibliometric-enhanced Information Retrieval (BIR)” at ECIR 2014 and co-organizer organizer of the “Knowledge Maps and Information Retrieval (KMIR)” workshop at DL 2014.

Aida Slavic: Aida Slavic is the editor-in-chief of the Universal Decimal Classification (UDC) and works on the development and the maintenance of this knowledge classification scheme for the UDC Consortium (The Hague). She holds a PhD in library and information studies from the University College London. Her doctoral research was on the use of knowledge classification in the networked environment and her research interest is in the field of knowledge organization, classification, metadata and semantic technologies. Aida is a member of the Scientific Academic Committee of the International Society for Knowledge Organization (ISKO) and member of the editorial board of the *Knowledge Organization Journal*. She is a co-leader of the Workgroup 1: Phenomenology of Knowledge Spaces of the COST Action KNOWeSCAPE and co-organizer of the first project workshop

Knowledge Orders and Science. Aida was co-organizer organizer of the “Knowledge Maps and Information Retrieval (KMIR)” workshop at DL 2014.

Preben Hansen: Preben Hansen is Associate Professor at the Department of Computer and Systems Sciences at the University of Stockholm. His research interests address collaborative information seeking and retrieval, task-based and context-based information access, patent IR, evaluation issues and user studies in real-life environments. Preben was involved in a number of EU-projects (Promise, ASSETS, Companions, DELOS NoE and Clarity). A particular focus of his research is on empirical studies of users and use of interactive information access systems. Preben was co-organizer of the “Knowledge Maps and Information Retrieval (KMIR)” workshop at DL 2014.