Visualizing a Discipline: An Author Co-Citation Analysis of Information Science, 1972–1995

Howard D. White and Katherine W. McCain
College of Information Science and Technology, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104. E-mail: whitehd@duvm.ocs.drexel.edu

This study presents an extensive domain analysis of a discipline—information science—in terms of its authors. Names of those most frequently cited in 12 key journals from 1972 through 1995 were retrieved from Social Scisearch via DIALOG. The top 120 were submitted to author co-citation analyses, yielding automatic classifications relevant to histories of the field. Tables and graphics reveal: (1) The disciplinary and institutional affiliations of contributors to information science; (2) the specialty structure of the discipline over 24 years; (3) authors' memberships in 1 or more specialties; (4) inertia and change in authors' positions on 2-dimensional subject maps over 3 8-year subperiods, 1972–1979, 1980–1987, 1988–1995; (5) the 2 major subdisciplines of information science and their evolving memberships; (6) "canonical" authors who are in the top 100 in all three subperiods; (7) changes in authors' eminence and influence over the subperiods, as shown by mean co-citation counts; (8) authors with marked changes in their mapped positions over the subperiods; (9) the axes on which authors are mapped, with interpretations; (10) evidence of a paradigm shift in information science in the 1980s; and (11) evidence on the general nature and state of integration of information science. Statistical routines include ALSCAL, INDSCAL, factor analysis, and cluster analysis with SPSS; maps and other graphics were made with DeltaGraph. Theory and methodology are sufficiently detailed to be usable by other researchers.

Introduction

Co-citation analysis shows that literatures cohere and change in intelligible ways over time, whether one defines them in terms of individual articles and books, whole oeuvres, or journals. Author co-citation analysis (ACA) is the subcategory that maps oeuvres, and, by implication, the people who produce them. The raw data are counts of the times that selected author pairs are cited together in articles, regardless of which of their works are cited. ACA synthesizes many such counts. Now that 15 years have passed since it was introduced by White and Griffith (1981), the present writers wish to explore this literature-based technique as a means for contributing to intellectual history. As in that earlier article, we shall use authors from information science to illustrate that, although ACA is applicable in any discipline (Bayer, Smart, & McLaughlin, 1990; Eom, 1996; Hoffman & Holbrook, 1993; McCain, 1986), readers of this journal will best be able to judge its validity when it is applied to their own field. Ultimately, we are interested in ACA as a way of visualizing a field through a representative slice of its literature, and we shall develop evidence from it for those who would define information science in terms of its specialties. Unlike McCain (1990) or White (1990a, 1990b), the present article is not an introduction to ACA but an exercise in domain analysis, in the sense developed by Hjørland and Albrechtsen (1995). On one level, it relates to the creation of an intellectual framework for information science (Buckland & Liu, 1995, p. 389); on another level, it relates to the visualization of linguistic data (Williams, Sochats, & Morse, 1995).

Because the data of ACA are merely noun phrases and associated citation counts, they produce history of the cliometric sort, which leaves out almost all the good parts, such as who had shouting matches, who slept with whom, and what actually gave rise to the most significant work. All ACA can do, for the historian of ideas or any other party, is to identify influential authors and display their interrelationships from the citation record. It is no substitute for extensive reading and fine-grained content analysis, if someone is truly interested in the intellectual history of a field. But with Social Scisearch data, which start in 1972, it is now possible to show author relationships over several multi-year periods—in the present study, three periods of 8 years each, overlapping most of the existence of the American Society for Information Science, which

Received December 3, 1996; revised March 6, 1997; accepted March 6, 1997.

replaced the American Documentation Institute in 1968. The three periods allow us to test our techniques with citations from wholly disjoint sets of writings. With three independent replications, one can examine stability and change in author placements overall, and in the specialty groupings they indicate. One can also see the emergence of authors representing new lines of work.

Further evidence for the historian includes the relative eminence of authors in terms of citedness, the brightening or dimming of reputations, the most heavily cited works (when these are traced), the specialties in which contributions are recognized, the major institutional bases of the discipline (traced through author affiliations), the extradisciplinary writers on whom information science regularly draws, and the precontemporary figures who remain influential. Authors may also be categorized demographically, e.g., by sex or by country of origin.

We defined the authors of information science as all those cited in 12 journals, as listed below. Authors were ranked in order of citedness for the entire period covered by Social Scisearch, 1972–1995. Co-citation data were retrieved for all pairs in the top-ranked 120, from which we produced:

- A factor analysis of the 120 authors for the entire 24-year span, 1972–1995, which reveals the specialty structure of the discipline. Factor analysis, unlike multidimensional scaling and clustering, can show an author's contribution to more than one specialty.
- Analyses of the 120 authors' mean co-citation counts, which indicate their standing and influence in the discipline as of 1972–1979, 1980–1987, 1988–1995, and at the end of the three periods combined.
- Two-dimensional maps of the top 100 authors in each of the 8-year periods (made with ALSCAL, the SPSS multidimensional scaling program).
- A map of authors whose ''citation images''—see below—changed markedly over the years of our study.
- A two-dimensional composite map of the authors who are in the top 100 in all three periods—some 75 in all. Their most cited works arguably make up the canonical literature of information science. Certain statistics generated by the mapping routine (INDSCAL, a part of ALSCAL) may bear on paradigm shift in the discipline.

It is well known that more than one disciplinary group has used the name ''information science'' for its activities. Our choice of journals operationalizes the field as it is construed in section 5 of Machlup and Mansfield (1983), where it is called ''library and information sciences.'' (We shall drop the plural and use the abbreviation L&IS.) Broadly speaking, this field concerns itself with modeling the world of publications, with a practical goal of being able to deliver their content to inquirers on demand. It is very much implicated with large, content-bearing linguistic structures like indexes, catalogs, and assemblages of full text. Other forms of information science, such as those discussed elsewhere in Machlup and Mansfield, involve symbol manipulations that are relatively content-neutral; they are infrastructural to information science in our sense.

The field in our sense has been mapped by Small (1981), Saito (1984), and Persson (1994), all using individual papers as data points; and by Tudjman, Tudor-Silovic, Boras, & Milas-Bracovic, (1988) and Karki (1996), using authors' oeuvres. White and Griffith (1982) included some information scientists in a map of key authors in studies of science, technology, and society. The most direct antecedent, White and Griffith (1981), mapped the positions of 39 information scientists in intellectual space on the basis of how their oeuvres had been co-cited by various writers in the journal literature during 1972–1979. The map showed authors in bibliometrics as central, flanked by authors in scientific communication (including citation researchers) and in information retrieval (including ''generalists''). It was partly on the evidence of this map that, in White and McCain (1989), we claimed information science to be split into two major specialties: Bibliometricians (including citationists) on the one hand, and retrievalists on the other—a claim since protested in Wilson (1996) as unnecessarily restrictive. Harter (1992) observes that the two camps are not well integrated intellectually.

This article sheds further light on that issue, among others. It is the fullest example of ACA to date, with the largest data exhibits. Even so, one may envision a time when many more authors can be routinely processed for display. ACA's present limits are computational, imposed by the set-forming capabilities of DIALOG or by the SPSS multidimensional scaling programs. It is still a labor-intensive methodology, and we look forward to the day when it can be performed online by vendor-provided software or offline in personal computers with downloaded data.

Authors' Citation Images

In any field of scholarship, writers make judgments as to who has written on what, using what methods, and they reflect the judgments in their citing practices. Aggregated over time, these practices assume definite structure: Writers show commonalities in how they judge the subject matter, methodology, and intellectual style of other writers; for example, they often attach the same meanings and significance to precedent works (Cozzens, 1985; Small, 1978). Call this structure, for which writers are jointly not singly responsible, the consensus on past literature. Small (1980) equates it with a Kuhnian paradigm. It is not, however, a consensus gained obtrusively, by getting the people around a table to agree. It is defined behaviorally, as the citing practices of many writers, and it is gained unobtrusively, through access to the citation data of the Institute for Scientific Information (ISI), such as those in Social Scisearch.

Author co-citation mapping with ISI data yields pictures from the history of the consensus. It suggests how authors are commonly viewed on two dimensions, often interpretable as *subject matter* and *style of work.* (The poles of the latter might be labeled, e.g., hard–soft, quantitative–qualitative, more mathematical–less mathematical). Author clusters placed on these two dimensions can be interpreted as specialties within a discipline (White, 1990a, 1990b).

The maps transcend the viewpoint—and the individual biases—of any one observer. For that very reason, however, some may dismiss them. Because the publication process is beset with lags, ACA shows the way a field looked some time ago, not necessarily how it looks today or will look tomorrow. Moreover, computational limits now constrain ACA to authors with citation records of a certain magnitude. Others no less interesting, including perhaps the newly fashionable, fall below the threshold. Thus, the consensus revealed may not be the present one, nor the field's open portal into the future (cf. DeMey, 1982).

What, then, is the best case to be made for ACA? Briefly, by highlighting the main conjunctions that citers have made among oeuvres, the maps and other displays simplify literatures to *writings related by use.* This is a more rigorous grouping principle than that of typical subject indexing, because it depends not on perfunctory indication of content by nonspecialists, but on repeated statements of connectedness by citers with subject expertise. Counts of these statements, obtainable from ISI citation indexes, are the means by which simplification is achieved. Writings wholly uncited fall away from consideration, as do writings cited below some minimal threshold, which the mapper can set and vary. Left standing are writings above the threshold, and they themselves can be ordered for presentation by their citation totals or averages.

The notion of ''the historical consensus'' in author maps can be made more precise. What is actually mapped is an author's *citation image.* Everyone ever cited has one, but only those who have been cited in many writings are likely to figure in ACA. In the latter case, the image has a constant part, the author's identity as it is rendered in successive reference lists. (To assemble this ''constant,'' one must know the different ways in which citers may cite the same author, the means for differentiating identically named authors, and the data entry conventions of ISI.) The image also has a variable part, the gradually increasing set of other author-names that co-occur with a given author in those lists. At the end of a time period, ACA sums up the record by mapping the author as a single point among other selected author-points on the basis of the repeated co-occurrences. Authors with similar profiles of co-occurrences are displayed close together. If the same authors are mapped for more than one period, one can look for trends and discontinuities. Such is the history of a field as told by ACA.

Note that the technique omits another important variable, the identities of the *works* that are co-cited. While these are recoverable through online searching, large-scale mapping with individual works, as opposed to oeuvres or journals, is feasible only with software to which few have access. We have not attempted here to reveal the individual works to which citers are responding. Some authors might be stably identified with one or two concepts (such as ''invisible colleges''); others might be more variegated in citers' eyes, with many different uses made of their works. Since the citation image reflects only the use of one's oeuvre by subsequent writers (including oneself), it rarely implies one's total subject range. For example, there are many prominent information scientists who could claim, by virtue of their publications, to have worked in both retrieval theory and bibliometrics. Nevertheless, their map positions may fail to reveal it, because their citing colleagues do not sufficiently use the full range of their works (or have not yet). Those mapped are like versatile actors who are typecast after success in a particular kind of part. The ''historical consensus'' on them captures only what citers have recognized them for, not what they have actually done.

However, let us not be unduly dismissive. It is no mean power to be able to automatically display authors in the aspect most salient to their peers. (Factor analysis, as mentioned, may suggest more than one aspect.) ACA focuses attention on author relationships most widely perceived by experts over time. As such, it should be an aid to any endeavor in which an overview of a literature or subliterature is useful before reading begins (as in online searching or library collection development), or in which simplification and synopsis are helpful to the newcomer (as in literature reviews or histories of a field; cf. White, 1990a, 1990b). Moreover, changes in some authors' citation images *can* be detected through ACA, as will be shown.

The decisive argument for ACA is that it enables one to see a literature-based counterpart of one's own overview of a discipline. In our experience, the agreement between the computed and the private overview is generally quite good. We thus have an answer for the person who looks at our graphics and says, ''I know all that already.'' If that indeed is the case, then we have made technical progress, since we can now reproduce much of the disciplinary expert's view on behalf of someone who does not know as much, and we can do it without benefit of the expert. Of course, ACA might also help the expert who wants visual corroboration, from external sources, of what would otherwise be private knowledge. It can automatically generate ancillae to insiders' histories and reviews, such as those in Vakkari and Cronin (1992) and Olaisen, Munch-Petersen, and Wilson (1996).

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
Journal of Documentation
Journal of Information Science
Library & Information Science Research (and *Library Research*)
Proceedings of the American Society for Information Science (and *Proceedings of the ASIS 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

This capability in ACA qualifies it, and similar kinds of domain analysis, as a form of artificial intelligence (AI). It is a form that takes advantage of the intelligence stored in repeated linkages of texts rather than in clever new programming, and so it is not likely to attract much attention in traditional AI. That leaves an opening for ASIS-type information scientists. It is our sense that huge amounts of domain knowledge remain to be exploited in bibliographic and textual databases—not the first place that typical AI researchers would look.

Methodology

Given current computational resources, the primary decision in author co-citation analysis is the choice of the authors to be mapped. The roster of authors for the present article was created from a broader base and with rather less subjective criteria than in White and Griffith (1981). Most of the 39 authors mapped there were chosen because they appeared in Griffith's (1980) ''key papers'' anthology; the remainder were his or White's personal additions. The number 39 was imposed by what a multidimensional scaling program could then handle (MDSCAL's 40 cases). That number has risen to 100 with ALSCAL, the multidimensional scaling routine in SPSS, and we wished to press the new capability to the fullest extent. At the same time, we wished to let ''the field'' dictate its top authors rather than choosing them ourselves. At present, the best way to do this is by operationalizing the field in terms of its journals.

The 1993 *Journal Citation Reports* for *Social Sciences Citation Index* (*SSCI*) lists 53 titles in the journal subject category ''Information Science & Library Science.'' In order to limit our universe as much as possible to authors writing in mainstream information science, basic or applied, we chose the 12 serials listed in Table 1. Three are sponsored by the American Society for Information Science (*JASIS*, *ARIST*, and the *Proceedings* of the ASIS annual meeting); two, by its British counterparts, the Association of Special Libraries and Information Bureaux (*Journal of Documentation*) and the Institute of Information Scientists (*Journal of Information Science*). We added three further English-language journals with strong IS orientations, as indicated by title and scope statements. Finally, we added four journals representing research in the development of automated library systems and electronic or digital libraries.

On Social Scisearch, the online *SSCI*, we retrieved the set of all articles from these 12 serials from 1972 through 1995 inclusive. Not all 12 have runs coextensive with Social Scisearch; some started earlier and some, later. When applicable, we searched serials under both present and former titles; the latter are parenthesized in Table 1. DIALOG's RANK command was then used to order the authors in the 12 serials by their citation counts over the 24-year period. These counts are thus *subsets* of the authors' total counts in Social Scisearch. In a study of information science per se, we do not necessarily want to list the latter, especially those of multidisciplinary behemoths like Thomas S. Kuhn, Robert K. Merton, and Herbert A. Simon.

Implemented in 1993, RANK allows searchers to retrieve a set of documents and then, within that set, to rank all the terms from a particular field by frequency of occurrence. European bibliometricians have long been familiar with its equivalents (e.g., ZOOM on ESA/IRS, EXTRACT on DIMDI, GET on ORBIT, and MEMS on Questel). For American researchers it is a godsend, in that it opens DIALOG's rich library of databases, including the citation databases of ISI, to co-occurrence analysis. DIALOG users who wish to gather citation counts for author or journal co-citation maps can now do so only semilaboriously, although there are still sharp limits on what can be done.

We downloaded the 300 highest-ranking authors and, as a matter of practicality, elected to base our co-citation analysis on the top 120. To obtain co-citation counts for 120 authors requires the ANDing of 120 (119)/2 = 7,140 distinct author pairs. This is no small task on DIALOG, even with its expanded retrieval set limit (from a maximum of 200 sets to at least 400 sets in 1996). In the early 1980s, it was possible to use an in-house DIALOG command called .INTERSECT to AND multiple pairs of terms relatively easily, but that command was withdrawn in 1983. Since then, online collection of co-citation data has been difficult (McCain, 1990). Fortunately, we were able to use a macro by one of our former students, Lauris Olson, which involves saving DIALOG sets online and then combining them in intersection statements created offline. The macro systematically uploads unique blocks of statements until, over the matrix of all possible author pairs, all nonduplicate pairs have been ANDed. This was done for each of our three time periods, 1972–1979, 1980–1987, and 1988–1995, in Social Scisearch in early

1996. Periods were identified by accession number ranges with the LIMITALL command.

Social Scisearch made changes in its coverage of the information science literature over the three periods, which would affect co-citation counts. We do not know the extent of the changes, but we note them as a source of uncontrolled variation in our data. Another source of uncontrolled variation is that, from 1972 to 1995, most of the oeuvres were themselves changing as their living authors brought out new works.

While our authors were chosen on the basis of their *citation* counts in 12 serials, we did not attempt to limit their *co-citation* counts to those occurring in the 12; the set-formations would have become prohibitively numerous. Rather, we took their co-citation counts from Social Scisearch as a whole and thus picked up the full range of references to our authors (including self-citations) after they have been paired. In the first step, sets were made of all articles in which a particular author is cited, for example, ''Select CA = Griffith B?'' where ''CA'' means ''cited author.'' ISI replaces cited authors' first and middle names with initials. The truncation symbol ? allows one to retrieve citing articles whether they cite an author by first-name initial only or by those for additional names. Some erroneously high counts result from this procedure—for example, ''CA = Wilson P?'' retrieves papers citing P. Wilsons other than the intended Patrick Wilson—but these are largely corrected in the second step when *pairs* of authors are ANDed, since only the ''right'' P. Wilson co-occurs in reference lists with such information scientists as T. Saracevic or M. Kochen. Special sets were formed for authors who have published under different names (e.g., Karen Markey or Drabenstott; Pauline Atherton or Cochrane) or whose surnames are complicated (e.g., Derek de Solla Price, cited as Price, Solla Price, and de Solla Price).

Once the co-citation counts were obtained, we used the computer techniques described in McCain (1990) to make the maps. As is well known, the closeness of author points on such maps is algorithmically related to their similarity as perceived by citers. We use Pearson *r* as a measure of similarity between author pairs, because it registers the *likeness in shape* of their co-citation count profiles over all other authors in the set. (Pearson *r* has been criticized for obscuring the relative *height* of profiles, but we have other ways of suggesting that.) White and Griffith's way of evaluating the diagonal in the raw data matrix before correlation was replaced by McCain's (1990) treatment of the diagonal as missing data. Little difference between the two treatments is observable on actual maps.

The raw co-citation counts were converted to Pearson *r* correlation matrices by the FACTOR routine in SPSS, and factors were extracted by principal components analysis with varimax rotation. The default criterion of ''eigenvalues greater than one'' determined the number of factors extracted. Missing data were handled through means substitution.

The Pearson *r* correlation matrices for ALSCAL and CLUSTER in SPSS were generated with another SPSS rountine, CORRELATIONS (cf. McCain, 1990). They were treated as nonmetric (ordinal) similarity data in ALSCAL and grouped by the complete linkage method in CLUSTER. Subdisciplinary groupings of the author-points on the maps are based on the dendograms from CLUSTER. Missing data were handled through pairwise deletion.

The ALSCAL maps for each 8-year period could have only 100 authors apiece, and so, for each period, the 100 authors with the highest mean co-citation counts were chosen from the pool of 120. As it happens, 75 authors have sufficiently high means to appear on all three maps. Of the remainder, some do not ''make the cut'' until the second or third period; others make it in the first or second and then fall below it. Coincidentally, the 120 authors originally chosen cover the full range of authors from the three separate periods.

Authors in the top 100 in all three periods—''the canonical 75''—were separately mapped with INDSCAL, a routine in the ALSCAL bundle that does a specialized kind of multidimensional scaling. The input data to INDSCAL are judgments on the similarity of a set of stimuli by a set of judges. INDSCAL reveals not only the judges' *composite* view of the stimuli in multidimensional space, but the weight each *individual* judge gives each dimension; INDSCAL is short for ''individual differences scaling.'' We used the individual weights in a new way (explained further below) to explore the notion of ''paradigm shift'' as it affects the canonical 75.

These ''all caps'' statististical routines like ALSCAL and INDSCAL appear as headings in SPSS manuals (e.g., *SPSS*, 1990), where their output is explained at length for generalist readers.

We used DeltaGraph charting software to map the author points in two dimensions according to their ALSCAL coordinates. There was no rotation of axes. When necessary, we reversed the polarity (positive and negative signs) of ALSCAL's default coordinates, so that the new maps would have the same left-right orientation as in White and Griffith (1981) and Persson (1994). This provides continuity of display without affecting substantive interpretation.

The two-dimensional space in which the authors appear is relative, not absolute, and it fails to capture certain relationships among oeuvres that appear in higher dimensionality. By rotating author points in three-dimensional space, as can now be done with microcomputer software, one can bring out features that two-dimensional projections may hide, but we decided to go with the default output on the ground that it is most replicable. It also resembles what online searchers would get if, in real time, software could put 2-D co-citation maps on their screens to guide literature retrieval.

We must add that we do not regard our 120 authors as wholly definitive of information science, past, present, or future. As in White and Griffith (1981), the *N* was set by computational limits and does not come close to exhausting the authors in information science, even in the limited sense of the field as defined by the journals we chose. Hundreds of other authors are cited in those journals, and their number could be extended by adding other journals in L&IS. A still larger group could be selected from journals outside L&IS proper. There are also new information scientists on the Internet who do not even publish (perhaps because they are so busy selling search-engine stock). For choosing representative authors from a much larger group, citedness above some threshold seems a good criterion, both fair and replicable. But it leaves out many information scientists whose work we know and respect, and colleagues could supply further omissions. We would simply claim that a set of 120 oeuvres exceeds those of past studies and should suffice for our main purpose, which is to convey, through representative authors' names, the main subdivisions of the field as they have evolved in the past quarter-century.

Results

The 120 Authors

Table 2 alphabetizes the 120 most-cited authors in the specified journals for 1972–1995. The additional information given for each author indicates some disciplinary and institutional bases of information science, broadly construed. The field is indeed multidisciplinary and would appear even more so had we listed the disciplines in which contributors were trained, instead of our best guess at those (often suggested by academic departmental names) in which they are or were employed. The disciplines in which the 120 hold advanced degrees extend across the arts and sciences. Their affiliations include universities, public organizations, and private firms.

Institutions placing at least three authors in the top 120 over the years include MIT (6), Michigan (5), UC Berkeley (5), UCLA (5), Illinois (4), Rutgers (4), Syracuse (4), Columbia (3), and Drexel (3). A fair number of other institutions have placed two. (Different assignments would produce slightly different results, e.g., assigning Paisley to Stanford would give Stanford three authors. But Table 2 has the most recent affiliations we could verify.) The authors are not necessarily from schools of library and information science—for example, Columbia's are all from its department of sociology, and MIT has never had such a school. Moreover, the list contains a number of widely influential researchers who are not information scientists at all (in the L&IS sense), but whose works have nevertheless been important to the discipline. These include figures who have not published directly in it, such as Bush, the Coles, Knuth, Kuhn, Merton, Shannon, Winograd, and Zipf. Others, such as Crane, Machlup, Shneiderman, Simon, and Zuckerman, have contributed to L&IS journals, but are known primarily for work in other fields.

Of those who teach in schools that offer degrees in librarianship, few are practicing librarians. In a separate article, we may address the degree of overlap between these authors of information science and the authors most cited in library science journals. It is considerable, but, as one would suspect, many authors in the latter list do not appear here.

Our selection rules to some degree privilege authors who have been around a long while, in particular those who were established enough in the 1970s to make the list during the early years of Social Scisearch (1972–). These include ancestral figures from the 1920–1950 period (Lotka, Zipf, Bradford, Shannon, Merton, Bush), as well as many more who had made their reputations between 1950 and 1970 (e.g., Price, Garfield, Brookes, Kochen, Salton, Sparck Jones). Note, however, that such persons had to remain well-enough cited during the entire 1972–1995 period to remain in competition. Note, too, that our rules were sufficiently inclusive to bring in dozens of persons who rose to distinction from 1970 on, some quite recently.

While probably few will be surprised by the early figures who made the list, we would remind readers to contrast the continuing citedness of, e.g., Lotka, Bradford, Bush, and Shannon, with that of other precursors, such as Donkers-Duyvis, whose contributions are in relative eclipse (cf. Buckland, 1996). Citation records and images are important to the intellectual historian not least because they raise questions as to why some authors remain influential for many years, while others, perhaps as deserving, do not.

In the vein of Mullins (1973) on theory groups, it would be interesting to trace mentor–student and ''trusted assessor'' links among persons in the top 120 (and beyond); some definitely exist, although that line of inquiry exceeds our scope here. We can, however, report that the male–female ratio on the list is exactly 5 to 1 (100 to 20).

Since this is Anglophone information science, our authors are overwhelmingly North American or British, but Braun, Egghe, Ingwersen, Moed, and Schubert represent continental Europe, and Lawani, Africa.

Specialties

The results of the factor analysis, incorporating 24 years' worth of data for the 120 authors, are presented in Table 3. They could hardly be clearer or more revealing as to the nature of the discipline.

Twelve factors were extracted; jointly (R^2), they ex-

TABLE 2. Names, disciplines, and institutions of 120 authors in information science

Name	*Discipline*	*Institution*	*Name*	*Discipline*	*Institution*
Thomas J. Allen	Management	MIT	Thomas S. Kuhn	Philosophy	MIT
Derek Austin	L&IS	British Library	F. Wilfrid Lancaster	L&IS	U Illinois
Henriette D. Avram	L&IS	Library of Congress	Stephen M. Lawani	L&IS	Intl Inst Trop Agric, NI
Marcia J. Bates	L&IS	UCLA	Ferdinand F. Leimkuhler	Indus Engineering	Purdue U
Nicholas Belkin	L&IS	Rutgers U	Maurice B. Line	L&IS	British Library
David C. Blair	Management	U Michigan	Alfred J. Lotka	Mathematics	Metropolitan Life Insur
Abraham Bookstein	L&IS	U Chicago	Hans Peter Luhn	Comp Sci	IBM
Christine L. Borgman	L&IS	UCLA	Fritz Machlup	Economics	Princeton U
Harold Borko	L&IS	UCLA	Richard S. Marcus	Info Systems	MIT
Charles P. Bourne	L&IS	DIALOG Info Service	M. E. Maron	L&IS	UC Berkeley
Samuel C. Bradford	L&IS	Science Library, UK	Ben R. Martin	Science Studies	U Sussex
Tibor Braun	Science Studies	Acad Sciences Lib, HU	James Martin	Info Systems	IBM
Bertram C. Brookes	L&IS	U London	John Martyn	L&IS	Central London Poly
Michael J. Buckland	L&IS	UC Berkeley	Joseph R. Matthews	L&IS	J. Matthews Associates
Vannevar Bush	Engineering	MIT	Katherine W. McCain	L&IS	Drexel U
Mark P. Carpenter	Science Studies	CHI Research	Charles R. McClure	L&IS	Syracuse U
Anthony E. Cawkell	L&IS	Citech, Ltd	Charles T. Meadow	L&IS	Toronto U
Ching-Chih Chen	L&IS	Simmons College	A. Jack Meadows	L&IS	Loughborough U
Daryl E. Chubin	Science Studies	Ofc Technol Assessment	Robert K. Merton	Sociology	Columbia U
Cyril W. Cleverdon	L&IS	Cranfield College, UK	Henk F. Moed	Science Studies	U Leiden, NE
Pauline Cochrane	L&IS	U Illinois	Michael J. Moravcsik	Science Studies	U Oregon
Jonathan R. Cole	Sociology	Columbia U	Philip M. Morse	Operations Research	MIT
Stephen Cole	Sociology	SUNY Stony Brook	Francis Narin	Science Studies	CHI Research
Michael D. Cooper	L&IS	UC Berkeley	Edward T. O'Neill	L&IS	OCLC
William S. Cooper	L&IS	UC Berkeley	Robert Oddy	L&IS	Syracuse U
Diana Crane	Sociology	U Pennsylvania	William Paisley	Communications	Knowledge Access
Bruce Croft	Comp Sci	U Mass Amherst	Miranda Pao	L&IS	U Michigan
Blaise Cronin	L&IS	Indiana U	Derek de Solla Price	History of Science	Yale U
Carlos Cuadra	L&IS	Cuadra Associates	Tadeus Radecki	Comp Sci	U Nebraska
Brenda Dervin	Communications	Ohio State U	Ronald E. Rice	Communications	Rutgers U
Martin Dillon	L&IS	OCLC	Stephen E. Robertson	L&IS	City U London
Tamas E. Doszkocs	L&IS	Natl Library of Medicine	William B. Rouse	Info Systems	Search Technologies
Karen Drabenstott	L&IS	U Michigan	Gerard Salton	Comp Sci	Cornell U
Leo Egghe	L&IS	Limburg U, BE	Tefko Saracevic	L&IS	Rutgers U
Robert A. Fairthorne	L&IS	Royal Aircraft Estab, UK	Andras Schubert	Science Studies	Acad Sciences Lib, HU
Jason Farradane	L&IS	U Western Ontario	Claude E. Shannon	Telecommunications	Bell Labs
Carol Fenichel	L&IS	Allegheny U	William M. Shaw, Jr.	L&IS	U North Carolina
Raya Fidel	L&IS	U Washington	Ben Shneiderman	Comp Sci	U Maryland
Edward T. Fox	Comp Sci	Virginia Polytechnic Inst	Herbert A. Simon	Economics, Cog Sci	Carnegie Mellon U
J. Davidson Frame	Management	George Washington U	Henry G. Small	Science Studies	ISI
Eugene Garfield	Science Studies	ISI	Linda C. Smith	L&IS	U Illinois
William Garvey	Psychology	Johns Hopkins U	Karen Sparck Jones	Comp Sci	Cambridge U
William Goffman	L&IS	Case Western U	Elaine Svenonius	L&IS	UCLA
Michael Gorman	L&IS	Calif State U Fresno	Don R. Swanson	L&IS	U Chicago
Belver C. Griffith	L&IS	Drexel U	Renata Tagliacozzo	Health Research	U Michigan
Stephen P. Harter	L&IS	Indiana U	Jean Tague-Sutcliffe	L&IS	U Western Ontario
Donald T. Hawkins	L&IS	Bell Labs	Robert S. Taylor	L&IS	Syracuse U
Robert M. Hayes	L&IS	UCLA	Carol Tenopir	L&IS	U Tennessee
Charles R. Hildreth	L&IS	U Oklahoma	C. J. van Rijsbergen	Comp Sci	U Glasgow
Peter Ingwersen	L&IS	Royal School Libnship, DK	Brian C. Vickery	L&IS	U London
Paul B. Kantor	L&IS	Rutgers U	Jan Vlachy	Science Studies	Natl Acad Sciences, CS
Jeffrey Katzer	L&IS	Syracuse U	Howard D. White	L&IS	Drexel U
E. Michael Keen	L&IS	U Wales Aberystwyth	Martha E. Williams	L&IS	U Illinois
Allen Kent	L&IS	U Pittsburgh	Patrick Wilson	L&IS	UC Berkeley
Myer M. Kessler	L&IS	MIT	T. D. Wilson	L&IS	U Sheffield
Frederick G. Kilgour	L&IS	U North Carolina	Terry Winograd	Comp Sci, Cog Sci	Stanford U
Donald W. King	L&IS	King Research	Clement T. Yu	Comp Sci	U Illinois Chicago
Donald E. Knuth	Comp Sci	Stanford U	George K. Zipf	Social Sciences	Harvard U
Manfred Kochen	Business	U Michigan	Harriet A. Zuckerman	Sociology	Columbia U
Donald H. Kraft	Comp Sci	LSU	Pranas Zunde	Comp Sci	Georgia Tech

plain 84% of the variance. After rotation, they appear, from left, in order of decreasing variance explained. The present writers labeled them in the column headings. Even if our labels are contested, these wholly automatic outcomes will probably make good sense to anyone familiar with ASIS-style information science. Whether one reads

TABLE 3. Factor analysis of 120 authors in information science.

Name	Experi-mental Retrieval	Citation Analysis	Online Retrieval	Biblio-metrics	General Library Systems	Science Commun-ication	User Theory	OPACs	Imported Ideas	Indexing Theory	Citation Theory	Commun-ication Theory
Yu	.96											
Radecki	.94											
Robertson	.93											
Croft	.92											
Dillon	.92											
Sparck Jones	.91											
Maron	.91											
van Rijsbergen	.91											
Bookstein	.88											
Luhn	.86											
Fox	.86											
W. S. Cooper	.86											
Kraft	.84											
Blair	.81											
Salton	.78											
Cleverdon	.78									.35		
Oddy	.78		.38				.31					
Shaw	.75											
Harter	.71		.47									
Katzer	.70		.42				.37					
Tague-Sutcliffe	.68			.55								
Doszkocs	.66		.60					.30				
Swanson	.66				.31		.33					
Keen	.61									.56		
Lancaster	.60		.38		.33							
Borko	.59		.33		.41					.35		
Belkin	.58		.31				.54					
Kantor	.57				.51							
Saracevic	.56						.52					
Zunde	.55			.49								
Moed		.92										
Moravcsik		.91										
Narin		.90										
Frame		.90										
Carpenter		.90										
B. Martin		.88										
Braun		.87										
Small		.87										
Schubert		.86										
Vlachy		.86		.38								
Cronin		.85										
Meadows		.84				.36						
Kessler		.84									.32	
Cawkell		.83			.30							
Lawani		.83		.43								
Griffith		.77				.45						
Chubin		.76				.52						
Price		.73				.48						
Line		.72		.36	.36							
McCain		.72									.51	
Garfield		.70				.39						
H. D. White		.70									.52	
Martyn		.65			.49							
Garvey		.65				.59						
Pao		.61		.47								
Kochen	.34	.53		.32	.33				.30			
L. C. Smith	.44	.50	.50									
Rice		.48							.31		.38	.37
Hawkins			.83									
Meadow			.82									
Marcus	.39		.80									
Fenichel			.80									

(*continued*)

TABLE 3. Factor analysis of 120 authors in information science (continued).

Name	Experi-mental Retrieval	Citation Analysis	Online Retrieval (cont.)	Biblio-metrics	General Library Systems	Science Commun-ication	User Theory	OPACs	Imported Ideas	Indexing Theory	Citation Theory	Commun-ication Theory
Tenopir	.35		.74									
Williams			.72		.35							
Borgman			.65				.40	.38				
Fidel	.33		.65				.44	.31				
Bourne			.59		.56							
Bates	.38		.57				.38	.42				
Fairthorne				.91								
Egghe				.90								
Leimkuhler				.89								
Bradford		.32		.84								
Goffman		.38		.79								
Vickery				.78	.33							
Brookes		.31		.75								
Zipf				.75					.48			
Lotka		.40		.70		.32						
Morse				.61	.51							
Kent					.81							
King					.77							
Hayes					.75							
Kilgour					.71			.33				
Chen		.36		.36	.63							
Buckland				.46	.62		.34					
McClure					.62							.40
M. D. Cooper	.40		.52		.56							
Avram					.49			.41				
Cuadra	.39		.45		.47		.30			.39		
Tagliacozzo		.30	.39		.46			.35				
Zuckerman		.52				.76						
Crane		.54				.75						
S. Cole		.57				.74						
Simon						.72						
Kuhn						.72						
J. R. Cole		.62				.68						
Merton		.38				.64						
J. Martin						.62			.58			
Dervin							.85					
T. D. Wilson							.82					
R. S. Taylor							.81					
Ingwersen	.36		.38				.76					
P. Wilson	.32				.44		.59					
Paisley		.34					.53					.42
Matthews			.44					.77				
Gorman					.37			.76				
Hildreth			.46					.75				
O'Neill				.48				.74				
Cochrane			.53					.66				
Drabenstott			.48					.62				
Svenonius	.41							.61		.36		
Winograd									.88			
Shannon				.33		.30			.75			
Machlup						.49			.68			
Knuth	.55								.66			
Shneiderman			.38						.63			
Rouse	.35		.31		.39				.46			
Bush		.31			.32	.37			.38			
Austin	.33							.38		.59		
Farradane	.38				.31					.57		
T. J. Allen						.51						.55

down the columns or across the rows, they seem both accurate and sensitive to nuance.

The first eight factors alone explain 78% of the variance. All have seven or more authors with loadings greater than 0.60 and may be interpreted as specialties within the discipline. The ninth factor captures a rather less coherent group of extradisciplinary authors whose ideas have been imported into information science (e.g., Winograd, Knuth, Machlup). The remaining factors pick up either interesting isolates (Austin, Farradane, Allen) or some residual variation in authors whose main loadings are elsewhere.

Authors are ranked in the specialty on which they load most highly. If they also have loadings above 0.30 on other factors, those are given as well. The latter can be read as recognition for contributions to more than one specialty. Readers of this journal will often know what these contributions are; also, why many authors load on one factor only. If not, searches in Social Scisearch may suggest answers by showing authors' most cited works.

The 0.30 reporting threshold is conventional. A stricter standard would put the threshold at twice the size of statistically significant Pearson *r*'s in the input matrix, which here would be roughly 0.50. If that is done, our factors come close to displaying ''simple structure,'' in which each variable (author) loads on only one factor (specialty). Psychometricians like simple structure because it makes the relation between variables and factors unambiguous. However, we find little to puzzle over when a 0.30 threshold shows authors loading on multiple specialties; the loadings are almost always meaningful in terms of what we know of their work. We invite readers, including authors in the table, to check our perceptions.

The two biggest specialties, obviously, are *experimental retrieval,* which focuses on the design and evaluation of document retrieval systems, and *citation analysis,* which focuses on the interconnectedness of scientific and scholarly literatures, usually with data from ISI. Almost half the authors in our study have their main loadings on one of these two factors—30 retrievalists and 28 citationists. Remarkably, there is almost no overlap of membership: None among those who load primarily with the retrievalists, and in only two cases among those who load primarily with the citationists. This does not mean that authors in the two groups (say, Price and Salton) are never co-cited; in fact, the raw data show a great deal of low-level co-citation among them. But to load above the 0.30 threshold in both specialties in the factor analysis requires relatively high levels of co-citation in both. The only authors in the entire set so recognized by citers—a considerable achievement—are Kochen and Smith.

The third biggest specialty we have labeled *practical retrieval.* Unlike the experimental retrievalists, the authors in this group, rather than working with content-neutral indexing theory, thought experiments, or document testbeds, have tended to discuss retrieval in terms of ''real world'' databases; terms such as ''INSPEC'' or ''DIALOG'' occasionally profane their pens. Of course, the separation between the experimental and the practical retrievalists is by no means pure; Table 3 shows many authors loading above 0.30 in both specialties, and the interesting thing is to see how the factor analysis bestows the primary identification. It would appear, for example, that citers take Lancaster as an experimental retrievalist first and a practical retrievalist second, while for Tenopir, the reverse is the case.

The next specialty we call *bibliometrics*—a word often used to subsume the specialty we labeled *citation analysis.* However, unlike the citationists, the authors who load primarily here, including the pioneers Lotka, Bradford, and Zipf, are most interested in mathematically modeling certain regularities in textual or bibliographic statistical distributions, irrespective of the literatures from which they come. Although they may use subject literatures in their examples, their bent is away from the disciplinary particulars that interest the citationists and toward claims that are relatively timeless, general, and abstract. Chronologically, much of their work predates the availability of data from the Institute for Scientific Information, and even present-day authors in this tradition, such as Egghe, are not known for their dependence on data from ISI.

General library systems theory is a not altogether satisfactory name for a body of writings on library automation, library operations research, library and information service policy, retrieval system evaluation, and many other interconnected topics. The factor's four top authors, Kent, King, Hayes, and Kilgour, are long-time leaders in the movement to computerize library-based information services. They brought to it encyclopedic interests (Kent, literally) that others in the specialty also display. Note, for example, that most of the authors loading on this factor have written textbooks or wide-ranging monographs for use in schools of library and information science. A tendency to integrate and synthesize may be seen not only among those whose main loadings are here, such as Buckland and M. D. Cooper, but also among those with lesser loadings, such as Lancaster, Kantor, Kochen, and Line.

The specialty we call *user theory* is appropriately headed by Dervin, author of a highly cited chapter on ''information needs and uses'' in the 1986 ARIST. Paisley, also here, wrote the ARIST chapter on the same topic in 1971. Those who load primarily on this factor, such as Taylor, T. D. Wilson, and Ingwersen, have developed the cognitive side of information science. The main loading of the field's most acute philosopher, Patrick Wilson, appears here as well.

The secondary loadings in the *user theory* specialty exemplify the factor-analytic technique's sensitivity to nuance. It will be seen that authors who write about literatures—the citationists, bibliometricians, and scientific communication people—never load above 0.30 on this factor, apparently because citers do not perceive their work as having the right psychological content. On the other hand, quite a few retrievalists load above 0.30, and this suggests the nature of the cognition involved. It has to do with problem-solving at the interface where literatures are winnowed down for users with: Question formulation, search strategies, information-seeking styles, relevance judgments, and the like. Note that the factor by no means picks out all retrievalists; those perceived as comparatively uninterested in user psychology, such as Maron, van Rijsbergen, Meadow, and Bourne, do not load here. But authors who perceptibly are interested, as shown by their writings and speeches, are picked out. Saracevic and Belkin load in the 0.50s; Borgman, Fidel, and Bates in the 0.40s; Oddy, Harter, Katzer, Swanson, Fenichel, Cuadra, and Buckland, in the 0.30s. This seems quite a discriminating selection from among the 100-plus authors who do not load mainly on the user-theory factor. Other factors may be examined for the same kind of sensitivity, with the same good results overall.

Authors loading mainly on *scientific communication* all have strong disciplinary identities outside L&IS—for example, in sociology. They may be thought of as explicating the social systems of science, including those in which formal publication of results is an important (but not the only important) part. The sociologists among them all have loadings, some quite high, in citation analysis, confirming their relevance to the study of scientific literatures. (It will also be noted that several citationists load high in scientific communication, e.g., Price, Garvey, Griffith, and Chubin.)

Kuhn and Merton, of course, are giants of the social sciences, and Simon is a Nobel laureate; they are read in several disciplines, and their citation counts run well into the thousands. But it may come as a surprise that they are orienting theorists in L&IS. This is partly attributable to their prominence in the journal *Scientometrics,* but it will be found that they are also well cited—particularly Simon—in other L&IS journals, such as *JASIS.* James Martin, the writer on computerized information systems, apparently loads here not because of strong conceptual ties with the others but because he happens to be cited with them in several review articles—an example of how the factor-analytic technique occasionally may mislead. But Allen, known for studies of communication in R&D settings, seems rightly placed in this column. (He loads slightly more on the last factor, which is largely his own.)

The design of computerized library catalogs, especially for subject searching, is the province of authors who load on OPACs (online public access catalogs). It makes sense that leading authors here, such as Matthews, Hildreth, Cochrane, and Drabenstott, load secondarily in practical retrieval, just as several of the primary authors there, such as Borgman and Fidel, also turn up here. We again note the discernment with which secondary authors were automatically picked out: For example, the presence of Doszkocs, Kilgour, Avram, and Austin on this factor are all readily interpretable. O'Neill's dual interest in bibliometrics and OPACs is captured. The loadings of Svenonius suggest the connections among OPACs, experimental retrieval, and subject indexing theory. Gorman is present—although his (1990) lampoon denies that information science exists—as the doyen of the Anglo-American Cataloging Rules, a major document in making OPACs possible.

As was said, the chief remaining factor seems a collection of authors in other disciplines from whom information science has imported ideas—e.g., cognitive science (Winograd), information theory (Shannon), computer science (Knuth)—that are all variously relevant to the central concern of information science, the human–computer–literature interface. An early symbol of this interface at its utopian best is the memex of Vannevar Bush, who loads on this and three others factors.

Most readers will have already detected nonrandom patterns of cohesion *across* the 12 factors when authors load above 0.30 on more than one. Experimental and practical retrieval cohere in this sense, and so do citation analysis and bibliometrics. In fact, as both author cross-loadings and the maps below suggest, almost all of the factors or specialties in Table 3 can be aggregated upward into two larger subdisciplines: (1) The analytical study of learned literatures and their social contexts, comprising citation analysis and citation theory, bibliometrics, and communication in science and R&D; and (2) the study of the human–computer–literature interface, comprising experimental and practical retrieval, general library systems theory, user theory, OPACs, and indexing theory.

The point to note is that authors with loadings on more than one factor tend to load in *one* of the subdisciplines: They are essentially ''literatures people'' or ''retrieval people.'' Thus, Pao is recognized as both a citationist and a mathematical bibliometrician; Harter, as both an experimental and practical retrievalist. Much rarer is recognition for work not only across specialties but in both subdisciplines. As mentioned, Kochen and Smith have it; Smith's loadings imply that she is about equally known as an experimental retrievalist, a practical retrievalist, and a citationist. Others in this company include Tague-Sutcliffe and Zunde (experimental retrieval and bibliometrics), the very versatile Tagliacozzo (citation analysis, practical retrieval, general library theory, and OPACs), and O'Neill (bibliometrics and OPACs). A number of authors connect the literature-oriented specialties of citation analysis and bibliometrics with the more retrieval-oriented specialities of general library theory and user theory: Line, Martyn, Vickery, Morse, Chen, Buckland, Paisley, and Bush. But that is about it, in a field of 120.

We agree with Harter (1992) that the two subdisciplines of information science are not yet well integrated. Thus, a would-be synthesist might look for clues toward integration in the works of authors who are highly cited in both. Given the authors named, however, 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. Kochen may be an exception.

White and Griffith (1982) contrasted authors with a high loading on a single factor and authors with lower loadings on several. The identities of the first were said to be ''crystallized,'' while the second were ''diffuse'' and ''pervasive.'' In the present author set, highly crystallized authors are exemplified by Yu and Radecki, or Moed and Moravcsik: Their citation images are, to date, very much fixed in a single specialty. Perhaps more interesting are authors who are protean influences. The champions here are Cuadra and Kochen, who load on five specialties each. Bates, Fidel, Tagliacozzo, Rice, and Bush load on four, and many other authors load on three. This does not necessarily mean that the diffusely loading authors write on more topics than those who load on one. It means that, irrespective of content, their oeuvres are cited in a greater variety of contexts, that is, with a greater variety of other oeuvres.

Counts of the number of factors on which an author loads are specific to the literature and the time period being analyzed. In White and Griffith (1982), which looked at studies of science, technology, and society—a field that overlaps information science—Merton loaded above the threshold on six factors; here, he loads above it on two. Only if we could simultaneously analyze the citation records for thousands of authors across scores of disciplines could we approach something like absolute knowledge of their citedness across specialties. And even that knowledge would be time-dependent. For example, separate factor analyses on our authors for 1972–1979, 1980–1987, and 1988–1995 show a few ''short-lived'' specialties, such as one in library operations research in the early period, that disappear in the analysis of data for the full 24 years. Our remarks on authors here should be read in this light; they result from many acts of judgment, not raw natural fact. Nevertheless, even with its present limitations, the factor-analytic approach to ACA seems to us useful, informative, and well worth pursuing with other literatures.

Eminence

Table 4 suggests the course of 120 reputations in information science, as here defined. It presents each author's mean number of co-citations with the other 119 authors for each of the 8-year segments and for the entire 24-year period. Authors are ranked high to low by the latter in the italicized column. Some authors, of course, had no citation counts during 1972–1979, so their rankings depend on their citedness later.

By comparing the italicized rankings, one may infer the eminence of authors. By observing growth, stasis, or decline in citedness across the three time periods, one may infer the changing influence of individual authors during 1972–1995. At the bottom of Table 4, we give as reference points two overall averages for each set of mean counts. Under the 1972–1979 columns, for example, 1.92 is the overall mean, and 1.52 the overall median, for the 120 mean counts from Garfield's to Knuth's.

Why not simply display authors' total citation counts? We have them as they appeared when we found the most-cited authors in 12 journals in early 1996. But we lack total counts for 120 authors in 12 journals (some with two titles) in each of the 8-year periods; subsetting on DIALOG can be taken only so far. Table 4 reflects what was feasible as a by-product of our co-occurrence data. Nevertheless, little is lost, since authors' mean co-citation counts and their total counts in the 12 journals for the entire 1972–1995 period yield much the same information. The linear relationship between the two is very strong ($r = 0.84$). Moreover, a few social scientists—the Coles, Crane, Kuhn, Merton, Simon, and Zuckerman—are intercited so much as a group that their co-citation means notably overpredict their total counts in information science. If these authors are removed from the 120, the linear relationship becomes very strong indeed ($r = 0.95$), suggesting that mean co-citation counts may substitute for total counts in gauging reputations.

These mean counts should not be confused with *annual* means, that is, with each author's citation count for a period divided by the period's number of years. Rather, for each period, they are the sum of each author's co-citations with all other authors divided by 119. (The excluded case is each author's co-citation count with himself or herself, which we treat as missing data.) If an article cited Salton with 11 other authors from among our 120, that would increment Salton's total count in our framework by 11.

Note that the counts being averaged are counts of *whole articles* in which pairs of authors are both mentioned at least once. They are not counts of all co-references *within* articles. For example, an article that cites one work by Salton and one work by Croft increases their co-citation count by one, and so does an article that cites three works by Salton and five works by Croft. This is a convention of DIALOG retrieval with intersections of the Cited Author field in ISI databases. One could score authors on the basis of every co-mention, but that would entail laborious recourse to ISI's printed citation indexes rather than online searches in DIALOG.

Ultimately, the means here depend on how many times one is cited overall within the context of the other highly cited authors in information science. Salton, for example, was co-cited with each of the other 119 authors about seven times, on average, during 1972–1977, about 12 times during 1980–1987, and about 21 times during 1988–1995, for an overall average of about 13.5. This record makes him not only a very eminent figure but one

TABLE 4. Mean co-citation counts for 120 authors in information science.

Name	1972-1979	1980-1987	1988-1995	*1972-1995*	*Name (cont.)*	1972-1979	1980-1987	1988-1995	*1972-1995*
Garfield	7.38	18	18.68	*14.69*	Allen	2.02	2.03	4.45	*2.83*
Salton	7.14	12.06	21.26	*13.49*	King	2.51	4.23	1.71	*2.82*
Price	7.12	17.2	15.56	*13.29*	Vlachy	1.08	4.34	3.04	*2.82*
Lancaster	6.97	12.36	13.51	*10.95*	J. Martyn	2.4	3.49	2.36	*2.75*
Merton	8.13	12.28	11.61	*10.67*	Borko	2.13	3.08	2.91	*2.71*
Kuhn	6.53	9.13	10.73	*8.8*	Machlup	1.24	3.09	3.82	*2.71*
Saracevic	3.8	5.26	14.66	*7.91*	Leimkuhler	2.84	2.82	2.45	*2.7*
Small	3.49	8.61	11.18	*7.76*	Paisley	1.32	2.08	4.69	*2.69*
Crane	6.63	8.17	8.14	*7.65*	Ingwersen	0	1.83	6.14	*2.66*
Bookstein	3.31	8.76	10.65	*7.57*	Fenichel	0	3.86	4.08	*2.65*
Brookes	5.82	8.51	7.22	*7.18*	Fairthorne	3.18	2.44	2.33	*2.65*
Sparck Jones	3.76	7.34	10.39	*7.16*	Carpenter	1	3.18	3.61	*2.6*
van Rijsbergen	2.32	7.01	11.5	*6.94*	Svenonius	0.87	2.88	3.99	*2.58*
Robertson	3.33	7.46	9.06	*6.62*	Radecki	0.49	3.33	3.86	*2.56*
Simon	3.97	6.65	8.92	*6.51*	McCain	0	1.12	6.46	*2.53*
Narin	2.95	7.45	8.82	*6.41*	Morse	1.9	3.67	1.76	*2.45*
Belkin	0.82	4.15	14.2	*6.39*	Kessler	2.09	3.11	2.14	*2.45*
S. Cole	4.56	7.42	7.14	*6.38*	Bourne	2.63	3.44	1.21	*2.43*
Zuckerman	4.13	7.33	7.55	*6.33*	Lawani	0.5	3.32	3.44	*2.42*
J. R. Cole	4.61	7.35	6.95	*6.31*	Frame	0.7	3.13	3.39	*2.41*
Bates	0.68	5.1	12.34	*6.04*	Katzer	0.98	1.55	4.69	*2.41*
W. S. Cooper	3.34	5.05	8.69	*5.69*	Oddy	0.52	2.96	3.72	*2.4*
Line	3.65	7.52	5.4	*5.52*	Kent	1.82	3.5	1.71	*2.34*
Croft	0.45	4.62	10.16	*5.08*	Tenopir	0.01	2.46	4.44	*2.3*
M. Williams	3.29	7.22	4.06	*4.86*	Kraft	2.01	2.95	1.82	*2.26*
Griffith	3.56	5.55	4.88	*4.66*	Tagliacozzo	1.89	2.46	2.34	*2.23*
Cleverdon	4.46	4	5.34	*4.6*	Hildreth	0	1.96	4.71	*2.22*
Goffman	4.11	4.13	5.28	*4.51*	Chen	1.32	3.49	1.78	*2.2*
Moravcsik	2.1	5.45	5.82	*4.46*	Cuadra	2.49	1.55	2.48	*2.17*
Swanson	2.23	3.71	7.3	*4.41*	Braun	0.04	2.38	4.07	*2.16*
Borgman	0.34	2.7	10.13	*4.39*	Moed	0	1.28	5.18	*2.15*
Harter	1.54	4.48	6.97	*4.33*	Fox	0	0.76	5.64	*2.13*
Buckland	2.64	4.5	5.45	*4.2*	M.D. Cooper	1.9	2.87	1.61	*2.13*
Drabenstott	0	4.52	7.81	*4.11*	Schubert	0	1.7	4.67	*2.12*
C. T. Meadow	1.51	6.13	4.61	*4.08*	Kantor	0.41	3.04	2.86	*2.1*
Bradford	3.14	4.39	4.65	*4.06*	Shaw	0.48	2.2	3.51	*2.06*
Kochen	2.7	5.06	4.28	*4.01*	Rouse	0.97	2.81	2.34	*2.04*
Chubin	1.34	4.87	5.82	*4.01*	Winograd	1.19	2.01	2.92	*2.04*
L. C. Smith	0.65	5.66	5.49	*3.93*	Rice	0.03	1.19	4.87	*2.03*
Vickery	4.25	3.74	3.55	*3.85*	Hayes	1.42	2.48	2.16	*2.02*
Garvey	2.71	4.62	4.2	*3.84*	B. R. Martin	0.02	1.73	4.17	*1.97*
Cochrane	1.25	5.57	4.26	*3.69*	Bush	0.74	2.3	2.87	*1.97*
Cronin	0	3.86	7.16	*3.67*	Keen	1.58	2.45	1.75	*1.93*
R. S. Taylor	1.37	3.08	6.38	*3.61*	T. D. Wilson	0.56	1.79	3.45	*1.93*
Maron	1.96	3.53	4.97	*3.48*	Matthews	0.01	2.09	3.55	*1.88*
Marcus	1.44	5.44	3.13	*3.34*	Zunde	1.52	2.15	1.87	*1.85*
A. J. Meadows	2.64	4.55	2.8	*3.33*	Egghe	0	0.14	5.34	*1.83*
Doszkocs	0.5	4.91	4.48	*3.29*	Tague-Sutcliffe	0.35	1.76	3.33	*1.81*
Hawkins	1.66	4.86	3.36	*3.29*	Shneiderman	0.24	1.71	3.29	*1.75*
Blair	0	1.73	8.03	*3.25*	J. Martin	0.97	1.98	2.09	*1.68*
Lotka	1.66	4.13	3.92	*3.24*	Austin	1.71	1.8	1.2	*1.57*
Pao	0.25	3.25	6.12	*3.21*	O'Neill	1.02	1.38	2.28	*1.56*
P. Wilson	1.05	3.39	5.01	*3.15*	Farradane	1.76	2.02	0.77	*1.52*
Fidel	0	2.22	6.86	*3.03*	Dillon	0.27	1.44	2.63	*1.45*
Zipf	2.02	2.97	3.92	*2.97*	Kilgour	1.84	1.34	0.82	*1.33*
Yu	1.17	3.72	3.77	*2.89*	Luhn	0.94	1.39	1.6	*1.31*
Cawkell	2.45	4.13	2.02	*2.87*	McClure	0.19	1.5	2.18	*1.29*
H. D. White	0.03	3.07	5.46	*2.85*	Gorman	0.29	1.71	1.71	*1.24*
Shannon	2.01	2.66	3.82	*2.83*	Avram	1.86	1.24	0.57	*1.22*
Dervin	0.04	0.9	7.56	*2.83*	Knuth	0.78	1.08	1.08	*0.98*
MEANS	**1.92**	**4.1**	**5.29**	**3.77**	**MEDIANS**	**1.52**	**3.33**	**4.27**	**2.83**


Garfield
Price
Bookstein
Lancaster
Small
L. C. Smith
Salton
van Rijsbergen
C. T. Meadow
Drabenstott
Narin
Bates
Doszkocs
Cochrane
Croft
Merton
Robertson
Marcus
M. Williams
Line
Cronin
Fenichel
Sparck Jones
Chubin
Moravscik
Belkin
Vlachy
Zuckerman
Hawkins
H. D. White
Pao
Harter
S. Cole
Radecki
Lawani
J. R. Cole
Brookes
Simon
Kantor
Kuhn
Yu
Lotka
Tenopir
Oddy
Frame
Kochen
Borgman
Braun
P. Wilson
Fidel
Carpenter
Chen
Matthews
Svenonius
Griffith
Hildreth
Garvey
A. J. Meadows
Buckland
Machlup
Rouse
Ingwersen
Morse
Blair
Shaw
King
B. R. Martin
W. S. Cooper
R. S. Taylor
Schubert
Kent
Cawkell
Maron
Bush
Crane
Swanson
Shneiderman
Saracevic
Gorman
Tague-Sutcliffe
McClure
Moed
Bradford
T. D. Wilson
Dillon
Rice
McCain
J. Martyn
Hayes
Kessler
J. Martin
M.D. Cooper
Borko
Zipf
Kraft
Keen
Dervin
Winograd
Bourne
Fox
Paisley
Shannon
Zunde
Katzer
Tagliacozzo
Luhn
O'Neill
Knuth
Farradane
Egghe
Austin
Goffman
Allen
Leimkuhler
Cleverdon
Kilgour
Vickery
Avram
Fairthorne
Cuadra
Comparing 1980-87 to 1972-79
Comparing 1988-95 to 1980-87
-4
-2
0
2
4
6
8
10
12
14
16

whose influence continued to increase during the period of our study.

It should be borne in mind, of course, that the co-citation counts of everyone in the matrix, including Salton, are distributed neither uniformly nor normally. In the familiar pattern, they exhibit high positive skew: Authors are co-cited often with relatively few others and seldom or never with many others (there are many zeros in the cells). Thus, their means are pulled upward by untypical values at the high end of the skew, and in that sense they misrepresent the actual distributions of the counts.

Nevertheless, we think the rank-ordering of means provides a plausible list of nominees for the 15 or 20 most distinguished contributors to information science in the contemporary period. (The substitution of medians for means produces a list with many tied authors, including some tied at zero when their means are non-zero, and thus is much less discriminating.) We expect most readers would intuitively supply much the same list if a poll were made, which is an essential test of an indicator. The only further consideration in the ranking is whether one would want to include figures such as Merton, Kuhn, Crane, and Simon, who are not primarily information scientists.

We shall limit our comments here to the top four names in the listing. All seem appropriately singled out. Garfield, although a somewhat special case because of his ability to cite himself in his weekly column in *Current Contents*, is unquestionably a major figure because of his founding of ISI and his voluminous inquiries into aspects of citation analysis, bibliometrics, scientific communication, and information retrieval. Salton and Price are the two names most likely to be eponymic in recent disciplinary history. Thus, one could speak of the ''Saltonian period'' of information retrieval, roughly 1965–1995, and everyone in the field would know what was meant. (Croft, 1996, uses ''Post-Salton'' to describe IR's present phase.) In the same way, one might cogently designate the years 1960–1983 as the ''Pricean period'' of bibliometrics. Finally, although we have not heard ''Lancastrean'' used to name the last quarter-century of L&IS, Lancaster *is* the author whose citation record most firmly links information science with library science, and probably the single best representative of the union of the academic discipline with professional practice.

We conclude that means like those in Table 4 are an acceptable way of assessing reputations. Authors can readily be compared by decile, quintile, and so on. The data may be gathered unobtrusively online, and they are obtainable for far more authors than could reasonably be covered in a poll.

Changes in Influence

Here we shall construe *eminence* as *influence* and treat changes in authors' mean co-citation counts as indicators of changes in the latter over time. We are well aware that the means are an imperfect indicator. Citedness is not the only form of influence, and, even if it were, our data do not capture the quality of individual citations; they do not, for example, distinguish between self-citation and citation by others, or between perfunctory and nonperfunctory references to an author's work. Nevertheless, we think that rendering abstract influence as concrete citedness is an operationalization worth retaining, in that it allows one to give evidence for judgments that might otherwise rest solely on intuition.

Across the three 8-year periods in Table 4, each author's means fall into a few standard patterns. One of the most common is for the mean to go *up* and then *up* again. The questions in that case are, how large are the increases and from how high a base? Another common pattern is *up* and then *down.* This is a pattern frequently observed when tracing the annual citation counts of papers, which sooner or later reach a peak and then decline (Aversa, 1985). Here, the *up–down* pattern is exhibited with oeuvres rather than papers and may betoken the peaking of an author's influence. (Those whose pattern here is *up* and *up* simply have not had this happen to them *yet.*) Rarer patterns are *down–up* and *down–down.* Occasionally, no change from one period to the next is seen.

By taking differences in means, we can show changes in influence over our time periods. Figure 1, a stacked bargraph, illustrates. For all authors in Table 4, two differences were taken: The first between their mean counts in the early and the middle periods, and the second between their mean counts in the middle and the late periods. Light gray bars depict the first difference; dark gray, the second. The longer the bar, the greater the difference. Authors have been ranked high to low by the first difference (which was positive for almost everyone), allowing the second to play freely. Note that the nonabsolute zero-point in Figure 1 merely denotes *no difference in means*, whereas the absolute zero-point in Table 4 denotes *absence of co-citation.*

Dark gray bars to the right of the light gray ones indicate the pattern *up* and *up,* characteristic of authors with late-period gains. For example, the top-ranked author, Garfield, had a very sizeable gain from the early to the middle period, but only a small gain from the middle to the late. We surmise that his reputation is large but fairly static.

Dark gray bars to the left of the zero-point indicate the pattern *up* and *down*, characteristic of authors whose citedness may have peaked during the middle period. For example, Price had the next largest initial gain, but his mean count decreases when the second difference is taken. A brilliant intellect, he probably would have remained highly influential had he not died in 1983.

← FIG. 1. Differences in mean co-citation of 120 authors over two periods.

The remaining patterns are seen mostly at the bottom of Figure 1. Absence of a bar (of either shade) indicates *no difference* between given periods (Knuth, Allen, and Leimkuhler furnish examples). When both light and dark gray bars are to the left of the zero-point, the authors' means went *down* and *down*, a sign of declining influence over the 24 years of our study. Cleverdon and Cuadra, on the other hand, exhibited the unusual pattern *down* and *up;* obviously, one must not rule out the possibility of revivals that restore authors to earlier levels or higher.

Of prime interest here are means that shot up for the period 1988–1995, since authors who had very marked gains in this period are often those who, on other grounds, have become prominent in the field in the last decade or so. Figure 1 makes them stand out dramatically. In descending order of gains, they are Belkin, Saracevic, Salton, Borgman, Bates, Dervin, Blair, and Croft, all primarily retrievalists. A second group, more of a mixed bag, includes McCain, Egghe, Fox, Fidel, van Rijsbergen, Ingwersen, Moed, and Rice. Saracevic, Salton, and van Rijsbergen, of course, were well established in the 1970s, but the others did not really come into their own until the 1980s and 1990s.

Historians may want to know not only who emerged but *what specialties produced them,* and it is instructive to relate these 16 names back to specialties in which they load strongly, as shown in Table 3. Thus, we find Croft, van Rijsbergen, Fox, Blair, Salton, Belkin, and Saracevic all representing *experimental retrieval,* which clearly is a viable area in all periods. Borgman, Fidel, and Bates represent *practical retrieval,* and Dervin and Ingwersen represent *user theory.* But note the contingent, retrievalists all, who also load on *user theory:* Belkin, Saracevic, Borgman, Fidel, and Bates. (Blair, 1990, shows that he, too, is interested in this area.) We regard this as evidence that the cognitive side of information science, which some retrievalists tend to ignore in favor of algorithms, has emerged during the 1980s and 1990s as a major enterprise.

Of the remaining emergent authors, McCain, Moed, and Rice represent a new unfolding of *citation analysis,* and Egghe, of mathematical *bibliometrics.* We stop arbitrarily with these 16, leaving readers to trace eminence and influence further with the ample data in Table 4 and Figure 1.

The Maps

Figures 2 through 4 are our 8-year period maps. We shall use them to explore the idea, introduced earlier, of *two subdisciplines* in information science. We operationalize this idea as the last two clusters joined in a complete-linkage clustering of 100 authors. These final clusters, which are brought together only after all closer ties have been exhausted, are separated by an angled line superimposed on each map.

We have not, as in the past, drawn lines around smaller clusters of authors corresponding to their specialties. The crowding of many names on the maps makes this difficult, and, besides, the specialties are better conveyed by the factor analysis of the earlier section. To a great extent, however, the authors forming specialties in the factor analysis will be found to have been placed near each other in the maps.

The space defined by author coordinates varies from map to map, and axes are not necessarily symmetrical. A small circle has been placed at each map's origin as a guide to the presence or absence of axial symmetry. The axes in the maps have not been labeled, but their interpretation will be discussed in a later section.

Two-dimensional solutions in placing the author-points are in all cases quite powerful. Bear in mind that these solutions are for 100 cases each—considerably more than we have seen in other reported uses of multidimensional scaling. The R^2 goodness-of-fit values for two dimensions are 0.87 (early period), 0.87 (middle period), and 0.91 (late period; respective values for Kruskal's Stress 1 measure are 0.18, 0.17, and 0.14).

As things turn out, information science looks rather like Australia: Heavily coastal in its development, with a sparsely settled interior.

We refer, of course, to the pronounced division between the two main groups of authors and the relative absence of figures useful to both. Three independent pictures of information science yield essentially the same result. Whether we examine the maps for 1972–1979, 1980–1987, or 1988–1995, those interested in aspects of retrieval form a large cluster at right, and those interested in aspects of literatures and communication form another at left. More specifically, the experimental retrievalists are the upper group at right; the practical retrievalists, writers on OPACs, and user theorists are the lower. General library systems theorists also appear on the right side, tending toward the center. At left, the citationists are above, and the writers on scientific communication are below. Across the top is a scattering of bibliometricians, while across the bottom are extradisciplinary figures from whom ideas are imported. The central areas always have something in common with what, in the early days of Australian exploration, was called ''the ghastly blank.''

The first finding to note is the overall stability of information science, as here defined. Some author-points undergo remarkable changes of position from map to map, but many more authors stay put in discernible specialties. Fully 75, moreover, persist through all three maps. In general pattern, and in many specific details, author placements resemble those in the original map of information science in White and Griffith (1981).

We conclude that author co-citation analysis is useful for rendering the *inertia* of fields. In other words, it objectively captures the slow-changing divisions on which

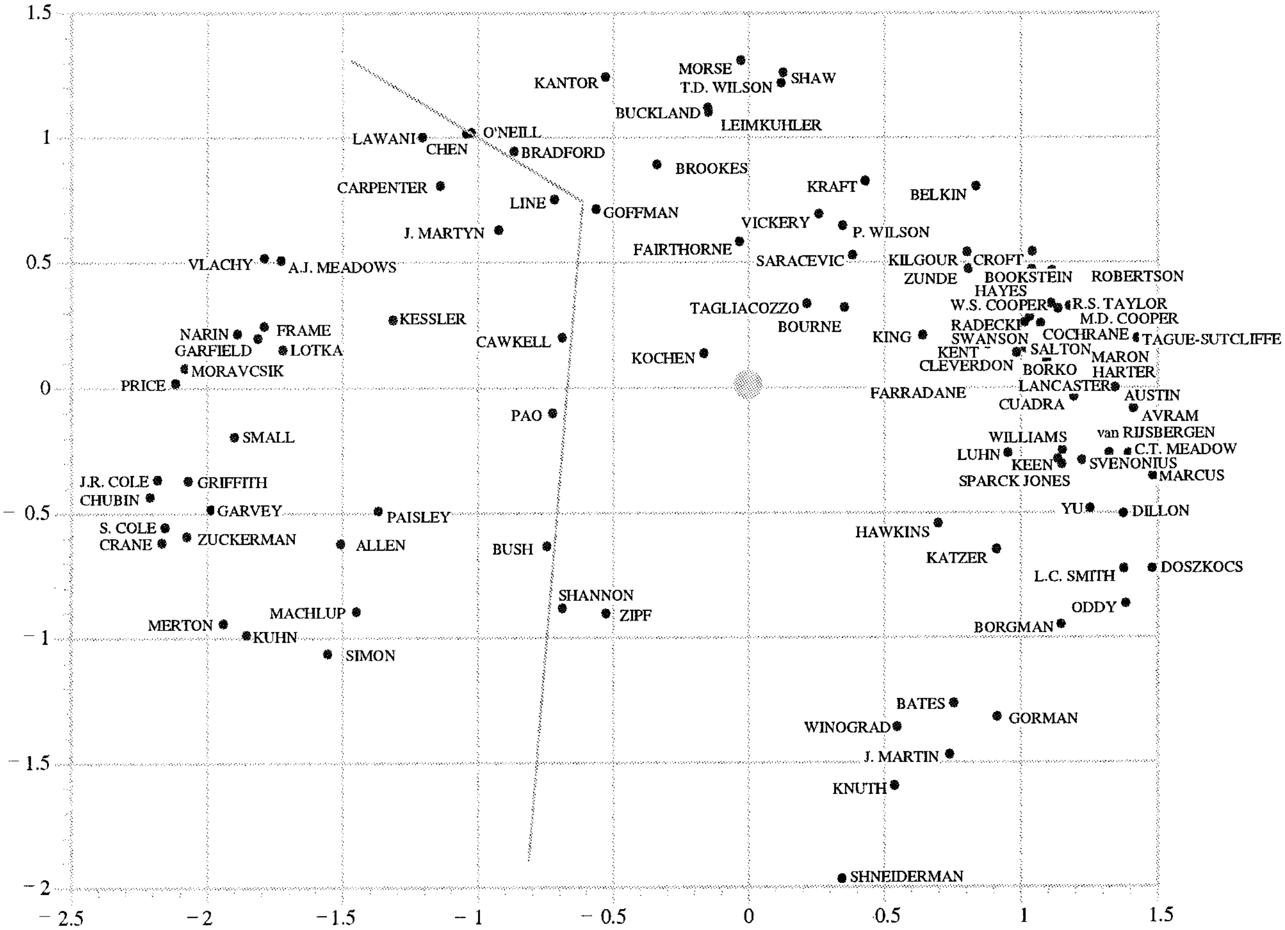


FIG. 2. Top 100 authors in information science, 1972–1979.

one's subjective sense of ''semi-permanent'' disciplinary structure rests. (This sense may well derive from long-term exposure to samples of the same data that go into the maps, that is, it is a by-product of ''monitoring the literature,'' which includes not only browsing and reading, but attendance at conference presentations.) Co-citation analysis of *papers*, as opposed to authors, captures disciplinary history at a different, faster rate, which may better suit fields with livelier research fronts than information science.

That a field exhibits considerable inertia may seem a truism when it is pointed out: After all, one thinks of disciplinary subject-matter as persisting through time, and oeuvres embody subject-matter. The specific configurations of Figures 2 through 4 may also seem predictable, once seen. However, *ready intelligibility* is hardly the same as *inevitable outcome*. It is conceivable, for example, that leading figures like Price and Salton might have been at the center of each map, closely surrounded by authors, such as Harter, Vickery, Borgman, and Griffith, who have published both as citationists and retrievalists. (Are the likes of Price and Salton not often said to be *central* to the discipline? And are citers not gradually citing more and more information scientists and commingling more and more ideas?) Nor is it inconceivable that specialties might resist easy identification because of their volatility. As it turns out, specialties are both easy to identify and durable. But that, we think, was not a foregone conclusion for information science, and it remains to be corroborated in most other fields. Moreover, Price, Salton, and other leaders turn out to be central to their specialties, but not to the field as a whole.

Thus, our second major finding, already conveyed, is that information science lacks a strong central author, or group of authors, whose work orients the work of others across the board. The field consists of several specialties around a weak center. While many other fields might also consist of loosely linked specialties, we think that at least some are dominated by authors who would be placed near the center of ACA maps. Technically, in order to be central, an author's co-citation profile must correlate highly with those of many other authors in the matrix. The profile that best matches this description over the full 24 years is Kochen's, whose writings and influence are discussed at length in Lancaster, Bushur, & Low, (1993). But Kochen seems to us more a multifaceted talent than

TABLE 5. Counts of authors in subdisciplinary clusters.

Cluster	Period 1972–1979	1980–1987	1988–1995
Left side	30	41	49
Right side	70	59	51
Total	100	100	100

a major integrator of the field. The same is true of the other figures in the central areas; whatever their merits, we cannot find evidence that they dominate work across both halves of the maps.

That evidence in our scheme would be an extraordinarily high average co-citation count, like Salton's or Price's. To be a *strong* central figure, in other words, it is not enough for an author's profile to correlate with those of others in all directions; the raw co-citation counts in the profile must be high as well. No one in information science has both attributes, though, of course, having even *one* is an accomplishment worth the historian's attention. (Bush floats not too far from the center of our maps because his memex still intrigues citers on all sides. But his raw counts are not high, perhaps because L&IS has pretty well assimilated his prophecies.)

Our third major finding is the changing composition of the two subdisciplines of information science. They alter rather markedly from the early period to the late. Table 5 has counts of the authors on either side of the line over the three periods, as shown in Figures 2 through 4.

Obviously, the authors to the left increase as the years go by, with the 1972–1979 period having the least balanced counts between sides. Note, too, that the map for 1972–1979 is asymmetrical. All the retrievalists fit on a horizontal axis that extends only 1.5 units to the right, whereas it must extend 2.5 units to the left to fit all the authors in citation research and scientific communication. The latter, fully 21 of whose points lie between −1.5 and −2.5, are less strongly correlated with the rest of the field in the early period than in later ones. (The vertical axis is also asymmetrical, to accommodate Knuth and Shneiderman from computer science.) In contrast, the maps for 1980–1987 and 1988–1995 are symmetrical in both axes.

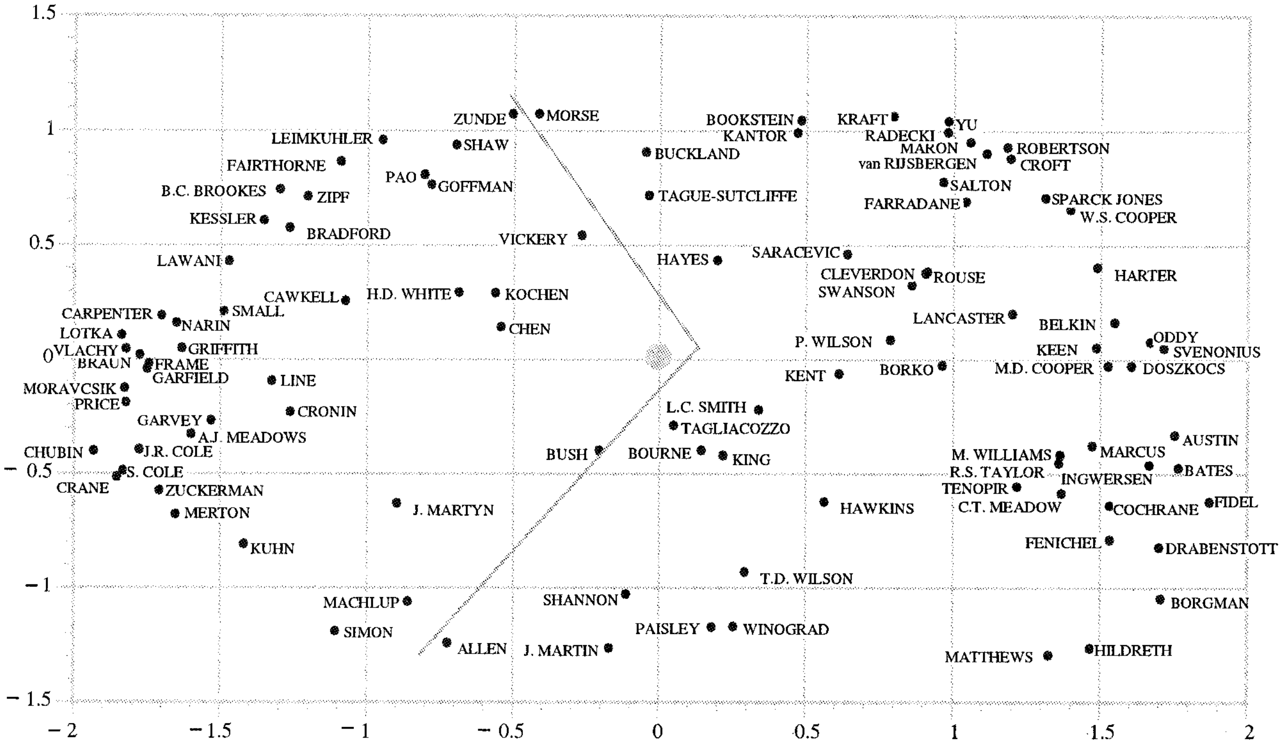
In the 1972–1979 map, the line runs between the overlapping points for Chen (left) and O'Neill (right) and separates Cawkell, Pao, and Bush from Shannon and Zipf. Thus, the huge majority on the right includes not only the densely clustered experimental retrievalists—the computer can hardly tell them apart—but the operations researchers, the library automators, the online search specialists, the early user theorists, the computer scientists, and, most strikingly, the Bradford-Zipf bibliometricians. According to the complete-linkage clustering routine, all are more closely related to each other than to the citationists and science communication authors, who are relatively far out at left.

The 1980–1987 map in Figure 3 shows three important changes taking place. One is the breakup of the retrievalists, from a large, undifferentiated mass into something like the specialties of Table 3. Another is the increasing number, and more even distribution, of authors across the midregion of the map, suggesting that citers are reading and finding connections between authors on both sides. (The new symmetry of the axes may result from this greater integration.) The most interesting is the shift of 10 bibliometricians, a coherent group including Bradford and Zipf, to the left of the dividing line, as citers connect them to the citationists and other writers on scientific communication. It is largely this shift, rather than the appearance of three newcomers, that increases the number of author-points at left.

All told, the left side has gained 13 authors and lost two, for a net gain of 11 (from 30 to 41), while the right side has gained nine authors and lost 20, for a net loss of 11 (from 70 to 59). Figure 3 includes the details. ''Entering'' authors are those who, not having made the top 100 in 1972–1979, appear in 1980–1987 for the first time. ''Exiting'' authors made it in 1972–1979 but not in 1980–1987. Authors who changed subdisciplines in the second period because of citers' revaluations are said to be ''shifted from right'' or ''from left'' of the dividing line.

A glance at the names of the seven authors entering on the right side shows the type of retrievalists who were rising to prominence in the 1980s. They are not algorithmists but people interested in practical retrieval, OPAC design, and linguistic or cognitive questions. It looks very much as if the discipline were rewarding researchers who focused on what Persson (1994) calls the ''soft'' side of retrieval. This is a lead that an intellectual historian might pursue.

Another lead is that, in the 1980s, the subdiscipline of information science at left appears to be strengthening its identity. The correlates are a journal of its own (*Scientometrics*), growing bodies of research in several other journals, a new professional association, and a series of international conferences. The name we and many others used for it was ''bibliometrics''—cf. Persson (1994), quoted below—though that, we grant, is probably too narrow a designation. In a review of its literature, we wrote, ''Including citation analysis, bibliometrics constitutes about half of information science—arguably the more vigorous half, given the continuing isolation of theoretical information retrieval from the database publishers, indexers, vendors, and users of the real world'' (White and McCain, 1989, p. 120). The mild gibe about theoretical information retrieval was aimed at the tradition of ''document testbed'' research summed up in such works as Sparck Jones (1981). As the present evidence shows, other types of retrievalists were prominent in the 1980s,



Left Side

Entering (3): Braun, Cronin, White

Exiting (0)

Shifted from right (10): Bradford, Brookes, Fairthorne, Goffman, Kochen, Leimkuhler, Shaw, Vickery, Zipf, Zunde

Right Side

Entering (7): Drabenstott, Fenichel, Fidel, Hildreth, Ingwersen, Matthews, Tenopir

Exiting (10): Avram, Cuadra, Dillon, Gorman, Katzer, Kilgour, Knuth, Luhn, O'Neill, Shneiderman

Shifted from left (2): Allen, Paisley

FIG. 3. Top 100 authors in information science, 1980–1987.

and so the gibe oversimplified. But the final clusters in Figures 3 and 4 support our claim that the ''bibliometric'' side of information science constitutes about half the discipline, even if we need a more inclusive name to capture its complexity.

What should the left side be called? In INSPEC, a commercial database covering the information science literature, bibliometric and citation studies are indexed under ''information analysis,'' and that might be taken as a new name for the subdiscipline. However, ''domain analysis,'' as put forward by Hjørland and Albrechtsen (1995), seems a more appropriate choice. It incorporates citation analysis and bibliometrics, but also a range of topics broader than what ''bibliometrics'' usually implies—for example, scholarly and professional communication, parts of sociology of science and sociology of knowledge, interdisciplinary linkages, discourse communities, and disciplinary vocabularies (cf. Beghtol, 1995). Thus, it encompasses all or most of the authors to the left of the line in Figures 2 through 4. We shall therefore use ''domain analysts'' as their generic name, parallel to ''retrievalists.'' The retrievalists and the domain analysts correspond approximately to Wilson's (1983) second and third categories of researchers in ''bibliographical R&D.''

Amending White and McCain (1989), we would now subsume about half of information science under ''domain analysis,'' including contributors to the ''sociology of texts'' proposed by McKenzie (1986) and endorsed by Wilson (1996). The author-names supplied for this half by ACA add some sociological weight to McKenzie's proposal, which is otherwise that of a humanist bibliographer.

Hjørland and Albrechtsen (1995) give examples of authors who are domain analysts and examples of authors who are not. Encouragingly, ACA places several of the former (e.g., Garvey, Griffith, White) to the left of the line in Figures 2 through 4, and several of the latter (e.g., Belkin, Dervin, Taylor) to the right, thus corroborating their judgment. Even so, ''domain analysis'' does not exclusively categorize the work of the authors at left, any more than ''information retrieval'' exclusively categorizes the work of the authors at right. Since some retrievalists contribute to the literatures of domain analysis, and some domain analysts to the literatures of information retrieval, the two headings merely indicate what citers choose to highlight in authors' total oeuvres; they suggest relative emphases. Among the retrievalists, one may further distinguish *commentators* on domain analysis (e.g., P. Wilson, Taylor) from occasional *practitioners* (e.g., Borgman, Bates).

Before turning to the map for 1988–1995, we must show how our results bear on claims in two other works, Persson (1994) and Vickery and Vickery (1987). Using data from *Social Sciences Citation Index* on CD-ROM, Persson mapped information science on the basis of the works most cited in *JASIS* during 1986–1990. He was led to conclude:

> The tradition of information science has two major branches, bibliometrics and information retrieval. Within the bibliometrics subfield there is one group of authors primarily related to citation analysis and another group associated with bibliometric distributions. The information-retrieval cluster, which [in his *JASIS* map] is twice as large as the bibliometric group, seems to be subdivided in one ''hard'' part working on algorithms and one ''soft'' part concentrating on the user–system relation. . . . The map of information science made in the early 1980s by White and Griffith has many of the names and relative positions in common with the map produced in this study, yet the two maps are based on quite different methodologies.

He had earlier drawn a similar contrast:

> The main difference is that the IR subfield is more pronounced in the *JASIS* map [than in White and Griffith], which in turn is a reflection of both editorial policies and the role *JASIS* is perceived to play by active researchers in the field. For example, if articles from *Journal of Documentation*, *Scientometrics,* and *Social Studies of Science* had been added, the resemblance with the bibliometric subcluster [in White and Griffith] would probably become larger and researchers that have studied scientific communication would eventually show up. On the other hand, the difference between our map and the one produced by White and Griffith may depend on the time scale, indicating a real change in which the classical studies of scientific communication have less influence today than 10 years ago.

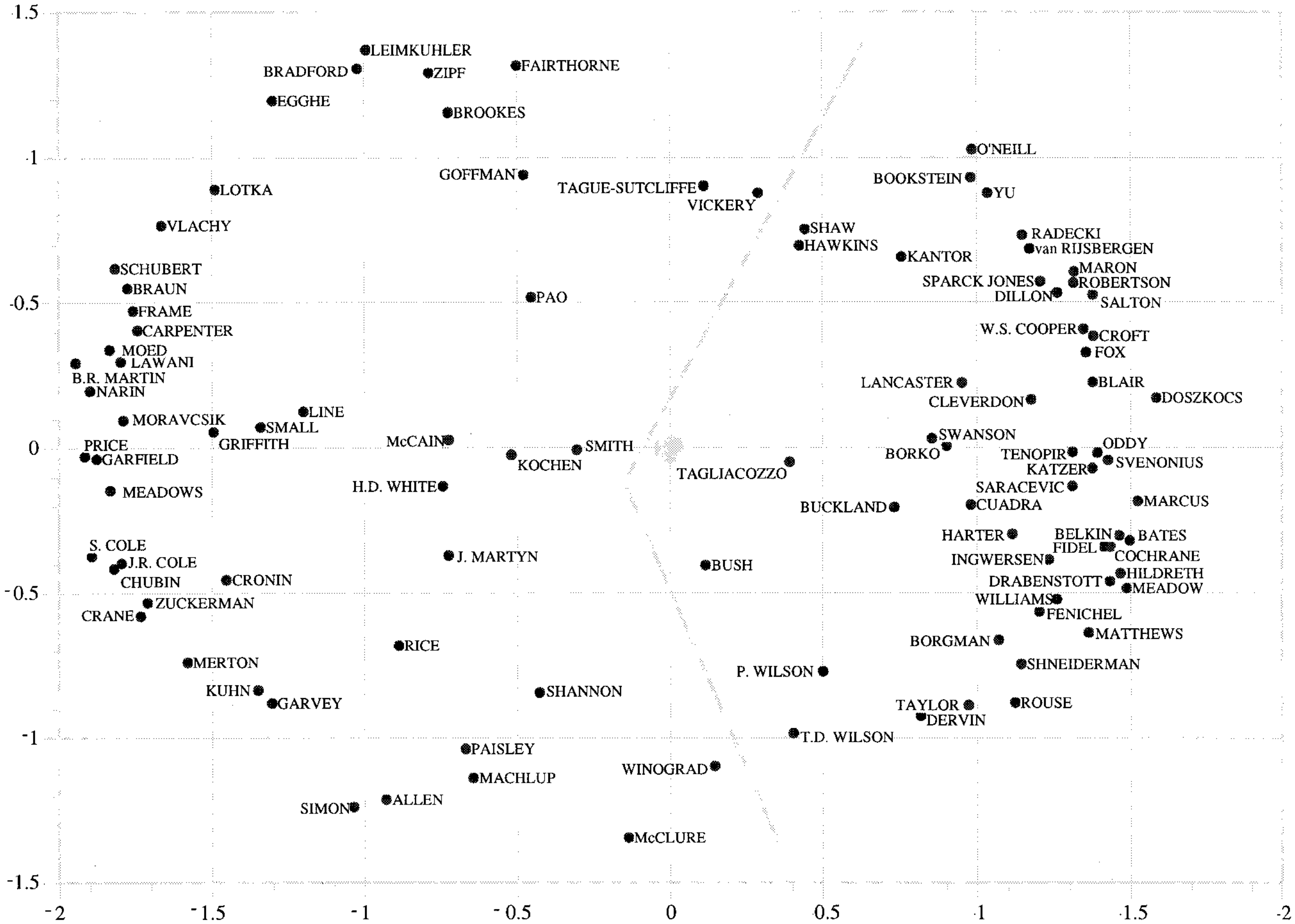
The data of the present study augment those in White and Griffith (1981) and resemble those envisioned by Persson. On that basis, we can say he is correct at every point in these two passages; despite the different methodologies, we are describing the same elephant. We agree with his claim about the main divisions of the field, including the ''hard'' and ''soft'' parts of IR, which appear on our maps much as on his. We added journals along lines he suggests and got a substantially enlarged ''bibliometric subcluster'' that includes the classic authors on scientific communication as in White and Griffith (1981). These latter persist through the entire period 1972–1995, but our Table 4 shows most of them to have already peaked in influence, which squares with Persson's expectation. Taking him as a domain expert, we show here that we can algorithmically reproduce part of his knowledge about authors in information science.

There is yet another opportunity to check our present maps against expert judgment. When Vickery and Vickery (1987), a textbook, reproduced White and Griffith's (1981) map of 39 information science authors, the Vickerys added seven other well-known names, placing them intuitively in the subject clusters but otherwise leaving the original unchanged. Four of the seven made our top 100 and so are mapped here. Despite the large increase in authors, it will be seen that, just as the Vickerys predicted, Cyril Cleverdon appears with the experimental retrievalists, A. J. Meadows with the writers on scientific communication, and R. S. Taylor not too far from the generalists (e.g., Saracevic; P. Wilson). The actual points for these authors conform reasonably well to the Vickerys' guesses. Philip M. Morse also appears more or less where they predicted (top center), but drawn to bibliometricians and library operations researchers (e.g., Buckland, Kantor) rather than the ''systems'' retrievalists of 1981. In analyses not shown here, we mapped the Vickerys' other three choices in the 1981 context, and we can assure readers that Colin Cherry does indeed appear with Shannon, Everett Rogers with the communication writers (e.g., Allen), and S. R. Ranganathan with the generalists, where the Vickerys placed them.

ACA's confirmation of expert judgments by Hjørland and Albrechtsen, Persson, and the Vickerys is consistent with the claim that citation databases can be exploited for non-experts in a form of AI.

To return to the period 1988–1995, the map in Figure 4 exhibits no dramatic breaks, rather, a continuation of the tendencies that emerged in the map for 1980–1987. The two subdisciplines, domain analysis and information retrieval, are now almost evenly balanced as clusters. The former has gained 13 authors and lost five, for a net gain of eight (from 41 to 49); the latter has gained or regained nine authors and lost 17, for a net loss of eight (from 59 to 51). Again, Figure 4 has the details.

During 1980–1987 and 1988–1995, most of the 10 newcomers entering domain analysis have been citationists of one sort or another (Braun, Cronin, McCain, B. R. Martin, Moed, Rice, Schubert, White). Mathemati-



Left Side

Entering (7): Egghe, McCain, McClure, B.R. Martin, Moed, Rice, Schubert

Exiting (4): Cawkell, Chen, Kessler, Zunde

Shifted from right (6): Allen, Paisley, Shannon, Smith, Tague-Sutcliffe, Winograd

Right Side

Entering (3): Blair, Dervin, Fox

Exiting (11): Austin, Bourne, M.D. Cooper, Farradane, Hayes, Keen, Kent, King, Kraft, J. Martin, Morse

Shifted from left (1): Shaw

Re-entering (5): Cuadra, Dillon, Katzer, O'Neill, Shneiderman

FIG. 4. Top 100 authors in information science, 1988–1995.

cal bibliometrics, as represented by Egghe, is a more rarified specialty, with fewer writers and citers. On the information retrieval side, all the newcomers have been ''user–system relation'' people, except for Fox, a computer scientist who studied with Salton. (Of course, for many computer scientists, even retrievalists with degrees in computer science are too concerned with the messy ''user–system relation'' to be mainstream; cf. Croft, 1996.)

There is some evidence of integrative forces at work in the recent period. The interior of the map continues to be populated, if thinly, by authors known to be interested in both sides of the discipline, e.g., Vickery, Pao, P. Wilson, Swanson, Buckland. Tague-Sutcliffe and Smith actually connect the sides by crossing the boundary from right to left; Shaw returns from left to right. The axes on which all 100 authors are distributed have the same extensions as before, which argues that the discipline is not flying apart.

The specialties of information retrieval seem to be coalescing, although they remain recognizable, in terms of leading authors, down the right side. Domain analysis at left is more diffuse, but it, too, retains recognizable specialties.

Nevertheless, there are questions about some of the automatic placements. Most have to do with writers along the bottom, where information science shades into other fields. The labels for the subdisciplines and specialties, while broadly accurate, vary considerably in their fit to particular authors. Why, for example, do Allen and Paisley swing across the line into the retrievalists' camp in 1980–1987? Why do Shannon and Winograd appear among the domain analysts in 1988–1995? Is Shannon really a domain analyst at all? And so on.

Coarse-meshed ACA yields only partial answers to such questions. In 1980–1987, Allen is strongly correlated only with Paisley and moves rightward with the latter because of Paisley's perceived relevance to user theorists on the retrieval side. In 1988–1995, both Shannon and Winograd continue to be used in philosophic attempts to characterize ''information'' and ''information science,'' and this increasingly links them with other authors used for this purpose—Kuhn, Simon, Merton, and, especially, Machlup.

To find out more, one must use the maps as leads for looking up the co-citation histories of individual works (or passages). For example, the continuing presence of Allen, Paisley, and Borgman and the emergence of Rice along the bottom of the maps serves as a reminder of the overlap between information science and communication studies (cf. Vakkari, 1996). Thus, it may only be Shannon's well-known model of a communication channel, and not full-scale information theory, that is being cited in analyses of information science as an intellectual domain.

Changes in Image

ACA can highlight the careers of individual authors as perceived by citers (cf. McCain, 1986). Figure 5 reveals changes in the citation images for selected oeuvres over the three 8-year periods. Aside from its substantive interest, it points toward a capability for tracking author images in any field; one can foresee personal citation histories done as animated color graphics (perhaps less chaotic-looking than Fig. 5).

The 75 authors who appeared in all three maps above were remapped separately for 1972–1979, 1980–1987, and 1988–1995. Their points for each of the three periods were then plotted in a common space. From the full set of 225 (75×3) points, Figure 5 presents those for 19 authors whom citers repositioned markedly. Arrows trace their placements early to late. Authors whose points migrated relatively little are excluded, e.g., Price, who always anchors the horizontal axis at mid-left. (Immobility may be quite as interesting as movement, of course.) A sketch of the movers recapitulates some earlier findings and adds a few new ones:

Five bibliometricians—Bradford, Brookes, Fairthorne, Leimkuhler, and Vickery—move left into the region of domain analysis in the 1980s and then up and right in the 1990s, as if their ties with the citationists and others there were weakening. Zipf, who starts in the lower central region (because of a strong link to Shannon), shoots up in the 1980s to join bibliometrics, as his name is increasingly conjoined with Bradford's. Vickery remains closest to the retrievalists. Shaw, like Vickery, a bridging author, first moves left with the bibliometricians and then right as if rejoining the retrievalists. In contrast, Tague-Sutcliffe moves from a position with the ''hard'' retrievalists to one much closer to the bibliometricians (cf. Tague-Sutcliffe, 1981). Also, Smith begins with the ''soft'' retrievalists and crosses the map in the direction of the citationists, largely because of her review article (Smith, 1981).

At a 1993 ASIS meeting, Kantor told the present authors that he was not fond of bibliometrics, the specialty in which some of his earlier work was done (e.g., Kantor, 1981). His point is seen traversing the upper region, which includes not only bibliometrics but library operations research, another of his specialties, to end with those of presumably more congenial retrievalists (cf. Kantor, 1994).

T. D. Wilson and Buckland are other authors who start in the general area of bibliometrics and library operations research. T. D. Wilson moves sharply downward to join user studies. Buckland's final point shows him recognized as a generalist with interests in retrieval, users, library systems, and OPACs. Saracevic, another generalist, has a similar career. Likewise P. Wilson, first identified with retrieval (e.g., clarification of concepts such as ''relevance'') but then with a broader range of user and policy issues. Belkin's trajectory is from relatively ''hard'' to relatively ''soft'' retrieval, again reflecting the field's growing interest in user psychology. (Recall also that Belkin gained the most in mean co-citation in 1988–1995, which accords with the attention his writings on cognition receive in, e.g., Olaisen et al., 1996.)

Hawkins published at least four articles in practical bibliometrics during 1976–1980. His elevator-like rise into that region shows that he is not much co-cited with bibliometricians until 1988–1995, when the names Bradford, Brookes, Leimkuhler, and Vickery crop up with his in reference lists. Martyn's descent traces the opposite course: Starting with ties to British bibliometricians and operations researchers, he is progressively co-cited with user theorists and citationists (cf. Martyn, 1974, 1975), emerging as a counterpart on his side to Buckland and T. D. Wilson on theirs.

Although Borgman is plainly a retrievalist, her point moved left in 1988–1995. We thought this might be because her publications in bibliometrics (e.g., Borgman, 1990) were co-cited with works by citationists. This is not so. A look at the authors with whom she is frequently co-cited reveals that she is simply being pulled toward user theorists such as Dervin. Paisley's point, too, was pulled toward user studies when his works were linked with those of user-oriented retrievalists (e.g., Belkin, Oddy, Taylor, T. D. Wilson) in 1980–1987. In 1988–

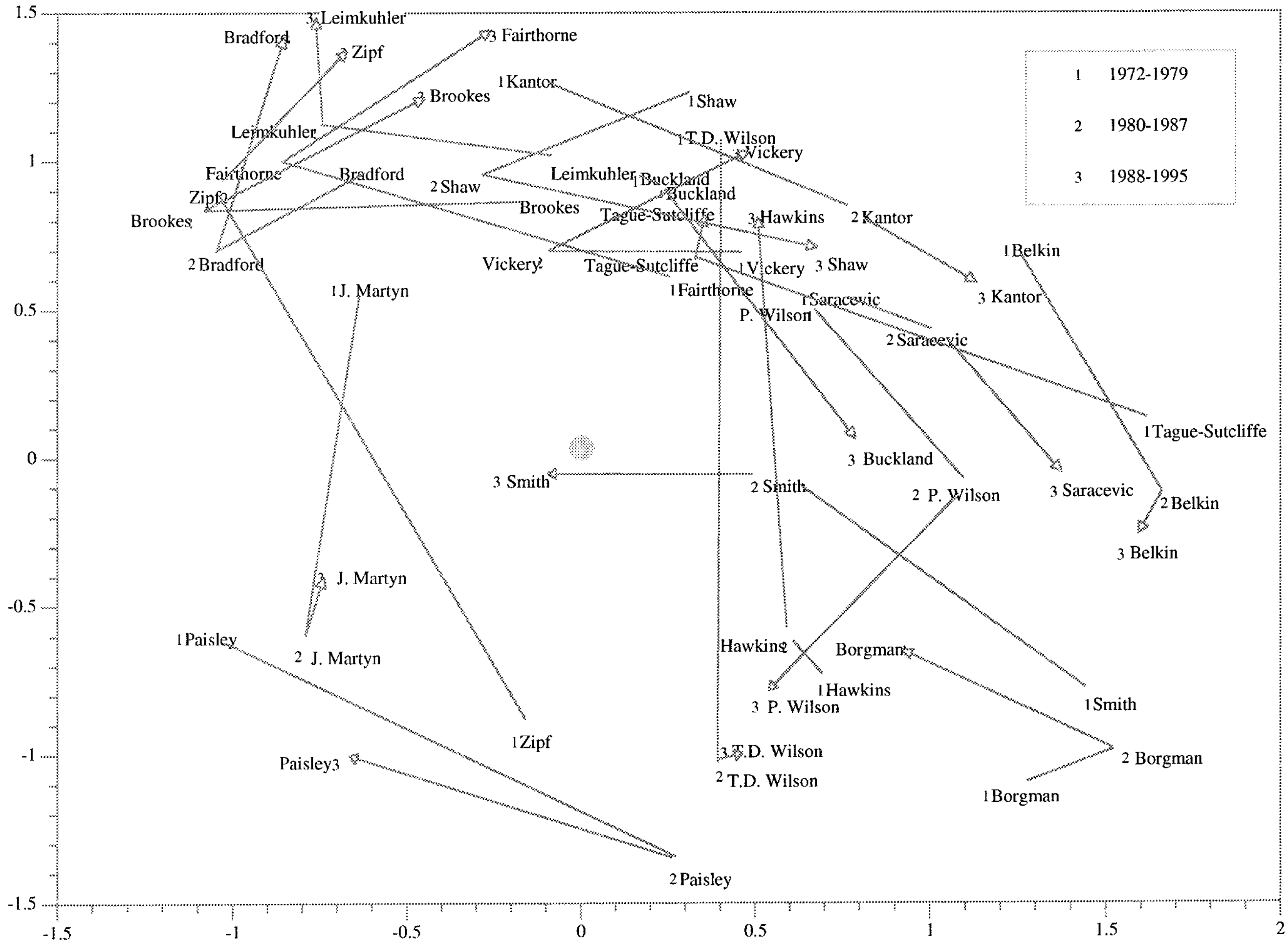


FIG. 5. Shifts in 19 authors' citation images.

1995, it swung back, owing to his new linkages with counterbalancing domain analysts (e.g., Crane, Garfield, Rice, White).

If authors are co-cited with several distinct companies of names over time, their points will move markedly. If they are co-cited with more or less the same company—that is, some new names but many old names as well—their points will move relatively little. Authors can create co-citation contexts for themselves in their own works, of course, but otherwise the matter is largely out of their control. For example, late in his career, the bibliometrician B. C. Brookes devoted several articles to a cognitive theory that Vakkari (1996) links with those of Belkin and Ingwersen. Nevertheless, Brookes is still far from those two on our maps. Or, again, note how citers have moved Tague-Sutcliffe into bibliometrics, but not Borgman, whose work in retrieval still receives much more attention.

Data like these even now yield clues to the reception of works and the transit of ideas—clues that can be immediately pursued to more specific levels through online searching with author names. In the future, computers may display such data as ''literature processes'' that appear one way when authors' images are marked by continuity and another when they are marked by change.

Canonical Authors

The 75 authors who remain in the top 100 for the entire period 1972–1995 have a fair claim to be considered canonical for information science. They are by no means the only possible candidates, but their case is strengthened by long and broad records of citedness. The step not taken here is to identify their most cited works in search of their most highly valued ideas.

A composite view of their positioning for the full 24 years is given in the INDSCAL map in Figure 6. It synthesizes the information in the three separate 75 × 75 correlation matrices for 1972–1979, 1980–1987, and 1988–1995. The R^2s for the two-dimensional solutions are, respectively, 0.80, 0.84, and 0.84; the average R^2 for the composite is 0.82 (average stress is 0.20). In other words, a solution in only two dimensions explains more than 80% of the variance in the author placements.

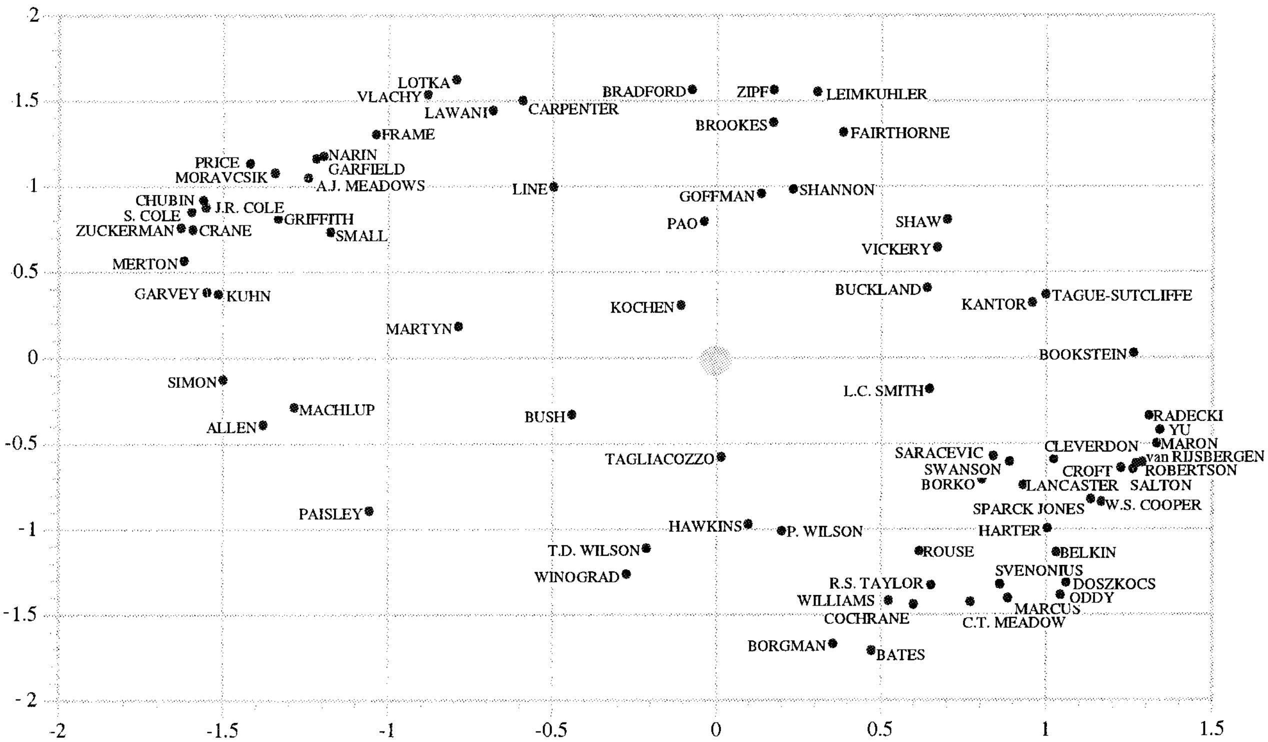


FIG. 6. INDSCAL map of 75 ''canonical'' information science authors.

The axes in INDSCAL maps are not subject to rotation and are supposed to be maximally interpretable. Thus prompted, we think the horizontal axis conveys, as in past studies, the *range of subject specialties* within the subdisciplines of domain analysis and information retrieval. This axis is asymmetrical, with the sociologists of science being somewhat farther out at left than their extreme opposites, ''arch-retrievalists'' like Radecki and Yu, at right. Coherent groups from left include the citationists, the arc of bibliometricians across the top and the philosophically orienting figures across the bottom, ''generalist'' writers such as Smith, Wilson, Saracevic, and Swanson, and the hard and soft retrievalists. The plot generally makes good sense. For example, it is easy to accept Bookstein, Tague-Sutcliffe, Kantor, Buckland, Vickery, and Shaw as transitional figures between the retrievalists and the bibliometricians.

The more interesting vertical axis reflects another subject-related continuum. Information science deals, we said earlier, with ''the human–computer–literature interface.'' If so, then the top pole represents a relative emphasis on *literatures* as objects of study, and the bottom, a relative emphasis on *people* or *users*. The same polarity can be inferred in earlier maps. Figure 4 showed that when a *literature* theoretician like Egghe enters, it is automatically at the top, whereas a *user* theoretician like Dervin is automatically placed at the bottom.

Applications of the computer pervade the field. Information science could, and did, exist without them, but in the past generation they figure inextricably in its progress. ''Bringing literatures to people through computers'' might be the current motto; or, in two words, ''literature systems.''

The polarization of quantitative versus qualitative researchers that is sometimes seen in other fields is not clearly visible in information science. There may be a few more mathematically-oriented writers in the upper half of Figure 6, but the lower has its share, too. So that dimension is not salient. Instead, one has only to look at the authors in the upper half of the map to infer a dominant interest in properties of literatures (especially scientific and technical literatures). Moreover, this orientation cuts across authors in all specialties, left to right, Price to Bookstein. The only real puzzle here is Shannon (always a problematic figure in L&IS), who has moved (following Zipf?) from the lower central position he has in all other maps. He rejoins his usual company (Winograd, Simon, Machlup, etc.) when scaled on the third dimension.

Although the authors in the lower half are a bit harder to categorize, they tend to put less emphasis on literatures per se than on people as users of literatures or, more generally, as users of scholarly and professional information systems. These authors discuss such topics as information-seeking behavior, user-friendly retrieval, optimal indexing, cognitive overload, human–computer interfaces, information policy issues, and the meaning of ''information.'' And this orientation, too, cuts across all specialties. At left, for example, Martyn, Paisley, and Allen

TABLE 6. Coordinates of ''polar'' authors on third INSPEC dimension.

Name	Historical pole	Name	Ahistorical pole
Frame	1.7975	Winograd	−2.1547
Carpenter	1.7465	Shannon	−1.8119
Line	1.5564	Rouse	−1.6746
Martyn	1.5420	Kantor	−1.5284
Narin	1.5156	Yu	−1.4682
Vlachy	1.4794	Croft	−1.4005
Moravcsik	1.4104	Simon	−1.3777
Meadows	1.3828	Maron	−1.3749
Hawkins	1.3352	T. D. Wilson	−1.3469
Garfield	1.2946	van Rijsbergen	−1.2704
Small	1.2435	Radecki	−1.2658
Lawani	1.2079	Belkin	−1.2627
Griffith	1.0565	Robertson	−1.2355
Tagliacozzo	1.0216	Sparck Jones	−1.1818
Price	1.0198	Machlup	−1.0461
Garvey	0.9857	W. S. Cooper	−1.0165
Pao	0.9513	Zipf	−0.9762
Smith	0.9270	Oddy	−0.9613
Chubin	0.9177	Salton	−0.9608
J. R. Cole	0.7499	Bookstein	−0.9506
Buckland	0.7434	Bates	−0.6525
S. Cole	0.6491	Taylor	−0.5729
Zuckerman	0.5716	Harter	−0.5529
Crane	0.5680	Kuhn	−0.4750
Paisley	0.5436	Doszkocs	−0.4525

all write about, or do, user studies; at right, where retrieval systems are a major concern, virtually everyone toward the lower pole is involved with social, behavioral, and cognitive aspects of design.

The oval plot of points in Figure 6 is tilted (compared to those seen earlier): The citationists and sociologists of science at left have shifted toward the ''literatures'' pole, while the retrievalists at right have shifted toward the ''people'' pole. The two camps remain in strict opposition. When the 75 authors are plotted in three dimensions rather than two, a new division appears that largely reinforces the opposition, rather than counteracting it. (At the same time, average R^2 rises to 0.87, and average stress drops to 0.15.) Table 6 rank-orders the 25 most extreme authors at either pole of this third dimension (and excludes the middle 25). It is immediately apparent that the authors at left who are *high* on the third axis are all, except for Hawkins, in the upper half of Figure 6, while the authors at right who are *low* on the third axis are all, except for Shannon and Zipf, in its lower half. What is the opposition to be inferred on this third axis? We think (echoing White and Griffith, 1981) it is between those whose interests are relatively *historical* (left) and those whose interests are relatively *ahistorical* (right).

The authors in the left column, that is, are not merely interested in properties of literatures; they tend to study the literatures (and other communications) of particular disciplines at particular points in time. Many are cliometricians of sorts, and some, such as Buckland and Zuckerman, actually write accounts of named individuals. Some also function like journalists, reporting on, e.g., the implications of new information technologies for scientific and technical communication. (Note that the more ahistorical bibliometricians are not in the top 25 on the positive half of this axis.) In contrast, the authors in the right column, many of whom are retrievalists, tend to write about systems or users (or both) without reference to particular times or disciplines. Examples of their subjects would be abstract typologies of information use, or evaluations of retrieval algorithms for users in general. They rarely write about individuals or do anything journalistic. The nonretrievalists among them, Winograd, Shannon, Zipf, Simon, and Machlup, apparently are cited in works on the nature of information itself—a rather timeless matter. Kuhn is also implicated with them, perhaps because the paradigms he writes about, though historical, are relatively permanent information structures.

In sum, most of the top domain analysts and retrievalists are perceived as being markedly different on all three dimensions. We leave open whether, and how, their ideas might be integrated in a single vision. According to ACA, Kochen's work is a plausible beginning. He is not only the most central author on the first two dimensions on the INDSCAL map, he is also at the exact middle of the third dimension (that is why in Table 6, which lists only the ''polar'' authors, he is not shown). But much of Kochen seems addressed only to mathematically literate specialists. The field still awaits a synthesist who can attract a broad readership.

Paradigm Shift?

We have suggested that ACA can be used to validate claims by historians and commentators. For example, in various ways our data have implied an increase of interest in the cognitive side of information science—and generally in user studies—since about 1980, the start of the second period. This independently corroborates claims to that effect by expert judges, such as Saracevic (1992), who calls it a paradigm shift, and Ingwersen (1996), who writes of it as ''the turning point 1977–1980.'' (ACA is a rearview mirror.) By way of contrast, note that the citation images of, say, the sociologists of science are relatively static in our maps from 1980 on.

Here we wish to explore, very tentatively, the possibility of capturing paradigm shift in numeric measures. As noted, the map in Figure 6 is a composite of citers' views of 75 authors in three different 8-year periods. The INDSCAL composite (''group stimulus space'') represents a solution in which, for all periods, the author placements and dimensions are fixed. However, INDSCAL is expressly designed to reveal *differences in the importance of each dimension* to whoever is judging the similarity of stimuli. In our use of INDSCAL, the stimuli are the 75

TABLE 7. Dimension weights for three periods.

Period	Weights		Weirdness index
	Dimension 1 Range of specialties	Dimension 2 Literature/users	
1972–1979	0.82	0.36	0.25
1980–1987	0.68	0.62	0.20
1988–1995	0.76	0.51	0.01

authors, and the three periods are regarded as three separate ''judges.'' (Rikken, Kiers, & Vos, 1995, take a related approach with another kind of bibliometric data.)

Usually, of course, *persons* are the judges in INDSCAL studies, and the ''derived subject weights,'' which are standard INDSCAL output, are taken to show the salience of each dimension to each person. In replacing individuals as judges with large numbers of citers, we are acting as if the citers collectively embodied the paradigm of information science in each 8-year period. Accordingly, we interpret the derived subject weights for each period as indicating the relative importance of the dimensions within the paradigm. Thus, we can probe a hidden aspect of disciplinary history—whether key dimensions of the field were given about the same weight in all periods. If not, that would be consistent with a perception of paradigm shift.

In our data, they were not. Table 7 has the dimensions as weighted in the two-dimensional solution at the end of each period. At the end of 1972–1979, the weight of the horizontal *range of specialties* dimension, 0.82, is more than twice that of the vertical *literatures/users* dimension, 0.36. But by the end of 1980–1987, a notable change has occurred. The two dimensions, now weighted 0.68 and 0.62 respectively, have become almost equally salient to citers. If the points in Figure 6 are reweighted by period, the authors form a long, narrow oval along the horizontal axis in 1972–1979, but a shorter, deeper oval in 1980–1987 when the weights of the two axes are more nearly equal. (The latter shape also accentuates the ''tilt'' of Figure 6, with domain analysts at upper left, retrievalists at lower right.)

Substantively, it is as if during 1972–1979 citers had regarded the *range of specialties* as by far the most important part of the information science paradigm, but then during 1980–1987 had taken much more cognizance of the differences in authors' orientation toward *literatures* or *users*. Changes in ''derived subject weights'' may or may not be useful measures of paradigm shift, but, in this case, they seem consistent with the perception that the 1980s were a time of change for information science, with ''user-oriented'' and ''literature-oriented'' authors increasingly differentiated across all specialties of the discipline. (Recall that weights for each period are based on independent data.)

The weights for the final period, 1988–1995, create a plot midway between the extremes of the first two: 0.76 for *specialties* and 0.51 for *literatures/users*, suggesting that, in recent years, the horizontal dimension has become about 1.5 times as important as the vertical in positioning authors. It looks almost as if the paradigm were exhibiting, not a clear trend, but a process of ''thesis–antithesis–synthesis.'' This may be an instance of ''pulsation'' or ''structural oscillation between expansion and contraction in the disciplines'' observed in Small (1993, p. 18).

This third set of weightings produces a plot very similar to the composite or ''most typical'' map in Figure 6. ''Weirdness'' in Table 7 is a routine INDSCAL measure that shows how much the view of each judging subject—here, each period—differs from that of the ''most typical'' subject. No period is especially weird by this index, which runs from 0 (for complete agreement with typicality) to 1 (for completely ignoring one dimension in favor of the other); but the third period is obviously least weird of all.

Figure 7 summarizes matters. The points correspond to the weights given the two main dimensions by each period (the diagonal line represents an equal weighting of dimensions). INDSCAL normalizes weights so that, when squared, they sum to the R^2 for the period. Since the R^2s for all periods are about the same (low 0.80s), the vectors representing the periods are all about the same length. The difference in the angles of the vectors as they project from the origin suggests a way of rendering paradigm shift—here, most marked between 1972–1979 and 1980–1987. Admittedly, we do not know whether the differences between periods are *unusually* great; since this is the measure's first appearance, there is no point of

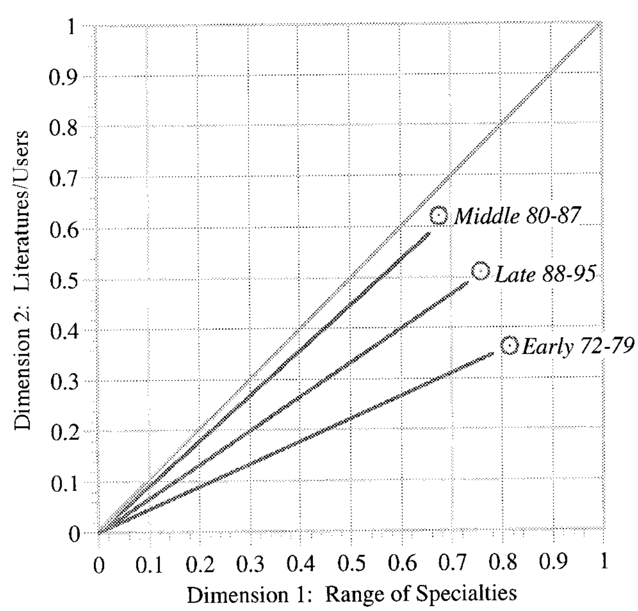


FIG. 7. Changes in importance of axes over three periods.

comparison. They must also depend on how the data are periodized.

When the same input data are analyzed in three dimensions, the pattern seen in Table 7 is preserved. The derived subject weights for the early, middle, and late periods are 0.79, 0.63, and 0.71 on the first dimension; 0.34, 0.61, and 0.50 on the second; and 0.29, 0.36, and 0.35 on the third. The split on the third, *historical/ahistorical* dimension, which reinforces splits on the other two, thus becomes slightly more pronounced with time. The respective R^2s are 0.82, 0.89, and 0.88. The respective weirdness indices are 0.19, 0.15, and 0.02.

Perhaps the main weakness of this INDSCAL measure is that it is so indirect—that is, not clearly connected to specific papers with specific claims about the world. One expects evidence of paradigm shifts to leap from main texts, not references; from writers, not citers. Accordingly, it seems best to use this measure when there are already grounds for believing that a shift has occurred. The present article in several ways supports the idea of a post-1970s shift—perhaps ''the cognitive revolution'' of other fields coming to information science (especially to some authors identified with retrieval), even as the less cognitive, ''bibliometric'' side was increasing in strength. But not all important authors participated in this change, and there is a question whether the discipline as a whole was impacted by it. It is therefore interesting that the INDSCAL measure shows a global change in perception along lines that other evidence or expert judgment predicts.

Which returns us to the measure's possible strength. Though it might be used to *discover* paradigm shift, we think it has more promise as a means of *confirming* one. The historian who has inferred a shift from reading is prompted to test whether it can be confirmed in the periodized citation record—the hardest part of the disciplinary paradigm to affect. A shift detectable *there* implies not only that authors are promoting new lines of inquiry, but that citers are responding in such a way that the overall map of the discipline is changed. The history of the paradigm grows to include not only authors' claims but citers' perceptions, as foreseen by Kuhn (1962, p. xi):

> If I am right that each scientific revolution alters the historical perspective of the community that experiences it, then that change of perspective should affect the structure of postrevolutionary textbooks and research publications. One such effect—a shift in the distribution of the technical literature cited in the footnotes to research reports—ought to be studied as a possible index to the occurrence of revolutions.

But is a Kuhnian revolution or paradigm shift the same as the natural ''pulsation'' of a discipline in Small's (1993) sense? Can ACA distinguish between them? Can any literature visualization technique? We do not yet know.

Conclusion

ACA shows two broad subdisciplines in information science, each with a certain internal coherence, each with authors who read, write, and are co-cited in both. Even so, ACA does not lead to an author whose work unequivocally unites both sides or magisterially portrays the whole discipline. There is no North American textbook that introduces a representative body of the authors mapped here. There is no serious popularization that puts Small and Sparck Jones, or Egghe and Cochrane, in a common framework (how many books have popularized AI?). An integrated account of information science as embodied by our ''canonical'' authors seems to us well worth having.

Toward that account, ACA simultaneously provides both breadth and focus. It provides breadth by forcing contemplation of multiple specialties: Once they are seen, it is difficult to reduce ASIS-style information science to anything less (some will claim more). It provides focus by forcing contemplation of particular authors, which is to say particular oeuvres and works. It also provides crude but unmistakable evidence of intellectual change.

Beneath the changing surface, however, there is a unifying concern. In our view, most of the authors mapped here are united in working on aspects of *literatures* as modes of communication. The intellectual problems they address arise from a particular kind of information system—literatures as content-bearing objects—rather than from ''information systems'' in general. Literatures are the distinguishing mark of domains likely to interest domain analysts, and at least part of their interest in literatures lies in features that are exploitable for retrieval. The partitioning of literatures on the basis of such features, so as to most usefully approximate an answer to a request, defines the retrievalists' enterprise.

Not all information scientists are clear on these points. The rich word ''information'' has seduced some into characterizations of their field that are, to date, overgeneral. Thus, ACA can help assess their pronouncements in the disciplinary press. It leads to a judgment of ''highly unfocused'' on definitions of the field such as those quoted in Rayward (1996, p. 4): ''The application of information science results in an information system. The role of information science is to explicate the conceptual and methodological foundations on which existing systems are based'' (Borko, 1968, p. 67). Or ''Information science is the study of the means by which organised structures (which we call 'information systems') process recorded symbols to meet their defined objectives'' (Hayes, 1985, p. 174). These definitions would have ASIS-style information science dealing with employers' payrolls, a housewife's grocery receipts, *Nightline*, arrival and departure listings in airports, the Grand Ole Opry, and color-coded vial caps used by crack dealers. Need we say that it does not?

Examine the works of authors revealed here by ACA.

It is doubtful that information in all senses is their true object of study (beyond explicating it as a word). Which of them, for example, does empirical work on information as a sustained psychological process, as in discourse comprehension or sensory perception? That is left to others. What they do study empirically, and uniquely, are problems associated with the human–literature barrier—the special difficulties of obtaining answers to questions from publications, in any medium, rather than persons. In other words, while many scholars seek to understand communication between persons, information scientists seek to understand communication between persons and certain valued surrogates for persons that literatures comprise (White, 1992). This study requires a conceptual scheme that encompasses properties not only of literatures (e.g., size, growth rate, age, dispersion, authority levels, degree of summarization, quality of indexing) but also of people (e.g., interests and concerns, vocabularies, social ties, knowledge of existing systems, search styles, editorial strategies, resource environments). In a way it is unfortunate that the field is not called *literature science*, but that, of course, would only create new definitional problems and misunderstandings, such as taking ''literature'' to mean nothing but belles-lettres or nothing but texts (cf. Buckland, 1991).

The bond between domain analysts and retrievalists is their common interest in the literature barrier and related phenomena on both sides. The barrier in action is exemplified by information overload and underload—recurring topics for authors in both subdisciplines because they require both *literatures* and *users* to be discussed in a single framework, as implied by the second dimension of our maps. Of course, to explicate this framework in the manner of Ingwersen (1996) or Vakkari (1996), we would need to move from the level of authors and works to the level of problem-statements and findings, claims and counterclaims. We conspicuously have not done that here. But we have identified and categorized a set of authors with whom an adequate topical account of information science, synchronic or diachronic, must deal.

We do not mean that the ideas of all 120, or 100, or 75, must be somehow be synthesized in one gigantic review. What is needed are accounts capacious enough to explain why the work of any of them is valued—and why some are more valued than others. Take Price, Lancaster, Brookes, and Winograd—or any roughly equivalent set of authors—as four points of an intellectual compass. In what history would they all fit comfortably by name? What is it that gives them *jointly* their significance? Author co-citation analysis makes it difficult for information science, a rather evasive field, to evade such questions. It can raise them in other fields as well. If the questions it raises are important, so may be the answers.

References

Aversa, E. (1985). Citation patterns of highly cited papers and their relationship to literature aging: A study of the working literature. *Scientometrics, 7,* 383–389.

Bayer, A. E., Smart, J. C., & McLaughlin, G. W. (1990). Mapping intellectual structure of a scientific subfield [sociology of the family] through author co-citations. *Journal of the American Society for Information Science, 41,* 444–452.

Beghtol, C. (1995). Domain analysis, literary warrant, and consensus: The case of fiction studies. *Journal of the American Society for Information Science, 46,* 30–44.

Blair, D. C. (1990). *Language and representation in information retrieval.* Amsterdam: Elsevier.

Borgman, C. L. (Ed.). (1990). *Scholarly communication and bibliometrics.* Newbury Park, CA: Sage.

Borko, H. (1968). The conceptual foundations of information systems. In E. B. Montgomery (Ed.), *The foundations of access to knowledge: A symposium.* Syracuse, NY: Syracuse University Press.

Buckland, M. (1991). Information retrieval of more than text. *Journal of the American Society for Information Science, 42,* 586–588.

Buckland, M. (1996). Documentation, information science, and library science in the U.S.A. *Information Processing & Management, 32,* 63–76.

Buckland, M. K., & Liu, Z. (1995). History of information science. *Annual Review of Information Science and Technology, 30,* 385–416.

Cozzens, S. E. (1985). Comparing the sciences—Citation context analysis of papers from neuropharmacology and the sociology of science. *Social Studies of Science, 15,* 127–153.

Croft, W. B. (1996). Remarks at *Information Retrieval 2000: Workplace Needs and Curricular Implications:* Symposium sponsored by College of Information Science and Technology, Drexel University, Philadelphia, PA, May 24, 1996.

De Mey, M. (1982). *The cognitive paradigm.* Dordrecht, The Netherlands: Reidel.

Eom, S. B. (1996). Mapping the intellectual structure of research in decision support systems through author co-citation analysis. *Decision Support Systems, 16,* 315–338.

Gorman, M. (1990). A bogus and dismal science, or the eggplant that ate library schools. *American Libraries, 21,* 462–463.

Griffith, B. C. (Ed.). (1980). *Key papers in information science.* White Plains, NY: Knowledge Industry Publications.

Harter, S. P. (1992). Psychological relevance and information science. *Journal of the American Society for Information Science, 43,* 602–615.

Hayes, R. M. (1985). The history of library and information science: A commentary. *Journal of Library History, 20,* 173–178.

Hoffman, D. L., & Holbrook, M. B. (1993). The intellectual structure of consumer research—A bibliometric study of author co-citations in the 1st 15 years of the *Journal of Consumer Research. Journal of Consumer Research, 19,* 505–517.

Hjørland, B., & Albrechtsen, H. (1995). Toward a new horizon in information science: Domain-analysis. *Journal of the American Society for Information Science, 46,* 400–425.

Ingwersen, P. (1996). Information and information science in context. In J. Olaisen, E. Munch-Petersen, & P. Wilson (Eds.), *Information science: From the development of the discipline to social interaction* (pp. 69–111). Oslo: Scandinavian University Press.

Kantor, P. B. (1981). Rigorous solution of the competitive advantage problem—Application to bibliometric problems. *Proceedings of the American Society for Information Science, 18,* 282–284.

Kantor, P. B. (1994). Information retrieval techniques. *Annual Review of Information Science and Technology, 29,* 53–90.

Karki, R. (1996). Searching for bridges between disciplines—An author co-citation analysis on the research into scholarly communication. *Journal of Information Science, 22,* 323–334.

Kuhn, T. S. (1962). *The structure of scientific revolutions.* Chicago: University of Chicago Press.

Lancaster, F. W., Bushur, S., & Low, Y. M. (1993). Kochen influence examined bibliometrically. *Library Trends, 41,* 549–566.

Machlup, F., & Mansfield, U. (Eds.). (1983). *The study of information: interdisciplinary messages.* New York: Wiley.

Martyn, J. (1974). Information needs and uses. *Annual Review of Information Science and Technology, 9,* 3–23.

Martyn, J. (1975). Progress in documentation: Citation analysis. *Journal of Documentation, 31,* 290–297.

McCain, K. W. (1986). Co-cited author mapping as a valid representation of intellectual structure [in macroeconomics and *drosophila* genetics]. *Journal of the American Society for Information Science, 37,* 111–122.

McCain, K. W. (1990). Mapping authors in intellectual space: A technical overview. *Journal of the American Society for Information Science, 41,* 433–443.

McKenzie, D. F. (1986). *Bibliography and the sociology of texts.* London: British Library.

Mullins, N. C. (1973). *Theories and theory groups in American sociology.* New York: Harper & Row.

Olaisen, J., Munch-Petersen, E., & Wilson, P. (Eds.). (1996). *Information science: From the development of the discipline to social interaction.* Oslo: Scandinavian University Press.

Persson, O. (1994). The intellectual base and research fronts of JASIS 1986–1990. *Journal of the American Society for Information Science, 45,* 31–38.

Rayward, W. B. (1996). The history and historiography of information science: Some reflections. *Information Processing & Management, 32,* 3–17.

Rikken, F., Kiers, H., & Vos, R. (1995). Mapping the dynamics of adverse drug-reactions in subsequent time periods using INDSCAL. *Scientometrics, 33,* 367–380.

Saito, Y. (1984). Identification of the specialties in library and information science using co-citation analysis. *Library and Information Science, 22,* 61–74.

Saracevic, T. (1992). Information science: Origin, evolution, relations. In P. Vakkari & B. Cronin (Eds.), *Conceptions of library and information science: Historical, empirical and theoretical perspectives* (pp. 5–27). London: Taylor Graham.

Small, H. G. (1978). Cited documents as concept symbols. *Social Studies of Science, 8,* 327–340.

Small, H. G. (1980). Co-citation context analysis and the structure of paradigms. *Journal of Documentation, 36,* 183–196.

Small, H. G. (1981). The relationship of information science to the social sciences: A co-citation analysis. *Information Processing and Management, 17,* 39–50.

Small, H. G. (1993). Macro-level changes in the structure of co-citation clusters: 1983–1989. *Scientometrics, 26,* 5–20.

Smith, L. C. (1981). Citation analysis. *Library Trends, 30,* 83–106.

Sparck Jones, K. (Ed.). (1981). *Information retrieval experiment.* London: Butterworth.

SPSS Base System User's Guide (1990; also later eds.) Chicago: Marija J. Norusis/SPSS.

Tague-Sutcliffe, J. (1981). The success-breeds-success phenomenon and bibliometric processes. *Journal of the American Society for Information Science, 32,* 280–286.

Tudjman, M., Tudor-Silovic, N., Boras, D., & Milas-Bracovic, M. (1988). A literature measure of scientific communication: Co-citation analysis of masters theses in information sciences in Yugoslavia, 1961–1984. In N. Tudor-Silovic, & I. Mihel (Eds.), *Information research; research methods in library and information science* (pp. 225–247). London: Taylor Graham.

Vakkari, P. (1996). Library and information science: Content and scope. In J. Olaisen, E. Munch-Petersen, & P. Wilson (Eds.), *Information science: From the development of the discipline to social interaction* (pp. 169–231). Oslo: Scandinavian University Press.

Vakkari, P., & Cronin, B. (Eds.). (1992). *Conceptions of library and information science: Historical, empirical and theoretical perspectives.* London: Taylor Graham.

Vickery, B. C., & Vickery, A. (1987). *Information science in theory and practice.* London: Butterworths.

White, H. D. (1990a). Author co-citation analysis: Overview and defense. In C. L. Borgman (Ed.), *Scholarly communication and bibliometrics* (pp. 84–106). Newbury Park, CA: Sage.

White, H. D. (1990b). Introduction. *Journal of the American Society for Information Science, 41,* 430–431.

White, H. D. (1992). External memory. In H. D. White, M. J. Bates, & P. Wilson, *For information specialists; interpretations of reference and bibliographic work* (pp. 249–294). Norwood, NJ: Ablex.

White, H. D., & Griffith, B. C. (1981). Author co-citation: A literature measure of intellectual structure. *Journal of the American Society for Information Science, 32,* 163–172.

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. (1989). Bibliometrics. *Annual Review of Information Science and Technology, 24,* 119–186.

Williams, J. G., Sochats, K. M., & Morse, E. (1995). Visualization. *Annual Review of Information Science and Technology, 30,* 161–207.

Wilson, P. (1983). Bibliographical R&D. In F. Machlup & U. Mansfield (Eds.), *The study of information; interdisciplinary messages* (pp. 389–397). New York: Wiley.

Wilson, P. (1996). The future of research in our field. In J. Olaisen, E. Munch-Petersen, & P. Wilson (Eds.), *Information science: From the development of the discipline to social interaction* (pp. 319–323). Oslo: Scandinavian University Press.