Network of Scientific Publications in Information Science

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

Summary

Bibliometric methods (bibliographic coupling and co-citation) implicitly postulate impact factor and influence factor as criteria for evaluation of „knowledge maps" or „intellectual structures". However, bibliometric methods can also be used for presentations of „Historio-Bibliography", that is genealogical tree of primordial publications. Visual presentation of history of key publications which E. Garfield (2001) named historiograph. Using corpus of bibliographic data from doctoral dissertations (1978-2009) we want to show the network of nodal publications from information science in the period from 1960 until today. Historiographs for presentations of Network of scientific publications could be generated with different methods (N. P. Hummon, P. Doreian, 1989): critical path method, bibliographic coupling method, citation method, co-citation method. We use these methods in order to show, by historiograph, historical overview of key publications in information science in Croatia. Chronological presentation of the development of scientific publication network also enables chronological analysis of certain authors' roles in scientific community from researchers and scholars to predecessors.

Key words: information science, scientific publications, network

Introduction

In all bibliometric researches Lotka's law is confirmed: a small number of authors holds the largest part of scientific production. Following that logic the most cited authors have the largest influence factor. That law is also confirmed in previous research on the corpus of doctoral dissertations in information science done at Croatian universities from 1978 to 2009. Based on gathered and analysed bibliographic data about the corpus of 170 doctoral dissertations, and 28,188 analysed references, an overview of (the most) cited authors and documents in information science according to the criteria of institutions, disciplines and time periods was retrieved. Indicators of our research follow

Lotka's distribution. Analysing 13273 cited authors it is established that the biggest number of authors (76.5%) is cited only once, and only a small number of authors is frequently cited. The most cited 43 authors (0.32% of authors) is cited 1339 times, which covers 6.1% of all citations. Citation frequency of the most cited 43 authors is in range from 18 to 68 times (in average 43 authors are cited 30 times). (Đ. Pečarić, 2010)

In this paper we wish to show the network of the most cited publications in information science using analysed corpus. Furthermore, we wish to retrieve the historiograph of key authors in information science. We will use the method which was inaugurated by E. Garfield (2001), because we believe that historiograph can be a good illustration of key publications and key authors in certain periods of information science. That is, by the selection of the most cited authors and their publications presented by historigraph we can not only understand the role of certain authors and their continual contribution to information science, but we can also present cognitive networks that show the development of information science.

Methods

The results of our research will be presented in three ways. First of all, the overview of the most cited author will be shown. The overview is retrieved by author's citation analysis of 28,188 references.

The second set of data is the overview of the most cited publications in analysed corpus. The most cited publications partially overlap with the indicators of the most cited authors. The difference is not the consequence of analysis by different methods, but rather it is consequence of the fact that the most cited authors receive citations for a series of their publications, and this also makes them relevant and influential authors. However, the criteria of citation frequency of certain publication alone is not sufficient for the construction of historiograph, neither according to disciplines nor according to time periods of information science development.

For the construction of historiograph Garfield used software for the analysis of huge amount of data. His overviews (2001) use different available methods (N. P. Hummon, P. Doreian, 1989), which we simulated by the combination of hand selection of data and an overview done by software. Since we think that historiograph be in function of cognitive network overview, together with the author we also selected: a) the year of publication's first edition, not later editions (the differences between the publishing year of original publication and the publishing years of translated edition or republished edition can be huge); b) for authors that receive citations for bigger number of their publications, we used the publication year of their oldest publication; c) in historiograph overview according to periods, we also used the year in which the oldest publication of the author cited in that period was published. We believe that in this way we can obtain more precise insight in the sequence of the appearance

and real logic of influence of certain authors on the development of information science.

Overview of the most cited authors

As we said before, in all researches that were done on the presented corpus of data we find confirmation of Lotka's law: a small number of authors holds the biggest part of scientific production. Following that logic we can conclude that the most cited authors have the largest influence factor. Indicators in our research also follow Lotka's distribution. We already indicated that a small number of authors 43 (0.32%) received even 6.1% of all citations. In another publication (Đ. Pečarić, 2010) the most cited authors in information science according to the criteria of time periods, scientific disciplines, and institutions in which doctoral dissertations were done are shown.

In table 1 we present 43 most cited authors according to citation frequency over researched period.

Table 1: Overview of 43 most cited authors in information science

Author	Citation frequency	Author	Citation frequency
Plenković, M.	68	Malović, S.	26
Srića, V.	61	Horvat, A.	26
Tuđman, M.	57	Gorman, M.	25
Novosel, P.	55	Brookes, B.C.	25
Klasinc, P.P.	43	Zelenika, R.	25
Strànský, Z.Z.	42	Prelog, N.	24
Verona, E.	41	Bauer, A.	23
Maroević, I.	40	Line, M.B.	22
Žiljak, V.	39	Mihajlov, A.I.	22
Martin, J.	38	Anić, V.	21
Lasić-Lazić, J.	37	Kržak, M.	21
Aparac, T.	36	Borgman, C.L.	21
Shera, J.H.	35	Lancaster, F.W.	21
Garfield, E.	34	Burrell, Q.L.	20
Saračević, T.	33	Price, D.J. de S.	19
Vreg, F.	32	Topolovec, V.	19
Lubetzky, S.	31	Belkin, N.J.	19
Žugaj, M.	30	Castells, M.	19
Brumec, J.	29	Buckland, M.K.	18
Šola, T.	27	Đorđević, J.	18

Author	Citation frequency		Author	Citation frequency
Foskett, D.J.	18		Žontar, J.	18
Petz, B.	18			

Data in table 1 can serve for the comparison with similar, or related researches. Unfortunately, there are only a few of them that pursue bibliometric analysis of information science in Croatia. Likewise, researches in other scientific environments are neither methodologically nor in time correlation with our research. However, although we cannot compare those research data wholly with our research, we can find mutual points of overlapping (M. Tuđman, et al., 1984 and 1988; H. D. White and B. C. Griffith, 1981; H. D. White and K. W. McCain, 1998). Mutual points of overlapping are obviously the most cited authors in different scientific communities that are research subjects in already mentioned researches. If we compare the most cited authors in the research M. Tuđman, et. al. (1988), which pursues cited analysis of 374 master theses done at CSLDIS[1] from 1961 to 1984, then among most cited authors in both researches are:

> *D. J. Foskett, E. Garfield, L. W. Lancaster, J. Martin, A. I. Mihajlov, V. Srića, J. H. Shera, E. Verona.*

The comparison of the most cited authors also partially made possible by the research of H. D. White and K. W. McCain (1998). This research analyses 120 most cited authors in 12 leading information science journals in English speaking areas. The overlapping of the most cited authors between our research and the research of H. D. White and K. W. McCain is the following:

> *N. J. Belkin, C. L. Borgman, B. C. Brookes, E. Garfield, M. Gorman, F. W. Lancaster, M. B. Line, J. Martin, D.J. de S. Price, T. Saračević.*

Also, overlapping among all three researches exists (M. Tuđman, et. al.; H. D. White i K. W. McCain; Đ. Pečarić):

> *E. Garfield, F. W. Lancaster, J. Martin.*

The comparison of most cited authors in different researches do not allow us to draw far-reaching conclusions. As we already point out, one reason is that researches include different scientific communities in different time periods, and the second reason is that researchers use different methodologies. However, from these indications it can be concluded that core authors mutual to all scientific communities i.e. information sciences communities exist, although those authors i.e. core members, change over a time. We can assume that an alteration in the sequence and frequency of authors' citation comes together

---

[1] Centre for the Study of Librarianship, Documentation and Information Sciences, University of Zagreb, Croatia.

with the alterations in research topics and interests in certain scientific communities.

Overview of the most cited publications according to disciplines

After we gained insight into the most cited authors, we can raise the question which publications received the most citations? By answering that question we can also obtain the answer to the question which authors are cohesive factors in information science, but also which publications homogenised the entire area of information science or certain disciplines.

For that reason, according to citation frequency of certain publications criteria, we extracted the most cited publications, not authors (previous tables were done based on the authors' citation frequency and not the frequency of their publications). The overview of the most cited publications according to disciplines, with marked criterion of selection in the brackets, follows.

Archival and documentation science (citation frequency bigger than 2):
- **Beuc, I.** Arhivistika.
- **Antoljak, S.** Pomoćne istorijske nauke.
- **Kolanović, J.** Arhivistika i povijest upravnih institucija
- **Tuđman, M.** Teorija informacijske znanosti.

Information systems (citation frequency bigger than 5):
- **Srića, V.** Uvod u sistemski inženjering.
- **Strahonja, V.; Varga, M.; Pavlić, M.** Projektiranje informacijskih sustava.
- **Srića, V.** Sistem, informacija, kompjutor.
- **Lazarević, B.; Jovanović, V.; Vučković, M.** Projektovanje informacijskih sistema.
- **Radovan, M.** Projektiranje informacijskih sistema.
- **Tkalac, S.** Relacijski model podataka.

Information science (citation frequency bigger than 3):
- **Anić, V.** Rječnik hrvatskoga jezika
- **Boras, D.** Teorija i pravila segmentacije teksta na hrvatskom jeziku.
- **Chomsky, N.** Aspects of the Theory of Syntax
- **Kržak, M.** Serbo-Croatian Morpho-Spelling.

Librarianship (citation frequency bigger than 4):
- **Verona, E.** Pravilnik i priručnik za izradbu abecednih kataloga.
- **Verona, E.** Abecedni katalog u teoriji i praksi.
- **Aparac, T.** Teorijske osnove knjižnične znanosti
- **Beck, H.** Klassifikation und Informationswiedergewinnung.

Communicology (citation frequency bigger than 5):
- **Vreg, F.** Društveno komuniciranje.

- **Eco, U.** Kultura, informacija, komunikacija.
- **Novosel, P.** Delegatsko informiranje.
- **Plenković, M.** Teorija i praksa javnog komuniciranja.
- **Supek, R.** Ispitivanje javnog mnijenja.
- **Plenković, M.** Demokratizacija masmedija.
- **Plenković, M.** Suvremena radiotelevizijska retorika.
- **Plenković, M.** Teorija i praksa javnog komuniciranja.
- **Supek, R.** Ispitivanje javnog mnijenja.

Museology (citation frequency bigger than 3):

- **Maroević, I.** Uvod u muzeologiju.
- **Strànský, Z. Z.** Pojam muzeologije.
- **Strànský, Z. Z.** Temelji opće muzeologije.

From previous data we can notice several indicators. Firstly, relatively small number of publications is cited frequently. Secondly, when we look at the titles we can see that the most cited publications are ones that give theoretical frames, that is "general theory" of disciplines, thematic areas, and also norms, regulations and guides for certain professions. Thirdly, the most cited publications are mostly publications done by domestic authors, that is, publications published in Croatian (Z. Z. Strànský, F. Vreg, U. Eco). Fourthly, some authors of the most cited publications are not among the most cited authors (e.g. S. Antoljak). Fifth, some of the most cited authors are not the authors of the most cited publications. Citation frequency of certain authors is large because larger number of their publications receive citations. That is why citation frequency of publications for some of the most cited authors is small. Just to illustrate that point, following authors are the outmost case of the authors that are among most cited authors, but not one of their publications is cited more than once[2]:

Archival science: P. P. Klasinc (26); M. Kovačević (12); M. Milošević (9); M. Modrušan (9); M. Novak (10); E. Pusić (10); V. Žiljak (10).

Information science: D. W. Allen (9); A. Bookstein (8); B. C. Brookes (16); Q. L. Burrell (19); Y. S. Chen (11); D. R. Cruickshank (10); L. Egghe (13).

Information systems: S. Dobrenić (13); P. P. Klasinc (17).

Communicology: V. Lamza (17).

Librarianship: C. L. Borgman (16).

Museology: J. Beneš (15).

---

[2] In the brackets, behind authors' name is indicator of citation frequency in certain information science disciplines.

Points four and five alert us that citation frequency as the only indicator cannot be sufficient for understanding impact factor. In other words, the data about most cited authors per se are insufficient for understanding influence and/or impact of those authors in both certain disciplines and in the entire area of information science.

Historiograph or overview of cognitive networks
Graph 1 shows the overview of the most cited authors in information science. And graphs 2 to 4 show the overview of the most cited authors in doctoral dissertations in information science done at the universities in Croatia from 1978 to 2009, according to periods. These tables differ from the overview of the most cited authors which we retrieved by the citation method, that is co-citation method on the same corpus of data (Đ. Pečarić, 2009, Đ. Pečarić, M. Tuđman, 2009). The clusters of the most cited authors retrieved by co-citation analysis group authors according to citation frequency and mutual relationships. By co-citation method we can discern poles made of key authors in certain cluster, from the authors with a small number of mutual relationships that are positioned on the cluster periphery.

Although, we used the same set of data as in co-citation analysis, by historiographs (graphs 1 to 4) allow us to obtain chronological overview of the most cited authors. Historigraph is the result of the selection of the most cited authors by citation method, but the data are shown in the time sequence of the publishing of scientific publications. We have already mentioned that certain correction of data was done. It was done in the following manner: the position of certain authors on time axis is defined according to the year of first edition of cited publication (not by the year of cited reprinted editions, republished editions or the year of cited translated edition). Using this method, historiograph allowed us to obtain an overview of the most cited authors in broader context: it points to the source year of the published publication, and it also shows and connects the most cited authors in the time sequence. In this overview citation frequency is not recognisable, but it is implicitly present on graph (by the selection of the most cited authors). Also co-citation relationships between certain authors disappear.

Graph 1: Historiograph – 43 most cited authors in information science from 1978 to 2009.
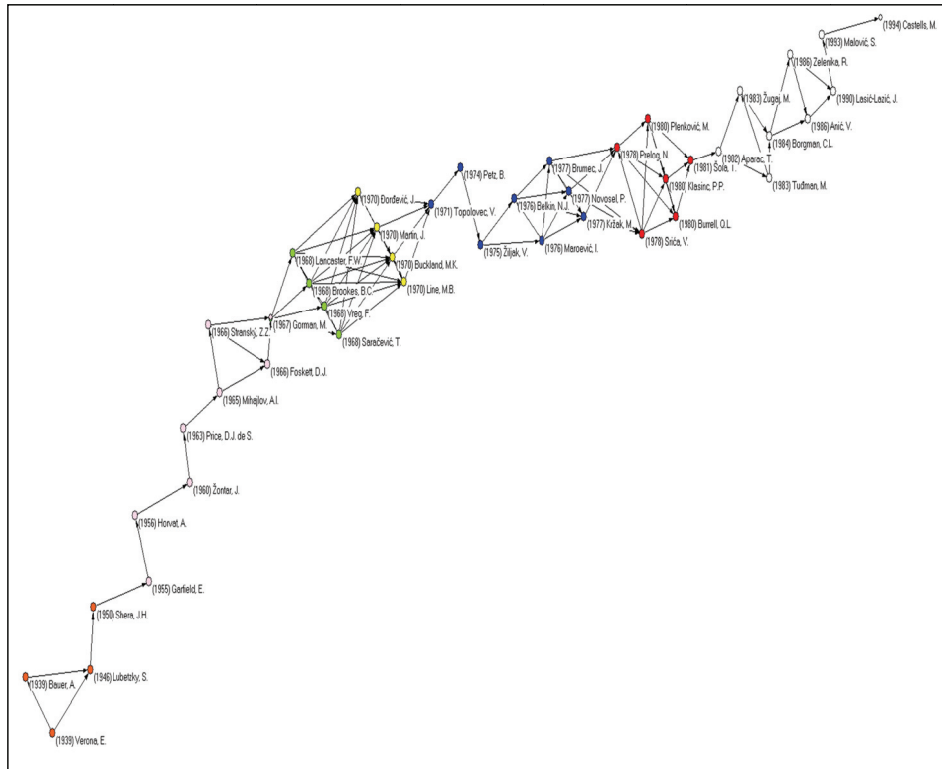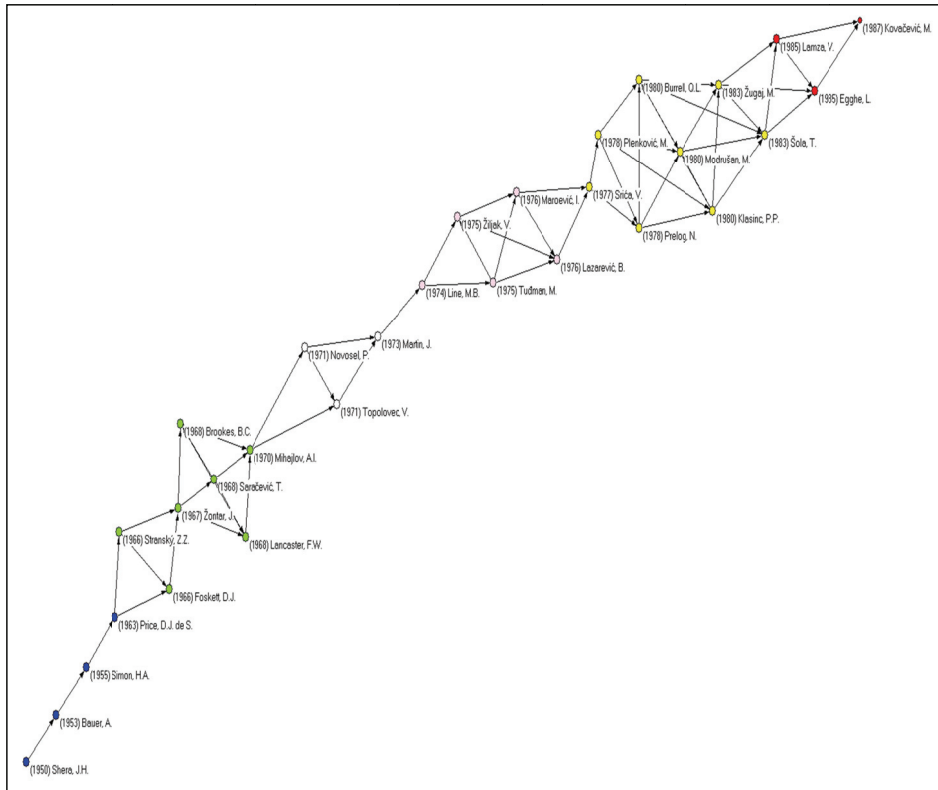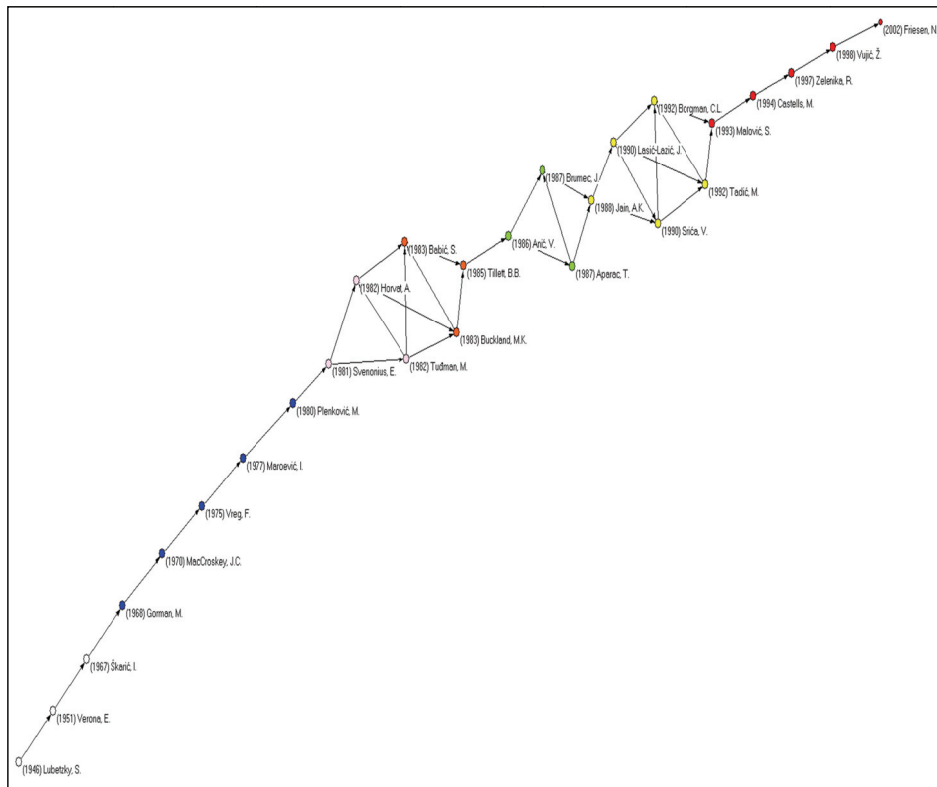
Graph 2: Historiograph – The most cited authors in information science in time period from 1978 to 1989.
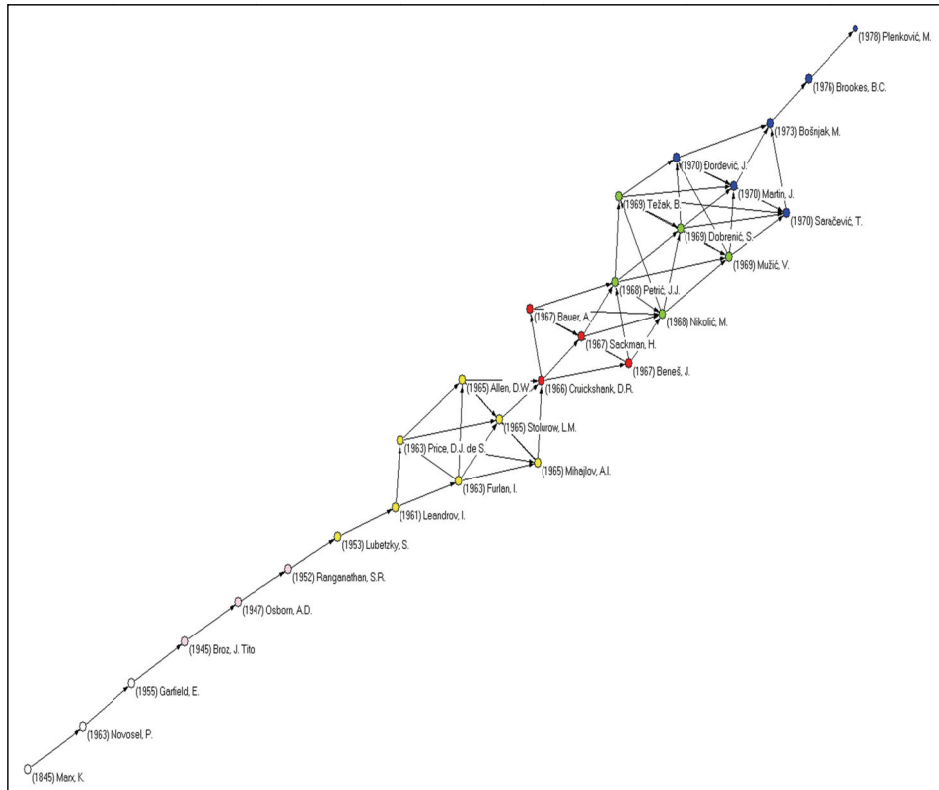
Graph 3: Historiograph – The most cited authors in information science in time period from 1990 to 1999.

Graph 4: Historiograph – The most cited authors in information science in time period from 2000 to 2009.

However, chronological links between authors are established, so we get an insight into the time of the appearance of certain authors and the life span of their influence on scientific community. That is why histograph can be used as an indicator of cognitive network (graph 1), and cognitive networks (graphs 2 to 4) if we follow authors' citation dynamics according to periods. In other words, histograph is the answer to the question who the most cited authors in certain scientific disciplines over time are. In combination with the previous overview of the most cited publications, historiograph can be used for the overview of the holders of the paradigm of certain disciplines, or key authors for some specific scientific area or topic.

As we cannot give detailed analysis of shown graphs here, we will point out several main characteristics.

The ratio of foreign and domestic authors on all graphs is similar (graph 1: from 43 authors 23 are foreign authors; graph 2: from 28 authors 15 are foreign;

graph 3: from 30 authors 13 are foreign; graph 4: from 27 authors 10 are foreign).

On graph 1 (shows historiograph for 43 most cited authors in the entire researched period) we can notice the sequence of authors that precede the time of establishment of information science. The most cited publication of domestic authors (A. Bauer and E. Verona) were done as early as 1939, and "global" predecessors H. Shera, E. Garfield, D. de Solla Price and A. I. Mihajlov are present in this corpus of data as well.

More authors published their first cited publication from 1968 to 1970, and again from 1976 to 1981. In the period from 1968 to 1970 foreign authors that published first cited publication are: T. Saračević, F. W. Lancaster, B. C. Brooks, F. Vreg, M. B. Line, M. K. Buckland, J. Martin, J. Đorđević. In the period from 1976 to 1981 published publications of domestic authors are mostly among the most cited publications: N. J. Belkin, I. Maroević, J. Brumec, P. Novosel, M. Kržak, V. Srića, N. Prelog, Q. L. Burrel, P. P. Klasinc, M. Plenković, T. Šola.

It is interesting that on all graphs all cited publications are older than the half-life of cited literature; half-life of analysed 28,188 references is 7.5 years. Citation half-life it remains 7.5 years for the analysed references. This confirms our postulate about the possibility of recognising the role of scientists as predecessors, scholars and researchers by bibliometric methods, and thus also by historigraps (Đ. Pečarić, M. Tuđman, 2009).

On graph 2 (the first period in the research of the development of information science: 1978-1989) among predecessors are both K. Marx and J. Broz Tito. This is yet another indicator that the ruling ideology took its toll on scientific production in one-party, undemocratic systems. The largest number of the most cited authors in this period published their publications from 1963 to 1970.

It is interesting that on graph 3 (the overview of the most cited authors in information science in period from 1990 to 1999) in the lower part of the graph we can notice the dominance of foreign authors (their publications were published from 1955 to 1974), and in the upper part of the graph dominant authors are domestic authors (their publications were published from 1975 to 1987)

These differences are not noticeable in the last observed period (graph 4), although only 10 foreign authors are present from overall 27 most cited authors. However, what can be noticed here is that the life span of the most cited authors is large: from 1946 to 2002.

If we would compare data in all four historiographs we could notice several other phenomena. First of all, we can notice the differences in the lists of the most cited authors according to periods. Some authors are repeated and some authors, in time, cease to be cited frequently. But the second indicator is also interesting. Authors that receive citations in different periods frequently receive citations for other publications in new period. This indicates double dynamic of

development: scientific areas develop and in the course of time new authors with new ideas and publications enter But at the same time authors themselves develop, because their publications cause interest of new researchers, and old publications – publications that were among the most cited publications in previous periods – are not cited any more.

Conclusion

The method of historiograph was illustrated by E. Garfield in 2001, however until today it has not found broader implementation, particularly not in Croatia.

In this paper we advocate the use of historiograph for the overview of scientific development, more precisely for the development of information science. We believe that this technique of historiography provides far better diachronical overview of the development of certain scientific areas than would be possible to show by clusters, both the clusters of co-authors and clusters of co-words.

There certainly exists a number logics and techniques to obtain historiographs. In this paper we presented only one of possible approaches. Our aim was not to reconstruct scientific paradigm of information science, i.e. key authors, as much as to advocate a new method and technique of research and presentation of the development of information science.

A particular challenge would be the development of appropriate software for the preparation of historigraph. It should be made in such a manner as to enable a larger number of researchers to use this method. Alike challenge is the development of historiograph by key words. This technique would improve the research of the development of scientific disciplines. We believe that this technique would enable scientists to follow topics and thematic areas over longer time periods.

References

Garfield, Eugene. From Bibliographic Coupling to Co-Citation Analysis via Algorithmic Historio_Bibliography. 27.11.2001. URL: http://EugeneGarfield.org (17.023.2011)

Hummon, N.O.; Doreian, P. Connectivity in a Citation Network: The Development of DNA Theory. ISocial Networks 11 (1989),1, 39-63

Pečarić, Đ. Relationship between scientific paradigm and research front. On Example of Information Science Research Production. // INFuture2009 – Digital Resources and Knowledge Sharing / Stančić, H.; Seljan, S.; Bawden, D.; Lasić-Lazić, J.; Slavić, A. (ed.). Zagreb: Department of Information Sciences, Faculty of Humanities and Social Sciences, University of Zagreb, 2009, 767-777.

Pečarić, Đilda. Razvoj informacijskih znanosti u Hrvatskoj. Bibliometrijska analiza doktorskih disertacija iz informacijskih znanosti 1978.-2007. (Doctoral dissertation). Zagreb : Filozofski fakultet Sveučilišta u Zagrebu, 2010.

Pečarić, Đ., Tuđman, M. Predecessors, Scholars and Researchers in Information Sciences. Contribution to Methodology for Bibliometrics Anaysis of Scientific Paradigms. // INFuture2009 – Digital Resources and Knowledge Sharing / Stančić, H.; Seljan, S.; Bawden, D.; Lasić-Lazić, J.; Slavić, A. (ed.). Zagreb: Department of Information Sciences, Faculty of Humanities and Social Sciences, University of Zagreb, 2009, 713-726.

Pečarić, Đ., Tuđman, M. Theoretical difference between impact factor and influence factor. //*CECIIS 2009 Proceedings*/ Aurer, B.; Bača, M.; Rabuzin, K. (ed.). Varaždin: Faculty of Organization and Informatics, 2009, 23-28.

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

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. // *Information Research. Research methods in library and information science/* Tudor-Šilović, N.; Mihel, I. (ed.). London : Taylor Graham, 1988, 225-247.

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

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 (1998), 4, 327-355.