# Modification of Dissemination Channels for Scientific Information*

STEPHEN A. ROSSMASSLER
Office of Standard Reference Data, Institute for Basic Standards,
National Bureau of Standards, Washington, D. C. 20234

Received October 24, 1968

**Traditional channels for the dissemination of scientific information are based on primary publication in a scientific journal, with reader guidance provided by an abstract journal. These channels are designed primarily to meet the author's requirement and are document-oriented. The present growth of scientific information calls for a redesign of the technical literature. It is proposed that the single piece of scientific information be recognized as the unit of flow into the distribution system. The resulting need for organized efforts for selection and evaluation of information, and increased emphasis on users' requirements lead to modification of dissemination channels. Possible changes in present publication practices are considered.**

The individual scientific paper in a learned journal is, by tradition, the basic transfer unit for scientific information. The simple and widely used dissemination channel based on this principle is the following:

I. Originator of information → Primary publication → Reader/user

More than 50 years ago, it became obvious that because of the growing volume of material available in journals, and the increasing number of the journals with overlapping coverage, readers and users needed a guide to their reading of the primary literature. Abstract journals came into being to perform this function. As a general practice, professional societies have assumed responsibility for issuing abstract journals appropriate to the interests of their membership. The dissemination channel familiar to most scientists, based on a combination of primary and abstract sources, is the following:

II. Originator of information → Primary publication → Abstract journal → Reader/user

The primary publication may, of course, equally well be a patent, or an internal company report, or a government report. The abstract journal may have some other title but it performs the same function.

In the flow of information through the first and second channels of dissemination, certain responsibilities are recognized. The author has responsibility for generating and assembling his information. Usually, he is under some constraints as to form of presentation, amount of detail, language, and style. The primary journal takes responsibility for a variable amount of screening and selection before a paper is permitted to continue through the channel. The primary journal also makes known the area of science or technology which it is covering, so that readers can use channel I instead of channel II if they have the time and energy, and a sufficiently narrowly defined field of interest. The abstract journal has responsibility for setting up comprehensive and continuing surveillance of a relatively large number of primary journals and for defining what that coverage is. It must also establish and announce a consistent set of categories and subdivisions of its contents, so that the user can be selective in his scanning of abstracts. The abstract journal also has responsibility for setting up and maintaining a consistent depth of analysis of the primary literature, so that the user can have confidence, based on continued successful use, in the screening which the abstract journal provides, and its ability to lead him to the references he wants.[1]

It is recognized that the abstract journal serves principally as a guide by which the reader/user can select and then study in detail those primary publications which are of real interest to him. In this sense, channel II can be modified:

III. Originator of information → Primary publication → Abstract journal → Primary publication → User

During the period since 1946, we have seen an expansion of the literature of science and technology which has been so rapid that it is commonly called an explosion. The number of articles written for original publication has grown greatly and the number of different journals handling original publications has increased correspondingly. Most busy scientists with a broad range of interest now find that channel I is no longer practical for them. The volume of material to be abstracted has required a continuing growth in the size of the abstract journals themselves,[2] and for many workers, channels II and III are hardly manageable any longer.

* Presented in Symposium on "Redesign of the Chemical Literature," Division of Chemical Literature, 156th Meeting, ACS, Atlantic City, N. J., September 11, 1968.

At the same time, librarians have become more deeply involved in science and scientists more interested in the handling of publications; the whole field of information science has come into a first stage of maturity. Computers now offer the capability of handling fragments of information at speeds which human hands, eyes, and brains cannot match.[3] Specialized libraries, selective dissemination of information, and a wide range of publicly available specialized information handling services are now generally accepted. These new activities have been extremely valuable to those who try to keep abreast of their field and have served to postpone the ultimate crisis. However, they are still in the traditional mold, in that they all relate themselves to the primary publication as the basic flow element of information dissemination. They are document-oriented systems.

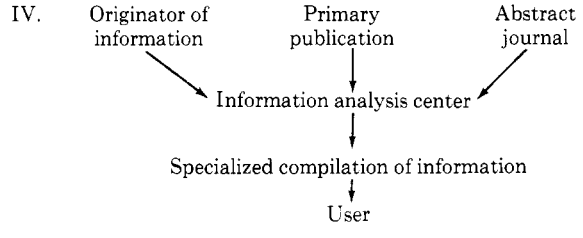
Let us look at the dissemination of scientific information in a somewhat different way. In a representative paper in the scientific literature, the author has—presumably with intention—placed a number of pieces of information which he wishes to communicate. These pieces of information can be considered to be the elements, $m_{ij}$, of a matrix, $M_A$ (a two-dimensional matrix is assumed for simplicity of typography). (The application of matrix concepts to information theory is treated by Meredith[4] as a general subject. The present analysis is much more specifically directed to certain elements of information in a paper.) The fact of publication of the paper may in itself fulfill certain of the objectives which the author had in mind when he prepared the paper. In addition, the existence of a written document on the topic may discharge some further responsibilities. It may, for example, constitute a concluding report which is necessary as one of the terms of a grant he received to support his experimental work. It may fulfill certain requirements of publication of thesis material for the granting of an advanced degree. However, the communication of the information elements is obviously not fully satisfied by a publication procedure which lacks explicit means for directing the information content to the person who should receive it.

Now, the potential user of some element of the information also has a matrix of information needs, $N_U$, with elements, $n_{kl}$, relating to a specific problem he may face. Some of the elements of the user's matrix may correspond to elements of the author's matrix. To the extent that $n_{ab}$ corresponds to $m_{cd}$, completion of the communication process may be as simple a matter as identifying the correspondence, locating and extracting $m_{cd}$ from $M_A$, and placing it in $N_U$. However, the correspondence is not frequently so exact that a single element can be extracted from one matrix and placed into another. Morever, $M_A$ is really not a self-contained entity, but is a part of a larger matrix of the author's information; construction of $M_A$ required elimination of many potential elements during the process of writing the paper. The mental selection by which the author decides what to include must derive from the purposes of his paper, but in the selection process, he may eliminate certain material and information at his disposal which would be highly valuable from the user's viewpoint.

Equally important, $N_U$ itself is a part of a larger matrix of needs. The user of scientific information is seldom so single-minded that he wishes to concentrate exclusively on and look for a particular matrix element at a time as he reads journals. Rather, his habits usually include a continuing process of accession and digestion of information elements on a variety of special subjects throughout his scientific work day.

The difficulties encountered in establishing the correspondence between $n_{ab}$ and $m_{cd}$ may go far to explain the comment of many scientists that they prefer to get necessary information from personal associates, whether by direct conversation or by telephone. The correspondence between $n_{ab}$ and $m_{cd}$ is easier to achieve in a two-way conversation,[5] and the neighboring matrix elements of both the source matrix and the need matrix can be readily visualized by both parties, so that the information exchanged is larger in quantity. Another reason for this preference is the relative inflexibility of the primary publication. The formal requirements of writing a scientific paper are such that the matrix, $M_A$, which the author presents is not really the most convenient information-transfer device from the user's point of view. It is rigidly constructed, and edited to contain only the elements essential to the main course of the originator's line of reasoning.

It is suggested, therefore, that we should give more recognition to the value of the individual information element as the unit of flow into the dissemination channel. A major modification of the channel results from such a change in emphasis. Writing, refereeing, publishing, indexing, abstracting, and distributing the whole document, while still important, are no longer sufficient. The specialized information analysis center[6] is added to the path between originator and user, as shown in channel IV.




The center concerns itself with a limited and well-defined scientific topic, and provides expert scientific judgment in this area. Out of the matrix, $M_A$, which comprises the entire information content of a single document, the scientists in the center may select a small subset, $m(1)$. From other papers they may select similar subsets, $m(2)$, $m(3)$.

The entire scientific literature is screened comprehensively in this manner. The chosen elements are subjected to close scrutiny and appraisal by the experts. If necessary, private correspondence or direct communication with the originating authors may be required to obtain other material which they did not include in their publication. The goal of the center can be defined as a specialized matrix which is a systematic and orderly presentation of selected, reliable information on a particular topic for direct utilization by the ultimate user. The new matrix is usually discipline-bounded and user-oriented in its content and format.[7]

This alternate path provides two immediate advantages:
a) the vast bulk of scientific literature is condensed by

selection and clearly structured so that the user has an unambiguous path to the information he wants, and b) the selection process provides increased reliability of information, because of the expert scrutiny and analysis which have been brought to bear on the original material. The analysis of the original publications and the synthesis of a new matrix can focus on any one of a variety of concepts, such as a unifying theory, a summary of experimental methods, techniques and apparatus, a critical review and analysis of the state of the art in a specialized field of science, or a compilation of critically evaluated numerical data on the properties of substances.

How does this alternate path affect the existing utilization of the chemical literature and the working habits of scientists? For the expert in his field, reading the traditional journal to keep up with that field, it affords no need for change. He wants to know what results the author obtained and how he went about obtaining them, as well as the conclusions he derived from them. He will be satisfied with nothing less than reading the papers of his peers in the way they were written. For the outsider, a novice in a specialized field, this new channel provides a reliable guide until the novice becomes an expert. For the problem solver, it provides direct and immediate answers of established reliability which he can use with confidence in his work. For the reflective scientist, it can provide, in a specified area, an orderly structuring of the well-known quantitative data, and thus afford a starting place for new theory and a stimulus to speculation. For the systematic transfer of knowledge from one field of science into another field, it provides convenient and rational packaging.

Information analysis is not, of course, a new activity. Analysis of the published literature, critical reviews of special topics, and development of unifying theories have long been recognized as vitally important scientific functions which demand the highest level of competence. Work of this kind has gone on for years as part of the overall scientific communication process, and no major modification of the technical literature has occurred.

It is suggested now, however, that changes are already taking place which require some substantial redesign of the literature. Abstract journals are becoming too expensive for individual scientists to buy, and too overwhelming in content for them to use. Subscriptions are not keeping up with the growth of the scientist population. Primary journals are becoming more and more specialized, and reading of technical material outside of one's direct field of specialization is a luxury few can afford.

At the same time, concerted efforts are being made, for the first time in 30 years, to develop a comprehensive information analysis and critical data evaluation system, at least for the numerical data of the physical sciences.[8] If these efforts can provide adequate coverage of the primary literature, and an effective dissemination channel for the important information elements selected from it, redesign will become feasible.

A number of redesign possibilities for changing the basic literature come to mind. Not all of them are compatible with one another. They are listed below as a basis for further consideration.

1. Primary publication can become a formal and highly concise statement of starting point and conclusions. All details of techniques, observations, and interpretation can be handled separately, going simultaneously to a repository (where they are available for public perusal) and to an appropriate information analysis center where the observations, numerical data, etc., are digested and become part of the input-output stream of the center. This possibility would permit considerable reduction in the volume of primary publication.
2. Publication in printed form may be reduced to abstract only. Papers would be written and refereed as they are now, but those portions which could be separated out without loss of unity would be sent in parts to a number of specialized analysis centers—one for new theory, one for new techniques or apparatus, and one for numerical data. Regularly scheduled output from these centers would be the major information resource for the user. The originator of the information would receive credit not only in the abstract, but also in each of the specialized outputs which made reference to his work. Admittedly, such dismemberment might not be applicable to all papers, but for a substantial number it would be possible.
3. Primary publication and abstracting can continue as they now exist. However, with much more extensive systematic exploitation of the results of past experiments, unnecessary duplication of measurements will be reduced. Scientists will know far more exactly just what areas provide the greatest opportunity for making an important contribution of new knowledge; and can plan their work much more efficiently.

This author is not at present advocating any one of these possibilities as the best single basis for redesigning the scientific literature. It is proposed, however, that a change in dissemination channels is already taking place, that increased recognition can and must be given to the user's need for ready access to just that information which applies to his problem, and that primary publication is no longer an adequate information channel for the majority of users.

## LITERATURE CITED

(1) "Science, Government, and Information," A Report of the President's Science Advisory Committee, pp. 15–16, January 10, 1963; Superintendent of Documents, Washington, D. C., 25 cents.
(2) "CAS Today," 60th Anniversary Edition, Chemical Abstracts Service, Columbus, Ohio, 1967.
(3) Tate, F. A., "Progress Toward a Computer-Based Chemical Information System," *Chem. Eng. News* **45**, 78 (1967).
(4) Meredeth, G. P., "Semantic Matrices," Vol. **2**, Proceedings of the International Conference on Scientific Information, Washington, D. C., 1958, National Academy of Sciences-National Research Council, Washington, D. C., 1959.
(5) Taylor, R. S., "Studies in the Man-System Interface in Libraries," 1967, AD **659468**, Clearinghouse for Federal Scientific and Technical Information, U. S. Department of Commerce, Washington, D. C.
(6) "Science, Government, and Information," A Report of the President's Science Advisory Committee, pp. 32–33, January 10, 1963; Superintendent of Documents, Washington, D. C., 25 cents.
(7) Garvin, D., and H. M. Rosenstock, "Two NBS Data Centers, Chemical Kinetics and Mass Spectrometry," J. CHEM. DOC. **7**, 31 (1967).
(8) Brady, E. L., and M. B. Wallenstein, "National Standard Reference Data System—Plan of Operations," NSRDS-NBS-1, Dec. 30, 1964; Superintendent of Documents, Washington, D. C., 15 cents.