IS JOURNAL PUBLICATION OBSOLESCENT?

by Simon Pasternack

STORAGE AND RETRIEVAL of information is currently a subject of considerable urgency and one that affects all branches of science. I know that PHYSICS TODAY readers are concerned about the problem, and I suspect that they are, as I am, rather uneasy about many trends in journal publication. I want to discuss some of them in detail, but first I would like very briefly to present some background information.

The United States government sponsors, directly or indirectly, an enormous amount of scientific research. It regards the scientific establishment as one of the country's great assets, and rightly so. In recent years the government has become increasingly concerned about the efficiency and effectiveness of the scientific establishment, and has been studying its operations with a view to strengthening it in various ways.

A growing critical problem

In particular, a great deal of attention has been focused on transfer of scientific information—a problem that has been growing steadily more critical because of the very rapid growth of scientific research, scientific conferences and scientific publication. The number of significant documents generated per year in the world's technical literature has been estimated to be over 600 000 in 1961, 900 000 in 1965 and nearly 1 200 000 by 1970. In biology, chemistry and engineering (each about ¼ of the total) it will soon be possible really to keep up with all developments by studying a thousand articles per day; to maintain reasonable current awareness by reading a thousand abstracts per day; or to maintain superficial current awareness by reading a thousand titles per day. Physics is much easier; there are only one fifth as many documents. It is no wonder that people are beginning to worry about science drowning in its own output and are beginning to take information storage and retrieval seriously.

Many aspects of this problem need careful analysis: the production of documents, their evaluation (not every document is significant in the sense that it is worth storing and retrieving), their improvement if necessary, their collection, their classification (indexing), their storage, their distribution, their use. Even the order of these aspects is important, as we shall see.

Now there are two major recent technological developments that affect the problem. One is the enormous improvement in duplicating copies of manuscripts and papers; it is now easy to duplicate freely and plentifully (and distribute widely) copies of one document at various stages of its development (progress reports, fragmentary preliminary reports, first draft of paper, second draft, . . ., submitted paper, revised versions, printed paper, errata, etc.). If each version is regarded as a significant document in itself, one either stores many versions or has to find some way to phase earlier stages (and references to them) out of the permanent storage system. Removing earlier stages would be easy except that other papers in the system contain references to these earlier stages. These references greatly complicate the retrieval problem.

The other major technological development, one that offers some hope of making information storage and retrieval controllable, is the tremendous growth of computer technology. It has become feasible to put much bibliographical and other information into computers, leading to such things as citation indexes, with which some readers may be familiar, to data compilation centers like the Sigma Center for neutron-cross-section compilation at Brookhaven National Laboratory, and even to such things as Project Intrex at the Massachusetts Institute of Technology, which envisions a

---

The author, who has been editor of *The Physical Review* since 1956 and on its staff since 1951, came to Brookhaven National Laboratory as a theoretical physicist. He was born in Canada and got his PhD from the California Institute of Technology. This article is based on talks before the American Physical Society and the Washington Academy of Sciences.

RIGHTSLINK

Orderly communication through research journals may be jeopardized by a developing national information system that is beginning to encroach on the domain of the primary publication system. The author also believes mass distribution of unedited, unreferred and often unproofread preprints, which has recently been proposed, would put journals out of business or transform them into depositories.

system whereby individual researchers at their own individual desk consoles can query a computer for bibliographic and other information. Indeed it seems to be feasible in the not too distant future to put whole articles into computers to be printed out on demand.

In addition to this development of what can be called the "hardware" of information storage and retrieval, the government has also become increasingly concerned with the overall process of storage and retrieval of information. As you may have already read in the January issue of PHYSICS TODAY,[1] COSATI (Committee on Scientific and Technical Information in the White House Office of Science and Technology) under William T. Knox has taken on the horrendous task of planning a national information system. The system would embrace all government agencies dealing with scientific and technical information (for example those in AEC, NASA, National Bureau of Standards, Department of Defense) and would also include information activities of the so-called "private sector" (the scientific societies and their journals, compilation groups, etc.). The program would seek to promote the greater efficiency and compatibility of all these diverse groups, in all fields of science.

Two aspects of the COSATI program are particularly interesting: 1. The rough budget estimate for government information programs for fiscal 1966 is $380 million. 2. The COSATI group sees its program as a partnership of the government and the private sector in fashioning the tools—the information systems—to meet our present and future needs. In this connection COSATI arranged last May to have a meeting with representatives of about 25 large scientific societies, at which Burton M. Adkinson of the National Science Foundation's Information Service presented the general concept of the national information system and asked that the societies consider the extent to which they could and would participate in its planning and development. The implication was clear that if societies could not or would not take effective action in the information field, government would.

An unintended takeover?

I am impressed by the COSATI program for a national information system—by its magnitude, imaginativeness, cost and potentialities for serving and improving the dissemination of scientific information. But I am simultaneously impressed by the inherent danger when an enormous system of effectively unlimited resources (by our standards) teams up with a number of small organizations with limited resources. I fear that the partnership may easily drift into an unintended takeover and that in the process many of the most valuable features of the existing systems may be destroyed.

There are already in my opinion several ominous indications that this is not an idle remark.

One such indication relates to a statement made in a report from COSATI,[2] that COSATI recognizes the importance of continuing dialogues with representatives of various nongovernment components of the information-system complex. (This remark is also made in Knox's PHYSICS TODAY article.) As an editor of the largest research journal in physics, as a member of the American Institute of Physics publication board, and as the designated representative of the American Physical Society for liaison with COSATI (and under the assumption that other societies have had the same experience), I can only say that such liaison has been woefully inadequate even on a simple information basis, let alone consultation in the formulation of the program.

I think the report itself suffers from lack of adequate consultation with the scientific community. It seems to regard the major problem as one of collection, classification, storage, and distribution of documents—*any* and all documents that deal with scientific and technical subjects. The question of quality control *seems* to be a secondary matter. I stress the word "seems" because I do not believe that COSATI is unaware of the question of quality control. But other aspects of the problem are more directly amenable to attack with computer technology, and COSATI wants to get on with the job. Yet the two aspects cannot be treated separately—the easy one first and then the

RIGHTSLINK

other—because of their complex interrelation. The organizational structure must take this interrelation into account if the job is to be done successfully.

Further, there are strong indications that the developing national information system, which can render such a valuable service to the research scientist through a major improvement in storage and retrieval for the rapidly increasing volume of world scientific literature, is beginning to be entranced by its power to collect and store and retrieve and distribute *anything* and is beginning to encroach on the domain of the primary publication system itself—without an adequate understanding of the nature of that system and with a real threat to its effectiveness and indeed its survival. It is that threat which in my opinion makes the title of this article not simply a rhetorical question on which to hang an obvious answer. I think that scientific journal publication as we know it could be destroyed.

Orderly communication of new information

Let me first make clear what I mean by scientific journal publication as compared with absence of scientific journal publication. This distinction does not lie in whether automated computer equipment is or is not used. We would expect a publishing organization like the American Institute of Physics to speed up its operations by making use of the most modern equipment available—even to putting all of its journal papers on computers. The distinction between scientific journal publication and its absence also does not lie in the form the distribution takes. It does not matter whether one has printed groups of papers or separates. (There exist some journals now, such as the Danish Matematisk-fysiske Meddelelser, that print papers as separates.) It does not matter whether the distribution is in printed papers or microfilm or microcards or computer tape—though I think there will always be a need for ordinary library and personal availability of printed papers. Scientists will still want to have on their shelves well organized, compact, easily accessible collections of papers relevant to their special interests, and they will want to have the opportunity of browsing in papers that are not of immediate necessity in their work. It does not even matter whether papers or groups of papers are distributed from some central repository. The real distinction between scientific journal publication and its absence lies in who controls the *input* to the stored body of scientific information—the scientists themselves, basing their acceptance on the scientific merit of the work presented, or an outside agency such as a library system that makes no value judgment but classifies and stores everything it gets. In the latter case one has a collection of documents, not a scientific journal.

The major purpose of a scientific research journal is the orderly communication of new information. The key word is "orderly." Through editing and refereeing, scientists seek to ensure that the results presented in the scientific literature are correct, new and significant, that they are presented in a reasonably clear and understandable way with proper tie-in to the existing body of knowledge through introductory discussion and references, with adequate description of the new work and new results and with reasonable evaluation of their significance, range of validity, etc. In short, they seek to ensure that the published literature is worth recording, worth saving, and worth using. We editors cannot of course do a perfect job of attaining these objectives because of limitations inherent in any system—for example, lack of omniscience in referees, pressure from authors and readers for speed in publication, and limitations imposed by small financial resources. But we believe that we do set minimum standards and that the system results in significant improvements in many papers and elimination from the literature of much wrong information. We also believe that the mere existence of a reviewing system makes authors devote considerable attention to the preparation of their papers prior to submittal—though we wish that they would be even more careful.

Physical Review statistics

I would like to illustrate my remarks with some crude statistics concerning the journal with which I am most familiar, *The Physical Review*, the biggest and (I believe) the best of the physics research journals. In 1965 we received about 2600 papers. Of these about 20% or 500 papers were not accepted for publication. Of these 500, about 200 were rejected as being incorrect or otherwise below our standards (most of these actually were returned with severe referee criticisms and were not resubmitted); about 150 were withdrawn by authors; about 100 were returned as being much more suitable in subject matter for other journals; about 50 were crackpot papers.

RIGHTSLINK

Of the 2100 papers accepted for publication, about 1000 were accepted by our referees with no change or with minor corrections which did not involve return of the paper to the author; about 700 were accepted after more or less straightforward modifications (that is, with changes that we felt did not need reëxamination by the referee); about 300 papers required reëxamination by the same referee before acceptance; about 70 went to more than one referee before final acceptance. Most of the papers accepted went in for publication with messages to the authors to check certain minor points relative to notation, references, changes in English, referee's corrections, etc.

Incidentally, in evaluating these 2600 papers we used about a thousand different referees. Of these, about 150 received more than six papers each for evaluation during the year. Many other referees received only one or two. One of the great strengths of a professional organization like the American Physical Society is the very generous response of its members to the ever increasing burden of refereeing papers for their journals—a task that they accept as a professional responsibility.

I have presented these statistics to emphasize that editing a journal like *The Physical Review* is not a matter of a few top experts making arbitrary yes-and-no decisions on the publishability of submitted papers. The evaluators are themselves competent working physicists—experts in the subject matter that they are asked to evaluate. A referee is chosen for a specific paper because we feel that he is sufficiently interested in that subject and sufficiently responsible to read the paper in detail and write down the items that bother him as he reads it. The value of such refereeing lies far less in the yes-or-no judgment of the overall paper than in the service rendered by pointing out flaws that would bother or mislead most other readers. These flaws include misleading claims, omitted details, ambiguous statements, minor errors in the argument, overlooked pertinent references, unrealized implicit assumptions, unrecognized limitations to the conclusions, obscurity, discursiveness. The final result is an article that is easier and quicker to read in detail and to understand and use (or even to decide not to use) and is more reliable than the original preprint. This procedure makes for orderliness of communication through scientific journal publication.

Erosion of the system

This orderliness in communication of scientific research has already been compromised by the precursors of the national information system. Authors used to distribute a few copies of their papers to friends and associates for their information and to solicit constructive criticism before publication. Such preprints were cited in advance of publication only with the express permission of the authors. Largely with government subsidy, however, many major scientific institutions have for a number of years been distributing advance copies of the papers of their staff members to large lists of recipients including the major institutional libraries. These impersonal distributions have to a considerable extent constituted prior publication of unedited, unrefereed and often unproofread papers. These so-called "preprints" or reports are cited (without permission from the authors) by the people who make use of them so that the major libraries have been forced to expand their space to keep rather complete files—even for material that is later published in essentially unaltered form in the regular literature. The issuing organizations themselves tend to regard these reports as publications for reference purposes and for priority claims—and limit their distribution only to the extent necessary to keep the regular journals from rejecting them.

Journal editors have for the most part been timid about opposing this evasion of the concept of prior publication. They have tried with only middling success to get authors to keep the number of preprints down to a reasonable number and to upgrade their report-literature citations to the regular literature in galley whenever the full published reference was available. But in many fields, especially the more active ones, the reader who does not have access to the report literature is at a considerable disadvantage.

We now have proposals for the government to subsidize free and wide distribution of a more complete set of preprints within one large area of physics through a central collection and distribution agency. To my mind such a formal distribution, to a worldwide audience virtually equal (within the field of interest) to that reached by *The Physical Review* and other comparable journals, constitutes publication in any reasonable interpretation of the word and would be in direct competition with established journals. Being free, it would constitute unfair competition since the journals cannot match that price and stay in business.

It is my personal opinion that research journals

RIGHTSLINK

should reject papers that are previously given such a formal, wide distribution on the grounds of prior publication. For if they ignore the facts of the case and publish such papers, they will rapidly become only archival depositories, to be supplanted in the not too distant future by a central machine depository. Their only virtue would be neater printing. The refereeing system would be destroyed. For when papers have already effectively entered the full stream of distributed research in their original form, how can anyone expect referees to devote time later to review them in detail for archival burial or expect the authors subsequently to tidy up the corpses or expect those who have already read and used the preprint to reread the published paper?

I also believe that the journals, to be consistent, will have to tighten up their restrictions on the allowable size of the current institutional preprint distributions. A committee of the American Institute of Physics publication board is presently studying this question to come up with a uniform and viable policy.

Personally I fail to appreciate the rationale behind the belief apparently held by a number of theoretical physicists that the effectiveness of their research would be greatly enhanced if they could only be continuously bombarded with every thought and morsel of information, in however disorganized a form, that they can scrounge from every other researcher in their field. Is this really the nature of physics research? If so, why not put an open intercom or television system between groups of laboratories, so that the researchers can talk to each other continuously?

Pressures against self-restraint

Some of my friends have argued that the views I have expressed place too high a value on the refereeing system—that scientists in general would exercise reasonable self-restraint in publishing if they were free to publish whatever they wished in whatever form they chose and that evidence to this effect is provided by the existence of scientific journals that have little or no refereeing or editing. But I believe that this is an illusion—that these journals exist within a framework of the edited journals, which set the pattern and the standard, and that most scientists prefer to publish their work in journals which have high standards and reputation. Certainly the many instances when authors have vehemently contested referees' negative evaluations bear witness to this attitude.

Also, responsible scientists are too prone to attribute their own attitudes to all other scientists. No large group of people has ever been able to exercise such remarkable self-restraint, and the pressures of the day—competition, impatience to go on to new activities, and the publish-or-perish syndrome—all push toward increasing the output of fragmentary and hastily prepared reports. Certainly our experiences in the few cases when effective controls temporarily broke down in recent history (for example, the early days of the Mössbauer effect and the flareup of publications on SU (6)) indicate that removal of controls could easily mushroom into uncontrolled proliferation—a situation that in biological systems is called cancerous.

The state of affairs in publication is analogous to that in law enforcement. Law enforcement can never prevent all violations of socially acceptable conduct, but the mere existence of law-enforcement agencies enables most people to live within a reasonable code of social behavior.

In summary, the threat as I see it is that the orderly communication of scientific information may be supplanted by a huge depository of information into which will flow all the outpourings of scientists and pseudoscientists—the competent and incompetent, responsible and irresponsible alike. From this vast storehouse of scientific food—some of it garbage, scraps, and half-baked items, some of it spoiled and even poisonous, some of it reasonably edible or edible in part and some of it consisting of gourmet delicacies—the storehouse keepers will supply on demand all portions whose contents fit a specified scientific classification. And all this in the name of efficiency!

I realize that the developing national information system will be and should be document-collection oriented in many areas of its activity. But I hope that in the handling of its services to the scientific community it will be aware of the danger of indiscriminate collection and that its primary mission will be regarded as one of strengthening the scientific literature, not diluting it—of strengthening it through the development of better and more efficient classification, storage and retrieval techniques. I hope that the information system will also recognize the necessity of coöperating with scientific institutions in helping them mod-

A Debate on Preprint Circulation

In June, PHYSICS TODAY will publish a three-part debate on the PIE proposal for rapid circulation of preprints. Michael Moravcsik will present the case for his PIE proposal. Simon Pasternack will state his opposition. Finally Moravcsik will offer rebuttal.

RIGHTSLINK

ernize and improve their own publication efforts without destroying their unique concept and purpose.

A frightening growth rate

The major problem of the regular journals is that, even with the controls imposed in the interest of orderly publication, the scientific literature is growing at a frightening rate. For example, *The Physical Review* has doubled in size in the past five years. Since this is a faster rate than that of the total physics literature, I could quote a well known commercial: "We don't know what it is, but we must be doing something right." This total growth in itself threatens the effectiveness of the journals because they are becoming too bulky and too expensive for wide individual distribution and use. We are trying hard to meet this problem, insofar as *The Physical Review* is concerned, in several ways. We are paying more attention to adequacy of titles and abstracts since they will be all that many physicists will read except in their specialties. We are paying more attention to the problem of indexing papers for future retrieval, and here of course we are vitally interested in activities of the national information system in promoting development of an overall physics classification scheme. We are beginning to involve authors in the classification of their own papers, and we hope that this will furnish a feedback on the adequacy of the index. We are grouping papers into sections, each containing one area of physics, so that physicists may be able to subscribe at a reasonably low rate to those portions of the journal that they would like to have on the shelves of their offices for ready referral. We are looking into ways of speeding up our handling of papers without compromising our standards. The American Institute of Physics is looking into ways of speeding up the printing process. It is also embarking on an ambitious set of projects, including development of a common index scheme for all of its physics journals, development of a multicoördinate index scheme for computer storage and retrieval and research into user needs.[3] The American Physical Society is looking into ways of stimulating an increase in the woefully inadequate supply of critical review articles that can help in reducing the need for referral to original research articles. In all of these activities we need the understanding and active support of the physics community. In many of these areas we need also the interest and understanding and help of the developing national system.

There are other broad areas of science publication that need the attention of responsible scientists. The relationship of a national information system to the international scientific community requires careful consideration. The problem of conference proceedings, which, like reports, often occupy an intermediate position between unpublished and fully published work, needs effective evaluation and control. The trend toward proliferation of small journals with inadequate distribution and lack of time or resources to participate in the developing overall information system may give way to a trend toward consolidation into larger, more viable units. (This problem is, of course, nowhere near as acute in physics as it is in biology, which has several hundred small journals.) The role of commercially sponsored journals needs to be evaluated. With the increasing demands that the information network is already beginning to make on the time and attention of editorial offices, there will have to be a reëxamination of the cost structure and financing of journal publication.

Need for involvement of scientists

In conclusion, research scientists will have to make an effort to become better informed concerning their publication problems and more actively involved in solving them if they are to mold the rapidly changing situation to meet their own needs and desires. Otherwise the government will try to do the job for them. I would like to see much more discussion of these problems—in journals and at society meetings. There should be more active representation and participation by research scientists in the planning and development of the programs of the national information system—for example in development of a universal classification scheme in each discipline, with compatibility between disciplines, which is the crux of the effectiveness of the national system in meeting the needs of scientists. And finally there must be a general awareness of and sharpening of the values that are worth preserving in the changing publication patterns of science. □

References

1. W. T. Knox, PHYSICS TODAY **19**, no. 1, 39 (1966).

2. COSATI Report PB 168 267 (AD 624 560). *Recommendations for National Document Handling Systems in Science and Technology*, available from Clearinghouse for Federal Scientific and Technical Information, US Department of Commerce, Springfield, Va. 22151.

3. V. Z. Williams, E. Hutchisson, H. C. Wolfe, PHYSICS TODAY **19**, no. 1, 45 (1966); P. A. Atherton, PHYSICS TODAY **19**, no. 1, 58 (1966).

RIGHTSLINK