Analyzing the network structure and gender differences of the “NKOS community”

Philipp Mayr
GESIS – Leibniz Institute for the Social Sciences
2018-09-13, Porto, Portugal

@article{karimi2018,
  title = {Analyzing the network structure and gender differences among the members of the Networked Knowledge Organization Systems (NKOS) community},
  url = {http://arxiv.org/abs/1803.04225},
  doi = {10.1007/s00799-018-0243-0},
  journal = {International Journal on Digital Libraries},
  author = {Karimi, Fariba and Mayr, Philipp and Momeni, Fakhri},
  year = {2018}}
Motivation

- European NKOS network has held a long-running series of annual workshops at ECDL & TPDL
- NKOS research output has never been subject to analysis
  - Esp. the (informal) research output of the NKOS community
- Most workshop output has never published as formal publications
  - Except 4 special issues
    (2* J. Dig. Inf., New Rev. Hypermed. Multimed, Int. J. Dig. Libr.)

http://hypermedia.research.southwales.ac.uk/kos/nkos/
http://nkos.slis.kent.edu/
NKOS workshop bibliography

• We checked all workshop published agendas and compiled a bibliography of past NKOS workshops (2000-2016)
• Manual author name disambiguation of each author entries
• The bibliographic data as bibtex files are available

https://github.com/PhilippMayr/NKOS-bibliography

For the follow-up IJDL paper we added all US NKOS workshops and other scattered events

Dataset for the paper includes: 123 papers with a sum of 256 distinct authors (just co-authored papers)

We describe the network with standard network analytic measures

- Global properties
- Centrality of the authors
- Gender differences in collaboration behavior
Results

- 44 isolated components
- Largest connected component (107 nodes)
- Men = purple, women = orange

Table 1  Overview of all NKOS papers sorted by years. In general, community shows a high average clustering in many years indica that there are many triangles in the network.

<table>
<thead>
<tr>
<th>Year</th>
<th>Nr. of papers</th>
<th>Nr. of authors</th>
<th>Nr. of links</th>
<th>Avg. cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>4</td>
<td>9</td>
<td>6</td>
<td>0.37</td>
</tr>
<tr>
<td>2002</td>
<td>3</td>
<td>10</td>
<td>13</td>
<td>0.8</td>
</tr>
<tr>
<td>2003</td>
<td>5</td>
<td>12</td>
<td>9</td>
<td>0.4</td>
</tr>
<tr>
<td>2004</td>
<td>13</td>
<td>39</td>
<td>47</td>
<td>0.65</td>
</tr>
<tr>
<td>2005</td>
<td>7</td>
<td>22</td>
<td>26</td>
<td>0.81</td>
</tr>
<tr>
<td>2006</td>
<td>11</td>
<td>33</td>
<td>39</td>
<td>0.73</td>
</tr>
<tr>
<td>2007</td>
<td>4</td>
<td>15</td>
<td>24</td>
<td>1.0</td>
</tr>
<tr>
<td>2008</td>
<td>7</td>
<td>15</td>
<td>9</td>
<td>0.2</td>
</tr>
<tr>
<td>2009</td>
<td>10</td>
<td>34</td>
<td>60</td>
<td>0.68</td>
</tr>
<tr>
<td>2010</td>
<td>8</td>
<td>21</td>
<td>19</td>
<td>0.61</td>
</tr>
<tr>
<td>2011</td>
<td>8</td>
<td>32</td>
<td>59</td>
<td>0.80</td>
</tr>
<tr>
<td>2012</td>
<td>6</td>
<td>26</td>
<td>56</td>
<td>0.92</td>
</tr>
<tr>
<td>2013</td>
<td>5</td>
<td>18</td>
<td>31</td>
<td>0.86</td>
</tr>
<tr>
<td>2014</td>
<td>6</td>
<td>16</td>
<td>13</td>
<td>0.85</td>
</tr>
<tr>
<td>2015</td>
<td>9</td>
<td>24</td>
<td>23</td>
<td>0.58</td>
</tr>
<tr>
<td>2016</td>
<td>17</td>
<td>60</td>
<td>114</td>
<td>0.75</td>
</tr>
</tbody>
</table>
Results

Largest component:
- Louvain clustering method identified 9 clusters
- Size of the nodes represent the degree
- Majority of European researchers
- Members of a cluster are often from the same institution / geographic proximity
Gender differences

- 97 (38%) women and 157 (62%) men and 2 unidentified names
- 46 women and 59 men in the largest component
- Homophily:
  - gender assortativity in this community is 0.1
    “there is a **positive tendency** among scholars in this community **to collaborate with the same gender**”
  - homophily among women is higher than the homophily among men in this network (see Slide 5)
Conclusions

- Limited dataset:
  - Focus on research paper presentations (organization activity not covered)
  - Influencial papers and standards are missing
  - Bibliometric data is missing

- Future work:
  - Complementing the dataset
  - Combining network and bibliometric measures
Thank you

Contact:

Dr Philipp Mayr  
GESIS - Leibniz Institute for the Social Sciences  
Germany  
Email: philipp.mayr@gesis.org  
Twitter: @philipp_mayr

Please consider to contribute to the NKOS Bibliography  
https://github.com/PhilippMayr/NKOS-bibliography

Special thanks go to my co-authors:  
Fariba Karimi and Fakhri Momeni  

https://doi.org/10.1007/s00799-018-0243-0