Toward a National Chemical Information Network*

WALTER M. CARLSON
Department of Defense, Washington, 25, D. C.
Received November 12, 1964

The privilege and the honor afforded by being asked to address this luncheon audience of so many of my friends and fellow workers is greatly appreciated. I am particularly pleased to have this opportunity to share with you a pattern of thinking about national information systems that has evolved recently. Further, it is timely to discuss with some realism where we seem to stand today on the prospects for a national chemical information network.

Most chemists have enough difficulty finding what information they need in the local facilities, and it seems farfetched and "blue-sky" to them for anyone to be talking about national systems. Whatever your views of the present situation may be, I think there is general agreement that more attention will be given in the next few years to the information network concept. The hardware capability for such a network is well assured; in fact, the capability exists today. The real question is when, and under what conditions, the chemical community will determine that an economic need exists for a network that will tie together a wide range of chemical information services.

Before addressing this question in any detail, I would like first to lay some groundwork on the concepts that enter into discussion of large, nationally important information systems.

The basic concepts may be analyzed three ways (Figure 1). The purpose of technical communication is to provide current awareness or to respond to some task-oriented search. The scope of technical communication may embrace scientific information (*i.e.*, new knowledge) or technical information (*i.e.*, application of knowledge). The orientation of the audience may be that of general interest, a mission or project interest, or a management interest. It turns out that the people who are talking these days about large information systems for chemists are referring almost exclusively to a system that is *task-oriented, scientific,* and *general.* There are other systems that can be formulated by other combinations of these concepts, but the one of importance to our discussion today is task-oriented, scientific, and general.

PURPOSE	SCOPE	ORIENTATION
CURRENT AWARENESS	SCIENTIFIC	GENERAL
	TECHNICAL	MISSION
TASK-ORIENTED SEARCH		MANAGEMENT

Figure 1.—Distribution of information basic concepts.

* Presented at the Division of Chemical Literature Luncheon, 148th National Meeting of the American Chemical Society, Chicago, Ill., Sept. 2, 1964.

The distribution of information becomes a highly useful analog for the technical communication process. In fact, the time has come to start treating information as a commodity—a resource. The most elementary distribution procedure is direct from producer to user. There are many ways in which this direct contact takes place. They range from 100% down to zero in ability to provide two-way communication as one goes from personal conversation to closed-circuit TV (Figure 2).

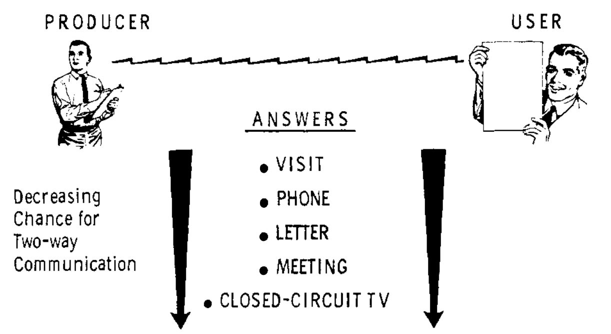

Figure 2.—Distribution of information (task-oriented, scientific, and general).

A study of chemists undertaken by Case Institute in 1958, with help from the American Chemical Society, showed more time spent on scientific communication than any other professional activity. Also, the study indicated that there is an average of one listener for each speaker in scientific communication involving chemists.

It becomes too expensive for all people to get all answers by direct contact, and converters are used to document the information for later use. The publishers of journals, reports, books, and patents are all involved as converters. There are factors which create a generalized hierachy of these conversion processes with respect to the amount of time it takes to convert new knowledge into the finished, recorded format (Figure 3).

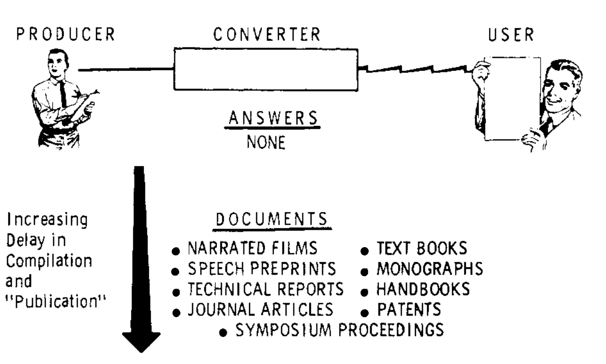

Figure 3.—Distribution of information (task-oriented, scientific, and general).

The same Case Institute study resulted in estimates that there are seven readers for each writer in task-oriented, scientific, general communication of chemical information.

When it becomes impractical for all users to deal directly with the publishers (or converters), a need arises for wholesale services. Any one reader sees less than ½ of 1% of the available journal articles, for example. Typical wholesalers of answers are the information analysis centers of industry and the government. Typical wholesalers of documents are the National Library of Medicine or the Defense Documentation Center. Typical wholesalers of references are Chemical Abstracts Service or the National Referral Center at the Library of Congress (Figure 4).

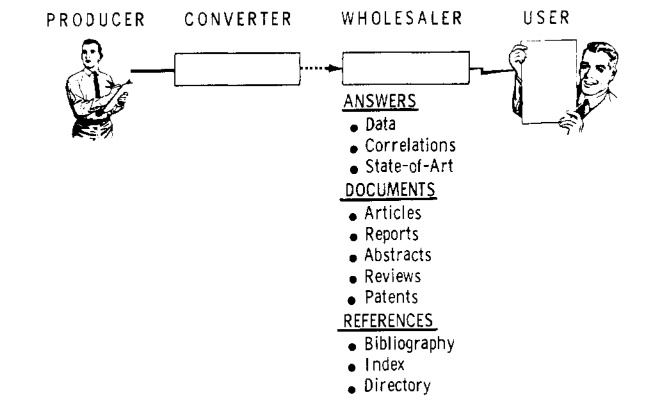

Figure 4.—Distribution of information (task-oriented, scientific, and general).

Finally, a retail service is needed when direct contact between user and wholesaler becomes uneconomical. The local technical specialist provides answers and accumulated know-how for his associates. The local library or document center provides a wide range of services aimed at finding documents for the individual user; they also serve to direct the searcher's steps to the proper source of documents or answers (Figure 5).

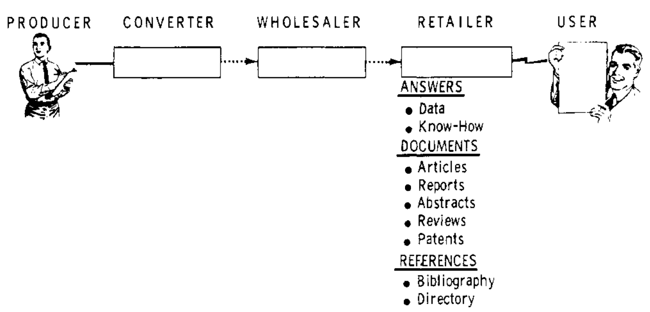

Figure 5.—Distribution of information (task-oriented, scientific, and general).

In actuality, all of these channels act simultaneously, each being used according to a host of factors determined by the user and the relative availability to him of the producer, converter, wholesaler, and retailer (Figure 6). The situation, then, is a complex mixture of the different channels operating simultaneously. I particularly wish to call your attention to the fact that we do not have any real measurement of how much technical communication takes place in each of these channels or, therefore, what

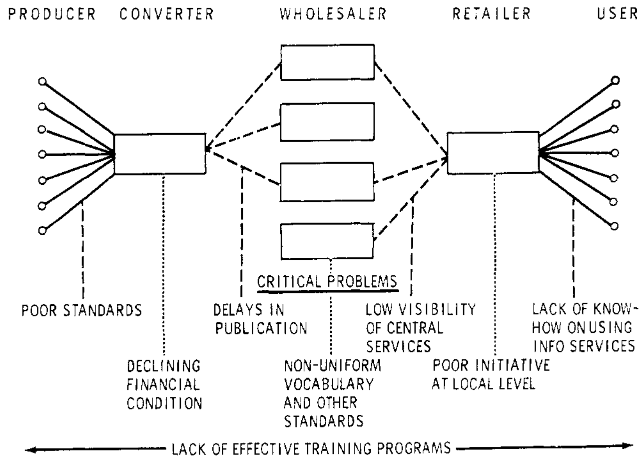

Figure 6.—Distribution of information (task-oriented, scientific, and general).

their relative importance may be. I sincerely hope that someone will accept the stimulating challenge of trying to find out more about the relative use of these different channels of distribution.

When looking at task-oriented, scientific, and general distribution in this manner, it is possible to identify where some of the critical problems are. The underlying problem, of course, is one of education or training across the board from producer to user. The user, in general, doesn't know how to use the information resources available to him. The retailer often lacks the initiative to seek out either the user or the wholesaler. There is low visibility of the central services even though the local people are trying to find them. The wholesalers each have their own procedures, and common standards are clearly for the future. The documents are frequently delayed in publication, partly because the publishers are suffering financial pains at the very time they have more material thrust upon them to be published. Finally, the producers have very little incentive to learn and adhere to a reasonable set of standards for presenting their material. Without doubt, you can provide many more problems, and you will have different opinions about the relative importance of the problems I have listed.

A national network for chemistry does exist (Figure 7). It is not mechanically or electronically linked together, however, today. The producers are the men and women in the laboratories, technical sales, and development organi-

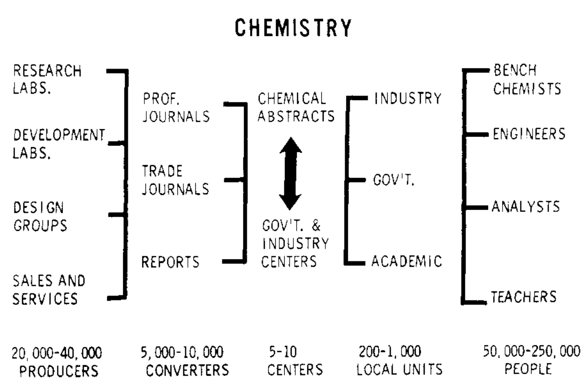

Figure 7.—National network (task-oriented, scientific, and general).

zations who write or speak about their new findings to a general audience. The converters are the journals and the report printer who provide the primary distribution of new chemical knowledge. The wholesalers are the Chemical Abstracts Service and the centers in government or industry to collect and disseminate chemical documents. The local units or retailers, wherever they may be, each serves a specific collection of people who are the users of chemical information. It is anyone's guess at the moment on how large the population of each function is at the present time in the chemistry network. Here again, there appears to be a fruitful area for study.

With this existing network in mind, we can now return to the question: when, and under what conditions, will the chemical community determine that an economic need exists for tying together the chemical information services into a network that utilizes modern machine capabilities? It is obvious that any effort to mechanize or to bring out new information products or services will have to be undertaken with this existing framework clearly in mind.

Perhaps we can approach the answer to the question by stating a few facts of economics:

1. The cost of any conceivable mechanized information system will be greater than existing information services; hence, the justification must come from beneficial use of the new or swifter services rather than from cost reduction within the information network.
2. The cost of developing a truly mechanized national chemical network will lie in the range of $20 million to $50 million, over a period of many years, counting expense for setting up the major centers as well as a central switching and reference service.
3. The inefficiency, if any, of the existing network has not yet been measured in cost terms that are either understandable or persuasive to management.
4. There exists an upper limit, as yet undefined, beyond which the allocation of resources to information handling, as a portion of total technical effort, will not be increased; within this limit, new information services will compete directly with existing services.

From these facts and from consideration of the normal rewards and penalties associated with the gathering of new technical information prior to or during a task, it seems likely that a mechanized national chemical information network is at least five years off. It may be much farther into the future unless the user of chemical information develops a greater dissatisfaction with existing services than he has today. The mechanized network could come much sooner if the government should decide that a large transfusion of funds is justifiable in the national interest.

My own view, with respect to government subsidy of the network, is that it is neither desirable nor likely. The government's support of chemical, petroleum, materials, and pharmaceutical research and development is only a small fraction of the support from industry; it does not seem plausible that the government will use such a small wedge to take over a responsibility that essentially belongs to industry. In addition, the strong statements in Congress about the need for national information services come from committees who have little or nothing to do with the appropriation of funds; the appropriations committees of Congress have shown only a token interest in the funding of new technical information services.

Thus, the impetus, the funding, and the timing of major new developments in the chemical information field will be determined primarily by the chemical processing industry.

Industrial management is well noted for its ability to avoid buying a pig-in-a-poke. The questions that they will raise and require being answered before they underwrite expenditures for a national system will be tough and penetrating: How much are we spending now for chemical information? Who uses it? What for? What do they get in return that helps our profits? Where is the evidence that we missed anything that costs us money? How much more will the new service cost? Who will benefit? Do the line managers agree? How about the chemists? Do they agree? And so on!

If my industrial training has taught me anything, it has demonstrated to me that getting answers to questions like these in the information area is practically impossible, because there are no measuring tools, no accounting records, and no reproducible basis for rational estimates. This merely means, then, that the answer on timing will be dependent upon management's ability to see the need for improvement and to see that greater improvement can be obtained by coordinated national effort rather than in-house effort within the individual company or organization.

It seems quite clear that there is no widespread recognition of this need in industry today, nor does it exist in government either. Thus, the timetable for developing a mechanically augmented version of today's national chemical information network will be set by the ability to come up with reliable and convincing numbers on cost and utility.

The hardware exists and the basic building blocks of technique exist or are just around the corner. The fully mechanized chemical information network will exist when someone finds a way to pay for it. This is the real challenge to which all of us must address our energies and our wisdom about the use of chemical information.