Qualitative and Quantitative Methods in Libraries (QQML) 5: 49-59, 2016

Co-authorship networks of scientific elite: case study of information science in Croatia

Dr. Đilda Pečarić and Prof. Dr. Miroslav Tuđman

The University of Zagreb, Faculty of Humanities and Social Sciences, Department of Information Sciences, Ivana Lučića 3, 10000 Zagreb, Croatia

Abstract: Relatively small number of scientists receives large number of citations. Researches of most cited authors who form scientific elite are often. What is rare are researches of co-authorship of most cited authors. In this paper we research co-authorship networks of most cited authors in information science in Croatia. Overview of social characteristics of co-authorship networks or whether authors and co-authors belong to the university or research institute is given. Also, position of most cited authors in co-authorship networks is researched: a) order of co-authors and their position in the scientific elite according to time periods; b) the institutional affiliation; c) international collaboration. We try to use bibliometric methods for detecting, identifying and visualizing social, institutional and intellectual networks and impact of scientific elite.

Keywords: Co-Authorship Network, Scientific Elite, Scientific Collaboration, Information Science, Citations, Bibliometric Analysis.

1. Introduction

From the research of history of scientific collaboration (Beaver and Rosen, 1978, 1979) we can conclude that starting interest was focused primarily on hard sciences, while the research of scientific collaboration in social sciences and humanities was in background. Reasons for that can be found not only in the specialization (Bush and Hattery, 1956) and professionalization of hard sciences (Beaver and Rosen, 1978, 1979), but also in the fact that the collaboration of universities and industry, that is research institutes and productions – was the crucial factor of industrial and technological development in 20th century. This had as its consequence the largest production of papers with multiple authorship primarily in hard sciences (Beaver and Rosen, 1979).

The research of scientific collaboration in social sciences is not that often and numerous, among other reasons because the indicators of multiple authorship of the papers in social sciences do not follow trends from (natural) sciences. In this

Received: 31.1.2016 Accepted: 21.3.2016 ISSN 2241-1925

paper we analyse institutional and international collaboration by co-authorship networks of scientific elite in information sciences in Croatia. The corpus of data used in this analysis is bibliographic database from 28,188 cited reference in 170 doctoral dissertations in information science done at Croatian universities from 1978 to 2009 (Pečarić, 2011). We start from assumptions that doctoral dissertations are original scientific papers which pass the process of rigorous scientific peer evaluation, and that is the reason why they are good and representative source of references from the field of information sciences. Based on that corpus of bibliographic data in previous researches we proposed bibliometric methods and techniques by which predecessors, scholars and researchers in information sciences can be identified (Đ. Pečarić, 2009). Also, we advocated methods and techniques by which key authors in scientific paradigm can be identified (M. Tuđman, Đ. Pečarić, 2009). In both cases we can talk about "cognitive collaboration" by identification of citation and co-citation networks, since those co-authors and co-word networks indicate methodological, theoretical and cognitive relationships among authors that are connected in clusters (Đ. Pečarić, 2010). The number of co-authors or the number of papers in co-authorship, mutual publishing, scientific connections measured by citations (citations and co-citations analysis) are used as indicators of researched scientific collaboration (F. Pehar, 2010 and W. Glänzel, 2003):

- Collaboration of individual scientists
- Collaboration inside research groups, departments, institutes
- International collaboration
- Collaboration among divisions (academic-private, public-academic, private-public etc.).

In this paper we want to research scientific collaboration in information science by analysing multiple authorship on the corpus of cited reference in doctoral dissertations done at universities in Croatia. We are also aware that the concept of collaboration cannot be simply or unambiguously measured through authorship? (J. S. Katz, B. R. Martin, 1997). That is why our goal was just to retrieve basic indicators about: a) the order of co-authors and their position in the scientific elite according to time periods; b) the institutional affiliation; c) international collaboration. Retrieved indicators could serve as the basis for the understanding of scientific collaboration character: is it a result of social networks, institutional networks, communicational networks or cognitive networks?

2. Methods and basic data

The corpus of bibliographic data consists of 28,188 cited references in 170 doctoral dissertations. The majority of citations (22,262 out of 28,188) is with authors (table 1.), and 5,926 or 21% is without author (patents, standards, and similar documents). The citations without author are not analysed in this paper.

	1 author	2 co-authors	3 co-authors	4 or more co-authors
Number	15993	3996	1402	871
Percentage of overall no. of citations	56.7	14.2	5	3.1
Percentage of no. of citations with authors	71.8	17.9	6.3	3.9

Table 1. The number of citations for papers with authorship according to the number of authors of the paper

Out of 22,262 cited documents with authorship, 72% have one author, and 28% have two or more authors (table 1). Table 2 shows the overview of the number of cited papers according to citations' periods and number of co-authors.

Periods	*Percentage of citations (from references with authors)*			
	1 author	*2 co-authors*	*3 co-authors*	*4 or more co-authors*
from 1978 to 1989	80.1	15.0	3.2	1.8
from 1990 to 1999	76.8	15.5	4.2	3.5
from 2000 to 2009	66.1	20.5	8.7	4.8

Table 2. The percentage of citations (from references with authors) according to citations' periods and number of authors of the paper

It can be noted from the data presented in Table 2 that the appearance of multiple authors in information science is similar to hard sciences; 80% of cited papers with one author in the 1980s dropped to 66% in the 2000s. Or in other words, in the period of 30 years the number of multiple authorship grew from 20% to 34%.

The percentage of papers with one author, that is single authorship, varies between 61% to 93% in different information science disciplines (table 3). The percentage of cited papers with multiple authorship in lexicography and museology is lower than 10%, unlike information systems and information science where over 1/3 of cited papers is with multiple authors.

Discipline	*Percentage of citations (from references with authors)*				**Total number**
	1 author	*2 co-authors*	*3 co-authors*	*4 or more co-authors*	
Archivistics and documentation	86.3	9.3	2.5	2.0	**1013**
Information systems	61.2	23.3	9.4	6.1	**5920**
Information science	64.1	22.0	8.3	5.5	**4565**
Communicology	73.8	18.3	5.1	2.9	**4245**
Lexicography	93.4	5.8	0.8	0	**121**
Librarianship	78.9	14.5	4.4	2.2	**4202**
Museology	90.5	6.8	1.7	1.0	**2217**

Table 3. The percentage of citations (from references with authors) according to information science disciplines and number of authors of the paper

In another place we point out that distribution of citation in this corpus of data follows bibliometrics laws (Đ. Pečarić, 2011): a small number of authors receives a large number of citations. For our topic of research the data about the number of authors that are cited only once in the analyzed corpus of data is indicative (table 4).

	1 author	*2 co-authors*	*3 co-authors*	*4 or more co-authors*
Number	6766	2061	757	885
Percentage of paper of authors cited once	*64.6*	*19.7*	*7.2*	*8.5*
Percentage of papers with authors	*32.2*	*9.8*	*3.6*	*4.2*

Table 4. The number of cited publications from authors cited once according to number of authors of the paper

The number of papers from authors cited only once is 49.8%. Out of 10,469 of those papers only 65% are signed by one author, 20% by two authors, and 15% by three or more authors. This points to the conclusion that the distribution of multiple authorship is independent of citation frequency, but is dependent on the characters of the production of papers. Since papers cited only once are insufficient for analysis of institutional and international collaboration, we

directed our analysis to "scientific elite", i.e. the papers of most cited authors as a reliable sample for the understanding of scientific collaboration.

3. Core authors, co-authorship and institutional scientific cooperation

We analyzed scientific collaboration in information science on the sample of 43 most cited authors. The most cited 43 authors, that is core authors, received 1,420 citations, which is 6.37% from the overall number of citations, and 11,69% from the citations with citation frequency larger than 1. The size of the sample of 43 authors is just 0.33% from all cited authors (13,162), that is just 1.38% (3,109) from authors that are cited more than once. From these 43 core authors, 966 papers are cited once or more times.

Table 5 shows an overview of the number of cited papers according to periods and co-authorship.

Periods	*Publications with single authorship*	*Publications with multiple authorships*	*No. of co-authors*	*Out of 43 most cited authors*
From 1978 to 1989	127 (78.4%)	35 (21.6%)	31	28
From 1990 to 1999	383 (78.3%)	106 (21.7%)	110	40
From 2000 to 2009	264 (71.5%)	105 (28.5%)	110	39
Total	**774**	**246**	**251**	

Table 5. The number of cited papers according to periods and co-authorship

It is obvious that the number of publications with single authorship keeps scaling down in information sciences too, or in other words, the number of publications with multiple authorships keeps scaling up (from 21% to 28%). Further analysis should take into consideration that the number and content of core authors varies in the period from 1978 to 200. This data open up the possibilities that different factors (not only those of social and technical nature, but also generation differences) can affect the increase in the number of papers with multiple authorships.

Authorship	*43 most cited authors on position*			**Total**
	1st author	2nd author	3rd author	
1 author	713			**713 (73.8%)**
2 authors	111	46		**157 (16.3%)**

Authorship	*43 most cited authors on position*			**Total**
	1st author	2nd author	3rd author	
3 authors	42	14	8	**64 (6.6%)**
4 authors	28	1	3	**32 (3.3)**
Total	**894**	**61**	**11**	**966 (100%)**

Table 6. The position of 43 most cited authors in 966 papers according to the order of authorship and the number of authors of the paper

Tables 6 and 7 show the order of authorship (i.e. 1 author, 2 authors, 3 authors etc.) for 966 papers that are (co)signed by 43 most cited authors. Most papers 713 (i.e. ¾) are signed by one author, 16% of papers are signed by two authors, and 10% of papers are signed by three or more authors (table 6). Core authors are the first authors on 894 papers (713 with single authorship and 281 papers with multiple authorship), they are second authors on 61 papers, and only in 11 papers are they third authors. Almost the same ratio of data, with minimal differences, is repeated in citation percentages (table 7) received by the papers shown in table (table 6).

Authorship	*43 most cited authors on position*			**Total**
	1st author	2nd author	3rd author	
1 author	1097			**1097 (77.2%)**
2 authors	151	53		**204 (14.4%)**
3 authors	59	16	8	**83 (5.8%)**
4 authors	32	1	4	**37 (2.6%)**
Total	**1339**	**70**	**12**	**1421 (100%)**

Table 7. The number of received citations for 966 papers (according to the position of 43 authors) and the number of authors of the paper

From collected data we can observe the number of co-authors that most cited authors attached to themselves. Only 8 authors (that is 18.6%) from 43 most cited authors have no cited papers in co-authorship. Remaining most cited authors collaborate with 196 co-authors, that is 189 different authors. According to the order of authorship of the paper, when core authors are on the first

position, they collaborate with 139 co-authors, 23 core authors on the second position collaborate with 54 co-authors, and on the third position 9 core authors collaborate with 24 co-authors. The ratio of co-authors of the most cited authors regardless of the position in the order of authorship is in the range from 1.22 to 3.17. The number of co-authors of the most cited authors is shown in table 8.

No. of core authors	No. of co-authors scale	***Percentage***
19	1-5	***44.2***
10	6-10	***23.3***
5	11-13	***11.6***
1	16	***2.3***

Table 8. The number of most cited authors and the number of co-authors

What kind of scientific collaboration do the data about co-authors clusters that are formed around core authors point to? It is not possible to answer that question without the observation of the institution in which co-authors of cited papers work. Table 9 shows institutional affiliation of authors for 284 papers with multiple authorship.

The data about their institutional affiliation are unknown for 13.4% authors of analyzed papers with multiple authorship. The majority of most cited authors and their co-authors work on universities and research institutes. Only 9.28% work in industry, etc. There is a good reason to ask what the data about institutional affiliation of authors of scientific papers indicate. Not only in this example but in general. Are these indicators the indicators of scientific collaboration or perhaps institutional organization of science and scientific production?

If we start from the assumption that the number of co-authors is an indicator of scientific collaboration, than it could be claimed that the growth of the number of papers with multiple authors is the argument and proof of the development of scientific collaboration. Those are trends confirmed by data in this research too: ever growing number of papers with multiple authorship. However, the fact is that papers with multiple authorship are signed typically by authors who work in same institutions.

Type of institutions	Number of authors	Percentage
Archives	3	0.62
Documentation centre	2	0.41
Educational institution	5	1.03
Industry	45	9.28
Library	8	1.65
Research institute	20	4.12
University departments	337	69.48
Unknown	65	13.40

Table 9. Institutional affiliation of the authors of 284 papers with Type of institutions Number of authors Percentage

The consequence of such kind of thesis about co-authors as the key indicator of scientific collaboration would be that the dominant characteristic of science is to develop institutionally, i.e. inside leading institutions (universities and research institutes). This is correct from the point of organization of scientific work and scientific production, but it is not precise enough because the influence of information and communication processes and networks on the development of science is not taken into consideration. H. D. White and B. C. Griffith (1981) at the very beginning of bibliometrics research of scientific collaboration search for "Intellectual Structure", that is "knowledge maps" (R. Capurro, 2006) as an overview of the development of science and scientific collaboration, independently from institutional affiliation of authors. That is why it would be advisable to use indicators about co-authorship and multiple authorship primarily as an indicator of the organization of science and scientific production. Also, those indicators should be used with caution as general indicators of scientific collaboration.

4. Multiple authorship: international cooperation or social networks

The number of co-authors or the number of papers in co-authorship, that is mutual publishing and institutional scientific connections, are used as an indicator of international scientific collaboration. Table 9 shows the cluster of co-authors relationship that exists on national and international level.

It is easy to note that the authors of papers with multiple authorship are mostly from the same country. The authors of papers with multiple authorship are rarely from different countries, what is indicated by weak relationship among clusters made by co-authors networks on international level. If we would disregard these occasional papers of co-authors that belong to different national scientific communities, what we would get is the sequence of independent clusters.

This does not mean that international scientific collaboration does not exist, but it is questionable what we can prove by indicators about co-authors and multiple authorship. Numerous co-word and co-citation analyses also indicate international scientific collaboration by the overview of "intellectual structure" and "knowledge maps". Furthermore, the overview of national domicile of the authors of papers that are published by international journals would already give far more precise image about international collaboration of scientific community.

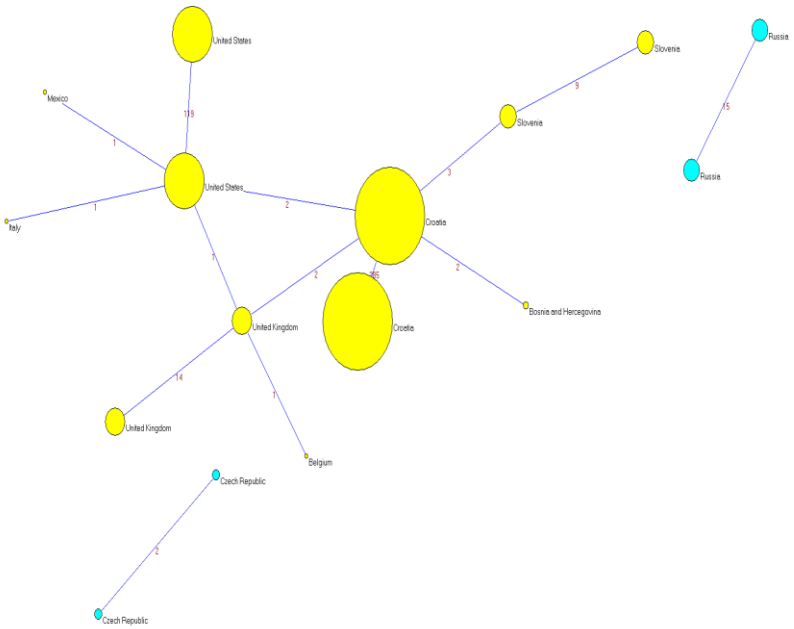

Figure Multiple authorship as an indicator of international scientific collaboration

We do not have enough data to precisely describe existing, although weak, international relationship of authors that publish papers together. Based on personal experience we could assume that those papers with multiple authorship are written more as a result of social and professional relationships of authors, than as a result of work on international projects of institutions they belong to. This assumption should yet be researched.

5. Instead of conclusion

The number of co-authors and the number of papers in co-authorship, mutual publishing, international scientific relationship measured by co-authorship networks, etc. are used as indicators for the research and overview of individual, institutional and international scientific collaboration (W. Glänzel, 2003). The data from our research of most cited papers with multiple authorship in

information science on the corpus of bibliographic data from doctoral dissertations done at universities in Croatia can also be used as a contribution to pre-existing belief that the data about multiple authorship are a valid proof of scientific collaboration.

However, as much as those data describe scientific collaboration, at the same time it remains insufficiently clear which aspect of scientific collaboration they reflect. That is why we, in conclusion, must go back to the question what „concept of scientific cooperation" is. In previous papers (Đ. Pečarić, 2010, 2011) we analyzed different dimensions of scientific collaboration: institutional, social, communicational and cognitive. We believe that by bibliometrics methods (by co-citation and co-word analysis) communicational and cognitive networks can be recognized as important dimensions of scientific collaboration. Co-authorship networks can be used for the overview of institutional and social networks that form scientific collaboration and the dynamics of scientific production. It would be pretentious and too ambitious to try to prove communicational and cognitive dimensions of scientific collaboration by co-authorship networks.

References

Beaver, D. de B., Rosen, R. (1978). Studies in Scientific Collaboration. Part I. The Professional Origins of Scientific Co-authorship., Scientometrics, Vol. 1. No. (1978) 65 - 84.

Beaver, D. de B., Rosen, R. (1979). Studies in Scientific Collaboration. Part II. Scientific Co-Authorship, Research Productivity and Visibility in the French Scientific Elite, 1799-1830., Scientometrics, Vol. 1. No. 2 (1979) 133 - 149.

Beaver, D. de B., Rosen, R. (1979). Studies in Scientific Collaboration. Part III. Professional and the Natural History of Modern Scientific Co-authorship., Scientometrics, Vol. 1. No. 3(1979) 231 - 245.

Bush, G. P., Hattery, L. H. (1956) Teamwork and creativity in research. Administrative Science Quarterly, 1 (1956) 361 - 362.

Capurro, R., Zins, C. (2006). Knowldege Map of Information Science. Rafael Capurro's responses to Chaim Zins. (URL: www.capurro.de 10.4.2009.)

Glänzel, W. (2003). Bibliometrics as a Research Field. A course on theory and application of bibliometric indicators. Course Handouts.

Gordon, M. D. (1980). A Critical Reassessment of Inferred Relations Between Multiple Authorship, Scientific Collaboration, the Production of Papers and their Acceptance for Publication., Scientometrics, Vol. 2. No. 3 (1980) 193 - 201.

Katz, J. S., Martin, B. R. (1997). What is Research Collaboration?. Research Policy, 26, 1- 18.

Pečarić, Đ. (2009). Relationship between scientific paradigm and research front. On Example of Information Science Research Production. Proceedings of The Future of Information Sciences: Digital Resources and Knowledge Sharing : Zagreb: Department of Information Sciences, Faculty of Humanities and Social Sciences, University of Zagreb, 767-777.

Pečarić, Đ. (2010) Scientific Collaboration and Development of Scientific Community – Contribution to Research of Development of Information Science in Croatia. //CECIIS 2010 Proceedings: Varaždin: Faculty of Organization and Informatics, 125 - 134.

Pečarić, Đ. (2011). Development of Information Sciences in Croatia. Bibliometric Analysis of Doctoral Dissertations in Information Sciences from 1978 to 2009. Saarbrücken: LAP LAMBERT Academic Publishing.

Pečarić, Đ., Tuđman, M. (2009). Predecessors, Scholars and Researchers in Information Sciences. Contribution to Methodology for Bibliometrics Anaysis of Scientific Paradigms. Proceedings of The Future of Information Sciences: Digital Resources and Knowledge Sharing : Zagreb: Department of Information Sciences, Faculty of Humanities and Social Sciences, University of Zagreb, 713-726.

Pehar, F. (2010). Komunikacijska uloga hrvatskih časopisa u polju informacijskih znanosti: Bibliometrijska analiza Vjesnika bibliotekara hrvatske i Informatologije. Faculty of Humanities and Social Sciences, University of Zagreb: Doctoral dissertation.

Tuđman, M., Pečarić, Đ. (2009). Prilozi dubinskoj analizi komunikacijskih obrazaca. // Informatologija 42, 2(2009), 87-92.

White, H. D., Griffith, B. C. (1981). Author Cocitaion: A Literature Measure of Intelectual Structure. Journal of ASIS 32 (2), 163-171.