DOCUMENT RESUME

ED 040 722 LI 002 039

AUTHOR Steere, William C., Ed.
TITLE Information Handling in the Life Sciences.
INSTITUTION National Research Counicl, Washington, D.C.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE Feb 70
NOTE 87p.

EDRS PRICE EDRS Price MF-$0.50 HC-$4.45
DESCRIPTORS Biological Sciences, Information Centers, *Information Dissemination, *Information Needs, Information Networks, *Information Processing, *Information Science, Information Services, Information Sources, Information Systems, *Information Utilization

ABSTRACT

Special problems in the handling of biological information arise from the diversity of biological subject matter and the complexity of biological approaches towards phenomena of the living world. This state-of-the-art report on communications of information in the biological sciences provides information on: (1) users of biological information, (2) informal transfer of information, (3) primary publications, (4) secondary information services and (5) libraries. Detailed suggestions relating to communication of biological science information include: (1) design of information systems to include the general user and (2) compatible interconnection of the three existing major organizations in biology - - the National Agricultural Library, the National Library of Medicine and the Biosciences Information Services of Biological Abstracts. Also outlined are various ways to improve: (1) informal transfer of information, (2) primary publications and (3) secondary information services. (MF)

Information Handling In the Life Sciences

Council on Biological Sciences Information

DIVISION OF BIOLOGY AND AGRICULTURE
NATIONAL RESEARCH COUNCIL

INFORMATION HANDLING
IN THE LIFE SCIENCES

Council on Biological Sciences Information

DIVISION OF BIOLOGY AND AGRICULTURE
NATIONAL RESEARCH COUNCIL
Washington, D. C.
February, 1970

The Council on Biological Sciences Information has been largely supported by the National Science Foundation through Task Order 127, Contract C-310, with the National Academy of Sciences.

CONTENTS

FOREWORD

This report was first conceived of as a staff paper to cover in some depth the state of the art of communication of information in the biological sciences, and was intended largely as a manual or reference work for the benefit of the members of the Council on Biological Sciences Information. It was started late in 1967 by Beverly L. Clark, a staff writer for the Council. Sections on special topics were contributed by several members of the Board of Directors, notably Robert E. Gordon, J. Roger Porter and the late Fred R. Cagle.

Early in 1968 Philip Handler invited the Council to prepare a review of information handling for the major report then being written by panels of his Committee on Research in the Life Sciences of the National Academy of Sciences. I undertook to prepare it, using as much of the previous material as possible. Also, as a member of the Committee on Scientific and Technical Communication of the National Academy of Sciences and the National Academy of Engineering, I had access to and familiarity with much up-to-date documentary material and many timely ideas in various aspects of scientific communication. I have drawn freely from these sources in the present report, and especially from my own notes on discussions that took place in meetings of SATCOM. The recent publication of the SATCOM report, "Scientific and Technical Communication: A Pressing National Problem and Recommendations for Its Solution," and its accompanying brief "Synopsis" (Cairns et al., 1969), now gives wide access to much of this background material.

The draft was edited by F. Peter Woodford, who made many improvements in it, added new material and drafted recommendations. Since then, it has received still further editorial attention and revision from Robert E. Gordon, who prepared the appendix on the history and development of the Council. A very thoughtful section on tables of basic biological data has been contributed by John T. Edsall. I am grateful, also, to the several interested persons who contributed helpful suggestions, improvements, and corrections for the August 20, 1969 draft of the report, which was circulated at the Second Roundtable on Information Problems in the Biological Sciences held at Burlington, Vermont on the same date.

WILLIAM C. STEERE, Editor

PROLOGUE

In 1866 Gregor Mendel published in an obscure Austrian periodical a classic paper reporting his experiments in plant hybridization and delineating the basic mechanisms of heredity. Mendel's paper was not just a suggestion or a finger pointing in the right direction; it was a profound and convincing report based on long experimentation and critical statistical analysis. It was one of the great biological papers of all time, economically written in perceptive and elegant language. But it was not widely disseminated. No abstracting and indexing services, no "current awareness" systems existed to bring the paper to the notice of other scientists who would realize its significance. Almost four decades passed before the full-fledged study of genetics could begin with the rediscovery of Mendel's work by Correns, DeVries, and Tschermak.

Another dramatic example from more recent times, of course, was the failure in the communication of Alexander Fleming's paper on the bacteriostatic properties of penicillin, which had to be "rediscovered" later.

Such failures in the communication of information has many modern parallels--the discoveries of cytochrome c and the technique of adsorption chromatography are two well-known examples--and there must be many of which we are unaware, for just these reasons. This kind of failure is one that a modern information system should be designed to prevent. But there are other problems in information transfer today, problems that stem from the vastly increased volume of publication in journals both prominent and obscure and in several other media. For any person in society, be he scientist, teacher, or administrator, only a small percentage of all the information being generated will be relevant and important to his work and intellectual development. Strong information systems must be designed to provide him adequately with such material and protect him from being deluged with the rest.

Although similar problems of information handling exist in all disciplines and although our ultimate aim must be an information system for all knowledge and all human activity, special needs arise within each discipline. This report attempts to delineate the needs for the biological sciences, to describe the present systems as they have grown up in the absence of any specific overall plan, and to put forward suggestions for improvement. We stress that these improvements cannot come about through the devoted efforts of a few; continuing contribution and refinement must be made by every biologist in the community (Orr, 1963).

INTRODUCTION

A Perspective

The object of any biological information system is to provide the user of biological information with the information he needs when he needs it, in the form most useful to him and at minimum expense. Most important, the system should ensure that the user is receiving full information on the topic of his interest, on a worldwide basis, as far back in time as desired, and in the appropriate degree of depth.

Technological developments greatly influence the form that an information system takes. The scientific work of antiquity could be disseminated only orally, by correspondence, or by laborious copying of manuscripts by scribes. Scholarly activities were largely restricted to the few centers where extensive collections of manuscripts were available. A modern parallel is the restricted dissemination of unpublished work within the "invisible colleges," which generally contain only a few "centers of excellence." Development of the printing press in the fifteenth century revolutionized the dissemination of knowledge, and a steadily increasing flow of publications resulted. The invention of the computer and the development of multiple and complex applications of it for communication, storage, retrieval, and even analysis of information have already created a greater revolution than the invention of printing with movable type--and the computer age has only just begun.

The burgeoning of scientific and technical information resulting from the tremendous worldwide growth of activities in all disciplines has been a matter of national concern for more than a decade. During this time, several prestigious groups have studied the problem and have recommended many approaches to improve the handling and dissemination of information, both by the government and by the relevant private services. Our concern is primarily with the non-governmental, or "private," sector of the community, although the sectors interlock and cannot be neatly segregated, even in theory.

As a background for our discussion of biological information systems, we summarize here the recommendations for all of science given by the Weinberg panel of the President's Science Advisory Committee early in 1963 (Weinberg et al., 1963).

- Authors must accept more responsibility for information retrieval.
- Unnecessary publication should be eliminated.
- American technical books must be improved.
- The technical community should give higher status to the writer of review articles.
- Modern psychological insights into communication should be exploited.
- Scientists and engineers must express themselves clearly.
- The technique of handling information must be widely taught.
- The technical documentalist must be recognized and supported.
- New "switching" methods must be explored and exploited.
- Centralized depositories are an attractive possibility.

- More and better specialized information centers are needed.

- Mechanization can become important but not all-important.

- Citation indexing should be useful.

- Secondary information services should be mutually compatible.

- Nongovernment technical publication will require government support.

Special Problems in the Handling of Biological Information

The special problems in the handling of biological information arise from the diversity of biological subject matter and the complexity of biological approaches toward phenomena of the living world. Thus the subject matter ranges from subcellular organization through primates to highly complex communities and ecosystems, and from the reaction of a single enzyme to the behavior of populations. The types of approach include the experimental manipulations characteristic of biochemistry and molecular biology, the observation, description, and synthesis of taxonomy, and the teamwork of many kinds of specialists in both observation and experimentation in behavioral biology and ecology. This diversity is reflected in the twenty major categories that the National Science Foundation finds necessary for biology, as opposed to four for chemistry and ten for physics; in the existence of hundreds of societies and associations in biology; and, perhaps most relevant of all in the present context, in the large number of journals that biologists have found it necessary to establish in order to satisfy real or imagined specialized needs. Of an estimated 26,000 important periodicals in science and technology (Barr, 1967), the number for biology alone is 13,000. (See Primary Publication, page 18, for justification of these figures.)

Although the problems of handling biological information are no different in kind from those in the physical sciences, factors of diversity and complexity make them vastly different in magnitude. In addition, since biologists are deeply involved with the most important of the world's concerns--man's health, the expansion of populations, and the effects of man's activity on the environment--any failures in the communication of biological information are liable to have the direst consequences. This is why the problems of handling biological information must be solved.

The Users of Biological Information

The design of any system must be based on a clear understanding of the needs of those who are to use it. The users of biological information and their needs may be roughly categorized in the following simple terms, which are by no means fully comprehensive.

- The biologist-scholar (researcher and advanced university teacher) needs specific and detailed information to carry on his current research, to plan future research, and to transmit new knowledge to his students.

- The practitioner needs information relevant to his particular application of biological knowledge and principles, whether to medicine, agriculture, horticulture, or forestry, and in a degree of depth depending on whether he is a planner, manager, executive, consultant, or industrialist.

- The secondary or elementary teacher needs up-to-date information adapted to his teaching responsibilities and to the level of his students.

- The policy-maker or administrator does not usually need biological information, as such, but he does need information about biology--the relevance of biology to educational, economic, and national planning, the management problems in working with biologists, and the availability of biologists as consultants.

- The <u>citizen</u> needs information, both biological and biology-based, pertinent to such major subjects as conservation, civic planning, pollution, and public health, and even for such pragmatic matters as the objective evaluation of advertising claims for new drugs and medicines.

Naturally, users in the various categories need information in different forms. The scholar desires the original publications, at least in theory; actually, some will accept other scholars' syntheses (i.e., review articles) from original documents. On the other hand, the practitioner does not usually need original publications and is content to have only statements or digests indicating how biological information can be used to improve his production or services.

The teacher requires highly consolidated texts and teacher aids providing current information adapted to his teaching responsibilities; however, some teachers do considerable reading in the basic original literature to keep themselves up to date on the kinds of methods used to arrive at the results they intend to teach.

The policy-maker and administrator have scant use for information as presented in primary biological journals, but they urgently need concise analyses of the relevancy of biological developments to their task. These analyses are normally published in such general-essay magazines as <u>Science</u>, <u>Nature</u>, <u>BioScience</u>, <u>Endeavour</u>, <u>Scientific American</u>, and <u>New Scientist</u>, or those concerned with management or planning.

The citizen depends primarily on the mass media for his information about biology. Herein lies one of the greatest challenges to an information system. At present, little or no attempt is made to record and provide later access to the transitory products of newspapers, radio, and television, and it is debatable whether it would be worth the effort to do so. However, the system should provide easy access to biological information for those responsible for newspaper columns and radio or television programs on science, especially as it relates to the population explosion, international agriculture, medicine and public health, and even politics. Certainly the less ephemeral literature about biology published in such thought-provoking nonscientific journals as Daedalus, Orbis, Atlantic Monthly, and Saturday Review, should be entered into the information system in such a way as to be retrievable at a later time. The annual index to The New York Times is also an invaluable guide to well-written articles on biological topics of general interest to the public.

Satisfying the information needs of each category of user entails solving the problem, in each case, of appropriate selection from the information available. For the biologist-scholar, the magnitude of the task increases with the breadth of his interests. The scholar with closely defined interests in a field that has been long established may state quite realistically that he has no serious information problem at present. On the other hand, the creative scholar interested in using new approaches and correlating work from many different fields has a more difficult task and undeniably needs the support of improved information services. The same is true of the most innovative administrator and the most inspiring teacher.

The biologist-scholar needs to learn of new ideas with minimum delay. The practitioner does not have so great a need for rapid dissemination of basic research information, but in an emergency he may need rapid retrieval of a specific fact or technique. Some delay in dissemination or retrieval may be tolerable to the school teacher, policy-maker, or citizen, but progress may be seriously inhibited if the introduction of new biological information into the educational system, the government, industry, business, or the mass media is subject to extremely long delays.

INFORMAL TRANSFER OF INFORMATION

Often condemned by those who love a neat and tidy world, informal scientific communication not only complements but nurtures the formal means of transfer of information. The questions faced by the designer of an information storage, retrieval, and dissemination system are: What kinds of informal communication should be integrated into the system, in what degree of detail, under whose control, and by what means? We can neither ignore the whole nor include the whole of informal communication in the system; where is the distinguishing line to be drawn? In this section we describe several kinds of scientific communication that are less formal than primary publication in journals. They differ greatly in degree of importance and in the kinds of problems they pose.

Meetings: Informal

Personal conversations and correspondence among scientists have always provided a vital link in the exchange and development of scientific thinking. Seminars and other discussion groups are valuable adjuncts to the educational process in universities and industry. Much information is exchanged in this way that is unlikely ever to be published: details of experiments and technique, descriptions of unsuccessful experiments, striking or fanciful analogies and inter-relationships far too tenuous for cold print but potentially enormously fruitful. Because scientists recognize the stimulating effects of informal discussions, they flock to meetings of all sizes and kinds in the valid expectation of benefiting from this kind of interchange.

In an attempt to capture and disseminate such conversations, symposia and discussions are sometimes tape-recorded and subsequently published almost verbatim. Some readers of the resultant volumes claim to be inspired by them, but often at much cost in time and patience. Because the discussants are rarely pressed for accurate documentation of their statements, the basis of their remarks frequently cannot be discovered and they may quote results inaccurately in the haste of making their point. Although such publications may be of sociological interest, we rank them low in priority for publication and inclusion in an information system.

Meetings: Formal

Besides the informal, fruitful conversations described above, a scientific congress or meeting sponsors the delivery of formal papers. Should these be gathered together as the proceedings of the congress and published? Although at first glance such publication seems highly desirable, the views of most scientists have hardened against it in recent years, and for good reasons.

- The work presented in a delivered paper is often already in the process of being published in a journal. There is nothing against its being described orally for twenty different groups if they are willing to listen, but there is everything to be said against its being published in half a dozen different places--the expense of multiple publication, the expense of entering multiple citations into the system and retrieving them, the trouble and expense of obtaining the documents from the citations retrieved, all for the sake of one piece of information.

- The work to be presented is rarely reviewed critically by a peer group, so that substandard or trivial papers that would be rejected by a journal are frequently presented orally.

- The publication of a complete Proceedings is almost always slower than that of the slowest journal.

- The discussion of the papers that seemed so pertinent and thought provoking at the time may emerge--usually for mechanical reasons--as trite or incomprehensible in print.

- The knowledge that his delivered paper is to be published too often inhibits the speaker from speculating imaginatively or even speaking intelligibly as he mumbles away with eyes glued to his prepared typescript.

Scientists should realize that oral and written communications serve various purposes by different means and that the information they contain should, therefore, be handled in different ways (Woodford, 1969). Thus, oral communication has the advantage over written in that it is more up to date, adaptable to many audiences with different interests in the information, subject to critical correction or reinforcement when the research described is at an early stage, and open to questioning and clarification by the audience. On the other hand, it is more redundant, less detailed, and more liable to misinterpretation, since the audience cannot examine every detail at leisure.

For all these reasons it seems unwise or unnecessary to enter the full proceedings of formal meetings into the information system via publication. One of the best ways to disseminate news of what occurred at such meetings is to have a leading participant at an advanced (that is, nonroutine) meeting prepare a succinct report of it for publication in such a journal as Science, Nature, or BioScience that has a short publication time. The report can accurately cite the participants' published work that led up to the talks presented and can convey some of the spirit or even the substance of informal discussions between formal sessions.

Some meetings, however, deserve publication in full. These are the carefully organized symposia in which a few speakers review their own and others' work in relation to a central theme. Such symposia, appropriately edited and even annotated by the organizers, can provide just the syntheses of an area that the educator or policy-maker needs--for whom, as we noted above, some delay in hearing about the original research may be tolerable. Deciding between what is an integrated and publishable symposium and what amounts to an unpublishable collection of barely connected research reports is, of course, a matter of editorial or professional judgment and conscience for which no infallible prescription can be provided.

Listings and Calendars of Meetings and Conferences

Whatever the ultimate claim of a meeting to be immortalized in print, there is no question that all meetings should be well publicized in advance and that details of forthcoming meetings should be listed so that juxtapositions or conflicts between similar ones are readily apparent. Such listings for all of science are given in the publications "World Meetings, United States and Canada,"* which is indexed in several different ways--subject, area of science, location, date--and "Scientific Meetings",** and in special sections of _Science_, _Nature_, and _Chemical and Engineering News_. For biology, the listing that appears in _Bioscience_ could be amplified and strengthened if the editors of all journals and magazines that publicize scientific meetings on request were to forward all copy for such publicity to the American Institute of Biological Sciences for further processing as soon as they receive it. It is essential that an up-to-date, comprehensive, and reliable clearinghouse be established (Baum, 1966).

Technical Reports

Technical reports may be described as "half-published." They describe work done for the government in its own laboratories or under contract by individuals or major corporations; many of them are produced in hundreds of copies and are given a wide but controlled circulation. Sometimes the work is also published formally; if not, it is difficult of access even though it may be essential to the mission-oriented biologist working in applied areas.

---

*Technical Meetings Information Service, 79 Dumlin Road, Newton Center, Mass. 02159

**Special Libraries Association, 235 Park Avenue South, New York, N. Y. 10003

To publish lists of the unclassified reports, to give them document numbers, and to make them available through a central clearinghouse seems to be a Federal government responsibility, since most of the technical reports originate with government agencies, and this ponsibility is now being assumed (Fuccillo, 1967; ssman et al. 1968). The Clearinghouse for Federal Scientific and Technical Information* now makes most reports of this kind readily available. Criteria for inclusion of documents in their files have not been disclosed; a possible criterion is discussed below in the section on newsletters (p. 13).

Information Exchange Groups

The verbatim accounts of roundtable discussions referred to above can be described as an attempt to capture scientific conversation for wide dissemination. The basic idea of information exchange groups, begun in 1961 and continued with U.S. government support until 1967, was to duplicate scientific correspondence for distribution among rather large numbers of scientists with similar interests. Any scientist working in the research area described by the title of the group could become a member of the group and be entitled thereby to contribute and receive memoranda, which were duplicated and distributed by a central office in the National Institutes of Health. It was originally intended that the memoranda should contain new, unreviewed results, improvements in technique, speculations, and free discussion--informal matter of the kind usually confined to letters between scientists. Unexpectedly, but in retrospect understandably, 80 percent of what was circulated consisted, in fact, of formal manuscripts ready for submission--or already submitted--to journals for publication. This abuse of the system and its cost soon brought the experiment abruptly to an end, although one or two information

---

*Clearinghouse for Federal Scientific and Technical Information, Springfield, Virginia.

exchange groups are continuing on a much reduced level without Federal support. Possibly, if the original purpose of the groups--facilitated diffusion of truly informal information--had been clearly kept in mind, such groups might be a flourishing part of the information system today. There are indications that the idea may be usefully revived under controls that would clearly distinguish such informal exchange from formal publication.

Circulation of Manuscripts Prior to Publication

The submission of formal manuscripts for circulation by the information exchange groups was the natural outcome of the custom among many scientists of sending to their colleagues, in lieu of letters, photocopies of manuscripts that are being considered for publication. The custom is in many instances a courteous and efficient response to the many colleagues who have requested to be "kept up to date with developments in your lab" and provides information many months in advance of its appearance in the formal literature as well as saving the recipient the trouble of searching for it.

The disadvantage to the recipient is that he reads, in what has the appearance of final form, material that has not been through the refining process that constitutes journal publication. Detailed, critical review may subsequently induce the author to revise the manuscript--including the conclusions--drastically before it is published. As the published paper is likely to bear the same title as the unrevised manuscript, the unwary reader may wrongly conclude that he has already digested its contents.

The dangers are not great provided the author and recipient recognize them. Any attempt to widen, formalize, and depersonalize the channel increases the dangers, for if the manuscript is circulated by anyone other than the author, it seems to gain outside sanction. Any proposals in this direction should, therefore, be resisted (Pasternack, 1966). Authors who discover that they are usually required to revise their manuscripts soon get into the habit of circulating only the final version, thus losing some of the timeliness of the communication but increasing its accuracy. To this practice there can be absolutely no objection.

Some publications provide preprints of meeting presentations in advance of appearance in print and their circulation to a limited mailing list is beneficial.

Newsletters and Other Ephemeral Publications

Much useful information is conveyed by leaflets, brochures, and newsletters that may be printed but can hardly be formally published. In this category come the brochures produced by manufacturers and suppliers of drugs, equipment, and chemicals. In general, the form in which this information is issued is too ephemeral to be acceptable in citations. Yet one cannot be too dogmatic about this as, for example, some method used in an experiment may be described in the equipment brochure and nowhere else.

At about the same level of informality are the unrefereed newsletters put out by many laboratories, societies, and institutions--for example, the picturesquely named Shellfish Soundings, Mosquito News, Mouse News Letter, The Worm Runner's Digest, The Eggsaminer, and The Milky Way. Such is the perversity of human nature that some scientists will "publish" in a newsletter important results that others may wish to quote in a more formal journal. The editor of the latter, however, must be firm: quotation must be dependent on the original author's permission and must take the form "personal communication" since the results reported cannot be regarded as satisfactorily documented or reviewed. With these restrictions, the newsletter can be regarded only favorably: it puts people with similar interests in touch, it circulates news and ideas, and it does not confuse the formal information system (Woodford, 1969).

What happens when the newsletter is so successful that it aspires to become a journal? Often, the desired change is signalled only by a change of name to "Bulletin" and the addition of a printed cover. At what point must the information system recognize it, acquire it, store it, and make it accessible? There are no easy answers to these questions. To apply the criterion of whether the articles in the bulletin have been subjected to review by peer scientists would be to exclude from the information system many publications that are widely regarded as admissible journals--for despite general agreement that such review is desirable, nowhere near all journals employ it. A more realistic criterion can be derived from the definition of primary publication recently arrived at by the Council of Biology Editors: citable publications are those in which sufficient details are given to enable peer scientists to assess the probable soundness of the results, to repeat the experiments if necessary, and to retrace the lines of reasoning that led to the conclusions. The criterion of "sufficient detail" for distinguishing formal publication from informal circulation cannot be applied without the exercise of judgment; nevertheless, it seems one on which any authoritative body should be able to reach satisfactory decisions.

This issue has been discussed at length because unnecessary difficulties arise when the line between formal and informal communication is allowed to become fuzzy. Both types of communication have important roles, but the roles are different and should not be confused.

The criterion of "sufficient detail" brings technical reports comfortably into the realm of formal publication. The degree of detail they contain is frequently not only sufficient but overwhelming. Now that these reports are under bibliographic control, that is, have a document number and a supplier* (the U.S. Government) who is unlikely to go out of business, there is no reason to exclude them from the formal literature. They are difficult to acquire and to catalog, but we can expect that libraries will eventually become as adept at handling them as they are at handling journals and books.

---

*Clearinghouse for Federal Scientific and Technical Information, Springfield, Virginia.

Listings of Research in Progress

An interesting development in informal communication has been the compilation and distribution of lists of research projects in progress in some such specialized field as systematic botany. The American Society of Plant Taxonomists has collaborated with the International Organization of Plant Biosystematists (Jackson, 1966, 1969) in producing the ASPT-IOPB Index of Current Taxonomic Research, which serves the purpose of announcing research under way but not yet published, thereby alerting scientists to major research projects that might overlap, duplicate, or even conflict with their own. Although this mechanism has faults, especially in allowing investigators to "stake out" claims to certain problems, at least their interest is brought to light and the depth of their involvement can be ascertained through correspondence or conversations.

The Science Information Exchange of the Smithsonian Institution began originally as an information center about ongoing research based only on grants made by Federal agencies. However, through the years its coverage has become comprehensive in the nonfederal area also; in fiscal year 1968, SIE registered 16,658 nonfederally supported research projects, which is probably the best coverage anywhere in the country.

Because research investigations change direction frequently, SIE attempts to keep listings as accurate and complete as possible by asking grantees to bring their summaries up to date each year, but if a grantee does not answer the questionnaire deficiencies can develop in the system. Our impression is that a conscientious research worker who attends well-chosen meetings can probably get more and better information from his colleagues working in the same specialty by discussing the problem personally and in depth. However, for the research planner, administrator, and policy-maker, who are less likely to attend specialized meetings, the SIE listings may be a highly useful guide, especially in correcting the biased impressions of research specialists advising from within their own narrow areas of competence.

Book Reviews

Book reviews are informal in the sense that they are unrefereed and often contain more unsupported opinion and speculation than do other portions of a journal. Nevertheless, because they often incorporate valuable ideas that the reviewer is making public for the first time, they deserve to be indexed or even abstracted, however briefly, and stored in the information system. For example, many specialized journals in various restricted areas of biology publish book reviews by leading specialists that are in themselves original contributions, even with citation of pertinent literature. Too often, these original ideas are overlooked, and they almost never reach the mainstream of our information systems. Excellent book reviews appear in Science, BioScience, Nature, Quarterly Review of Biology, American Scientist and other journals of a more general character.

A remarkable instance of how the rich content of book reviews can be extracted is provided by the annual Mental Health Book Review Index, in which are recorded bibliographic data on reviews of books dealing with mental health. From this compilation and a consideration of the book reviews themselves, the small committee responsible for the publication distills an editorial that points out and comments on the trends in mental health research and current attitudes to it, as reflected in the book reviews of the preceding year.

Book reviews have their negative aspects, and the uncritical reader may be led astray by a careless or irresponsible reviewer. But because they are potentially highly stimulating and informative and because they are usually published in accessible journals presenting no technical storage problems, they deserve, in our opinion, a place in the information system. Further exploration of the approach of the Mental Health Book Review Index in other fields seems potentially rewarding.

PRIMARY PUBLICATION

The Scientific Journal

Journals were first established as a means of scientific communication more efficient than oral presentation or correspondence. One of the first, the Philosophical Transactions of the Royal Society in England, was begun in 1665 as an outgrowth of the regular meetings of the society and of the voluminous correspondence conducted by its secretary (Porter, 1967). The fact that meetings and letters still play an important part in scientific communication does not mean, as some maintain, that journals fail to fulfill their purpose, but that the two lines of communication are complementary.

Correspondence and frequent meetings between scientists are feasible only when their numbers are small. For the large global community of scholars, the printed journal, circulated in large numbers, is essential. With the growth of the scholarly community, which produces information as well as consuming it, the number of primary journals has increased and the number of articles in each per year has also tended to increase. These increases in volume create unrest and dissatisfaction among researchers and teachers, largely because of the frustration induced by the knowledge that however diligently they read "the literature" they will still never be able to read all that is potentially valuable in their work (Kennedy and Parkins, 1969).

How serious are the problems faced by users of biological journals, and how may they be alleviated? In discussing these questions, the following themes will be developed:

- The growth rate and actual volume of scientific publication has been exaggerated in some quarters, but even the current volume is large enough to justify a reappraisal of the purpose, functions, and use of journals.

- Scientists will have to be educated both to search the literature more efficiently and to improve the quality of their contributions.

- Means will have to be devised to induce and enable journals to raise and maintain even higher standards than at present.

- Although research into alternative means of fulfilling the purpose and functions of journals, for example, by using computer memories and networks, should be vigorously pursued, the journal will remain the most economical means for several decades.

- If the journal is superseded, high-quality input to the system will remain preeminently important and will only intensify the need for skilled communication by scientists and quality control by network "editors."

- Secondary services, discussed in a later section--including abstracting and indexing services for the whole of biology, specialized information services for small parts of it, and critical review articles--must be developed further to aid the individual scholar in his work, but cannot solve all the problems on their own.

Authors, editors, and referees must all make strenuous efforts to improve primary publication, which is the fountainhead from which most scientific information flows.

Number of Biological Journals

In recent years the estimated number of scientific publications has been greatly exaggerated, mainly because of uncritical extrapolations of Price's statement (Price, 1961) that the number of scientific journals has been doubling every ten to fifteen years from an approximate figure of 10,000 in 1900. Even if we take fifteen years as the doubling time, the predicted number of journals for 1960 would be 160,000, whereas Gottschalk and Desmond (1963) estimated the number in 1963 to be only 35,000. The fallacy that led to this discrepancy was, of course, that the first calculations were based only on the number of journals <u>founded</u> and ignored their mortality rate, which according to Barr (1967) was 33-40 percent during the period under discussion. This percentage must still be too low, or the rate of doubling has been overestimated, since Barr's count in 1965 was 26,000--lower than either Gottschalk and Desmond's or another recent one by Hutton (1967).

These figures are for the whole of science and technology. If we accept Gottschalk and Desmond's percentage breakdown into disciplines, the life sciences would account for about 50 percent (13,000) of the total--20 percent for agriculture, 13 percent medical, 4 percent basic life sciences, and 10 percent technology. Biological Abstracts abstracted some 7,500 periodicals in 1968, but critical examination of the list shows that the majority are unlikely to publish anything that will materially advance the progress of science (BIOSIS, 1969). It is possible to identify--although probably no one would dare do so in print--a list of about a thousand journals in which upwards of 90 percent of the significant original work in biology appears.

The growth of the number of published pages in the major biochemical journals from 1950 through 1965 has recently been reported (Biochimica et Biophysica Acta, 1966). The number of equivalent pages published increased from 12,300 in 1950 to 47,800 in 1965, a fourfold increase in a period of fifteen years.

All numbers connected with publication tend to be "soft"--that is to say, they are too imprecise to stand critical examination or they have to be hedged about with so many qualifications concerning the inadequacy of the sample that their significance remains doubtful. We deplore the lack of a firm base of information about scientific serials, and recommend that an accurate list of all primary publications in biology be drawn up, complete with all data necessary for information analysis and research, not only for one-time publication but for computer storage and continuous updating. Contrary to popular belief, such a definitive list does not exist at present and never has existed in the past; the lists currently maintained by the National Library of Medicine, the National Agricultural Library, the Library of Congress, Biological Abstracts and Chemical Abstracts cannot simply be merged since they are in forms that are both conceptually and technically incompatible. Preliminary work by the Council on Biological Sciences Information has shown that the degree of overlap between medical, agricultural, and biological listings is smaller than expected; that the labor of merging the lists is considerable but not overwhelming; and that predictions and decisions about total information in biological sciences must remain inspired guesswork until a definitive list is produced (Gurtowski, 1968).

Whatever the accurate numbers may be, the scientist user of information is aware that an increased fraction of his time is now necessary to cover in even the most elementary way fields that he used to be able to study at leisure. Primary publications could lighten his burden in many ways, all of which may be regarded as having the reader, not the author, as prime consideration.

Improvement of Quality in Scientific Journals

The Editorial Process

The conventional screening process that is applied by the editor, his editorial board, and outside reviewers, all of whom are usually scientist peers of those who submit manuscripts, aims at selecting publishable material in the best interests of the scientific community. The purpose of this process is to afford rapid publication of important work that materially advances the understanding of nature; to permit the recording of competent, interesting work that will assist others to make important advances; to set a low priority on competent work that seems to provide no basis for theoretical or practical advances; and to suppress the publication of incompetent or incorrect work. Most journals attempt to publish all work in the first three categories and to exclude that in the fourth (Woodford, 1969).

The question is, does the present system suppress publication of work in the fourth category? With the multitude of journals available to the biological scientist and the great pressure to publish exerted by an academic system that prizes publication as an indicator of competence and diligence, authors do have a tendency to submit rejected manuscripts to journals with progressively lower standards, or smaller backlogs, until the work is accepted for publication. It must be recognized that this reaction is not universal; conscientious scientists whose manuscripts are rejected often repair or extend their experimental work until it meets the high standards of the journal originally selected, usually utilizing the criticisms of the expert reviewers to redesign their work. But the pattern of "going down the pecking order" of journals is unfortunately well established among the more cynical or lazy. We should, however, point out that the existence of more than one outlet is actually healthy--editors and reviewers are human and therefore fallible, and may mistakenly reject work that is so new and original that its significance has escaped them. It would be unfortunate and unjust if that work then had no chance of reaching the biological community through another channel.

Clear lines of demarcation between "good" and "bad" are always difficult to draw, but most scientists agree that, pragmatically speaking, some existing journals may definitely be regarded as superior, some inferior. It seems not beyond the bounds of possibility that a group of scientists in each country could select, within each discipline, journals to be designated "approved," some perhaps even "superior." We are well aware of the difficulties and dangers in doing so, but they are no different in kind from those inherent in any decision making by peer groups in, for example, the award of research grants by Federal agencies. The principle of peer evaluation is widely accepted for many decisions; why not for journals? The advantages would be many, one being that editors of journals not on the approved list would surely make strenuous efforts to upgrade their publication until it was admitted to the select group. Implicit in this recommendation is the idea that the approved list would be continually reviewed, and evaluated annually. Further, the criteria for selection would be published for criticism and possible amendment before being first applied.

Besides selection, the editorial process serves to improve accepted manuscripts before they are published. This improvement falls into two categories: additional experimentation suggested by the reviewers, and greater effectiveness of communication, usually suggested by the editor. As for the first, one obviously cannot improve the quality of the results, or facts, but one can sometimes improve the quality of the information by requiring a confirmatory experiment. This aspect is at present rather well taken care of for the better journals by conscientious reviewers. On the side of greater effectiveness of communication, however, much needs to be done to enable editors to help authors in achieving more logical organization, increased precision and clarity of writing, better arrangement of tables and display data, correct nomenclature and usage of technical terms and internal consistency. Many deficiencies are attributable to the fact that most editors of biological journals are volunteers and are inadaquately assisted. When the overall efficiency of the information system is considered, failure to provide adequate assistance to editors, both in staff and funds, is a glaring example of false economy. Some recommendations for the education and training of authors and editors to this end are given below.

Let us consider first, however, what effect the editorial process has on the speed of information transfer. By comparison with the informal process, formal publication is almost intolerably slow, an average of seven months for biological journals between receipt of manuscript and publication (Stewart, 1964). But we must remember the difference in quality of the information transferred. The published material is archival; it is permanent and must be complete, available for validation, and repeatable in detail when the author, for geographic or temporal reasons, is not available for questioning. Some of the time required for the editorial process is time exceedingly well spent. That it could be reduced is indubitable; the difficulties in the way of reduction are mainly financial and could be alleviated by larger, better trained editorial staffs, liberal use of the telephone to speed review, and more paid editors. There will always remain an irreducible minimum of time that is necessary for careful consideration by reviewers, who ideally consist of scientists in the same field and whose primary activity is not the review of others' work but the pursuit of their own research.

Increased efficiency in the production of journals once the editorial process is completed is a different matter. Computerized processing and typesetting offer some improvement in the near future, but again there is an irreducible time factor to be recognized; authors' examination of proof and the attendant mailing time--which last seems to be increasing! The daily production of newspapers, so often cited in contrast to the three-month production of a journal, depends on the presence of a resident staff of writers, editors, and proofreaders who take complete responsibility for the product.

The Need for Education and Training of Authors and Editors

Scientists as a class are notoriously poor writers. We suggest that this reflects a deficiency in their education, and that the remedy lies not in an increased number of English courses but in specific instruction in scientific writing, at the time when they are learning all the other necessary techniques for scientific research--in graduate school (Wilson, 1969; Woodford 1967, 1968, 1969). The technique of writing a journal article, in particular, is intimately bound up with the principles of experimental design and of scientific method, not merely of exposition, and deserves a place in the very center of a scientist's education. Regrettably, only a few graduate schools in the United States offer a course in scientific writing, let alone require it. Other universities apparently rely on individual guidance by the graduate students' supervisors, who have themselves seldom received any instruction in scientific writing. The "technical language that is at once graceful, easy to write, and easy to comprehend" that Weinberg (1967) yearns for does exist, but it is only just beginning to be taught, and only a minute percentage of today's research scientists can wield it with anything like mastery.

Editors are traditionally the guardians of the printed word, yet they may receive less training for their prestigious posts than does an unskilled laborer. They are usually appointed or elected in recognition of their preeminence in research, sometimes coupled with skill in exposition. They accept the appointment with reluctance, as a professional obligation, and are appalled to discover how much more there is to editing a journal than editing a series of manuscripts. Quite apart from such managerial considerations as selecting and prodding reviewers, there are matters of layout, format and type size; record-keeping and statistics; recommendations and conventions for nomenclature and abbreviations; characteristics of titles, abstracts, key-words and indexes from the point of view of information retrieval; close cooperation with abstracting and indexing services, and so on. The editor does his best in the too small fraction of his time that he can allot, concentrates on certain aspects of the job, and lets the rest go. Herein lies an important clue to the ragged diversity, designated only by optimists as rugged individualism, in our biological journals.

The Council of Biology Editors, formed in 1956, has brought some degree of order into the chaos by publishing a Style Manual (Porter, 1964) in which important standards in editorial practice are laid down and by bringing editors together for an annual conference at which experienced editors can help to educate the neophytes. What editors need most, however, is well-trained assistants of high caliber who will be paid to devote the necessary time to studying all aspects of the journal that have a bearing not only on recording but also on the storage and retrieval of information. These assistants would restore the efficiency of the journal as an instrument in the total information transfer process. They will need to have had advanced scientific training, including research experience, onto which they will graft a knowledge of information processing through further study and training.

Costs of Improved Quality

Not all the innovations we recommend are costly. Paid editorial manpower at salary levels corresponding to those of responsible academic positions constitutes a new expense that publishers--be they learned societies, university presses, or commercial firms--will have to assume and pass on to the subscribers or authors in the form of increased subscription rates or page charges. Because these costs ultimately come, in most cases, out of research grant funds, the mechanisms already exist for the indirect governmental subsidy of journals that the Weinberg report (Weinberg, 1963) predicted would be necessary.

Training in scientific writing is an added expense and a responsibility that universities should take. Some faculty time must be invested in teaching a way of thinking that will be serviceable for a lifetime.

Editorial practices that will ensure high quality without delaying publication have been mentioned. They are regarded in conservative editorial offices as inordinately expensive but are already in operation in the larger and more progressive journals. They include the addition of staff and the efficient use of the telephone and telegraph. The investment is minuscule when viewed from the standpoint of the total information system, but we realize that it cannot be made without radical changes in attitudes as to the relative importance of the different components of that system.

New Forms of Primary Publication

Some experiments are being conducted with primary publication, each aimed at solving one of the many problems that beset it. One problem is that in certain fields of biology long papers cannot be avoided; long lists of names of species or of geographic locations form an intrinsic part of the data. Although these lists are of lasting value, they may be of immediate interest only to a few; it seems wasteful to print and distribute a conventional journal solely in order to give them permanence. A depository of this kind is available at the American Society for Information Science. One journal, Wildlife Disease, has adopted publication in microfiche exclusively. Other journals issue a microcard, microfilm, or microfiche edition in addition to the conventional form in order to relieve the pressure on libraries' limited storage space. The recent development of greatly improved equipment for reading microfiche and obtaining normal-sized copies--macroprints or "hard copy"--of selected pages will doubtless lead to wider acceptance of this form of publication.

A second problem is that individual subscribers to journals of wide scope find that the fraction of material in each issue that interests them gets continually smaller, and they begrudge the shelf space and mailing costs for the remainder. The suggestion has been made several times, most recently by Brown et al. (1967), that journals continue to produce bound issues for libraries as at present but, in addition, to make every article available as a separate, either to be sold on demand or to be dispatched to individual subscribers in sets determined by previously specified profiles of interest. Custom tailoring has always been more expensive than mass production, but it is possible that automation of the selection, packaging, and mailing of separates from a large journal might be economically feasible. Alternatively, a large number of journals produced by a single printer might combine in order to provide the numbers necessary to justify the initial cost of automation. This scheme has not yet been tried in actual practice.

The new publication, Communications in Behavioral Biology, from Academic Press, attacks several problems simultaneously. The abstracts and the index of the journal are sold separately from the journal; this procedure allows the scanner of the abstracts to order the separate articles. The pages of abstracts are made more attractive by the inclusion of prepublication abstracts from many other co-operating journals, from which separates are not, however, available. The publication is set by computer, and from the tapes several other types of information services can be generated. The publication contains many features that will undoubtedly be further developed later on. It combines with primary publication some characteristics of secondary services, and it uses computer technology.

Future Forms of, or Alternatives to, Primary Publication

The rapid development of computer technology in recent years, especially in the on-line manipulation of large bodies of data stored in a computer memory that is accessible to many users at different consoles, has naturally stimulated thoughtful observers to speculate whether conventional publication is obsolescent. When each research department and perhaps each scientist has a desk computer console, connected to a central store of all the world's information, will anyone need to go through the time consuming business of printing a cumbersome ink-on-paper journal, distributing it, cataloging and shelving it, and above all carrying or mailing it around every time its contents are to be consulted?

There is no question but that the electrical transmission of information is quicker than are present means of its transportation. It will, however, be more expensive, at least for the foreseeable future. Computer networks being planned at Project MAC and EDUCOM will take perhaps ten years to construct in such a way as to place one to two terminals in all major U.S. research institutions, another ten to provide a terminal for each small group of scientists, and another ten to link with other continents. For three decades at least the printed journal seems likely to remain an important vehicle of communication.

Storage and retrieval are relevant to all disciplines, and we can probably look to a time when an author will compose his "manuscript" directly on computer tape and transmit it, or rather its contents, over the network to the editor, who will transmit it similarly to the reviewers. When he receives the reviewers' comments, the author will revise his text, using a computer program to make insertions and deletions. When the revised version is approved, the editor will "publish" it instantaneously by releasing it to the network with a notation signaling its approval by a specific review group. The network will automatically feed the work to its own "secondary services" where it will be indexed and classified, perhaps in hundreds of different ways for different users, for alerting and retrieval purposes.

This postulated system, which is largely borrowed from Licklider (1965), has one aspect that we wish to emphasize: the continuing role of editor and reviewers. Indiscriminate release of information to a computer network would probably be even more disastrous than is indiscriminate publishing today. Although in a system like this journals of the present form may disappear, their equivalent will remain; the Computer Network Biological Chemistry Review Committee, for example, will replace the editorial board of the Journal of Biological Chemistry, the Mammalogy Review Committee that of the Journal of Mammalogy, and so on. Release of unapproved material will probably also be allowed, but with the label "Unapproved." This will correspond to informal transfer of information, and the material will automatically be purged from the system either after a fixed time or when newer material displaces it from the memory space.

If the real system of the future resembles at all the one described, economic factors will almost certainly require more rigorous screening of the input than in the present system, which means the application and maintenance of even higher standards. It is imperative, therefore, for the future as well as the present, to concentrate on improving both our writing and our editing and on defining and making explicit our criteria of quality in all aspects of primary publication.

REVIEW ARTICLES

As the volume of primary publications increases and biological disciplines continue to become more and more splintered, there is growing demand for condensation of the literature in a given area for the convenience and enlightenment of nonspecialists in other fields. Preparation of critical reviews will carry greater and greater prestige as the need for them increases.

The review differs from an abstract or a group of abstracts in being a systhesis of papers, not just an aggregation. At their worst, some reviews present the gist of each individual paper in such undigested form that the review is little better than the corresponding sections of an abstracting journal. At their best--and as an ideal to be encouraged--comprehensive critical reviews may contain much new information contributed by the author, who may present such brilliant new interpretations and ideas that the review itself becomes a primary contribution to research. See, for example, Lwoff's review on lysogeny (1953).

The _Quarterly Review of Biology_, _Physiological Reviews_, and related publications have been filling an urgent need for many years. In addition to occasional summaries and comprehensive reviews that appear in the standard biological journals, several special serials have been created to fill the information gap. The series of "Annual Reviews of," "Advances in," and "Yearbooks of" this or that subject are excellent tools for the biologist (for a list of these see Bottle and Syatt, 1967). In addition, many organizations have put out Festschrift volumes to celebrate anniversaries; examples are _Fifty Years of Botany_ (Steere, 1958) celebrating the fiftieth anniversary of the Botanical Society of America, _Fifty Years of Plant Physiology_ (Weevers, 1949), _A Century of Progress in the Natural Sciences, 1853-1953_, published by the California Academy of Sciences (1955), _The Scientific Endeavor_, to celebrate the Centennial of the National Academy of Sciences (Revelle _et al_., 1965), and many others. A bibliography of these "one shot" works would be useful indeed.

The characteristic of "one shotness" leads to the criticism that reviews, like textbooks, are out of date before they are disseminated. Automation of information creates the technically feasible possibility of producing a review that may be stored on tape, continuously updated and reproduced either in full text or in part as needed to answer specific inquiry. The preparation, updating, and dissemination could be a function of a specialized information center. The Brain Information Service, University of California at Los Angeles, has recently announced the production of the first such review (Amacher *et al*., 1969).

Inducements to Authors of Review Articles

It is easy to understand why most active scientists are reluctant to use their time and energies in the production of comprehensive surveys and reviews. The work seems repetitious to anyone who has been in the forefront of the field, and while he is writing, he is losing ground on his basic research program. The inducement of financial reward sufficient to allow the author to take sabbatical leave has been tried, so far without great success--applicants for governmental and private awards of this kind are few. Provision of a huge number of reprints seems to constitute more effective persuasion, presumably by satisfying a more fundamental urge. Most effective of all, however; might be to provide the prospective author of a review article with all the documents that he will need and to relieve him of all such mechanical chores as checking and arranging the long list of references. Way (1968) has made a similar suggestion for aiding the compilers of handbooks and critical lists of tables. Exhaustive search of the literature and acquisition of the actual documents concerned can occupy at least half of the time necessary to write a review, and these tasks are frustrating and repugnant to an inventive mind. The effectiveness of providing this kind of service to the authors of reviews is now being explored by the Brain Information Service of the National Institute for Neurological Diseases and Blindness.

Computer Checking of References

If references to the scientific literature were cited in all journals in a standard form, their accuracy could be checked by computer. Such automation is urgently needed in many steps of primary publication. Now that an American standard for the abbreviation of titles of periodicals has been agreed upon (U.S.A. Standards Institute, 1966) and the corresponding abbreviations for biological serials have been published by Biological Abstracts (BIOSIS, 1969), agreement on a standard form for the complete reference should be in sight. We recommend that the American National Standards Institute should continue its sustained efforts in developing this standard. In spite of these cooperative efforts and attempts to bring uniformity and standardization into the handling of abbreviations of the titles of biological periodicals on a national as well as an interdisciplinary basis, we still find major new works appearing that offer their own individual systems for abbreviations, without reference to the American standard abbreviations (Lawrence _et al_., 1968).

A wholly new kind of machine-readable designations of journal titles has been devised by the Association for Testing and Materials in Philadelphia, which has established the concept of a five-letter unique code or CODEN for each serial publication: for example, the CODEN for _Bacteriological Reviews_ is BAREA. Thousands of CODENs already assigned have been published in reference volumes. This organization has also accepted the responsibility for encoding new periodicals as they appear. The utility of the CODEN in computer techniques is obvious, and it can effect economy of space even in standard printing procedures. The _Biological Abstracts_ 1968 and 1969 lists of serials (BIOSIS, 1969) show the CODEN assigned to each biological serial. A standard for 8-digit Standard Serial Numbers, which will allow the assignment of unique codes to serials for the entire foreseeable future, is now under consideration by the American National Standards Institute.

Computer Techniques for Handling the Names of Organisms

Sooner or later, every biologist has to concern himself with the correct name of at least one organism, if not many. The systematic biologist, who is most concerned with the names of organisms, knows the sources that will enable him to verify the spelling and other aspects of the names of plants or animals that he must use in his research. However, even the systematic biologist can be as helpless as the non-systematist if he has to find or verify names of organisms far outside his own area of specialization. There is a very real need by all biologists--as well as librarians and editors--for comprehensive lists or "telephone directories" covering the names of all organisms, for general reference purposes that go far beyond biology. Surprisingly, the existence of some lists of, or indexes to, groups of organisms seems to compound the problem instead of solving it, partly because published supplements may not be found easily and partly because of the almost immediate obsolescence of such a work in a rapidly developing field once it is printed.

All this points up the need for computerized handling of the names of organisms and for an up-to-date data bank containing such information as the names of all species, as well as of all higher groups, the authors, and the date and place of publication. With the open system made possible by electronic data processing and computer techniques, new information can be interpolated at will and corrections made in original entries when necessary, thus abolishing the need for supplements. Computer print-outs of any special group of organisms could then be made at the request of an individual or institution--and eventually major institutions would have direct access to the central data bank through individual computer stations or consoles.

The concept of bringing modern machine techniques to the ordering of names of plants and animals is by no means new; adequate "hardware" and programs are available or in the process of development (Crovello and MacDonald, 1969; Creighton and King, 1969; Suszynski, 1969).

The work of Sydney W. Gould (Gould, 1962; Gould and Noyce, 1965), supported by the National Science Foundation, has shown that this urgently needed but monumentally large task cannot be performed by any one person. Only a major institution can provide the necessary manpower, library resources, and continuity to develop and maintain a program of this dimension.

Currently, a major activity in American botany is the Flora North America, initiated in 1966 by the American Society of Plant Taxonomists, after appropriate feasibility studies. This international project was conceived and organized to produce a new synthesis of knowledge about the vascular plants of the United States and Canada during the next 15 years. It will depend largely on electronic data-processing and computer techniques (Shetler et al., 1969) to produce a taxonomically-structured data bank from which the information necessary for the actual published treatises can be derived. The feasibility and planning studies have been funded by the National Science Foundation.

TABULAR COMPILATION OF SCIENTIFIC DATA AND THEIR EVALUATION

In the physical sciences, tabulation and critical evaluation of data are of major importance. Compilations such as "International Critical Tables," Landolt-Börnsteins's "Tabellen," and the vast chemical handbooks of Beilstein and Gmelin, form an indispensable repository of data for physicists and chemists. Biologists frequently have occasion to use the information contained in such volumes. The value of this information for the user depends not only on a comprehensive search of the literature by the compilers for all available data, but also on a careful, critical evaluation of the relative accuracy and reliability of the data reported by different investigators. The importance of promoting such work, and coordinating it on an international basis, has recently been recognized by the formation of the ICSU Committee on Scientific Data (CODATA), which is made up of representatives of the International Scientific Unions and of the major countries that are contributing signicantly to work in this field. In furtherance of its mission, CODATA has recently published an "International Compendium of Numerical Data Projects," (1969) which attempts to list all data projects and important compilations of data throughout the world.

Hitherto biologists have for good reason taken relatively little part in activities of this sort, since quantitative data in biology have been relatively few, and in general not highly precise. However, the situation is rapidly changing, as indicated for example by publication of the "Biology Data Book" by FASEB (Altman and Dittmer, 1964) and of the second edition of the "Biochemishes Taschenbuch" (Rauen, 1964). The first edition of the "Handbook of Biochemistry" (Sober, 1968) includes comprehensive data on the physical and chemical properties of a large number of biologically important molecules. It also includes such other data of importance to biologists as genetic linkage maps for a variety of organisms, from Escherichia coli to the mouse, with related data for man. Another useful compilation is the "Atlas of Protein Sequence and Structure" (Dayhoff, 1969), which contains extensive information of increasing significance for evolutionary biology and taxonomy, as well as for biochemistry. It is interesting to note, as an indication of the rapid growth of the field, that the 1969 edition of this Atlas contains nearly twice as much information as the previous edition published only one year earlier.

As this brief outline indicates, the major impulse for promoting compilation and evaluation of scientific data in biology has come from the biochemists and biophysicists. The need for such compilations, prepared by experts who are qualified to evaluate critically the reliability of the data in the literature, will certainly continue to increase, for working biologists of every sort. As in the preparation of critical reviews, competent and dedicated workers will be needed to carry on this work, and they will deserve similar aids and inducements to those we have suggested should be offered to authors of reviews. The use of computers should aid the initial process of compiling the data, and also should permit the rapid feeding of new data into the information system, after they have been critically evaluated.

SECONDARY INFORMATION SERVICES

Introduction

Secondary information services consist of all media, techniques, and activities by which the scientists is made aware of and assisted in obtaining access to primary scientific information. The traditional and still standard components of secondary communication services are abstracts, indexes, and bibliographies, which continue to be valuable tools. But an increasing number of new activities and methodologies are appearing as part of our communication revolution. Among them are current awareness projects, specialized custom-tailored reference services, computerized indexes, automatic notification of the publication of articles of interest on the basis of previously established interest profiles, provision of photocopies of full text, question-answering services, and the preparation of summaries of groups of papers.

At first sight, the number and diversity of secondary services seems chaotic (Swanson, 1967):

> Access to subject and subject-control over recorded knowledge is now attempted largely through the aggregate of nearly two thousand indexing and abstracting services throughout the world. In no sense, however, do they function as a single system and it is certainly among the most important of requirements that at some future date subject interrogation should become possible in a single operation, with a request being guided by the system itself into the appropriate index where a search can be conducted. As matters stand now, a specific piece of information must be sought in literally many dozens of places before any reasonable assurance can be obtained that even an appreciable fraction of relevant material is actually located.

A more detailed analysis shows that the situation is somewhat less frightening. The actual figures in the following passage (Parkins, 1966) have undoubtedly changed since it was written, but the general conclusions are the same:

> At present there are about 425 abstracting and indexing services in the United States that provide our scientists and technologists with approximately 2,000,000 abstracts and title listings annually dealing with virtually all fields of science and technology. Although this is a large number of abstracting and indexing services, closer examination shows that 214 or 50% are merely abstract sections of primary research-reporting journals, whose average coverage each year is relatively small.
>
> Actually . . . 9 abstracting services, or only 2.4% of the total number, provide 50% of the abstracts published per year; 247 abstract services, or 65%, publish only 5% of the total published. Similarly, 4 title-listing services, or only 7% of the total of such services, publish 50% of the titles listed.

In what follows, we examine the advantages and disadvantages of the diversity of scope and approach among the many secondary services in biology, in terms of the functions that these organizations serve.

Functions and Desirable Characteristics of a Secondary Service

Secondary information services in biology, as in other disciplines, have three major functions:

- Abstracting and indexing--to provide abstracts that will, if necessary, substitute for original documents and to classify them conveniently.

- Storage and retrieval--to store references to publications in a form suitable for rapid retrieval of specific information or to allow browsing in a defined field of interest.

• Alerting or current-awareness--to announce the appearance of a document of potential value to an individual or group with special interests.

Until the advent of computers these functions were fulfilled almost exclusively by the regular issue of printed collections of abstracts, or of titles without abstracts, indexes to the collections being separately provided at a later date. Both abstracting and indexing required intellectual effort. Storage and retrieval consisted merely of shelving the volumes of abstracts and then gaining access to the information they contained via the annual and cumulative indexes. The alerting function was served as the subscriber browsed through the abstracts in each issue under the broad subject headings that interested him most.

The major biological publication of this type is Biological Abstracts--Excerpta Medica, and International Abstracts of Biological Sciences supplement it to some extent. Compilations consisting of indexed listings of titles only, without abstracts, characterize the Zoological Record, the Bibliography of Agriculture, and Index Medicus. Chemical Abstracts also abstracts a large number of biological journals, and its coverage in biochemistry and basic medical sciences overlaps that of Biological Abstracts to a considerable degree. Not long ago, scientists would customarily subscribe to one of these publications and scan or read each issue on its arrival to discover news in their fields of interest. With the rapid growth of science the publications have become too expensive for individual subscription, and with the increased pace of science the "news" tends to be stale if it arrives 4 months to a year after publication of the abstracted document. For some kinds of retrospective search, indexed abstracts can still be more effective than computer tapes, but in general the new technology is displacing traditional methods.

Whether all the changes in the secondary services brought on by computers are in fact improvements can be profitably examined only if we first ask "What do biologists today require? The following summary statements are compiled from the answers to questionnaires sent by Biological Abstracts during the past few years (Parkins, 1966, 1969).

● Timeliness. Computers, with their ability to alphabetize, sort, arrange, and rearrange verbal material rapidly and accurately, can help greatly to make abstracts and indexes more up to date and to restore an alerting function to secondary services. Yet equally important savings in time could be made by innovative publishing methods that provide for early transfer of information from editors of primary publications to the secondary service. Research and experimentation with such methods should be vigorously pursued.

● Ease of Access. The information provided by the secondary service should be close at hand. Until there is a computer link with the information system in every library, printed books will continue to fulfill this requirement better than a computer-based system.

● Flexibility. Every biologist thinks he desires access only to the information directly pertaining to his interests; but because these interests inevitably change during the course of his career he actually needs a comprehensive source of information large enough to supply his varying needs and to be readily tapped in many different ways.

● Links with Other Disciplines. Even a source of information embracing all of biology would not necessarily satisfy all the requirements of biologists. To obtain information on physicochemical or mathematical methods employed in their research, for example, they will need access, preferably through the same access point, to portions of the information system relevant to other disciplines.

● Freedom in Phrasing Questions. The biologist wants to be able to pose his inquiries in familiar terms--natural language; he should not be required to conform to fixed or unnatural terms or codes set by the system, or to "speak" the machine language.

● Indications of Reliability. Readers judge whether published work is likely to be correct by applying several criteria that include the author's reputation, the standards of the journal or publisher, and internal evidence from the document itself. Surprisingly, many users expressed the desire that the secondary service also pass judgment on the material abstracted. Such action is at present contrary to the practice of all secondary services. Perhaps the majority of biologists would prefer that an information service include only works that reach a certain level of scientific merit. This possibility should be explored.

Together, these requirements suggest a total secondary system that features specialized, narrow-scope abstracting and/or information services for well-defined groups of users--the research personnel in a single company, a single subdiscipline or a scientific mission such as air or water pollution--should continue. Specialized expertise is extremely valuable in preparing information for other specialists. These specialized information centers should, however, feed into generalized, wide-sweep services such as the BioSciences Information Service of Biological Abstracts (BIOSIS) to a much greater extent than at present. The larger services should accept the specialists' contribution and incorporate it without modification into their own information bases, but in addition they should examine the same material from the general point of view and meet the requirement of variable scope and depth. These larger services must also adapt themselves to being part of a still larger system embracing chemistry, physics, and mathematics, and must solve the problems of compatibility, technological and conceptual, with the systems developed in other disciplines. Some of these problems can be avoided altogether by careful planning, and the emerging cooperation between the two giants among secondary services, Biological Abstracts and Chemical Abstracts, is gratifying. To achieve full freedom in phrasing questions may necessitate interposing an information specialist between the scientist and the secondary information system, to translate his question into terms the system can use. Even should this be done it will be imperative for the scientist to understand how the system works and what it can do, as well as to develop insight into what he actually wants from it. Otherwise he will find himself in a situation analogous to that in which an experimenter depends too heavily upon a technician, who does not understand the purpose of the investigation, merely because the technician knows how to use the equipment. If that happens, the scientist will not only use the system inefficiently, he will provide no useful feedback to the designers of new systems, which will therefore unknowingly and by default develop in undesired directions.

Inadequacies in the present secondary services are largely due to the failure of biologists in the past to express their requirements.

BioSciences Information Service of Biological Abstracts is the world's largest abstracting and indexing organization serving the biosciences. Biological Abstracts, its major publication, is issued twice monthly. Abstracts of more than 135,000 research papers, organized into approximately 500 different subject specialities, are published annually. BioResearch Index, published monthly, was introduced in 1965 to provide access to additional research reports that could not be presented in the traditional manner in Biological Abstracts. It provides a list of journals reported in each issue and a complete bibliographic reference for each paper indexed from these journals (BIOSIS, 1969). A detailed description of the present activities of BIOSIS, including many important innovations, has been published by Parkins (1966,1969) and has been reviewed by Gray, Werdel, and Weyl (1967) and so will not be repeated here.

Tools of the Secondary Services

Abstracts

An abstract is a tool that enables the user to cope with a great volume of world literature, to become aware of potentially valuable literature, and to decide whether or not he wants to read the abstracted document in full. He may even use the abstract as the actual source of information if the parent document is difficult to obtain or is in a language he cannot read. In many developing countries whose national and university libraries cannot yet afford to acquire the several thousand journals reporting original research in the life sciences, Biological Abstracts play an unanticipated and unplanned role in replacing them, and, in fact, is itself regarded as a primary publication!

When the chief function of secondary services was the provision of abstracts, most of the services employed experts in each field of inquiry to write them. These professional or semiprofessional abstractors were scattered widely, and in the interests of timeliness the larger abstracting services have decreased the use they make of professional abstracters and now often accept without modification the abstract provided by the author of the article. Editors of most leading U.S. biological journals now require that "heading abstracts" be provided by the author and print them at the beginning of his article. Placing them at this point impresses upon the author that the abstract must be intelligible quite apart from the article itself.

Of the world's biological journals, however, only about 30 percent provide heading abstracts, and of these not all are sufficiently informative to be used without augmentation. Articles in the remainder suffer unnecessary delay before they are reported in abstract form. Even in those journals where the average standard of the authors' abstracts is acceptable, many abstracts omit important information reported in the paper. Editors and reviewers should take responsibility for the adequacy of abstracts, yet in the experience of many editors the two parts of a submitted paper that almost invariably escape critical comment by reviewers are the abstract and the title--the very components that are most crucial for information retrieval over the long-term future.

We urgently recommend that:

- Editors require a succinct but fuller informative abstract from the author, to be printed at the head of each article;

- Editors furnish the abstracts to interested secondary services at the earliest possible stage of publication--for instance, by sending page proofs;

- Editors and reviewers pay particular attention to the adequacy of authors' abstracts, specifically in considering their usefulness when divorced from the text;

- Authors become aware of the immense importance of their abstracts in information retrieval, and write them not only for readers of the journal but for searchers of the literature;

- Teachers of scientific writing devote adequate time to teaching the requirements and the art of writing abstracts that will adequately describe the contents of the document abstracted.

Some may argue that bypassing the professional abstractor in order to increase publication speed incurs too much of a sacrifice of quality, that the author's necessarily more limited experience must lead to an inferior product. We believe that his more intimate knowledge of the subject partially compensates for this, and that the reviewers and editor should contribute the vision and greater publication experience necessary for a broadly useful abstract. This provides another strong reason for more professionally trained editors.

Titles

The day of the whimsical but meaningless title is over, and only the sentimental egotist will regret it. Computers, with their capacity for rapid permuted-title indexing (Parkins, 1963) have brought the importance of precise wording of the title into prominence, but even before their advent the title had long been used for alerting purposes--Current Chemical Papers, Chemical Titles, Current Contents. It is surprising that many authors are still so unskilled in devising clear, informative titles and reviewers so blind to their deficiencies.

As with abstracts, the situation could be immediately improved if authors, reviewers, and editors would school themselves to consider this component of an article in terms not of the immediate reader but of the searcher, for whom this short sequence of words may well be the sole signal that leads him to the article.

Products of the Secondary Services

Indexes

Classification and indexing are the key activities of secondary services. Opinions differ on whether articles should be indexed by professional indexers or the title supplied by the author taken as the starting point for computer-aided indexing. In the production of permuted-title indexes, the computer is programmed to drop insignificant words from the title and to provide an index entry for the remainder. The index can be arranged either as a KWIC (Key Word in Context) index or as a KWOC (Key Word Out of Context) index.

Biological Abstracts has adopted computerized indexing, with the result that a KWIC index (B.A.S.I.C.) is distributed with each semi-monthly issue of the publication. The speed of production of the index constitutes one of its most appealing features. The subtlety and refinement of the conventional index is lost, but one compensating advantage is that the index is not tied to a preconceived hierarchy of terms or a printed thesaurus; it evolves naturally with time as the technical language itself evolves. Some limitations of the technique are readily apparent from our earlier discussion of the inadequacies of authors, reviewers, and editors in producing the titles that form the basis of the index. Even if these producers of titles are alert and skillful there can be a further difficulty. Within a title of reasonable length it may be impossible to refer to all subjects in the article that are worthy of indexing. In recent years, Biological Abstracts has also developed several other kinds of indexes (BIOSIS, 1969) so that in each issue the reader has access to his subject of interest not only by a permuted title index (B.A.S.I.C.) but also an author index, a biosystematic index, and an index to subject specialities (CROSS). However, deficiencies in titles can readily be remedied by reinforcement--"management" or "enrichment"--of the title by adding supplementary key words or indexing terms before the permutation of the title begins (Parkins,1963). At present this is usually done by staff at Biological Abstracts, who select additional key words from the abstract and the original article, but there is no reason why the editor or author should not add the amplifying key words at the time of publication. Indeed, several biological journals, following the example set by engineering journals, are now inserting such indexing aids. The key words are usually printed in the journal under the heading abstract as an aid to the reader and guide to the indexer. We recommend that these practices become general.

Only time and experience will tell whether the choice of permuted-title indexing over that requiring extensive and time consuming intellectual input has been wise. In conventional indexing the relationships between the terms of the entries can be made more precise, and in general the process seems less crude. On the other hand, it must be recognized that no matter how skilled and experienced an indexer may be his final product is necessarily limited by the sum total of his educational and technical experience. There is at least a chance that the combination of an author's special knowledge and an editor's experience and training will match the special talents of the indexer. From the user's point of view, the elegant hierarchical index, especially if it is compendious, may be difficult to use unless the user actually has an indexer's training; whereas the natural flow of language in the original title is readily comprehended.

Two major generalized indexing services, Index Medicus (National Library of Medicine) and the Bibliography of Agriculture (National Agricultural Library), retain traditional indexing methods, as does Chemical Abstracts, and assign indexing terms to each document independently of the title. In some titles, regrettably, not one of the words will provide an indexing term. The terms that can be assigned are predetermined--for Index Medicus, they are listed as Medical Subject Headings, MeSH--, i.e., a "controlled vocabulary" is used, and the citations are listed in the monthly issues of Index Medicus under relatively unchanging headings and subheadings. The list of terms is updated annually, but in the opinion of many the list is not extensive enough and changes too slowly with time. Nevertheless, the printed index is one of the clearest and easiest of all to use, and has the further advantage that the scope of documents indexed is far from narrowly medical. Indeed, a search for citations containing chemical information may in some instances actually be easier via Index Medicus than via the index to Chemical Abstracts.

That the philosophies of indexing adopted by the major generalized services differ so much from one another is actually of advantage to the user who understands the systems--he has that many more alternative pathways to explore if his first approach proves fruitless. The moral here is not that hundreds of secondary services should be encouraged to maintain their individuality and incompatibility, but that the overall information system that is eventually developed should allow the user as much diversity in asking questions as possible. Efficient switching from one mode of inquiry to another is, of course, essential; whether this will be provided by human intermediaries or automatically by complex computer programs remains to be seen. The several different kinds of indexes developed by Biological Abstracts provide an excellent example of the diversification of access pathways needed (BIOSIS 1969).

A new mode of inquiry and a new method of indexing have been provided recently by the introduction of Science Citation Index (Garfield,1964). This indexing method is independent of language--that of the title, of an indexer, or of the country where the document was produced--and depends on the principle of searching for recent work on a given subject by looking for papers that have cited an important earlier article on that subject. The method is easier to apply across the boundaries of disciplines than are methods that depend on indexing terms, since specialists in one field are often unaware of the terms that will attract the attention of their colleagues in other disciplines and of the different connotations that each group may attach to a given term. The different meanings of "plasma," "transformations," "nucleus," and "circulation" to a physicist and to a clinician exemplify the interdisciplinary indexing problems created by vagaries of language.

Grouped Abstracts for Browsing

Besides providing an index, abstracting services supply abstracts, and usually group them in such a way that abstracts on related topics fall close together. Many years ago, subscribers to Biological Abstracts would read through whole sections of such grouped abstracts to gain an overview of progress in a subfield; few have time to do this now that each section is so lengthy. However, the specialized indexes are now helpful for this purpose (BIOSIS,1969).

If in the future seekers after knowledge rely exclusively on computerized searches for closely defined topics, sectional grouping will no longer serve a useful purpose and many be discontinued. There will then be a danger that the searcher will become too narrow in his interests, and will no longer benefit from those lucky accidents of happening upon information of great value to him, although he was unaware of needing it. This danger has been much exaggerated. At least for the foreseeable future, indexing and search techniques that depend on language with all its synonyms and ambiguities will continue to produce enough material of marginal relevance to the question asked to satisfy the most avid browser.

Selective Dissemination of Information

In principle, computerized secondary abstracting services are ideally set up to provide a different kind of service from that possible before computers were used, i.e., the dissemination of selected abstracts in accordance with a known "profile of interest," rather than distribution of all the abstracts accompanied by an index. In practice, the cost of separate packaging and the difficulty of accurate definition of interest profiles raise formidable problems that are currently under study, although this problem will surely eventually yield to computer control.

A major step forward in "repackaging" abstracts for selective dissemination was made by BioSciences Information Service of Biological Abstracts in 1966 when it inaugurated the publication Abstracts of Mycology, originally as an experiment supported by the National Science Foundation (BIOSIS,1967). Now in its third year, Abstracts of Mycology has proven to be economically viable, so that BIOSIS is presently considering other important yet restricted areas of the biological sciences for comparable publications. To date, no other major secondary service has followed this promising model.

Selective distribution of references only--titles plus bibliographic details--poses fewer problems, and computerized alerting services based both on key-word analysis of titles and on the citation-index principle are in operation (Garfield,1964), notably by the Institute for Scientific Information (I.S.I.).

Retrospective Searches on Request

Once information has been placed on computer tapes for the production of sets of abstracts or indexes, the tapes can be used later for searching the literature retrospectively, usually with great savings of time compared to manual searching. The National Library of Medicine conducts such retrospective searches through its computerized files, which go back to mid-1963. The search formulation, prepared by an information specialist on the basis of the searcher's written request, is geared to the MeSH indexing terms assigned when the document was first indexed. BIOSIS has also developed a retrospective search by means of its computerized back files that provides not merely titles but abstracts of the documents retrieved by the search; the service has been designed in accordance with responses to experimental and pilot services of different kinds submitted to various groups of users (Jacobus et al., 1966; Schultz, 1968). We regard the introduction of soundly designed search facilities like this by all major secondary services as of paramount importance.

One challenging intellectual problem for retrospective searching stems from the fact that technical terms and concepts evolve in time; searching the literature after 30 years is an entirely different proposition from searching after three years and requires considerable historical knowledge and judgment. For the moment, magnetic tape files do not go back more than a few years and the problem does not arise, but it will have to be faced eventually. This conceptual problem should be acknowledged by those who criticize the slow progress in computerizing back files and who imagine that a mere mechanical extension of present work is all that is involved. Both private and governmental services shrink from the great expense of the task on several other grounds, among them the realization that much of the older published work is obsolete or superseded. On the other hand, the application of modern search procedures to the older literature could bring much valid but forgotten work to light.

In the foregoing discussion of the products of secondary services, two tacit assumptions have been made for the sake of simplicity. Neither are wholly true. The first is that the only services worth considering in depth are those provided by large scale organizations. Quite to the contrary, specialized secondary services have important roles to fill, and are discussed in the following section.

The second assumption is that the only documents of interest to the user of secondary services are journal articles or technical reports related to a single investigation. Again, this is hardly true for the specialist researcher, much less the policy-maker, administrator, practitioner, or teacher. A notable weakness in the computerized indexing systems is the shallowness of indexing of review articles and monographs. In response to this criticism it can be argued that these documents represent condensations of the literature and that the serious searcher is expected to read them thoroughly before turning to the computer search facilities, and that extensive indexing is, therefore, not necessary. It seems to us, on the contrary, that reviews are so important that every subject they deal with--and this is often far more than the title can indicate--should constitute an access point for current awareness or retrospective search. For monographs, at least the subject of each chapter should constitute an access point for current awareness or retrospective search. We urge reexamination of the indexing philosophy here.

Specialized Information Centers

Many researchers need an analysis of the literature in greater depth than is provided by the large generalized information services. For some fields of biology, particularly the taxonomic sciences, coverage of the major primary journals that may yield, say, 85 percent of the significant information is not enough.

The growth of literature and, particularly for biology, the dissemination of papers relevant to a small field in diverse broad spectrum journals has led Alvin Weinberg (1967) to state:

> A division of labor between those who create or discover the facts and those who sift, absorb, and correlate the facts seems to be inevitable . . . The scientists who collect and sift the facts would in our modern terminology make up a "specialized information center," that is, a group of scientists who, in a narrow, well-defined field of science, collect, sift and interpret information for other workers in the field. The key word here is <u>interpret</u>: the scientists who man such centers not only handle the relevant documents, but they also glean the scientific gems from the documents and try to make sense out of them. The people who run the center must be recognized scientific leaders in their fields of science, for in the process of examining the data and finding new correlations, they create new science.

A specialized information center provides, in comparison to the generalized information service, (1) access to the literature of a subset of biology, with deliberate effort to include the more expensive "last 15 percent" of the information, and (2) indexing to greater depth, made possible by the presence of active subject-oriented scientists. It should also, as Weinberg suggests, play a role in synthesis--the generation of critical reviews and perhaps standard reference data such as is now being attempted at the National Bureau of Standards (1968).

The essential point in establishing a specialized information center is environment. Centers of this type will flourish only in contact with active research. Their usefulness to the community will be measured by the ability of those who process the literature to apply their knowledge of changing connotations in a given field, and thus provide the critical in-depth indexing and synthesis needed by their peers.

Inasmuch as the function of the specialized scientific society is principally to enhance the exchange of information, oral and written, in a relatively small area the establishment of specialized information centers is a natural extension of the society's role--witness, for example, the action of the Entomological Society of America (Foote,1967). Where a research institution has a particular focus or mission in a given area, the development of an information center--perhaps at first to sharpen the focus of its own scientists--should be recognized as the nucleus of a service to the biological community (Jacobus et al.,1966; Schultz, 1968).

But the danger of development in isolation, with attendant problems of inefficiency, unplanned duplication, and dissemination of effort, as is clearly the history of biology's primary journals, must not go unattended. Specialized information centers must be developed in active liaison and cooperation with existing generalized secondary services, so that mutual sharing of resources will permit the biologist to approach a problem not only in the context of his self-limited research area, with consequent limited answers, but also with full knowledge that where information in a tangential field exists, his needs may be served.

Problems for Secondary Services

Coverage

The question of what articles, from what journals, should be indexed and abstracted is one of the most difficult and fundamental of those any secondary service faces. Unless resources are unlimited, a selection from the total must be made. For the specialized information centers catering to a well defined clientele the choice may be relatively easy, but for any service with pretensions to serving the whole community the choice may be agonizing.

Nor is the choice entirely free. Great expense may be incurred in following up unfulfilled requests for copies of documents; postal and political difficulties in certain nations may preclude delivery of some issues of a journal; some documents may arrive years after publication. Other frustrating circumstances may combine to make consistent and rapid recording of all published documents impossible to achieve.

It has been estimated that the expense of acquiring and abstracting the more accessible 85 percent of the literature in biology is equalled by the expense of the same processes for the remaining 15 percent. Most scientists believe that the informational value of the second group is less than that of the first, and one can take the view that this information should be ignored. This simplistic solution contains two grave fallacies: first, that mere accessibility is a measure of ultimate value; and second, that without acquiring 100 percent of the documents theoretically available one can judge when he has reached the 85 percent level. And what is the magic of the number 85 percent? Should Biological Abstracts, for example, make no effort to acquire documents that do not find their way automatically--by the action of editors and publishers--into its mailbox? This attitude would clearly be absurd and, of course, is not BIOSIS policy.

In general, Biological Abstracts and such large scale indexing services as Index Medicus, cover the serial publications to an extent that satisfies their subscribers, and they are diligent in issuing annual lists of the periodicals covered. However, much more meaningful assessments of coverage could be made by the scientific community if there were available a complete list of biological periodicals against which these annual lists could be matched.

Lack of Feedback From the Community

More assessments of the performance of large scale services are needed, requiring a far greater involvement of working scientists in the information process than is evident at present. Probably the sheer size of the endeavor restrains those who give some thought to the problems from expressing their opinions, yet they should realize that their opinions are highly valued. The educational efforts recommended elsewhere in this report should lead to greater knowledgeability and active participation of biologists in the work of the secondary services.

Foreign Languages

The great majority, 70 percent, of the world's literature in basic and applied biological sciences is published in English, French, German (Conrad, 1965; Kennedy and Parkins, 1969), and translation of the two last-named languages into English presents little difficulty. The amount and value of the Russian literature has been much exaggerated; it constitutes only 10 percent of the total. Cover-to-cover translations of Russian periodicals have proved to be more expensive than seems warranted; abstracts in English coupled with the availability of full translations, if needed, on demand, are all that the vast majority of scientists need.

At the risk of sounding self-centered to the non-English speaking world, we state here our belief that English should be accepted as the second language most appropriate to a scientist, and therefore suggest that publishing an abstract in English with a paper in any other language is the most effective way of disseminating its message to the rest of the scientific community.

Form of Storing the Information

Some exploration is being made of the possibilities of storing in an information system not only titles with citations, supplementary indexing words, and abstracts, but even the full text of articles. The necessary increase in memory space that all this would demand is so colossal that the mere idea was considered ludicrous only a few years ago; today some specialized services in industry are actually operating such a system. All this brings us one step closer to the ultimate ideal: a system that will deliver to the inquirer not merely a reference to a document or a précis of it, but a copy of the document itself. Problems of storing not only words but tables, line drawings, and even half tones are being investigated and solved; the cost-benefit ratios have yet to be assessed.

LIBRARIES

If a scientist were to read through all the world's technical journals in his own field, he would have very little time left for research--see for example, Glass's estimate of reading ability (1962). To the difficulty of finding time to read all relevant documents is added the difficulty of obtaining them in the libraries, whose holdings have been increased sometimes to unmanageable proportions, with the aid of generous federal grants--about $60 million to 1900 colleges and universities in the last three years--that were, ironically, intended to facilitate the diffusion of knowledge.

Some biologists attempt to circumvent the problems associated with the formal use of libraries by attempting to keep abreast of current work via extensive reprint exchange. Useful departmental libraries may be built up this way, but as their holdings increase, so do the time consuming tasks of indexing and cataloging, until the members of the department turn back to the main library as less time consuming in the long run.

No library, however well endowed, can acquire more than a fraction of what is published. While the increasing diversification of research compels libraries to acquire ever greater quantities of materials, the world-wide acceleration in the production of these materials daily widens the gap between what a general research library might acquire, organize, and hold, and what it actually can do.

Although arrangements for sharing resources among libraries--such as cooperative acquisition programs, union catalogs, interlibrary lending, and photocopying services--do exist, they are at present limited and inefficient. In addition, individual library and comprehensive bibliographies are woefully incomplete and inadequate.

What are possible solutions to these seemingly overwhelming problems?

One is to improve the techniques for inexpensive reproduction and dissemination of materials to enhance local self-sufficiency. This approach seems to offer genuine possibilities of contributing not only to massive increases in the collections of libraries at much lower costs than at present, but also to the effective cataloging of collections, the disposition of lesser-used materials in the interest of efficient use of storage space, the preservation of deteriorating materials, and the improvement of union catalogs.

Improved technology and methodology can also facilitate the sharing of resources among libraries. Better bibliographic apparatus by which the existence of materials is made known, and better arrangements for securing copies after identification, are being studied.

However, the most useful of all developments for the purpose of sharing would be the assignment of specific responsibility for excellence in certain subjects to particular libraries, each of which would be obliged to acquire comprehensive holdings in the assigned subjects, organize and publish its bibliography, and render a nationwide, if not worldwide, lending or photocopy service. Examples of this approach are the national libraries for agriculture and medicine and a more highly specialized holding is the Brown University Library which is outstanding in mathematics. For many years several of the major university libraries in the Midwest have divided up among themselves on a geographical basis the responsibility for acquiring complete holdings of the proceedings of local academies and of the smaller national academies in Europe. With these precedents, biologists might well encourage some of the major academic libraries to concentrate on the subdivisions of the biological sciences, as for example: Tulane or Yale in zoology, Indiana or Notre Dame in environmental biology and others in botany, genetics, and animal behavior. The impending development of inter-university communication by computer networks would make possible a combination of the principle of local self sufficiency with an extension of the sharing principle.

LOOKING FORWARD

A national information network in science with interlocking public and private components is slowly evolving. This network will develop from individual systems, more or less compatible with each other, each representing an adaptive peak on a continuum involving major scientific areas or large scale organization operations. Such systems are in the process of development for chemistry, physics (Herschman, 1969), engineering and for the Federal sector, the latter embracing all types of operations. If biology is to take its proper place in the network, and if it is to meet the challenge of its increasing information problems, a national plan must be conceived and developed.

The pattern that seems to be emerging for the other disciplines, with the possible exception of engineering, relies heavily on the existence of a single coordinating organization--the American Institute of Physics, or the American Chemical Society. The diversity of the biological sciences, both in organizational structure and in subject matter, parallels that of the Federal system; no single "monolithic" biological organization yet exists.

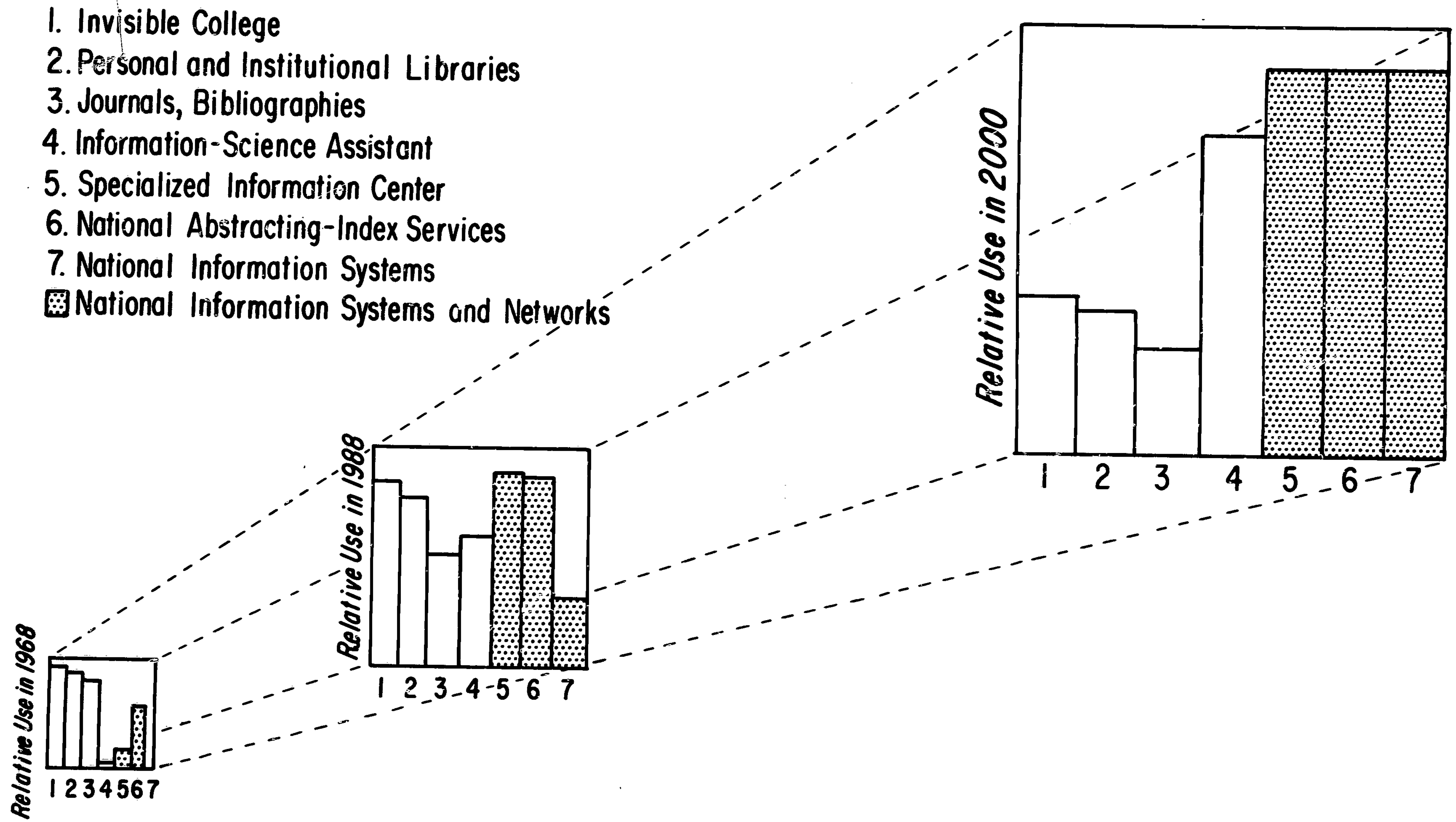
Unless one is willing to abandon existing institutions, the information system for the biological sciences must involve both public and private sectors, and must be compounded from the cooperation of several organizations (Gordon, 1969). Theoretically, it might be possible to assemble within a single organization the intellectual manpower needed to process scientific information, but the diversity of the biological sciences does not lend itself to this approach. In practice it seems best to follow existing historically developed lines, but to modify structure and operations to meet the new demands.

Thus for the biological sciences, the three existing generalized information services must function as central points of reference for the three major subsets of the community. For biology proper, the BioSciences Information Service of Biological Abstracts will serve the role of a generalized information center, and is the only major abstracting service in this field. In the applied fields, the National Agricultural Library will continue to be the central focus for agricultural users, and the National Library of Medicine will function as the central point for biomedicine.

Generalized information centers have a major role to play as switching points between specialized systems. In addition, the major center must assume the responsibility of serving the needs of the practitioner and the policy-maker or administrator, since no smaller center could provide the necessary interdisciplinary approach.

Particular needs of some researchers in biology will best be served by a series of smaller specialized information centers. These smaller units can harness the intellectual manpower required to evaluate data and information in a well defined area. These units must, however, interact closely with the generalized center so that inquiries addressed to either center can be appropriately switched according to the degree of specificity in the information desired. Prototypes of this type of cooperation, e. g., that between the American Museum of Natural History and BIOSIS, already exist in the biological sciences in the private sector (Schultz, 1968). The development of a toxicological information program within the National Library of Medicine is another example (Handler et al., 1966).

The organizational diversity of biology, so long considered a liability in forming a biological information system, becomes its most important asset when viewed against the above picture. Major societies, museums, or other institutions, currently serving a significant segment of the community, could develop a specialized information center providing coverage in depth and intellectual screening not achievable by a larger service. Where disciplines within the biological sciences lack structural organization, the appropriate generalized center could serve their needs.



Changing Patterns of Information Sources for the User

The role of computer networks such as EDUCOM in the information system of the future is obvious and need not be belabored here. We visualize the day when, in place of a telephone call for a single answer from a member of the "invisible college," the call will be made to the system to evoke a comprehensive answer to the query.

For this system to be developed adequately, the biological community must accept its responsibilities and coordinate its efforts. The information crisis confronts each biologist, be he neophyte or established "edge-cutter." Within his sphere of influence let him demand action of his society or institution.

Thus a forward look to the year 2000 reveals an inversion in existing practice and operations, and the addition of new elements, to cope with man's expansion of knowledge. We predict that journals, personal and local libraries, and the operation of the "invisible college" will decrease in importance compared to more highly organized services, some of which are now in their infancy. If we are correct, much thought and effort must go into planning those systems.

(See, in this connection, the diagram of changing patterns of information sources for the user, opposite page.)

RECOMMENDATIONS

The recommendations of the Weinberg panel, summarized on page 2, provide a framework for the following more detailed suggestions relating to communication of biological sciences information.

General

• Information systems of the future for biology should be designed not only for the research scholar but with the more general users of the system clearly in mind: the practitioner, the schoolteacher, the policy-makers and administrators, and all those concerned with educating the public about biological science--the science writer for newspapers and magazines and the writer of radio and television programs.

• The U.S. information system for the biological sciences, in the absence of any one monolithic information service, cannot be provided by a single governmental or private organization but should interconnect in a compatible manner the three existing major organizations in biology--the National Agricultural Library, the National Library of Medicine, and the Biosciences Information Service of Biological Abstracts--as well as several specialized information centers. These latter should use their particular expertise to provide specialized services, which should be developed in close conjunction with one of the three generalized services and feed into them. The generalized services should be alert to develop systems that are compatible with those in other scientific disciplines, with the ultimate aim of being able to switch questions in any field of science automatically to the appropriate data bank.

Informal Transfer of Information (See page 7)

• From the point of view of a system of recorded information, retrievable by computer or via printed documents, a clear demarcation between informal and formal transfer of information should be made. The proceedings of informal meetings should, in general, not be recorded or published.

● With rare exception, papers presented at formal meetings should not be published in Proceedings volumes; they should be submitted to journals for refereed publication according to established practice. Timely reporting of the essence of formal meetings can be achieved by means of conference reports in a journal that has a short publication time. A major exception to this recommendation concerns symposia highly organized around a central theme on a broad base, in which each of the presented papers constitutes an orally delivered review article, usually backed up with a carefully prepared manuscript.

● Listings of forthcoming meetings should be strengthened by giving them wider publicity and possibly subsidizing their distribution. Organizers of meetings should be encouraged and enabled to forward all information about forthcoming meetings to a central organization; the American Institute of Biological Sciences seems most appropriate.

● Technical reports should continue to be catalogued and distributed by the Clearinghouse for Federal Scientific and Technical Information and are somewhat reluctantly recognized by us as formal publications even though few are refereed.

● Informal exchange groups should be regarded as useful mechanisms for the transfer of information, provided they employ mechanisms to ensure that the information exchanged is truly informal and cannot be confused with formal publications.

● Newsletters should be regarded as useful, but also as ephemeral, unrefereed publications, and treated as such.

● Adequate provision for listing biological research in progress supported by predominantly Federal grants is available through the Science Information Exchange of the Smithsonian Institution. The Council on Biological Sciences Information does not agree with the President's Science Advisory Committee view that another "unpublished research information program" is desirable, at least in the present state of biological sciences.

● Indexes of book reviews should be included in the recorded information store for retrieval with the title and author of book.

Primary Publications (See page 17)

● Standards of quality in primary publications should be raised and rigorously enforced. This requires the immediate institution of programs of education and training for biologists--authors, editors, and referees.

● To be maximally effective, editors should be paid enough to allow them to devote a sizeable proportion of their time to editing, or should be provided with scientist assistants who are highly trained in editorial work.

● All graduate school training for biologists should include formal instruction in scientific communication, particularly in the writing of scientific articles, oral presentation of scientific work, and methods of searching the literature by traditional and modern techniques. The communications instructor should himself be a scientist.

● A definitive list of biological periodicals should be prepared at once, published and continuously updated.

● Associations of editors and other scientists should study the task of assigning biological journals to the categories "superior," "approved," and (tacitly) "not approved." Criteria for these designations should be published for consideration by the scientific community before application. The purposes of this categorization would be to encourage producers of journals to upgrade them into a higher category and to enable secondary services to provide approximate indicators of quality in their distributed listings.

Review Articles (See page 29)

- Inducements to potential authors of review articles should be increased, including the award of prizes and the provision of extensive assistance in the gathering of material and the more mechanical chores such as the ordering and checking of references.

- A standard form for citing references should be devised by the American National Standards Institute in cooperation with such international bodies as the ICSU Abstracting Board.

Secondary Information Services (See page 35)

- Biologists should be better informed about available services and participate intelligently as new services are designed.

- Editors of journals should place an informative abstract at the head of each article, and should furnish such abstracts to Biological Abstracts at the earliest possible time.

- Editors and reviewers should examine an author's abstract and title even more critically than other parts of the submitted article, because of their central importance in information retrieval.

- Authors should become aware of the immense importance of abstracts and titles in information retrieval, and consciously design them for searchers of the literature as well as for readers of the journal.

● Supplementary key words or indexing terms should be provided by the authors of articles and critically examined by reviewers and editors. Some journals already publish titles amplified by the addition of these key words. It is especially important to index new methods or new applications of old methods.

● The Biosciences Information Service of Biological Abstracts should continue to give high priority to its development of retrospective search services.

● Medical Subject Headings used in Index Medicus and in MEDLARS "demand" searches at the National Library of Medicine should be increased in number and revised more frequently, with the aid of appropriate scientific societies or widely representative members of the user community. Use should also be made of the Agricultural/Biological Vocabulary and its Supplement, published by the U.S. National Agricultural Library.

● Computerized information services should be capable of receiving inquiries in several modes. They should, if possible, utilize searching by title, by author's names, by index terms, and by citation.

● Review articles and monographs should be indexed in greater depth than they are at present to allow for easier and more dependable retrieval of these very important sources of consolidated information.

REFERENCES

Altman, P. L., and D. S. Dittmer, eds. 1964. Biology Data Book. Prepared under the auspices of the Committee on Biological Handbooks, Federation of American Societies for Experimental Biology, Washington, D.C.

Amacher, Peter, Victor E. Hall, and Allen C. Norton. 1969. Preventing obsolescence of scientific reviews: an updated-review project. Science 165:1029-1030.

Barr, K. P. 1967. Estimates of the number of currently available scientific and technical proceedings. J. Documentation 23:110-116.

Baum, Harry. 1966. A clearinghouse for scientific and technical meetings: Organizational and operational problems. Amer. Documentation 17:28-32.

Biochimica et Biophysica Acta. 1966. A statement of policy by the managing editors. Biochim. Biophys. Acta 121: 223-227.

BIOSIS. 1967. Abstracts of Mycology. Fungi in Biochemistry, Cytology, Genetics, Microbiology--Medical and Industrial, Pathology--Plant and Animal, and Systematics. (Vol. 3, 1969.) BioSciences Information Service of Biological Abstracts, Philadelphia, Pa.

BIOSIS. 1969. Biological Abstracts: 1968 List of Serials With Title Abbreviations. 88 p. BioSciences Information Service of Biological Abstracts, Philadelphia, Pa. (NOTE: Before 1968, the Annual List of Serials with Title Abbreviations was published in Biological Abstracts, instead of separately, and may generally be found in the last issue of each year.)

BIOSIS. 1969. Guide to the Indexes for Biological Abstracts and BioResearch Index. BioSciences Information Service of Biological Abstracts, Philadelphia, Pa.

Bottle, R. T., and H. V. Syatt. 1967. The Use of Biological Literature. Archon Books, Hamden, Conn. (Butterworths, London, England. 1966.)

Brown, W. S., J. R. Pierce, and J. F. Traub. 1967. The future of scientific journals. A computer-based system will enable a subscriber to receive a personalized stream of papers. Science 158:1153-1159.

Cairns, R. W. (Chairman). 1969. Scientific and Technical Communication. A Pressing National Problem and Recommendations for Its Solution. A Report by the Committee on Scientific and Technical Communication (SATCOM) of the National Academy of Sciences--National Academy of Engineering. 322 p. Also: A Synopsis. 30 p. National Academy of Sciences--National Research Council, Washington, D.C.

California Academy of Sciences. 1955. A Century of Progress in the Natural Sciences. 1853-1953. Published in celebration of the Centennial of the California Academy of Sciences. California Academy of Sciences, San Francisco.

Conrad, G. M. 1965. Symposium on information science. Part III. Changing patterns of scientific periodical publication. Bacteriol. Rev. 29(4):523-533.

Creighton, R., and R. King. 1969. The Smithsonian Institution Information Retrieval (SIIR) system for biological and petrological data. Smithsonian Institution Information Systems Innovations 1(1):1-25. Washington, D.C.

Crovello, Theodore J., and Robert D. MacDonald. 1969. Index of EDP-IR projects in systematics. Taxon 18 (in press: December 1969 issue).

Dayhoff, M. O. 1969. Atlas of Protein Sequence and Structure. Vol. IV. National Biomedical Research Foundation, 11200 Lockwood Drive, Silver Spring, Maryland.

Foote, R. H. 1967. Entomology looks at its mission. Information storage and retrieval for entomology. Bull. Entomol. Soc. Amer. 13(2):99-104.

Fuccillo, D. A., Jr., ed. 1967. The Technical Report in Biomedical Communication. Report of a Workshop held April 24, 1967, at the National Institutes of Health. National Institutes of Health, Bethesda, Maryland. (This may not be available as it is marked "For limited distribution.")

Garfield, E. 1964. Science citation index--a new dimension indexing. Science 144:649-654.

Glass, B. 1962. Information crisis in biology. Bull. Atomic Scientists 18(8):6-12.

Gordon, R. E. 1969. Toward an information system for biology--community activity. BioScience 19(7):628-629.

Gottschalk, C., and W. F. Desmond. 1963. World centers of scientific and technical serials. Amer. Documentation 14:188-194.

Gould, S. W. 1962. Family Names of the Plant Kingdom. 111p. International Plant Index, New Haven, Conn.

Gould, S. W., and D. C. Noyce. 1965. Authors of Plant Genera. 336 p. International Plant Index, New Haven, Conn.

Gray, E. E., J. Werdel, and F. J. Weyl. 1967. Scientific and Technical Communications. Secondary Services in the Private Sector. A Digest of SATCOM Reviews and Discussions, February 1966-September 1967. 66 p. National Academy of Sciences-National Research Council, Washington, D.C.

Gurtowski, C. G. 1968. Preliminary Report to the Council on Biological Sciences Information on the Preparation of a Core List of the Substantive Primary Periodicals in the Biological Sciences. 7 p., with appendixes. FASEB, Bethesda, Md.

Handler, P. (Chairman). 1966. Handling of Toxicological Information. A Report of the President's Science Advisory Committee. 21 p. U.S. Government Printing Office, Washington, D.C.

Herschman, A. 1969. A Program for a National Information System for Physics: 1970-1972. American Institute of Physics, New York.

Hutton, R. S. 1967. ASLIB, past and future. ASLIB Proceedings 19:19-28.

International Compendium of Numerical Data Projects. A Survey and Analysis. 1969. Produced under the auspices of the Committee on Data for Science and Technology of the International Council of Scientific Unions. Springer-Verlag, Berlin, Heidelberg, New York (1969). (In English.)

Jackson, R. C. 1966. ASPT-IOPB index of current taxonomic research. Regnum Vegetabile 43:1-72. Utrecht, Netherlands.

Jackson, R. C. 1969. American Society of Plant Taxonomists Index of Current Taxonomic Research. No. 5, 74 p. (multigraphed.) Botany Department, University of Kansas.

Jacobus, D. P., R. R. Gulick, L. Schultz, and P. V. Parkins. 1966. Direct user access to the biological literature through abstracts: A cooperative experiment in customized service. BioScience 16(9):599-603.

Kennedy, H. E., and P. V. Parkins. 1969. Biological literature. In Encyclopedia of Library and Information Science 2:537-551. Marcel Dekker, Inc., New York.

Lawrence, G. H. M., A. F. G. Buchheim, G. S. Daniels, and H. Dolenzal, eds. 1968. B-P-H. Botanico-Periodicum Hunteanum. Hunt Botanical Library, Pittsburgh, Pa.

Licklider, J. C. R. 1965. INTREX: Report of a Planning Conference on Information Transfer Experiments. M.I.T. Press, Cambridge, Mass.

Lwoff, A. 1953. Lysogeny. Bacteriol. Rev. 17:269-337.

National Bureau of Standards. 1968. National Standard Reference Data System News. Technical News Bulletin 52(5).

Orr, R. H., ed. 1963. Communication Problems in Biomedical Research: Report of a Study. National Academy of Sciences-National Research Council, Washington, D.C. (Reprinted in Federation Proc. 23:1164-1176. 1964.)

Parkins, P. V. 1963. Approaches to Vocabulary Management in Permuted-Title Indexing of Biological Abstracts. In Automation and Scientific Communication, Part 1, p. 407-408. American Documentation Institute (now American Society for Information Science), Washington, D.C.

Parkins, P. V. 1966. BioSciences Information Service of Biological Abstracts. Science 152:889-894.

Parkins, P. V. 1969. BioSciences Information Service of Biological Abstracts. In Encyclopedia of Library and Information Science 2:603-621. Marcel Dekker, Inc., New York.

Passman, S. (Chairman). 1968. COSATI Panel Report on the Role of the Technical Report in Scientific and Technological Communication. Federal Council for Science and Technology, Washington, D.C. (Available from Clearinghouse, PB 180 944.)

Pasternack, S. 1966. Is journal publication obsolescent? Physics Today 19(5):38-43.

Porter, J. R. 1967. The scientific journal--300th anniversary. Bacteriol. Rev. 28:211-230.

Porter, J. R. (Chairman). 1964. Style Manual for Biological Journals. (Second edition.) Prepared by the Committee on Form and Style of the Conference of Biological Editors. 117 p. American Institute of Biological Sciences, Washington, D.C.

Price, D. de S. 1961. Science since Babylon. Yale University. New Haven, Conn.

Rauen, H. M., ed. 1964. Biochemisches Taschenbuch, 2nd ed., 2 vol. Springer-Verlag, Berlin, Göttingen, Heidelberg, New York.

Revelle, R. (Chairman. 1965. The Scientific Endeavor. Published in connection with the Centennial of the National Academy of Sciences. Rockefeller University Press, New York.

Schultz, L. 1968. New developments in biological abstracting and indexing. Library Trends 16(3):337-352.

Shetler, S. G., J. H. Beaman, M. E. Hale, L. E. Morse, J. J. Crockett, and R. A. Creighton. 1969. Pilot Data Processing Systems for Floristic Information. 47 p. Preprint of paper given in symposium, "Advances in Data Processing for Biology and Geology," September 24-26, 1969, Cambridge University, by S. G. Shetler (in press, Academic Press, New York and London, 1970). Available from Department of Botany, Smithsonian Institution, Washington, D.C. 20560.

Shetler, S. G. (ed.), L. E. Morse, J. J. Crockett, S. Rakosi, and E. R. Shetler. 1969. Preliminary Generic Taxon Catalog of Vascular Plants for Flora North America. 69 p., multigraphed. Smithsonian Institution, Washington, D.C.

Sober, H. A., ed. Handbook of Biochemistry. Selected Data for Molecular Biology. The Chemical Rubber Co., 18901 Cranwood Parkway, Cleveland, Ohio 44128.

Steere, W. C., ed. 1958. Fifty Years of Botany. Golden Jubilee Volume of the Botanical Society of America. xiii, 638 p. McGraw-Hill Book Co., New York.

Stewart, I. E. 1964. Characteristics of Scientific Journals (1949-1959). Report of the National Science Foundation, NSF 64-20. National Science Foundation, Washington, D.C.

Suszynski, N. J., Jr. 1969. Recent advances in source data automation. Smithsonian Institution Information Systems Innovations 1(3):1-18. Washington, D.C.

Swanson, D. R. 1967. Crucial National Problems in the Planning of Future Research Libraries. SATCOM Working Paper, National Academy of Sciences-National Research Council, Washington, D.C.

U.S.A. Standards Institute (name now changed to American National Standards Institute). 1966. Revised and Enlarged Word-Abbreviation List for USASI Z39.5--1963 American Standards for Periodical Title Abbreviations. Section

Committee Z39 on Library Work and Documentation: Subcommittee on Periodical Title Abbreviations. American National Standards Institute, New York.

Way, K. 1968. Free enterprise in data compilation. Science 159:280-282.

Weevers, T., ed. 1949. Fifty Years of Plant Physiology. 308 p. Chronica Botanica Co., Waltham, Mass.

Weinberg, A. M. (Chairman). 1963. Science, Government, and Information: The Responsibilities of the Technical Community and the Government in the Transfer of Information. A Report of the President's Science Advisory Committee. U.S. Government Printing Office, Washington, D. C.

Weinberg, A. M. 1967. Reflections on Big Science. M.I.T. Press, Cambridge, Mass.

Wilson, J. H., Jr. 1969. Better written journal papers--who wants them? Science 165:986-987.

Woodford, F. P. 1967. Sounder thinking through clearer writing. Science 156:743-745.

Woodford, F. P., ed. 1968. Scientific Writing for Graduate Students. A Manual on the Teaching of Scientific Writing. 190 p. Rockefeller University Press, New York.

Woodford, F. P. 1969. Improving the communication of scientific information. BioScience 19(7):625-627.

APPENDIX

THE COUNCIL ON BIOLOGICAL SCIENCES INFORMATION

The need for more efficient organization to improve handling and transfer of scientific information became apparent in the 1950's when the scientific community--its population and activity--expanded exponentially under the influence of Federal funding. With a 300-year tradition of organization of information in terms of scientific discipline, based on professional societies, and a clear recognition that publication and subsequent packaging of research results was an integral part of the research process, massive aid was provided to those scientific disciplines whose historical organization lent itself to the development of a more or less unified modern information system. Thus rapid strides were made in the development of a chemical data system by the Chemical Abstracts Service under the sponsorship of the American Chemical Society; similar plans were being developed in physics, engineering, mathematics and to a lesser extent in social sciences. Unfortunately, the more complex biological societies, not organized to the same degree, suffer both in development and support.

Brief History of the Council

In 1965, the Trustees of Biological Abstracts, noting the historical splintering of the biological sciences into many autonomous organizations, each with an interest in scientific communication, established an _ad hoc_ committee* charged with planning the development of a more unified approach to the handling of biological information. The Committee recommended the organization of a conference whose participants would embrace major organizations in the biological community to discuss possible approaches to the problem.

---

*The planning committee included: Theodore C. Byerly, Floyd S. Doft, Donald S. Fredrickson, Robert E. Gordon, Philip Handler, Phyllis V. Parkins, William C. Steere, Kenneth V. Thimann and Raymund L. Zwemer (Chairman).

The Conference, subsequently called "The Cherry Hill Meeting" because of its location at Cherry Hill, New Jersey, was held November 22-23, 1965*. The Conference was sponsored by the Board of Trustees of Biological Abstracts and supported by a grant from the National Science Foundation; it was chaired by William C. Steere, by virtue of his position then as president of the Trustees of Biological Abstracts. Conferees numbered 44 from a variety of biological institutions and programs embracing the governmental and private sector. The Conference was opened with an address by Burton W. Adkinson, Head, Office of Science Information Service, National Science Foundation. Dr. Adkinson explained the complex and interlocking Federal bodies in the Executive Branch of the Government that deal with scientific and technological information. He reviewed the steps planned, or being taken, by major scientific fields other than biology to develop, coordinate, and strengthen information services on a discipline-wide basis. He emphasized the urgency for biologists to develop a coordinated and unified plan for an integrated information system to serve the whole discipline of biology.

The Conference divided into five working groups: Oral Communication, Primary Publications, Secondary Services, Data Collection and Information Centers, and Libraries. On the second day, the recommendations of the working groups were presented to the Conference as a whole.

From each working session a variety of recommendations were made. But the five working groups were unanimous in recommending the need for a single organization that would speak with a more unified voice for the information needs of the field. The Conference addressed itself to the nature of such an organization, its location and function**.

---

*Steere, W. C. (Chairman), et al. 1965. Report. Conference on Communications for Biology. Cherry Hill, New Jersey, November 22-23, 1965. BioSciences Information Service of Biological Abstracts, Philadelphia, Pa. (Available from BIOSIS on request)

**Steere, W. C. 1966. Council on Biological Sciences Information. Biological Abstracts 73(3): Editorial, 2 pages, not numbered.

A definitive resolution was proposed including the following elements:

- That the organization be named the Council on Biological Sciences Information

- That the purpose of the Council be defined to include information activities in agriculture, medicine, engineering and other related biological sciences areas

- That each of the following organizations: The American Association for the Advancement of Science, the American Institute of Biological Sciences, the Council of Biology Editors, and the Federation of American Societies for Experimental Biology, be invited to be initial sponsors and to name two representatives each to form the Council on Biological Sciences Information; and that the President of the National Academy of Sciences be requested to lend his good offices in convening the representatives for their formal establishment of the Council

Two representatives from each of the above-named initial sponsors met with appropriate staff members of the National Research Council, under the chairmanship of A. G. Norman on May 14, 1966. A basic organizational charter was established and a proposal for financial support of the organization and its activities was drafted.

The organizational charter, ratified by the initial sponsors as the founding member organizations, provided for the establishment of a Board of Directors to be named by the President of the National Academy of Sciences from a list of ranked nominees presented from the member organizations, and later from the existing Board.

The Council's Board of Directors was constituted in late 1966 and the first meeting convened on January 21, 1967. Officers were elected and an Executive Committee established to work out initial organizational matters, to pull together a statement of goals and objectives, and to propose specific mechanisms for action. A list of Board members and Officers appears on page 77.

Objectives of the Council

The organizational charter specifies the functions of the Council:

- To provide a coordinated voice for the nation's biologists in the field of scientific communication;

- To coordinate activities among the existing components and develop short-and-long range plans;

- To conceive and effect recommendations for improvement of biological communication in line with broad plans and specific goals;

- To serve as a forum for discussion and implementation of ideas;

- To provide liaison for its constituents and others;

- To obtain and dispense funds in support of its activities, through incorporation or close association with another organization.

These general functions were detailed and elaborated in six assumptions adopted by the Board in its first operational meeting on March 10, 1967, as follows:

- The Council represents the biological sciences, as constituted in the research areas of agriculture, biology and medicine, in matters of information.

- The Council may speak for the biological sciences without specific delegated authority from existing biological organizations.

- The Council is primarily concerned with information processes, activities and products involving the private sector as distinct from governmental interests. These processes, activities and products in member organizations are open to evaluation.

- The Council will function as a planning body for specific action, and may delegate funds and specific tasks to other organizations.

- The Council will survey and evaluate components of the existing information system(s) serving the biological sciences, assessing their effectiveness to their users.

- The Council will formulate plans for an effective communication system that is both compatible within the biological sciences and also with other systems in a national network.

The last two items have been regarded as the short term and long term objectives, respectively, of the Council.

Steps Toward the Objectives

Feasibility Study for a Serials Record Center. For nearly ten years various individuals and organizations have noted the need for dependable data concerning the number of primary journals serving the biological sciences. Although several published lists are available, none is complete, nor does any one distinguish between titles that are current and those that have been laid to rest or superseded by different titles. Under the supervision of Dr. Raymund L. Zwemer, the Council has made a study of four major lists preparatory to a proposal for the establishment of a serials record center to serve the biological community. The Zwemer report, "Selected List of Current, Primary Serial Publications in the Biological Sciences," based on 27,000 cards, is published as COBSI Working Document Number 1.

Survey of Existing Activities. To coordinate activities among existing components, one must have some measure of the projects and products in which the components are involved. Inventories of the kinds of activities conducted by the more than 100 different societies and an even greater number of research laboratories and institutions serving the biological community are being developed. The input is organized into six categories: Primary Journals, Secondary Services, Libraries, Data and Information Centers, Oral Communications, and System Development.

Preparation of Working Documents. To enable the Board to make recommendations and to provide a base for a systems analysis, state-of-the-art documents on each of the above mentioned six categories are in various stages of preparation. Members of the Board have been assigned, as appropriate, to each of the six categories.

Input to the Committee on Research in the Life Sciences. The survey of research in the life sciences conducted by the Committee on Research in the Life Sciences, chaired by Dr. Philip Handler, includes a section on the state of scientific communication. Preparation of material for Dr. Handler's Committee occupied much of the Council's attention during 1969.

Liaison Activities. Members of the Board of Directors of the Council participated in the First Roundtable on Information Problems in the Biological Sciences at Columbus, Ohio, September 5, 1968. The Council, with the assistance of the Biological Sciences Information Project, sponsored the Second Roundtable at Burlington, Vermont, on August 20, 1969. Both of these activities were held in conjunction with the annual meeting of the American Institute of Biological Sciences. These roundtable discussions provide the opportunity for specific organizations both to present their on-going activity and to gain insight into that of sister institutions.

Officers and Members of COBSI Board of Directors*

John M. Brookhart, Chairman, Department of Physiology, University of Oregon Medical School, Portland (Board Member, 1969).

Vernon Bryson, Institute of Microbiology, Rutgers University, New Brunswick, New Jersey (Board Member, 1968-71; Secretary, 1968-70).

Fred R. Cagle, Vice President, Tulane University, New Orleans, Louisiana (Board Member, 1967-68; Executive Committee Member, 1967-68).

---

*As of February 1, 1970.

LaMont C. Cole, Langmuir Laboratory, Cornell University (Board Member, 1969-73).

Martin M. Cummings, Director, National Library of Medicine, Bethesda, Maryland (Board Member, 1967).

John T. Edsall, The Biological Laboratories, Harvard University, (Board Member, 1967-70; Executive Committee Member, 1968-69).

Sidney R. Galler, Assistant Secretary (Science), Smithsonian Institution, Washington, D.C. (Board Member, 1967-70; Executive Committee Member, 1969-70).

Ralph W. Gerard, Dean of Graduate Division, University of California at Irvine (Board Member, 1967-68).

H. Bentley Glass, Academic Vice President, State University of New York at Stony Brook (Board Member, 1967).

Robert E. Gordon, Department of Biology, Notre Dame University, Notre Dame, Indiana (Board Member, 1967-73; Chairman, 1967-68; Past Chairman, 1968-70).

Karl F. Heumann, Executive Editor, Federation of American Societies for Experimental Biology, Bethesda, Maryland (Board Member, 1969-72).

Theodore L. Jahn, Department of Zoology, University of California at Los Angeles (Board Member, 1969-72; Chairman Elect, 1969-70; Chairman, 1970-71).

Foster Mohrhardt, Director, National Agricultural Library, USDA, Washington, D.C. (Board Member, 1967).

Phyllis V. Parkins, Director, Biological Abstracts, Philadelphia, Pennsylvania (Board Member, 1967-68, 1969-73).

J. Roger Porter, Chairman, Department of Microbiology, College of Medicine, University of Iowa, Iowa City (Board Member, 1967-70; Chairman Elect, 1967-68, 1968-69; Chairman, 1969-70).

William C. Steere, Director, The New York Botanical Garden, Bronx, New York (Board Member, 1967-69; Executive Committee Member, 1967-68; Chairman, 1968-69).

Robert E. Stowell, Department of Pathology, School of Medicine, University of California, Davis (Board Member, 1968-71).

F. Peter Woodford, The Rockefeller University, New York (Board Member, 1968-71).

Raymund L. Zwemer, Director, Translation Project, Federation of American Societies for Experimental Biology, Bethesda, Maryland (Board Member, 1967-69; Secretary, 1967-68).