Primordial Concepts, Citation Indexing, and Historio-Bibliography*

EUGENE GARFIELD

The Perennial Dichotomy—Storage Medium or Information Stored?

In several papers (1), I have described the *Science Citation Indexes* and the Unified Index to Science (2) as preliminary steps toward achieving the dream of universal bibliographical control which H. G. Wells symbolized in the *World Brain* (3). To some, the Wellsian term "World Brain" might seem less appropriate than "Memex," the term chosen by Vannevar Bush (4) to symbolize the ideal information retrieval device. However, there is a world of difference between *Memex* and a *World Brain*—essentially the difference between hardware and software—between a communication carrier and the intellectual-message-carrier problem. The *World Brain* symbolizes the information stored—*Memex*, the storage device. In designing any bibliographic system, it is imperative to make these distinctions. It distresses me when this dichotomy is glossed over in vague generalizations about the so-called Information Explosion.

What Is a Subject?

One of the most frequently expressed criticisms of the citation index is that it is not a "subject" index. What is really meant, however, is that the citation index is not a "word" index. Therefore, the question, "What is a subject?" is fundamental not only in evaluating citation indexing, but all types of systems for subject analysis. By attempting to establish why a citation index *is* a subject index, some important philosophical insights concerning subject and content analysis are obtained.

Primordial Word-Document Events

Theoretically every word in the dictionary could be traced to an important historical event—the time and place that word first occurred.

* Based on a talk presented at the Eighth Annual Summer Symposium on "The Foundations of Access to Knowledge," Syracuse University School of Library Science, Syracuse, N.Y., July 30, 1965 (In Press). See reference No. 3.

Reprinted from THE JOURNAL OF LIBRARY HISTORY 2 (3), 235-249 (1967)

Especially in science, newly coined words (neologisms) can usually be traced to, or identified with, a particular paper or book. Frequently, the "main theme" or subject of this *primordial* document is the same topic symbolized by the new word (5).

Example of Lederberg's Euphenics

For example, in 1962, Professor J. Lederberg coined the word "euphenics" which first appeared in a paper entitled, "Molecular Biology, Eugenics, and Euphenics" published in *Nature* 198, 428 (1963) (6). Most people would agree that the main "subject" of that paper is indeed euphenics. As long as this paper was the only one in the literature on euphenics, there was effectively a one-to-one equivalence between the word "euphenics" and the citation which identifies the document in which it appeared. The word "euphenics" and the citation "*Nature* 198, 428" are both symbols. They are equivalent symbols for the main topic discussed in Lederberg's paper. The "subject" of the paper is the same whether we symbolize it by the word "euphenics," by the short citation "*Nature* 198, 428" or by the longer citation "LEDERBERG J, 63, NATURE 198, 428" as it would be identified in the *Science Citation Index*. (See Figure 1).

Now suppose that some other author discusses the subject of euphenics in a subsequent paper. It is the usual custom in scholarly research, when using new terms, to provide a footnote reference to the source of the new term. As a result, in a citation index system, the second citing paper is indexed under the "term" "LEDERBERG J, 63, NATURE 198, 428."

If a theme of the second paper is euphenics, one expects the document to be indexed under the term "euphenics" in a conventional word indexing system. Both methods have achieved identical objectives —to make information on euphenics retrievable. In the latter system, the word is the access or starting point; in the former, the citation is the starting point.

Dialogue between Librarian and Scientist

The expert geneticist probably does not have to be told that Lederberg has written on euphenics; nor will he have to be told other key events in the development of his particular specialty. On the other hand, the reference librarian is of necessity a generalist and cannot be expected to remember that which has become second nature to the specialist. This means that the reference librarian must usually engage the scientist in a dialogue, the purpose of which is to simplify the use of the indexing systems available. If I ask my librarian to "find me

1965 CITATION INDEX

LEDERBERG J--------•	•NASACR57029 DENDRAL6¤				
LEDERBER.J	P NAS US		65	53	134
	• • •				
----------57-	J BACT----------			73	144
MOOKERJE.S	I J EX BIOL		65	3	1
----------58-	J BACTERIOL--------			75	143
KAWAKAMI M	BIOC BIOP R		65	18	716
PLAPP R	ARCH MIKROB		65	50	171
----------59-	SCIENCE----------			129	1649
MAKELA O	ANN MED EXP		64	42	152
NISONOFF A	ANN R BIOCH	R	64	33	355
WHITNEY PL	P NAS US		65	53	524
----------62-	NSG816051 NASA REP--¤				
BOTAN EA	SPACE SCI R	R	64	3	715
----------63-	CIBA S MAN FUTURE---¤				
SWAMINAT.MS	CURRENT SCI		65	34	108
----------63-	MAN FUTURE---------¤				267
MURRAY JE	ANN NY ACAD		64	120	545
----------63-	NATURE LOND--------			198	428
MCKUSICK VA	J CHRON DIS		64	17	1077
----------64-	COMPUTATION MOLECULA¤				
LEDERBER.J	P NAS US		65	53	134
TUNNICLI.DD	ANALYT CHEM		65	37	543
----------64-	N6421426 STAR SCIENT¤				
TUNNICLI.DD	ANALYT CHEM		65	37	543
----------64-	SUBALGOL PROGRAM CAL¤				
TUNNICLI.DD	ANALYT CHEM		65	37	543

Figure 1

Sample of 1965 Science Citation Index illustrating synonomous citations.

papers on euphenics," there are a number of reasonable questions that can or must be asked by the librarian. If the librarian is embarrassed to do so, or feels that any sacrifice of additional time on the part of the patron is unjustified, then he pays a heavy price. He employs a series of bibliographic twists and turns taught to him in library school or which have become second nature through experience.

In the Process of Question Translation, What is Reasonable?

It all boils down to this: How do you translate the user's question into the form a given system can cope with? For example, what can the system do with a question like "Find papers on engineering human development." The answer to the question happens to be the same as the previous question "Find me papers on euphenics." Is it unreasonable of the reference librarian to ask, "Do you know of anyone working in this field, or do you know of alternate terminology used to symbolize this subject?"

Is the Completely Up-to-date Thesaurus Necessary?

Anyone familiar with Lederberg's Ciba Foundation talk on euphenics would probably be able to translate the phrase "engineering human development" into the equivalent word "euphenics." But how can the librarian answer a question concerning this concept if he does not know the term "euphenics" exists? At that moment, it would have been almost impossible for the librarian to use any existing system successfully. The general reference and lexicographic apparatus in libraries is quite far behind the advance of scientific terminology. Four years after Lederberg's original paper, the word "euphenics" still cannot be found in a dictionary or encyclopedia. Even if it were, how would the librarian know that euphenics is the synonym for "engineering human development"? One might conceivably find the term through a Roget-type thesaurus if it were up-to-date. If indexing services like *Index Medicus* issued subject heading authority lists which accounted for every new biomedical word, phrase, or eponym, one could identify such concepts more easily. In addition, this up-to-date thesaurus would also have to contain "terms" from German and other languages. Lederberg's concept of euphenics, no matter how or in which language it is expressed, is consistently symbolized by the citation "LEDERBERG J, 63, NATURE 198, 428." This does not imply that any and every paper related to euphenics or the engineering of human development will contain such a reference. The reference may be omitted deliberately or unwittingly. However, any scholarly paper on this topic is bound to contain the Lederberg citation or some near-synonym for it such as Crow's "Modifying Man." (7)

All Existing Systems Make Demands on User

The previous analysis illustrates how a citation is a "subject" and that in any existing non-ideal system, some accommodation must be made to the system by the user. Any system which makes *excessive* demands on the user will fail. However, no system yet available makes *no* demands on the user. If such a system existed, it would be a *World Brain*—probably a community thinking machine—a network of human brains linked by telepathy.

The Ideal Gas Theory vs. Network Theory of Bibliographic Organization

The traditional philosophy of indexing system design implies that individual documents can be treated as though they were independent entities. This basic fallacy not only results in the loss of important informational links, but it is basically inefficient. This is illustrated by the

example of the identical document published in two different journals. The same indexing procedure will be followed for "both" papers as though they were two entirely different documents. The indexers would select subject headings to describe the "main theme" of each paper. In practice, we know this does not occur consistently even for the same indexer. Little or no effort is made to establish a possible relationship between the document being indexed and the documents already indexed in the collection. There are exceptions to this rule, but generally the building-block development of human knowledge is not perceptibly reflected in traditional indexing systems. But frequently in literature research, one would like to trace an unambiguous and uncluttered path from the first occurrence of a new concept to its *subsequent* occurrence in the literature. In conventional word indexing systems, the indexer cannot afford the time to establish such linkages between concepts. He treats the literature as a series of independent events, like molecules of gas. This may also account for much probabilistic philosophy in the literature as, e.g., the work of Mooers (8).

Network Model

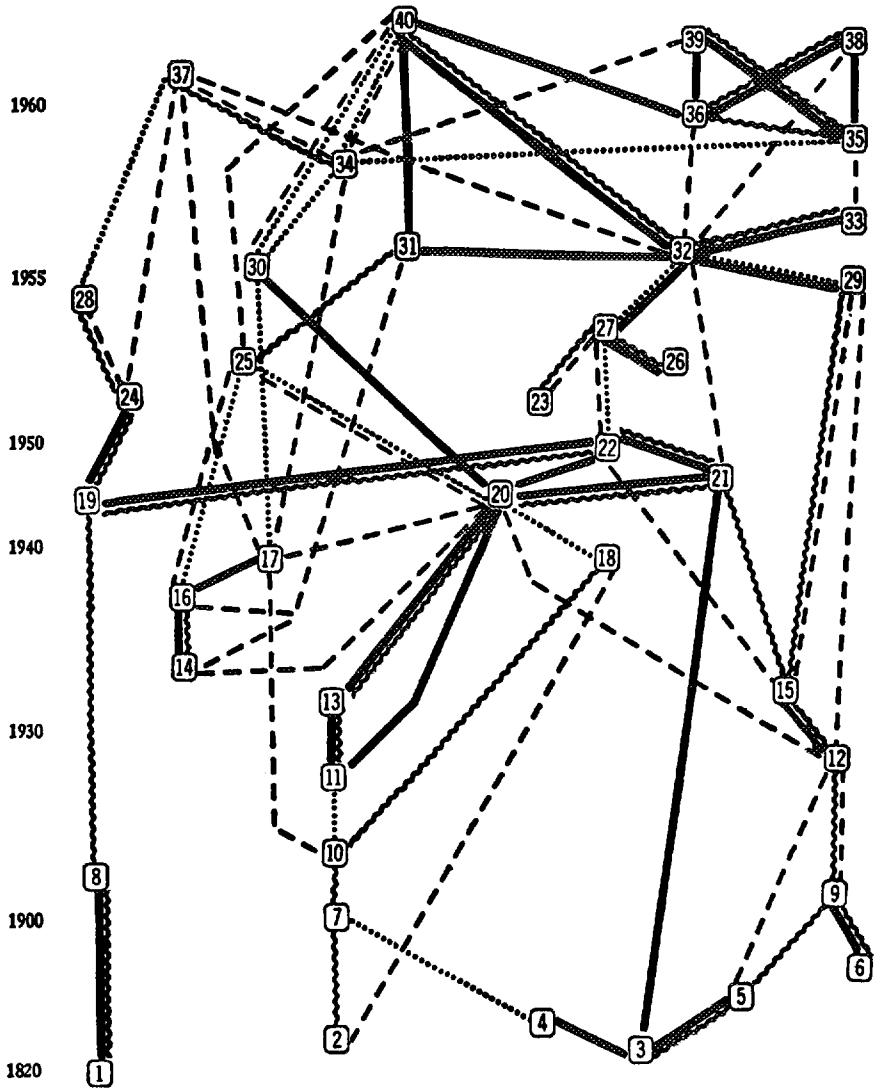
But the literature is not an "ideal gas." Libraries consist of collections of highly interrelated documents. The literature is a heavily cross-linked network. The clearly visible linkages are those ordinarily provided by authors in the forms of explicit citations. Less clearly seen are implicit references as in eponyms and neologisms. Almost invisible linkages exist in the natural language expressions which obscure the relationships, especially to the unskilled observer.

Graph Theoretic Model

Conventional bibliography is essentially a simple listing or inventory of publications which disregards most of the interrelationships between the items in the inventory. In contrast, citation indexing integrates this necessary and useful listing in a huge graph or network. In this graph, each document is a node or vertex in a huge multi-dimensional network. By analogy, this model of the literature (man's knowledge) is like a large road map in which the cities and towns share varying degrees of connectivity. Even the smallest hamlets are nodes on the citation map of science.

Historio-Bibliography—A New Methodology

In previous work, I have referred to this type of graph as an historical map (9). Since each document is, in fact, an "event" and bears a



Network diagram for history of DNA based on Asimov's book, The Genetic Code. Composite of six network diagrams as reported in E. Garfield, I.H. Sher, and R.J. Torpie, The Use of Citation Data in Writing the History of Science (Philadelphia: Institute for Scientific Information, 1964), 76 pages.

FIGURE 2

date, a subject bibliography, displayed in graphical form, is seen to be an approximate history of the subject covered. In conventional bibliography, a simple chronological listing of publications gives the faintest hint of the historical development of a particular subject. In a citation index, a second critical aspect of historical description is provided. Conventional and citation indexing provide a simple inventory of the events. But the citation index, and the network that can be derived from it, display the interrelationships among events. (See Figure 2) The experiment on "The Use of Citation Data in Writing the History of Science" (10), demonstrated that such an historio-bibliographic model is not only possible, but also provides a legitimate starting point for the historian. The citation-network technique can eliminate much of the drudgery associated with scholarly historical writing. The historian can now devote more time to the *evaluation* of documents and less to

KEY

Direct citation connections

Indirect citation connections

Asimov's specified historical connections

Asimov's implied historical connections

NODE	DATE	NAME	NODE	DATE	NAME
1	1820	Braconnot	21	1947	Chargaff
2	1865	Mendel	22	1950	Chargaff
3	1871	Meischer	23	1950-51	Pauling and Corey
4	1879	Fleming	24	1951-53	Sanger
5	1886	Kossel	25	1952	Hershey and Chase
6	1891	Fischer and Piloty	26	1953	Wilkins
7	1900	DeVries	27	1953	Watson and Crick
8	1907	Fischer	28	1953	DuVigneaud
9	1909	Levene and Jacobs	29	1955	Todd
10	1926	Muller	30	1954-56	Palade
11	1928	Griffith	31	1955-57	Fraenkel-Conrat
12	1929	Levene, Mori and London	32	1955-56	Ochoa
13	1932	Alloway	33	1956-57	Kornberg
14	1935	Stanley	34	1957-58	Hoagland
15	1935	Levene and Tipson	35	1960-61	Jacob and Monod
16	1936-37	Bawden and Pirie	36	1960	Hurwitz
17	1938-39	Caspersson and Schultz	37	1961	Dintzis
18	1941	Beadle and Tatum	38	1961-62	Novelli
19	1943-44	Martin and Synge	39	1962	Allfrey and Mirsky
20	1944	Avery, MacLeod and McCarty	40	1961-62	Nirenberg and Matthaei

FIGURE 2

searching. Furthermore, the initial preparation of the *framework* of his historical analysis is greatly simplified.

Bibliographies Annotated by Network Coordinates

Similar capabilities may also become available to any library patron in the future. When conducting a literature search, he will receive not only a conventional bibliography, but also suitable notations for each item indicating the interrelationship with other items in the bibliography. In addition, he will receive a graph which shows these relationships more clearly. The graph will be drawn by a plotting device attached to the computer or displayed on a TV tube. For a short bibliography, this can be done with relatively inexpensive equipment. In fact, a useful map could be prepared by a conventional typewriter or line printer once the drawing instructions had been completed on the computers. This is very similar to the computer methods used in drawing PERT diagrams.

Citations display all the properties of word indexing terms because citations are, in fact, alternative and usually unambiguous symbols for concepts traditionally codified by headings. Citations can symbolize "simple" one-word concepts, like euphenics, as well as "complex" terms, like conduction-through-thin-films. But citations overcome the syntactic difficulties involved in traditional and even modern Boolean systems which often can neither distinguish between "dog bites man" and vice versa or the multiple meanings of homonyms such as plasma, aging, stress, etc.

Perfecting the Citation Method

Citation indexing can, of course, be significantly improved. The example of euphenics illustrates how this might be done. Consider Crow's book review, "Modifying Man: Muller's Eugenics and Lederberg's Euphenics," in *Science* (7). By simple title word indexing, this work would be found under "Euphenics" in a KWIC, KWOC, or any other type of title word index. However, Crow's review does not contain a direct explicit reference to Lederberg's original paper in *Nature* (6).

There is a reference to the multi-authored book in which the full talk is reprinted. The *Nature* article is a condensation of the original talk (11). By a more elaborate and expensive editing process than is now done in preparing the *Science Citation Index* (1), the editors could have established this more direct citation. At present, this is not economically sensible. The user would, of course, locate the pertinent item by Crow if he searched the index under the entry for Wolstenholme, the editor of the book reviewed (11).

Common Linguistic Properties of Citations and Words

For this practical reason and others, we must often rely on a combination search involving word and citation indexes just as we must search by a series of near synonyms in a word index. That citations share properties ordinarily associated exclusively with conventional subject headings is neither obvious nor trivial. These properties need to be carefully studied. In the process, we will learn much about the general problem of subject and content analysis. For example, O'Connor (12) has shown that word indexing, like citation indexing, can involve *a posteriori* indexing. Any indexing that involves the establishment of new correlations on the part of the indexer is *a posteriori*. This is precisely where the computer is, as yet, unable to match the intellect of man. Therefore, computer indexing is, by definition, *a priori* indexing.

Is a Perfect Bibliographic System Possible?

If we lived in a perfect world, could every paper published be relied upon to provide a "perfect" bibliography? To provide a perfect bibliography, one needs a perfect retrieval system. This is a vicious circle! To achieve a perfect retrieval system through word indexing is probably impossible and certainly expensive. Similar ideas simultaneously discovered will always be expressed in various languages and in various nomenclatures. "Perfection" through citation indexing is not possible either because similar concepts may be expressed via other communication channels than documents. Ideas are in the air and even the most pedantic and rigorous scholar may not realize the true origin of an idea. In real retrieval systems, we compromise. We approach an ideal. Since the "natural" language of words is the way people speak and communicate and the less ambiguous "unnatural" language of citations is not, we must inevitably bridge the gap between words and citations. This may not be a serious problem for the scientist. He knows the classical references—sometimes even remembering page numbers; but the librarian may not. The librarian or the student needs a link between the usually more ambiguous word or phrase and its citation synonym. That link, at present, is provided by the basic reference armamentarium of dictionaries, encyclopedias, card catalogs, word indexes, and the memory of the scientist himself.

Word-Citation Relationships

Most people tacitly assume that "in the begining there was the word," *then* there was the citation. This is not true for primordial terms.

However, for many "concepts," there may not have been a primordial term but there usually was a "citation." For the purpose of this discussion of the fundamental symbolic and unambiguous nature of citations, we need a rigorous definition which establishes the relationship between necessarily ambiguous (13) words and unambiguous citations.

A reference citation is a specification for the address of a linguistic term (word, phrase, sentence, etc.) which precisely specifies its context unambiguously. Thus, one could say, "in the beginning there was a word *and* a citation."

The closest approximation, to my knowledge, in so rigorously specifying the "citation" is found in biblical exegesis where one specifies chapter, verse, and line. It also occurs in the quotations from statutory law and other legal documents.

An unambiguous and precise citation will completely specify the serial location of any lexical occurrence. Traditional citations have been quite adequate even though they are not completely unambiguous. They have been adequate because, in most instances, some abstract "main theme" concept was referred to by the citing document. One class of exceptions to this "adequacy" is found in numerous references to books. In such cases, one frequently wishes the citing author would be more precise by specifying a particular page or paragraph. For most journal articles, this does not present a problem.

Similarity and Coupling

I would like to conclude my comparison of word and citation indexing systems by introducing the concepts of similarity (14), bibliographic coupling (15), and information content (16). The information content of a library is not a function of the number of books shelved. It is a function of the indexing or cataloging done to produce the catalog. If two libraries are using the same indexing technique and subject heading authority list, each may contain essentially the same information content even though one may contain more books than the other. A critical element is the probability of occurrence of descriptors. A library with no duplicates contains more "information" than a library with the same number of volumes but a high number of duplicates, that is, high degree of redundancy.

Descriptor Coupling

Information content is a function of the probabilities of occurrence of each descriptor. The document is defined in the search system by the set of descriptors or headings used to catalog the document. Theoretically, the actual document does not exist in the search system. If one

document is described by a given combination of descriptors and another document is described by the same set of descriptors, the two documents are equivalent. If the indexing procedures are absolutely consistent, both versions of a paper published in two different publications should always be retrieved together. We can say they are 100% similar. Their information content would be equal. For the purpose of this discussion, I shall call the degree of similarity "descriptor coupling." The degree of similarity between two or more documents is a function of their descriptor coupling. This applies to all word systems whether they are called "descriptors," "subject headings," "uniterms," etc.

Bibliographic Coupling

The similarity of two documents can also be measured by bibliographic coupling. Kessler (15) studied bibliographic coupling extensively, though it was Fano who first expressed the notion in 1956 (17). Of course, the idea of grouping similar documents by citation relationships is the essence of citation indexing. Nevertheless, one can employ bibliographic coupling to measure similarity regardless of whether one wishes to prepare a citation index. Thus, for any bibliographic system, if one needs a method for determining the degree of similarity between two documents, one can examine the number of reference or footnote citations they share in common. Just as the document is defined by its set of indexing terms in descriptor coupling, the document, in bibliographic coupling, is described by its bibliography—the set of reference citations the author has used in documenting his work. Each bibliographic citation is equivalent to a descriptor.

To test the equivalence of citations to words, one can index a given source document by using words taken from the titles of the cited papers (18) or subject headings used by an indexer to index the cited papers (1). This procedure provides additional insight into the problem of automatically or algorithmically identifying what is really new in a given source document. If the set of words extracted from the cited titles is compared to the text of the source document, then new words, such as names of new chemical compounds, stand out. They cannot appear in the list of old words compiled from the bibliography. Thus, in our previous example involving "engineering human development," this phrase would not appear in a list of the words used to index the papers Crow cited, but the word "euphenics" would occur if the book he reviewed were thoroughly indexed.

An extremely exciting application of bibliographic coupling is observed in the ASCA system (19). In the ASCA system, the user creates a bibliography of about 50 papers or books. This bibliography is his

field-of-interest profile. Suppose a paper were found which cited all or most of these 50 references. You can be certain that the new paper would have a direct relationship to the research of the user. This method can be used to uncover cases of "duplicate" research. In an ASCA report, one frequently finds that several items in a profile will be cited in a current paper. The degree of citation coupling is an indication of similarity between the retrieved paper and the interest profile.

Relevance is Subjective

I have carefully avoided the term "relevance." In general, the degree of bibliographic coupling, when three or more profile items are cited, will invariably turn up a "relevant" paper. However, even a coupling "strength" of one or two may be associated with an even "more" relevant document. Relevance is a highly subjective factor which only the user can evaluate (20). In fact, two patrons may have similar profiles but disagree on the relevance of any retrieved document (21). Relevance is not discernible on an *a priori* basis. Similarity can be measured objectively, whereas relevance is purely subjective.

Salton measured similarities based on frequency of occurrence of a term in the document itself. He concluded, in comparing similarity coefficients derived from term analysis and from citation analysis, that "citations provide a large number of relevant index terms not originally available with a given document collection, and thereby create a much more flexible retrieval process" (14).

Borko has confirmed the notion that grouping of documents "according to similarity of word content" facilitates browsing and retrieval. He did not attempt to confirm experimentally a similar claim for citation analysis (22).

Word Coupling

I have described how citation coupling is used in the ASCA system. What of word coupling? Naturally this is possible and is now employed in the ASCA system mentioned earlier. However, the natural ambiguity of language makes word coupling difficult to use unambiguously. The success of word coupling will be determined in part by the type of terminology peculiar to a given field. "Euphenics" is so specific and rare that one can rely on its low probability of occurrence (high information content) and discrimination value to turn up "relevant" information whenever it occurs in a title. A word like "films," however, is so highly ambiguous it may create as much noise as music even if other terms in combination with "films" are employed.

In the ASCA system, one can select documents based on key words appearing in titles. The use of words and word stems is restricted to those occuring in article titles. However, indexing words in titles is not equivalent to the depth and specificity of indexing achieved by citation indexing. The depth or degree of specificity of citation indexing, in some instances, can be matched by conventional non-title deep indexing as done by *Index Chemicus* or *Chemical Abstracts.* However, even these indexing systems cannot cope with the complexities of indexing mathematical formulas, complex methodology, etc. That is why these systems cannot readily answer such questions as "Where has Smith's equation XYZ, as modified by Ford, been used?"

The Indivisibility of Knowledge

Citation indexing in combination with judicious word indexing systems are essential to the notion of a unified index to science (2). It is increasingly recognized that it is a temporary expedient to segregate knowledge into discipline-oriented compartments. Every university library administrator knows the difficulties in establishing departmental libraries. Even if one could create the "perfect" classification system, it would not satisfy most users (23).

Literature of Science vs. the Literature of Interest to Scientists

There is a tendency to confuse the literature of science with the literature of interest to scientists. For example, I recently attended a meeting concerned with the documentation of oceanographic literature. There is an important, though small, segment of the scientific literature which can be called "pure" oceanography. An experienced group of catalogers could identify this literature, most of which occurs in a small number of journals. This literature of oceanography was compared to the literature "of interest" to oceanographic scientists who use the ASCA service. These scientists, however, like most other scientists, are interested in subject matter or concepts which may be found anywhere in a hard core of scientific journals. In one sense, they couldn't care less about the literature of pure oceanography, to which they may contribute, because they are in regular touch with this literature through personal contacts. It should be obvious that the chemistry of water is pertinent to oceanography, but it is also pertinent to a vast array of other problems in biology, physics, chemistry, and hundreds of applied fields. Knowledge is interdisciplinary, and our many existing fragmented approaches to bibliographical control in the past have been compromises dictated by bibliographic poverty in the midst of research affluence. To achieve uni-

versal bibliographical control requires that we think and dream, as did Wells and Bush, on a larger, though not necessarily extravagant, scale.

In this discussion, I have avoided the consideration of hardware—the problem symbolized by "Memex." This is not to minimize its importance, but it is a useful dichotomy. There is a tendency to believe that the computer is going to solve our problems. On the contrary, it has only heightened the need for a more perceptive understanding of the basic philosophy of all indexing systems.

REFERENCES*

1. E. Garfield, "*Science Citation Index*—A New Dimension in Indexing," *Science* 144, 649-654 (1964).
2. E. Garfield, "A Unified Index to Science," *Proceedings of the International Conference on Scientific Information, 1958* (Washington, D.C.: National Academy of Sciences-National Research Council, 1959), pp. 461-474.
3. E. Garfield, "World Brain or Memex?—Mechanical and Intellectual Requirements for Universal Bibliographic Control," *Proc. Symp. on "Foundations of Access to Knowledge,"* Syracuse, 1965. (In press)
4. V. Bush, "As We May Think," *Atlantic Monthly* 176, 101-108 (1945).
5. For an interesting discussion of citation and cognate "historical" methods see R. A. Fairthorne, *Towards Information Retrieval* (London: Butterworths, 1961), p. 192.
6. J. Lederberg, "Molecular Biology, Eugenics and Euphenics," *Nature* 198, 428-429 (1963).
7. J. F. Crow, "Modifying Man: Muller's Eugenics and Lederberg's Euphenics," *Science* 148, 1579-1580 (1965). ("It is a measure of their impact . . . that the word has already come into general usage.")
8. C. N. Mooers, "Zatacoding Applied to Mechanical Organization of Knowledge," *American Documentation* 2, 20-32 (1951).
9. E. Garfield, "Citation Indexing: A Natural Science Literature Retrieval System for the Social Sciences," *American Behavioral Scientist* 7(10), 58-61 (1964).
10. E. Garfield, I. H. Sher, and R. J. Torpie, "*The Use of Citation Data in Writing the History of Science* (Philadelphia: Institute for Scientific Information, 1964), 76 pp. See also D. J. deS. Price, "Networks of Scientific Papers," *Science* 149, 510-515 (1965).
11. J. Lederberg, "Biological Future of Man," in G. Wolstenholme, Ed., *Man and His Future* (Boston: Little, Brown and Co., 1964), pp. 263-273.
12. J. O'Connor, "Mechanized Indexing Methods and Their Testing," *Journal of the Association for Computing Machinery* 11, 437-449 (1964).
13. S. Gorn "Treatment of Ambiguity and Paradox in Mechanical Languages," *Proceedings of Symposia in Pure Mathematics* 5, 201-218 (1962).
14. G. Salton, "Associative Document Retrieval Techniques Using Bibliographic Information," *Journal of the Association for Computing Machinery* 10, 440-457 (1963). (Quote from p. 456).
15. M. M. Kessler, "Bibliographic Coupling Between Scientific Papers," *American Documentation* 14, 10-25 (1963).
16. E. Garfield, "Information Theory and Other Quantitative Factors in Code Design for Document Card Systems," *Journal of Chemical Documentation* 1, 70-75 (1961).
17. R. M. Fano, "Information Theory and the Retrieval of Recorded Information," in J. H. Shera, A. Kent and J. W. Perry, Eds., *Documentation in Action* (New York: Reinhold Publishing Corp., 1956), pp. 238-244.

18. M. E. Stevens and G. H. Urban, "Training a Computer to Assign Descriptors to Documents: Experiments in Automatic Indexing," *AFIPS Conference Proceedings*, Volume 25, *1964 Spring Joint Computer Conference* (Baltimore: Spartan Books, Inc., 1964), pp. 563-575.
19. I. H. Sher and E. Garfield, "New Tools for Improving and Evaluating the Effectiveness of Research," *in* M. C. Yovits, D. M. Gilford, R. H. Wilcox, E. Staveley, and H. D. Lerner, Eds., *Research Program Effectiveness*, Proceedings of the Conference Sponsored by the Office of Naval Research, Washington, D.C., July 27-29, 1965 (New York: Gordon and Breach, 1966), Chapter 7, pp. 135-146.
20. L. B. Doyle, "Is Relevance an Adequate Criterion in Retrieval System Evaluation?" SP-1262 (Santa Monica, California: System Development Corporation, 1963).
21. R. S. Taylor, "The Process of Asking Questions," *American Documentation* 13, 391-396 (1962). As cited by Doyle in Reference No. 20.
22. H. Borko and M. D. Bernick, "Toward the Establishment of a Computer Based Classification System for Scientific Documentation," *Technical Memorandum TM-1763* (Santa Monica, California: System Development Corporation, 1964), 47 pp.
23. H. Marron, "Science Libraries—Consolidated/Departmental?" *Physics Today* 16(7), 34-38 (1963).

* For editorial reasons the form of the citations accompanying this article do not conform to JLH style.—Ed.