large, the institutions have not kept up with the technology, we have undesirable (unplanned) "fall-out" from the technology. The name which we have given to this fall-out is pollution and the problem of how to fight pollution is one of the most serious which we face. Now the explosion of information, as indicated by the rapid growth of primary journals, is an off-shoot of the technological explosion. The development of information systems is, in this analogy, the development of social institutions to cope with this technological problem. If we allow these institutions to be "driven" by technology in their form of development, rather than plan them in a broader social context, we will increase rather than decrease pollution. In this case, we will degrade the quality of our scientific environment and damage science itself.

LITERATURE CITED

(1) de Solla Price, Derek J., "Science Since Babylon," *passim*, Yale University Press, New Haven, Conn., 1961.
(2) de Solla Price, Derek J., "Little Science, Big Science," *passim*, Columbia University Press, New York, 1963.
(3) Gottschalk, Charles M., and W. F. Desmond, *Am. Doc.* **14**, 188–194 (1963).
(4) Pasternack, Simon, *Phys. Today* **19**(5), 38–43 (1966).
(5) Ziman, J. M., "Public Knowledge, An Essay Concerning the Social Dimensions of Science," *passim*, Cambridge University Press, London, 1968.
(6) Skolnik, Herman, J. CHEM. Doc. 8, 190 (1968).
(7) Merton, Robert K., *Am. Scholar* **38**, 197–255 (1969).
(8) Slater, M., and S. Keenan, American Institute of Physics, Information Division, Rept. AIP/CPP-1 (available as PB 178 368), Sept. 1967.
(9) Libbey, M., and G. Zaltman, American Institute of Physics, Information Division, Rept. AIP/SDD-1, Aug. 1967.
(10) Stewart, Ileen E., *Federation Proc.* **22**, 1002–7 (1963).
(11) Moravcsik, M., and S. Pasternack, *Phys. Today* **19**, 60 (June 1966).
(12) Kronick, David A., "A History of Scientific and Technical Periodicals, The Origins and Development of the Scientific and Technological Press, 1665–1790," *passim*, Scarecrow Press, New York, 1962.
(13) Garvey, William D., and B. C. Griffith, *Science* **157**, 1011-16 (1967).
(14) Garvey, William D., and B. C. Griffith, *Am. Psych.* **21**, 1019 (1966).
(15) Moravcsik, M., *Phys. Today* **21**, 48 (Oct. 1968).
(16) Herring, Conyers, *Phys. Today* **21**, 27 (Sept. 1968).
(17) Branscomb, Lewis M., Editorial, *Rev. Mod. Phys.* **41**, 1–2 (Jan. 19, 1969).
(18) Branscomb, Lewis M., *Sci. Res.* **3** (11), 49 (1968).
(19) Debate on Preprints, *Science* **154**, 332–35 (1967); *Phys. Today* **19**(8) 12–16 (1966).
(20) Bernal, J. D., *J. Doc.* **13**(4), 195–208 (1957).
(21) Luhn, H. P., *Am. Doc.* **12**(2), 131–38 (1961).
(22) Hoshovsky, A. G., *et al.*, USAF Office of Aerospace Research, Rept. OAR 67-012 (AD 668 072), Sept. 1967.
(23) Phelps, Ralph H., and J. P. Herlin, *UNESCO Bull. Libr.* **14**, 61–75 (Mar.–Apr. 1960).
(24) Gannett, E. K., "Technical Journals and the Information Explosion," Intern. Tech. Comm. Conf. Soc. Tech. Writers and Publishers, March 1967.
(25) Sjoberg, G., *Phil. Sci.* **23**(2), 90 (1956).
(26) Brown, W. S., J. R. Pierce, and J. F. Traub, *Science* **158**, 1153–59 (1967).
(27) Swanson, D. R., *Am. Psych.* **21**, 1005 (1966).
(28) Weinberg Report, President's Science Advisory Committee, Science, Government, and Information, U. S. Govt. Printing Office, Washington, D. C., 1963.
(29) Bryand, James E., *Newsletter Am. Soc. Inform. Sci.* **7**(4), 2 (1968).
(30) Simpson, G. S., Jr., *Battelle Tech. Rev.* **17**, 28 (May–June 1968).

Work supported by the National Science Foundation under Grant No. GN-710.

New Developments in Primary Journal Publication*

JOSEPH H. KUNEY
American Chemical Society Publications, 1155 Sixteenth Street, N. W., Washington, D.C. 20036

Received June 17, 1969

The form and function of the journal as a medium of communication between scientists, that is, one scientist talking to another, have not changed greatly these past 300 years. But during that time, the journal has evolved into a highly effective medium for meeting the scientists' needs for publication, distribution, storage, and retrieval of scientific information. Kessler describes the journal as "the most successful and ubiquitous carrier of scientific information in the entire history of science." And the scientific paper, he adds, "is such a marvelous and accommodating invention that we seem to take it for granted and forget that it has form and structure that fit its function."[1]

The notion seems to be growing that the journal is approaching obsolescence as a medium for the publication and distribution of scientific information. In his recent ACS Presidential Address, Dr. Cairns used the term, "unintentional obsolescence,"[2] to describe not only the fate of journals, but the shortcomings of our present information system in general. The report of the National Academy of Sciences Committee on Scientific and Technical Information[3] concludes that the primary journal faces trouble because of delays in publication, impact of less

*Presented before the Division of Chemical Literature, 157th Meeting, ACS, Minneapolis, Minn., April 14, 1969.

In the immediate future, efforts will be accelerated to provide users with more selective groupings of articles. These attempts will vary from the availability of single copies of articles on an "on-demand" basis to the publication of customized packages of articles produced through the use of computerized profile-matching techniques. In these efforts and in the current developments in computerized typesetting, the groundwork is being laid for a computer-based system of disseminating and using primary information. Papers will be put into machine-readable form and all subsequent editorial processing will be within the computer. In turn, users will have access to the content of the system via remote consoles, which will permit on-line access for fulfillment of the variety of user needs that the journal system now provides.

formal modes of communication—particularly unrefereed, "semi-informal" publication—and the reader's need to skim much that is irrelevant and useless in order to read material of interest. The latter has been fostered by the economics of journal production, which makes it possible to put articles in users' hands at relatively low cost-per-article-published, although at a rapidly increasing cost-per-article-read.

Severe demands are being placed on all journals, as they attempt to cope with the rapid and pervasive growth of specialization in new subject matter, together with an increasingly large and literate total user population. These demands are reflected in financial problems, which are being met by increasing subscription prices and page charges to maintain the present system of publishing journals. The financial problem is less urgent for some journals than for others, and it is clear that for some time we can continue to operate the present journal system and to support its growth by passing on the added costs to authors and subscribers. But it is equally clear that the whirl of problems created by publishing a growing volume of literature for an audience with a widening list of specialized scientific interests will have to be solved by routes other than simply increasing prices to users.

Publishers are constrained in their efforts to test and implement innovations by the knowledge that they have a responsibility to provide for the whole spectrum of functions that the journal fulfills. If readers are to receive a more efficient selection of articles to meet their demands for current awareness, as suggested by Brown, Pierce, and Traub,[4] the publisher must also see that such functions as creative stimulation through browsing, archival storage, and retrospective search capability are not lost in the process. The publishers of journals are also well aware that journals are as important in meeting the scientist's need for publication outlets as they are in meeting the scientist's need for information.

We have the know-how and the equipment to implement sophisticated computer-based systems which can provide users with a current awareness service based on more selective packaging of primary papers than is now achieved by most scientific journals. In addition, functions such as storage, retrieval, and retrospective searching can be made available through computer technology on a basis separate from the primary publication function. But before we implement such programs, we must test user acceptance and determine whether or not the user is willing to pay the costs, both in dollars and in increased personal participation in the information transfer process.

To meet the challenge, publishers of journals are taking two broad approaches. One might be thought of as a print-based approach, wherein the various information functions of the journal will be isolated and offered to users with the aim of providing a more useful selection of information services. In effect, such programs will incorporate into the primary system some of the services now available at the secondary level, in an effort to afford users greater selectivity in access to the content of the primary literature. These attempts vary from mechanisms to make single copies of articles or manuscripts available on an "on-demand" basis to the publication of customized packages of articles based on matching of group profile requirements.

The second approach recognizes the information-seeking patterns that have evolved between user and journal, and seeks to incorporate those patterns into a machine-based information system concept. Experience to date with computer technology shows that we can develop techniques which will provide the user with all he might ask of an information system—whether that request be for current awareness, for specific data, or for a retrospective search. Such a computer-based system can supply information to users on the basis of an interest profile, or via some form of on-line interaction with a data bank. The computer can be used to build personal files or for access to the world's store of chemical and scientific information. In short, it offers all the capability that is possible from the present primary journal system, but which must be mechanized if we are to cope with the growing volume of literature.

Computer-based systems offer a dramatic and basic change from print-based systems. At the precise moment of encoding a manuscript in a form that a computer can read, manipulate, and transmit, the information content of the manuscript becomes available for use throughout the total system. Any particular bit of encoding information is there in the system to be exploited for whatever use required, at any time following the moment of primary input.

Thus, there is a need for concern with processing primary information for computer input at least equal to that accorded the development of systems for accessing the information. The objective becomes, in terms of a total information system, to input primary information in a manner that will permit its full and immediate use in any and all subsequent phases of information transfer process. By operating to the fullest extent possible at this point in the system, we maximize the availability of the information to the user, we reduce to a minimum the possible duplication in handling and intellectual processing, and we begin at an early point in the system to maximize its efficiency and utility.

PRINT-BASED APPROACHES

Does the user of scientific information require or even want the whole spectrum of information functions that the journal provides? In an effort to answer this question, publishers are isolating the various information functions and offering users the opportunity to select only those parts they want. For example, many publishers are considering expansion of current programs that make available on demand single copies of journal articles or manuscripts.

A new type of service in this area is the Mathematical Offprint Service (MOS) offered by the American Mathematical Society. This service offers to the individual on a continuing basis reprints and/or title listings of articles which satisfy the criteria specified in his interest profile. In addition to an interest profile, the subscriber indicates authors whose work he specifically wishes to receive or exclude, fields of primary and secondary interest, languages, and journals from which he does not wish to receive articles or titles.

Approximately 70 journals from among those reviewed in *Mathematical Reviews* send to MOS galley proofs from which the computer and keyword listings are prepared, based on a subject classification that is essentially the same as used in *Math Reviews*. About 450 articles are processed per month, and the computer matching process is done once a week.

At the end of 1968, the system had about 1200 subscribers who pay $30 for 100 reprints. Ten title listings equal one reprint. With each listing, the subscriber receives an accounting on the unused portion of his subscription, and when the figure reaches a critical level he is billed for another $30. The average reprint is distributed to 10 subscribers with a range from a low of 1 or 2 up to 50. The current program is subsidized by an NSF grant, with actual cost estimated at about $75 per 100 reprints or equivalent.

Another area attracting the interest of journal publishers is the concept of a selective journal system, whereby the entire content of a limited journal would be intended to meet the needs of a specialized audience. The goal of such a system is to repackage journal articles for limited, highly select audiences, and thus provide users with a high percentage of relevant articles per issue of the journal. Under such a system, articles of wide interest would be published in as many of the selective units as necessary to meet all needs. In addition, such a journal system might expand its current awareness function by including, as part of the content of each issue, abstracts or other announcement-type material covering articles in peripheral areas of subject interest.

This is a concept somewhat different from a selective dissemination approach based on the computer matching of individual subject interest profiles with the titles and/or keywords of journal articles. Profile matching is costly and it is unlikely that publishers of journals could provide such individualized service to large numbers of subscribers. However, it does seem feasible for publishers to provide a selective distribution of journal articles, based on identifying groups of scientists that are relatively stable and can be identified with subspecialties of scientific subjects or disciplines.[5] Providing this type of service is seen as an evolutionary development of the journal function, with a somewhat sharpened and more intensive effort toward identification of the groups and their literature needs. Various techniques have been proposed for doing this, ranging from computer matching of profiles to personal selection by users. Working out a classification scheme based on group profiling will be no easy task.

But the most prevalent current interest seems to lie in the role of the short paper as a variation of or supplement to traditional primary publication. We do not have, as yet, any agreement as to what a short paper is. To some it is a long abstract, to others a paper in which the experimental discussion has been omitted, to still others a communication or letter. Certainly, some concept of space limitation is involved, and this limitation promises real savings in terms of publication cost and user time.

The evidence is sparse, but seems to show that "communications" and "letters" as a mode of primary publication are finding increasing favor with authors and acceptance by users. In theory, the communication or letter is supposed to serve as a preliminary announcement of work to be reported in full as a journal article at some later time. But the experience of *Physics Reviews Letters* indicates that only about 50% of the items in *Letters* ever appear as full journal articles.

The ACS publications have begun exploration of a system whereby authors might be required to submit both a short form and a paper containing all the details of the work. The short form would be published as rapidly as possible in journals of wide circulation. The full papers would be made available in microfilm editions or in some form of printed master journal for use in libraries or information centers. We plan to explore the feasibility of such an approach, including a study of authors' willingness to prepare two versions of their papers, the economics of such a system, the abstracting and indexing problems, and related problems.

MICROFILM PUBLICATION

In what is seen as a transitional step between print-based and computer-based systems, publishers have begun to make journals available in microform, mainly as microfilm. In terms of information system design, microforms may be thought of as an interim step toward the ultimate form of microstorage—the computer—or as an answer to full-text storage problems. But the ease of use spurred by development of the 16-mm. cassette and the reader-printer has created a demand which is resulting in increased availability of journals in microform.

The marketing and pricing of microform publications are also providing clues to the requirements of marketing and pricing computer-based information systems. For example, the ACS primary publications and *Chemical Abstracts* on microfilm are made available under a license arrangement. The program is considered an experimental one and is designed to provide a constant flow of in-use reactions. It is also designed to provide users a means of exploring the fullest possible use of microforms without concern for possible copyright restrictions. Out of such programs, publishers can expect to gain considerable insight into the problems of marketing future information systems.

COMPUTER-BASED SYSTEMS

The development of computer-based systems of handling primary information was started as an evaluation of computer-oriented composition systems able to handle all the complex requirements of scientific manuscripts. Out of these efforts have emerged results and design data which enable us to move toward development of a highly flexible typesetting system designed and programmed for the present generation of computers. For example, the procedures developed by the ACS have been previously described.[6,7] They include computer-based techniques for keyboarding of textual information, chemical structure diagrams, and mathematical expression, as well as the hyphenation and justification routines required in the composition of chemical text. These procedures are being tested and refined in a pilot production system that now handles the regularly scheduled production of four journals, *Journal of Chemical Documentation*, *Journal of Chemical and Engineering Data*, *I&EC Product Research and Development*, and *I&EC Process Design and Development*.

An essential aim of the ACS pilot production system is to guide the continuing development of a unified system for reliably recording published information in a machine-readable form as part of the publication process. Here, editorial and keyboarding operations are being combined into a single flow, producing a machine-readable record with little variation from one publication to another. Yet the over-all system permits each of the publications to maintain its own character. Only in the final composition phase are the alternatives in form and format of the individual printed publication produced under programmed controls with well-defined parameters.

Despite five years of effort in keyboarding manuscripts for input to a computer-based typesetting system, the problems associated with the keyboarding process remain a major bottleneck in evolving an efficient and economic process. For straight text requirements, the input problem has not prevented the ever-increasing implementation of computer typesetting systems. It has been estimated that near the end of 1968 there were almost 700 such installations in the United States[8] and Canada. The majority of these were for use in newspaper production.

But the problem of using approximately 45 keys to produce codes from which more than 1000 characters can be produced has not been solved to the full satisfaction of users. To the large variety of characters must be added the complex spatial problems to be solved in setting mathematical expressions and chemical structures. The latter can be handled adequately, using as input one of the chemical typewriter devices,[7] but thus far we have not made equivalent progress in the handling of mathematics. Our present computer methods are about as ponderous as the methods used in hot-metal typesetting. Our capability in the area of math setting must be improved if we are to realize the full economic advantages promised by computer-aided typesetting. With the increase in the use of computers in typesetting, more attention is being given to keyboard problems, both in terms of more versatile equipment and more sophisticated software. We anticipate that out of a combination of these efforts will come the solutions we need.

As the volume of the material being made available in machine-readable form increases, the structure of the files of primary information becomes more important. Straight sequential filing of articles on magnetic tape in order of publication is inadequate for a computer-based system. The basic file structures now being developed for working, master, and archival files are information-oriented, as differentiated from journal-oriented; they must be assessed in light of the emphasis to be placed on the treatment of primary information for inclusion into a chemical information data bank within the framework of a total information system.

In the months ahead we will begin work on the definition of design specifications for the next phase of development for preparing primary chemical information as input into an ACS data bank, which, in turn, will be a subset of a total chemical information data bank. In this effort, we will focus on the relationship of the primary system to the total information system, and the techniques and procedures required to input material into the system to take fullest advantage of integrating primary and secondary information systems.

In our progression toward integration of the primary and secondary information systems, plans include the writing of programs which will enable Chemical Abstracts Service and other secondary processors to receive, via magnetic tape, selected information encoded by primary journal publication systems. Illustrative of the type of problem to be encountered in moving toward an integration of systems was the need to analyze the differences in style and usage between the two systems and to move to eliminate these differences, either by editorial agreement or through the development of computer capability to handle the necessary translations. This will be a continuing effort throughout the development and implementation of primary-secondary interfaces—wherever an interchange of machine-readable data takes place.

Redesign and programming efforts, both in progress and planned, are now definitely more information-system-oriented, as a result of our experience to date. Our computer-based design efforts seek to accommodate the processing of the primary literature as an information subsystem that will integrate with a total chemical information system. It follows, then, that we must develop means of effectively treating and manipulating information after it has been input to the system. We propose to develop and to implement more efficient, immediate, and facile modes for accessing and treating primary information in its flow through the information system. Advantages to be gained include the following:

(1) Immediate and constant access to our data files.

(2) Observation of both the corrections to be edited into the data and the effect resulting from the same.

(3) Use of the display devices to input other kinds of information: e.g., information flagging and job controls.

(4) Experience gained from accessing information through a remote terminal.

Future work will include exploration of additional areas which must be developed if we are to evolve an efficient, workable, over-all computer-based system for recording and processing primary information.

1. Editing and correction techniques, which can be applied at any stage of the manuscript-processing process, as the

foundation for the basic over-all system for production, storage, and access to information.

2. On-line capability, that is, to tie input devices directly into the computer, including the study of various modes of entering information into the system, whether they be manual keyboarding or optical scanning.

3. And finally, *man-machine capability*, whereby *all* information input to the system will become a part of the information system and can be revised, updated, or rearranged at any time during the computer handling cycle.

What has been described here is development actually being designed or tested by those concerned with the handling of primary information. These steps are illustrative of the type of detail which must be worked out and tested before implementation into the over-all system. It is much easier to write about than to do. The traditions associated with the journal system are firmly set, and it will take strong evidence of the value of any innovations before we can win approval for their testing, let alone adoption.

ACKNOWLEDGMENT

Much of the ACS work on computer-based systems is supported by the National Science Foundation.

LITERATURE CITED

(1) Kessler, M. M., "Some Very General Design Considerations," in "TIP System Report," Appendix H, Massachusetts Institute of Technology, Cambridge, Mass., Oct. 1967.
(2) Cairns, Robert W., "ACS Responsibilities for Communication," *Chem. Eng. News*, pp. 48–52 (Nov. 11, 1968).
(3) National Academy of Sciences–National Academy of Engineering, "Scientific and Technical Communication. A Pressing National Problem and Recommendations for its Solution." A report by the Committee on Scientific and Technical Communication (SATCOM), Washington, D. C., 1969.
(4) Brown, W. S., J. R. Pierce, and J. F. Traub, "The Future of Scientific Journals," *Science* **158**, 1153–1159 (1967).
(5) Swanson, Don R., "Scientific Journals and Information Services of the Future," *Am. Psychologist* **21**, 1005–1010 (1966).
(6) Kuney, J. H., B. G. Lazorchak, and S. W. Walcavich, "Computer-Aided Typesetting for the Journal of Chemical Documentation," J. CHEM. Doc. **6**, 1–2 (1966).
(7) *Ibid.*, "Computerized Typesetting of Complex Scientific Material," Proceedings, Fall Joint Computer Conference, Vol. **29**, Sparten Books, Washington, D. C., 1966.
(8) *CIS Newsletter*, Composition Information Services, Los Angeles, Calif., Oct. 15, 1968.

Subject Indexing as a By-Product of Electronic Composition*

E. R. LANNON
Consumer Protection and Environmental Health Service, Public Health Service, U. S. Department of Health, Education, and Welfare, Washington, D. C. 20204

Received June 17, 1969

Electronic composition in the federal government has progressed rapidly since it was instituted in October 1967. During the first year of operation, it saved $900,000 in composition, printing, and binding costs, and reduced the bulk by 900 tons. Using electronic composition allows production of an index with reduced keystroking, a decreased time cycle of publication, a more meaningful index, and an established data base for additional processing. An algorithmic approach to indexing is now being developed.

Within the federal government, we talk of electronic composition as the output of composing systems using electronic or electro-mechanical components that produce typographic quality composition at speeds greater than 600 characters per second for eight-point characters. By "typographic quality" we refer to a resolution of not less than 800 lines per inch and proportionally spaced characters. We established the speed factor of 600 characters per second by determining the theoretic composing speed we could obtain from a state-of-the-art computer line printer through to a film output suitable for preparation of a printing plate. I am not aware of any operational line printer system that can achieve a theoretic speed of 600 characters per second through to film. However, it was possible to produce a camera system that would process hard copy output from a computer line printer to film at such speed. In actual operation, a line printer operating at 1100 lines per minute has an effective character rate in the range of 200 characters per second through to film. The drop-off is a function of the normal photographic process.

Indeed, it was this fact of life that was instrumental in bringing into being the first high-speed composing system as a part of the MEDLARS system of the National Library of Medicine in 1964.

Now, are we talking about the world we live in, or that better world computer salesmen are prone to say lies in the immediate future?

We are talking about a world that became real in October, 1967, with the installation of the electronic composing system of the Government Printing Office. This system has as its base the Linotron 1010. In terms of hardware and software, I believe this system is the most

*Presented before the Division of Chemical Literature, 157th Meeting, ACS, Minneapolis, Minn., April 14, 1969.