Editorial for the 17th European Networked Knowledge Organization Systems Workshop (NKOS 2017)

Philipp Mayr[1], Douglas Tudhope[2], Koraljka Golub[3], Christian Wartena[4], and Ernesto William De Luca[5]

[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 website[6] 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 Organization 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 [5].

This editorial describes the workshop outline and overview of presented papers at the 17th European Networked Knowledge Organization Systems Workshop (NKOS 2017)[7] which was held during the TPDL 2017 Conference in Thessaloniki, Greece.

---

[6] http://hypermedia.research.southwales.ac.uk/kos/nkos/
[7] https://at-web1.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2017/

2   Workshop outline

The NKOS workshop at TPDL 2017 was in collaboration with the German ISKO[8]. 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 TPDL 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.

As the workshop at TPDL 2016 [6], the workshop at TPDL 2017 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. 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 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

---

[8] http://isko-de.org/

made complementary? What should be the focus of annotation, if full text retrieval is available?

3 Overview of the papers

The workshop featured an introduction and four paper sessions. The NKOS organizers have accepted 7 regular papers for presentation. All papers are included in the proceedings to be published with CEUR. In the following we will shortly summarize each workshop paper.

Stahn, Dahlberg, and De Luca in their paper titled "Knowledge Organisation for Digital Libraries" [8] discuss the highly relevant problematics of integrating knowledge organization systems in the Semantic Web for improving retrieval. They describe conversion of two major lexical resources, Information Coding Classification (ICC) and MultiWordNet, into the EuroWordNet RDF/OWL format. Then, they present and discuss the way to map the two lexical resources. Finally, they show how a domain-specific Knowledge Organisation System (KOS) can be enriched with relevant information from the ICC and EuroWordNet, using Semantic Web standards; with the purpose of improving information retrieval which they demonstrate on a use case.

Neubert investigates the possibilities to use Wikidata as a hub for the alignment of authority files, especially for author names [7]. The *ZBW – Leibniz Information Centre for Economics* in Germany uses two different personal name authorities in economics. For a small part of these authority files a mapping exists. The existing mapping can be stored easily in Wikidata by assigning two identifiers to the same item. Neubert shows how a number of simple, semi-automatic operations in combination with various existing mappings and clever heuristics could be used to extend both, the original mapping and the coverage of identifiers form each of the authority files for persons in Wikidata. The ZBW expect that other organizations might benefit from making the mapping freely available, but also hope that more organizations will contribute to the maintenance and extension of the mapping. Finally, the paper "Wikidata as a linking hub for knowledge organization systems?" discusses the possibilities to follow a similar approach for other types of data, like geographic names or concepts. The conclusion of Neubert is, that this is much more difficult than using the items of Wikidata to connect personal name authorities for various reasons. However, there are a lot of potentials and many techniques used for the person authorities can be used again.

Lüschow and Wartena present an implementation of a kNN algorithm on automatic classification of medical documents in their paper titled "Classifying Medical Literature Using k-Nearest-Neighbours Algorithm" [4]. They use close to 20,000 catalogue records of the Hanover Medical School Library. From the records they extracted existing classes from universal classification systems such as Dewey Decimal Classification (DDC) and used them to determine one of the total of 514 relevant classes of the National Library of Medicine (NLM) classifi-

cation system. Using the k-nearest-neighbours (kNN) algorithm, they achieved accuracy level of 55.7%.

Stellato et al. [9] introduce VocBench 3, the third chapter in the history of the open source collaborative web platform for thesauri development. The paper describes the evolution of VocBench, from its first stirrings inside FAO as the Agrovoc Concept WorkBench to the new umbrella provided by the ISA2 programme of the EU. Particular relevance is given to the set of new features and improvements brought by the new version, by describing the criticalities observed in the past version, the demand of the users, the objectives and requirements for the new version and how they contributed to its realization.

In the paper "Project Coli-conc: Mapping Library Knowledge Organisation Systems" [1], Balakrishnan and Akter present their KOS mapping project Coli-conc which develops different tools for storing (e.g. JSKOS) and building up of terminology mappings (e.g. Cocoda - mapping tool). The tools are also capable of supporting the automatic generation of mappings and integrating the results into the Coli-conc infrastructure.

The report "The Publication of a Knowledge Organization Classification System as a Linked Data Vocabulary" [3] by Lavrenova and Pavlov presents an implementation of the Russian Library Bibliographic Classification as a Networked Knowledge Organization System. The project goal is to support the content and indexing; enable information search based on structured knowledge; provide a semantic road map for the user; offer the tools to ensure the system's interoperability with other resources of the same nature (Linked Data Vocabularies) in the global network environment.

In the paper "Aquacold - a crowdsourced query understanding and query construction tool for the linked data web" [2], Collis and Frommholz introduce Aquacold (Aggregated Query Understanding And Construction Over Linked Data), a novel Linked Data query tool. Aquacold provides a simple browsing interface to guide users in the query formulation phase and allows to store a natural language representation of the underlying SPARQL query. The process of interacting with Aquacold includes: selecting a Linked Data source, building the results grid using filters, labeling the results, generating a template, searching and voting the results. The authors present an overview of the system and discuss an early prototype.

4   Acknowledgment

---

[9] https://at-web1.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2017/proposal.html
[10] http://hypermedia.research.southwales.ac.uk/kos/nkos/

References

1. Balakrishnan, U., Akter, M.: Project Coli-conc: Mapping Library Knowledge Organisation Systems. In: Proc. of the 17th European Networked Knowledge Organization Systems Workshop (NKOS 2017). CEUR-WS.org, Thessaloniki, Greece (2017)
2. Collis, N., Frommholz, I.: Aquacold – a crowdsourced query understanding and query construction tool for the linked data web. In: Proc. of the 17th European Networked Knowledge Organization Systems Workshop (NKOS 2017). CEUR-WS.org, Thessaloniki, Greece (2017)
3. Lavrenova, O., Pavlov, V.: The Publication of a Knowledge Organization Classification System as a Linked Data Vocabulary. In: Proc. of the 17th European Networked Knowledge Organization Systems Workshop (NKOS 2017). CEUR-WS.org, Thessaloniki, Greece (2017)
4. Lüschow, A., Wartena, C.: Classifying Medical Literature Using k-Nearest-Neighbours Algorithm. In: Proc. of the 17th European Networked Knowledge Organization Systems Workshop (NKOS 2017). CEUR-WS.org, Thessaloniki, Greece (2017)
5. 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`
6. Mayr, P., Tudhope, D., Golub, K., Wartena, C., Luca, E.W.D.: Editorial for the 15th european networked knowledge organization systems workshop (NKOS 2016). In: Proceedings of the 15th European Networked Knowledge Organization Systems Workshop (NKOS 2016) co-located with the 20th International Conference on Theory and Practice of Digital Libraries 2016 (TPDL 2016), Hannover, Germany, September 9, 2016. pp. 1–6 (2016), `http://ceur-ws.org/Vol-1676/editorial.pdf`
7. Neubert, J.: Wikidata as a linking hub for knowledge organization systems? Integrating an authority mapping into Wikidata and learning lessons for KOS mappings. In: Proc. of the 17th European Networked Knowledge Organization Systems Workshop (NKOS 2017). CEUR-WS.org, Thessaloniki, Greece (2017)
8. Stahn, L.L., Dahlberg, I., De Luca, E.W.: Knowledge Organisation for Digital Libraries. In: Proc. of the 17th European Networked Knowledge Organization Systems Workshop (NKOS 2017). CEUR-WS.org, Thessaloniki, Greece (2017)
9. Stellato, A., Turbati, A., Fiorelli, M., Lorenzetti, T., Costetchi, E., Laaboudi, C., Gemert, W.V., Keizer, J.: Towards VocBench 3: Pushing Collaborative Development of Thesauri and Ontologies Further Beyond. In: Proc. of the 17th European Networked Knowledge Organization Systems Workshop (NKOS 2017). CEUR-WS.org, Thessaloniki, Greece (2017)