Scientific Collaboration and Development of Scientific Community

Contribution to Research of Development of Information Science in Croatia

Đilda Pečarić
Faculty of Humanities and Social Sciences
Department of Information Sciences
The University of Zagreb
Ivana Lučića 3, 10000 Zagreb, Croatia
dpecaric@ffzg.hr

Abstract. *Field of information science in Croatia includes seven disciplines: archivistics and documentation, communicology, information science, information systems, lexicography, librarianship and museology. Until now, research about scientific collaborations and correlations on individual, institutional, national or international level within these disciplines does not exist in Croatia. This research shows scientific collaborations among disciplines in the field of information science using bibliometrics methods, by indicators about collaboration, in order to establish reciprocity of scientists and their influence and connections. On the corpus of 22,210 cited bibliographic units retrieved from doctoral dissertations in information science, scientific collaboration and influence of authors' productions and co-authors connections to the development of scientific community in the field of information science in Croatia is shown. Scientific collaboration is followed through three indicators: a) the distribution of citations according to disciplines and periods in order to identify the factor of connections and coherence among certain disciplines; b) by the analysis of the number of co-authors, more precisely, by the analysis of co-authors' connections, level and form of scientific collaboration inside certain scientific disciplines as well as overall field of information science is followed; c) by the analysis of most cited authors in conceptual knowledge zone, the dynamics of the development of scientific paradigm and their dominant authors in thirty year researched period in is shown, both according to periods and according to disciplines.*

Keywords: Bibliometrics, Cohesion, Coherence, Scientific Collaboration, Information Science, Social Network, Institutional Network, Communication Network, Cognitive Network

1 Introduction

In bibliometrics multiple authorship or co-authors' relationships are used as an indicator to measure level, form and intensity of scientific collaboration. In broader concept, the indicator of scientific collaboration are two or more researchers (from two or more organizations or countries) working together (V. Diodato, p. 47.). H. D. White and B. C. Griffith (1981.) are among the first researchers that used bibliometric data to research "Intellectual Structure", that is, "knowledge maps" which were the base for the presentation and analysis of scientific collaboration. So, we have a good reason to raise the question: how does scientific collaboration influence the development of scientific disciplines within one or among more scientific fields?

Since the corpus of bibliometric information that we have access to does not permit such comprehensive research and search for the answer to the question raised, we will limit our research to the following methodological problem: which set of bibliometric information is useful for the analysis of scientific collaboration?

More precisely, how can bibliometric indicators be used to analyze interdisciplinary and multidisciplinary activities in information science? That question is important because there is no consensus about the scope and domain of information science. Are we dealing with information science or with information

sciences? Therefore, we have to monitor cohesion and diversity – processes which influence the structure and dynamics of scientific development. Cohesion and diversity can serve as theoretical framework for analyzing scientific collaboration.

Examples will be elaborated on bibliometric information corpus on doctoral dissertations done at Croatian Universities in the period from 1978 to 2007. That analysis will include co-word analysis and citation and co-citation analysis of 22,210 references cited in 134 doctoral theses (Đ. Pečarić, 2009.).

2 About the methodological approach

Bibliometrics survey papers and textbooks start from premises that collaboration and connection of individuals, institutions and states are some of the basic principles of contemporary science (F. Pehar), and that bibliometrics measures scientific collaboration by means of co-publication statistics (W. Glänzel).

The number of co-authors or the number of papers done in co-authorship, joint publications, scientific connections measured by citations (citation and co-citation analysis) are used as indicators for the research of scientific collaboration (F. Pehar, W. Glänzel):

- Collaboration between individual authors
- Collaboration among research groups, departments, institutes
- International collaboration
- Collaboration between sectors (university-private, industry-academic, private-public etc.)

According to our question about the influence of the scientific collaboration on the cohesion and diversity of information science development, we believe that the following dimensions of collaboration can be followed and researched by using bibliometric indicators: a) social networks, b) institutional networks, c) communication networks, and d) cognitive networks.

These dimensions of scientific collaboration are not independent from each other. They are mutually related and dependent, but at the same time they are not identical. Social networks of scientists are not necessarily identical and they do not entirely overlap with institutional networks. Communicational networks are far bigger according to scope and temporal range than institutional networks. Cognitive networks ("intellectual structures" or "cognitive maps" of sciences - R. Capurro) are under the influence of social and institutional networks of scientists, but they usually follow the logic of scientific paradigm: dominant authors independent of their social or institutional position.

At our disposal are the bibliometric indicators that we retrieved from the corpus of data on 134 doctoral dissertations in information sciences. Our objective is to recognize and understand cohesive processes and diversity in the development in information science domain. We shall analyze the usefulness of bibliometric indicators for the presentation of social, institutional, communicational and cognitive networks. Our starting point is hypothesis that by combining sets of different indicators, social, institutional, communicational and cognitive functions of sciences can be researched. If that hypothesis is correct, it means that each scientific community has it is own history of development, and therefore bibliometric indicators cannot be used as a measure of collaboration: they are only instruments for analysis, but not a measure which can be used for measuring the level of development of a certain scientific community and scientific collaboration.

3 The influence of scientific collaboration on cohesion and diversity of information science development

3.1 Social networks

We can start with theorem that social scientific networks precede, or at least are the background for the development of new scientific community, and/or new scientific discipline. That was the case with the appearance and development of information science in Croatia. The Centre for the Study of Librarianship, Documentation and Information Sciences (CSLDIS) was the first institution which "produced" Masters of Arts in Information Sciences (442 M.A. in the period from 1961 to 1984), and this was precondition for the institutional development of information sciences in Croatia (M. Tuđman et al., 1984, 1988.).

The first PhD theses in information science could have been done at the University of Zagreb – primarily thanks to social network of scientists, i.e. mentors. Those mentors were not institutionally networked, because they worked at different faculties and institutions. The cohesive impact of the first social network, whose members were the predecessors of information science in Croatia, was founded on the enthusiasm and scientific interest of scientists from different sciences who focused and dedicated their work to new scientific area (Đ. Težak).

Sixteen mentors were responsible for 21 PhD candidates in the period from 1978 to 1989. Those 16 mentors were members of seven different faculties and scientific institutes, because at the time main Faculty for Information Science did not exist. Four mentors were from the Faculty of Economy; three were from the Faculty of Organization and Informatics (FOI) and Faculty of Humanities and Social Sciences (FHSS); 2 mentors were from the Faculty of Political Sciences, and one from: Faculty of Kinesiology, Faculty of Science, Institute for Lexicography and Museum Documentation Center.

According to the indicators about mentors' domicile institutions, we can conclude that in that period we have institutional dispersion, since mentors were located or dislocated at seven different addresses. But according to their activities we can conclude that efficient social network existed. In other words, there was no institutional concentration of academic community in the emerging field of information science, but the interest and enthusiasm of scientists from different institutions for the new scientific area acted as the driving and cohesive force on establishing the information science community (Đ. Pečarić, M. Tuđman, 2010).

The development of scientific community at the same time implies establishing institutional and social network. Therefore the following indicators about the role of the mentors in the process of establishing information science community can be understood as indicators of social networks (J. Ardanuy, C. Urbano, L. Quintana) as well as indicators of institutional networks.

Fig. 1 shows the most cited authors[1] – whose papers were published in Croatian – cited in PhD theses that were done at FHSS and FOI.

[1] All Figures are in the Appendix

Our starting point cannot be that the most cited authors form a social network: because the network of citations is primarily the result of thematic interest of PhD candidates. However, we can assume that 55 mentors and 134 PhD candidates have also developed social relationship as a form of social network. The data from the relatively small sample, given in Fig. 1, show that social networks influence citation maps.

From 32 most cited authors in Fig. 1, half are mentors (16) that supervised 70 out of 134 PhD candidates. Based on that information we can conclude that social network exists among the most cited authors. That social network is not visible and it cannot be retrieved from bibliometric data about citation and mentorship.

At the same time it should be pointed out that scientific collaboration between a mentor and a PhD candidate (i.e. mentor's role in social network which exists in academic community) does not guarantee the position on the map of most cited authors to any mentor. That conclusion is confirmed by bibliometric indicators about citation of the mentors in PhD theses, because as many as 11 mentors in 14 PhD theses were not cited. In 38 theses 20 mentors were cited from 1 to 3 times. Only a small number of mentors, 8 of them, are cited more then ten times in one third of all theses. This is indicated by the data in Fig. 1: 16 out of 32 most cited authors were not mentors. Among 45 most cited authors (from all cited authors) only 14 were mentors (13 of those working in Croatia and one outside Croatia - Đ. Pečarić, M. Tuđman, 2010). Therefore, mentors are cohesive determinant of information science community. However, since only 25% of mentors are among 45 most cited authors, and 2/3 of the most cited authors are not mentors, several other factors have impact on the cohesion and coherence of scientific community, and not only mentors' social network.

3.2 Institutional networks – or about the institutionalization of information sciences

Data shown in Fig. 1 are a good indicator of the influence that institutions and institutional networks have on the shaping of the map of most cited authors. What can be confusing on this map is institutional "affiliation" of certain most cited authors. Three authors (P. Klasinc, I. Maroević,

M. Tuđman) are cited in both institutional domains (FOI and FHSS). Also, it can be confusing that A. Bauer, I. Maroević, P. Novosel, B. Petz, M. Plenković, T. Šola, V. Žiljak are cited in the domain of FOI since they are key authors in the field of communicology and museology, and they also work in other faculties.

To understand these data it is important to know that until 1990 the Faculty of Organization and Informatics was the only institution that had postgraduate studies in information sciences, which means that PhD theses could be done in all information science disciplines. So, at FOI, apart from 46 PhD theses in information systems and 16 theses in communicology, were done: 2 theses in archivistics, 4 in informatology and 3 in museology. Based on that information it can be concluded that institutional networks influence the mapping of scientific fields. How strong and influential the impact of institutional networks is yet to be discovered.

Mentors can be members of different institutions which form infrastructure of scientific community, but at the same time they act as a cohesive factor inside the field of information sciences. The data about mentors according to disciplines (Table 2) leads us to that conclusion.

Out of 55 mentors, two were mentors in three disciplines, 6 were mentors in two disciplines, and the rest 47 mentors were mentors in only one discipline (Table 2).

Eight mentors who were mentors in more than one discipline supervised 43 PhD candidates. This indicator combined with other data can contribute to better understanding of the linkages and collaboration among disciplines that form the field of information sciences. Those interconnections can be followed more precisely by co-word analysis, because citation and co-citation analysis are not precise enough.

The data about the renewal and growth of scientific community provide information for the understanding of the development of institutional network of information scientists. Fourteen out of 55 mentors (or 25% of mentors) took their doctoral degree in information sciences at Universities in Croatia. So, 10% of PhD candidates became mentors. They supervised 35 out of 134 PhD candidates in information sciences. This process was gradual, which can be seen from time period in which mentors finished their PhD theses to the time when they began working as mentors.

These 14 mentors, who had done their doctoral degree in information sciences, supervised only one PhD candidate in the first period (until 1989). In the second period (until 1999) they supervised 12 PhD candidates, and in the third period (until 2007) 22 PhD candidates. That is one of the indicators of the collaboration inside the scientific community. But it is also an indicator of the development of institutionalization of information science and development of academic network.

Table 2. Mentors according to disciplines

Mentor	Number of disciplines	Number of PhD theses
Stipčević, A.	3	4
Boras, D.	3	4
Kržak, M.	2	3
Novosel, P.	2	11
Lasić-Lazić, J.	2	6
Srića, V.	2	7
Prelog, N.	2	4
Žiljak, V.	2	4

3.3 Communication network

It is reasonable to raise the question if the co-citation maps are an indicator of scientific collaboration. We can repeat S. Katz's (1997) question: "how closely researchers have to work together in order to constitute a 'collaboration'"? On one end, the answer is that the whole international community of researchers are an example of scientific collaboration because everyone contributes to the development of science. On the other end, the answer is that collaboration exists only when scientists work together on mutual research (S. Katz).

According to H. White, cohesion and coherence are the concepts which can be used to understand relationship between two texts (entities) which are „topically related" or „more or less on the same subject" (H. White, 2002.). On this occasion we advocate interpretation that the cohesion is communication factor that binds and connects communication components in one whole. Coherence refers to cognitive linkages and relationships among components (comprehensive bibliography exists about coherence and cohesion – W. Bublitz). In other words, cohesive influence of cited authors can be followed by co-citation analysis of cited references in PhD theses.

Co-citation maps offer the conclusion that the authors are „topically related“, or more or less on the same subject“, because they are grouped in clusters. In other words, co-citation networks established by the citation of authors form thematic topics. But in the best case, co-citation maps only provide information about who is present in thematic topic, but not what the meaning of their presence is.

Fig. 3 to 6 present the maps of co-citation of most cited authors. The selection is made from a set of figures (Đ. Pečarić, 2009.) with the intention to illustrate the role of co-citation networks as indicators of communication linkages. Co-citation networks inform us about information traffic within scientific community. Co-citation networks are not a picture of collaboration within scientific community. Pairs of cited authors form clusters of thematic topics that differ from one citing community to another or from one group of institutions and time periods to another.

Fig. 3 shows co-citation of most cited authors in doctoral dissertations done at FHSS. If we ask what the nature of relationship between cited authors J. Lasić Lazić – C. L. Borgman, I. Maroević – F. W. Lancaster, M. Tuđman – M. K. Buckland, E. Verona – M. K. Buckland, or any other pair is, we can conclude that the overview of co-cited authors does not refer to mutual communication of cited authors, that is co-cited pairs are not the result of mutual citation of those authors. Some of these authors (most probably) have never been in mutual communication, because some of them are not alive, and others do not belong to the same information science discipline etc.

The nature of relationship referred to by co-citation cannot be primary communication because of time, space, institutional or social barriers that disable contact among those authors. The conformation for this is found on Fig. 4 (pairs: A. Horvat – E. Verona, E. Svenonius – M. Tuđman, M. Gorman – J. Lasić Lazić), although pairs like I. Maroević – Ž. Vujić, D. Boras – I. Škarić exist, or on Fig. 5 (pairs: N. Prelog – M. Plenković, I. Maroević – T. Šola, J. Brumec – M. Žugaj, P. Novosel – M. Plenković) that can be indicators of the existence of both social and institutional networks.

We will call the overview of those co-citation authors communication networks, because the authors of PhD theses are the ones that communicate with available and relevant reserve of scientific knowledge. More precisely, co-citation networks refer to information networks because they transfer the information about structuring of scientific interest and the exchange of those pieces of information the content of which is presented by the cited authors.

Therefore, Fig. 6 is the illustration of most cited authors in PhD theses done at FOI. Doctoral candidates establish communication with cited authors by citation, and Fig. 6 is the overview of that communication.

Communication networks refer to the cohesion of scientific field, but they are not a presentation of cognitive maps. In the best case, they mirror information interest of a certain group of scientists, and that scientific interest changes over time, and/or according to the logic of mission of scientific institutions.

Communication networks retrieved by co-citation of most cited authors exist, in our case just because they are generated by one group of scientists. Without these groups of scientists, in our case PhD candidates, the networks disintegrate and do not exist. In other words, every scientific community cites different authors to resolve their problems, tasks, issues or interests and by that establishes different communication networks.

3.4 Cognitive networks

Coherence is much more defined by the relationship of PhD candidate toward the topics of their interest, than by their relationship toward the mentor and cited authors. That is why coherence is a cognitive relationship among entities (subjects) that are topics of scientific research. Coherence as cognitive relationship is established during the research process. Coherence is the result of cognitive process, and cohesion is the result of communication process. Cohesion and coherence are two relationships that are not necessarily reciprocally conditioned.

From data shown on Fig. 7 to 10 we can see the overview of subjects i.e. “cognitive maps” made by 134 analyzed PhD theses in information sciences.

By co-word analysis (key words) we can follow thematic topics and the dynamic of the development of particular field and differences in the development of scientific field in different environments.

We believe that it is possible to advocate following hypothesis: a) two scientific communities can use (generate) two **different**

communication networks (that can be identified by co-citation analysis, that is, clusters of most cited authors); b) two scientific communities can generate **same** or **similar** cognitive networks (that can be identified by co-word analysis). It is more difficult to assume that the opposite can be true: that two different scientific communities use same communication, but different cognitive networks. These postulates are conferred by Fig. 7 to 10.

The co-word analysis on Fig. 7 refers to the set of clusters that greatly overlaps with disciplines (sciences) of information sciences, as Prof. B. Težak defined the field in late 1960s (Đ. Težak).

Cognitive structures that are described by co-words (Fig. 8 to 10) also indicate the diversity of the development of information science field. New field *E-learning* appears on FHSS as a separate thematic discipline. Two new thematic topics appear on FOI as consequences of development and specialization: *programming* and *modeling methods.*

If we had enough space to present the development of information sciences in Croatia according to time periods and disciplines, then we would be able to follow precisely the inside dynamic and development of cognitive networks. Thematic topics can be very precisely identified, that is scientific disciplines that are studied at FHSS (Fig. 8 – archivistics and documentation, librarianship, museology, information science); and at FOI (Fig. 9 – information systems, programming, modeling).

Different networks of data, on Fig. 7 to 10, refer to conclusions that the coherence of whole researched corpus (Fig. 7) exists. But at the same time we can recognize the diversity of cognitive units inside the field of information science (Fig. 10).

For example, traditional field of librarianship is fragmented into cognitive units: *libraries, theories of librarianship, catalogs, bibliometrics* and *e-learning.* Whether bibliometrics and e-learning will be developed as separate disciplines is yet to be seen. However, bibliometrics and e-learning should already be recognized as separate cognitive units according to available data.

4 Instead of conclusion

We started from the postulate that scientific collaboration is the precondition for the development of scientific activity. The question we raised at the beginning of our research was whether the indicators of bibliometric connections and relationships inside the field of information sciences could be useful for the understanding of scientific collaboration. In other words, can bibliometric indicators follow interdisciplinarity and development of the field of information sciences? Our research shows that bibliometric data cannot offer the answer to that question. More precisely, bibliometric data can offer only fragmented answers based on quantitative indicators about authors and mentors' citations, co-citations of authors according to disciplines, multiple authorship, etc. What is missing in bibliometric approach to scientific collaboration is conceptual framework: what is the measure, what is the instrument for measuring, and what is the measuring unit with which we can follow scientific collaboration.

We believe that the postulates that co-citation analysis is not an image of cognitive structures and that co-word analysis is nearer to cognitive networks, are basically correct, but not precise enough. Future research should define that by co-word analysis it is possible to precisely follow the dynamic of the development of research topics, the development of scientific fields and conceptual relationships, and the relationships among scientific disciplines.

In this paper we advocate thesis that cohesion and coherence of scientific field should be conceptual framework, i.e. criterion for the research of scientific collaboration. With such an approach we can research scientific collaboration realized through scientific networks, institutional networks, communicational networks and cognitive networks. All these forms of collaboration, i.e. all these networks can influence the structure and dynamics of the development of (information) sciences. In that context, bibliometric data used as indicators of cohesion and coherence of information science are not only quantitative indicators, but could also be used as quantitative data of qualitative indicators that we could define by new conceptual framework.

To conclude. In this paper we did not change bibliometric indicators, we only pointed to the possibility that they could be used differently than up to now. We advocate their use as indicators of cohesion and coherence as a measure of scientific collaboration.

References

[1] Ardanuy, J., Urbano, C. & Quintana, L. **The evolution of recent research on Catalan literature through the production of PhD theses: a bibliometric and social network analysis**. *Information Research*, **14**(2), 2009, paper 404. URL: http://InformationR.net/ir/14-2/paper404.html (9.04.2010)

[2] Bublitz, Wolfram. **A Bibliography of Coherence and Cohesion**. URL: http://www.philhist.uni-augsburg.de/lehrstuehle/anglistik/sprachwissenschaft/bibliography/ (17.02.2010.)

[3] Capurro, Rafael; Chaim Zins. **Knowldege Map of Information Science. Rafael Capurro's responses to Chaim Zin**s. (2006)(URL www.capurro.de 10.4.2009.)

[4] Diodato, Virgil. **Dictionary of Bibliometrics.** Haworth Press, New York: 1994.

[5] Glänzel, W. **Bibliometrics as a Research Field. A course on theory and application of bibliometric indicators.** Course Handouts, 2003.

[6] Katz, Sylvan; Martin, Ben. **What is research collaboration?** *Research Policy*, 26, 1997, p.p. 1-18.

[7] Pečarić, Đilda. **Razvoj informacijskih znanosti u Hrvatskoj. Bibliometrijska analiza doktorskih disertacija iz informacijskih znanosti 1978. - 2007.** Zagreb: doktorska disertacija, rukopis, 2009.

[8] Pečarić, Đilda; Tuđman, Miroslav. **Položaj mentora u znanstvenoj zajednici. Bibliometrijska analiza mentora u informacijskim znanostima u Hrvatskoj od 1978. do 2007. godine.** XVII. Međunarodni znanstveni skup "Društvo i tehnologija 2010", Zadar 28.-30. lipnja 2010. godine.

[9] Pehar, Franjo. **Komunikacijska uloga hrvatskih časopisa u polju informacijskih znanosti: Bibliometrijska analiza Vjesnika bibliotekara hrvatske i Informatologije.** Filozofski fakuletet, Zagreb: Doktorska disertacija, 2010.

[10] Đurđica Težak i suradnici. **Profesor Božo Težak, lučonoša znanosti** / Zagreb : Hrvatska sveučilišna naklada, 2007.

[11] Tuđman, M., Milas, M., Šilović, N., Boras, D. **Bibliometric Analysis of Master Theses in the Information Sciences (Postgraduate Studies in Librarianship, Archivistics, Museology and Information Science, 1961-1984, Zagreb, Yugoslavia.** *Informatologia Yugoslavica,* (6), 1984, p.p. 31-40.

[12] Tuđman, M., Tudor-Šilović, N., Boras, D. Milas-Bracović, M. **A Literature Measure of Scientific Communication: Co-Citation Analysis of Master Theses in Information Sciences in Yugoslavia, 1961-1984**. In: Neva Tudor-Šilović and Ivan Mihel (ed.): "*Information Research. Research methods in library and information science*" Taylor Graham, London, 1988, p. 261.

[13] White, Howard. **Cross-Textual Cohesion and Coherence**. URL: http://www.pliant.org/personal/Tom_Erickson/DA_White.pdf (10.04.2010)

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

Appendix

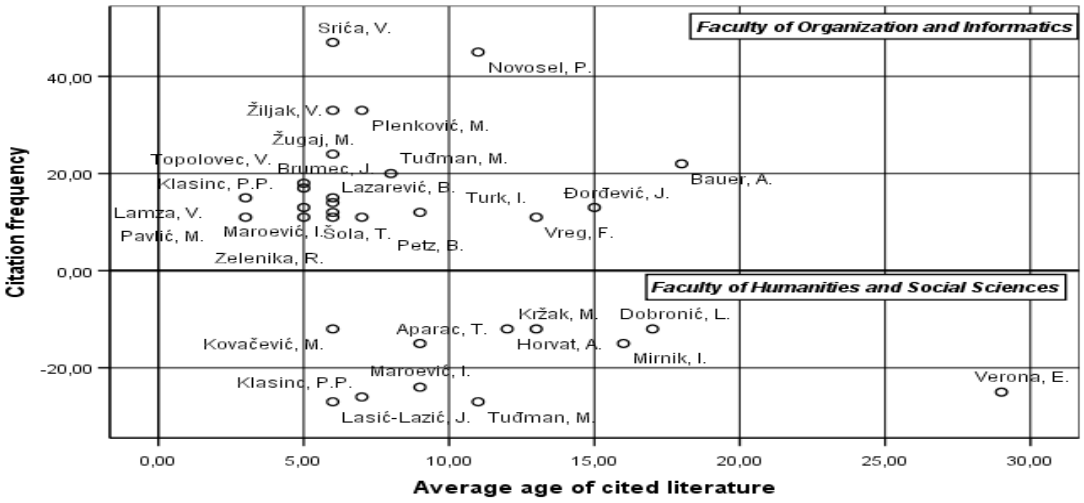

Figure 1. Most cited authors (FHSS and FOI)

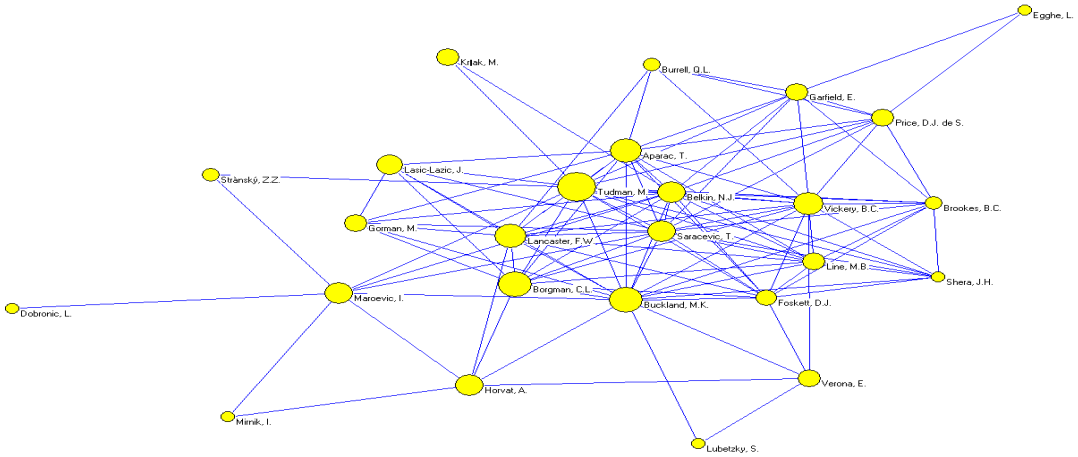

Figure 3. Co-citation overview, of the most cited authors from PhD theses done at FHSS

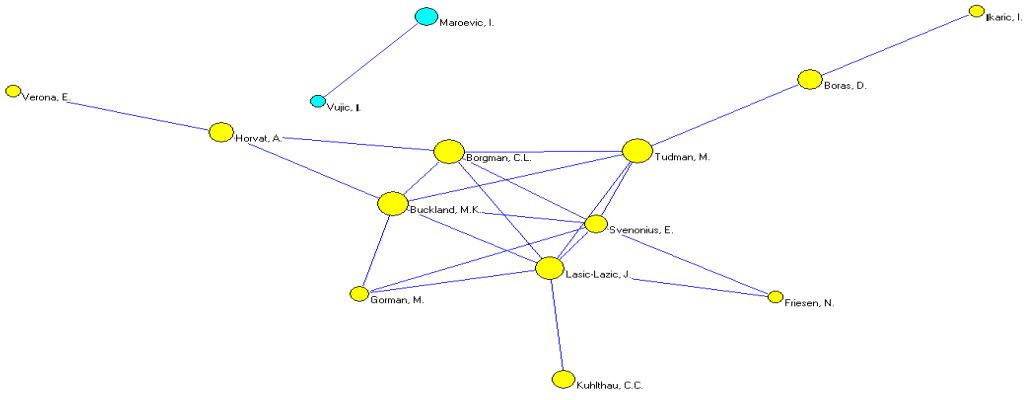

Figure 4. Co-citation overview, the most cited authors from PhD theses done at FHSS from 2000 to 2007

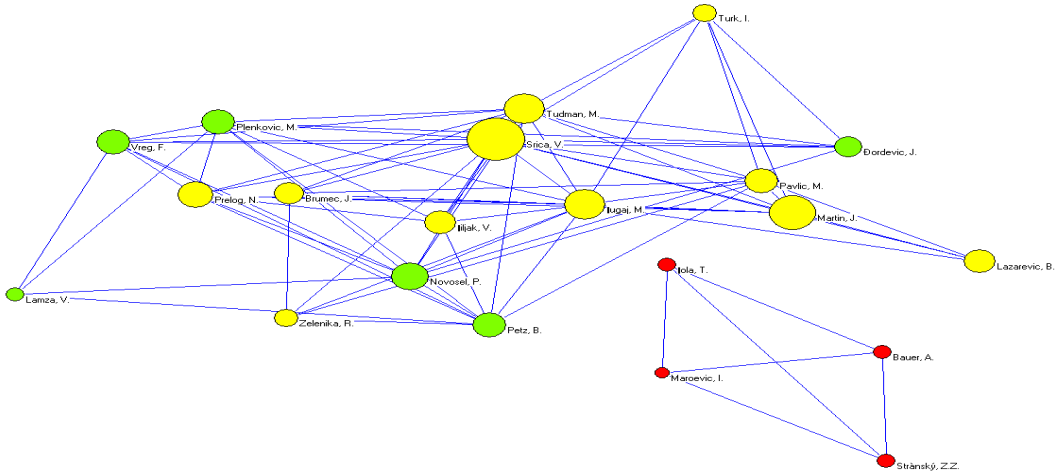

Figure 5. Co-citation overview, the most cited authors from PhD theses done at FOI

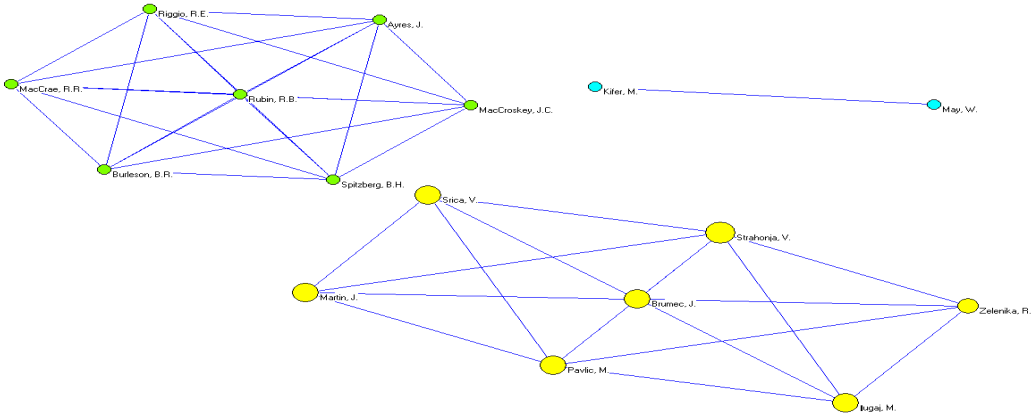

Figure 6. Co-citation overview, the most cited authors from PhD theses done at FOI from 1997 to 2007

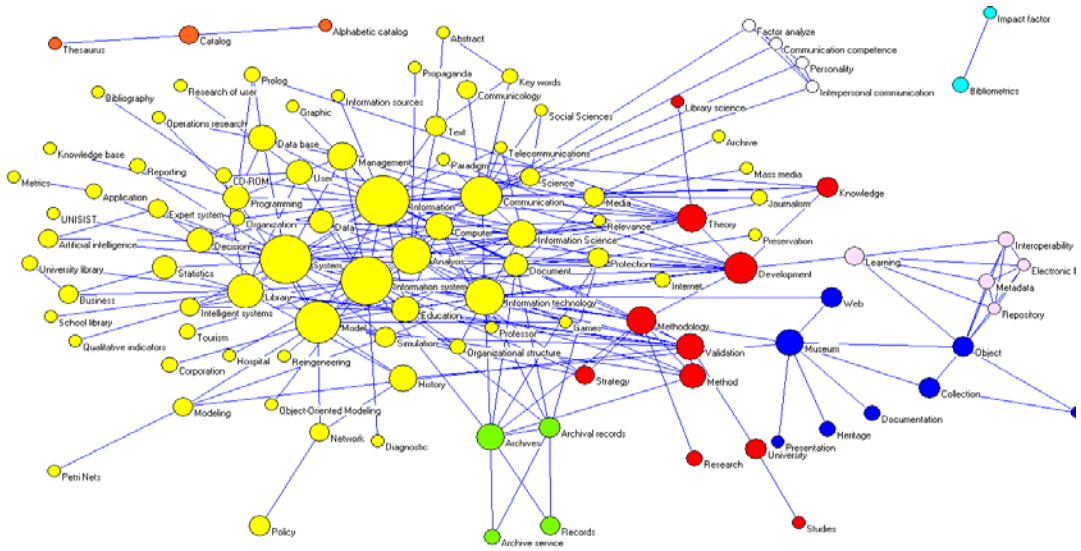

Figure 7. Overview of the most cited pairs of key words from indexed doctoral dissertations (1978 – 2007)

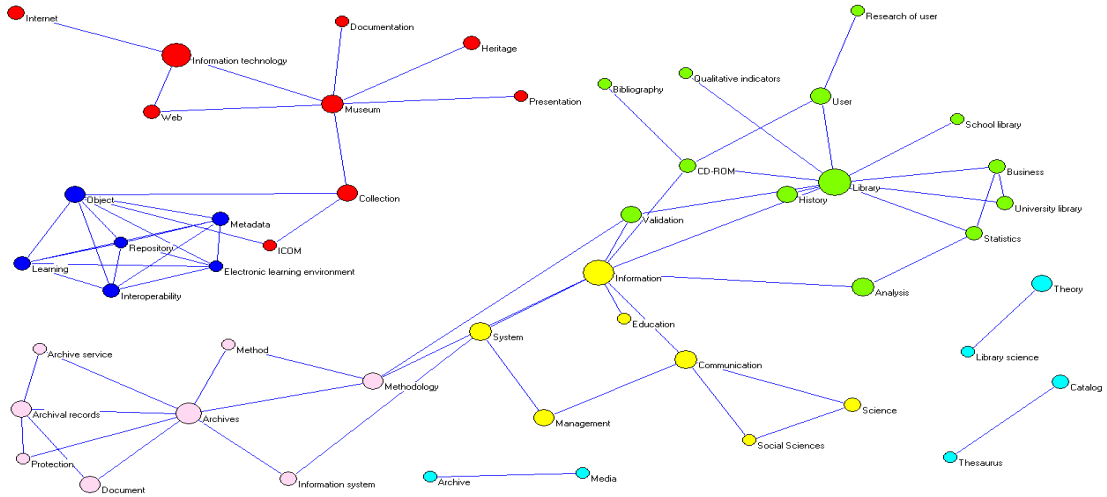

Figure 8. Overview of pairs of key words – indexed dissertations done at FHSS

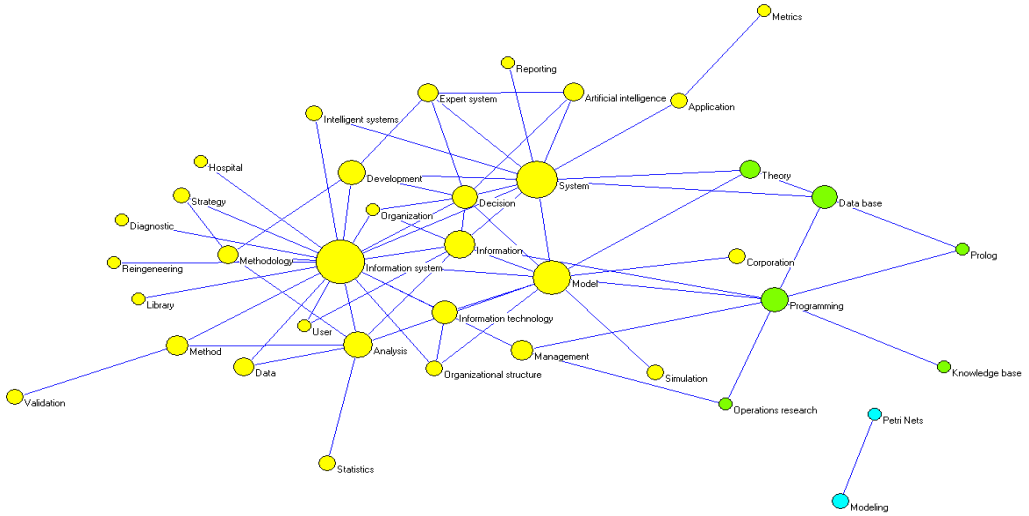

Figure 9. Overview of pairs of key words - indexed dissertations done in information systems

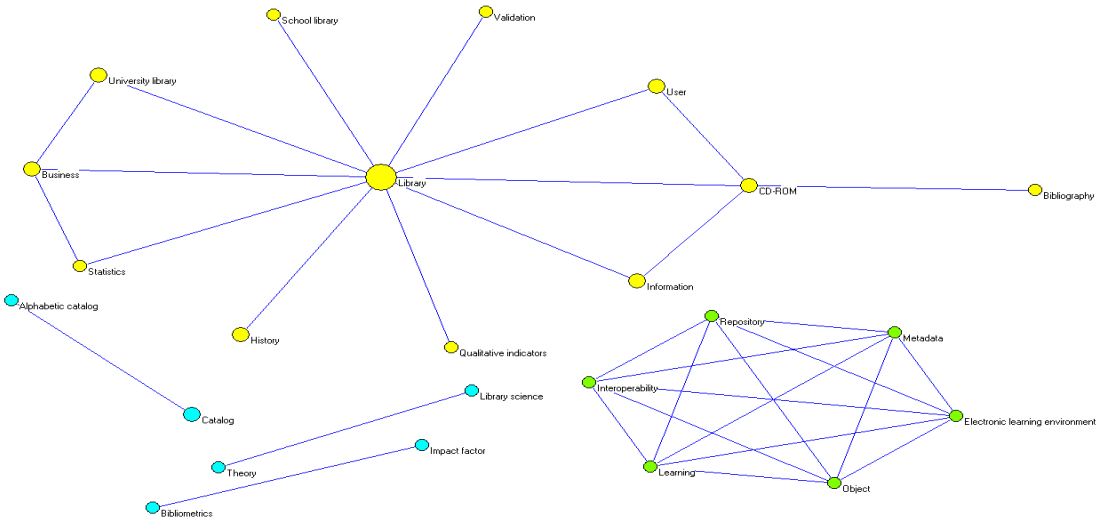

Figure 10. Overview of pairs of key words - indexed dissertations done in librarianship