Essays of an Information Scientist, Vol:1, p.98-99, 1962-73
Current Contents, #14, April 8, 1970

"Current Comments"

What Does Automation of Citation Mean?

April 8, 1970

Recently, we reprinted in *Current Contents*®[1] a paper I published in 1965, "Can Citation Indexing Be Automated?"[2]. Reprinting the article may have puzzled some of our readers. They know that this year, 1970, the *Science Citation Index*® is close to ten years old; and they know too that computers and computer-driven photocomposition devices are used extensively in compilation of the *SCI*®. It probably surprised most readers even more to learn that my answer to the question above was, and still is, "No!"

The title of my original article was ambiguous, if not confusing, not because of any deliberate loose use of words, but because in English there is so little *verb*al strength left in what remains of, or what must be used for, *verb*al nouns. (Oh, for a good solid Latin gerund or Arabic masdar!) A "citation" may be a thing, but "citation" may also be the act or process of effecting that thing. The title does not mean, "Can machines be used advantageously in assembly of a citation index?" The answer to that is obviously "Yes!" The question is, "Can machines (computers) simulate the human critical processes necessary to the act of citation?" The answer to that is, perhaps not so obviously, "Not yet!" I definitely don't intend that qualification as a lead-in to discussion of a computer's ability to "think" or to "simulate the processes of human thought".

A citation is a descriptor, in the same sense that a conventional subject heading is a descriptor. There should be little doubt of that. It is also, as a substantive, uniquely descriptive of the paper which uses it to cite another publication. And it shares none of the vagueness of the conventional subject heading and none of the inaccessibility of subject headings expressed in different national languages. As such, a citation is, in the strictest sense of the word, *explicit*, and a group of citations, as pointed out in the paper, can, linguistically and bibliographically, correspond to the document that they have been used to describe. The strength of citation indexing flows from this fact–for every document there is a unique set of explicitly descriptive citations. But have they been supplied?

Most publications also have, as most readers realize, *implicit* citations–those citations the reader feels must be added to complete the citation-description provided by the author. Implicit citations are the ones that aren't there but "ought" to be–to verify, to cor-

roborate, to explain, to amplify. The question considered by my paper concerned the prerequisites for an automatic procedure for supplying citations when they are missing. What characteristics of a scientific text imply the need for citations? Are these characteristics susceptible to some linguistic transformational or other automatic analysis that will enable the "machine" to identify them? Can we expect in the future, for example, that a patent examiner or inventor will be able to have patent applications scanned by an artificially intelligent machine in order to indicate all pertinent prior art? Can the journal editor or referee expect the same performance for research articles? The question is neither trivial nor merely speculative. The fundamental question involved provides a clue to the significance of proposals like H.G. Wells's "World Brain" and Bush's "Memex", [3] which will be discussed in the near future.

(1) Garfield, E., "Can criticism and documentation of research papers be automated?" *Current Contents/Life Sciences* 13 (9), 4 (1970).

(2) Garfield, E., "Can Citation Indexing Be Automated?" in M.E. Stevens *et al.*, Eds., *Statistical Association Methods for Mechanized Documentation, Symposium Proceedings, Washington 1964* (National Bureau of Standards Miscellaneous Publication 269, December 15, 1965) pp. 189-92.

(3) Garfield, E., " 'World Brain' or 'Memex?' Mechanical and Intellectual Requirements For Universal Bibliographic Control" in E.B. Montgomery Ed., *The Foundations of Access to Knowledge--A Symposium* (Syracuse University Press, Syracuse N.Y., 1968) pp. 169-96.