

Editorial for the 15th European Networked Knowledge Organization Systems Workshop (NKOS 2016)

Philipp Mayr¹, Douglas Tudhope², Koraljka Golub³, Christian Wartena⁴, and Ernesto William De Luca⁵

¹ GESIS - Leibniz-Institute for the Social Sciences, Cologne, Germany
philipp.mayr@gesis.org

² Hypermedia Research Group, Faculty of Computing, Engineering and Science, University of South Wales, UK
douglas.tudhope@southwales.ac.uk

³ Department of Library and Information Science, School of Cultural Sciences, Faculty of Arts and Humanities, Linnaeus University, Sweden
Koraljka.golub@lnu.se

⁴ Hochschule Hannover, Abt. Information und Kommunikation, Hannover, Germany
⁵ Georg Eckert Institute, Braunschweig, Germany

1 Introduction

We have proposed a full-day workshop of research projects and development related to next-generation Networked Knowledge Organization Systems/Services (NKOS) in digital libraries. This workshop builds on the well-attended NKOS workshops at previous ECDL, TPDL, JCDL conferences (see NKOS website⁶ for details).

Knowledge Organization Systems (KOS), in the form of classification systems, thesauri, lexical databases, ontologies, and taxonomies, play a crucial role in digital information management and applications generally. Carrying semantics in a well-controlled and documented way, Knowledge Organisation Systems serve a variety of important functions: tools for representation and indexing of information and documents, knowledge-based support to information searchers, semantic road maps to domains and disciplines, communication tool by providing conceptual framework, and conceptual basis for knowledge based systems, e.g. automated classification systems. New networked KOS (NKOS) services and applications are emerging, and we have reached a stage where many KOS standards exist and the integration of linked services is no longer just a future scenario [4].

This editorial describes the workshop outline and overview of presented papers at the 15th European Networked Knowledge Organization Systems Workshop (NKOS 2016) in Hannover, Germany.

⁶ <http://hypermedia.research.southwales.ac.uk/kos/nkos/>

2 Workshop outline

The NKOS workshop at TPDFL 2016⁷ was in collaboration with the German ISKO⁸. In the workshop we explored 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 provided an opportunity to discuss projects, research and development activities, evaluation approaches, lessons learned, and research findings. A further objective was to systematically engage in discussions in common areas of interest with selected related communities and to investigate potential co-operation.

The workshop allowed projects to report results, newcomers to interact with established people in the field and discussion of topical issues, requiring consensus or coordination, including standards efforts, to take place. Thus previous workshops have seen focused discussion on early drafts of BSI and ISO KOS standards, the W3C SKOS standard, the interface between traditional Library Science vocabularies and Semantic Web efforts, KOS linked data, social tagging and its relation to established vocabularies, KOS metadata and the different types of KOS. The TPDFL venue affords participation by KOS researchers and developers from different perspectives (reflecting the different conference threads), such as KOS design and construction, API and service developers, user oriented issues, management of KOS in registries. In 2015 we have published a special issue on Networked Knowledge Organisation Systems and Services in the international journal IJDL [4]. This special issue evolved from the 13th Networked Knowledge Organization Systems (NKOS) workshop held at the joint Digital Libraries conference 2014 in London.

The workshop 2016 has three themes as the main focus, together with topical presentations arising from the workshop call for papers.

1. KOS Alignment. KOS alignment or terminology mapping plays a vital role in NKOS for many years and fits very well to the general theme Overcoming the Limits of Digital Archives of TPDFL 2016. 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.
2. 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.
3. 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

⁷ <https://at-web1.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2016/>

⁸ <http://isko-de.org/>

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?

4. KOS-based recommender systems. The suggestion of the right meaningful concepts is a mission critical phase for searchers in modern DL.
5. Meaningful Concept Display and Meaningful Visualization of KOS.
6. Standards developments.
7. Evaluation of KOS-based systems methods and practical experience.
8. KOS in e-Research metadata contexts - intersection between research data, KOS, Semantic web.
9. Social tagging. What is the role of social tagging and informal knowledge structures versus established KOS? (How) can tagging be guided and informed by KOS?
10. Users interaction with KOS in the online environment.
11. KOS and learning. What is required to use KOS effectively to convey meaning, to assist users to express their information needs to assist in sense-making and learning?
12. Multilingual and Interdisciplinary KOS applications and tools.
13. Specific domains, such as environmental, medical, new application contexts, etc.

Supporting material for the workshop would, following standard NKOS practice, be available via the NKOS website. This would include abstracts of presentations, information on participants, list of resources, projects and plans for the workshop before the conference. After the workshop, copies of presentations will be made available on the website and via the main NKOS network website.

Authors presenting significant results at the workshop will be encouraged to submit papers for consideration in future issues of the International Journal on Digital Libraries, the New Review of Hypermedia and Multimedia and Knowledge Organization (The ISKO Journal) depending on the candidate papers.

3 Overview of the papers

The workshop featured an introduction and four paper sessions. The NKOS organizers have accepted 5 long and 2 short papers for presentation. All papers are included in the proceedings. In the following we will shortly summarize each workshop paper.

In their paper "Analyzing the research output presented at European Networked Knowledge Organization Systems workshops (2000-2015)" [5], Momeni and Mayr performed a network analytic study of the research output which has been presented at the past NKOS workshop and special issues on NKOS. They

provide with their paper an open dataset, the "NKOS bibliography", which includes 14 workshop agendas (ECDL 2000-2010, TPDFL 2011-2015) and 4 special issues on NKOS (2001, 2004, 2006 and 2015). The dataset covers 171 papers with 218 distinct authors in total. Momeni and Mayr visualized the co-authorship network of the NKOS community and calculated typical network measures like degree and betweenness. They found that 15% (with degree=0) of authors had no co-authorship with others and 53% of them had a maximum of 3 cooperations with other authors. 32% had at least 4 co-authors for all of their papers. In their preliminary analysis of the NKOS co-authorship network they concentrate on the results from the largest network component which includes most of the central NKOS authors.

Voß [8] deals with knowledge organization systems (KOSs) at two levels. The main topic of the paper is an investigation into the possibilities to extract a knowledge organization system from the class structure of Wikidata. This is exemplified by extracting a class hierarchy of different types of KOSs. The paper discusses the differences between a carefully designed system and the Wikidata structure that is growing bottom-up driven by the use of the classes in Wikipedia and other Wikimedia projects. Voß proposes an iterative process of extracting and analyzing hierarchies and modifying wikidata based on the insights from the analysis.

Gnoli, Pusterla, Bendiscioli and Recinella [1] describe a practical application of a knowledge organization system. The university library of Pavia is a merger of a number of smaller libraries, each of which having their own classification system and tradition. In order to enable a uniform search interface to all libraries a system was developed using the Dewey Decimal Classification (DDC) as its main classification system along with mappings to other classification systems. The DDC is widely used in Pavia and in other Italian scientific libraries, and therefore is an obvious choice. However, the strong monohierarchical organization and the lack of associative relations makes it unsuited for a search and browsing interface. The authors therefore add associative relations and show how these can be used effectively in the user interface of the joint catalogues of the University Library of Pavia.

Manguinhas et al. in their paper titled Linking subject labels in Cultural Heritage Metadata to MIMO vocabulary using CultuurLink [3] discuss challenges of harmonising subject metadata values from different multilingual knowledge organisation systems (KOS), as part of the Europeana Sounds project. The approach they take is an existing KOS alignment tool, CultuurLINK, which combines automated and manual mapping approaches. The sample included a total of 10,406 metadata records with the subject on musical instruments terms from 4 data providers: the British Library, Centre de Recherche en Ethnomusicologie, Maison Méditerranéenne des Sciences de l'homme, and the Netherlands Institute of Sound and Vision. The target KOS was the multilingual Musical Instruments Museums Online. The data providers found that applying a simple matching technique was enough to identify more than 50% of all possible alignments for

musical instruments, leading to a consensus that they were able to understand and work with the CultuurLINK tool with a good level of success.

Veira, De Brito, El Hadi and Zumer [7] seek to transform traditional library cataloguing by proposing to draw on images as possible indexing elements rather than keywords or descriptors. This effort is located within the FRSAD (Functional Requirements for Subject Authority Data) model and the associated body of work that elaborates what is meant by aboutness. Within FRSAD, a work may contain subject(s) (thema), which in turn have appellation(s) nomen. Key images can be seen as nomens within the FRSAD conceptual model. In the authors' imagetic model for a proposed iOPAC, key images can be tagged with concepts from thesauri or from folksonomies and represented via standards, such as XML and RDF. An illustrative example is discussed using the Library of Congress Classification Scheme. The authors suggest that indexing with images has the potential to be more intuitive and may be helpful for specific needs such as deafness and illiteracy and for social and cultural interoperability more generally.

The role of KOS in Natural Language Processing is taken up by Husevag [2] in an investigation of Named Entity Recognition (NER) in both subtitles and metadata records of TV programs, with a view to informing automatic subject indexing of subtitles based on NER. The paper explores the density and the frequency of different types of named entities in Norwegian language subtitle texts from the archives of the Norwegian Broadcasting Corporation. The paper begins with a useful review of relevant literature on NER and the origins of the TORCH project. The outcomes of the new investigation of news items are compared with previous project results on book and media TV programs. The investigation analysed the relative distribution of the different Named Entities (including Persons, Organizations, Locations, Work, Event and Other) in different TV programs and genres and in subtitles vs metadata. Similar results were found for the news programs compared to the earlier Norwegian NER research. The most frequent entities types (in both subtitles and news metadata) were personal names, geographical names and names of organizations. The analysis found that density of NEs in metadata records is higher than in subtitles. NEs with high frequencies in the subtitles are more likely to be mentioned in the metadata records.

The paper "Re-designing Online Terminology Resources for German Grammar" by Suchowolec et al. [6] is a project report presenting 'Grammis' which is a specialist hypertext resource hosted by the Institute for German Language (IDS) in Mannheim. The authors describe the state of the art of their resource and address the re-design principles of the terminology management discussing different options for evaluating new tools, looking into commercial native terminology, native thesaurus, and hybrid solutions. Suchowolec et al. propose questions regarding ramification of Linked Open Data and Semantic Web approaches for their re-design decisions.

4 Acknowledgment

We are indebted to the referees who contributed to the review process of this workshop⁹ and previous NKOS events¹⁰.

References

1. Gnoli, C., Pusterla, L., Bendiscioli, A., Recinella, C.: Classification for collections mapping and query expansion. In: Proc. of the 15th European Networked Knowledge Organization Systems Workshop (NKOS 2016). CEUR-WS.org, Hannover, Germany (2016)
2. Husevåg, A.S.R.: Named entities in indexing: A case study of TV subtitles and metadata records. In: Proc. of the 15th European Networked Knowledge Organization Systems Workshop (NKOS 2016). CEUR-WS.org, Hannover, Germany (2016)
3. Manguinhas, H., Charles, V., Isaac, A., Miles, T., Lima, A., Néroulidis, A., Ginouvès, V., Atsidis, D., Hildebrand, M., Brinkerink, M., Gordea, S.: Linking subject labels in Cultural Heritage Metadata to MIMO vocabulary using CultuurLink. In: Proc. of the 15th European Networked Knowledge Organization Systems Workshop (NKOS 2016). CEUR-WS.org, Hannover, Germany (2016)
4. Mayr, P., Tudhope, D., Clarke, S.D., Zeng, M.L., Lin, X.: Recent applications of Knowledge Organization Systems: introduction to a special issue. *International Journal on Digital Libraries* 17(1), 1–4 (2016), <http://link.springer.com/10.1007/s00799-015-0167-x>
5. Momeni, F., Mayr, P.: Analyzing the research output presented at European Networked Knowledge Organization Systems workshops (2000-2015). In: Proc. of the 15th European Networked Knowledge Organization Systems Workshop (NKOS 2016). CEUR-WS.org, Hannover, Germany (2016)
6. Suchowolec, K., Lang, C., Schneider, R.: Re-designing Online Terminology Resources for German Grammar. In: Proc. of the 15th European Networked Knowledge Organization Systems Workshop (NKOS 2016). CEUR-WS.org, Hannover, Germany (2016)
7. Vieira, S., de Brito, M., Mustafa El Hadi, W., Zumer, M.: Developing imaged KOS with the FRASAD model: a conceptual methodology. In: Proc. of the 15th European Networked Knowledge Organization Systems Workshop (NKOS 2016). CEUR-WS.org, Hannover, Germany (2016)
8. Voß, J.: Classification of Knowledge Organization Systems with Wikidata. In: Proc. of the 15th European Networked Knowledge Organization Systems Workshop (NKOS 2016). CEUR-WS.org, Hannover, Germany (2016)

⁹ <https://at-web1.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2016/proposal.html>

¹⁰ <http://hypermedia.research.southwales.ac.uk/kos/nkos/>