Information Processing & Management, Vol. 17, pp. 39–50, 1981
Printed in Great Britain

0306-4573/81/010039-12$02.00/0
Pergamon Press Ltd.

THE RELATIONSHIP OF INFORMATION SCIENCE TO THE SOCIAL SCIENCES: A CO-CITATION ANALYSIS†

HENRY SMALL
Institute for Scientific Information, 3501 Market Street, University City Science Center, Philadelphia, PA 19104, U.S.A.

(*Received* 3 *April* 1980)

Abstract—A co-citation cluster analysis of a three year (1975–77) cumulation of the *Social Sciences Citation Index* is described, and clusters of information science documents contained in this data-base are identified using a journal subset concentration measure. The internal structure of the information science clusters is analyzed in terms of co-citations among clusters, and external linkages to fields outside information science are explored. It is shown that clusters identified by the journal concentration method also cohere in a natural way through cluster co-citation. Conclusions are drawn regarding the relationship of information science to the social sciences, and suggestions are made on how these data might be used in planning an agenda for research in the field.

INTRODUCTION

One way of gaining insight into the state of a research field or discipline is to examine the publications produced by its practitioners. To the extent that practitioners in the field publish the results of their investigations, this mode for assessing the state of a field can reflect with great specificity the content and problem orientations of the group. Of the many ways that publications can be analyzed and counted, perhaps the most revealing kind of data are the references cited by the practitioner group in their publications. References to earlier literature tell us about the author making them as well as the items being cited. When references are cumulated over a significant volume of source literature such as in the *Science Citation Index* (*SCI*)®and *Social Sciences Citation Index* (*SSCI*)®, the collective patterns reveal the concerns of the field as symbolized by the documents and authors cited. This is how earlier cited literature can inform us on the current conceptual framework: the act of citing involves an association of a notion or idea expressed in the text with a cited document[1]. Hence, each reference is connected to a concept. The cumulative pattern of such contexts provides a representation of the cognitive structure of the research field.

The objective of a study currently underway at the Institute for Scientific Information is to examine the structure and development of the field of information science using the published literature of information science as data, and the techniques of citation analysis. We know very little about how the field of information science has developed over the past several years. On the one hand the field might be viewed from a technological standpoint, the primary accomplishments of which are creation of machine readable data bases and retrieval systems. From another perspective, however, the field of information science can be seen as an investigation into the nature of information, the theoretical basis for retrieval, the evaluation of retrieval, and the way that human beings use and transmit information. We expect that the literature of information science will reflect both the conceptual and technical concerns of the field. By using the statistical techniques of citation analysis we hope to get a picture of how the field has developed, the main lines of research in the field, its principal foci of interest, and where the field appears to be going.

Some studies have attempted to use published literature to arrive at insights into the structure of the field. Saracevic reviewed the first five volumes of the *Annual Review of Information Science and Technology* using bibliometric techniques and concluded that ARIST was biased toward the technology and practice of information science and against fundamental

†Research supported by NSF grant IST-78-16677.

research in the area[2]. In a breakdown of most cited authors in ARIST Saracevic found about 10% theoreticians, 40–50% experimenters, and the rest developers.

Donohue has provided the most ambitious bibliometric analysis of the "information science" literature, although his choice of journals to include in the database was somewhat idiosyncratic (e.g. inclusion of the *Journal of the Acoustical Society of America* as an information science journal)[3]. In general, the main corpus used in Donohue's study represented information science in its more technical, mathematical and engineering sense, rather than information science related to documentation, retrieval and library science. Therefore, the structure of "information science" in Donohue's study cannot be taken as a guide to what we expect to find in our study. (For example, the list of clusters Donohue obtained using bibliographic coupling among documents included groups designated as acoustics, computing, cybernetics, engineering, logic, numerical mathematics and statistics).

Salton has undertaken a citation analysis of individual researchers in information science using two comprehensive bibliographies and two points in time a decade apart[4]. His main conclusion is that the field has changed significantly over the decade and that this perhaps indicates not only intellectual ferment, but perhaps also a lack of focus on key problems. From Salton's study we would anticipate a lack of lasting theoretical orientations and a predominance of experimental work without the benefit of a theoretical framework.

A recent survey of concerns in the field of information science by Pratt, based on a qualitative analysis of the ARIST series which does not use bibliometric techniques, finds the major topics of current research to be[5]:

(1) Library problems
(2) Economics of information
(3) The nature of information
(4) Techniques of measurements.

All of the studies mentioned provide some hints of what we could expect in our citation analysis, but certainly none stands out as a definitive benchmark against which to compare our results.

Our study takes two different approaches to the analysis of information science as a field. First, we are using a special citation file extracted from the *Social Sciences Citation Index* database, consisting of a core set of journals in information science over a nine year period (1969–77), to explore the internal structure of information science and its development. The second approach is to show how information science links to the other fields and disciplines in the social sciences. To accomplish this we use existing cluster data files at ISI. It is this latter work which I report on in the present paper.

CLUSTER ANALYSIS OF THE SSCI

The starting point for this study was a cluster analysis of a special three year cumulation (1975–77) of the *Social Sciences Citation Index* (*SSCI*). Earlier we had performed a similar analysis of the *SSCI* for the period 1972–74[6, 7], and the new analysis of the 1975–77 file will allow us, eventually, to examine rates of specialty change within the social sciences. The results I will report here are only for the 1975–77 file and do not attempt to assess change over time.

The clustering procedure was identical to that used in the original study. I will review only the essentials here. First, all documents in the file cited ten or more times during the three year period are selected. Table 1 presents some statistics on the cluster analysis, and indicates that of the over two million cited items in the three year file, about 25,000 were cited ten or more times. This is a fairly weak criterion for selection when compared with our usual threshold of 15 citations per document per year for cluster analyses of annual SCI's[8]. Following selection of these highly cited items, all co-citations among the 25,000 were determined, that is, the number of times any pair of them is cited together in the three year period. As indicated in the Table there were 1.8 million unique pairs of co-cited items thus formed. The raw co-citation counts for each pair were normalized by dividing by the sum of citation frequencies for the two items minus the number of co-citations. This is essentially the fraction of citations to the two items that are co-citations, which is equivalent to the so-called Jaccard coefficient used in numerical taxonomy[9]. These coefficients were the basis for the cluster analysis. The clustering agorithm used (called single-link clustering)[10] requires only that we specify a threshold for

Table 1. Statistics on clusters from 1975–77 Cumulative *Social Sciences Citation Index* (*SSCI*)

1. total citations (1975-1977)	3,399,058
2. distinct cited items	2,196,127
3. highly cited items (≥ 10 citations)	24,954 (1.14%)
4. distinct co-cited pairs of cited items	1,846,585
5. distinct co-cited pairs at level 22%	10,418
6. clusters at level 22% (≥ 2 cited items)	2,095
7. mean cited items per cluster	4.1
8. mean citing items per cluster	39.9

the normalized co-citation strength to generate a set of disjoint clusters containing the cited items. Setting this threshold at 0.22 generated about 2000 clusters each containing two or more cited items. The average cluster size was 4.1 cited items. This set of 2000 clusters in the social and behavioral sciences formed the universe from which we selected clusters on information science topics. Of course, information science is expected to represent only a small fraction of the clusters in this file, which is dominated by fields such as psychology (experimental and social), sociology, economics, psychiatry, and so on.

THE SELECTION OF INFORMATION SCIENCE CLUSTERS

The procedure used to select information science clusters from the 2000 1975–77 *SSCI* clusters was to define a set of information science journals which appear as source journals in the *SSCI*. Fifty journals were selected and are listed in Table 2. This list should not be regarded

Table 2. Information science journal subset

1. American Archivist
2. Annual Review of Information Science and Technology
3. Aslib Proceedings
4. Bulletin of the Copyright Society of the U.S.A.
5. Bulletin of the Medical Library Association
6. Canadian Journal of Information Science
7. Canadian Library Journal
8. College and Research Libraries
9. Drexel Library Quarterly
10. Government Publications Review
11. IEEE Transactions on Information Theory
12. IEEE Transactions on Engineering Management
13. IEEE Transactions on Professional Communication
14. Information and Control
15. Information Processing & Management
16. Information Sciences
17. Information Scientist
18. International Classification
19. International Forum on Information and Documentation
20. International Journal of Computer & Information Sciences
21. Journal of the American Society for Information Science
22. Journal of Chemical Information & Computer Sciences
23. Journal of Documentation
24. Journal of Education for Librarianship
25. Journal of Librarianship

Table 2 (*Contd*).

26.	Journal of Library Automation
27.	Journal of Library History Phiosophy & Comparative Librarianship
28.	Journal of the Patent Office Society
29.	Law Library Journal
30.	Library & Information Science
31.	Library Resources & Technical Services
32.	Library Trends
33.	Library Quaterly
34.	Libri
35.	Methods of Information in Medicine
36.	Nachrichten Für Dokumentation
37.	Nauchno-Teknischeskaya Informatsiya. Seriya 1. Organizatsiya I Metodika Informatsionnoi Raboty
38.	Nauchno-Teknicheskaya Informatsiya. Seriya 2. Informatsionnye Protessy I Sistemy
39.	On-line Review
40.	Proceedings of the American Society for Information Science
41.	Pattern Recognition
42.	Program-New of Computers in Libraries
43.	Review of Public Data Use
44.	Social Science Information
45.	Social Studies of Science
46.	Special Libraries
47.	Unesco Bulletin for Libraries
48.	Wilson Library Bulletin
49.	Zeitschrift für Bibliothekswesen Und Bibliographie
50.	Zentralblatt für Bibliothkswesen

as a definitive list of journals in the field, but rather as one way of defining the field, which is subject to an empirical test later in our analysis. It should be noted, for example, that computer science journals were intentionally not included in the list, and that we were slanting the list toward the library/information science direction. Some journals included were concerned with the mathematical study of communication (information), but the majority deal with the more traditional view of information science as an off-shoot of documentation.

The procedure to select clusters, similar to that used in an earlier study[7], was to calculate the fraction of source (citing) papers for each cluster which fall in the specified journal set. A distribution of these fractions is obtained (see Fig. 1) which ranges from clusters having 100% of their citing papers in journals which are members of the set, to clusters which have none of their citing papers in these journals. For purposes of comparison a similar distribution for the field of chemistry is included on the same graph (from *SCI* not *SSCI* cluster data). The chemistry distribution shows clearly a group of disciplinary clusters centered on about 80% concentration in the journal sub-set for chemistry, and a smaller interdisciplinary group of clusters centered at 55%. The up-swing of the distribution to the left shows all the clusters which are not, or only marginally, in the field. The information science distribution shows similar disciplinary and interdisciplinary peaks, though on a much smaller scale: there are only eleven clusters of the 2000 which have 30% or more of their citing papers in information science journals, and only 22 with 10% or more. The latter group is listed in Table 3.

CHARACTERISTICS OF THE INFORMATION SCIENCE CLUSTERS

Information science by our definition, therefore, comprises at most one percent of the clusters in the three year social and behavioral sciences database (about 20 of 2000

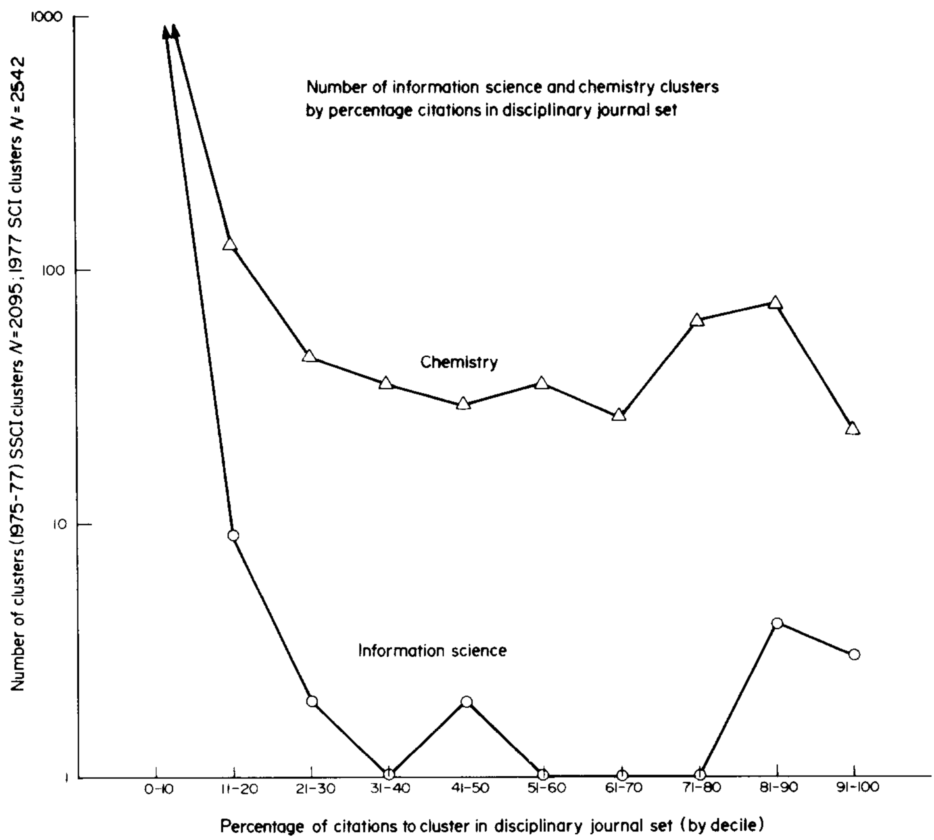

Fig. 1. Number of information science and chemistry clusters by percentage citations in disciplinary journal set.

clusters). Nevertheless, we can examine the clusters identified to see what they can tell us about the field. Referring to Table 3, each of the clusters with 10% or more of their citing papers in information science journals has been listed in descending order by percentage concentration. The order of clusters in Table 3 could be interpreted as the degree of disciplinary purity. This ranges from 100% for the "Precis" cluster to 14.3% for "copyright law" and "medical use of computers".

The cluster number in the left column is an arbitrary identification number assigned by the computer to each cluster as it is generated. An approximate name was given to each cluster based on an examination of the titles of the cited and citing documents. In the last two columns the size of the cluster is given both in terms of the number of cited items and the number of items citing them.

Table 4 lists all documents cited ten or more times which comprise the clusters having at least 40% of their citations from information science journals. The number of times each document was cited in the *SSCI* from 1975 to 1977 is indicated in the right-hand column. Of course, citations to an item can come from any source journal in the *SSCI* coverage, not just those journals listed in Table 2.

Most of the clusters are very small (the smallest number of cited items a cluster can have is two). In addition, a number of clusters have been given the same name (e.g. there are "on-line retrieval and data bases" clusters "a" and "b"). This redundancy occurs because it was not possible to distinguish the content of these clusters based on the titles of the papers, and they were therefore given the same name. This suggests that these areas are being fragmented at this level of association (22% normalized co-citation) and we will show in a moment that this is indeed the case. Both the small size of the clusters and their fragmentation tell us that information science has a weaker structure than areas such as psychology, sociology or economics which emerge as larger and more coherent specialties at this level. If the co-citation threshold had been lowered, the fragmented information science areas would have congealed, but for other subject areas, composite macro-clusters would have formed, indicating too low a threshold in their case.

Table 3. Information science clusters (greater than 10% participation in journal set)

Cluster Number	Percent Participation	Cluster Name	# Cited Items	#Citing
1. 239	100.0	Precis	2	31
2. 1140*	91.7	On-line retrieval & data bases (a)	2	24
3. 606*	91.1	Bradford's law & bibliometrics (a)	7	45
4. 311*	88.9	On-line retrieval & data bases (b)	2	18
5. 563*	87.5	Theory of indexing & retrieval (a)	2	16
6. 1999*	82.6	Theory of indexing & retrieval (b)	3	23
7. 777*	81.2	Bradford's law & bibliometrics (b)	2	16
8. 1998*	80.0	Theory of indexing & retrieval (c)	2	50
9. 824*	69.6	Library serials planning	2	23
10. 1299*	47.3	Citation analysis (a)	4	55
11. 1300*	41.2	Citation analysis (b)	2	17
12. 101*	29.4	Sociology of Science (a)	3	34
13. 2021*	24.1	Zipf's law	2	29
14. 1239*	20.8	Sociology of science (b)	3	24
15. 406	20.3	Fuzzy systems	12	64
16. 548*	20.0	Sociology of science (c)	2	20
17. 1096	20.0	Symbolic interactionism	2	15
18. 1844*	19.2	Sociology of science (d)	2	26
19. 2013	18.3	Information theory	2	60
20. 898*	18.2	Sociology of science (e)	5	55
21. 1354	14.3	Copyright law	3	21
22. 1421	14.3	Medical use of computers	3	21

* on single-link network

We also investigated how the clusters in information science compare with cluster samples in other fields in terms of size, age of cited material, and percentage of cited items that are books. The results are shown in Table 5 for the comparison fields of sociology, economics, psychology and particle physics. The last named field was included as an extreme case of a large and fast moving specialty in the physical sciences, obtained of course from the *SCI.* It is clear that information science clusters are very small compared with clusters in these other areas. Regarding the recency of literature cited, information science falls roughly between sociology and economics, when measured by either the mean publication date of cited items or Price's index[11] (the percentage of items which fall within the last five years). Somewhat surprisingly, information science clusters contain relatively few books as cited items, about the same percentage as psychology. This indicates the tendency for important contributions in the field to appear in journal article form.

THE STRUCTURE OF INFORMATION SCIENCE CLUSTERS

The next question was how the selected clusters in information science relate to one another? To investigate this we make use of residual co-citation linkages—the linkages between documents at levels lower than the clustering level. In other words, when clusters are formed at level 22% we utilize only the strongest co-citation links among the highly cited documents. But there are many weak links below the clustering threshold among documents in different clusters, and summing these residual links enables us to measure the strength of inter-cluster relationship. When a matrix of these intercluster links is analyzed using the technique of multidimensional scaling[12], a two-dimensional configuration of points (each representing a cluster) is obtained in which high co-citation between clusters corresponds to proximity in the

Table 4. Highly cited documents in information science clusters (with greater than 40% of their citations in information science journals)

Cluster Name / Cluster Number	Document	times cited 1975-1977
Precis		
239	Derek Austin, "The Development of Precis: A Theoretical and Technical History," Journal of Documentation 30: 47-101, 1974.	19
	Derek Austin, Precis: A Manual of Concept Analysis and Subject Indexing (London: Council of the British National Bibliography, 1974).	19
On-line retrieval and data bases		
311	A. Stephanie Barber, Elizabeth D. Barraclough and W. Alexander Gray, "On-line Information Retrieval as a Scientists Tool," Information Storage and Retrieval 9: 429-440, 1973.	10
	Jeffrey Katzer, "The Cost-Performance of an On-line, Free-Text Bibliographic Retrieval System," Information Storage and Retrieval 9: 321-329, 1973.	12
1140	Stanley A. Elman, "Cost Comparison of Manual and On-line Computerized Literature Searching," Special Libraries 66: 12-17, 1975.	20
	Barbara Lawrence, Ben H. Weil and Margaret H. Graham, "Making On-line Search Available in an Industrial Research Environment," Journal of the American Society for Information Science 25: 364-369, 1974.	10
Bradford's Law and bibliometrics		
606	S.C. Bradford, Documentation (Washington, D.C.: Public Affairs Press, 1948).	14
	B.C. Brookes, "The Derivation and Application of the Bradford-Zipf Distribution," Journal of Documentation 24: 247-259, 1968.	19
	B.C. Brookes, "Numerical Methods of Bibliographic Analysis," Library Trends 22:18-43, 1973.	10
	B.C. Brookes, "Bradford's Law and the Bibliography of Science," Nature 224: 953-956, 1969.	15
	Ferdinand F. Leimkuhler, "The Bradford Distribution," Journal of Documentation 23: 197-207, 1967.	17
	B.C. Vickery, "Bradford's Law of Scattering," Journal of Documentation 4: 198-203, 1948.	16
	Elizabeth A. Wilkinson,"The Ambiguity of Bradford's Law," Journal of Documentation 28: 122-129, 1972.	11
777	Mark P. Carpenter and Francis Narin, "Clustering of Scientific Journals," Journal of the American Society for Information Science 24:425-436, 1973.	12
	Francis Narin, Mark Carpenter and Nancy Berlt, "Interrelationships of Scientific Journals," Journal of the American Society for Information Science 23: 323-331, 1972.	12
Theory of indexing and retrieval		
563	Abraham Bookstein and Don R. Swanson, "Probabilistic Models for Automatic Indexing," Journal of the American Society for Information Science 25: 312-318, 1974.	11
	M.E. Maron and J.L. Kuhns, "On Relevance, Probabilistic Indexing and Information Retrieval," Journal of the Association for Computing Machinery 7: 216-242, 1960.	10
1998	Gerard Salton, Automatic Information Organization and Retrieval (New York: McGraw-Hill, 1968).	41
	Gerard Salton, ed. The Smart Retrieval System: Experiments in Automatic Document Processing (Englewood Cliffs, New Jersey: Prentice-Hall, 1971).	23

Table 4 (contd)

Cluster Name / Cluster Number	Document	Times Cited 1975–1977
1999	Gerard Salton and M.E. Lesk, "Computer Evaluation of Indexing and Text Processing," Journal of the Association for Computing Machinery 15: 8–36, 1968.	10
	Gerard Salton and C.S. Yang, "On the Specification of Term Values in Automatic Indexing," Journal of Documentation 29: 351-372, 1973.	12
	Karen Sparck-Jones,"A Statistical Interpretation of Term Specificity and its Application in Retrieval," Journal of Documentation 28: 11-21, 1972.	10
Library serials planning		
824	Ching-Chih Chen, "The Use Patterns of Physics Journals in a Large Academic Research Library," Journal of the American Society for Information Science 23: 254-264, 1972.	16
	Alexander Sandison, "Densities of Use, and Absence of Obsolescence, in Physics Journals at MIT," Journal of the American Society for Information Science 25: 172-178, 1974.	18
Citation analysis		
1299	E. Garfield, I.H. Sher and R.J. Torpie, The Use of Citation Data for Writing the History of Science (Philadelphia: Institute for Scientific Information, 1964).	10
	B.C. Griffith, H.G. Small, J.A. Stonehill, and S. Dey, "The Structure of Scientific Literatures II: Toward a Macro- and Microstructure for Science," Science Studies 4: 339-365, 1974.	18
	H. Small, "Co-citation in the Scientific Literature: A New Measure of the Relationship between Two Documents," Journal of the American Society for Information Science 24: 265-269, 1973.	18
	H. Small and B.C. Griffith, "The Structure of Scientific Literatures I: Identifying and Graphing Specialties," Science Studies 4: 17-40, 1974.	28
1300	E. Garfield, "Citation Indexes for Science," Science 122: 108-111, 1955.	10
	M. Weinstock, "Citation Indexes," in: Encyclopedia of Library and Information Science (New York: Marcel Dekker, 1971) pp.16-40.	12

configuration. Figure 2 shows the results of the scaling for the clusters having 10% or more journal concentration (clusters 1096, 1354 and 1421 had no links to other areas and were excluded from the analysis). Each point (cluster) has ben labeled by the percentage concentration and all non-zero inter-cluster linkages have been drawn as lines connecting the points.

This analysis clearly shows that the redundant cluster names are the result of a fragmentation of larger clusters which appear on this map as closely grouped points. The more "pure" information science clusters fall on the right and the lower middle portion of the map. These include "Precis", "theory of indexing", "on-line retrieval", "Bradford's law", and "serials planning". The less pure clusters are located toward the upper left and include a grouping of "sociology of science" clusters, "citation analysis", "Zipf's law", "information theory", and "fuzzy systems". The last two are the most mathematical of the clusters in the set. "Citation analysis" and "Bradford's law" appear to play mediating roles between the more pure information science clusters and the sociologically oriented areas.

Another interesting question concerning these selected clusters is the extent of their linkages to areas outside the information science set. This should roughly parallel their per cent concentration in the journal set since the latter presumably measures the cluster's degree of interdisciplinarity. Table 6 shows the ratio of linkages a cluster or cluster group has to other clusters within the information science set or outside that set. For example, the group of three

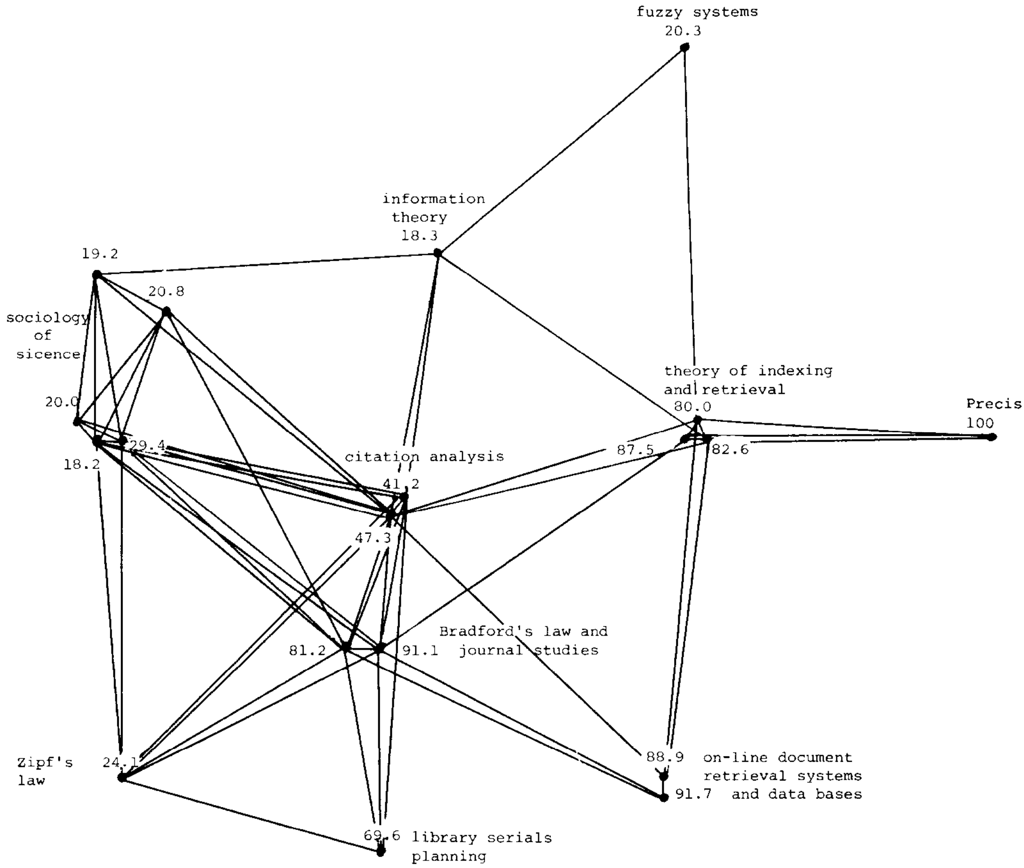

Fig. 2. Multidimensional scaling of information science clusters (clusters with more than 10% citations in information science journals).

clusters on "theory of indexing and retrieval" has six of its eleven linkages to clusters within the information science set, and five linkages to clusters outside the set. Similarly, "on-line retrival" and "Bradford's law" have a majority of their links to clusters in the set. Hence, these are inward looking clusters, By contrast, "fuzzy systems", "information theory", and "citation analysis" are outward looking clusters having a majority of their links to clusters outside the set. As expected, these in/out ratios are closely correlated with the percentage concentration of articles in the journal set. These results strongly suggest that information science is not entirely separate from the fabric of the social sciences as a whole, and that it might be interesting to systematically explore the linkages which lead outside the set of selected information science clusters.

Table 5. Comparison of information science clusters with clusters in other fields

	Information Science	Sociology	Economics	Psychology	Particle Physics
Number of clusters	13	36	44	277	19
Mean size (cited items)	2.8	5.5	5.3	9.1	12.4
Mean size (citing items)	29.3	69.1	65.9	96.8	153.8
Mean year of publications of cited items in clusters	1967.9	1966.6	1968.4	1968.8	1972.4
Percentage of cited items that are books	13.9	39.0	24.5	14.5	0.9
Percentage of cited items published in last 5 years (Price's Index)	47.2	46.5	52.8	57.9	72.9

THE NETWORK OF CLUSTERS AROUND INFORMATION SCIENCE

Up to this point we have examined the information science clusters identified by journal concentration in isolation from all other clusters in the social sciences. This isolation is of course not the case. To explore these external connections we use the full index of inter-cluster linkages, specify a starting cluster and an inter-cluster link threshold. The starting cluster was No. 563 "theory of indexing and retrieval," but it is important to point out that the final network obtained is independent of starting point. All linkages from the starting cluster to any other clusters in the file which had summed normalized co-citations of one or more were followed. These clusters were in turn treated as starting clusters, and their strong links followed. The resulting network is shown in Fig. 3, where clusters are represented as circles containing a cluster identification number and the lines connecting them represent summed co-citation links of one or more.

The first finding was that we eventually reached a number of points where the number of linkages leading to other clusters became very large and, in effect, exploded into other larger networks. These are indicated on the Figure by short lines emanating from the cluster. Upon inspection of the document titles in these clusters, it was evident that one set of clusters was branching off into social psychology while another set was leading into sociology. Links from all the other clusters on the network were exhaustively followed and many of these clusters turned out to be identical to ones identified by the independent measure of journal concentration. In fact 16 of the 22 clusters selected by having 10% of their citing articles in the

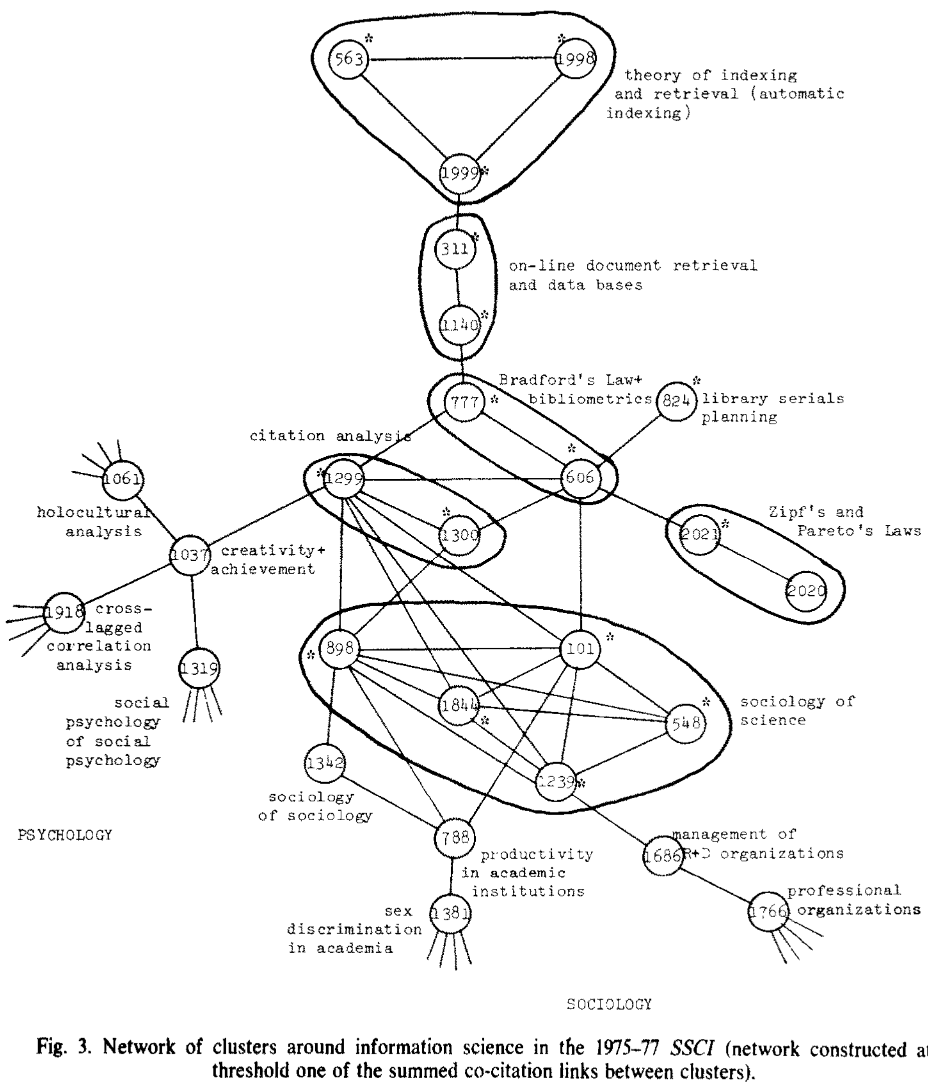

Fig. 3. Network of clusters around information science in the 1975–77 *SSCI* (network constructed at threshold one of the summed co-citation links between clusters).

Table 6. In/out ratios for selected clusters and cluster groups

Cluster or Cluster Group(cluster numbers)	Ratio In	Out
1. on-line document retrieval and data bases (311,1140)	5/6 (.83)	1/6 (.17)
2. Bradford's law and bibliometrics (606,777)	9/13 (.69)	4/13 (.31)
3. theory of indexing and retrieval (563,1998,1999)	6/11 (.55)	5/11 (.45)
4. citation analysis (1299,1300)	13/42 (.31)	29/42 (.69)
5. information theory (2013)	5/33 (.15)	28/33 (.85)
6. fuzzy systems (406)	2/22 (.09)	20/22 (.91)

information science journal set, appear on the network. (Their cluster numbers are starred on both Table 3 and Fig. 3). Furthermore, all clusters on the network which were selected by journal concentration are contiguous on the network, i.e. are not separated from one another by non-selected clusters. This result corroborates our selection of the journals for the information science set in the sense that the literature appearing in these journals behaves as a coherent whole.

Information science, as represented by this network, appears poised somewhere between psychology and sociology, with a very strong link to sociology via the sociology of science, and a more tenuous link to psychology through a cluster called "creativity and achievement". At the same time, information science, at least in the context of the social and behavioral sciences, appears somewhat isolated. It certainly is not the central discipline, with strong linkages to many diverse fields, that many would like it to be.

One way to suggest possible areas with which information science could strengthen its intellectual ties is to examine some of the weak links its clusters have with other clusters which are not in the field. A few examples will suffice to give an idea of the possibilities for inter-disciplinary linking that exist. "Fuzzy systems" has weak links with a Bayesian decision making cluster and a small linguistics cluster. "Copyright law" has a weak link with a legal cluster on the right to privacy. "Information theory" weakly links with diverse areas such as information processing in schizophrenia, urban modelling, and Bayesian estimation. Finally, citation analysis is linked to multidimentsional scaling, one which I have exploited in this paper. It is, of course, not surprising that such weak links exist, and information scinence is by no means unique in this respect. What is more surprising is the paucity of such links when the field is compared with others in the social sciences.

CONCLUSIONS

Our analysis has been descriptive, and to describe the state of a field is not to say what it should be. However, some observations of an evaluative kind seem in order. My purpose is not to argue that information science is a social science, although no one would want to deny the relevance of social and behavioral sciences for a field which deals with human beings generating and interacting with information. Rather my purpose has been to examine the place of information science within the social sciences, and its internal and external connections. What the field's position is with respect to the natural and physical sciences remains to be seen. For example, its connections with computer sciences remain for future study[13]. A second caveat is that we have examined the state of information science in only one time period (1975–73) providing a kind of snap shot of the field, but not revealing its dynamics. A later report will deal with the way the clusters have changed over the last nine years (1969–77) and will hopefully provide insights on how and why the field has changed.

It appears from the present data that information science is relatively isolated from the intellectual framework of the social and behavioral sciences. This may make it difficult for the field to contribute to social science knowledge generally and to adequately draw upon and utilize the findings generated by other social science fields. One possible strategy for better integration with the social sciences is to foster and encourage the strengthening of weak links which already exist between information science and other fields.

It is, however, equally possible to argue from another, perhaps not contradictory, position that information science should concentrate its efforts on building its own internal structure of knowledge and achieve intellectual self-sufficiency. Here we should perhaps look to the highly cited works in our core clusters as the pillars on which to build the next generation of significant findings. In principal, there is no reason why this building from within on the best of past work could not go forward at the same time that interdisciplinary links are being forged and strengthened. Only one path seems barred to us: to radically alter or shift the structure of the field from what it is to something entirely different. Revolutions, if they occur, must come from within, as part of the natural evolution of the field.

REFERENCES

[1] H. SMALL, Cited Documents as concept symbols. *Social Studies of Science* 1978, **8**, 327–340.

[2] T. SARACEVIC, Five years, five volumes, and 2345 pages of the annual review of information science and technology. *Inform. Stor. Retr.* 1971, **7**, 127–139.

[3] J. C. DONOHUE, A bibliometric analysis of certain information science literature. *J. ASIS*, 1972, **23**, 313–317.

[4] G. SALTON, On the development of information science. *J. ASIS*, 1973, **24**, 218–220.

[5] A. D. PRATT, Libraries, economics and information: recent trends in information science literature. *College Res. Libs.* 1975, **36**, 33–38.

[6] B. C. GRIFFITH and H. SMALL, The structure of the social and behavioral sciences literature. *Proc.* 1 *Int. Conf. on Social Studies of Science* (Cornell University) 4–6 Nov. 1976.

[7] H. SMALL and D. CRANE, Specialities and disciplines in science and social science: an examination of their structure using citation indexes. *Scientometrics* 1979, **1**(5–6), 445–461.

[8] H. SMALL, Structural dynamics of scientific literature. *Int. Classiffication* 1976, **3**, 67–74.

[9] P. H. A. SNEATH and R. R. SOKAL, *Numerical Taxonomy*, p. 131. (Freeman, San Francisco. 1973).

[10] J. A. HARTIGAN, *Clustering Algorithms*, p. 199. Wiley, New York (1975).

[11] D. PRICE, Citation measures of hard science, soft science, technology and nonacience. In: *Communications among Scientist and Engineers*, (Edited by C. E. NELSON and D. POLLOCK). D. C. heath, Lexington, Mass. (1970).

[12] J. B. KRUSKAL, Multidimensional scaling by optiminzing goodness of fit to a nonmetric hypothesis. *Psychometrika* 1964, **29**, 1–27.

[13] G. SALTON and D. BERGMARK, A citation study of computer science literature. *IEEE Trans. Professional Commun.* 1979, **PC-22**(3), 146–158.