To Contents Page

Epilogue: The Future of Citation Indexing

The primary focus of this book is the past and present of citation indexing. But what about the future?

Citation indexing promises to be at least as robust in the future as it is in the present. The popularity of citation indexes as a tool for conducting retrospective searches of the scientific literature will continue to increase, and probably at an accelerated rate.

The acceleration will powered by several factors. The increasing availability of on-line searching facilities throughout the world will place the *SCISEARCH®* * system at the fingertips of most working scientists and scholars. Improvements in on-line software will increase the use of *SCISEARCH* by those already familiar with it.

The increased awareness that citation indexes are used as part of the research evaluation process will induce more scientists to learn how to use *SCI*. Hopefully, good refereeing will prevent those who would from compromising legitimate reference practices.

Access to the *SCI/SSCI* data bases will become more convenient. Microstorage techniques or minicomputers could eventually make it economically practical for individual scientists and departmental groups to have their own copies of the data bases. Alternatively, the increase in accessibility will come from continuing reductions in the cost of providing on-line, remote links to data bases stored in central computers.

In addition, ISI is thoroughly examining the idea of producing disciplinary citation indexes whose source coverage would be based on the journal literature of a single discipline. However, the cited and citing material would reflect the full scope and diversity of interests in the field. This would include all the references by the publishing authors of the discipline, as well as all citations to the field by authors outside the discipline.

The critical problem that must be solved to do this remains the one faced in the design of the *Genetics Citation Index* (Chapter Two): how to define the literature in a way that minimizes irrelevant material without interfering with the ability of a citation index to reflect the disciplines's interaction with other disciplines. One way, tested in the *Genetics Citation Index* project, is to define a subset of cited authors in the combined multidisciplinary *SCI/SSCI* data base that meets some disciplinary criteria and then from that to work back to the source material (Chapter Two). Another way, suggested by the journal citation studies done in recent years (Chapter Nine), is to use the *Journal Citation Reports* to identify the source journals of the discipline first and then use straightforward methods to compile a citation index to the material they publish. Either way, the sharper focus of disciplinary citation in-

*A registered trademark of the Institute for Scientific Information.

dexes will produce search tools priced and sized to the requirements of departmental and personal libraries.

Such disciplinary indexes would be cumulated for long periods. For example, a geosciences citation index for 20 years is now under evaluation. Although some fields have a greater dependence on the older literature than others, a large-scale multiyear cumulation provides any field with an important historical perspective that is otherwise difficult to obtain. That is why I have set an ISI goal of extending the *SCI* back through the first 60 years of the twentieth century.

Not only will the use of citation indexes for literature searching accelerate within the scientific community, but it also will spread to the arts and humanities community. That has already begun to happen as the scholars in that community begin to respond to the *Arts & Humanities Citation Index.*(1)

The use of citation data to measure performance and for historical and sociological studies will also accelerate and spread (2). Despite the controversy that presently surrounds some of the applications of citation data (Chapter Ten), it has become a standard tool for exploring the social systems of science. At the present time, the literature on citation studies is growing quite rapidly. At least 100 papers per year are published—most independent of the various ISI studies published in *Current Contents* and elsewhere.

When Cole, Rubin, and Cole studied the peer review procedures on which the National Science Foundation bases its grant awards, citation counts were one of the variables that were examined to determine whether grant proposals were being prejudged (3). The fact that they were included suggests that they are increasingly accepted as a legitimate indicator of research significance. The fact that the citation counts were found to have an insignificant correlation with actual grant awards demonstrates that a potential for abuse is not equivalent to actual abuse. I hope this will not change in the future.

ISI is working to make citation data as precise as possible. The recent publication of the first comprehensive comparison between primary-author and all-author data (4) was an important step in this direction. This data not only showed how careful one must be in compiling citation data, but it also demonstrated some of the subtleties of interpreting the data. Consider the matter of high rates of coauthorship. In the case of a science administrator who coauthors numerous papers with members of the research staff, a very high coauthorship rate raises questions about the ethics of the authorships. In the case of a teacher who coauthors papers with numerous graduate students, a very high coauthorship rate suggests a measure of the teacher's impact. Both types of cases were found in the data.

Several key developments presage a substantial expansion of the use made of citation data for sociological studies. One is an NSF grant to ISI to compile a citation index of the 1920s journal literature of physics. The data base will be used to determine whether citation analyses can provide some new and sociologically useful insights into a decade of activity that has come to be called the "golden age of modern physics." Plans call for the eventual development of citation networks (5–8) and maps of co-cited clusters (9) from the data base that historians and sociologists can study.

Of even greater importance to the scholarly community is the recent agreement signed by ISI and NSF. This agreement gives any NSF-supported scholar access to the complete *SCI* data base for program planning, science policy studies, and so on. I am also hopeful that a similar agreement will be made with most foreign nations. The Japanese Ministry of Education has already leased ISI files and I expect the U.S.S.R. will do so in the near future. The National Research Council of Italy has been using *SCI* files for a variety of information purposes for several years.

Another development important to the expanded use of citation analysis is the *Arts & Humanities Citation Index* (1). The concept of producing a citation index to the arts and humanities literature dates back to at least 1955. In November of that year, I presented a paper at the American Documentation Institute Annual Meeting in which I outlined the concept of a citation index to the Bible that would be useful in studying the interaction between science and the humanities (10). The *Arts & Humanities Citation Index* should facilitate and stimulate new studies of that critical but little understood relationship.

A fourth development that can be expected to increase the sociological impact of citation analyses is Price's general theory of cumulative advantage processes (11). Cumulative advantage processes are ones that operate in situations where success breeds success. Merton (12) and then Cole and Cole (13) have described such processes at work in the social systems of science. Price's general theory for these processes proposes a statistical model that predicts the distributions they produce. What makes this highly significant is that the distributions predicted by the model fit those derived from such empirical laws as the Lotka distribution for scientific productivity, the Bradford law for journal use, the Pareto law of income distribution, and the Zipf law for literary word frequencies. What Price calls the "cumulative advantage distribution' also fits the empirical results of citation-frequency analysis. The theory, then, appears to provide a unifying conceptual framework for all of the empirical laws and citation data that make up the study of bibliometrics.

This promises to have a major impact on both the field of bibliometrics and on the use of citation data within the field. For one thing it appears to represent an important step forward in the attempt to determine what it is that citation data and other bibliometric measures define. It also should provide the type of conceptual foundation needed to increase the rate at which useful empirical generalizations and underlying theories are derived from the rich lode of citation data that exists. Although it is not possible to predict the outcome of such advances, it is reasonable to speculate that they could have a very significant impact on the practice of science.

As long as scientists and scholars continue to use the instrument we call "papers" as a primary communications medium, citation indexing and analysis will play an increasingly significant role in the management of mankind's knowledge and the processes by which that knowledge is produced. The future of the scholarly paper has often seemed in jeopardy. Whether the technological changes available to the next generation of scholars will undermine the role of the paper in the process of scholarship remains to be seen. My innate optimism gives me hope that it will not, and that citation indexing will have an increasingly strong, positive influence on scholarship.

References

1. **Garfield, E.** "Will ISI's *Arts & Humanities Citation Index* Revolutionize Scholarship?" *Current Contents* No. 32:5-9 August 8, 1977.
2. **Garfield, E. Malin, M.V. and Small, H.** "Citation Data as Science Indicators." In Elkana, Y., Lederberg, J., Thackeray, A. Merton, R.K., and Zuckerman, H. (eds.). *Toward a Metric of Science.* (New York: John Wiley & Sons, Inc., 1978). Pp. 179-207.
3. **Cole, S. Rubin, L. and Cole, J.R.** "Peer Review and the Support of Science." *Scientific American,* **237**(4):34-41, 1977.
4. **Garfield, E.** "At Last: The 300 Most Cited Authors, 1961-1976, Including Co-Authors." *Current Contents,* No. **28**:5-17, July 10, 1978.
5. **Garfield, E.** "Citation Indexes in Sociological and Historical Research." *American Documentation,* **14**:289-291, 1963.
6. **Garfield, E. Sher, I. H. and Torpie, R.J.** *The Use of Citation Data for Writing the History of Science.* (Philadelphia: Institute for Scientific Information, 1964). 86 pp.
7. **Garfield, E.** "When Citation Analysis Strikes Ball Lightning." *Essays of an Information Scientist,* Vol. 2 (Philadelphia: ISI Press, 1977). Pp. 279-490.
8. **Cawkell, A. E.** "Search Strategy, Construction, and Use of Citation Networks With a Socio- Scientific Example: Amorphous Semiconductors and S.R. Ovshinsky." *Journal of the American Society for Information Science.* **25**:123-130, 1974.
9. **Small, H. and Griffith, B.C.** "The Structure of Scientific Literatures. I: Graphing Specialities." *Science Studies,* **4**:17-40, 1974.
10. **Garfield, E.** "Citation Indexes—New Dimension in Documentation (Citation Index to tne Old Testament)." Paper presented at American Documentation Institute Annual Meeting, Philadelphia, November 1955.
11. **Merton, R.K.** "The Matthew Effect in Science." In Storer, N.W. (ed.). *The Sociology of Science* (Chicago: University of Chicago Press, 1973). pp. 439-459.
12. **Cole, S.** "Professional Standing and the Reception of Scientific Discoveries." *American Sociological Review,* **76**:286-306, 1970.
13. **Zuckerman, H.** "Stratification in American Science." *Sociological Inquiry,* **40**:235-257, 1970.
14. **Zuckerman, H. and Merton, R.K.** "Age, Aging, and Age Structure in Science." In Riley, M.W. et al. (eds.) *A Sociology of Age Stratification* (New York: Russell Sage Foundation, 1972). pp. 292-256.
15. **Cole, J.R. and Cole, S.** *Social Stratification in Science* (Chicago: University of Chicago Press, 1973).
16. **Allison, P.D. and Stewart, J.A.** "Productivity Differences Among Scientists: Evidence for Accumulative Advantage." *American Sociological Review,* **39**:596-606, 1974
17. **Zuckerman, H. and Cole, J.R.** "Women in American Science." *Minerva,* **13**:82-102, 1975.
18. **Zuckerman, H.** *Scientific Elite: Nobel Laureates in the United States* (New York: Free Press, 1977). Chapters 3 and 8.
19. **Price, D.J.D.** "A General Theory of Bibliometric and Other Cumulative Advantage Processes" *Journal of the American Society for Information Science,* **27**:292-306, 1976.

To Contents Page

To Cited Author Index