

Information Science during the First Decade of the Web: An Enriched Author Co-citation Analysis

Dangzhi Zhao¹

School of Library and Information Studies, University of Alberta, Edmonton, Alberta, Canada T6G 2J4.

Phone: (780) 492-2814, Fax: (780) 492-2430, Email: dzhao@ualberta.ca

Andreas Strotmann

School of Business, University of Alberta, Edmonton, Alberta, Canada T6G 2J4.

Phone: (780) 492-6924, Fax: (780) 492-3325, Email: andreas.strotmann@ualberta.ca

Abstract

Using an enriched author co-citation analysis (ACA) approach, we map Information Science (IS) for 1996-2005, a decade of explosive development of the World Wide Web, to examine its development during these years compared to previous decades examined by the influential study of IS by W&M98, which demonstrated the power of ACA in visualizing disciplines.

The Web, we find, has had a truly profound impact on IS, driving the creation of new disciplines and the revitalization or obsolescence of old ones, but most importantly, beginning to bridge at last the ancient chasm between the “literatures” and “retrieval” IS camps. Simultaneously, the development of the field towards cognitive aspects has intensified.

We take the opportunity of this enriched ACA study of IS, which employs both orthogonal and oblique rotations in the Factor Analysis (FA), and which reports both pattern and structure matrices for the latter, to compare the relative merits of these several FA methods in ACA. Each provides interesting information not available from the others, we find, especially when FA results are also visualized in the novel manner we introduce here. We verify these methodological findings in a brief ACA of the XML research field.

¹ Corresponding author

Introduction

Although it has had a long history, the mapping and visualization of knowledge domains has attracted great interest in recent years both within and outside of the Library and Information Science / Studies field (LIS). This is probably due to a number of factors. For example, written knowledge has become available in digitized forms in recent years with the explosive development of the World Wide Web and its plethora of online databases, which provides huge amounts and large varieties of data for this type of studies; and the increased computing power has enabled routine analysis and visualization of such vast networks of information for social science researchers (Shiffrin & Börner, 2004).

Three main components of network analysis are (a) the objects of analysis, (b) the relationships between these objects, and (c) the mapping or visualization tools, where the former two correspond to nodes and edges of a network, and the latter to methods for representing the resulting network in an intuitively graspable manner in two- or three-dimensional space for simple viewing or interactive exploration. Traditionally, individual articles or patents, authors' oeuvres, and scholarly journals have been used as objects of analysis, citation links and word or phase co-occurrences as measures of relatedness between these objects, and multivariate analysis techniques such as Cluster Analysis, Multidimensional Scaling (MDS), and Factor Analysis as analysis and mapping tools. All have recently been adapted to the Web environment (e.g., websites viewed as journals and hyperlinks as citation links), giving rise to Webometrics (Thelwall, Vaughan, & Björneborn, 2005). New types of objects and relationships have appeared in recent studies, e.g., genes and DNA sequences, and so have new mapping methods (Börner, Maru, & Goldstone, 2004; Boyack, Börner, & Klavans, 2007; Chen, 2006; Henzinger & Lawrence, 2004; Van Eck, Waltman, & Van Berg, 2006; Wilkinson & Huberman, 2004).

Visualization of LIS

LIS is perhaps the most frequently mapped knowledge domain in the LIS literature, which is not surprising given the amount of expert domain knowledge required by this type of studies. Among these studies, White & McCain (1998), referred to as W&M98 from now on in this paper, is probably one of the most thorough studies, covering the field of Information Science (IS) for the years 1972-1995. It defined the IS field by its 12 core journals, and collected citation and co-citation data from the ISI databases through DIALOG search facilities. It used authors' oeuvres as the objects of analysis, and factor analysis and MDS to visualize the field at several stages of its development. Naturally, this

well-cited study also provided an excellent review of earlier visualization studies of the LIS field. Following the publication of this study, there was a long quiet time period with relatively few visualization studies of the LIS field, especially considering the surging interest in the visualization of knowledge domains we find in this study, until Jansens, et al. (2006) and Astrom (2007).

Jansens, et al. (2006) analyzed full-text articles from five core LIS journals covering the years 2002-2004, and Astrom (2007) articles from 21 LIS journals covering years 1990-2004 and references found in these articles as indexed in Web of Science. Both used individual articles as objects of analysis, but one used word links and the other co-citation links to measure the relationship between articles. Both used MDS and cluster analysis techniques to visualize relationships.

These two studies provided two views of the LIS field represented by individual papers. An ACA study of the field should contribute to an even more complete view of the field, and, as importantly, of studies on visualization of knowledge domains, as it can help researchers contrast the two article-based studies to compare different types of visualization studies.

The present study examines the IS field defined in the same way as W&M98 for the years 1996-2005, i.e., the decade following the time period examined in that study using ACA techniques. This will allow us to compare with that study in order to see how the IS field changes, in addition to a static view of the field during this decade. We also take this opportunity to attempt to clarify a few confusions that exist in ACA, and to introduce a novel visualization technique for factor analysis results in ACA.

Author Co-citation Analysis (ACA)

Since its introduction by White & Griffith (1981), author co-citation analysis (ACA) has gained great renown in the study of intellectual structures of scholarly fields and of the implied social structures of the corresponding communities, and as a result has become an important area of information science (White & McCain, 1998; Zhao & Strotmann, 2007).

ACA is one particular type of co-citation analysis. In co-citation analysis, a set of items (documents, authors, journals, etc.) is selected to represent a research area which can be as large as an entire science or as small as a single specialty within a research field (Small, 1999). Relationships between these items are then analyzed by using their co-citation counts as similarity measure and multivariate analysis techniques such as Cluster Analysis and Multidimensional Scaling (MDS) as analysis tools in order to study the intellectual structure of this research field and to infer some of the

characteristics of the corresponding scientific community. In general, two items are counted as co-cited when they appear together in the same reference list of an article. ACA takes the author as the unit of its analysis, uses co-citation counts to measure the relationships between oeuvres of representative authors, and applies multivariate analysis techniques such as Factor Analysis to study the underlying structure of a research field as represented by these oeuvres.

Most ACA studies have applied the fairly consistent general steps and techniques of classic ACA (De Mey, 1982; McCain, 1990a) to different research fields with little or no modification. However, in recent years, interest in exploring alternatives has resurged. Some studies have proposed new techniques for mapping author clusters (White, 2003; Chen, 2006); others have explored how to go beyond first-author counting in ACA with the support of newly available citation analysis data sources (Schneider, Larsen, & Ingwersen, 2007; Zhao, 2006); still others have challenged some of the statistical techniques used in classic ACA, two examples of which being Ahlgren, Jarneving, & Rousseau (2003) on whether Pearson's r or cosine measures should be used to define similarity in ACA, and Leydesdorff & Vaughan (2006) on whether a symmetric co-citation matrix or an asymmetrical citation matrix should be used as input for mapping purposes.

These studies, including Ahlgren, Jarneving, & Rousseau, 2004a & 2004b; Bensman, 2004; White, 2003 & 2004, have contributed to the clarification of some existing measures and techniques and to the development of new ones, which may result in improved ACA studies. The present study aims to further contribute to this area of research by providing empirical evidence to help clarify a few issues in the use of Factor Analysis in ACA.

Factor analysis in ACA

Factor Analysis (FA) has been used in ACA from the very beginning (White & Griffith, 1981). Its goal is to seek a small number of meaningful underlying factors that explain the relationships among a larger number of items, and is used in ACA to reveal specialty structures of research fields and authors' memberships in one or more such specialties. FA applied in ACA has been shown to provide clear and revealing results as to the nature of a discipline (White & McCain, 1998).

In addition to the specialty structure of a research field that other multivariate analysis techniques such as Cluster Analysis and Multi-Dimensional Scaling (MDS) can also identify, FA can assign individual authors to more than one specialty, can indicate the level of association between

authors and the specialties to which they belong, and can provide measures of “the degree of relationship between specialties” (White & Griffith, 1982, p. 260).

A researcher faces a number of methodological decisions when designing a FA study in ACA, regarding the selection of a base set of citing documents for co-citation counting, the selection of authors to analyze as representative of the field of interest, the type of data to input to the FA routine, the method for extracting factors and the number of factors to extract, which factor rotation method to apply, and which factor loadings to consider meaningful for interpreting factors. Different decisions may well produce different results (Leydesdorff & Vaughan, 2006). It is therefore important to understand the options that are available (e.g., types of input data, methods of extracting factors, factor rotation methods), the different results that may be obtained using different techniques, and when and why certain techniques may be more appropriate than others.

An examination of previous ACA studies shows some common practices in the application of FA techniques in ACA. Principle Component Analysis has been the major method used for extracting factors; Kaiser’s rule of eigenvalue greater than one the main technique for deciding the number of factors to select; both raw co-citation frequency matrices and Pearson’s *r* correlation coefficient matrices have been used as input data to the Factor Analysis procedure; both orthogonal and oblique rotations have been applied in ACA to facilitate the interpretation of results; and various levels of significance of factor loadings have been reported for interpretation, ranging from 0.3 to 0.5.

However, studies appear to be largely silent on why a particular technique was chosen instead of other techniques of the same type, nor do they compare ACA results obtained using different techniques (e.g., oblique vs. orthogonal rotation) on identical data sets to help researchers choose among them. The present study attempts to address this issue. We will focus our examination on the rotation methods of FA in ACA where confusion exists in the literature, and look forward to seeing a clarification of other aspects of ACA methodology in future research publications.

There are two types of rotation methods in FA: orthogonal and oblique. Theoretically, an orthogonal rotation assumes that resulting factors are not correlated, and works best for revealing independent dimensions of the underlying structure being studied. An oblique rotation, by contrast, is not restricted to keeping the extracted factors independent of each other, and therefore works better at separating out factors when it can be expected theoretically that the resulting factors would in reality be correlated (Hair, et al., 1998). In the case of ACA, large factors are interpreted as specialties within

a research field, and the correlation between factors can therefore be expected to be fairly high. Indeed, high correlations between factors were found in different fields, with the highest correlation between two factors reported in a given study ranging from 0.39 to 0.65 (McCain, 1990a; White & Griffith, 1982; Zhao, 2006). An oblique rotation, therefore, appears theoretically more appropriate for ACA. As an additional advantage, an oblique rotation produces a component correlation matrix to indicate the degree of correlation between resulting factors (McCain, 1990a; White & Griffith, 1982).

In practice, however, both oblique and orthogonal rotation methods have been used in ACA studies, even by the same authors, and both types of studies have demonstrated how useful and informative the factor-analytic approach to ACA is. For example, McCain (1990a) and White & Griffith (1982) applied an oblique rotation method and discussed its advantage of providing measures on the inter-relationships between factors. Later studies by these authors, such as W&M98, however, switched to an orthogonal rotation instead, without discussing why they made this choice. And yet, the orthogonal rotation in that study still produced clear and revealing results as to the nature of the IS field (White & McCain, 1998).

We have become curious about whether different rotation methods would produce different specialty structures in ACA, and whether such differences are sufficient to justify one over the other.

To summarize, the present study therefore addresses the following research questions.

- What is the intellectual structure of the IS field, 1996-2005, the first decade of the Web? How has it changed from the decades before this time period?
- What are the differences in factor analysis results in ACA between the use of an oblique rotation and an orthogonal rotation?

We hope to contribute to a more complete mapping of the IS field, and to a better understanding of the factor analysis techniques when applied to ACA, and consequently to improved ACA studies.

Methodology

In order to address the first research question, we collected citation data in the Information Science (IS) research field, and processed the data in a way compatible with W&M98 in order to better compare our results with theirs for the study of how the IS field changed during the first decade of the Web. In order to help address the second research question, we also collected and processed citation data in the XML research field in order to verify findings in the IS field.

Data collection

We define the IS field by its twelve main journals as listed in Table 1, taken with minor updates from W&M98 (p. 330). We used the Web of Science to download records of all articles published in these journals during the years 1996 to 2005. We chose “Full Record + Cited Refs” as the record format, and saved the file in “Field tagged (plain text)” format. This way, we collected 4,422 records of source papers that have references. These papers included 110,785 references altogether, i.e., 25 references per source paper on average.

Table 1: Journals used to define information science*

Information science

- *Annual Review of Information Science and Technology*
- *Information Processing & Management (and Information Storage & Retrieval)*
- *Journal of the American Society for Information Science and Technology*
- *Journal of Documentation*
- *Journal of Information Science*
- *Library & Information Science Research (and Library Research)*
- *Proceedings of the American Society for Information Science and Technology (and Proceedings of the ASIST Annual Meeting)*
- *Scientometrics*

Library automation

- *Electronic Library*
- *Information Technology and Libraries (and Journal of Library Automation)*
- *Library Resources & Technical Services*
- *Program—Automated Library and Information Systems*

* Taken with only minor updates from White & McCain (1998, p. 330)

It is, however, difficult to define the XML research field by its core journals. XML is a newly-emerging interdisciplinary research area with a core in Computer Science and a wide range of application areas both within and beyond Computer Science. Although there are now conferences and journals devoted to this research area, many XML studies also appear in journals and conferences of many other computer science areas. We therefore defined the XML research field by papers retrieved from a keyword search for “XML” or “eXtensible Markup Language” in Web of Science. This difference in defining a research field should not affect our study because results from the two research fields are not compared. Instead, the different statistical properties of rotation methods in FA are compared within each research field, and the XML field is merely used for verifying findings from the IS field.

Using Web of Science, we downloaded all records retrieved from a search for “XML” or “eXtensible Markup Language” in the topic areas (i.e., “TS”) for the time period 2001- 2005. We downloaded these records in the same way as in the IS field described above. We obtained 2475 papers² after removing articles that were not published in journals in any of the ISI Computer Science categories. Each of these papers cited about 19 references on average.

We developed Python programs to parse these downloaded records, and to store the resulting data fields such as authors, publishing sources and years of both source papers and cited references in a data structure that was convenient for later data analysis such as counting citations and co-citations, and to produce various co-citation matrices.

Data analysis

We conducted an ACA using FA to study the intellectual structure of the IS field as perceived by the authors of the 4,422 IS publications as citers. We followed the commonly accepted steps and techniques of ACA mentioned earlier (McCain, 1990a; White & McCain, 1998), except that we applied both an orthogonal and an oblique rotation in order to address our research questions. We will briefly describe these steps and techniques, and refer our readers to earlier studies for additional detail.

Given the base set of citing documents described above, a set of core authors was selected to represent the IS field based on “citedness” – the total number of citations they received as first authors. Citedness above some threshold is a good criterion for selecting authors to include in ACA although the resulting authors may not be “wholly definitive” of the research field being studied (White & McCain, 1998, p. 332). As there are no strict rules regarding thresholds for citation-based author selection in classical ACA studies (McCain, 1990a), the present study chose the 120 most highly cited individual authors to be included in the final factor analyses, even though the more authors studied the better a research field may be represented. This number is the same as in W&M98 which is higher than in most if not all earlier ACA studies and should suffice for addressing our research questions.

In classic ACA, two authors are considered co-cited when at least one document from each author’s oeuvre occurs in the same reference list, an author’s oeuvre being defined here as all the works with the author as the first author (McCain, 1990b). Based on this definition, we developed a Python program to determine co-citation frequencies in this data set of the 120 highly-cited authors

² We did this search in February 2006 and chose “2001 to present” as the date limit. The resulting data therefore also included a few papers published in 2006.

and to record them in a co-citation matrix which was used as input to the factor analysis routine in SPSS with the diagonal values treated as missing data and replaced by the mean in that routine.

Factors were extracted by Principal Component Analysis (PCA), and the number of factors extracted was determined using Kaiser's rule of eigenvalues greater than 1 because the model fit was good as indicated by total variance explained, communalities, and correlation residuals (Hair, et al., 1998). The resulting 12-factor model explains 84% of the total variance, and the differences between observed and implied correlations were smaller than 0.05 for the most part (99%). The communalities range from 0.62 to 0.94, only 4 (or 3%) of which are below 0.7 and 18 (or 15%) of which below 0.8.

We first applied an orthogonal rotation (SPSS Varimax), and then an oblique rotation (SPSS Direct OBLIMIN) to this factor model. Results are presented both as two-dimensional maps in figures and as tables in the appendices. These will be discussed in corresponding sections later in the paper. Here, we will first explain how the factor analysis results are represented in these maps and tables.

Visualization of factor structures

In previous ACA studies, ACA factor structures have been presented in table formats and cluster structures on Multi-Dimensional Scaling (MDS) maps (McCain, 1990a; White & McCain, 1998). Earlier studies using factor analysis simply placed the authors included in the studies into the factors where they load higher than a chosen threshold (e.g., Eom & Farris, 1996; White & Griffith, 1981). W&M98 introduced an improved table format of presentation, which is much more informative but also uses significantly more space. In this table format, factor labels are headings of table columns, authors are listed in the factor on which they load most highly, and are ranked within each factor by the strength of their loadings in it; and loadings higher than 0.3 on additional factors, if any, are also presented, indicating the contributions of these authors to more than one specialty.

We followed this format in the present paper as shown in the appendices, except that factor numbers appear as column headings in the tables to represent the factors, and their correspondence to the meaningful factor labels is shown in the caption.

The factor analysis results are also presented as two-dimensional maps based on a novel way of reporting and visualizing factor analysis results in ACA that we briefly introduce here. This visualization allows us to see factor structures in a more condensed visual way to help with interpreting the results.

To prepare these maps, the factors are labeled as usual upon examining the frequently cited articles written by authors in the corresponding factors. On the maps, authors are represented by square nodes and factors by circular nodes, each appropriately labeled. The thickness of a line that connects an author with a factor is proportional to the value of the loading of this author on this factor, as is its grey-scale value, with darker lines corresponding to higher loadings. Dashed lines indicate loadings of opposite signs. Throughout, only sufficiently high loadings of absolute value 0.3 or higher are considered. In this way, a map (e.g., Fig. 1) preserves all the relevant information that is contained in a large table with author loadings on specialty factors as described above (e.g., App. I).

The sizes and colors of nodes in these maps carry auxiliary information. An author node is proportional in size to the author's total loading on all factors combined, and the size of a factor node corresponds to the sum of the loadings on this factor by all authors who load sufficiently (i.e., 0.3 or higher in this case) on it. In this way, the size of a node is an approximate indicator of its overall significance in the map. In addition, node colors in this map represent the number of links to or from a node to other nodes in the map (i.e., its degree). The color of factor circles therefore represents the number of authors that load sufficiently on it, while the color of a square node representing an author indicates the number of factors that this author loads on with a value of at least 0.3 each: yellow for authors who only load sufficiently on a single factor, green for those who co-load on two factors, red for three factors, and blue, for four factors.

In all cases, the layout of our maps is produced using Pajek's implementation of the Kamada-Kawai graph layout algorithm (Batagelj & Mryar, 2007), using loadings as similarity measure between author and factor nodes. The result is a map that is visually informative and true to the factorization it represents, but is not equivalent to the MDS author cluster maps produced in traditional ACA studies.

We will use both formats of presentation in our analyses of the IS field to test how and to what degree this visualization of factor structures helps in the interpretation of results.

Verification of results

We conducted an ACA of the XML research field in the same way as in the IS field described above, and used the results to verify findings regarding rotation methods of FA in ACA that resulted from studying the IS field. We chose this particular field because we have expertise in it and have previously studied it using ACA, which eases interpretation of results.

A set of 100 core authors was selected to represent the XML field based on “citedness”. A matrix of co-citation frequencies of these highly-cited authors was produced by one of our Python programs. This matrix was input into the factor analysis routine in SPSS where factors were extracted by Principal Component Analysis (PCA), and the number of factors extracted was determined based on Kaiser’s rule of eigenvalues greater than 1. This resulted in a 12-factor model which explains 82% of the total variance, and the differences between observed and implied correlations were smaller than 0.05 for the most part (98%). The communalities range from 0.51 to 0.92, only 5 (or 5%) of which are below 0.6 and 10 (or 10%) of which below 0.7. The model fit is apparently very good.

As with the IS data set, we first applied an orthogonal rotation (SPSS Varimax), and then an oblique rotation (SPSS Direct OBLIMIN) to this factor model, and the results are presented both as two-dimensional maps in figures and as tables in the appendices. These results will be discussed in corresponding sections later in the paper.

Intellectual structure of Information Science 1996-2005, orthogonal rotation

With large factors interpreted as specialties, the results of the factor analyses reveal the specialty structure of the IS field and the associated authors’ memberships in one or more specialties as judged by citing authors in the base document set (White & McCain, 1998). We will examine the intellectual structure of the IS field as revealed from an ACA of its main literature as published in the years 1996-2005, and how this structure compares to that found by W&M98 for 1972 to 1995. We will first analyze the results from a factor analysis of the co-citation matrix with an orthogonal rotation in this section because that is the method that was used in W&M98, and then compare them with the results from a factor analysis of the same co-citation matrix with an oblique rotation in the next section.

Results from a factor analysis, with an orthogonal rotation (SPSS Varimax), of the co-citation matrix of 120 highly-cited authors in the IS field, are presented as a two-dimensional map in Fig. 1 and as a table in Appendix I, the latter formatted as in W&M98. The factors are labeled upon examining the corresponding authors’ frequently cited articles, and are listed in Table 2.

Specialties

As in W&M98, which also employed an orthogonal rotation, some factors have primary loadings from a considerable number of authors, and may therefore be interpreted as specialties of the IS field,

whereas other factors only have primary loadings from very few authors, and “pick up either interesting isolates or some residual variation in authors whose main loadings are elsewhere” (p. 336).

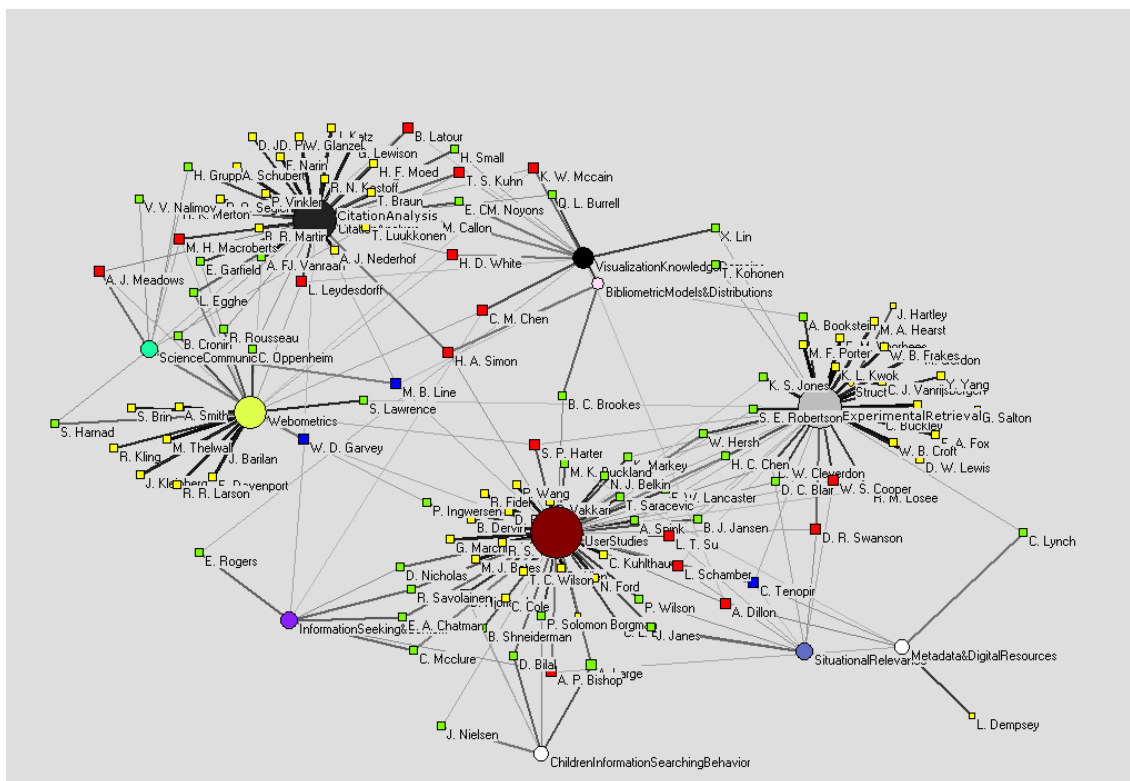


Figure 1: Results of factor analysis with orthogonal rotation (Varimax) – IS

Table 2: Factors and their labels (orthogonal rotation)

| | |
|---------------------------------------|--|
| 1. User studies | 7. Users' judgments of relevance (situational relevance) |
| 2. Citation analysis | 8. Information seeking and context |
| 3. Experimental retrieval | 9. Children's information searching behavior |
| 4. Webometrics | 10. Metadata & digital resources |
| 5. Visualization of knowledge domains | 11. Bibliometric models & distributions |
| 6. Science communication | 12. Structured abstracts (academic writing) |

Among the 5 specialties in our study that have primary loadings from 5 or more authors, *User studies* is the most prominent with 36 authors or 30% of the 120 highly-cited authors, followed by *Citation analysis* (27, or 23%) and *Experimental retrieval* (24, or 20%). The *Webometrics* specialty is

quite visible as well with primary loadings from 13 authors (11%). The smallest is *Visualization of knowledge domains* with primary loadings from 5 authors.

Experimental retrieval and Citation analysis were the largest two specialties identified in W&M98, and they retain their prominence in our results. The relatively small User theory specialty in W&M98 grew dramatically over 10 years and became the largest specialty in the IS field. White & McCain's (1998) observation that cognitive aspects of IS were gaining in importance during their period of study thus appears to be vindicated in quite a dramatic fashion, possibly aided by the intense interest in Web use and usability that this new medium generated during the decade. This specialty encompasses all aspects of user studies, from users' interaction with information systems to users' general information behaviors. Studies in this area have proliferated so much that many of the authors who had secondary loadings in this area in W&M98 are now placed primarily here, e.g., Belkin and Saracevic from the Experimental retrieval specialty, and Borgman, Fidel, and Bates from the Online retrieval specialty. As one would expect, many authors continue to load on both this and the *Experimental retrieval* specialty. While the *User studies* specialty has significant overlap with *Experimental retrieval*, neither overlaps with *Citation analysis* at all. The two authors in W&M98 who loaded with both retrievalists and citationists did not make the list of highly-cited authors in our study.

Webometrics, which focuses on the quantitative study of websites and web traffic, was not one of the specialties identified in W&M98, of course, as it only emerged with the enormous recent impact of the Web on information creation, transmission, and use. This newly-emerging specialty is prominent and distinct in our study as seen from the large number of authors in the group (13) and from the high loadings of its authors (from 0.6 to 0.93). As one might expect, given the hypertext link network nature of the Web, several of these authors (e.g., Cronin, Egghe, Oppenheim, & Rousseau) have secondary loadings on the *Citation analysis* specialty that studies citation networks between documents.

The small specialty *Visualization of knowledge domains* is another newly recognized area of IS although studies on this topic have been around for a long time, dating back to at least the introduction of the co-citation concept in 1973 (Marshakova, 1973; Small, 1973). This new recognition as a separate research field is an indication of recently increased activity in this area, which is probably partly due, as we speculated earlier in this paper, to the fact that written knowledge has largely become available in digitized form on the Web, which provides huge amounts and large varieties of data for this type of studies, and because computing power has increased dramatically to easily

accommodate the types of large-scale network analysis required by this type of studies. Authors in this specialty all have secondary loadings on at least one other specialty, and many authors in other specialties have secondary loadings here as well, some of which higher than 0.4 (e.g., Noyons, Callon, Kuhn, and Small). This further supports the analysis that this new recognition is an indication of boosted activities in a long-existing research area.

Among the small factors in our study, two appear to be a diminished version of previously fairly large specialties identified in W&M98. The *Science communication* specialty in W&M98 has shrunk to only three authors: Meadows, Garvey, and Line, two of whom also have loadings in the *Citation analysis* specialty, as did most of those who appeared in this specialty in W&M98. The pioneers in the *Bibliometrics* specialty in W&M98 did not show up among the 120 most highly cited authors any more (e.g., Bradford, Zipf, Lotka). Brookes is the only author remaining from that specialty, and is joined here by Burrell and Simon, to form the group we labeled as *Bibliometric models and distributions*.

The remaining small factors include Dempsey and Lynch on metadata, Bilal and Large on children's information searching, and Hartley on academic writing / structured abstracts.

Some of the specialties identified in W&M98 have essentially disappeared as separate specialties from our study, including Online retrieval, OPACs, and General library systems. Authors in these specialties are largely excluded from the top 120 most highly-cited authors in the base document set (e.g., all those in OPACs and most in General library systems). Those who remain are mostly regrouped, with low loadings, into either specialties on user studies where they previously had secondary loadings (Bates, Borgman, and Fidel) or the specialty labeled as "*Metadata and digital resources*" (Tenopir and Buckland). This appears to indicate some major shifts of research focus in the IS field, due perhaps to the swift and radical shift to the Web paradigm in all aspects of online access to information during this time period.

The group of authors labeled in W&M98 as "Imported Ideas" largely disappeared as well (except Schneiderman). This may well be a side effect of the burgeoning of the user studies area at the expense of other areas in which these authors used to be cited, perhaps sufficiently so to prevent them from making the list of the 120 most highly cited any more.

Two-camp division

The observation in Harter (1992) and in W&M98 regarding the division of IS into two large sub-disciplines that are not yet well integrated continues to be supported by our data: the “literatures people” (*Bibliometric models and distributions, Citation analysis, Visualization of knowledge domains, Webometrics, Science communication*) do not normally load with the “retrieval people” (*Experimental retrieval, User studies, Metadata*), while co-loadings within these two sub-disciplines occur frequently.

There are a few exceptions: Ingwersen and Harter are recognized primarily as retrievalists, but have secondary loadings on the *Webometrics* specialty; Lawrence, Lin, Kohonen, Garvey, and Line are primarily recognized as literatures people (i.e., *Webometrics, Visualization of knowledge domains, or Science communication*), but have secondary loadings on either the *Experimental retrieval* specialty or the *User studies* specialty. It appears that the two newly-emerging specialties, *Webometrics* and *Visualization of knowledge domains*, have been drawing on both of the two otherwise still largely separate camps of IS, and thus have become two areas that have perhaps begun integrating the IS field. Contributing factors to this phenomenon may include the import of citation analysis principles and techniques from the literatures camp into modern Web search engines such as Google which are at the heart of the retrieval realm of study, and the application of information visualization technology in both information retrieval user interfaces and the mapping of knowledge domains.

Perhaps surprisingly, the small group in the retrieval camp labeled *Information seeking and context* appears to bridge the two camps as well. It has loadings both from people in the retrieval camp who study information behaviors of marginal populations or in everyday life (e.g., Chatman and Savolainen), and from people in the literatures camp who study communication patterns or information behaviors of scientists and innovators (e.g., Garvey and Rogers). It appears that the study of information seeking in an information user's natural context is a common theme of this group. The fact that this group appears to be bridging the divide may be an indication that the influence of cognitive studies is gaining a foothold on the other side of the divide (i.e., in the literatures camp).

A few authors (e.g., Brookes, Simon and Bookstein) load on both the *Bibliometric models and distributions* group and the *Experimental retrieval* specialty, which at first blush seems to indicate that they, too, bridge the two camps of IS. However, this appears to be a case as observed in W&M98 where authors loaded on both camps of IS, but in fact were recognized by each camp separately without genuinely bridging the divide: “... such recognition probably reflects versatility in carrying out

different projects at different times in their careers, rather than successful syntheses already published on their part.” (p. 337-338)

Two-dimensional maps

The five-specialty structure of the IS field is clearly visible in Fig. 1. The five large factors representing specialties are seen as the five largest circles with many and often prominent lines connecting to them. The two camps of the IS field are placed visibly apart, with the literatures people towards the upper left and the retrieval people towards the lower right. Connections (i.e., co-loadings) within each camp are dense whereas those between the two camps are remarkably sparse. Newly-recognized specialties or groups discussed above that bridge the two camps of IS (e.g., *Webometrics* and *Visualization of knowledge domains* in the literatures camp and *Information seeking and context* in the retrieval camp) are placed towards the respective other camp, indicating a certain degree of connection with it. Other small groups that have little connections with the opposite camp (e.g., *Children’s information searching behavior* on the retrieval side) are placed away from the other camp. Similar observations can be made easily from the map regarding individual authors: some authors build bridges between the camps (e.g., Lin, Kohonen, Simon, Brookes, Line, Lawrence, Garvey, Rogers), while others stay within their own camp.

It thus appears that our novel method for visualizing factor analysis results depicts the overall specialty structure of the IS research field very well, at least in the case of an orthogonal rotation. It also represents authors’ memberships in specialties quite well, especially for multiple memberships. The advantages of this method over a table representation include: (a) it is more visual, which makes it more informative for people who are more visually oriented; (b) it shows the interrelationships between specialties or groups quite clearly; and (c) it requires significantly less space without sacrificing much information. The downside is that some author names are hidden in the dense network, especially when a large number of authors are included in the study, which makes it still necessary to consult the table format of representation when studying these authors’ memberships.

Intellectual structure of Information Science 1996-2005, oblique rotation

Fig. 2 / App. II present the FA results from the pattern matrix of the oblique rotation (SPSS Direct OBLIMIN). The factors and their labels are listed in Table 3.

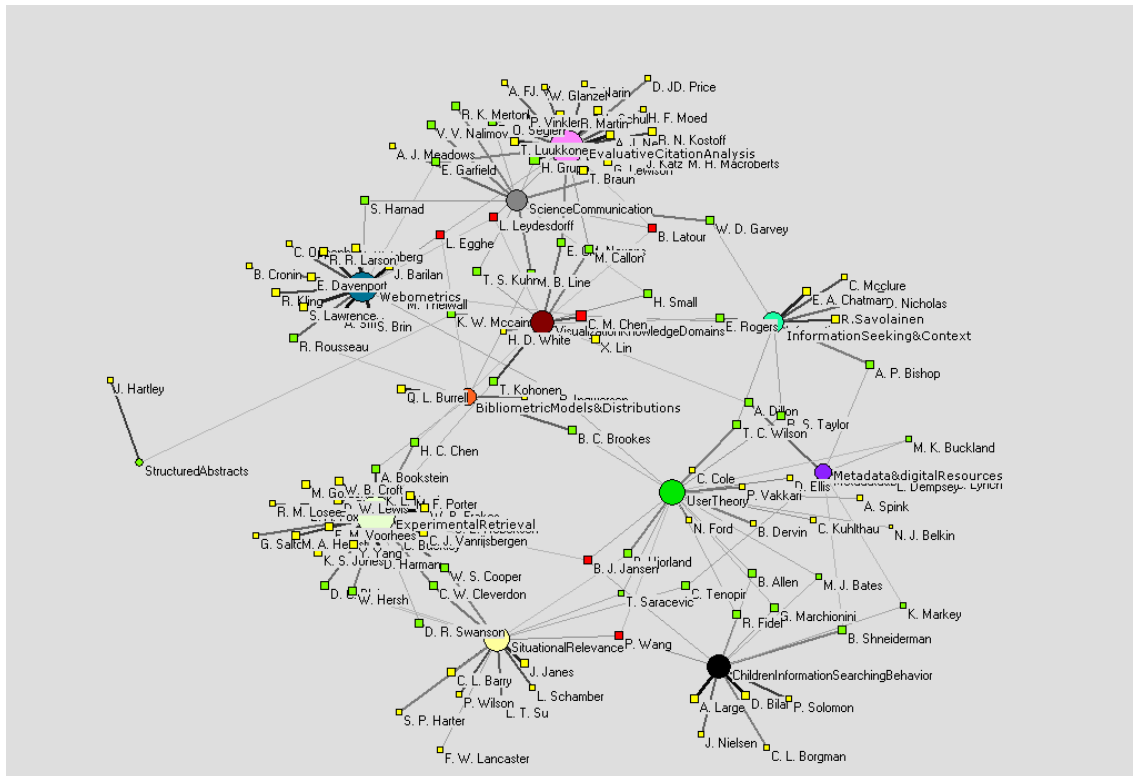


Figure 2: Pattern matrix of factor analysis with oblique rotation (Direct OBLIMIN) – IS

Table 3: Factors and their labels (oblique rotation)

| | |
|---------------------------------------|---|
| 1. User theory | 7. Information seeking and context |
| 2. Evaluative citation analysis | 8. Metadata & digital resources |
| 3. Experimental retrieval | 9. Bibliometric models & distributions |
| 4. Webometrics | 10. Children's information searching behavior |
| 5. Science communication | 11. Users' judgments of relevance (situational relevance) |
| 6. Visualization of knowledge domains | 12. Structured abstracts (academic writing) |

Compared with Fig. 1 / App. I, which present the FA results from the view of an orthogonal rotation, Fig. 2 / App. II suggest a somewhat different picture.

Number of specialties

There are now ten rather than five factors that have primary loadings from five or more authors. The five new additions to those identified by the orthogonal rotation are *Metadata & digital resources*, *Children's information searching behavior*, *Users' judgments of relevance or situational relevance*, *Science communication*, and *Information seeking and context*. We consider some of these labels as

somewhat tentative, especially the last one, as we originally found it quite difficult to label a factor that appears to include people researching such diverse topics: Chatman on information behavior of marginal populations, Savolainen on everyday life information seeking, McClure on information policy and library management, Rogers on diffusion of innovation, and Bishop on social informatics. On closer analysis, it now appears to us that the study of information seeking behaviors in users' natural contexts is a common theme of this group. Readers of this journal who have a deeper understanding of these authors' works may be able to make better sense of this. It is quite common in factor analysis studies that one or more factors remain undefined (Hair, et al., 1998).

Size of specialties

The specialties in Fig. 2 / App. II are more evenly distributed in size, unlike those in Fig. 1 / App. I that had both very large and very small factors.

The *Experimental retrieval* and *Webometrics* specialties remain essentially unchanged. The small groups in Fig. 1 / App. I (3 authors or fewer), including *Science communication*, *Visualization of knowledge domains*, *Metadata*, and *Children's information searching behavior*, have doubled or even tripled their respective sizes in Fig. 2 / App. II by recruiting members from the largest two specialties: *User studies* and *Citation analysis*, and as a result, the latter have become much smaller.

Nature of specialties

As a result of membership change, the focus (and thus, perhaps, the nature) of the recognized specialties appears to have shifted as well. The citation analysis group in Fig. 2 / App. II, for example, has become more centered on evaluative citation analysis after Merton, MacRoberts, Latour, and Nalimov moved to the *Science communication* specialty, Callon, Noyons, Small, and Kuhn to the *Visualization of knowledge domains* specialty, and Egghe to the *Webometrics* specialty. The *User studies* group is now more focused on the theorists such as Wilson, Dervin, Ellis, Kuhlthau, Taylor, Ingwersen, Saracevic, and Belkin. Among the authors who have left this group, some now load high (0.5 or higher) on the *Children's information searching behavior* specialty (e.g., Solomon, Borgman, Shneiderman, Fidel), and some who have secondary loadings in the specialty *Users' judgment of relevance or situational relevance* in Fig. 1 / App. I (e.g., Janes, Barry, Schamber, P. Wilson, and Su) now have separated out to form a distinct specialty.

Overlap between specialties

It appears that the specialties in Fig. 2 / App. II overlap less than those in Fig. 1 / App. I do, partly because the focus of some specialties has shifted. Most authors in the *User studies* specialty in Fig. 1 / App. I who had secondary loadings in factor 3 – *Experimental retrieval* or in factor 7 – *Users' judgment of relevance or situational relevance* (e.g., Spink, Belkin, Schamber, Su), no longer have secondary loadings. Authors loading in Factor 2 – *Citation analysis* or in Factor 3 – *Experimental retrieval* in Fig. 2 / App. II have far fewer secondary loadings than their counterparts in Fig. 1 / App. I. Many authors in the *Visualization of knowledge domains* and / or *Science communication* specialties in Fig. 1 / App. I load on three factors whereas those in Fig. 2 / App. II only load on two factors. All three authors in the *Children's information searching behavior* specialty in Fig. 1 / App. I load high on the *Users studies* specialty, but no longer have secondary loadings in Fig. 2 / App. II.

Individual authors' memberships

Among the six authors who have loadings of 0.5 or higher in the *Children's information searching behavior* specialty in Fig. 2 / App. II, four came from the *User studies* specialty in Fig. 1 / App. I. Three of the four (i.e., Solomon, Borgman, and Fidel) have done research on children, and among their highly cited papers, it is their research on this particular user population that was the most highly cited. For example, Solomon's paper on "Children's information retrieval behavior: a case analysis of an OPAC" was cited 20 times while his next most highly cited paper "Discovering information behavior in sense making I: Time and timing" was cited 11 times and "Discovering information behavior in sense making II: The social," 9 times. The most highly cited papers by Borgman ("Children's Searching Behavior on Browsing and Keyword Online Catalogs: The Science Library Catalogue Project") and by Fidel ("A visit to the information mall: Web searching behavior of high school students") were also on children's (or young adults') searching behavior. It thus appears that the oblique rotation recognized their most frequently used and often most unique work while the orthogonal rotation emphasized their more general research area.

Similarly, all authors except one in the *Situational relevance* specialty in Fig. 2 / App. II, come from the *User studies* specialty in Fig. 1 / App. I. All but one of these authors (namely, Janes, Barry, Schamber, P. Wilson, and Su) have high secondary loadings on the *Situational relevance* specialty in Fig. 1 / App. I. That means that their research focusing on users' judgments of relevance or situational

relevance was recognized by both rotation methods but in different ways: as a separate group representing this specialty in the results from the oblique rotation, and as high secondary loadings in this specialty in the results from orthogonal rotation. It appears that a tightly focused group can be recognized more clearly by oblique rotation than by orthogonal rotation. This reinforces the earlier observation that the oblique rotation recognized scholars' unique contributions while the orthogonal rotation emphasized their more general research area.

This observation finds yet more support from the *Citation analysis* specialty. In results from both oblique and orthogonal rotations, Nalimov, Merton, MacRoberts, Latour have loadings in both the *Citation Analysis* and the *Science communication* specialty, and Noyons, Callon load both on the *Citation analysis* and the *Visualization of knowledge domains* specialty. Again, the orthogonal rotation emphasized more their general research area of citation analysis and placed them there in Fig. 1 / App. I, whereas the oblique rotation identified their special research foci on *Science communication* and *Visualization of knowledge domains*, respectively, by placing them there in Fig. 2 / App. II.

Overall structure

The two camps of the IS field are not as clearly separated visually by the pattern matrix of the oblique rotation (Fig. 2) as they were by the orthogonal rotation (Fig. 1), although they each are still located on one side of the map (i.e., upper-left vs. lower-right). In order to see the overall intellectual structure of the IS field, results from the Component Correlation Matrix (Table 4) and from the structure matrix (Fig. 3 / App. III) produced by the oblique rotation need to be examined instead.

Table 4: Component Correlation Matrix – IS

| Component | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|-----------|--------------|--------------|-------------|-------------|--------------|-------|--------------|--------------|--------------|--------------|--------------|-------|
| 1 | 1.000 | -.179 | .238 | -.006 | .039 | .051 | .354 | .292 | -.103 | -.481 | .462 | .048 |
| 2 | -.179 | 1.000 | -.174 | -.209 | -.351 | .228 | -.129 | -.238 | -.340 | .184 | -.134 | .126 |
| 3 | .238 | -.174 | 1.000 | -.013 | .122 | .164 | -.150 | .232 | .005 | -.145 | .314 | .041 |
| 4 | -.006 | -.209 | -.013 | 1.000 | .303 | -.171 | -.025 | -.122 | .223 | .022 | -.035 | -.117 |
| 5 | .039 | -.351 | .122 | .303 | 1.000 | -.207 | -.111 | -.055 | .241 | -.058 | -.069 | -.262 |
| 6 | .051 | .228 | .164 | -.171 | -.207 | 1.000 | .022 | .046 | -.149 | -.041 | .035 | .130 |
| 7 | .354 | -.129 | -.150 | -.025 | -.111 | .022 | 1.000 | .243 | -.065 | -.316 | .216 | .166 |
| 8 | .292 | -.238 | .232 | -.122 | -.055 | .046 | .243 | 1.000 | -.010 | -.364 | .301 | .204 |
| 9 | -.103 | -.340 | .005 | .223 | .241 | -.149 | -.065 | -.010 | 1.000 | -.034 | -.090 | -.100 |
| 10 | -.481 | .184 | -.145 | .022 | -.058 | -.041 | -.316 | -.364 | -.034 | 1.000 | -.310 | .016 |
| 11 | .462 | -.134 | .314 | -.035 | -.069 | .035 | .216 | .301 | -.090 | -.310 | 1.000 | .222 |
| 12 | .048 | .126 | .041 | -.117 | -.262 | .130 | .166 | .204 | -.100 | .016 | .222 | 1.000 |

The Component Correlation Matrix (Table 4) provides indicators of how closely factors are statistically related to each other. We highlight correlation coefficients higher than 0.3. The numbered components there correspond to the factors with the same numbers in Table 3.

Unsurprisingly, the *User theory* specialty (factor 1) is shown to be closely related to the two specialties that apply these theories and models to study specific populations or users' interactions with information systems: factor 10 – *Children's information searching behavior* (0.48), and factor 11 – *Users' judgments of relevance* (0.46). In fact, most of the authors of these two specialties were placed into a single factor by the orthogonal rotation method (labeled as *User studies*) as seen in Fig. 1 / App. I. It also has a high correlation with factor 7 – *Information seeking and context* (0.35), as one would expect. The close relationship between *Children's information searching behavior* and *Metadata and digital resources* (0.36) may be through their common interest in usability and interface design.

As one would expect, *Evaluative citation analysis*, which largely concentrates on a quantitative study of science, is significantly correlated with *Science communication* (0.35) and with *Bibliometric models and distributions* (0.34). Other close relationships (correlation 0.3 or higher) include those of *Experimental retrieval* with *Users' judgments of relevance* (0.31), *Webometrics* with *Science communication* (0.3), *Children's information searching behavior* with *Information seeking and context* (0.32) and with *Users' judgments of relevance* (0.31), and *Metadata and digital resources* with *Users' judgments of relevance* (0.3).

As seen, all of these close relationships are essentially within each of the two camps of IS, and most if not all of them make good sense. These interrelationships between specialties can also be observed in the structure matrix of the oblique rotation (Fig. 3 / App. III), labeled as in Table 3.

When an orthogonal rotation is used in Principal Components Analysis (the typical kind of factor analysis used in ACA), only one component matrix is produced, and it is subsequently used for interpretation. An oblique rotation, however, produces two distinct matrices: the structure matrix and the pattern matrix. Loadings in the pattern matrix represent the unique contribution of individual authors (variables) to specialties (factors), whereas loadings in the structure matrix, which are "simple correlations between variables and factors," are determined both by an author's unique contribution to each factor and by the correlation among factors (Hair, et al., 1998, p. 113).

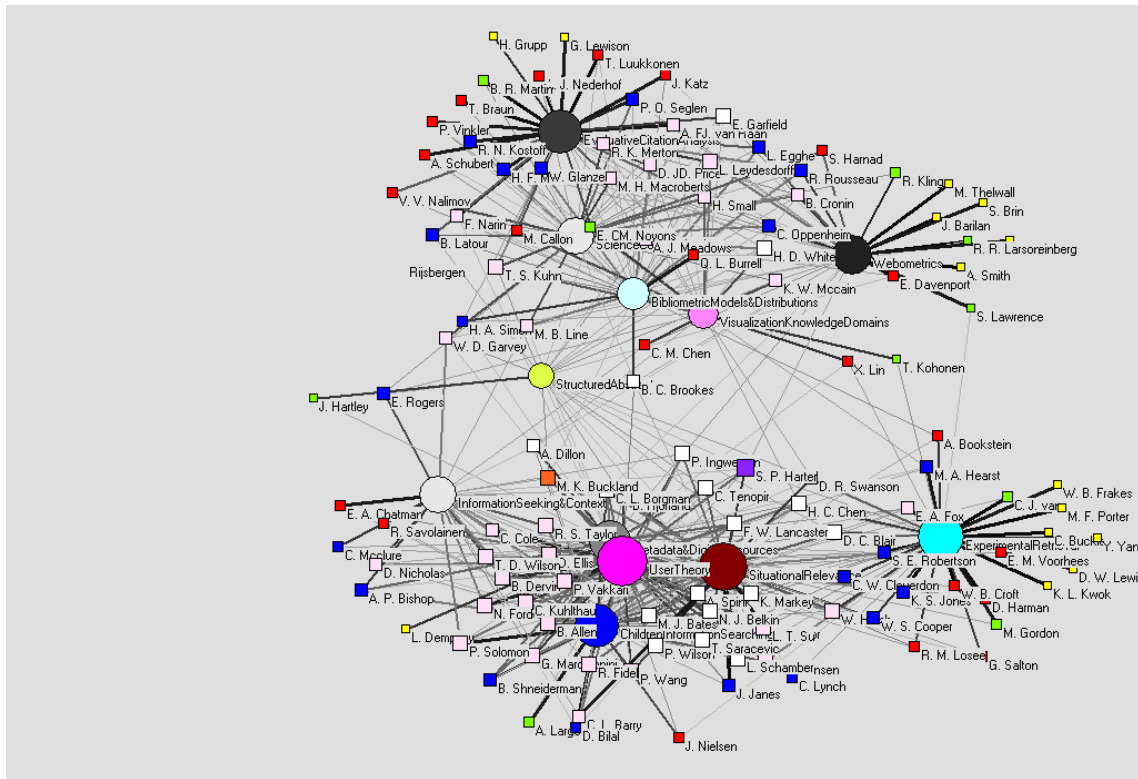


Figure 3: Structure matrix of factor analysis with oblique rotation (Direct OBLIMIN) – IS

Because the correlation among specialties of the IS field is quite high as shown in Table 4, the structure matrix (App. III) has a significantly more complex structure and is much denser than the pattern matrix (App. II): almost all 120 authors (except Barilan, Brin, Buckley, Dempsey, Frakes, Grupp, Kleinberg, Kwok, Lewis, Lewison, Porter, Smith, Thelwall, Yang) have loadings higher than 0.3 on more than one specialty, and many authors load highly on several specialties. For example, most authors in factor 1 – *User theory*, also have fairly high loadings on factor 10 – *Children’s information searching behavior*, factor 11 – *Situational relevance*, factor 8 – *Metadata*, and factor 7 – *Information seeking and context*; most authors in factor 2 – *Citation analysis* also load on at least 2 other factors (mostly factor 5 – *Science communication* and factor 9 – *Bibliometric models and distributions*); and many authors in factor 3 – *Experimental retrieval* have loadings higher than 0.3 on factors 1 and 11.

This co-loading structure is consistent with the factor correlation structure in Table 4: the factors on which the same authors load are significantly correlated, and the higher the correlations, the higher the co-loadings. For example, most loadings of authors in factor 1 on factors 10 and 11 are 0.6 or higher, and factor 1 has a correlation of -0.48 with factor 10 and 0.46 with factor 11. Similarly, most

authors in factor 2 have loadings ranging from 0.3 to 0.5 on factors 5 and 9, and the correlation between factor 2 and factor 5 is -0.35 and that between factor 2 and 9 is -0.34 which are lower than those between factor 1 and factor 10 or factor 11.

Fig. 3, which represents the structure matrix of the oblique rotation, shows this dense network visually. Here, we can clearly see the two camps of the IS field towards the upper and lower ends of the map, respectively, visually separated through a pattern of very dense connections within each camp and sparse connections between them. This visualization also shows the specialties or groups of authors belonging to each camp, as well as those that bridge the two camps as discussed earlier.

It is interesting to see in Fig. 3 that the small factor labeled as *Structured abstracts* is virtually a bridge between the two camps. This factor is represented by studies on academic writing by Hartley, especially on structured abstracts, but has low loadings from many authors in both camps. It appears that this area of study has been found useful for both literature analysis and user and retrieval studies.

Another observation specific to the structure matrix in Fig. 3 is that the *Webometrics* specialty in the literatures camp and the *Experimental retrieval* specialty in the retrieval group are a little off the rest of their respective camps with many high loadings (i.e., thick and dark lines) specific to each of these two specialties. This indicates that they are quite distinct within their respective camps.

It is, however, difficult to see the individual authors' memberships in the specialties on this map due to the extremely dense connections within each camp. This can be explained by the nature of the structure matrix in the case of an oblique rotation as discussed above. The loadings represent a combination of unique contributions of authors to specialties and the correlation between specialties. As the specialties identified in the LIS field are highly correlated as shown in Table 4, it is very difficult to distinguish which authors contribute uniquely to each specialty from a structure matrix (Hair, et al., 1998), or which authors have memberships in which specialties.

To summarize, we find that an oblique rotation tends to be better at identifying authors' unique research contributions in the pattern matrix and, indirectly, the interrelationships between specialties in the visualization of the structure matrix, whereas an orthogonal rotation tends to elicit their general research areas. Consequently, results from an orthogonal rotation show the major specialties of the IS field more clearly, whereas those from an oblique rotation provide a simpler and clearer structure showing all recognizable individual specialties and their interrelationships as well as authors'

memberships in these specialties by the pattern matrix, and the overall structure of the IS field through visualization of the structure matrix.

In other words, different rotation methods have different strengths. Using an orthogonal rotation alone one may overlook a researcher's unique research contributions as well as specialties that appear too small to be significant in interpretation, whereas using an oblique rotation alone one may miss the forest for the trees as major specialties split into their different subspecialties. A complete ACA study using factor analysis would require a combination of the two rotations in order to obtain a view of both the commonalities and the uniqueness of different research areas and authors within a field.

Similarly, the different matrices produced by an oblique rotation provide different kinds of information, with the pattern matrix extracting authors' unique major contributions to their individual core specialties, and the structure matrix documenting the overall relationship between the specialties and authors by combining authors' unique contributions to specialties with the correlations between specialties, from which the overall intellectual structure is nicely visible using our visualization technique. Although in other social sciences, "most researchers report the results of the factor pattern matrix" (Hair, et al., 1998, p. 113), we find that it is the combination of the pattern matrix and the structure matrix that allows a detailed examination of the intellectual structure of a research field in ACA as they each emphasize different aspects of the network.

Verification of findings regarding rotation methods of FA in ACA

Although some of these findings are consistent with the theoretical underpinnings of factor analysis, others are at odds with them. We therefore examine another research field to verify whether the findings from the IS field are specific to that particular field or whether they generalize to others.

For this purpose, an ACA comparable to that in the IS field described above was carried out in the XML research field. Figures 4-6 depict the factor analysis results. The factors are labeled as usual upon examining the frequently cited articles written by authors in the corresponding factors. The table representations are not shown due to space limits.

We observe a very similar pattern here in the XML field as we did in IS.

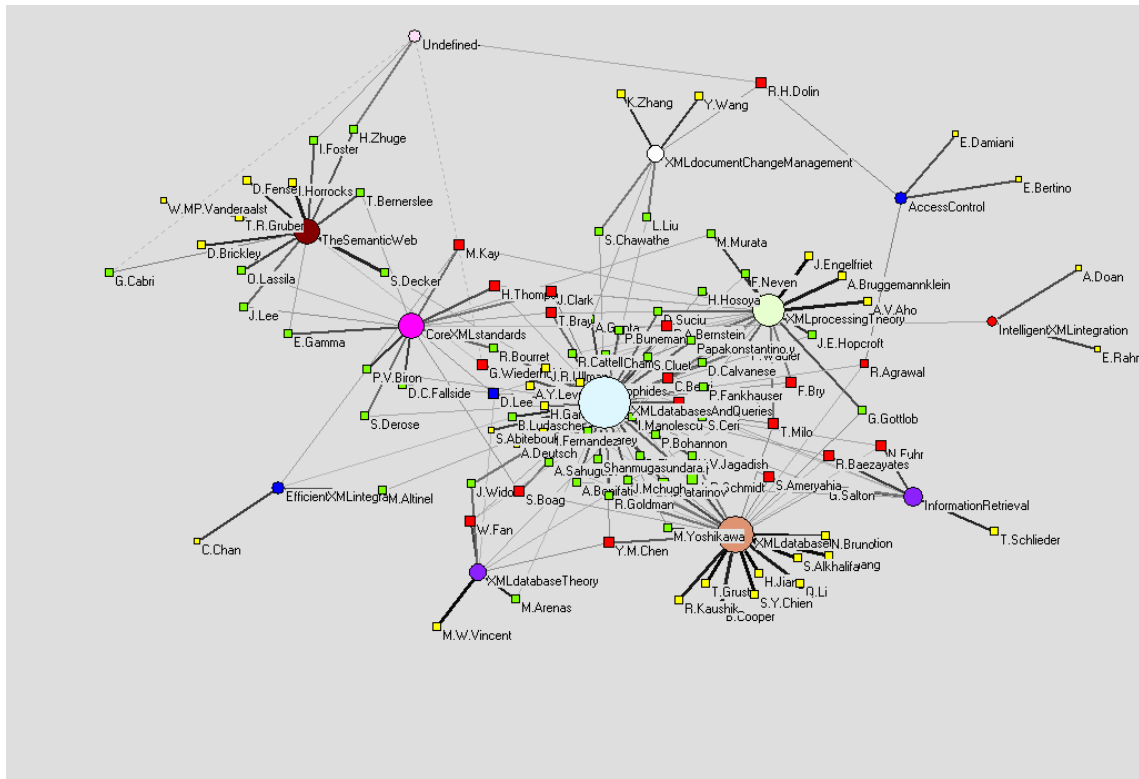


Figure 4: Results of factor analysis with orthogonal rotation (Varimax) – XML

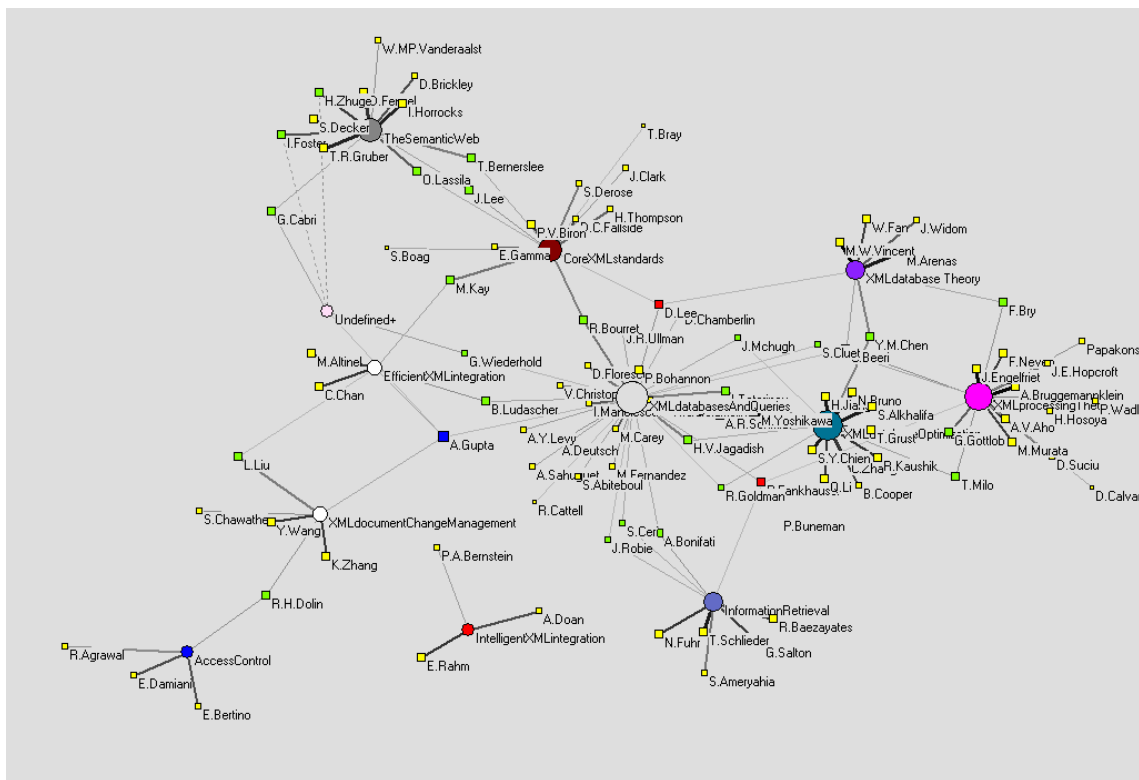


Figure 5: Pattern matrix of factor analysis with oblique rotation (Direct OBLIMIN) – XML

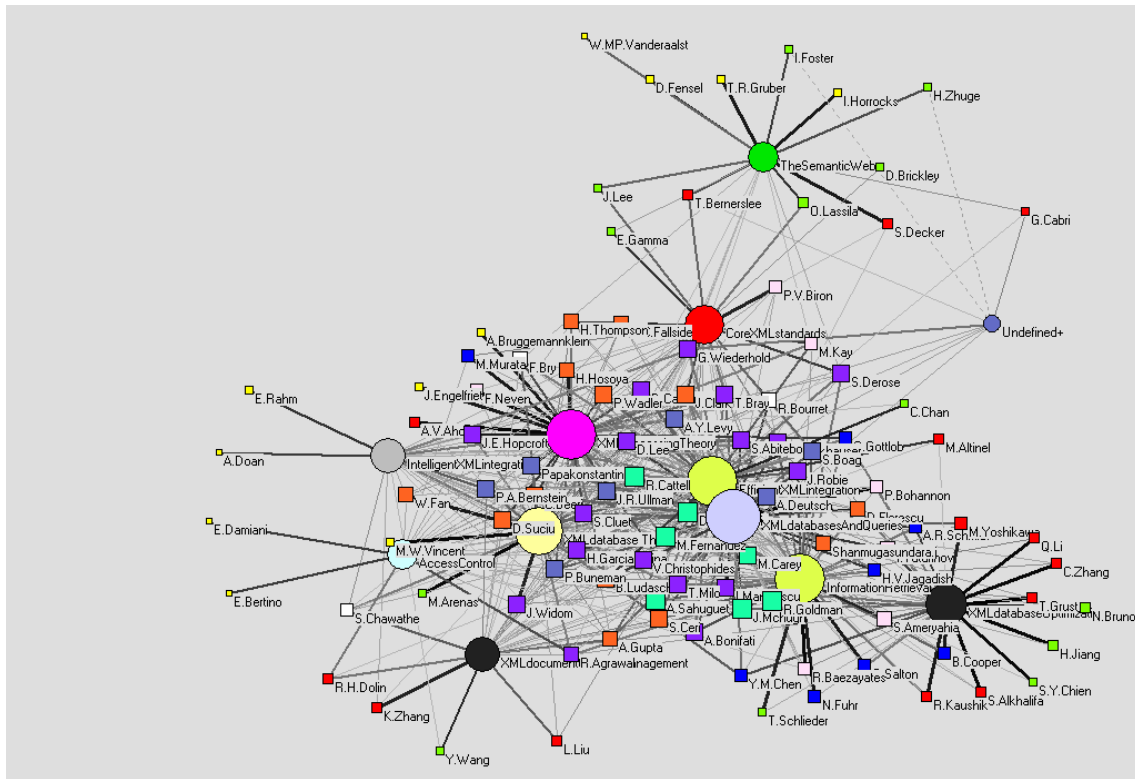


Figure 6: Structure matrix of factor analysis with oblique rotation (Direct OBLIMIN) - XML

The orthogonal rotation results in a dominant factor (labeled as *XML databases and queries* in Fig. 4) with which many factors have connections. This is a general area on XML and databases. This area becomes much smaller in Fig. 5 which represents the pattern matrix of the oblique rotation, after a number of authors moved away to other factors. As a result, the factors become more evenly distributed in size, and have fewer inter-factor connections due to their shifted foci. The structure matrix (Fig. 6) shows the overall intellectual structure of the XML field: a dominant area centered around the study of XML and databases, and a small and distinct group on the Semantic Web. One specialty in the dominant area, *Core XML standards and recommendations*, serves as the connecting point between these two areas; so does a very small research focus (labeled as “Undefined+”) that appears to be about one aspect of mediation.

We examined more closely the authors who were placed into the general XML databases research area by the orthogonal rotation (Fig. 4), but moved away from this specialty in the results from the oblique rotation (Fig. 5). Papakonstantinou, Beer, Calvanese, and Cluet moved to *XML processing theory*; Bonifati, Robie, and Ceri moved to *Information retrieval*; Boag and Bray to *Core*

XML standards; and Wiederhold moved to his own separate factor. All these moves are essentially to those factors where these authors had high secondary loadings before. In other words, their primary contributions to the XML research field picked out by the orthogonal rotation are considered secondary by the oblique rotation, and *vice versa*. This indicates that the oblique rotation picks out these authors' individual unique contributions while the orthogonal rotation identifies their general research area, i.e., XML and databases, as all research areas in the XML field except *The Semantic Web* are related to this central theme in one way or another.

Observations in the XML field are thus consistent with those in the IS field, and therefore confirm our findings there regarding rotation methods of factor analysis in ACA. Note that the high-level intellectual structure of the XML field is quite distinct from that of the IS field. The latter is separated into two camps of comparable size, with few connections bridging the gap, while the former has a single huge main component centered on the database concept, and a very small outlier component (Semantic Web) that has little if anything to do with databases, connected this time via a specific factor, the Core Standards group, because Semantic Web authors heavily contributed to these standards. Our findings therefore appear to be quite robust with respect to different intellectual structure types of research fields revealed by different rotation methods of FA in ACA. Differences in time-spans covered or in field delineation methodology do not appear to affect these results, either.

Conclusions

Information Science during the first decade of the Web

We conducted a full-scale ACA study of the Information Science (IS) field covering the years 1996-2005, i.e., the decade subsequent to the time period examined in W&M98, the last highly-cited comprehensive ACA study of IS. This time period is particularly significant in that it is the first decade of the rise to prominence of the World Wide Web, and allows us to glimpse its effects on the IS field.

We find that the Web has had a truly profound effect on the intellectual structure of IS. It has spawned new research areas such as *Webometrics*, helped revitalize such classic IS areas as the *Visualization of knowledge domains*, and obsoleted other previously strong areas such as *OPACs* and *Online retrieval*. More importantly perhaps for IS as a field, the Web appears to be catalyzing the bridging of an ancient chasm between the "literatures" and "retrieval" camps within the IS field (Harter,

1992; W&M98), as classic “literatures” principles and techniques (e.g., citation analysis) famously begin to drive Web “retrieval” engines, and visualization of information finds interest in both camps.

We also find that a major trend that was beginning to have noticeable effects in earlier periods (White & McCain, 1998), namely the increasing prominence of the cognitive approach in IS, has accelerated to the degree that the corresponding research areas have surpassed others in prominence in the time period we studied. Again, this development may have been boosted unexpectedly by the importance of Web use and Web usability studies in the user studies area, especially in *Children's information searching behaviors* that largely study their use of the Web.

ACA methodology

We took the opportunity afforded us by this study to employ a more thorough factor analysis methodology than previous ACA studies have generally used, in the hope of contributing to a better understanding of the pros and cons of the different choices available to researchers performing ACA studies. In particular, rather than using either an orthogonal or oblique rotation in a factor analysis, we applied both to the same sets of data, and reported and compared the results thus obtained. To this end, we introduced a novel visualization technique for factor analysis results in ACA that we find can provide more intuitive insight into the intellectual structure of a field, and can simplify considerably the comparison of results between different approaches in ACA.

We find that the two rotation methods result in significantly different factorizations of the IS field into specialties. An orthogonal rotation appears to elicit a less detailed picture that emphasizes a small number of general research areas of IS (e.g., user studies, information retrieval, citation analysis), while an oblique rotation provides a detailed map of specializations by the pattern matrix (e.g., children's information searching behaviors, information seeking in context), a map of the overall structure of the IS field (e.g., two-camp structure) by visualization of the structure matrix, and a matrix of correlation strengths between recognized specialties. Oblique rotation thus provides quite detailed information on both the unique contributions of an author to his or her specific area(s) of specialization and the interrelationship between recognized research areas, while orthogonal rotation elicits a picture of authors' memberships in their general areas of expertise.

Thus, both orthogonal and oblique rotations in factor analysis have their own respective strengths when applied to ACA, and both the pattern and the structure matrix provide significant

information when an oblique rotation is used. Especially in conjunction with a visualization of all three factor matrices as two-dimensional maps of the kind that we introduced here, each in its own way offers significant contributions to a more thorough understanding of the field being studied.

We verified these results by performing an equivalent study in a second research area, namely XML. Despite significant differences in size, maturity, and structure between the two research fields, we find a very similar pattern in this field as we did in IS when it comes to comparing the relative merits of the two rotation methods of factor analysis in ACA and those of the two factor matrices produced by an oblique rotation.

We therefore conclude that there is considerable value in performing a more thorough factor analysis in ACA, of the kind we did in this paper, in order to obtain a more complete picture of the intellectual structure of a research field. This more thorough analysis takes into account both orthogonal and oblique rotations of the factor analysis results, and in the case of oblique rotations, reports both the resulting pattern and structure matrices instead of just one as in previous studies. Depending on the research question being asked, a particular one of these might suffice to answer it when ACA is applied for other purposes than the mapping and visualization of knowledge domains.

References

- Ahlgren, P., Jarneving, B., & Rousseau, R. (2003). Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology*, 54, 550-560.
- Ahlgren, P., Jarneving, B., & Rousseau, R. (2004a). Author cocitation and Pearson's r. *Journal of the American Society for Information Science and Technology*, 55(9), 843.
- Ahlgren, P., Jarneving, B., & Rousseau, R. (2004b). Rejoinder: In defense of formal methods. *Journal of the American Society for Information Science and Technology*, 55(10), 936.
- Astrom, F. (2007). Changes in the LIS research front: Time-sliced cocitation analysis of LIS journal articles, 1990-2004. *Journal of the American Society for Information Science and Technology*, 58 (7), 947-957
- Batagelj, V., & Mrvar, A. (2007). *Pajek: Program for analysis and visualization of large networks*. Retrieved on August 3, 2007 from <http://vlado.fmf.uni-lj.si/pub/networks/pajek/doc/pajekman.pdf>.

- Bensman, S.J. (2004). Pearson's r and author cocitation analysis: A commentary on the controversy. *Journal of the American Society for Information Science and Technology*, 55(10), 935-936.
- Börner, K., Maru, J.T., & Goldstone, R.L. (2004). The simultaneous evolution of author and paper networks. *Proceedings of the National Academy of Sciences of the United States of America* 101 (SUPPL. 1), 5266-5273
- Boyack, K., Börner, K., & Klavans, R. (2007). Mapping the Structure and Evolution of Chemistry Research. *Proceedings of the 11th International Conference of the International Society for Scientometrics and Informetrics*, June 25-27, 2007, Madrid, Spain
- Chen, C.M. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359-377
- De May, M. (1982). *The Cognitive Paradigm*. Boston, MA: D. Reidel. [now University of Chicago Press]
- Eom, S.B., & Farris, R.S. (1996). The contributions of organizational science to the development of decision support systems research subspecialties. *Journal of the American Society for Information Science*, 47, 941-952
- Hair, J.F. Anderson, R.E., Tatham, R.L., & Black, W.C. (1998). *Multivariate data analysis* (5th edition). Upper Saddle River, NJ: Prentice Hall.
- Harter, S.P. (1992). Psychological relevance and information science. *Journal of the American Society for Information Science*, 43, 602-615.
- Henzinger, M., & Lawrence, S. (2004). Extracting knowledge from the World Wide Web. *Proceedings of the National Academy of Sciences of the United States of America* 101 (SUPPL. 1), 5186-5191
- Jansens, F., Leta, J., Glanzel, W., & De Moor, B. (2006). Towards mapping library and information science. *Information Processing and Management*, 42, 1614-1642
- Leydesdorff, L., & Vaughan, L. (2006). Co-occurrence matrices and their applications in information science: Extending ACA to the Web environment. *Journal of the American Society for Information Science and Technology*, 57(12), 1616-1628.
- Marshakova, I. V. (1973). A system of document connections based on references. *Scientific and Technical Information Serial of VINITI*, 6, 3-8.

- McCain, K. W. (1990a). Mapping authors in intellectual space: a technical overview. *Journal of the American Society for Information Science*, 41, 433-443.
- McCain, K. W. (1990b). Mapping authors in intellectual space: population genetics in the 1980s. In C. L. Borgman (ed.), *Scholarly Communication and Bibliometrics* (pp.194-216). Newbury Park, CA: Sage.
- Schneider, J.W., Larsen, B., & Ingwersen, P. (2007). Comparative study between first and all-author cocitation analysis based on citation indexes generated from XML data. *Proceedings of the 11th International Conference on Scientometrics and Informetrics*, June 25-27, Madrid, Spain.
- Shiffrin, R.M., & Börner, K. (2004). Mapping knowledge domains. *Proceedings of the National Academy of Sciences of the United States of America* 101 (Suppl 1): 5183-5185
- Small, H. (1973). Cocitation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24, 265-269.
- Small, H. (1999). Visualizing science by citation mapping. *Journal of the American Society for Information Science*, 50, 799-813.
- Thelwall, M., Vaughan, L., & Börneborn, L. (2005). Webometrics. *Annual Review of Information Science and Technology*, 39, 81-135
- Van Eck, N.J., Waltman, L., Van Berg, J.D. (2006). A novel algorithm for visualizing concept associations. *Proceedings - International Workshop on Database and Expert Systems Applications*, DEXA 2006, art. no. 1508306, pp. 405-409
- White, H.D. (2003). Author cocitation analysis and Pearson's r. *Journal of the American Society for Information Science and Technology*, 54, 1250-1259.
- White, H.D. (2004). Replies and a correction. *Journal of the American Society for Information Science and Technology*, 55, 843-844.
- White, H. D. & Griffith, B.C. (1981). Author cocitation: A literature measure of intellectual structure. *Journal of the American Society for Information Science*, 32, 163-171.
- White, H. D. & Griffith, B.C. (1982). Authors as markers of intellectual space: Co-citation in studies of science, technology and society. *Journal of Documentation*, 38, 255-272.
- White, H. D. & McCain, K.W. (1998). Visualizing a discipline: An author co-citation analysis of information science, 1972-1995. *Journal of the American Society for Information Science*, 49, 327-355.

- Wilkinson, D.M., & Huberman, B.A. (2004). A method for finding communities of related genes. *Proceedings of the National Academy of Sciences of the United States of America* 101 (SUPPL. 1), 5241-5248
- Zhao, D. (2006). Towards all-author co-citation analysis. *Information Processing & Management*, 42, 1578-1591.
- Zhao, D., & Strotmann, A. (2007). All-author vs. first-author co-citation analysis of the Information Science field using Scopus. *Proceedings of the American Society for Information Science and Technology 2007 Annual Meeting*, October 19-24, 2007, Milwaukee, Wisconsin, USA

Appendix I. Factor analysis of 120 authors in information science (1996 – 2005)

(orthogonal rotation)

1. User studies (information seeking / searching behaviour, user-centered approach to IR, users and use);
2. Citation analysis (Scientometrics; evaluative bibliometrics);
3. Experimental retrieval (algorithms, models, systems, evaluation of IR);
4. Webometrics;
5. Visualization of knowledge domains (ACA; co-citation analysis);
6. Science communication;
7. Users' judgments of relevance (situational relevance);
8. Information seeking and context;
9. Children's information searching behaviour (usability; interface design);
10. Metadata & digital resources;
11. Bibliometric models & distributions;
12. Structured abstracts (academic writing)

| Author | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|-------------------|------|------|------|------|------|------|------|------|------|------|----|----|
| P. Vakkari | 0.95 | | | | | | | | | | | |
| B. Allen | 0.92 | | | | | | | | | | | |
| D. Ellis | 0.91 | | | | | | | | | | | |
| C. Kuhlthau | 0.9 | | | | | | | | | | | |
| R. S. Taylor | 0.9 | | | | | | | | | | | |
| P. Wang | 0.9 | | | | | | | | | | | |
| B. Dervin | 0.89 | | | | | | | | | | | |
| T. D. Wilson | 0.89 | | | | | | | | | | | |
| R. Fidel | 0.89 | | | | | | | | | | | |
| N. Ford | 0.88 | | | | | | | | | | | |
| M. J. Bates | 0.88 | | | | | | | | | | | |
| G. Marchionini | 0.85 | | | | | | | | | | | |
| A. Spink | 0.84 | | 0.35 | | | | | | | | | |
| N. J. Belkin | 0.83 | | 0.36 | | | | | | | | | |
| T. Saracevic | 0.82 | | 0.35 | | | | | | | | | |
| C. Cole | 0.8 | | | | | | | | | | | |
| P. Ingwersen | 0.79 | | | 0.4 | | | | | | | | |
| P. Solomon | 0.78 | | | | | | | | 0.47 | | | |
| L. T. Su | 0.78 | | 0.37 | | | | 0.38 | | | | | |
| B. Hjørland | 0.76 | | | | | | | | -0.3 | | | |
| C. L. Borgman | 0.75 | | | | | | | | | | | |
| P. Wilson | 0.75 | | | | | | 0.46 | | | | | |
| L. Schamber | 0.73 | | 0.32 | | | | 0.45 | | | | | |
| K. Markey | 0.72 | | 0.39 | | | | | | | | | |
| C. L. Barry | 0.71 | | | | | | 0.5 | | | | | |
| F. W. Lancaster | 0.71 | | 0.4 | | | | | | | | | |
| S. P. Harter | 0.7 | | 0.35 | 0.34 | | | | | | | | |
| B. J. Jansen | 0.68 | | 0.42 | | | | | | | | | |
| M. K. Buckland | 0.68 | | | | | | | | | 0.33 | | |
| B. Shneiderman | 0.68 | | | | | | | | 0.33 | | | |
| J. Janes | 0.67 | | | | | | 0.58 | | | | | |
| R. Savolainen | 0.65 | | | | | | | 0.59 | | | | |
| D. Nicholas | 0.65 | | | | | | | 0.47 | | | | |
| A. P. Bishop | 0.54 | | | | | | | 0.47 | | 0.36 | | |
| C. Tenopir | 0.51 | | 0.33 | | | | 0.32 | | | 0.43 | | |
| A. Dillon | 0.48 | | | | 0.33 | | | | | 0.48 | | |
| A. J. Nederhof | | 0.94 | | | | | | | | | | |
| B. R. Martin | | 0.93 | | | | | | | | | | |
| T. Luukkonen | | 0.92 | | | | | | | | | | |
| R. N. Kostoff | | 0.91 | | | | | | | | | | |
| P. Vinkler | | 0.89 | | | | | | | | | | |
| T. Braun | | 0.89 | | | | | | | | | | |
| H. F. Moed | | 0.87 | | | | | | | | | | |
| A. F. J. van Raan | | 0.87 | | 0.31 | | | | | | | | |
| A. Schubert | | 0.86 | | | | | | | | | | |
| F. Narin | | 0.86 | | | | | | | | | | |
| G. Lewison | | 0.86 | | | | | | | | | | |
| W. Glanzel | | 0.85 | | | | | | | | | | |
| P. O. Seglen | | 0.85 | | | | | | | | | | |
| J. Katz | | 0.84 | | | | | | | | | | |
| D. J. D. Price | | 0.82 | | | | | | | | | | |
| R. K. Merton | | 0.77 | | | | 0.42 | | | | | | |
| L. Leydesdorff | | 0.76 | | 0.4 | 0.31 | | | | | | | |
| M. H. Macrobarts | | 0.76 | | 0.31 | | 0.39 | | | | | | |
| H. Grupp | | 0.75 | | | | -0.4 | | | | | | |

Appendix I. (continued)

| Authors | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|----------------------|------|------|------|------|------|------|------|------|------|------|------|------|
| E. Garfield | | 0.73 | | 0.38 | | | | | | | | |
| M. Callon | | 0.72 | | | 0.51 | | | | | | | |
| E. CM. Noyons | | 0.7 | | | 0.55 | | | | | | | |
| B. Latour | | 0.67 | | | 0.33 | 0.34 | | | | | | |
| V. V. Nalimov | | 0.67 | | | | 0.45 | | | | | | |
| H. Small | | 0.66 | | | 0.44 | | | | | | | |
| T. S. Kuhn | | 0.62 | | | 0.47 | 0.4 | | | | | | |
| L. Egghe | | 0.62 | | 0.53 | | | | | | | | |
| K. L. Kwok | | | 0.95 | | | | | | | | | |
| C. Buckley | | | 0.95 | | | | | | | | | |
| C. J. van Rijsbergen | | | 0.93 | | | | | | | | | |
| W. B. Croft | | | 0.92 | | | | | | | | | |
| E. M. Voorhees | | | 0.92 | | | | | | | | | |
| M. F. Porter | | | 0.91 | | | | | | | | | |
| D. Harman | | | 0.91 | | | | | | | | | |
| K. S. Jones | 0.31 | | 0.89 | | | | | | | | | |
| W. B. Frakes | | | 0.89 | | | | | | | | | |
| D. W. Lewis | | | 0.88 | | | | | | | | | |
| E. A. Fox | | | 0.87 | | | | | | | | | |
| S. E. Robertson | 0.37 | | 0.86 | | | | | | | | | |
| M. Gordon | | | 0.85 | | | | | | | | | |
| M. A. Hearst | | | 0.84 | | | | | | | | | |
| A. Bookstein | | | 0.83 | | | | | | | | 0.33 | |
| Y. Yang | | | 0.83 | | | | | | | | | |
| W. S. Cooper | 0.34 | | 0.79 | | | | 0.36 | | | | | |
| R. M. Losee | 0.31 | | 0.77 | | | | | | | | | |
| D. C. Blair | 0.45 | | 0.75 | | | | | | | | | |
| C. W. Cleverdon | 0.43 | | 0.74 | | | | 0.38 | | | | | |
| G. Salton | | | 0.71 | | | | | | | | | |
| W. Hersh | 0.52 | | 0.71 | | | | | | | | | |
| H. C. Chen | 0.5 | | 0.67 | | | | | | | | | |
| D. R. Swanson | 0.41 | | 0.58 | | | | 0.47 | | | | | |
| J. Barilan | | | | 0.93 | | | | | | | | |
| A. Smith | | | | 0.93 | | | | | | | | |
| M. Thelwall | | | | 0.9 | | | | | | | | |
| E. Davenport | | | | 0.9 | | | | | | | | |
| S. Brin | | | | 0.88 | | | | | | | | |
| R. R. Larson | | | | 0.88 | | | | | | | | |
| J. Kleinberg | | | | 0.85 | | | | | | | | |
| S. Lawrence | | | 0.3 | 0.83 | | | | | | | | |
| R. Kling | | | | 0.81 | | | | | | | | |
| C. Oppenheim | | 0.44 | | 0.78 | | | | | | | | |
| B. Cronin | | 0.45 | | 0.7 | | | | | | | | |
| R. Rousseau | | 0.48 | | 0.69 | | | | | | | | |
| S. Harnad | | | | 0.6 | | 0.44 | | | | | | |
| X. Lin | | | 0.39 | | 0.73 | | | | | | | |
| T. Kohonen | | | 0.47 | | 0.73 | | | | | | | |
| C. M. Chen | | | | 0.38 | 0.69 | | | 0.32 | | | | |
| K. W. McCain | | 0.4 | | 0.31 | 0.58 | | | | | | | |
| H. D. White | | 0.45 | | 0.36 | 0.53 | | | | | | | |
| A. J. Meadows | | 0.49 | | 0.39 | | 0.59 | | | | | | |
| W. D. Garvey | 0.34 | 0.31 | | | | 0.53 | | 0.37 | | | | |
| M. B. Line | 0.38 | 0.37 | | | | 0.51 | | | | | 0.32 | |
| E. A. Chatman | 0.58 | | | | | | | 0.64 | | | | |
| C. McClure | 0.51 | | | | | | | 0.58 | | | | |
| E. Rogers | | | | | 0.37 | | | 0.59 | | | | |
| D. Bilal | 0.61 | | | | | | | | 0.64 | | | |
| A. Large | 0.59 | | | | | | | | 0.68 | | | |
| J. Nielsen | 0.37 | | | | | | | | 0.56 | | | |
| L. Dempsey | | | | | | | | | | 0.74 | | |
| C. Lynch | | | 0.4 | | | | | | | 0.54 | | |
| Q. L. Burrell | | 0.45 | | | | | | | | | 0.76 | |
| B. C. Brookes | 0.49 | | | | | | | | | | 0.62 | |
| H. A. Simon | 0.43 | 0.52 | | | | | | | | | 0.55 | |
| J. Hartley | | | | | | | | | | | | 0.68 |

Appendix II. Factor analysis of 120 authors in information science (1996 – 2005)
(oblique rotation – pattern matrix)

1. User theory; 2. Evaluative citation analysis; 3. Experimental retrieval;
4. Webometrics; 5. Science communication; 6. Visualization of knowledge domains;
7. Information seeking and context; 8. Metadata & digital resources;
9. Bibliometric models & distributions; 10. Children's information searching behavior;
11. Users' judgments of relevance (situational relevance); 12. Structured abstracts (academic writing)

| Author | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|--------------------|------|------|------|-------|-------|------|------|---|-------|-------|------|----|
| C. Cole | 0.73 | | | | | | | | | | | |
| N. Ford | 0.65 | | | | | | | | | | | |
| P. Vakkari | 0.57 | | | | | | | | | | | |
| T. D. Wilson | 0.52 | | | | | | 0.39 | | | | | |
| B. Dervin | 0.52 | | | | | | | | | | | |
| B. Hjørland | 0.51 | | | | | | | | | | 0.33 | |
| D. Ellis | 0.5 | | | | | | | | | | | |
| C. Kuhlthau | 0.46 | | | | | | | | | | | |
| R. S. Taylor | 0.46 | | | | | | 0.4 | | | | | |
| B. Allen | 0.44 | | | | | | | | | -0.37 | | |
| B. J. Jansen | 0.44 | | 0.3 | | | | | | | -0.39 | | |
| P. Ingwersen | 0.41 | | | -0.39 | | | | | | | | |
| A. Spink | 0.41 | | | | | | | | | | | |
| T. Saracevic | 0.41 | | | | | | | | | | 0.35 | |
| M. J. Bates | 0.38 | | | | | | | | | -0.34 | | |
| N. J. Belkin | 0.37 | | | | | | | | | | | |
| G. Lewison | | 0.96 | | | | | | | | | | |
| T. Braun | | 0.92 | | | | | | | | | | |
| H. Grupp | | 0.92 | | | 0.371 | | | | | | | |
| T. Luukkonen | | 0.9 | | | | | | | | | | |
| A. J. Nederhof | | 0.89 | | | | | | | | | | |
| B. R. Martin | | 0.87 | | | | | | | | | | |
| A. Schubert | | 0.84 | | | | | | | | | | |
| P. Vinkler | | 0.83 | | | | | | | | | | |
| J. Katz | | 0.81 | | | | | | | | | | |
| R. N. Kostoff | | 0.77 | | | | | | | | | | |
| W. Glanzel | | 0.74 | | | | | | | | | | |
| P. O. Seglen | | 0.73 | | | -0.35 | | | | | | | |
| H. F. Moed | | 0.72 | | | | | | | | | | |
| F. Narin | | 0.71 | | | | | | | | | | |
| A. FJ. van Raan | | 0.66 | | | | | | | | | | |
| D. JD. Price | | 0.55 | | | | | | | | | | |
| E. Garfield | | 0.5 | | -0.33 | | | | | | | | |
| L. Leydesdorff | | 0.47 | | -0.33 | | 0.34 | | | | | | |
| K. L. Kwok | | | 0.99 | | | | | | | | | |
| C. Buckley | | | 0.99 | | | | | | | | | |
| M. F. Porter | | | 0.98 | | | | | | | | | |
| W. B. Frakes | | | 0.95 | | | | | | | | | |
| C.J.van Rijsbergen | | | 0.91 | | | | | | | | | |
| Y. Yang | | | 0.9 | | | | | | | | | |
| D. W. Lewis | | | 0.9 | | | | | | | | | |
| E. M. Voorhees | | | 0.88 | | | | | | | | | |
| W. B. Croft | | | 0.87 | | | | | | | | | |
| D. Harman | | | 0.86 | | | | | | | | | |
| A. Bookstein | | | 0.83 | | | | | | -0.39 | | | |
| E. A. Fox | | | 0.82 | | | | | | | | | |
| M. Gordon | | | 0.8 | | | | | | | | | |
| K. S. Jones | | | 0.8 | | | | | | | | | |
| S. E. Robertson | | | 0.79 | | | | | | | | | |
| M. A. Hearst | | | 0.78 | | | | | | | | | |
| R. M. Losee | | | 0.69 | | | | | | | | | |
| W. S. Cooper | | | 0.66 | | | | | | | | 0.5 | |
| G. Salton | | | 0.59 | | | | | | | | | |
| D. C. Blair | | | 0.59 | | | | | | | | 0.32 | |
| W. Hersh | | | 0.58 | | | | | | | | 0.31 | |
| C. W. Cleverdon | | | 0.58 | | | | | | | | 0.54 | |
| H. C. Chen | | | 0.53 | | | 0.3 | | | | | | |

Appendix II. (continued)

| Authors | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|------------------|------|------|------|-------|-------|------|------|------|-------|-------|------|------|
| J. Barilan | | | | -0.97 | | | | | | | | |
| A. Smith | | | | -0.95 | | | | | | | | |
| M. Thelwall | | | | -0.92 | | | | | | | | |
| S. Brin | | | | -0.92 | | | | | | | | |
| E. Davenport | | | | -0.88 | | | | | | | | |
| J. Kleinberg | | | | -0.88 | | | | | | | | |
| R. R. Larson | | | | -0.86 | | | | | | | | |
| S. Lawrence | | | | -0.84 | | | | | | | | |
| R. Kling | | | | -0.77 | | | | | | | | |
| C. Oppenheim | | | | -0.72 | | | | | | | | |
| R. Rousseau | | | | -0.64 | | | | | -0.32 | | | |
| B. Cronin | | | | -0.62 | | | | | | | | |
| S. Harnad | | | | -0.48 | -0.46 | | | | | | | |
| L. Egghe | | 0.35 | | -0.46 | | | | | -0.31 | | | |
| A. J. Meadows | | | | | -0.63 | | | | | | | |
| V. V. Nalimov | | 0.35 | | | -0.59 | | | | | | | |
| M. B. Line | | | | | -0.58 | | | | -0.37 | | | |
| W. D. Garvey | | | | | -0.54 | | 0.38 | | | | | |
| R. K. Merton | | 0.44 | | | -0.53 | | | | | | | |
| M. H. Macrobarts | | 0.47 | | | -0.5 | | | | | | | |
| B. Latour | | 0.32 | | | -0.39 | 0.33 | | | | | | |
| T. Kohonen | | | 0.34 | | | 0.78 | | | | | | |
| X. Lin | | | | | | 0.77 | | | | | | |
| C. M. Chen | | | | -0.33 | | 0.71 | 0.4 | | | | | |
| E. CM. Noyons | | 0.53 | | | | 0.63 | | | | | | |
| K. W. McCain | | | | | | 0.59 | | | | | | 0.31 |
| M. Callon | | 0.46 | | | | 0.57 | | | | | | |
| H. D. White | | | | | | 0.52 | | | | | | |
| T. S. Kuhn | | | | | -0.42 | 0.46 | | | | | | |
| H. Small | | | | | -0.31 | 0.46 | | | | | | |
| E. A. Chatman | | | | | | | 0.85 | | | | | |
| R. Savolainen | | | | | | | 0.81 | | | | | |
| C. McClure | | | | | | | 0.7 | | | | | |
| E. Rogers | | | | | | 0.36 | 0.64 | | | | | |
| D. Nicholas | | | | | | | 0.63 | | | | | |
| A. P. Bishop | | | | | | | 0.54 | 0.37 | | | | |
| L. Dempsey | | | | | | | | 0.86 | | | | |
| C. Lynch | | | | | | | | 0.61 | | | | |
| A. Dillon | | | | | | 0.31 | | 0.6 | | | | |
| C. Tenopir | | | | | | | | 0.44 | | | 0.38 | |
| M. K. Buckland | 0.31 | | | | | | | 0.38 | | | | |
| Q. L. Burrell | | | | | | | | | -0.85 | | | |
| B. C. Brookes | 0.35 | | | | | | | | -0.68 | | | |
| H. A. Simon | | | | | | | | | -0.65 | | | |
| A. Large | | | | | | | | | | -0.97 | | |
| D. Bilal | | | | | | | | | | -0.94 | | |
| P. Solomon | | | | | | | | | | -0.75 | | |
| J. Nielsen | | | | | | | | | | -0.67 | | |
| C. L. Borgman | | | | | | | | | | -0.52 | | |
| B. Shneiderman | | | | | | | | 0.33 | | -0.51 | | |
| R. Fidel | 0.36 | | | | | | | | | -0.5 | | |
| G. Marchionini | 0.36 | | | | | | | | | -0.48 | | |
| P. Wang | 0.33 | | | | | | | | | -0.45 | 0.45 | |
| K. Markey | | | | | | | | 0.33 | | -0.37 | | |
| J. Janes | | | | | | | | | | | 0.88 | |
| C. L. Barry | | | | | | | | | | | 0.78 | |
| L. Schamber | | | | | | | | | | | 0.73 | |
| P. Wilson | | | | | | | | | | | 0.7 | |
| L. T. Su | | | | | | | | | | | 0.66 | |
| D. R. Swanson | | | 0.38 | | | | | | | | 0.6 | |
| S. P. Harter | | | | | | | | | | | 0.5 | |
| F. W. Lancaster | | | | | | | | | | | 0.41 | |
| J. Hartley | | | | | | | | | | | | 0.74 |

Appendix III. Factor analysis of 120 authors in information science (1996 – 2005)

(oblique rotation – structure matrix)

- | | | |
|--|--|--|
| 1. User theory; | 2. Evaluative citation analysis; | 3. Experimental retrieval; |
| 4. Webometrics; | 5. Science communication; | 6. Visualization of knowledge domains; |
| 7. Information seeking and context; | 8. Metadata & digital resources; | |
| 9. Bibliometric models & distributions; | 10. Children's information searching behavior; | |
| 11. Users' judgments of relevance (situational relevance); | 12. Structured abstracts (academic writing) | |

| Author | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|----------------------|------|------|------|-------|-------|------|------|------|-------|-------|------|------|
| P. Vakkari | 0.87 | | | | | | 0.56 | 0.37 | | -0.66 | 0.62 | |
| N. Ford | 0.85 | | | | | | 0.54 | 0.38 | | -0.66 | 0.42 | |
| D. Ellis | 0.84 | | | | | | 0.56 | 0.45 | | -0.65 | 0.61 | |
| C. Cole | 0.84 | | | | | | 0.51 | 0.39 | | -0.41 | 0.41 | |
| B. Dervin | 0.82 | | | | | | 0.58 | 0.43 | | -0.62 | 0.55 | |
| T. D. Wilson | 0.81 | | | | | | 0.67 | 0.36 | | -0.61 | 0.5 | |
| R. S. Taylor | 0.81 | | | | | | 0.66 | 0.43 | | -0.58 | 0.61 | |
| B. Allen | 0.81 | | | | | | 0.51 | 0.53 | | -0.74 | 0.6 | |
| C. Kuhlthau | 0.8 | | | | | | 0.57 | 0.46 | | -0.67 | 0.59 | |
| M. J. Bates | 0.78 | 0.37 | | | | | 0.43 | 0.53 | | -0.71 | 0.65 | |
| A. Spink | 0.77 | 0.44 | | | | | 0.36 | 0.48 | | -0.64 | 0.67 | |
| T. Saracevic | 0.76 | 0.43 | | | | | 0.36 | 0.45 | | -0.59 | 0.71 | |
| N. J. Belkin | 0.76 | 0.45 | | | | | 0.41 | 0.49 | | -0.62 | 0.67 | |
| B. Hjørland | 0.72 | | | | | | 0.46 | 0.42 | | | 0.67 | 0.45 |
| P. Ingwersen | 0.7 | | | -0.44 | | | 0.39 | 0.4 | | -0.53 | 0.6 | |
| B. J. Jansen | 0.69 | | 0.5 | | | | | 0.33 | | -0.66 | 0.47 | |
| A. J. Nederhof | | 0.96 | | | -0.4 | | | | -0.34 | | | |
| B. R. Martin | | 0.93 | | | -0.48 | | | | | | | |
| T. Luukkonen | | 0.93 | | | -0.37 | | | | -0.31 | | | |
| T. Braun | | 0.91 | | | -0.33 | | | | -0.38 | | | |
| P. Vinkler | | 0.9 | | | -0.46 | | | | -0.37 | | | |
| G. Lewison | | 0.9 | | | | | | | | | | |
| R. N. Kostoff | | 0.89 | | | -0.45 | 0.39 | | | -0.37 | | | |
| A. F.J. van Raan | | 0.88 | | -0.45 | -0.46 | 0.33 | | | -0.48 | | | |
| J. Katz | | 0.88 | | -0.33 | | | | | -0.43 | | | |
| H. F. Moed | | 0.88 | | -0.35 | -0.49 | | | | -0.48 | | | |
| A. Schubert | | 0.88 | | | -0.32 | | | | -0.46 | | | |
| W. Glanzel | | 0.88 | | -0.34 | -0.33 | | | | -0.52 | | | |
| F. Narin | | 0.86 | | | -0.43 | 0.31 | | -0.3 | -0.36 | | | |
| P. O. Seglen | | 0.85 | | -0.39 | -0.61 | | | | -0.36 | | | |
| D. JD. Price | | 0.81 | | -0.38 | -0.48 | 0.4 | | | -0.55 | | | |
| H. Grupp | | 0.79 | | | | | | | | | | |
| L. Leydesdorff | | 0.76 | | -0.53 | -0.49 | 0.53 | | | -0.4 | | | |
| E. Garfield | | 0.75 | | -0.49 | -0.37 | 0.34 | | | -0.45 | | | 0.31 |
| M. Callon | | 0.69 | | | -0.4 | 0.67 | | | | | | |
| C. Buckley | | | 0.95 | | | | | | | | | |
| K. L. Kwok | | | 0.95 | | | | | | | | | |
| W. B. Croft | 0.37 | | 0.95 | | | | | | | | 0.42 | |
| C. J. van Rijsbergen | | | 0.94 | | | | | | | | 0.35 | |
| E. M. Voorhees | 0.31 | | 0.94 | | | | | | | | 0.38 | |
| D. Harman | 0.32 | | 0.93 | | | | | | | | 0.41 | |
| K. S. Jones | 0.38 | | 0.91 | | | | | 0.34 | | | 0.56 | |
| M. F. Porter | | | 0.91 | | | | | | | | | |
| E. A. Fox | 0.35 | | 0.9 | | | | | 0.44 | | -0.36 | 0.37 | |
| S. E. Robertson | 0.45 | | 0.9 | | | | | 0.34 | | | 0.52 | |
| D. W. Lewis | | | 0.89 | | | | | | | | | |
| W. B. Frakes | | | 0.88 | | | | | | | | | |
| M. A. Hearst | 0.34 | | 0.88 | | | 0.38 | | 0.33 | | | | |
| M. Gordon | | | 0.86 | | | | | | | | 0.42 | |
| Y. Yang | | | 0.82 | | | | | | | | | |
| A. Bookstein | | | 0.81 | | | | | | -0.43 | | 0.35 | |
| W. S. Cooper | 0.35 | | 0.81 | | | | | 0.31 | | | 0.71 | |
| R. M. Losee | 0.41 | | 0.79 | | | | | | | | 0.49 | |
| D. C. Blair | 0.46 | | 0.79 | | | | | 0.46 | | -0.31 | 0.66 | 0.34 |
| C. W. Cleverdon | 0.41 | | 0.77 | | | | | 0.38 | | | 0.76 | |
| W. Hersh | 0.52 | | 0.76 | | | | | 0.44 | | -0.47 | 0.63 | |
| G. Salton | 0.31 | | 0.73 | | | | | | | | 0.39 | |
| H. C. Chen | 0.54 | | 0.73 | | | 0.39 | | 0.47 | | -0.5 | 0.41 | |

Appendix III. (continued)

| Author | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|-----------------|------|------|------|-------|-------|------|------|------|-------|-------|------|------|
| J. Barilan | | | | -0.94 | | | | | | | | |
| A. Smith | | | | -0.94 | | | | | | | | |
| M. Thelwall | | | | -0.92 | | | | | | | | |
| E. Davenport | | | | -0.91 | -0.41 | | | | | | | 0.32 |
| R. R. Larson | | | | -0.91 | | | | | -0.3 | | | |
| S. Brin | | | | -0.89 | | | | | | | | |
| J. Kleinberg | | | | -0.86 | | | | | | | | |
| S. Lawrence | | | 0.31 | -0.86 | | | | | | | | |
| C. Oppenheim | | 0.46 | | -0.84 | -0.56 | | | | | | | 0.36 |
| R. Kling | | | | -0.83 | -0.45 | | | | | | | |
| B. Cronin | | 0.45 | | -0.77 | -0.55 | 0.34 | | | -0.4 | | | |
| R. Rousseau | | 0.51 | | -0.76 | -0.37 | | | | -0.54 | | | |
| L. Egghe | | 0.62 | | -0.63 | -0.42 | | | | -0.55 | | | |
| A. J. Meadows | | 0.43 | | -0.48 | -0.85 | | | | -0.34 | | | 0.47 |
| R. K. Merton | | 0.73 | | -0.32 | -0.76 | 0.39 | | | -0.31 | | | |
| M. H. Macrobets | | 0.73 | | -0.44 | -0.76 | 0.35 | | | | | | 0.31 |
| W. D. Garvey | | | | -0.31 | -0.71 | | 0.52 | 0.31 | | | | 0.46 |
| M. B. Line | 0.31 | | | -0.36 | -0.68 | | 0.39 | | -0.55 | | | |
| V. V. Nalimov | | 0.62 | | | -0.67 | | | | -0.44 | | | |
| T. S. Kuhn | | 0.52 | | | -0.67 | 0.62 | | | -0.46 | | | 0.47 |
| B. Latour | | 0.62 | | | -0.64 | 0.49 | | | | | | 0.42 |
| S. Harnad | | | | -0.62 | -0.63 | | | | | | | 0.34 |
| T. Kohonen | | | 0.47 | | | 0.79 | | | | | | |
| X. Lin | | | 0.42 | | | 0.77 | | 0.3 | | | | |
| C. M. Chen | | | | -0.43 | | 0.75 | 0.41 | | | | | |
| E. CM. Noyons | | 0.69 | | | | 0.72 | | | | | | |
| K. W. McCain | | 0.33 | | -0.4 | -0.48 | 0.7 | | | | | | 0.47 |
| H. D. White | | 0.38 | | -0.45 | -0.52 | 0.67 | | | -0.36 | | | 0.41 |
| H. Small | | 0.61 | | -0.41 | -0.59 | 0.64 | | | -0.39 | | | |
| E. A. Chatman | 0.48 | | | | | | 0.88 | | | -0.35 | | |
| R. Savolainen | 0.55 | | | | | | 0.87 | | | -0.39 | | |
| D. Nicholas | 0.53 | | | | | | 0.78 | 0.35 | | -0.39 | 0.3 | |
| C. Mcclure | | | | | | | 0.76 | 0.4 | | -0.31 | 0.41 | |
| A. P. Bishop | 0.3 | | | | | | 0.72 | 0.59 | | -0.5 | | |
| E. Rogers | | | | | -0.44 | 0.4 | 0.69 | | | | | 0.4 |
| L. Dempsey | | | | | | | | 0.77 | | | | |
| C. Lynch | | | 0.44 | | | | | 0.72 | | | 0.32 | 0.32 |
| A. Dillon | 0.4 | | | -0.31 | | 0.35 | 0.3 | 0.72 | | -0.48 | | |
| C. Tenopir | 0.35 | | 0.37 | | | | | 0.68 | | -0.35 | 0.65 | 0.46 |
| M. K. Buckland | 0.6 | | | | | | 0.4 | 0.64 | -0.32 | -0.39 | 0.56 | 0.44 |
| Q. L. Burrell | | 0.45 | | -0.31 | | | | | -0.87 | | | |
| B. C. Brookes | 0.48 | | | -0.34 | -0.33 | | | | -0.77 | | 0.37 | 0.31 |
| H. A. Simon | 0.32 | 0.42 | | | | 0.36 | | | -0.76 | | | |
| A. Large | 0.37 | | | | | | | | | -0.9 | | |
| D. Bilal | 0.39 | | | | | | 0.3 | | | -0.89 | 0.3 | |
| P. Solomon | 0.57 | | | | | | 0.55 | 0.34 | | -0.88 | 0.36 | |
| R. Fidel | 0.76 | | | | | | 0.39 | 0.53 | | -0.81 | 0.61 | |
| B. Shneiderman | 0.59 | | | | | | | 0.58 | | -0.76 | 0.32 | |
| C. L. Borgman | 0.58 | | | -0.32 | | 0.36 | 0.54 | | | -0.75 | 0.53 | |
| G. Marchionini | 0.73 | | | | | 0.43 | 0.53 | | | -0.79 | 0.51 | |
| P. Wang | 0.74 | | | | | 0.42 | 0.37 | | | -0.75 | 0.7 | |
| J. Nielsen | | | 0.33 | | | | | 0.45 | | -0.7 | | |
| K. Markey | 0.6 | | 0.47 | | | | | 0.65 | | -0.65 | 0.63 | 0.31 |
| J. Janes | 0.49 | | | | | | | 0.37 | | -0.41 | 0.92 | |
| L. Schamber | 0.58 | | 0.38 | | | | 0.35 | 0.42 | | -0.43 | 0.88 | |
| C. L. Barry | 0.53 | | | | | | 0.38 | 0.37 | | -0.47 | 0.87 | |
| L. T. Su | 0.64 | | 0.45 | | | | | 0.45 | | -0.6 | 0.87 | |
| P. Wilson | 0.59 | | | | | | 0.45 | 0.38 | | -0.37 | 0.86 | 0.32 |
| S. P. Harter | 0.58 | | 0.4 | -0.39 | | | 0.3 | 0.48 | | -0.44 | 0.77 | 0.31 |
| D. R. Swanson | 0.33 | | 0.59 | | | 0.38 | | | | | 0.75 | 0.4 |
| F. W. Lancaster | 0.58 | | 0.46 | | | | | 0.59 | | -0.53 | 0.75 | 0.42 |
| J. Hartley | | | | | | | | 0.34 | | | | 0.73 |