The Creation of the *Science Citation Index*

Paul Wouters

NEWS Contact: Mrs. Joan E. Shook
RELEASE
INSTITUTE FOR SCIENTIFIC INFORMATION 33 SOUTH SEVENTEEN STREET PHILADELPHIA, PA phone/locust4-4400 cable/currcon twx/ph 803

For Immediate Release

$300,000 GRANT TO PROBE INFORMATION RETRIEVAL AWARDED TO INSTITUTE FOR SCIENTIFIC INFORMATION BY NATIONAL INSTITUTES OF HEALTH AND NATIONAL SCIENCE FOUNDATION . . .

THREE YEAR PROJECT TACKLES CITATION INDEX TECHNIQUES FOR SCIENCE

Research scientists will soon be consulting a more precise and specific literature index that links together subject material that would never be collated by usual indexing systems. Concerned with new starting points for scientific literature searches, the unique concept uncovers sometime-buried associations, relating important works and authors, yet keeps the researcher abreast of the masses of current published scientific information. This new approach to information retrieval is called the Citation Index.

A $300,000 grant extending over a three-year period has been awarded to the Institute for Scientific Information, Philadelphia, Pennsylvania, to study the practicability of citation indexes and to test their techniques of preparation. The project, under joint sponsorship of the National Institutes of Health and the National Science Foundation, is aimed at producing a unified citation index for science, including the publication of a genetics index.

Two years after the above press release, the *Genetics Citation Index* was published (Garfield & Sher, 1963). It was quickly followed by the first volume of the proper *Science Citation Index (SCI)* (Garfield, 1963). Since then, the *SCI* has become part of the world of science. This is not to say that the *SCI* was immediately applauded. Initial responses were mixed (Wouters, 1999). These mixed feelings were nevertheless far more positive than the reactions Eugene Garfield, the inventor of the *SCI*, had received in the preceding years. He had been actively propagating the idea of a citation index for science since he became acquainted with it in 1953. Hardly anyone had responded. Even a few years before getting the decisive grant from the National Institutes of Health (NIH) and the National Science Foundation (NSF), referees of his proposal were quite critical, some

even hostile. An undated and anonymous overview of the referees comments can be found in Eugene Garfield's personal archive in Philadelphia, Pennsylvania.

This resistance to the idea of investing in citation indexes of the scientific literature may come as a surprise to the present-day user of the *SCI* and *Social Science Citation Index*. After all, scientists must acknowledge their peers and must share their ideas and resources with their colleagues. Therefore, it seems rather obvious to use the footnotes of a scientific article or the bibliography of a book as an entry in a literature-searching procedure. The fact of the matter is, however, that the concept of the citation index did not come to science as naturally as this interpretation suggests. To start with, the *SCI* has its roots not in science but in law.

Stumbling over the Citation Concept

Citation indexes were already old hat for American lawyers when the history of *SCI* begins. In the second half of the nineteenth century Frank Shepard in Illinois deemed it useful to know whether a legal case was still valid. He produced gummed paper with lists of the cases that cited the case at hand. Lawyers in Illinois glued them to their dossiers so enthusiastically that Shepard set up a commercial business in 1873. His company, Shepard's Citations, Inc., had the monopoly of producing the one and only citation index "to serve the Bench and Bar." First in Chicago, later in New York, and in the 1950s in Colorado Springs, Colorado, a staff of highly qualified lawyers produced the Shepard's Citator by hand, covering all judicial decisions in the United States. Shepard's was a respectable firm, proud of its supreme reliability. Its product was grounded in the norms and procedures of the legal system. Shepard's was purchased in 1996 by Reed Elsevier and the Times Mirror Company. As William C. Adair, former vice president of the company, explained to the readers of *American Documentation* in 1955:

> The lawyer briefing a case must cite authorities to back up his arguments. So must the court in writing its opinions. This is because of the doctrine of "Stare Decisis," which means that all courts must follow precedents laid down by higher courts and each court generally also follows its own precedents. . . . The lawyer, however, must make sure that his authorities are still good law, that is, that the case has not been overruled, reversed, limited or distinguished in some way that makes it no longer useful as a valid authority. Here is where the use of Shepard's Citations comes in. (Adair, 1955)

The searching procedure was simple. First, the lawyer located a case similar to his own, then he looked up Shepard's to see whether later cases had cited it. He would immediately see whether the decision was still valid and which other cases had made use of it. A lowercase "r" before the case meant that it was reversed. Adair told his audience that important law suits were won "on the strength of a case located by the use of Shepard which no other method of research disclosed" (Adair, 1955). "Shepardizing" the legal literature was based, since 1873, on the authority-centered norms in the United States legal system: The most recent decision of the highest court is valid. The way of indexing by citation tied in perfectly with this value system. One can hardly think of a sharper contrast with supposedly ruthless scientific criticism. This hierarchical indexing style served nevertheless as the model for the Institute of Scientific Information's *Science Citation Index*.

Retired, running a cattle ranch in Colorado Springs, and still eager to act, William Adair sometime in 1953 read in the local newspaper that the scientific world "was being swamped in a sea of literature." It was a report on a conference organized by the Welch Medical Indexing Project at Johns Hopkins University in Baltimore, Maryland. This project had been sponsored by the Army Medical Library since 1948 (Larkey, 1949; Miller, 1961). The main task of the project was to find out whether, and if so how, machines could be used to improve the efficiency of indexing and retrieving medical literature. The indexing itself was supposed to be the good old subject indexing form. In this respect the Welch Medical Library was not very innovative. Within these boundaries, the staff had to devise new systems of indexing, subject headings, and using machines to solve "the literature problem." Adair wrote a letter to Sanford Larkey, supervisor of the project. He told Larkey about the citation indexing system, informing him that "if the whole body of American Law can be classified so that a knowledge of one case can be used as a key to locate all other cases in point, the same thing can be done with medical articles" (W. C. Adair to S. V. Larkey, personal communication, 10 March 1953). Adair offered his expertise: "I have retired from Shepard's and am now free to undertake and organize such a project." He received a reply from a 25-year-old junior member of the staff—Eugene Garfield—who did not know anything about citation indexing. He wrote Adair that his suggestion would be investigated but kept him at a distance. "We do not have any positions open for staff members," Adair was told (E. Garfield to W. C. Adair, personal commu-

nication, 16 March 1953). Nothing happened. Adair's initiative had no impact on the Medical Indexing Project.

More than a year later, after he had been fired by Larkey, Garfield (Eugene Garfield, personal interviews, 27 January 1992 and 4 February 1992) resumed contact with Adair "with the idea of writing a paper to be published in one of the learned society journals" (E. Garfield to W. C. Adair, personal communication, 11 June 1954). Having browsed through *Shepard's Citations* at the public library, Garfield was intrigued with the idea and had even written a paper on "Shepardizing the Scientific Literature" while he was a student at Columbia University (Garfield, 1954). At first, Garfield was not certain whether citation indexes could be applied to science:

> Without knowing exactly what you had in mind I do not feel it is fair for me to be discouraging at the outset. But the one thing that must be kept in mind when comparing the field of science with that of law, is that there are anywhere from one to three million articles each year appearing in the scientific journals. (E. Garfield to W. C. Adair, personal communication, 11 June 1954)

Garfield did not yet think of building a citation index. Working as a consultant in automation, he mainly focused on possible uses of computers (E. Garfield to W. C. Adair, personal communication, 11 June 1954). He perceived an opportunity in automating the production of citation indexes. Garfield and Adair decided to write two papers on the subject (Adair, 1955; Garfield, 1955). Through their correspondence, Garfield learned about the way Shepard's produced its citator, which familiarized him with the ins and outs of citation indexing.

The Citation Introduced to Science

Garfield's article was published in *Science*:

> In this paper I propose a bibliographic system for science literature that can eliminate the uncritical citation of fraudulent, incomplete, or obsolete data by making it possible for the conscientious scholar to be aware of criticisms of earlier papers. It is too much to expect a research worker to spend an inordinate amount of time searching for the bibliographic descendants of antecedent papers. It would not be excessive to demand that the thorough scholar check all papers that have cited or criticized such paper, if they could be located quickly. The citation index makes this check practicable. (Garfield, 1955, p. 108)

The index would be very handy for the working scientist: "It is best described as an association-of-ideas index, and it gives the reader as much leeway as he requires." In this respect the citation index would be, Garfield stressed, far superior to the traditional subject indexes that by nature restrict the interpretation of the article to a predefined number of topics (Garfield, 1955). Not only did Garfield focus on the information needs of the scientist, but he also translated the concept of the citation index in terms of the subject indexes with which both scientists and librarians were more familiar. His and Adair's articles did not, however, attract much attention.

Garfield, now an independent documentation consultant who advised, among others, Smith, Kline & French, was not deterred by the silence that followed his proposal. He undertook several initiatives to make the citation index more popular, which increased his grip on the intellectual and practical difficulties in compiling such an index. Together with Margaret Courain, supervisor of the Research Files Division at Merck, Garfield produced an experimental citation index to patents that he presented at the Minneapolis meeting of the American Chemical Society on 16 September 1955 (Garfield, 1957). At the December 1955 meeting of the American Association for the Advancement of Science in Atlanta, he made a strong plea for a centralized national documentation center (Garfield & Hayne, 1955). Partly as a personal exercise Garfield prepared a citation index to the Old Testament that he presented in 1956 to the American Documentation Institute in Philadelphia (Garfield, 1956). In this talk Garfield presented a new idea, interpretative citation indexing.

In January 1957 Garfield received the first serious support from a scientist. "Dear Mr. Garfield," wrote geneticist Gordon Allen, then at the Department of Health, Education, and Welfare of the NIH:

> Since the appearance of your article in *Science* two years ago, I have been eagerly looking for some news of steps toward a citation index. I have urged the American Society of Human Genetics to take some initiative in the matter [Allen had done this in 1956, when he talked to Sheldon Reed, then president of the society, (G. Allen, personal communication, 9 April 1959)], but they are already involved in the construction of a subject index in human genetics. The references I have seen to your suggestion (for want of a citation index, I probably have not seen all of them) have been disappointingly cool, and I wonder if you have received any personal letters that were

> more enthusiastic. If a group of interested persons were brought together, they might be able to make some headway. (G. Allen to E. Garfield, personal communication, 24 January 1957)

Stimulated by this support, Garfield submitted a proposal to NSF (Garfield, 1958) in August 1957. Its goal was "to determine the utility of citation indexes for science in terms of general usefulness, invariance in time, minimizing the citation of poor data, identification of the 'impact factor,' and provision for individual clipping services." The study also should develop "a suitable technical design for citation indexes." Its motivation followed Garfield's line of reasoning, albeit with an added emphasis on the index's potential for "the encyclopedic integration of scientific statements." Garfield also emphasized the potential of John Desmond Bernal's 1948 proposal for a central clearinghouse. The project was meant to be a two-year study, starting September 1958, and would restrict itself to compiling the index. One month later NSF turned down Garfield's proposal but expressed interest in a citation index (Dwight Gray to E. Garfield, personal communication, 23 October 1958).

Garfield took this as a flat refusal, proving once again in his mind NSF's inability to deal with the tasks at hand. He did not stop his campaign, though. In November he made his plea for a "unified index to science" at the National Academy of Science's Conference on Scientific Information (Garfield, 1959). In this presentation the idea of integrating scientific knowledge, already mentioned in his 1958 article, was further developed (Garfield, 1959, p. 674).

In May 1959 Garfield received a letter from geneticist Joshua Lederberg, which would prove the turning point in the history of the *SCI*:

> Since you first published your scheme for a "citation index" in *Science* about 4 years ago, I have been thinking very seriously about it, and must admit I am completely sold. In the nature of my work I have to spend a fair amount of effort in reading the literature of collateral fields and it is infuriating how often I have been stumped in trying to update a topic, where your scheme would have been just the solution! I am sure your critics have simply not grasped the idea, & especially the point that the author must learn to cooperate by his own choice of citations + thus he does the critical work. Have you tried to set this out in an adequate experiment? Would you look for support from the NSF? Of course you have to count on opposition from the established outfits, which have already succeeded in blocking any progressive centralization of the Augean tasks. (J. Lederberg to E. Garfield, personal communication, 9 May 1959)

As Lederberg later explained (29 July 1960) to Garfield, Lederberg's initiative was prompted by a science policy debate in the Genetics Study Section of NIH. The administration wished to evaluate its actual impact on research and proposed, in the words of Lederberg (J. Lederberg to E. Garfield, personal communication, 29 July 1960) "a number of rather fancy and inefficient schemes." Lederberg recognized that a citation index would accomplish the purpose "at a negligible additional cost" and decided to contact Garfield. Garfield's reaction was enthusiastic:

> I hope you won't be embarrassed by a show of emotion, but your memo almost brought tears to my eyes. It then seemed that over six years of trying to sell the idea of citation indexes had not been completely in vain. (E. Garfield to J. Lederberg, personal communication, 21 May 1959)

He told Lederberg the whole story of his pleas for citation indexes, the support of Gordon Allen, and the resistance he had met since 1954 (E. Garfield to J. Lederberg, personal communication, 21 May 1959). Lederberg was shocked by Garfield's letter and replied that he was "absolutely astonished that citation indexes are not long since a standard feature at the Patent Office" (J. Lederberg to E. Garfield, personal communication, 18 June 1959). He advised Garfield in the same letter to resubmit his proposal "to all the agencies who could be interested." The NIH he said, "would be an excellent target," since it "is anxious to evaluate its 'impact' on scientific progress, and how better do this than through your scheme." Lederberg proposed Garfield to "jump in" and ask NSF's assistance in organizing a scientific committee as suggested by NSF's last letter on the subject.

This resulted in the proposal to NSF (Garfield, 1960a) to construct a citation index by scanning a list of a thousand journals and processing all references to forty-three specified genetics journals as well as all references to twenty-two specified general science journals. The proposal to NIH, on the contrary, entailed the processing of "references from the specified journals, punching them into IBM cards, and mechanically sorting these. Only then would the citations to genetics journals be selected and printed" (Garfield, 1960b). The index would be published by turning over to the editors of the journals "an individual journal citation index" that they could publish as a yearly supplement. This was only an "interme-

diate mechanism," though. Garfield wished to keep open the option of a separate publication and wrote that he was "in correspondence with editors on the publication problem." On 26 December 1960 he could at last break the big news (the "notification and statement of grant award" granting $49,450 per year for three years) to Lederberg: "Dear Josh, the official note that NIH approved our grant came in the other day. This was quite nice Xmas present to say the least" (Brewer to E. Garfield, 15 December 1960, received December 23, 1960; E. Garfield to J. Lederberg, personal communication, 26 December 1960).

And to Allen: "Dear Gordon, Santa Claus was very good to us. We learned that NIH approved its half of the revised budget which NSF asked me to submit based on $100,000 per year for three years" (E. Garfield to G. Allen, personal communication, 26 December 1960). The NSF grant was approved two months later.

Building the Index

"I think you're making history, Gene!" wrote Lederberg on 24 January 1962. Building the *SCI* turned out to be a bigger project than even Garfield expected. It took more time, more money, and was technically more complicated than foreseen in the contracts. Constructing the index was not only a technical endeavor, but a political enterprise as well. Joshua Lederberg perceived the *SCI* as a means to open up the clogged communication channels in science. Building the index required not only a technical or library expertise but political acumen as well. Through the intense cooperation of Garfield and Lederberg, this enterprise became part of the science policy debate in the United States on the now-famous Weinberg report (PSAC, 1963). The history of the *SCI* can hardly be understood without an appreciation of these technical difficulties and political dimensions.

Science policy provided the context in which Lederberg remembered Garfield's 1955 proposal and decided to write him. *SCI*'s political relevance was directly related to its bibliographic properties. Yet the connection between the *SCI* and science policy was rather loose. From October 1961 onward this changed. The building of the *SCI* became intimately involved in the debate on the future of scientific information in the United States after Joshua Lederberg was appointed a member of the Panel on Science Information (PSAC). His assignment was to rewrite the general introduction to the Weinberg report, and he saw this as an opportunity to push for a radical overhaul of the "anarchic" way scientific information was organized. Lederberg advocated a centralized information system, modeled after John Desmond Bernal's pleas from 1948. In the end Lederberg argued, in a letter to Garfield on 8 October 1961, that this should result in the abolition of the traditional journals.

Lederberg took this political development as an opportunity also to promote the *SCI* itself and asked Garfield to be his "informal consultant" and give him background information on "detailed proposals that you consider reasonably intelligent." Garfield showed no hesitation (E. Garfield to J. Lederberg, personal communication, 11 October 1961).

The PSAC issued its report in 1963 with an array of proposals and calls for action, directed to the federal government, to the scientific community, to individual scientists, and to the libraries. PSAC (1963) has since been seen as a landmark in the history of documentation (Schneiders, 1982, p. 176). In this report the information crisis was not merely a question of keeping the individual scientist informed, as it had been formulated in the 1958 Baker report (PSAC, 1958). The crisis threatened the very identity of science. The panel opened the report with the following sweeping statement:

> Science and technology can flourish only if each scientist interacts with his colleagues and his predecessors, and only if every branch of science interacts with other branches of science; in this sense science must remain unified if it is to remain effective. The ideas and data that are the substance of science and technology are embodied in the literature; only if the literature remains a unity can science itself be unified and viable. Yet, because of the tremendous growth of the literature, there is danger of science fragmenting into a mass of repetitious findings, or worse, into conflicting specialties that are not recognized as being mutually inconsistent. This is the essence of the "crisis" in scientific and technical information. (PSAC, 1963, p. 7)

The PSAC report called for drastic action and for major changes in the scientific system and the behavior of individual researchers. One of the recommendations was the development of a new searching tool, the citation index, about which the panel was "particularly impressed."

These recommendations were, at least partly, the result of an intense correspondence between Lederberg and Garfield about the solution to the problem of scientific information while they were building the *SCI*. Garfield had laid out a comprehensive scheme comprising "three levels of reporting: title, abstract, full paper." The basic idea was that 1 percent of the papers would

be published in a "national or international organ" (for example, a daily science newspaper); the next 10 to 25 percent would be published in "a series of select journals," whereas the vast majority of the papers would be put in a central depository. This would put an end to the proliferation of new journal titles. The newspaper would also publish lists of all papers (E. Garfield to J. Lederberg, personal communication, 10 November 1961). The national documentation center, which would be the central axis, should distribute "a series of abstract journals." Moreover, a "prompt translation service" would provide for fast international communication.

Garfield envisioned his *Current Contents* or its successor as the place to "publish by title," whereas the newspaper would have a daily citation index section. All in all the system would be a drastic improvement for timely access to all available information. "An important factor," Garfield stressed in a 10 November 1961 letter to Lederberg, "is that a man's personal bibliography should have the same publication value regardless. A reference to a paper that does not get into the major primary organ or in the journals should be considered equally." Garfield reiterated Bernal's ideas in a letter to Lederberg, dated 15 November 1961, as he made clear by urging Lederberg to look into Bernal's papers.

In the first draft of his "Notes on a Technical Information System," Lederberg reconstructed the main problem as follows:

> As members of the scientific community we have a deeply rooted obligation to interact with the "literature." Not so much the size but the dispersion and formlessness of the institution make this an ever more hopeless aspiration. . . . The present system has generated two responses: the defeat of neurotic frustration for some, the compromise of narrow specialism for others. I feel the survivorship of humanistic science demands a better solution. (Lederberg, 1962a, p. 1)

Lederberg proposed a central depository together with "select journals":

> A centralized repository would provide the range of materials that I would specify as being required for my immediate and retrospective information requirements. Concurrently, select journals with high standards of selection and editorial quality would maintain my contact with the breadth of scientific culture. (Lederberg, 1962a, p. 7)

The depository would be built according to a set of ground rules. One of these would be that no paper could be withdrawn once deposited: "As with journal publications the author's reputation is permanently attached to it." Papers would be distributed and refereed "promptly." Moreover, an updated citation index would be attached to the articles. The principal advantage of the repository scheme was, according to Lederberg, the "prompt and widespread availability" of contemporary findings. "That contributions can take a full year to come out in print is an absurdity of modern science" (Lederberg,1962a, p. 4). The repository would "discourage the redundancy implicit in peripheral publication and in the irresponsibility of gossip and 'invisible colleges.'" It would facilitate the publication of "expensive archival documents" like taxonomies. Last but not least, it would stimulate the journals to "revert to being *select* journals: they are broadsides on which I would rely to bring me *unasked* the best or overtly most interesting of contemporary science." The user would be more central than in the prevailing system, Lederberg felt. He expected the "journal output" to decrease to "about 10% of its current level."

The central problem with realizing this radical overhaul was that it needed a certain critical mass. Hence, the idea of a daily newspaper, which Garfield and Lederberg had discussed in their conversations about the *SCI*, also became a strategic item in realizing their information revolution. The already existing publication system, with its vested interests, seemed the main obstacle (E. Garfield to J. Lederberg, personal communication, 10 November 1961). Unfortunately, Lederberg noted, "one of the serious shortcomings of the *OSIS* in NSF is that it really has neither the staff nor the mandate to consider such large scale systems propositions." Garfield had the same experience. He had sent his proposal for a unified index to science in newspaper format to NSF but received no response (E. Garfield to J. Lederberg, personal communication, 6 March 1962).

Lederberg's scheme differed fundamentally from conventional publication in scientific journals. First of all, the primary responsibility for seeking editorial criticism would be shifted to the author. Second, the need for primary journals would disappear. "Relieved of the unnatural responsibility for primary archives and communication," the scientific societies and other journal sponsors could devote themselves to too-often neglected services "especially in review and interpretation." At this stage of scientific communication Lederberg wished more opportunities for commercial initiatives (Lederberg, 1962b, p. 4). Third, authors would also be responsible for the production of abstracts, since "manpower requirements" prevented their central production. Leder-

berg acknowledged the possibility that "peripheral agencies" also might be able to continue their abstracting services. Fourth, the government would have the primary responsibility for financing the whole system. Fifth, the system would be oriented to innovation, looking to "the future development of data handling and telecommunication systems to replace the techniques of the present proposal."

These policy discussions stimulated Garfield to see the Institute for Scientific Information's (ISI's) role in terms of shaping the future of scientific communication:

> I have been thinking "big" down here in terms of ISI's future. I hope to incorporate this thinking into a series of proposals that tie in with your proposals on Science Advisory Committee. . . . I am convinced we are only five to ten years away from bridging the existing artificial gap between technical science writing and writing for the laymen. In fact, there is probably a greater need than you and I realize for a citation index "structure" that would relate a conventional clipping service with our scientific clipping service. (E. Garfield to J. Lederberg, personal communication, 9 July 1962)

This view tied in with the problem of how to publish the *SCI*. While the computer programs, data files, and citation indexing procedures were being developed, the question of publishing the resulting index became more pertinent. Garfield proposed to Ralph O'dette in a 17 September 1962 letter that NSF test "the newspaper format" for a daily citation index, to achieve a "low cost per reading." The newspaper should have the format of the *New York Times*, initially comprise sixteen pages, and contain reprints of original research papers and review articles (four pages), a daily updated author bibliography (five pages), a citation index (six pages), and a subject index (one page) (Garfield, 1962). The author bibliography would contain 750 papers per day and was vital for the use of the indexes. Garfield expected that in one year three million citations would have been listed this way. The "Daily Scientist" as it was called should be a throw-away paper: "The philosophy behind a daily dissemination technique is that the information comes in small segments. The daily newspaper is quickly scanned and then discarded" (Garfield, 1962, p. 2). Garfield estimated that scientists would be prepared to pay a subscription fee of thirty dollars a year. He proposed that NSF test the idea by sending 25,000 scientists consecutive daily issues for two months. If the NSF would give initial support, the experiment could be expanded with the help of NIH, NASA, and AEC. A one-year experiment would cost around $500,000, Garfield estimated. Most of this money would be necessary to produce a unified citation index to science anyway. Therefore, Garfield argued, his proposal would "bring a vast amount of information to the individual scientist at a phenomenally low cost" (Garfield, 1962, p. 2).

The NSF, however, was not prepared to fund the production of the *SCI*, which resulted in Garfield's decision to publish the index on a for-profit basis. The risky adventure nearly bankrupted ISI, and it was mainly on the profits generated by other products, primarily *Current Contents*, that the *SCI* had a chance to become profitable.

Conclusions: Translating the Citation Concept

Automation

The citation index NIH and NSF supported and the *SCI*, as it would be published from 1964 onward, did not look like *Shepard's Citations*. Technically, the idea was still the same. Because of this, Garfield's proposal to NSF could state that most of the uses of the *SCI* were "analogous to their use in legal research." This statement nevertheless concealed essential dissimilarities. The fundamental change was in the meaning of a citing relation. The outlook of the index differed as well. Moreover, ISI's way of producing the index would be the complete reversal of Shepard's. The production of the *SCI* was therefore not a matter of simply applying a ready-made tool in a novel area. Developing the *Science Citation Index* required both a new way of looking at the scientific literature and a new conception of citation indexing: "The brilliant utility of the citation index approach is that it cuts across the problem of meaning by an automated procedure" (J. Lederberg to E. Garfield, personal communication, 9 November 1962).

It would have been impossible to make a database such as the *SCI* without computers because it would have been far too expensive. Even with existing computers, it was a risky business. It was computerized processing that made possible the migration of the citation concept from the legal to the scientific context. The corresponding devaluation of labor made the production of the *SCI* possible within the budgets available for these kinds of enterprises in the United States at the time.

Comprehensiveness

By automating the production of the *SCI*, Garfield, Lederberg, and Allen could tackle the enormous task of

indexing the scientific literature while retaining its complete coverage of science. William Adair had been aware of the problems of scale as well but did not think of automation. Instead, he proposed to index separately the various scientific specialties or disciplines. It was a familiar solution; *Shepard's Citation* also was fragmented according to the structure of the legal system in the United States. The idea was not strange in the world of science either. After all, most scientific journals were limited to a narrowly defined specialty. Moreover, several other citation index projects were constructed along the same lines. Garfield had been the principal propagandist for citation indexing, but he was not the only one involved. In the early 1960s NSF supported several citation index research projects. The *SCI* project was, however, the only attempt to produce a comprehensive citation index covering, in principle, all of science.

Two principal "competitors" had opted explicitly in favor of a monodisciplinary approach. Statistician and leading citation index researcher John Tukey studied and built a citation index of the statistics literature at Princeton University (Tukey, 1962; Tukey, n.d. b; Tukey, n.d. a). At the Massachusetts Institute of Technology, the inventor of bibliographic coupling, Michael Kessler (1961), was constructing a complete information system of the physics literature. He did not consider a citation index strong enough to sustain a pilot model system in itself, although it would be a useful element to add once the model was constructed, because citation was "a low probability event" (Kessler & Heart, 1962; Kessler, 1965).

The objective of the *SCI* to cover all the scientific literature was underpinned by "the unity of science." Without the possibility of going beyond the boundaries of the academic disciplines, a citation index would add practically nothing to traditional subject indexes. After all, the researcher could be relied upon to know the literature in his or her own specialty. The *SCI* should be able to locate relevant research in unexpected places, and this only seemed possible if the *SCI* was not structured along disciplinary lines. The *SCI* was also expected to change the citing behavior of the scientist, which was not the case with Shepard's Citator, only a registering device. The citing behavior of attorneys and judges was fairly standardized, which made it possible for indexers to classify citations with a fairly restricted set of symbols. In contrast the reason a scientist cites an article is not restricted at all. In fact references to scientific papers play divergent roles. Even the same citation can change meaning in the course of time. The makers of the *SCI* expected to exert a positive influence on the scientists' citing behavior. In its turn this would increase the value of the *SCI*.

The Information Crisis

In the 1950s science had been growing too fast to cope with its results. It made some parts of the scientific community gradually receptive to innovations in handling the literature. This "information crisis" is a key factor in the birth of the *Science Citation Index*, playing social as well as cognitive roles. It shaped the way the central problems in the realm of science, science management, and science policy were defined. Government agencies provided funds to find solutions to this information crisis and thereby created a new labor market for people with both scientific and librarian skills. This new field was where people as diverse as a documentation specialist, a researcher in human genetics at the National Institutes of Health, a Nobel laureate in bacterial genetics, and a retired vice president of Shepard's could meet each other. The crisis, made more urgent by the *Sputnik* surprise, eventually gave citation indexing the official approval it needed to take off.

It was a debate at NIH about the evaluation of NIH-funded research that reminded Lederberg of Garfield's 1955 paper in *Science* and prompted him to write his memo in 1959. Once a citation score is transformed into a measure of the impact of a paper, all sorts of policy-related studies can be easily imagined if the database is large enough. The sociological use of the *SCI* was an outgrowth of this capability and of the network approach. Notwithstanding, the central motive for scientists like Lederberg and Allen was, and would continue to be, the literature-searching capabilities.

Innovative Outsiders

Without the drive, perseverance, social capacities, and technical expertise of Eugene Garfield, the immense task of building the *SCI* would probably not even have been thinkable. It is not only a matter of personal traits, but also of being in the right place at the right time. Garfield was an outsider in more than one respect, which made it possible for him to think about solutions other people would reject immediately. Garfield was well prepared for information services, *Current Contents* being the proof of that. Not coincidentally, the two scientists who reacted to Garfield's 1955 article in *Science* were geneticists. The structure of the new science of genetics made coping with the literature more pressing for Lederberg and Allen than for, say, the nuclear physicists. Genetics

still had unclear boundaries. On the other hand, the professional societies in human and bacterial genetics stuck to the old subject indexing, as did all relevant institutions. Thus the personal histories of Lederberg and Allen have been important factors as well.

In the process of translating the citation concept to the world of science, the funding agencies and Eugene Garfield learned to cope with each other. Garfield was an outsider to the academic world, which does not mean he did not have many contacts with researchers and science policy officials. On the contrary networking was one of Garfield's strong points. He was asked to review proposals to NSF on indexing projects on a regular basis (E. Garfield, personal communication, 12 September 1959). He was running his own company, Eugene Garfield Associates, with *Current Contents* as its main product. He was not affiliated with a respectable academic institution, which created additional hurdles.

An intellectual problem existed as well: Citation indexing was an unknown entity. Garfield's proposals showed this, and naturally he wanted to keep open as many options as possible. But the funding agencies were also uncertain and wanted to know more precisely what they were supposed to support. Allen's and Lederberg's support made Garfield's undertaking more respectable. Moreover, they taught Garfield how to deal with agencies like the NSF and NIH, while transforming the citation concept in this process.

Success as well as Failure, and Yet a Success

The experimental genetics citation index appeared in 1963, the *SCI* in 1964. Since then the *SCI* and its associated products are a well-known feature in most scientific libraries over the world. ISI almost went bankrupt because of the *SCI*, but in the end it turned out to be profitable (E. Garfield, personal interviews, 27 January 1992 and 4 February 1992, Philadelphia). It seems a classic American success story, with log cabin (Garfield's chicken coop in New Jersey where he started producing *Current Contents*) and all. And a success the *SCI* surely is.

But it is also a story of failure. Lederberg was not only thinking about a bibliographic tool when he pushed the case of citation indexing in the courts of science policy, but he also set out to revolutionize the whole publication system of science. In 1959 Lederberg had adopted Bernal's program of doing away with all scientific journals as a primary channel of publication. As a member of the PSAC panel on scientific information, he was impressed by Derek Price's book *Science since Babylon* (1961) and pressed for abolishing the anarchical way of publishing. All commercial publishers should be pushed out of the business of primary publication. The process of scientific communication should be made "efficient, systematic, anxietyfree, reliable." Papers would be available on request, and their existence would be announced via abstract services. The refereeing system would be completely eliminated, authors being responsible for their own products. Retrieval of literature would be rationalized with machine-driven indexes, citation indexing being one of them. A daily journal of science would be the central medium of mass communication in the whole system.

With Garfield acting as his informal consultant on the matter, Lederberg advanced these innovative—and perhaps radical—ideas. In their hands the *SCI* would not be merely a searching tool but a revolutionizing instrument, profoundly changing the world of science. In this respect their enterprise was a failure. The birth of the *SCI* did not make any immediate changes in the scientific community, nor did it profoundly influence scientists' behavior. By limiting the scope of the *SCI*, the existing institutions successfully defended the traditional way of publishing.

And yet on a more fundamental level the *SCI* is a success, but in a different way than its creators expected. While the *SCI* did not trigger immediate changes in the scientific system, it did shape a whole new set of signs of science. The citation indexing concept that Garfield and Lederberg had forged from its original legal citator predecessor became the cornerstone of a novel social science specialty—scientometrics—as well as the building block of an intricate maze of science and technology indicators. As I have argued elsewhere (Wouters, 1999), this development has created a set of fundamentally novel representations of science and technology that has influenced both science policy and the production of scientific knowledge at all levels.

References

Adair, W. C. (1955). Citation indexes for scientific literature? *American Documentation 6*, 31–32.

Garfield, E. (1954). *Shepardizing the scientific literature.* Unpublished manuscript, Columbia University Library School, New York.

Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of ideas. *Science 122*, 108–111.

Garfield, E. (1956). Citation indexes to the Old Testament. Unpublished.

Garfield, E. (1957). Breaking the subject index barrier—a citation index for chemical patents. *Journal of the Patent Office Society 39*, 583–595.

Garfield, E. (1958). *A general feasibility study of citation indexes for science: A proposal for research.* Philadelphia: Eugene Garfield Associates.

Garfield, E. (1959). A unified index to science. *Proceedings of the International Conference on Scientific Information, Washington D.C., November 16–21, 1958,* (pp. 461–74). Washington, D.C.: National Academy of Sciences–National Research Council.

Garfield, E. (1960a). Citation index to genetics and general science literature: research proposal. Philadelphia: Institute for Scientific Information.

Garfield, E. (1960b). Citation index to genetics literature: Application for research grant. Philadelphia: Personal Archive Eugene Garfield.

Garfield, E. (1962). Proposal: Application of newspaper publishing methods to the dissemination of scientific information. Philadelphia: Institute for Scientific Information.

Garfield, E. (1963). *Science Citation Index: An international interdisciplinary index to the literature of science* (Vol. 1: aabe–capl). Philadelphia: Institute for Scientific Information.

Garfield, E., & Hayne, R. L. (1955, December). Needed—a national science intelligence & documentation center. Paper presented at the American Association for the Advancement of Science, Atlanta, GA.

Garfield, E., & Sher, I. H. (1963). *Genetics citation index—experimental citation indexes to genetics with special emphasis on human genetics*. Philadelphia: Institute for Scientific Information.

Kessler, M. M. (1961). An experimental study of bibliographic coupling between technical papers. Preprint.

Kessler, M. M. (1965). The MIT technical information project. *Physics Today,* 28–36.

Kessler, M. M., & Heart, F. E. (1962). Concerning the probability that a given paper will be cited. Supported by a grant from NSF.

Larkey, S. V. (1949). The Army medical library research project at the Welch Medical Library. *Bulletin of the Medical Library Association 37*, 121–124.

Lederberg, J. (1962a). Notes on a technical information system. Philadelphia: Personal Archive Eugene Garfield.

Lederberg, J. (1962b). SCITEL: A central scientific communication system. Philadelphia: Personal Archive Eugene Garfield.

Miller, E. (1961). The National Library of Medicine index mechanization project. *Bulletin of the Medical Library Association 49*(1), 1–96.

Price, D. J. De S. (1961). *Science since Babylon.* New Haven, CT: Yale University Press.

PSAC (1958). *Improving the availability of scientific and technical information in the United States: A report of the President's Science Advisory Committee*. Washington, DC: The White House.

PSAC (1963). *Science, government and information: A report of the President's Science Advisory Committee*. Washington, DC: The White House.

Schneiders, P. (1982). *De bibliotheek—en documentatiebeweging 1880—1914; bibliografische ondernemingen rond 1900.* Doctoral dissertation, University of Amsterdam, The Netherlands.

Tukey, J. W. (1962). Keeping research in contact with the literature: Citation indices and beyond. *Journal of Chemical Documentation 2*, 34–37.

Tukey, J. W. (n.d. a). *The citation index and the information problem: Opportunities and research in progress.* Princeton, N.J.: Statistical Techniques Research Group, Princeton University.

Tukey, J. W. (n.d. b). *A citation index for statistics: An introduction and a progress report.* Philadelphia: Personal Archive Eugene Garfield.

Wouters, P. (1999). *The citation culture*. Doctoral dissertation, Science and Technology Dynamics, University of Amsterdam, The Netherlands.