The Qualitative Approach to Scientific Information*

By G. S. SIMPSON, Jr.
Battelle Memorial Institute, Columbus, Ohio
Received September 1, 1961

In January, 1951, we received a government-sponsored project which required, among other things, the eventual preparation of many monographic or state-of-the-art reports. These reports were to be exhaustive studies and evaluations of the scientific and technical literature. The technical areas concerned were many and varied, ranging from the metallurgy of steels and high-temperature alloys through polymer chemistry, and electronics, to areas of physics. In all some thirty different specialties of the physical sciences were involved, and still are.

As you can well imagine, we ran head on into that problem now called the scientific information problem. In 1951, the scientific information problem was not so widely bemoaned as it has been since October, 1957, Sputnik. But to us, it was there in 1951 and very real. We had to find some solution and do it in an economical but effective manner.

Our story is concerned with how we solved that problem, and how we are utilizing the experience gained.

With such a wide variety of physical sciences to be surveyed, it became clear rather early that it would be less expensive to attack the whole rather than approach the job piece-meal. The broad approach, it turned out, was indeed sound, because the now well-recognized interdisciplinary interest of scientists and engineers became immediately obvious. Once the decision was made to consider our information problem as one unit, a technique for over-all implementation had to be established. Based on some experiences accumulated during 1945 to 1950, a system was designed which emphasized the importance of the scientist and engineer user. A number of the professional staff associated with our project examined the "user" system as well as other information systems such as are built around and suggested by the words Dewey Decimal, punch cards, photoreduction, and many others. After much debate—argument really—the "user" system was adopted.

The "user" system considers that these specifications must be met, either totally or with proper modifications depending upon the nature of the information that can be acquired (foreign or domestic, formal or informal, unclassified or classified, non-proprietary or proprietary, copyrighted or unrestricted, *etc.*):

(1) The qualitative control of the entire system must be by the users.

(2) Literature study as a continuing *exclusive* activity is stifling to scientists and engineers, therefore: (a) All routine operations of processing must be accomplished by other than the user. (b) A stimulating rather than thought-restricting retrieval technique must be used. (c) The information must be easily readable, and rapidly available. As many users can testify, among the most frustrating aspects of literature searching and study is to be handed illegible copy, or to find that something you might want is not available until three hours or the day after tomorrow.

(3) Scientists and engineers are inclined to be mostly concerned about that information that relates to the job immediately at hand, and therefore the system must compensate for this very natural circumstance. We consider this a case of protecting the user against himself.

In 1951, the qualitative approach to scientific information was a different approach and, strangely, it is not used very often today. The Battelle information systems that have been designed for users since 1951 have been based on these specifications, appropriately modified for the input available and the objectives of the analysis to be conducted.

As a result of the qualitative approach, and the specifications just mentioned, our information centers:

(1) Employ a "clue word in context" initial search approach. The clue word in context approach is supervised by the users.

(2) Employ three interdependent integrator files which serve as automatic collators. These files are stores of information arranged alphabetically by authors, facilities, and contract numbers. These files are supervised by the Information Center, not the user. These are the gold mines. If you already know the authors or facility in which you are interested you can short-cut the clue-word approach. Many regular users of information know precisely who is working in the specialties of interest to them and where they work.

(3) Use extracts or abstracts rather than titles. Over the years *Chemical Abstracts* has been and is one of our best sources. Multiple reviews of extracts and abstracts by users and Information Specialists lead to least volume, cut out redundancy, padding, and vagaries, and aid in the fundamental problem of literature searching, namely, IDENTIFICATION.

(4) Conduct aggressive continuous collection programs. Information of potential value, while found normally in sources such as standard periodicals, can be procured through numerous other sources. Trips, letters, and conferences, such as this one, develop information which is timely and of value. It is necessary, of course, to maintain a conscious effort to integrate the information received from other than periodicals. By so doing, a considerable amount of information redundancy is eliminated.

*Presented before the Division of Chemical Literature, ACS, National Meeting, St. Louis, Mo., March, 1961