Communicating Via Computers*

WALTER M. CARLSON
IBM Corporation, Armonk, N. Y. 10504

RUTH M. DAVIS
Lister Hill Center for Biomedical Communications, Bethesda, Md. 20034

Received May 27, 1970

Prospects are substantial for use of computing machines to improve the efficiency of existing methods for technical documentation. Pioneering efforts to date have demonstrated the ability of computers to assist in such functions as retrieval of literature or patent references and abstracts, storage and retrieval of data on chemical structures, and analysis of data relating to synthesis, manufacture, testing, distribution, or use of chemicals. However, computers have not yet provided much assistance to methods for communicating chemical knowledge or technical know-how. Experiments now in progress are teaching that much more must be known about human communications processes before machines can be programmed to play a significant role in such processes. The most important early conclusion from current work is that combinations of men and computers operating interactively will be more efficient than computer methods now in use.

It is especially fitting and appropriate that a paper should be given on this subject at a symposium in Toronto, the city in which Marshall McLuhan first received international attention.[1] Our experiences in the use of computer-based methods for improving chemical and biochemical documentation have led to some conclusions that may sound as though they came straight from McLuhan. While he was not the source of our conclusions, it neither surprises nor distresses us that striking similarities occur between his exposition of the instantaneous, simultaneous world of electronic information media and our examination of communicating via computers.

For the foundation of the structure in this paper, we have chosen a perceptive and pertinent paragraph that appears in "Information Sciences 1968," published by the Air Force Office of Scientific Research to summarize the work of its Directorate of Information Science in 1968.[2] The paragraph reads:

> "The problem really begins when someone, we will assume altruistically motivated, creates something called data. Much of this data is in tabular, graphic, or symbolic form. At this instant a highly predictable, and absolutely irreversible, process takes place. A convention of long standing dictates that the results of this labor be written up for presentation at meetings and publication in numerous journals. As a result, additional data, often referred to as verbiage, is generated to connect the original data together. Behold—a document has been created."

The contention here is that this process, because of its irreversibility so far, has relegated the role of computers to the improvement of methods for handling the resultant documents. And, unfortunately, there is very little evidence yet available to show in what ways, if at all, improved document handling results in improved communication. Other authors presumably will discuss how document handling affects chemical communication.

For our purposes, the word "communication" is used in its original sense: the transfer of knowledge from one person to another. Until at least one human being has received, assimilated, and understood a message containing meaningful data or symbols, there has been no communication. The writing of this paper, in itself, did not produce communication except as each author enjoyed and marveled at the discovery of the similarities and differences of our two views on this subject. The presentation of this paper at a symposium, in itself, does not constitute communication unless and until at least one listener obtains a new insight or a new piece of data from hearing the paper. Similarly, the publication, abstracting, and distribution of this paper does not constitute communication until at least one reader assimilates and understands what we are saying.

These introductory remarks can be concluded by a summary of what follows. Computers are being used and will be used extensively in chemical documentation systems. They have a demonstrated value in assisting complex chemical calculations, including the manipulation of data on organic structures. They have also been especially useful in the collection and analysis of data generated from an enormous variety of activities ranging from laboratory research or properties measurement to plant operation or product distribution. However, computers have not yet provided any significant improvement in the communication of chemical knowledge or technical know-how.

* Presented at the ACS-CIC Joint Meeting in Toronto, Canada, May 27, 1970.

Moreover, it is not clear that computers will participate in chemical communication processes until much more is learned about interactive combinations of men and information processing machines.

These conclusions may be illuminated with some examples of work under way.

LABORATORY AUTOMATION

The state of the art in laboratory automation provides a good starting point.

Computers are becoming increasingly useful tools for the collection and analysis of experimental data and test results.[3] The assembly of such data in a machine-readable form provides a golden opportunity to intercede before that "highly predictable and absolutely irreversible process" toward documentation begins to take place. The use of the computer to communicate the data and the knowledge gained from his experiments and tests between a chemist or engineer and his peers, without resorting to documentation, now looms as a distinct possibility.

First, however, laboratory automation will have to progress through several successive stages of sophistication. The technology now exists for connecting sensors and transducers to both analog and digital computers for purposes of data capture and data reduction. The key questions today are related to economic justification.

A next stage that is obvious and that has come into limited operation today involves storage of the data in either raw or reduced form for later analysis or comparison with results of other experiments or tests. Here again, technical feasibility is not at issue; the key questions are programming costs, file maintenance costs, and storage costs and their economic justification.

The next stage of sophistication is also obvious, but it has not yet achieved routine, operational status. Here the experimental and test data are stored and then formatted for comparison with completely different kinds of data. A typical application is the use of computer-stored test data to identify compounds through comparison with computer-stored physical property data, such as infrared or NMR spectra of known compounds or elements. At this stage, feasibility has been demonstrated, and tentative cost estimates can be made for installing such applications. The economic justification begins to take on a new flavor, however. At the previous two levels, the primary benefit is in reduced clerical labor requirements. At this level, there is still a clerical labor factor, but the new element is an improvement in the ability of the professional chemist or engineer to test a far wider range of hypotheses and to do so automatically. We now enter the realm of the productivity of the professional individual as a source of justification. We fully expect to see widespread use of computers in laboratories for this analytical, multifile function within a few years.

The final stage of the sequence is only obvious when one accepts an interaction between the chemist or engineer and the computer-stored data bases that have just been implied. One can then visualize that many individuals will begin to interact with the multiple files of data. When this begins to occur, there will be communication via computers. The data collected, assimilated, and understood by one professional person will be available electronically to another person not only in raw form but also in any of the several stages of insight employed by the person for whom the data were orginally collected.

Files derived from laboratory automation projects exist today in formats that could be exploited by several people from their own terminals for their several purposes. We do not know of anyone who has yet written the programs to close this kind of loop. Our suspicion is that the first programs, when written, will perform rather trivial and highly stylized inquiries. The real payoff will now be found not only in the greater productivity of individual professionals but also in the better over-all performance of the group that has access to the data bases and is communicating via computers. It is painfully clear, however, that much more must be known about the human communication process to define how the laboratory automation systems at this final stage are to be developed and installed.

AUTOMATION OF RECORDS FOR MANAGEMENT

One of the most expensive record keeping operations today is that for medical care. Keeping the necessary records for MEDICARE, MEDICAID, and the many health insurance plans is time-consuming, costly, and is normally still manual. Handwritten records in physicians' offices or hospitals become typed records in a prescribed form. These are mailed to central locations, generally retyped, and the same data are then pulled off many times onto different forms for different purposes. There is good reason to suspect that all the generated forms are kept in filling cabinets for extended time periods. Errors are introduced at every rewriting and retyping. Studies have even been made which indicate the frequency by type of errors made through retyping and recopying.

What is perhaps most discouraging is that during this rather long drawn-out process of record making, identification of a given data element on a series of individuals is difficult, if not impossible, to retrieve. And these records are an excellent source of statistics on drug prescription habit, drug dosage, hospital-stay patterns by hospital and community, medical costs in various locations throughout the country, and the like.

The potential transfer of knowledge, the potential improvement in communication is hampered by the means employed for document or record handling.

Fortunately, automated handling of medical records is being furthered by a few individuals and organizations concerned with medical care, both in terms of quality and cost. Some of the principal benefits of these innovations will be the reduction of hard copy documents, the utilization of computers as means of communication of information between individuals, and hopefully a resultant improvement in the transfer of knowledge from one person to another.

The key to such proper use of communications via computers is the use of input and output terminals connected to remotely located computers. The terminals, normally keyboard or cathode ray tube terminals, substitute for the traditional paper and pencil mechanisms for record making. Since people find it difficult to make large transitions rapidly in behavior patterns, most new automated medical record systems make hard-copy duplicates of the transmitted records available to the physician or paramedical personnel.

As an example of this automation of medical records, Blue Cross and Blue Shield of Virginia announced in March 1970 a system in which medical data terminals in doctors' offices connect with an IBM 360/50 in Richmond by telephone lines.[4] Installation of the system started in May 1970 permits a doctor to enter data via the keyboard terminal in his office and receive a voice response. Blue Cross estimates a doubling of its capacity to handle claims and a saving of $1 million a year in Virginia's medical care costs. If the service spreads nationwide, an estimated savings of about $60 million a year can be anticipated.

The improvement in communications and transfer of information is more subtle but still highly visible. Subject, of course, to all those privacy caveats now imposed on data maintained manually, but free of the existing constraints of the inaccessibility of manually-kept records, the computer can now manipulate in any desired combination the data in the records and communicate the results to researchers and clinical investigators at their own locale. The savings in gross time, in researchers' time, the improvement in accuracy, and the gain to health care practices is exciting, even if impossible now to predict. The remaining restraint is the ability of people to adapt to this new form and source of information and communications.

AUTOMATION OF LITERATURE BIBLIOGRAPHIES, INDICES, AND ABSTRACTS

The transfer of knowledge and the communication of information between individuals is hampered more than most people realize by inaccessibility of information about the content of books, documents, journals and audiovisuals. There are over 60 published sources of abstracts, indices, and bibliographies of biomedical holdings.[5] These include Epilepsy Abstracts, Excepta Medica, BIOSYS, Chemical-Biological Activities, RINGDOC, Medical Digest, Mechanized Clinical Data, European Nuclear Documentation System (ENDS), British Medical Abstracts, etc.

In most cases, the customer who is attempting to find those documents most pertinent to his inquiry is a passive recipient of these guides or directories to literature holdings. He cannot combine the sources or manipulate index terms or keywords to produce a different set of relevant documents. In those cases where the citations are contained in computer data banks such as Science Citations Index, MEDLARS, and Excerpta Medica, the computer data banks are not directly accessible to the customer. He must use handwritten or typed requests, communicate with human intermediaries, receive hardcopy outputs, and then register his satisfaction or dissatisfaction via voice or new round of written communications. The results are the relegation of the computer to a secondary document handling role and the total lack of use of available communications directly from the computer to the customer.

Some attempts have been made recently to institute direct communications from the computer to the customer searching for relevant literature. In the early 1960's the Air Force established a computer data bank in California of over 200,000 citations to foreign scientific and technical literature.[6] Authorized customers with computer terminals were allowed to directly interrogate the data bank from wherever their location within the United States.

Dr. Kessler at MIT established a similar data bank of computerized citations to the 34 most used physics journals.[7] He established this data bank on the Project MAC computer system, and its customers were allowed access to the physics citations. The project was known as the Technical Information Project (TIP).

One group of scientists and technologists under the guidance of Dr. Davis obtained access in 1967 to both data banks from Bethesda, Maryland. Here communications permitted transfer of information from a computer in California and another in Massachusetts to the same room in Bethesda equipped with very simple, inexpensive computer keyboard terminals. This group found it possible to switch from one data bank to the other, narrow the search, or widen it through keyword connections to include different sets of journals.

The results were exciting enough to validate the hopes of many that computers and scientists could intercommunicate information and expedite the transfer of knowledge in the library world. Although the increase of such readily accessible computer communication has proceeded at a snail's pace, there is no denying its eventual important, perhaps dominant, role in information communication.

PROGNOSIS

While it is clear that computers have not yet had a significant part to play in the human communication of technical knowledge, we are pleased to be part of the movement toward using computers to expedite the transfer of knowledge. We hope that our enthusiasm for the opportunities that lie ahead will affect the decisions of those who read our words.

Much remains to be done. There is a need for empathy with humans and the human decision processes. Whatever the direction of future work, we predict that it will be exciting for those who undertake to find a method for communicating via computers.

LITERATURE CITED

(1) Compton, N., "Cool Revolution," *Commentary* 39:79–81, January 1965.

(2) Air Force Office of Scientific Research, "Information Sciences 1968."

(3) *IBM Journal of Research and Development*, January 1969 (entire issue devoted to laboratory automation).

(4) "An Information System Expects to Save $1 Million Per Year," *Biomedical Electronics*, March 25, 1970.

(5) "Abstracting-Indexing Service: An Updated Report of Coverage," *NFSAIS News* 12:1, Feb.-April 1970.

(6) Barrett, R. P., "CIRC-Centralized Information Reference and Control," Proceedings of the Workshop on Working with Semiautomatic Documentation Systems, Warrenton, Va., May 2–5, 1965, J. J. Maher, Ed., System Development Corporation, Santa Monica, Calif., AFOSR 65–1699 (AD 620 360), pp. 15–31, 1965.

(7) Kessler, M. M., "The MIT Technical Information Project," *Physics Today* 18:28-36, March 1965.