from nonuse of data. Much more support is needed for retrieving and organizing data from the unorganized primary literature—even if the data must be left unevaluated. Current, evaluated data are, of course, ideal. And while we are organizing data, we should also plan for genuine information retrieval, including theories, propositions, and conclusions, and their evaluation.

Information within arm's length and the rebirth of the personal library are important objectives in redesign of the technical literature. Techniques are here for centralized cataloging, indexing, and abstracting for personal collections. The several papers that emphasize nearby information are indication of the profound significance attached to this aspect of redesign.

Whether it is called continuing education, adult education, keeping up with the literature, preventing functional obsolescence, or upgrading technical personnel, there is tremendous interest and importance in the new and valuable techniques for painless, perfect learning as a way of keeping abreast. Audio-visual and computer-aided instruction have a rapidly expanding role in redesign.

And finally, the precise ways in which computers and related equipment can serve us best are becoming clear, thanks to papers, such as these.

Condensed Technical Literatures*

CHARLES L. BERNIER†
The Squibb Institute for Medical Research, New Brunswick, N. J. 08903

Received July 25, 1968

Scientists have too many words, too many languages, and too many journals to read. Condensation by evaluation, done by authorities, such as authors, editors, and reviewers, can be expressed in sentences like aphorisms stating conclusions, results, intentions, etc. Condensed literatures of organized collections of terse statements would not be substitutes for other literatures, but would provide evaluated guides to them. Backlogs of reading might be avoided because the past would be embodied in the statements that could help professional people to keep up in fields one or two orders of magnitude broader than they now can read.

"But the problem in fact is how to pass through multiplicity so as to transcend it, and not at all how to escape it."[1]

Professional people have too many words to read in the time available. Too many of the words are in languages other than English and the words appear in too many journals, reports, and books. This is the problem that many of us have been trying to solve for at least a quarter of a century. Chemists, biologists, engineers, and medical men, among others, are finding it increasingly difficult to keep up. Serendipity supplants complete searches. Greater specialization to reduce the amount to be read is successful only to a limited extent. In multidisciplinary fields, such as medicine and pharmaceutical research, increased specialization can lose vital information. Complete searches often yield too many documents to be read. Reading this year's output of 250,000 to 500,000 medical papers, for example, at one hour a day would take two and a half to five centuries. At this rate a medical researcher can read only 0.2–0.4% of the literature. The approach to the primary literature through indexed abstracts is also unsatisfactory, in this respect, because the primary literature retrieved is just as large as before. Plans for condensation of medical and other literatures and the use of computers for selection of references or documents have not solved the problems of too much to read and too little time in which to read and digest it.[2, 3, 4, 5, 6]

The purpose of this paper is to present an approach to solving the problem of too much to read in the limited time available, and to invite comments that will enable evaluation and improvement of this approach.

We need condensed technical literatures[7] that aid us to keep up with pertinent literature in fields perhaps one or two orders of magnitude broader than we now can read. Since we cannot read all necessary words, the literature must be condensed in some way. All condensation requires evaluation—the part selected is considered to be of greater value than the part rejected. It is also apparent that selection, evaluation, and condensation must be done by those knowledgeable in the subject field—otherwise nobody would take time to read the condensed product—nor should they read it. Knowledgeable people who see the results of new work first are: author, editor, and reviewers. These authorities are in a position to produce the promptest condensations. These knowledgeable people will have opinions about the manuscripts that they read. They will have come to conclusions about the manuscript. These opinions and conclusions can usually be expressed in a carefully written sentence. The sentence may resemble an aphorism ("Exclusively of the abstract sciences: the largest and worthiest portion of our knowledge consists of aphorisms." Funk and Wagnalls New Standard Dictionary.) of Hippocrates—e.g., "Persons who are naturally very fat are apt to die earlier than those who are slender."[8] A similar aphorism from current actuarial data might read, "For males between 15 and 69, those who are 30% over the average weight have an excess

* Presented before the Division of Chemical Literature, Symposium on Redesign of the Technical Literature, 156th Meeting, ACS, Atlantic City, N. J., September 1968.

† Present address: State University of New York, Buffalo, N. Y. 14214.

mortality of 42% over standard experience" (derived from actuarial data of the Metropolitan Life Insurance Co.). As a personal example, after I read the Surgeon General's Report on Smoking and Health,[9] I came to a conclusion: "The prudent course for me is not to start smoking cigarettes." For my future conduct, in this respect, I found that I could condense this large report into one terse sentence. Later reports, perhaps from the tobacco industry, might cause me to alter my conduct. These later reports could probably likewise be condensed into other terse statements.

The kinds of terse sentences or aphorisms generated from the medical literature, for example, would not usually be in the form of admonitions, but in the form of conclusions, results, intentions, etc. The authority writing the aphorism could be instructed—e.g., as follows: "What is the one most important conclusion (result, intention, etc.) derived from your experience and knowledge as related to the paper you have just read or written? State the conclusion, etc., as one terse, informative sentence. In order to make the sentence informative, be specific by omitting generalizations and pronouns. Make the sentence intelligible in itself without dependence upon other sentences or references." Examples of aphorisms expressing conclusions are:

1. Sunlight is now established statistically as the number one physical carcinogen.
2. Smoking of cigarettes may lead to lung cancer and other diseases.
3. KWIC indexes omit entries, scatter like subjects, provide inadequate internal and external guidance, and have unnecessary entries.
4. KWIC indexes cost the user more time and more lost entries than do standard subject indexes.

Examples of aphorisms expressing findings or results are:

1. About 10% of all lung-cancer patients have never smoked or have used tobacco only rarely.
2. Of 250,000 white U. S. veterans, cigarette smokers have a higher death rate from lung cancer and almost every other disease than non-cigarette smokers.
3. A person with one cancer is six times more apt to develop a second cancer than would be expected by chance alone.
4. Norhydroxyprogesterone caproate, corresponding to a clinical dose of 1 gram, caused regressive changes in adenocarcinoma of the uterus in vitro, consisting of severe cell necrosis and disappearance of the recently established cell layer.
5. Cyclophosphamide and nitrogen mustard *N*-oxide inhibited DNA synthesis in vivo in crypts of Lieberkuhn at much lower doses than those inhibiting glycolysis.

Examples of aphorisms expressing intention are:

1. I will stop smoking cigarettes pending further data.
2. Test the synergism between hydroxyurea and urea in therapy of mouse leukemia.
3. Study actuarial data on ischemic cardiovascular disease in relation to dietary sucrose as indicated by Yudkin's data on elevated sucrose in diets in arteriosclerosis and coronary embolism as described in *Lancet*, **4**, July 1964.
4. Prepare an experimental Thesaurus of Aphorisms on Cancer for distribution to oncologists to determine its usefulness.

Condensed literatures consisting of collections of aphorisms would not be substitutes for the primary and other literature. They would serve as critical indexes to the other literature if they were not accepted, at least temporarily, as valid without further study. The terse sentences would resemble critical or evaluative annotations. The aphorisms would bear coded or abbreviated guides to references to the primary documents; these documents would supply the substantial information needed before taking action based upon the statements. This kind of condensation should enable sharing more of the wisdom accumulated on a subject than does the primary literature or abstracts. Expert opinion, judgment, advocacy, and guidance could be obtained conveniently from the different kinds of condensed literatures. Omitted from condensed literatures could be: proof of statements, experimental details, references to other work, historical introductions, data, results, procedures, graphics, etc.

The terse, evaluative statements could be collected periodically, say daily, and then organized and published. A newspaper of such terse statements in a subject field would seem appropriate, although collections in periodicals and books would be especially useful for edited cumulations of the most valid statements. Such cumulations would resemble reviews. The statements could be cumulated, revised, deleted, reorganized, and updated. Identification of authors of statements could be optional. Contradictory statements would be published and would indicate need for further research or discussion.

Categorization of the evaluative statements would facilitate comparison and selection. Unorganized statements would all have to be read to ensure complete coverage. Statements could be organized by Universal Decimal Classification (U.D.C.) numbers assigned to each statement. Such numbers would enable organization by every aspect of the statements. Alternatively, the statements could be organized by automated clustering techniques based upon words, word frequency, and syntax. For presentation of expected correlations among statements, newspapers or reference works containing organized collections would seem most efficient. For making unexpected correlations among large numbers of statements, computers would seem more appropriate.

Authorities capable of writing valuable statements that would be accepted as valid by those working in the subject field include: authors, reviewers, and editors. Abstractors who were subject experts could also provide terse statements; however, their statements would be delayed by the time required for publication of the primary literature.

While the terse, authoritative statements discussed here are intended to be valid, they could not replace the experimental confirmation or the reverse that would appear and generate succeeding statements. Equivalent or identical statements could be combined and given a number to indicate how many were combined. This would give a measure of consensus and perhaps of confidence that could be placed in the statement. Validity of the statements could be increased by: 1) using only those derived from manuscripts or published documents; 2) careful selection of reviewers and authors by editors; 3) selection of journals; and 4) nomination of other authorities by professional associations. Usefulness and reliability of the literatures formed by organized collections of terse statements could be evaluated through actual use. As confidence in condensed literatures was established through experienced reliability and effectiveness, it is possible that

these literatures would be consulted first by research workers in keeping up in their specialties.

A study of about a thousand aphorisms on cancer-research results gave an average statement of about 20 words and a five-component U.D.C. number. The 20 words represent a condensation of about two orders of magnitude in the primary literature. For reading all statements in a collection with each statement repeated for each component of its U.D.C. number, the condensation would amount to about twenty to one. In the field of medicine, for example, all statements derived from research papers on cancer could be collected. Within this collection, the statements could be arranged by kind of cancer—e.g., monocytic leukemia. A statement about calcium in epidermal tumors could be classified into three groups: those about calcium, skin, and the specific kind of tumor. For the entire field of medicine, since there are a quarter to a half a million medical papers published each year, a newspaper of aphorisms for all of medicine might have 1000 to 2500 statements published daily (a five-day week). Doubtless, some medical papers would not generate aphorisms. If a statement were repeated an average of five times to enable placement into all relevant groups, then the daily edition of the newspaper might have a total of five to 12,000 statements. If the newspaper for all of medicine, for example, were found to be too bulky, then specialized newspapers could be created to cover—e.g., neoplasms, cardiology, or infectious diseases. The classification of the aphorisms in the newspaper would enable the medical reader to limit his reading those statements for which he had the time. The user of a newspaper of medical aphorisms would have no backlog to read because each statement would present the current state-of-the-art by embodying the experience of the authority combined with information from the medical paper being reviewed. The medical papers that stimulated the aphorisms would be referenced in the back pages of the newspaper. Translation costs would be saved because not all papers would need to be translated, and translation of terse annotations would be less costly than translation of abstracts or extracts.

We have the technology to tailor technical newspapers composed of categorized, terse statements to fit the needs of each professional person, should such products be found useful. There is some question as to whether newspapers designed for individuals could follow their careers with sufficient accuracy to avoid loss of pertinent information. Experience with selective-dissemination systems should be valuable here. However, before tailormade publications are created, we will probably want to try newspapers for all or parts of some profession.

Production of aphorisms for a collection could be initiated by professional societies and by editors of technical journals. Authors and reviewers would participate by writing terse, evaluative statements of requested kinds according to rules. Statements by authors could be made a special feature of their technical papers.

It would seem that solutions to problems of science and technology could be brought promptly, by means of collections of the condensed literatures, to the attention of those who need them. New courses of action advocated by authorities in a field might be accepted more readily because they would normally be anticipated by a series of statements that logically and gradually lead readers to the final statement of action advocated. Thus, the condensed literatures might motivate readers and reduce cultural lag. Although action, other than reading, might seldom be based directly upon the condensed literatures, they could serve as guides to the primary literature.

In conclusion, I believe that there is a need for condensed technical literatures that aid those in professions in keeping up with pertinent information from fields perhaps one or two orders of magnitude broader than they now have time to read. For acceptability, the condensed literatures should embody evaluation by authorities in the subject field. Evaluative literatures comprising terse statements would give reactions of authorities to primary literature as guides to it. Evaluative literatures would condense by evaluation. Such condensed literatures support the objective of action as a proper product of research and development more closely than does the primary literature with its detailed presentations in many languages of facts that have now become too numerous to read. Reading is the minimum action desired for the verbal output of research and development. Cogent, condensed literatures containing perhaps a little conflict gradually resolved harmonize well with the concept of action as a desirable result of applied research and development. Pooling of aphorisms from different authorities in a specific subject area should gradually generate cogent literatures of consensus. Such terse, evaluative literatures offer hope of helping professional people to keep up with the growing literature.

ACKNOWLEDGMENT

I wish to thank Erwin Di Cyan, L. C. Kingsland, J. C. Leondar, D. B. Remsen, and L. M. Rosa for their contributions to the drafts of this manuscript.

LITERATURE CITED

(1) Marcel, Gabriel, "Being and Having—An Existentialist Diary," p. 86, Harper & Row, New York, 1965.
(2) Bush, V., *Atlantic Monthly* **176**, 101 (July 1945).
(3) Heilprin, L. B., and F. L. Goodman, Analogy between Information Retrieval and Education, Proceedings of the Symposium on Education for Information Science, p. 13, Spartan Books, Washington, D. C., 1965.
(4) Page, I. H., *Mod. Med.* **28**, 22 (1967).
(5) "CAS Today: Facts and Figures about Chemical Abstracts Service," 60th Anniversary Edition, Chemical Abstracts Service, Columbus, Ohio 43210.
(6) Orr, R. H., and A. A. Leeds, "Biomedical Literature: Volume, Growth, and Other Characteristics," *Federation Proc.* Part I, **23** (6), 1310 (Nov.-Dec., 1964).
(7) "Science, Government, and Information—The Responsibilities of the Technical Community and the Government in the Transfer of Information," A Report of the President's Science Advisory Committee (Weinberg Report), The White House, Jan. 10, 1963, p. 2, Superintendent of Documents, USGPO, Washington, 25, D. C., 25 cents.
(8) "The Genuine Works of Hippocrates," translated by Francis Adams, The Sydenham Society, London MDCCCXLIX, Vol. **2**, 713.
(9) Smoking and Health, Report of the Advisory Committee to the Surgeon General of the Public Health Service, Public Health Service Publication **1103**, USGPO, Washington, 25, D. C., $1.25.