

Analyzing the research output presented at European Networked Knowledge Organization Systems workshops (2000-2015)

Fakhri Momeni and Philipp Mayr

GESIS - Leibniz Institute for the Social Sciences,
Unter Sachsenhausen 6-8
50667 Cologne, Germany
`firstname.lastname@gesis.org`

Abstract. In this paper we analyse a major part of the research output of the Networked Knowledge Organization Systems (NKOS) community in the period 2000 to 2015. We focus on the paper output presented at the European NKOS workshops in the last 15 years. Our open dataset "the NKOS bibliography" includes 14 workshop programmes (ECDL 2000-2010, TPDL 2011-2015) and 4 special issues on NKOS (2001, 2004, 2006 and 2015) which cover in total 171 papers with 218 single authors. A focus of the analysis is the development of collaboration, core authors and topics in this interdisciplinary field.

Keywords: NKOS workshops, Special issues, Output analysis, Network analysis, Central authors, Collaboration

1 Introduction

The European NKOS network has held a long-running series of annual workshops at the European Conference on Digital Libraries (ECDL), latterly reformed as the International Conference on Theory and Practice of Digital Libraries (TPDL). Typically, recent advances of KOS have been reported at the NKOS workshops, e.g. including the Simple Knowledge Organization System (SKOS) W3C standard, the ISO 25964 thesauri standard, the CIDOC Conceptual Reference Model (CRM), Linked Data applications, KOS-based recommender systems, KOS mapping techniques, KOS registries and metadata, social tagging, user-centred issues, and many other topics. A comprehensive and well cited review article on KOS was published in 2004 [3]. Special issues on Networked Knowledge Organization Systems (NKOS) have been published in Journal of Digital Information in 2001 and 2004, in New Review of Hypermedia and Multimedia in 2006 and in the International Journal of Digital Libraries in 2015 [1].

The motivation of this paper is to analyse the research output of the NKOS community. We are focusing on the informal part of this output, the presentations given the workshops. The specialty of this output is that these research

papers typically are not published in journals or conference proceedings, these papers appear just as oral presentations at the workshop and are documented on the website. To our knowledge nobody has done an analysis on this part of the research output before.

2 NKOS workshop bibliography

For our analysis we have compiled an open dataset the "NKOS bibliography"¹ which includes 14 workshop programmes with all presented papers at ECDL 2000, 2003-2010 and TPDL 2011-2015. We added papers from 4 special issues on NKOS which have been published in the same period.

In a first step we have extracted all paper titles presented at the NKOS workshop websites. We manually disambiguated author names. We added the papers from the special issues. These papers are the only formal publications in our analysis.

Our dataset covers in total n papers with n single authors.

Table provides an overview of all workshop papers.

venue	papers	authors
ECDL 2000		
ECDL 2003		
ECDL 2004		
ECDL 2005		
ECDL 2006		
ECDL 2007		
ECDL 2008		
ECDL 2009		
ECDL 2010		
TPDL 2011		
TPDL 2012		
TPDL 2013		
TPDL 2014		
TPDL 2015		

Table provides an overview of all papers in the special issues.

venue	papers	authors
SI 2001		
ECDL 2004		
ECDL 2006		
ECDL 2015		

3 Analysis

In order to analyze the collaboration in the community of NKOS workshops we built a network of all authors in this community and computed the centrality of

¹ The NKOS workshop bibliography is maintained in the following github repository: <https://github.com/PhilippMayr/NKOS-bibliography>

each author. For this purpose we utilized some centrality measures in Pajek². The network is composed of pairs of author names. Each pair means that two authors cooperated for writing a paper. For a paper with n authors, there are $n*(n-1)$ pairs. These pairs built the network for analysis in Pajek. to avoid repetition of pairs, we gave weight to pairs and it is equal to number of cooperations for two authors in different papers. Two used centrality measures of authors are degree and betweenness. Degree is the number of nodes that a focal node is connected to, and measures the involvement of the node in the network[2]. In our authorship-network it specifies the sum of co-authors for all papers that each author has written. betweenness, assess the degree to which a node lies on the shortest path between two other nodes, and are able to funnel the flow in the network[2]. In the authorship-network the author with high betweenness has a large influence on the transfer of information.

4 Results

Figure 1 demonstrates a general view of the network. In this view each author has at least one co-author. From the network illustrated in this Figure we selected the largest component that is represent in Figure2. 68 authors(31% of all authors) are in this component.

² A program for analysis and visualization of very large networks(<http://mrvar.fdv.uni-lj.si/pajek/>)

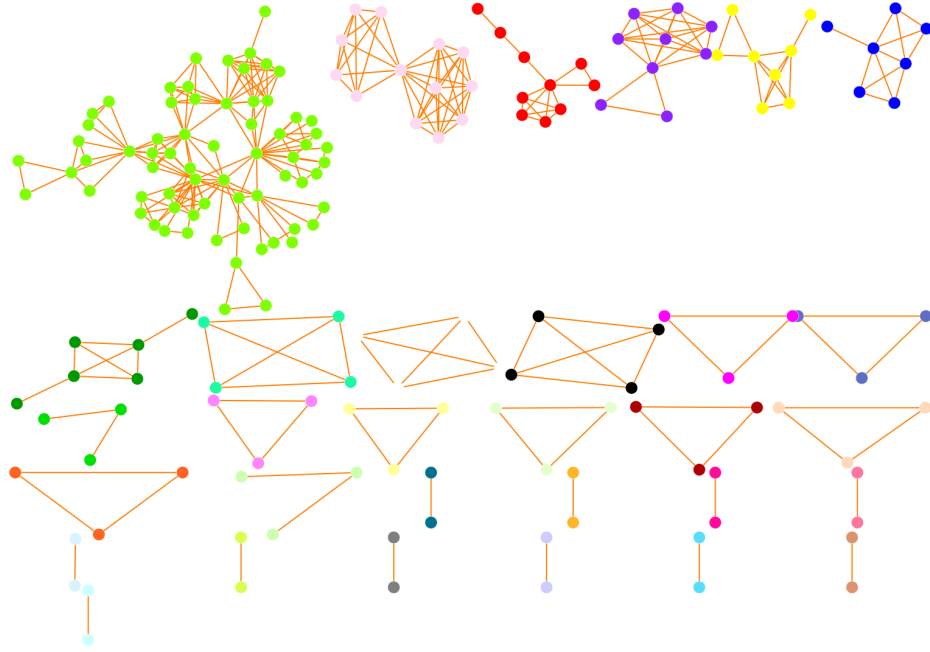


Fig. 1. The whole network of authors in community of NKOS workshops visualized in Pajek

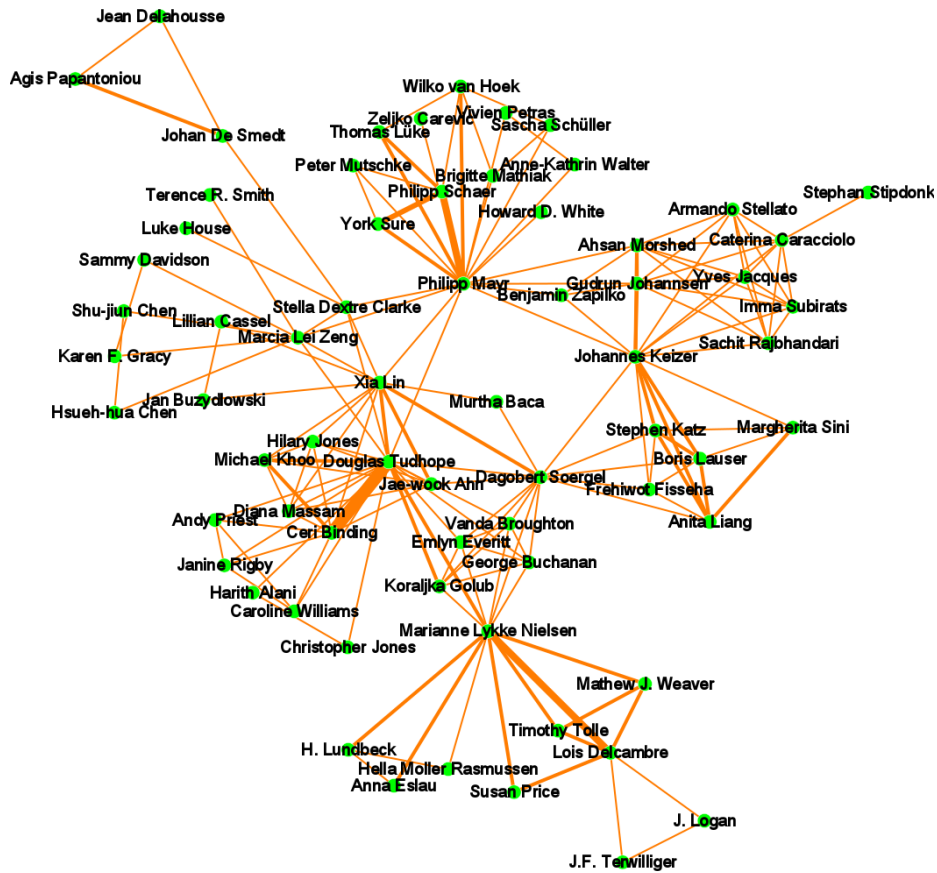


Fig. 2. argest component in the NKOS authorship-network

To show the quantity of collaboration in the community we measured the degree centrality for each author. Figure3 shows Percentage of authors with different degrees. From the Figure we see that 15% (with degree=0) of authors had no collaboration with others and 53% of them had just maximum 6 cooperations with other authors. Only 32% had with at least 8 co-authors for all their papers.

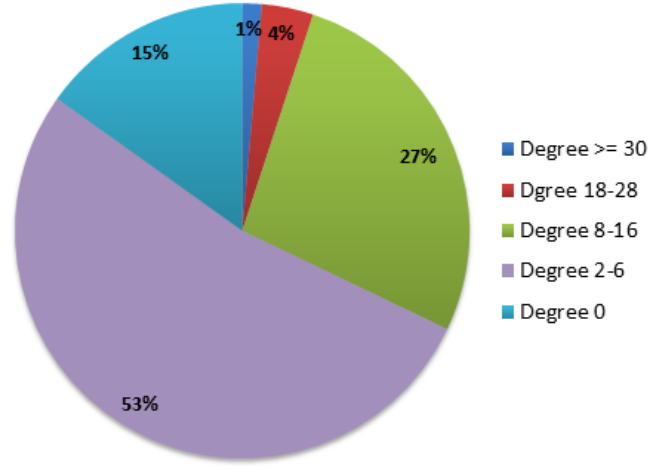
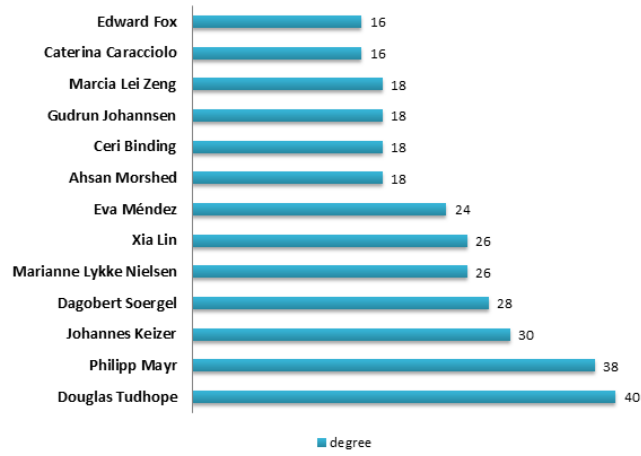
**Fig. 3.**

Figure 4 shows the authors with most degree (more than 16) in the network.

**Fig. 4.** Authors with degree more than 16 in the community of NKOS workshops

To detect the influence of authors on information exchange we calculated the betweenness centrality of authors. Figure 5 indicates the authors with most betweenness (more than 0.001). comparing betweenness with degree we can see, the ranking of authors has changed in betweenness. Some authors despite their high

cooperation with other authors, have lower ranking in betweenness comparison to others with lower degree.

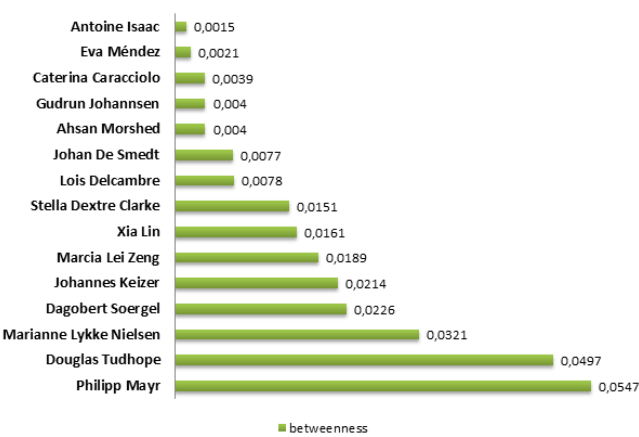


Fig. 5. Authors with betweenness more than 0.001 in the community of NKOS workshops

5 Conclusion

...

6 Acknowledgment

...

References

1. 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>
2. Opsahl, T., Agneessens, F., Skvoretz, J.: Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks* 32(3), 245–251 (2010), <http://dx.doi.org/10.1016/j.socnet.2010.03.006>
3. Zeng, M.L., Chan, L.M.: Trends and Issues in Establishing Interoperability Among Knowledge Organization Systems. *Journal of the American Society for Information Science and Technology* 55(3), 377–395 (2004)