Relationship Between Scientific Paradigm and Research Front. On Example of Information Science Research Production

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

Summary

Empirical indicators about time distribution and „obsolescence" of cited references are retrieved from doctoral dissertations in Information Science from 1978 to 2007 at Croatian universities. Cluster of most cited authors who are recognized as key authors for Information Science paradigm is gained from citation and co-citation analysis. Domination of these authors constitutes "conceptual knowledge zone" which is placed according to time distribution at the end of axes of obsolescence of cited literature. As opposed to that, research front, which is the period of most intensive research activity, is placed in the first time quarter (from zero to four years old cited literature), and there are 10% of most cited authors. Our research follows transit of researcher from research front to zone of paradigmatic knowledge. Hypothesis is that new authors who enter conceptual knowledge zone suppress "old" authors; so it can be concluded that incomers in conceptual knowledge zone are holders of new theoretical approaches and solutions for new problems. Duration and importance of "old" authors ensure paradigm and methods for solution of "old" problems, respectively for the production of professional papers and traditions of profession.

Key words: Information Science, Research Front, Scientific Paradigm, Obsolescence of Cited Literature, Conceptual Knowledge Zone

Introduction

At the beginning of 1960's E. Garfield and Irving H. Sher defined „journal impact factor" as a criterion for journals selection for Science Citation Index, which Institute for Scientific Information (ISI) in Philadelphia began publishing in 1961. Since then, the impact factor has become one of the basic, most prevailing and most used criterion in the Information Science for the evaluation

of scientific journals and scientific papers. After that strong pressure on researchers, scientific journals and scientific institutes begins, because their "effectiveness" and relevance in scientific community is determined by measuring impact factor and journal citation frequency. Entire scientific community lives under pressure that they have to be cited in the shortest possible time, because impact factor became key criterion for promotion of scientists, evaluation of scientific institutes, financing of scientific journals[1].
On the other hand, the research of scientific development ie. "maps of sciences" and "cognitive structures of science" display and recognise the most cited authors in sciences and scientific disciplines as key authors in prevailing scientific paradigms[2]. Paradox is that the Information Science has not examine relationships between first and second group of authors: research front determined by impact factor, ie. speed of citation on one side, and continuity of dominant authors in certain scientific paradigm, on the other side. Our interest is to explore this relationships and alterations inside first and second group of authors respectively within "research front" and "zone of conceptual knowledge".
Research is done on 134 doctoral dissertations in Information Science at Croatian universities from 1978 to 2007 (Đ. Pečarić, 2009).[3]

About constants in scientific communication and about differences in communications models

In another research, we explored, by citation analysis, features of communication models that are dominant in scientific communication (M. Tuđman, Đ. Pečarić, 2009.). The corpus of 22,210 cited bibliographic units is analyzed[4]. On the basis of citation frequency according to the age of cited

[1] See E. Garfield, 2006.; Respectively, M. Jokić (2005.) „Fundamentally, impact factor is ratio between citations and recent citable items published in the same period."

[2] Overview of those research are in H.D. White i K.W. McCain (1998)

[3] In the period from 1978 to 2007 at Croatian universities 134 doctoral dissertations were done in seven different Information Science disciplines: 20 in librarianship, 21 in information science, 53 in information systems, 22 in communicology, 9 in museology, 8 in archivistics and documentation, 1 in lexicography. The majority of doctoral dissertations were made at the Faculty of Organization and Informatics in Varaždin (FOI) – 69, followed by the Faculty of Humanities and Social Sciences in Zagreb – 49 doctoral dissertations. According to the periods of production: 21 doctoral dissertations were made until 1989; 62 doctoral dissertations from 1990 to 1999; 51 doctoral dissertations from 2000 to 2007

[4] From total number of citation (22,210), there are 17,178 cited units with authors, that is, the total number of cited authors is 10,683. There are 8,296 (77.65%) authors that are cited just once, and 2,387 or 22.34% of authors are cited more then once. Those 2,387 authors, that are cited more than once, hold 51.71% of citation. The rule that a small number of authors are often cited repeats again: 451 authors that are cited 5 or more times hold 23.61% of citations; 118 authors that are

literature we determined existence of constant in scientific communication models.

Communication model[5] has several unchangeable characteristics. Regardless of the variable used (citation frequency, self-citations, citations according to languages or distribution of citation that are cited only once) citation distribution curve is always the same or similar. This is also confirmed when the data are fragmented according to scientific discipline, as well as time periods or faculties on which doctoral dissertations are made.

Second, when we know cited half-life (***t/2***), period in which 50% of documents are cited, then first 25% of documents are cited until half of cited half-life (***t/4***). In time period ***t/4*** maximum frequency from overall number of cited documents is reached. Therefore, citation curve has log-normal distribution, with maximum in time period ***t/4***.

Third, perceived regularity is that in time period between ***t/2*** and ***t***, i.e. second time period of cited half-life, following 30% of documents are cited. After double cited half-life - in which 80% of documents are cited, the last 20% of documents are cited. For those documents, ages cannot be statistically predicted.[6]

Based on previously described regularity we could identify three communication zones based on nature of citation usage. These zones are shown in table 1.

We named the first zone **empirical knowledge zone**[7], which is sequential and extends through entire communication process. This zone consists of citations of authors and documents that are cited only once. This group holds 60% of citations. Their distribution is equally distributed and presented during entire communication process.

Second zone is named **research front zone** and it is placed in communication space and time that we marked as **t/4**. In time **t/4** first 25% authors and documents are cited. In this period maximum frequency of overall document's citation is reached. Attendance and citation of authors in research front zone implicate their understanding of problem and communication with everyone in their surroundings relevant for the problem. This is the space of authors bidirectional communications in which empirical and conceptual knowledge are

---

cited 10 or more times hold 11.71% of citations. First 29 most cited authors hold 5% of citations, that is, first 49 authors hold 7% of citations (M. Tuđman, Đ. Pečarić, 2009.)

[5] See R. Vrana (2003.)

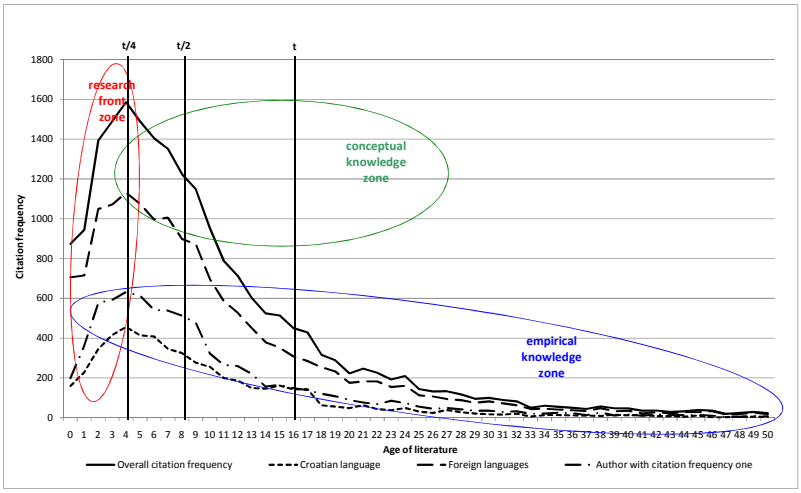
[6] See M. Tuđman, Đ. Pečarić, 2009.

[7] Basic concept comes from Capurro's (2006) definition of empirical knowledge: it is information that is the result of the process of selection in communication process. The documents and authors that are cited only once have value as empirical or theoretical information. Since these citations are used only once, we assumed that their value is more empirical than conceptual.

being overlapped, compressed and reinterpreted. In the nature of research activities, it is typical that research time is shorter from document citation half-time, although scientific, formal and informal communication can last much longer.

We named third zone **conceptual knowledge zone.** The most cited authors are in this zone, obviously because of their influence. It is logical that influence is bigger as it is more permanent. And that is why it is not strange that the age of cited literature of the most cited authors is older then citation half-live. Kuhn thesis[8] implicates that the most cited authors are cited primarily for referencing on dominant theories, for solutions of scientific problems. Referencing on mutual scientific paradigm, which is defined by influential scientist, binds members of certain communication community.

Table 1. Citation frequency and knowledge zones.



Authors alterations in conceptual knowledge zone

In table 2, 45 most cited authors in Information Science from 1978 to 2007[9] are displayed. If we know that half-time of cited documents is 7.5 years[10] it can be

[8] See Kuhn, T. 1999.

[9] Table 2 shows authors that are cited 14 or more times. In the right top corner of the table authors cited in foreign languages are shown. Authors cited in Croatian language are shown in the left bottom corner of the table. In the bottom right corner of the table there are publications that are published in SFRJ, but outside of Croatia.

[10] Certain differences exist according to the type of cited documents: for monographs’ cited half-life is 9.1 years, for journals it is 7.2 years, and for semi-publications it is 9.3 years. The

concluded that the most frequently cited authors are those whose cited publications are from 7.5 to 30 years old.
Since we are interested in the relationship between authors in research front zone and conceptual knowledge zone, primarily we have to establish how authors' alteration in conceptual knowledge zone occurs. That is why we analyzed the most cited authors according to periods of development of Information Science in Croatia from 1978 to 2007. The data are compared with several criteria: according to Information Science disciplines and according to faculties on which doctoral dissertations are made. In this paper we give only basic determinants, in order to indicate trends and to make conclusion about authors' alteration, if it exists, in the conceptual knowledge zone.

Table 2: The 45 most cited authors in Information Science from 1978 to 2007.

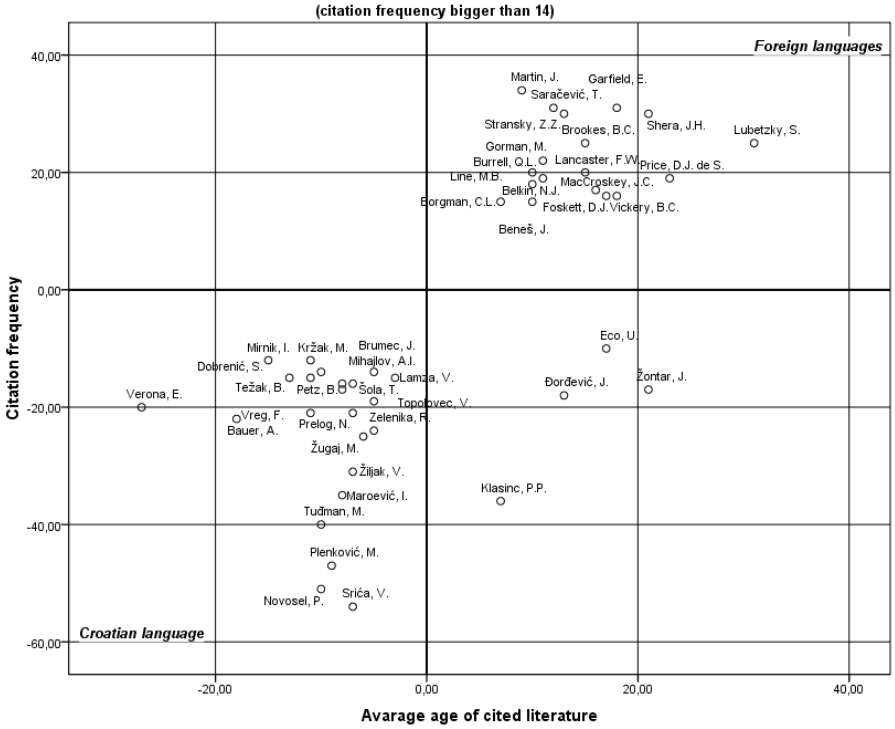
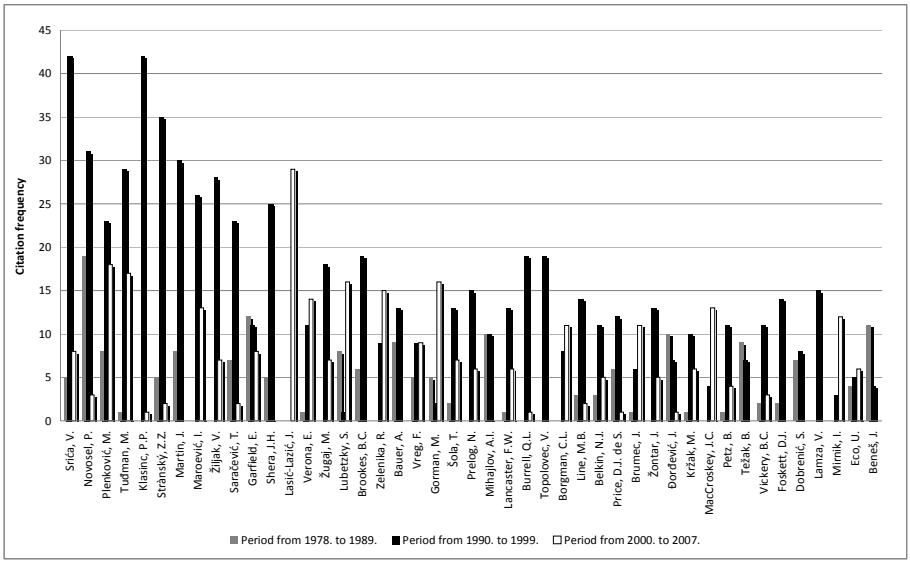


Table 2 displays 45 most cited authors in Information Science from 1978 to 2007 in 134 doctoral dissertations. Clearly, the display of the most cited authors according to disciplines, or according to time periods, will be different than it is

difference exists also among disciplines. The shortest cited half-life is in information systems 5.9 years, and the longest in museology 12.6 years. For information science it is 7.1 years, for librarianship it is 7.8 years, for communicology it is 8.5 years, and for archivistics and documentation it is 8.6 years.

shown in table 2. Yet our interest is not in cited authors, but in regularity upon which authors' alteration in certain knowledge zone happens.
This paper will discuss primarily the differences that occur during three different time periods[11]. Time periods are arbitrarily divided into 10-year periods. In the first period (from 1978 to 1989), 31 authors out of the 45 most cited authors are cited. It is important to notice that these 31 most cited authors' hold 6.28% citations from overall number of cited documents in that period. Since most citations are older than 7.5 years, it means that those authors hold more than 10% of the citation in second part of cited half-life[12]. In this period first 7 the most cited authors are: P. Novosel, E. Garfield, J. Beneš, A. I. Mihailov, J. Đorđević, A. Bauer, B. Težak.

Tabel 3: 45 most cited authors according to periods



In the second period (from 1990 to 1999), 44 out of the 45 most cited authors are cited. Even 8.19% citations from all citation in this period are held by these 44 authors, i.e. 0.9% authors out of 5094 authors cited in second period. Since most citations are cited in second part of citation obsolescence half-life it means that 44 authors hold almost 15% of citation. First 7 of the most cited authors in this period are: V. Srića, P. P. Klasinc, Z. Z. Stránský, P. Novosel, J. Martin, M. Tuđman, V. Žiljak.

[11] All of the following data are from Đ. Pečarić (2009, manuscript)

[12] In this period (1978.-1989.) overall number of cited authors is 1952, out of which 1582 authors are cited only once, which means that from overall number of cited authors 1.6% (from 31 most cited authors) of authors hold 6.28% of citation.

In the third period (from 2000 to 2007), 39 out of 45 the most cited authors are cited. Those 39 authors from overall of 4611 authors cited in third period hold 4.28% citations from overall number of citation. As the majority of these authors are cited in the second part of citation obsolescence half-life, we can see again that a small number of authors, 0.8% of the most cited authors, hold 8% of citation in the second part of citation obsolescence half-life. First 7 of the most cited authors in this period are: J. Lasić-Lazić, M. Plenković, M. Tuđman, S. Lubetzky, M. Gorman, R. Zelenika, E. Verona.

Authors' space and placement in conceptual knowledge zone are neither constant nor lasting. In analyzed range of 30 years, only 22 authors out of 45 of the most cited authors are cited in all three periods. This means that 50% of authors are not cited in all periods. In the first period, 31 out of 45 authors are cited, i.e. 14 authors are not cited, some of which are among the most cited authors in following periods (e.g. P. Klasinc, I. Maroević, J. Lasić-Lazić, V. Žiljak, N. Prelog). In the second period, 44 out of 45 most cited authors are cited. And in the third period, among 34 cited authors there are not present 11 authors, some of which are founders of Information Science (such as J. Martin, J. Shera, B. C. Brookes, A. Bauer, A. I. Mihailov, B. Težak, D. J. Foskett, S. Dobrenić).

These quantitative determinants cannot be the foundation for taking qualitative conclusions, because data can be re-arranged according to different criteria. For us it can be interesting to know which of 45 of the most cited authors are cited in doctoral dissertations from different disciplines in Information Sciences. Only one author (M. Tuđman) is cited in all seven disciplines; six authors (V. Žiljak, T. Saračević, J. Lasić-Lazić, N. Prelog, N. J. Belkin, and B. Petz) are cited in five disciplines. Ten authors are cited in four disciplines, while five authors are cited in three different disciplines. Remaining 23 out of 45 the most cited authors are cited in one or two disciplines.

These data alone indicate that neither authors' "lasting", i.e. their presence in all three periods, nor citation frequency are sufficient argument for evaluation of certain authors influence. The same rule applies for authors' citation in different Information Science disciplines. In fact, the most cited authors are not cited in all disciplines; authors that are cited in most disciplines are not cited in all periods. Exactly this authors' "vicissitude" is what we want to detect and describe.

Researchers, Scholars and Predecessors

Table 4 shows 22 out of 45 most cited authors that are cited through all three periods. When the data are analyzed according to average age of cited literature within certain period, it can be concluded that every analyzed period consists of three time zones. This can be illustrated by following examples.

Predecessors' time

In all three periods after the obsolescence of cited literature in time (**t**) i.e. double half-life of citation obsolescence, which is 15 years for our corpus of cited literature, the authors highly relevant for the development of Information Science are cited. The fact that they are present and cited after the time of obsolescence of cited literature gives us right to call this group 'predecessors of Information Science'.

It is clear that group of authors that belong to predecessors are not the same from one period to the next, in spite of the fact that there is often overlapping. Therefore, in the first period the following authors belong to this group: E. Garfield, A. Bauer, F. W. Lancaster, Z. Z. Stranski, E. Verona, S. Lubetzky, B. C. Vickery, D. de S. Price.

In the second period 13 authors belong to the group of predecessors, of which the "oldest" according to citations are: E. Verona, D. de S. Price, J. H. Shera, S. Lubetzky, A. Bauer, B. C. Vickery, E. Garfield, etc.

In the third period 11 authors (from 34 most cited authors) belong to the group of predecessors, among which the "oldest" according to citations are: S. Lubetzky, D. de S. Price, Z. Z. Stranski, E. Verona, etc.

Table 4: Authors cited through three periods according to average age of cited literature in certain period.

	Author	Average age of cited literature		
		Period from 1978 to 1989	Period from 1990 to 1999	Period from 2000 to 2007
1	Price, D.J. de S.	19	23	38
2	Vickery, B.C.	17	21	6
3	Lubetzky, S.	17	22	39
4	Verona, E.	15	27	31
5	Stràsnký, Z.Z.	14	14	37
6	Lancaster, F.W.	13	13	21
7	Garfield, E.	13	21	21
8	Gorman, M.	10	8	12
9	Eco, U.	10	14	24
10	Đorđević, J.	9	16	23
11	Novosel, P.	8	11	8
12	Vreg, F.	8	12	13
13	Belkin, N.J.	8	11	14
14	Saračević, T.	8	13	17
15	Srića, V.	7	6	11
16	Plenković, M.	7	8	13
17	Line, M.B.	7	9	14
18	Petz, B.	5	10	4
19	Šola, T.	4	5	9
20	Brumec, J.	2	3	7
21	Tuđman, M.	2	9	11
22	Kržak, M.	0	10	14

Time of researchers

In every period authors whose publications are not older than cited half-life (in our corpus a half-life of citation obsolescence is 7.5 years) can be found. We recognize this time as time of research, and the authors in this period as the group of most cited researchers. This period contains 11 authors that we recognized as the group of researchers, from overall of 31 most cited authors in this period. The youngest citations of authors are in this group: M. Kržak, M. Tuđman, J. Brumec, T. Šola, etc.

In the second period, the group of researchers, i.e. the most cited authors according to half-life of citation obsolescence, consists of 11 researchers also, and the youngest are (in sequential order): J. Brumec, V. Lamza, T. Šola, V. Topolovec, V. Srića, etc.

In the third period, the group of researchers consists of 8 out of 34 most cited authors: Q. L. Burrell, R. Zelenika, B. Petz, C. L. Borgman, J. Lasić-Lazić, etc.

It is clear from these indicators alone that the group of authors in "time of researchers" can only be formally determined. For more profound content analyses it is necessary to determine if the authors are cited really new publications or just new translations or reprints of old publications.

Time of scholars

On time scale between time of researchers and time of predecessors, which is between citation half-life and life of literature's obsolescence, third group of authors, which we named scholars[13], is positioned. **A**uthors that belong to the group of scholars in three analyzed periods can be recognized in table 4. It is visible that alterations of authors do not happen only from one period to the next, but from one group of authors to the next. Usually, path of scholars goes from researchers to the group of scholars in order to end up in the group of predecessors that future researchers and scholars will refer to.

Conclusion

With analysis of cited bibliographic references from 134 doctoral dissertation made in Croatian universities from 1978 to 2007, several zones in scientific communication are recognized. Three zones are permanently present: empirical knowledge zone, conceptual knowledge zone and research knowledge zone. In this paper particularly is discussed authors' relationship between research knowledge zone and conceptual knowledge zone. We identify, by citation frequency and percentage in overall number of citation, which authors are dominant in conceptual knowledge zone. In addition, we identify alterations of dominant authors in conceptual knowledge zone. On time axis determined by the age of cited literature, we identify three groups of authors in conceptual

[13] Scholar – in the authentic meaning of the word, it is a person who has improved knowledge, a learned person, scientist. (Anić, 2003.)

knowledge zone. According to chronological order first group of authors consists of predecessors, i.e. those authors that precede scholars and researchers. Publications of those authors are mostly older than double obsolescence half-life of cited literature. The second group consists of scholars, i.e. authors whose publications are most cited in period between cited half-life and time of knowledge obsolescence. The third group of authors consists of researchers, i.e. authors that are most cited in knowledge obsolescence half-life[14]. Based on empirical data it can be concluded that the influence of certain authors from researchers via scholars to predecessors does not depend on the publication obsolescence time, but on the sequence of factors that were not the topic of our analysis. We identify three different groups of authors in conceptual knowledge zone, as well as regularities i.e. why some authors can occur in a certain group, but not necessarily in all groups that we identified in this analysis. Just by looking at the titles of most cited authors in conceptual knowledge zone we can confirm Kuhn’s hypothesis about scientific paradigms that “incomers” suppress “old” authors, regardless of whether they work on old scientific problems in a new way or they deal with new problems. In this paper this hypothesis is shown only by quantitative indicators on alteration of authors in described zones. Only qualitative analysis of the publications of most cited authors would prove our hypothesis completely.

References

Anić, Vladimir. Rječnik hrvatskoga jezika. Zagreb: Novi Liber, 2003.

Capurro, Rafael; Chaim Zins. Knowledge Map of Information Science. Rafael Capurro's responses to Chaim Zins. (2006). www.capurro.de (2009.)

Garfield, Eugene. The History and Meaning of the Journal Impact Factor . //JAMA. 295 (2006.) 1, 90-93

Jokić, Maja. Bibliometrijski aspekti vrednovanja znanstvenog rada. Zagreb: Sveučilišna knnjižara, 2005.

Kuhn, Tomas S. Struktura znanstvenih revolucija. Naklada Jasenski i Turk. Hrvatsko sociološko društvo. Zagreb. 1999.

Tuđman, M., Tudor-Šilović, N. Boras, D., Milas-Bracović, M. A literature measure of scientific communication: Co-citation analysis of masters theses in informatin sciences in Yugoslavia. 1961-1984. In N. Tudor-Šilović, & I. Miheal (eds.), Information research methods in library and information science (pp. 225-247). London: Taylor Graham, 1988.

Pečarić, Đilda. Razvoj informacijske znanosti u Hrvatskoj. Bibliometrijska analiza doktorskih disertacija iz informacijskih znanosti 1978.-2007. Zagreb: Filozofski fakultet, doctoral dissertation, manuscript, 2009.

Tuđman, Miroslav; Pečarić, Đilda. Prilozi dubinskoj analizi komunikacijskih obrazaca. // Society and Technology 2009 Zadar (Informatologija, in press)

Vrana, Radovan. Utjecaj mrežnih izvora informacija na razvoj znanstvene komunikacije u društvenim znanostima u Hrvatskoj. Zagreb: Filozofski fakultet, doctoral dissertation, manuscript, 2003.

[14] See Kuhn, T. 1999.

White, Howard D. McCain, Katherine W. Visualizing a discipline: An Author Co-Citation Analysis of Information Science, 1972-1995. Journal of the American Society for Information Science 49 (4)(1998) 327-355