Chapter Two

A Historical View of Citation Indexing

As with so many other developments, the concept of indexing the scientific literature by citations grew out of an effort to achieve something else. Soon after World War II, federal expenditures in research and development turned sharply upward, beginning a trend that was to continue for two decades. Responding to the stimulus of government spending, the pace of scientific research accelerated and the scientific literature, which reported the activity, increased at an explosive rate. The government, concerned that the systems for information exchange among scientists would be unable to handle the growing volume of literature, sponsored a variety of projects aimed at expanding and improving the facilities and methods for distributing and managing scientific information. One such project, sponsored by the Armed Forces Medical Library (now the National Library of Medicine), was a study at the Johns Hopkins Welch Medical Library on the role machines might play in generating and compiling indexes to the medical literature (1). As a member of the study team, I became interested in whether and how machines could be used to generate indexing terms that effectively described the contents of a document, without the need for the intellectual judgments of human indexers. The concept of a citation index, which eliminates this need, was a product of this interest. As it turned out, machines, per se, were not required.

BITS AND PIECES

The concept of a citation index to the scientific literature was synthesized from bits and pieces of experience and insight accumulated over a period of a few years. The accumulation started with two years of work on the Welch Medical Library project. The study of indexes produced insights into the nature of the intellectual judgments made in the indexing process and the functional relationship between the index and the person using it. There were additional insights into the structure and functional characteristics of both alphabetic and classified subject indexes. While I was doing voluntary abstracting work for *Chemical Abstracts* I recognized that the utility of abstracts was increased by including in them references to abstracts of key papers cited by the authors. This recognition was followed by a recommendation that *Chemical Abstracts* develop an index from their "C.F." ("see also") statements. More importantly, at the suggestion of Chauncey D. Leake, I made a special study of review articles. They play a special role in providing scientists with guideposts to the literature. I soon came to the realization that nearly every sentence in such review articles was supported by a reference. And the sentence itself was an unusually definitive indexing statement of the cited work. And in 1953 I learned, through William C. Adair, a former vice president of *Shepard's Citations,* that there was an index to the case literature of the law that used citations.

Shepard's Citations is the oldest major citation index in existence; it was started in 1873 to provide the legal profession with a tool for searching legal decisions. It does this by listing the citations to precedents used in the cases decided by federal and state courts and various federal administrative agencies. The legal "citator" system provided a model of how a citation index could be organized to function as an effective search tool. It should be noted that at that time it was not an index to legal periodical literature as it is today.

Other bits and pieces came, over the next few years, from the formal study of library sciences and structural linguistics. At first, it seemed that the basic problems to be solved in machine indexing were the ones of identifying the key words in a paper and stringing them together to form useful indexing terms or phrases. Studies of various parts of the literature, primarily in the disciplines of chemistry and medicine, and the indexes to them showed that the problem was more complicated than that.

It is not particularly difficult to identify key words.Theoretically, the simplest approach to generating index headings by machine would be to match the words of a text against a dictionary of indexing terms. This procedure would produce a large number of useful indexing entries; unfortunately, however, it also would produce an even larger number of entries whose relevance to the subject matter of the document was superficial at best. The "noise" created by the large number of irrelevant entries would be so great as to make the index just about useless—all of which means that selectivity is an important characteristic of the indexing process.

Limiting the vocabulary analysis to titles or abstracts would seem to be a way of achieving the degree of selectivity desired. Presumably titles and abstracts deal with only the essence of their documents; and, in fact, studies performed as part of the Welch Library project produced statistical evidence to support the theory that the nature of their selectivity is quite close to the selectivity practiced in the process of indexing by subject headings. Using detailed indexing records of *Chemical Abstracts* and the *Current List of Medical Literature* (predecessor to *Index Medicus*), we found that a significant number (60 to 90%) of the terms used in a given index appear in, or are implied by, the titles and abstracts of the documents covered.

But there is another side to that coin. Titles, by themselves, do not provide all the terms used in an index. Abstracts supply some of the terms missed by the titles but still do not match the full power of the indexing vocabulary, and what they contribute in added detail is gained at the expense of losing some of the selectivity so important to an index (2). Not all key words appearing in the titles and abstracts are useful indexing terms for those particular documents. Elaborate indexing entries are not always paraphrases of sentences in the abstract, and not all significant points covered by an abstract necessarily end up being indexed.

These findings led the research back to the full, original text as the best starting point for an automated indexing process. Since earlier work had shown that simple vocabulary analysis was not selective enough to extract only the important material, obviously, more powerful analytical techniques were needed. Structural linguistics seemed to offer some (3, 4).

The analytical methods of structural linguistics are based on syntax, the grammatical relationships between the words of the sentence. Used mostly in research on mechanical translation and other types of text-analysis problems (5), some of them appeared to offer capabilities close to what is needed for automated indexing. For example, they seemed to be capable, in theory at least, of emulating the classification judgments made in cataloging material for a library, and they provided a way of reducing the text of a document to a statistical summary of classifications that could conceivably be useful in making human indexers more consistent and thorough.

Nevertheless, though syntactic analysis went farther than vocabulary analysis, it did not go far enough. Unable to make qualitative distinctions, it was insensitive to the differences between old and new information and important and trivial information—the qualities on which an indexer's intellectual judgments are largely based.

In addition, even the limited potential of syntactic analysis in an automated indexing process was more theoretical than real. Before the techniques could be automated, someone had to figure out some way for the machine to determine whether the word "paid" in the sentence "They have paid witnesses" is being used as part of the verb phrase "have paid" or the noun phrase "paid witnesses." At the time, this unsolved problem was the main obstacle in the way of automated translation, and it was just as important to the use of syntactic-analysis techniques in automated indexing. Whether the machine was supposed to translate or index, it had to first understand, unambiguously, what it was reading. Since that time, considerable progress has been made in solving these kinds of problems. For one thing, it has been found that when a machine is dealing with a collection of documents, it can make these judgments with a reasonable degree of accuracy by looking at the frequency with which a given phrase occurs.

A SHIFT IN OBJECTIVES

During the accumulation of these bits and pieces, it became obvious that a human indexer does considerably more than identify and string together key words. In assigning terms, a good indexer makes subtle judgments about the relative quality of information and identifies relationships between concepts that go beyond the explicit content of the document being indexed. If it were to succeed, automated indexing would have to match that capability.

There was a question, however, about whether the capability was worth matching. It was difficult, if not impossible, to study the indexing process thoroughly without recognizing that it had a number of shortcomings that drastically limited its utility. One of the major ones was the practice of building indexes around the viewpoints and terminology of individual disciplines, when research was growing increasingly multidisciplinary, and the need for bibliographic tools that stretched across disciplinary lines was growing apace. How could a subject index produced by specialists in a particular discipline from the specialized vocabulary of that discipline be made to interpret material from the viewpoints of all other disciplines?

The question was not original. It was implied by those who advocated more comprehensive indexes (ones that covered more disciplines), more classified indexes, and standardization of terminology and nomenclature—all measures aimed at improving the conventional subject index. Unfortunately, however, as useful as all these measures may be, none of them solve the basic problem of interpreting the work in one subject from the viewpoint of all other subjects.

More comprehensive subject indexes can be nothing more than collections of a number of individual subject indexes, tied together by some central cross-referencing mechanism (6). Classified indexes are nothing more than subject indexes that are arranged hierarchically rather than alphabetically; though they may provide more logical insights than alphabetic indexes, they do not offer any broader interpretation of the material. And terminology and nomenclature reflect, for the most part, the specialization of disciplines. The most that standardization in this area can achieve is to ease communication at the points where disciplines meet; it cannot eliminate the differences in viewpoints, which are inherent and real.

The greatest need, apparently, was for some new type of index that would do a better job of spanning the gap that often existed between the viewpoints of those reporting a particular bit of information and those searching for it. Within this context, a citation index, produced by any means, began looking more useful than a subject index that could be produced by machine.

What looked best about a citation index was the diversity of the insights it provided about the literature of a particular subject and the efficiency and stability with which they could be stated. By using author references to index documents, the limited ability of a subject indexer to make connections between ideas, concepts, and subjects was replaced by the far superior ability of the entire scientific community to do the same thing. This meant that a citation index would interpret each of the documents it covered from as many viewpoints as existed in the scientific community. If an index is looked at as an attempt to represent as much detail of the real world as possible, a citation index would be to a conventional subject index what a full-color photograph would be to a black-and-white line drawing.

The other attraction of using reference citations in place of subject headings was their efficiency and semantic stability. A single reference citation represents as many subject headings as scientists have reasons for citing it. For example, the reference citation KNOWLES W.S., *CHEM TECH*, 2, 590, 1972 represents the subject heading ASYMMETRIC HYDROGENATION because some author cited it for the material it had on that subject. Another author cited it for material on the use of phosphine catalysts, so it also represents the subject heading PHOSPHINE CATALYSTS for anyone using it to find material on that subject. It also has been cited for material on alpha-amino acids, olefin hydrogenation, and ligands, which means that it functions as a heading for all those subjects. And if more people cite it for material on other subjects, the Knowles citation will take on those meanings too. In other words, each reference citation is associated with as many subject meanings as other scientists attribute to it (7, 8).

The semantic stability of citations comes from the fact that KNOWLES W.S., *CHEM TECH*, 2, 590, 1972 identifies a complex set of ideas with a relatively stable meaning, regardless of any changes that may take place in the terminology used later on to describe that subject.

A LOOK AT PRACTICALITY

Despite these kinds of advantages, there were questions to be answered about citation indexing too. Most of them had to do with the practicality of a citation index to the scientific literature. What form should it take? How should it be produced? How big should it be? Should it be selective or comprehensive in terms of disciplines covered?

Most of these questions were answered at a theoretical level in a paper published in *Science* in 1955 (8) in which a citation index to the literature of science was visualized as following the model of *Shepard's Citations.* A numerical coding system would be used to identify all citations, both reference (cited) and source (citing). At the front of the index would be an alphabetic listing of all journals covered, along with their code numbers. The main portion of the index would consist of a coded listing of reference citations, arranged in numerical order, and the source citations for each. Next to each source citation would be some designation showing the nature of the source: original article, review article, abstract, patent, translation, note, etc.

Such an index was visualized as being 50 to 100 times bigger than *Shepard's Citations* if it were to cover all of the journal literature of science. At the time *Shepard's* contained some 1.1 million citations from 30,000 cases a year. It was suggested, however, that not all journals had to be covered, and that by being selective in the coverage, the number of source citations listed each year could be brought down to about 1 million, which was well within the realm of practicality.

Production did not seem to pose any particular problem. The biggest job would be the coding of all the journals and articles covered. Then, following the *Shepard's Citations* model again, a separate index card would be prepared for each reference made in each source article. The cards would show, in code, the article being cited, the source article, and the nature of the source article. The rest of the process could be automated, using tabulating equipment, first, to put the cards into the numerical order of the reference citations and, then, to sort the cards associated with each reference citation into the numerical order of the source citations. After that, the cards would be run through a printer to prepare a master copy from which reproduction-quality page proofs could be prepared. No special subject expertise would be required, and, except for the supervisory level, neither would any special bibliographic expertise.

The utility of a citation index to the scientific literature was explored in a very tentative way by compiling all the references made in the *Journal of Clinical Endocrinology* during a five-year period to the notable paper by Hans Selye on the "General Adaptation Syndrome." Twenty-three references to the paper were found by scanning 500 articles published in the journal during that time.

An analysis of the articles that cited the Selye paper showed a number of interesting things. Although all of the articles were indexed in the *Quarterly Cumulative Index Medicus,* none of them was listed under the subject heading ADAPTATION; no indexer had made the connection or, at least, no one thought that the connection was worth identifying. The subject matter of the citing articles was extremely varied. And a number of the articles confirmed some of Selye's claims.

All of these findings said something positive about the utility of a citation index. It led to papers that would have been missed searching under the most obvious subject heading in a conventional subject index. The range of papers found provided an objective measure of the impact of the Selye paper, and the papers found provided documentation of the confirmation of some of Selye's claims.

In all, this tiny citation index showed that the technique was capable of identifying useful aspects of a paper that were missed by the conventional indexing services. If the book is considered the macro unit of scientific thought, and the journal article regarded as the micro unit, as had been suggested by Ranganathan, the citation index seemed to reach down to the molecular level, which certainly was a useful degree of specificity that was beyond the economic, if not conceptual, capabilities of the subject indexes.

PILOT TESTS

During the late 1950s and early 1960s a series of more elaborate citation indexes were developed that tested the feasibility and utility of the idea more thoroughly. In 1959, the *Journal of the American Statistical Association* published a cumulative citation index to its volumes 35 through 50. The *Annals of Mathematical Statistics* did the same thing in 1962 for its first 31 volumes. Also published in 1962 was a citation index to the *Bibliography of Non-parametric Statistics.* All three of these indexes were one-time efforts that were selective in the references listed. The indexes to the two journals listed only references to the journals; the one to the bibliography listed only references to other items in the bibliography.

What turned out to be the mainstream of development, because it eventually led to the only continuing, comprehensive citation index to the full spectrum of the scientific literature, were two pilot tests conducted by Eugene Garfield Associates, the consulting firm that later became the Institute for Scientific Information® (ISI®)*. The first, conducted in the mid-1950s, involved the development of a citation index to 5000 chemical patents held by two pharmaceutical companies (9). The reference citations were the prior patents cited by examiners to support their decisions (mostly involving rejection or restriction) on applicant claims. In all, some 30,000 source citations to approximately 20,000 patents were compiled.

The connections made by the index were analyzed and compared with the classification decisions made by the Patent Office and the indexing decisions made by *Chemical Abstracts* on the same patents. The analysis showed that the citation index consistently identified subject matter in the source patents that did not show up in either the patent classifications (which, it should be recognized, are based solely on claims that have been granted) or *Chemical Abstracts.*

This test too, then, confirmed the potential utility of a citation index, based on its ability to identify details missed by other systems. The capability was particularly intriguing as it applied to patents, because the information identified by the citation index had to do mostly with the reasons for making negative decisions. It would seem that the ability of a citation index to reach across the lines of patent classifications to identify reasons for rejecting or limiting a particular claim could improve considerably both the effectiveness and the efficiency of the examination process. And if the index included the file of patents that had been completely rejected, which are not classifed, it would be even more useful.

The 1960s test of the value and practicality of a citation index to the scientific literature was on a considerably larger scale, and of broader significance. Working under a grant from the National Institutes of Health, the Institute for Scientific Information produced a series of three pilot citation indexes to the genetics literature (10).

The program started out to test the feasibility and utility of a narrow, discipline-oriented citation index. Genetics was chosen to be the test discipline because it interacted with many other disciplines and, presumably, would gain the most from a bibliographic tool with the multidisciplinary reach expected of citation indexing. The initial plan was to index all the articles published in a list of hard-core genetics journals—ones whose titles included the words "genetics," "heredity," or comparable terms—that had been compiled by an advisory committee of geneticists working with the study group. This strategy made the first point in favor of a comprehensive citation index over a selective one by demonstrating the difficulty of defining a discipline by the primary journals that served it.

The year was 1961. Molecular biology was just beginning to be recognized as a new and important field of study, one that was having a particularly strong impact on genetics. The *Journal of Molecular Biology,* which had been in existence only since 1959, was not included on the list of hard-core genetics journals, but had published genetics papers. Other journals not included on the list also were publishing genetics papers: *Nature,* for example, had published the landmark paper by Watson and Crick on DNA. Obviously, some way had to be found to define the

**Registered trademark of the Institute for Scientific Information.*

genetics literature that would pick up these papers. The question was how to do it without opening the index to a lot of material that plainly did not qualify as genetics.

The answer took the form of a decision to compile a comprehensive, multidisciplinary data base of citations and to develop a set of criteria that would permit a computer to extract all the material that could be considered relevant to genetics. The rationale behind this approach was that by compiling a comprehensive, multidisciplinary data base, we would pick up all the material relevant to genetics, regardless of where it had been published, and then we could separate it from the rest by applying more elaborate criteria than would be practical in defining initial coverage.

Though it was neat, this resolution of the problem was still unsatisfactory in one sense: it left unanswered the original study questions about the practicality and utility of a selective citation index compiled in a more direct way. Consequently, it was decided to produce two more indexes based on the more limited material available in the list of hard-core genetics journals provided by the advisory committee of geneticists. One would cover the output of 38 journals for a period of five years, from 1958 through 1962. The other, to determine the effect of time span on the utility of a citation index, would cover the output of three hard-core journals over a period of 14 years, from 1949 through 1963.

The revised strategy broadened the scope of the study considerably—and with it, the number of things that could be tested. Besides testing the concept of a citation index, we also would be able to test the relative merits of selective versus comprehensive indexes, and the feasibility of using machine methods to define the literature of a given discipline.

Work started with the compilation of the multidisciplinary data base, which consisted of citations of all the material published in 1961 by the 600 journals covered by *Current Contents*® * and of all the references in that material. The source journals included general ones, such as *Science* and *Nature,* plus those specializing in a broad range of disciplines that included clinical medicine, experimental biology, instrumentation, physics, virology, and, of course, genetics. In all, the data base contained 1.3 million references to 890,000 unique authored items. The 890,000 items cited stretched back in time a full 100 years, though most of them were no more than 10 years old. The average number of references to each cited item was 1.52.

From this comprehensive data base, the computer extracted the citations that applied to genetics for the compilation of the *1961 Genetics Citation Index.* This subset amounted to 19% of the full data base, and consisted of 246,000 references to 146,000 items. The size of the subset was not left to chance. Economic considerations made it necessary to limit the subset to 20% of the data base, and the selection criteria were tailored to accomplish this. The procedure for extracting the subset is shown in flowchart form in Figure 2.1.

Three dictionaries were developed for the extraction process. One was a dictionary of geneticists compiled from a number of sources: the membership lists of the American Society of Human Genetics and the Genetics Society of America;

**Registered trademark of the Institute for Scientific Information.*

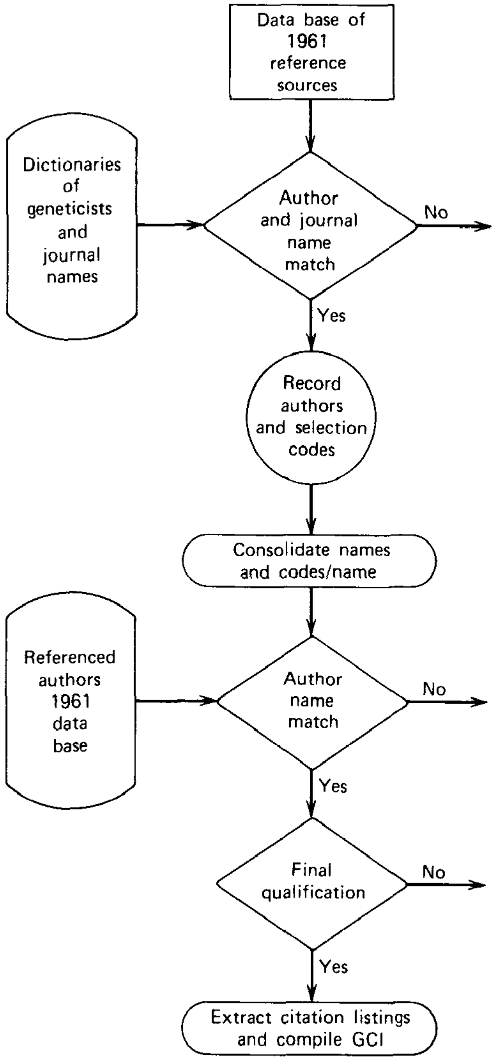

Figure 2.1 Procedure for extracting genetics subset from multidisciplinary data base.

the mailing list of the *Microbial Genetics Bulletin*; the "Bibliography of Human Genetics" from the *American Journal of Human Genetics*; bibliographies on "Mankind Evolving" by T. Dobzhansky, "Human Genetics" by R. R. Gates, and "Medical Genetics" by V. A. McKusick; and an unofficial list of geneticists supplied by NIH. These sources provided a total of 26,000 names.

The other two dictionaries were of hard-core genetics journals. One contained the 43 journals originally selected by the advisory committee of geneticists; the other contained an additional 28 genetics journals included in the source journals covered by the multidisciplinary data base.

The extraction process started with an analysis of the data base citations to determine which source and reference authors and journals were listed in the dictionaries. Every time a match was found, both the source and reference authors involved were recorded on magnetic tape, along with a code that showed the reason for selection.

All the codes listed for each author were then consolidated to produce a list of authors who met at least one of the following criteria:

- Were recognized geneticists.
- Had published in a recognized genetics journal.
- Had cited a recognized geneticist.
- Had cited a recognized genetics journal.
- Had been cited by a recognized geneticist.
- Had been cited by a recognized genetics journal.

This list of authors was run against the list of reference authors in the data base. The codes of those who were represented as cited authors were then checked to see who met the final selection criteria that had been set up to limit the genetics subset to 20% of the full data base. The basis for qualification was being a recognized geneticist, having published in a hard-core genetics journal, or meeting two of the other criteria concerning citing or being cited by a recognized geneticist or genetics journal. All the material pertaining to those cited authors who met at least one of those qualifications was then extracted from the data base and compiled to create a *1961 Genetics Citation Index.*

As was expected, this index provided a much more comprehensive view of the genetics literature than either of the other two. Though it was an index to only the material published in 1961, over 50% of the items cited were more than six years old.

Nevertheless, each of the other two indexes was interesting in its own right. Their relatively narrow scope of coverage was compensated for by the depth provided by their 5- and 14-year time spans. The 14-year span, which covered 1949–1963, provided a remarkable historical view of the field, even though it was based on the work published in only three journals.

CONCLUSIVE ANSWERS

The pilot study that produced the genetics indexes conclusively answered all the questions that had been surrounding the practicality of indexing the scientific literature by citations. It showed, for one thing, that the technique was economically practical. The major manual effort required to code all the entries, as *Shepard's Citations* did, turned out to be unnecessary. On the other hand, a great deal of attention had to be given to the problem of standardizing the author and journal names used in references and unifying all the variations of those names. Fortunately, it was possible to develop special computer programs that accomplished this unification with adequate, though not perfect, consistency. When all the pluses and minuses of citation indexing were added up and compared to conventional subject indexing, citation indexing turned out to be less expensive.

The study also showed that a citation index was eminently useful. A thousand copies of the three citation indexes to the genetics literature were published and distributed to scientists in the field for evaluation. Their responses, and an analysis of the genetics indexes and the multidisciplinary data base from which the 1961 index was derived, showed that a citation index produced a view of the literature that was deeper, more specific, and considerably broader in scope than conventional subject indexes. These qualities made it a more efficient and productive search tool.

The comparison between the genetics indexes and the multidisciplinary data base was particularly enlightening. Though the advantages of a citation index held up whether the index was narrow and selective or broad and comprehensive, they grew considerably more pronounced as the coverage of the index increased. In fact, a comprehensive, multidisciplinary index possessed a dimension of utility that went beyond its role as a search tool. It also provided a view of the literature that threw much light on such murky and important subjects as the bounds of particular disciplines, the interaction between disciplines, the structure of the journal network that is the primary mechanism for exchanging information in the scientific world, the historical development of scientific thought, and the implications and impact of individual pieces of scientific work.

These findings were compelling enough for ISI to decide to publish, on its own, after the government rejected the recommendation, the multidisciplinary data base as the first broad, comprehensive citation index to the scientific literature, and to produce it on a continuing, annual basis. Named the *Science Citation Index*® (*SCI*®)*(11), the index has steadily increased its coverage (see Figure 2.2) to the point where the 1977 edition contained 7.4 million references to 3.8 million items, reflecting the total literature published by 2655 journals and 1400 other sources reporting on all the disciplines of science in all the major scientific languages. In 1973, ISI brought out a companion *Social Sciences Citation Index* ™(*SSCI* ™)†(12). In 1978, it further expanded the application scope of citation indexing as a tool of scholarship by bringing out the *Arts & Humanities Citation Index* (A&HCI) ™†(13).

Probably because *SCI* is so broad and comprehensive, few additional citation indexes to the scientific literature have been published. One that has been is the *Citation Index for Statistics and Probability,* a cumulative one-time effort that covers the journal literature of the field from its inception, early in the twentieth century, through 1966. In addition, the *Journal of Histochemistry and Cytochemistry* published monthly, from October 1966 through September 1973, a citation index to articles published in some 2300 journals that had cited its material. In a sense, this index is a demonstration of the broad range of utility of the *Science Citation Index,* since it is based on information supplied from the *SCI* data base as part of an alerting service offered by ISI under the name *Automatic Subject Citation Alert*® (*ASCA*®)*(14).

The *Citation Index for Statistics and Probability,* compiled by Dr. J. W. Tukey of Princeton University, and published in 1973 as part of the "Information Access Series" of the R & D Press, provides comprehensive coverage (all published articles) of some 40 hard-core journals in the field of statistics and selective coverage of an

**Registered trademark of the Institute for Scientific Information.*

†Trademark of the Institute for Scientific Information.

	1970	1971	1972	1973	1974	1975	1976	1977
Source Journals	2,192	2,277	2,425	2,364	2,443	2,540	2,717	2,655
Source Journal Issues	17,992	18,976	19,384	20,493	20,719	21,390	22,697	22,480
Source Non-Journals								1,393
Journal Source Items								455,977
Non-Journal Source Items								38,884
Anonymous Source Items	11,320	9,639	10,191	9,806	8,096	7,284	6,405	6,578
Authored Source Items	350,555	354,851	367,423	397,137	392,875	411,617	438,146	488,283
Total Source Items	361,875	364,490	377,614	406,943	400,971	418,903	450,956	494,861
Citations to Authored Items*	4,041,165	4,302,885	4,579,183	4,938,132	5,148,630	5,446,889	6,080,275	7,271,526
Citations to Anonymous Items	55,357	61,037	62,884	61,989	65,265	70,332	74,208	95,065
Citations to Patents	11,425	15,783	17,048	17,299	17,815	18,747	22,070	31,435
Total Citations from Source Items	4,107,947	4,379,705	4,659,115	5,017,420	5,231,710	5,535,968	6,176,553	7,398,026
Unique Cited Authors*	619,872	645,505	688,320	710,992	730,001	772,500	812,974	907,517
Average Number of Citations To Cited Authors	6.52	6.67	6.65	6.95	7.05	7.05	7.48	8.01
Unique Authored Items Cited*	2,340,128	2,449,573	2,596,663	2,729,968	2,817,833	3,005,712	3,245,632	3,776,247
Average Number of Citations To Authored Cited Items	1.73	1.76	1.76	1.81	1.83	1.81	1.87	1.92

Figure 2.2 Growth of *Science Citation Index* over the years.

additional 100 journals through the first 66 years of this century. It contains some 300,000 references from approximately 35,000 articles and reviews.

A third citation index, this one outside of the scientific area, completed the cycle of interaction that led to the development of citation indexing as an established method of managing the scientific literature. In 1968, *Shepard's Citations* introduced the *Shepard's Law Review Citations.* This is a continuing index to more than 100 law reviews and journals that shows references to any legal article written since 1947. As the *Science Citation Index* was inspired in part by the model of *Shepard's Citations,* so was *Shepard's Law Review Citations* inspired in part by the model of *Science Citation Index.*

REFERENCES

1. **Field, H.G., Himwich, W.A., Garfield, E., Whittock, J.M., and Larkey, S.V.** "Final Report on Subject Headings and on Subject Indexing" (Baltimore: Welch Medical Library Indexing Project, Johns Hopkins University, 1955). 47 pp.
2. **Resnick, A.** "Relative Effectiveness of Document Titles and Abstracts for Determining Relevance of Documents."*Science,* **134:**1004–1006, 1961.
3. **Harris, Z.S,** *Structural Linguistics* (Chicago: The University of Chicago Press, 1951; 4th impression, 1960). 384 pp.
4. **Chomsky, N.** *Syntactic Structures* (The Hague; Mouton and Company, 1957).
5. **Sager, N.** "Syntactic Formatting of Science Information." *Proceedings of the Fall Joint Computer Conference, 1972,* **41:**491-800, 1973.
6. **Garfield, E.** "A Unified Index to Science." In *Proceedings of the International Conference on Scientific Information, Washington, D.C., November 16-21, 1958.* 2 Vols., 1637 pp. (Washington, D.C.: National Academy of Sciences—National Research Council, 1959). Vol. 1, pp. 461–474.
7. **Garfield, E.** "Citation Indexes—New Paths to Scientific Knowledge." *The Chemical Bulletin* **43:**11, 1956.
8. **Garfield, E.** "Citation Indexes for Science."*Science,* **122:**108–111, 1955.
9. **Garfield, E.** "Breaking the Subject Index Barrier: A Citation Index for Chemical Patents." *Journal of the Patent Office Society,* **39:**583–595, 1957.
10. **Garfield, E. and Sher, I.H.** *Genetics Citation Index; Experimental Citation Indexes to Genetics with Special Emphasis on Human Genetics* (Philadelphia: Institute for Scientific Information, 1963). 854 pp. Cf. introductory material, pp. i-xviii.
11. **Garfield, E.** "Science Citation Index—A New Dimension in Indexing." *Science,* **144:**649–654, 1964.
12. **Weinstock, M.** "ISI's *Social Sciences Citation Index*: A New, Comprehensive, Multidisciplinary Information Retrieval System." Paper presented at the National Convention of the American Society for Information Science, Washington, D.C., October 23–26, 1972.
13. **Garfield E.** "Will ISI's *Arts & Humanities Citation Index* Revolutionize Scholarship?" *Current Contents.* No. 32, 5–9, August 8, 1977.
14. **Garfield E.** *"ASCA (Automatic Subject Citation Alert).* A New Personalized Current-Awareness Service for Scientists." *The American Behavioral Scientist* **10:**29–32, January 1967.