COMMUNICATION IN SCIENCE
Documentation and Automation

ANTHONY DEREUCK AND JULIE KNIGHT, EDITORS

Boston, Little, Brown and Company, 1967, 274 pp. $12.50

A tension exists between the title and subtitle of this book that makes the discussion of the papers far more fascinating than the papers themselves. Because though documentation and automation (i.e., information theory) may be a science, communication, whether in science or in the arts, is a human interaction so complex it escapes the parameters of scientific discourse to run into questions of perception and prestige, promotions and privacy.

This book is an edited transcript of 16 papers and their discussions by the 24 participants in the CIBA Foundation Symposium on Communication in Science: Documentation and Automation, held in London, in November, 1966. The papers fall into three areas: communication systems in science in general, formal scientific communication networks and the technology of automated storage and retrieval.

Garvey and Griffiths' "communication in psychology" is meant to typify communication systems in the sciences. The writers found a remarkable degree of orderliness in a seemingly disordered complex. Through individual adherence to professional information-sharing norms, and through social patterns of conference and publication, psychologists had developed a system that could be diagrammed, studied and changed. New procedures could be consciously introduced into the system to bring information out of restricted channels and into public channels in a much shorter time.

In discussion, however, a medical scientist challenged the emphasis on rapid presentation of research findings. He believed this might lead to "an intolerable degradation of science," stressing that accuracy was more important than speed. Conversely, the case for rapid dissemination of hitherto restricted information was presented by a chemist. He noted that industrial and governmental sources pay for most chemical research, and that they considered prompt access to findings only fair in return for sponsorship. Does the concern of the medical scientist for accuracy reflect awareness that premature information may lead to direct, personal, human disasters, and that the value of human life, at least in the medical context, is too precious to be sacrificed for possible technologic efficiencies? When the chemical piper is almost entirely paid not by his scientific institution or university but by government or industry, is the chemical tune likely to be a bit more discordant than the sweet harmony of pure science?

Menzel's paper on "planning the consequences of unplanned action in scientific communication" offered convincing arguments for the continuing value of interpersonal scientific communication in the age of automation. Interpersonal communication continuously screens, evaluates, synthesizes, extracts implications, transmits the "ineffable" and provides "unlooked-for information" that is often of inestimable value. Of course, the interpersonal cannot substitute for documentation in scientific communication on a global scale, but Menzel offers suggestions for achieving the maximum from its special efficiencies. On the other hand, discussion revealed that many scientists of junior status or at smaller institutions are, in effect, excluded from the "invisible college" of leading scientists who already have the best opportunities for interpersonal communication. Because discovery and progress do not limit themselves by status or location, better communication in science requires not only improvements in documentation but improvements in interpersonal channels as well.

In Price's paper on "the ends of scientific communication," he suggests that "just as 80 percent of communication lies outside the scientific paper, . . . 80 percent of the value and function of the sci-

entific paper lies outside the realm of communication. The paper and communication overlap each other by only 20 percent." The noncommunicative functions include establishing an "intellectual copyright," strengthening credentials in the academic marketplace, and justifying the expenditures of research grants. Thus even the most seemingly objective fact about scientific information—the meteoric rise in the volume of reports and papers—proves loaded with at least three major socioeconomic elements beyond documentation and information processing.

This volume contains a number of reports outlining the structures of operating communication networks: in disciplines—biomedicine, physics, chemistry; in national communities—Sweden, the Soviet Union; in international groups like the European Space Research Organization. In almost every paper on any system certain questions tended to recur. For instance, whether a network was nominally national or international, much of its information gathering and dissemination was international. This always raises questions of language dominance, translation and national secrecy. That, in turn, raises questions about the capabilities of automated systems: will the computer require standardization and unity of language, or will the computer be so flexible and efficient that standardization will be unnecessary?

Until machine translation becomes practical the situation of scientists in minor language countries such as Sweden, Czechoslovakia or Ghana remains difficult. They must receive and disseminate information in a "foreign" language. But even in Canada where language is not a problem *per se,* Canadian scientists must publish in American, British or French journals to reach a significant audience.

Coblan's paper on "the mechanization of documentation—a tentative balance sheet"—indicates quite clearly that although the mechanization of catalogue and index production has been a great success, "each time [an automatic indexing and abstracting system based on the frequency or syntactical relationship of combinations of words] has been taken out of the laboratory and subjected to the real world the results have been uniformly bad." So far, the complexity of semantics has foiled every system of computer translation

and "key word in context" searching. Indeed, Cleverdon's paper reports a typical J-curve distribution between precision and recall. "If one wishes to increase the number of relevant documents which are being retrieved, this can only be done by increasing to a greater extent the number of irrelevant documents also retrieved." After all, an innate ambiguity exists in human communication. Because each man's experience differs from that of every other man, each man's language contains unique ideolectical elements. In a purely taxonomic index, the paper on the Chemical Registry of compounds revealed, perfect recall is possible—but Price compares it with the perfect recall of the telephone directory, long famous for its monumental cast of characters and complete absence of plot.

Knox looks toward the future impact from computerization that none can immediately foresee. The application of the computer only to those procedures now handled by librarians and typists is to miss the point. With full recognition of the value of direct human communication, he envisions "the inspiring potential of the remote-terminal, direct-access, time-shared computer system [to give many users] immediate access—in their own language—to the total store of information in the computer memory." But as he had stated earlier in the conference, "it will require real ingenuity and intelligence to exploit the potentialities of the new technology."

Ingenuity and intelligence are qualities frequently invoked as needed for progress in scientific communication, but almost always as too expensive and in too short supply. At the simplest level, the "intelligence gap" produces complaints of ambiguous or misleading titles to papers and generally poor writing in abstracts, summaries and reports. At the most complex level, however, the critical "intelligence gap" is at a policy level seemingly beyond the scope of this volume. It is the gap between what scientists know about the frequently destructive relationship between social, economic and military power and science, and what they know about coping with it. "Science clearly is supranational," Price states, "and we may or may not have international systems to cope with it." The scientist's natural desire is to disseminate information as widely and as effi-

ciently as possible, but "information is part of the competitive armament of our industrial society," Brée admits.

The papers in this volume inform the reader that automation is making simple many of the formerly complex problems of documentation, and will soon find solutions to formerly insoluble problems. The discussion, however, reveals that social problems like the "publish or perish" reward system, economic problems like the allocation of the largest proportion of "scientific" funds to applied technology, and military problems like the secrecy forced upon military-funded research projects have as great—or perhaps greater —influence on communication in science than the increasing capacity of the computer to refine symbols, store them, and retrieve them on command.

STEWART A. SELBY