

NATIONAL SCIENCE FOUNDATION

WASHINGTON, D.C. 20550

Mc
file

Office of Science Information Service

March 9, 1970

STATINTL

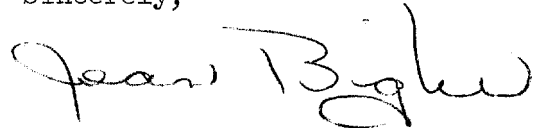
[REDACTED]
Deputy Director of Computer Services
Central Intelligence Agency
Washington, D. C. 20505

STATINTL

Dear Mr. [REDACTED]

Dr. Bamford asked me to forward to you the enclosed briefing for
your information.

Sincerely,



Enclosure

Jean Bigler
Secretary to the
Program Director
for Information Systems

**A Briefing
on
Scientific Communication
and the
Office of Science Information Service**

National Science Foundation

March 1970

**SCIENTIFIC
COMMUNICATION
AND THE OFFICE OF
SCIENCE
INFORMATION
SERVICE**

-1-

Our purpose here today is to place the process of scientific communication in the context of the total scientific enterprise and to show how the Office of Science Information Service contributes to that process.

It should be understood that when we speak of the Office we do so in the extended sense of staff, consultants, cooperating reviewers, and the Science Information Council, which shares the statutory basis of the Office.

THE MESSAGE

scientific communication

role of osis

osis strategy

evolution of system

-2-

If we are successful, when we conclude one-half hour from now we shall leave you with four thoughts:

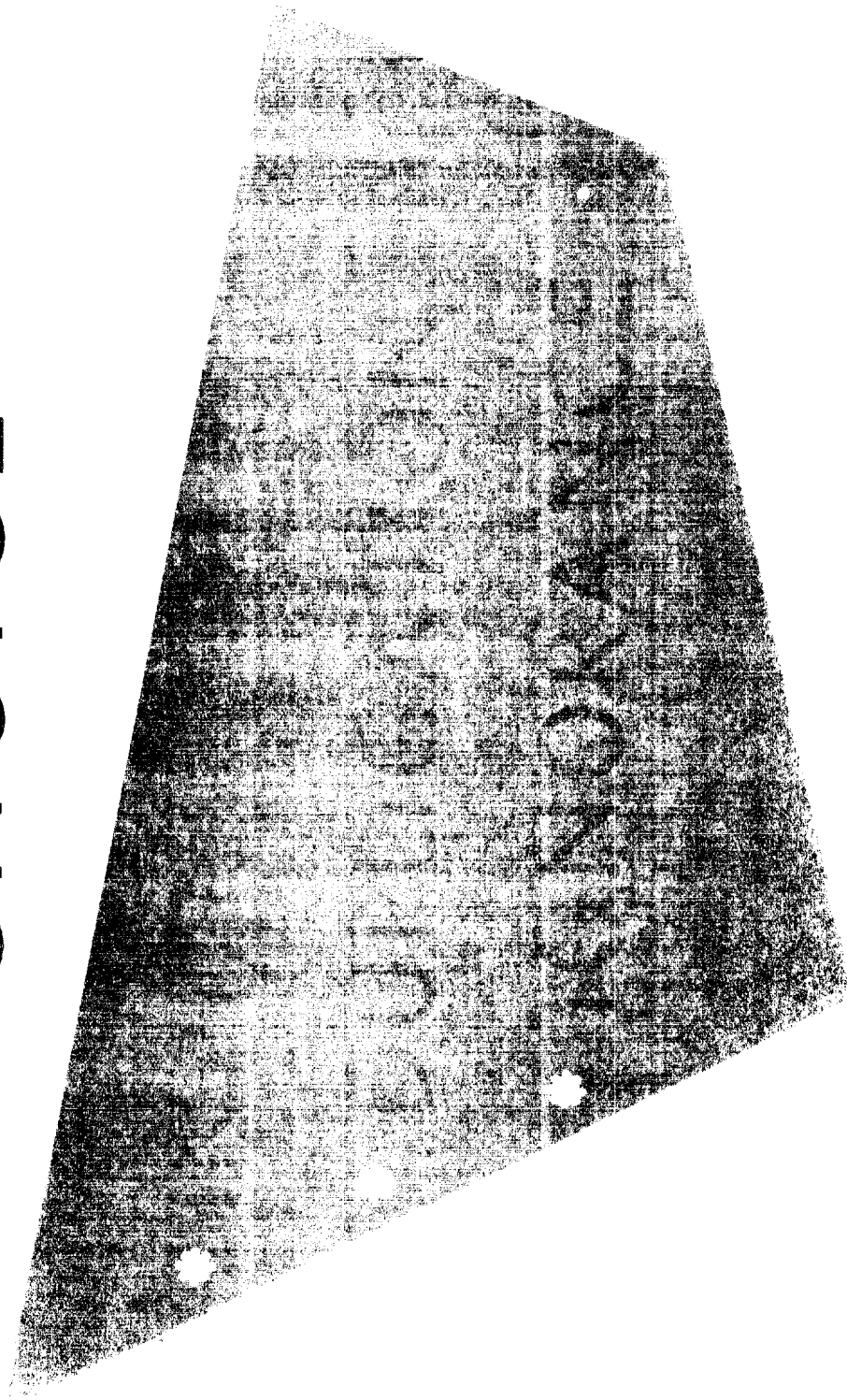
1. Communication is not an optional adjunct of science. It is an essential of the scientific enterprise.
2. The Office plays a key role in the process of scientific communication.
3. The Office is guided by a strategy which we believe to be forward-looking and well conceived, and
4. We may look forward to orderly growth and improvement of the world's scientific communication system.

To begin, let us consider "scientific communication". It involves the flow of scientific information from producers to consumers, two populations which as we know are not altogether distinct. The flow may vary in its formality, the information may vary in its currency, and the messages may vary in their physical form--they may be oral or written.

BEST COPY

Available

THE CONSUMER'S CHOICE



Producers and consumers have their own needs for communication.

-4-

Take the consumer. Whether his occupation be research, education, administration, or technology, his choice is to obtain information produced by others, to bear the burden of producing the needed information himself, or to accept the penalty of ignorance. That penalty is greatest in research, where each investigator builds upon what has gone before and what has been learned strictly conditions what can be investigated at any given time.

FOR THE PRODUCER

MOTIVATION

QUALITY CONTROL

SATISFACTION

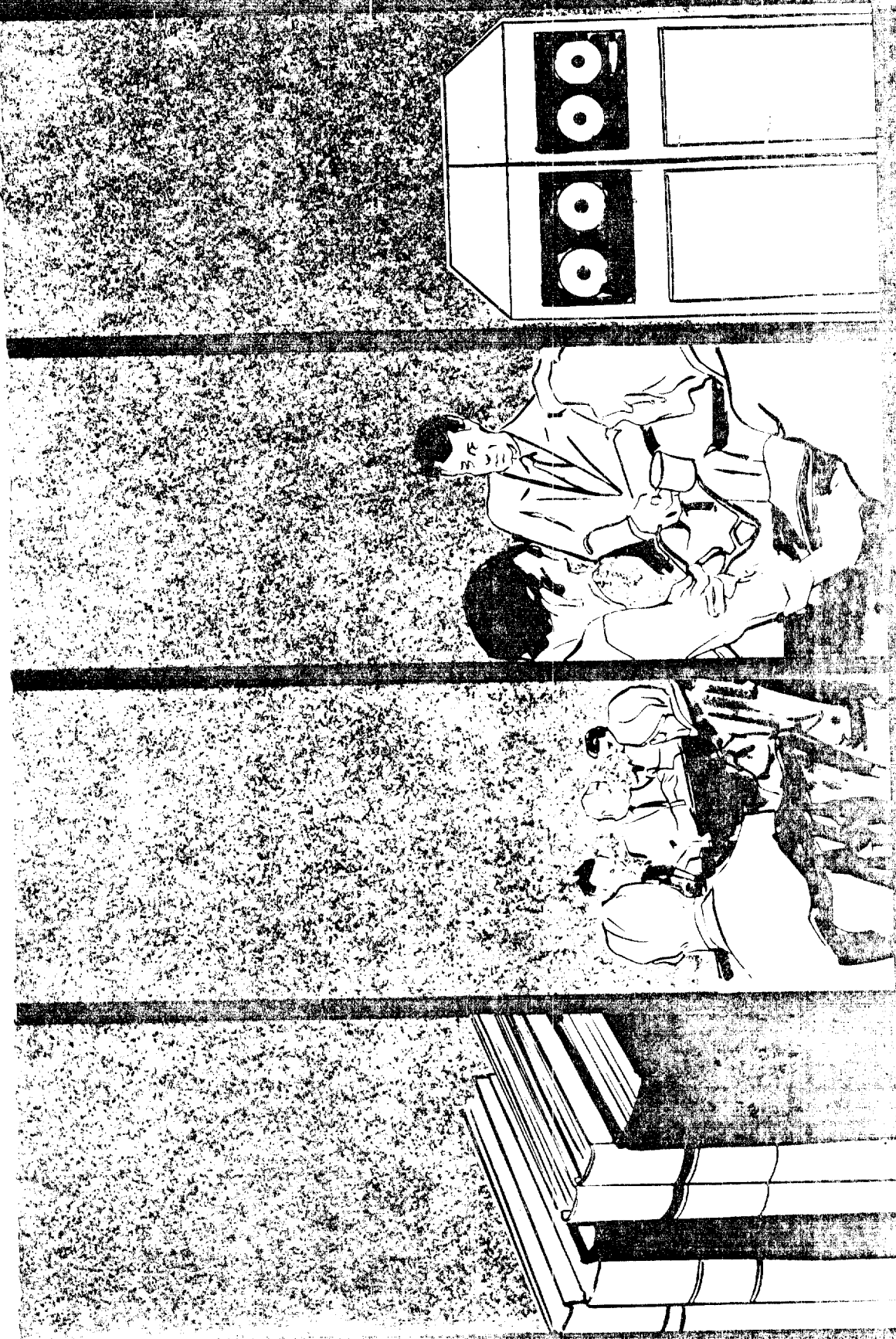
-5-

Shifting our attention now to the producer of scientific information, we cannot doubt his need to disseminate his product. It is fashionable to joke about and deprecate the rule of "publish or perish" under which the academician lives, but that rule has served science well in two respects. It has provided much of the motivation for the prodigious growth of scientific knowledge in this century. And it has made possible the quality control of scientific research through feedback of responsible criticism by the investigator's peers. Indeed, this is the real meaning of "discipline" in science.

The scientist may also take satisfaction in adding to man's store of knowledge, and perhaps hope for a measure of immortality. Be that as it may, his need to disseminate information is as real and as urgent as his need to receive it.

The scientific community and society at large have a stake in the conditions of effective scientific communication. If the effectiveness of communication were reduced, the scientific enterprise as a whole would suffer. By the same token, science can be strengthened by facilitating its communication.

THE COMMUNICATION SYSTEM OF SCIENCE



-7-

In communicating, scientists employ a variety of media. Notable among these are journals and other publications, conventions and other meetings, and telephones and other services. Not to be overlooked are informal written correspondence and conversations over coffee. To assist in using these media, there are abstracting and indexing services, libraries, data analysis centers, and informal referral networks. The totality of these things is what we refer to when we speak of the communication system of science, a complex network of interacting subsystems. No one denies that such a system exists or that it works.

At the same time no one pretends that the system works perfectly or is beyond any need for improvement. In this connection it is instructive to consider what an ideal communication system for science might be like. What would be the characteristics of a system which fully satisfied the communication needs of both producer and consumer of scientific information?

SYSTEM REQUIREMENTS: CONSUMER

*ACCESS

*SELECTIVITY

*SPEED

*FORM

-8-

From the consumer's point of view, four system requirements can be recognized:

- Access: The ideal communication system would provide the consumer with any information in existence which he might need, in any field, and even if he didn't happen to know that it existed, or where it could be found.
- Selectivity: Equally important, it would present only that information which was needed. The consumer would not need to screen his input for relevance, since it would all be relevant.
- Speed: From the onset of a need for information to the presentation of that information there would be no delay.
- Form: The needed information would be presented ready for use. It would not have to be translated, collated, plotted, or otherwise pre-processed by the consumer.

**SYSTEM
REQUIREMENTS:
PRODUCER**

***USEAGE**

***SPEED**

***PREPARATION**

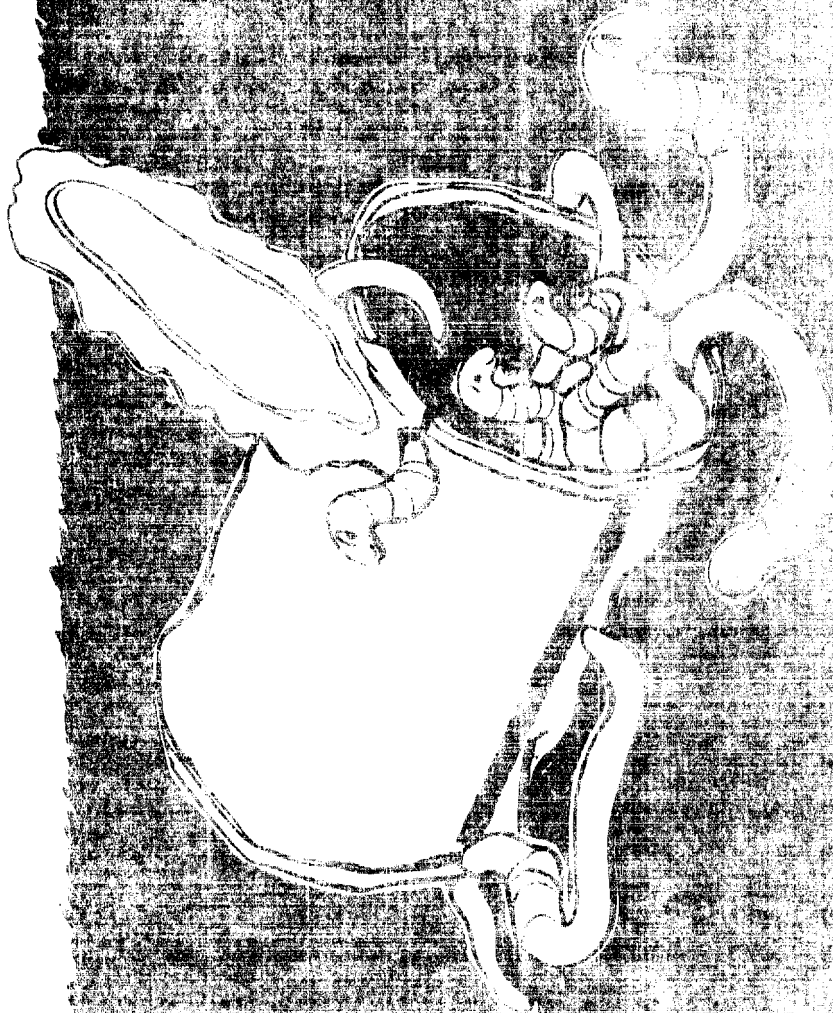
-9-

Similar system requirements may be identified from the producer's point of view:

- Usage: The ideal communication system would disseminate the producer's output to every consumer who needs it, whatever his field, whenever his need might arise, and whether or not his existence is known to the producer.
- Speed: There would be no delay between the production of information and its presentation for use by those who need it.
- Preparation: The producer would be relieved of all the routine functions involved in preparing his output for dissemination. Such functions as formatting, tabulating, plotting and proof-reading would be no more his responsibility than the typing of manuscript is today.

Against the background of these requirements, let us now consider the scientific communication system as it exists today. That system involves the participation of many kinds of organizations performing many different functions and producing a multitude of information products and services. It is comprised of the people, organizations, and facilities which perform functions and services contributing to the flow of scientific and technical information.

THE EXISTING SYSTEM

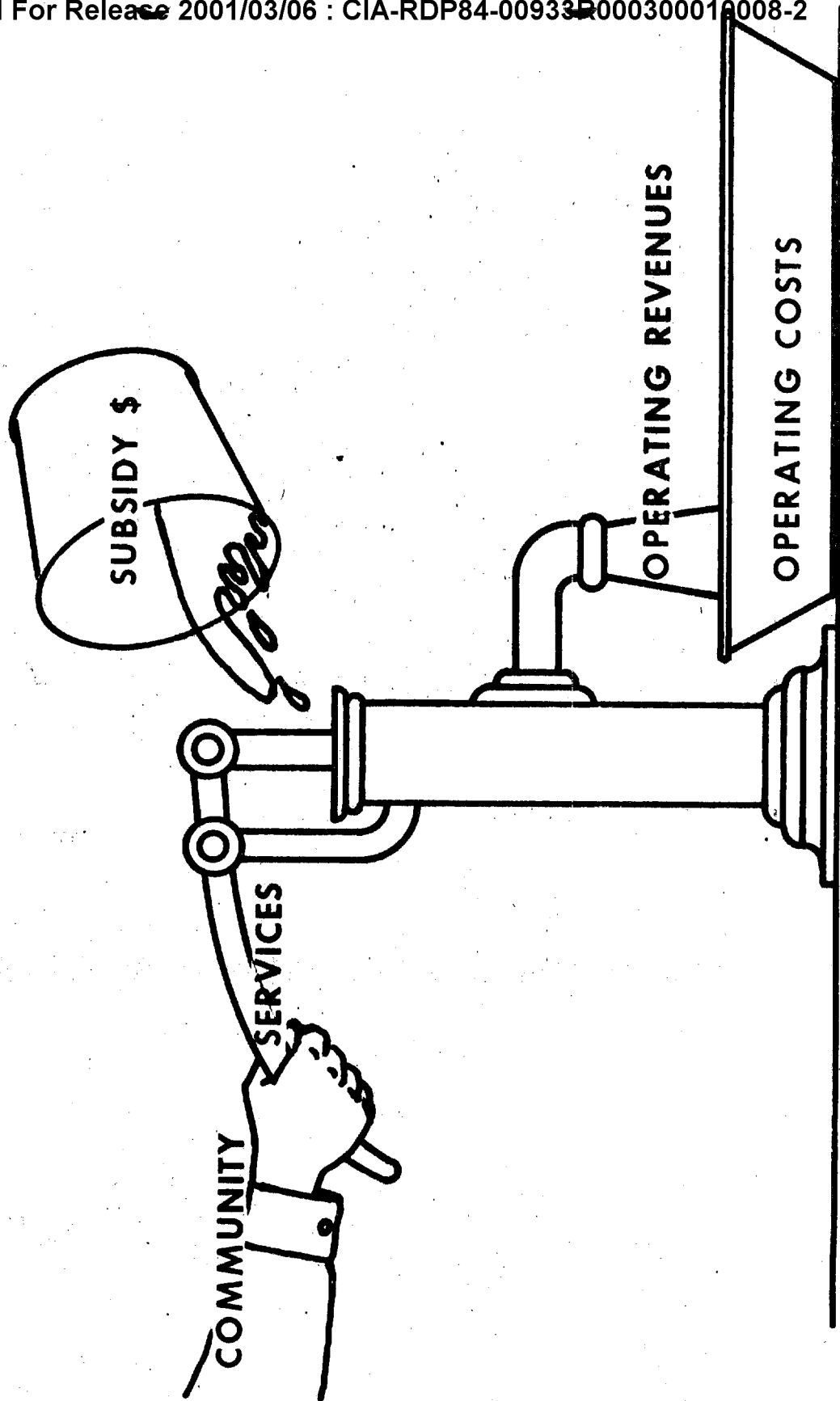


-12-

As indicated here the existing system operates without any central management or control and without much concern over its overall efficiency and effectiveness. It is the result of many years of evolutionary change in the structure of science, in the behavior of the scientific community, and in the technological and economic factors of the communication process. The wide gap between the needs now being served and those which would be met by the ideal system is the primary basis for our program, which is aimed at closing that gap.

The remainder of this presentation will focus on what we in the Office of Science Information Service are doing and intend to do in pursuit of our basic goal, closing that gap. We shall consider in turn the Office's role with respect to each of its four major program areas --namely, system operations, system improvement, research and studies, and the organizational environment of the communication system.

SYSTEM OPERATIONS



-21-

With respect to our first program area, system operations, our goal is for the communication system of science to provide information services which collectively can pay their own way. To this end, we subsidize innovative services during the transitional period in which operating costs are not covered by operating revenues. Established services are also subsidized during periods of deficit to prevent their interruption or degradation before they are able to regain self-sufficiency. In both cases, the subsidy is in the nature of pump-priming and is not intended to continue indefinitely.

SYSTEM
OPERATIONS

INFORMATION
SERVICES

PUBLICATIONS

TRANSLATIONS

-22-

Over the past five years we have committed a third of our funds to transitional operational support.

- In these 5 years the greatest impact has been on the bibliographic activities of the major abstracting and indexing services. The coverage of these services has increased by two-thirds, their speed has come to be measured in months instead of years, and most of them have achieved self-sufficiency through increased revenues and increased operating efficiency.
- In addition to our support of scientific journals and monographs there have also been a number of innovations in the 5-year period. Among them may be mentioned Communications in Behavioral Biology, a multi-purpose scientific journal with selective dissemination features, and the Mathematical Offprint Service, which distributes individual copies of papers from the world's current mathematical literature on a personal basis.
- Seventeen scientific and engineering societies have been assisted in producing scientific articles translated from foreign languages. These articles have been published in 31 different translation journals, 19 of which have become self-sufficient in the five-year period. (In this connection it should be noted that 30% of the world's scientific literature is produced in languages understood by fewer than 5% of United States Scientists.)

-23-

In addition to transitional support we have also served 13 agencies of the Federal Government. Under Executive Order we have provided the coordination and procurement necessary to implement the translation activities of Public Law 480. In the five-year period 1966 through 1970, 29 major bureaus of the 13 agencies have been provided materials processed from foreign languages. In this period, the Office obligated \$6½ millions of Foundation funds and administered \$4 millions on behalf of other agencies. The translations procured in this program are available to the scientific community through the Clearinghouse for Federal Scientific and Technical Information.

SYSTEM
IMPROVEMENT
COMPUTERIZING
DATA BANKING
NETWORKING
UNBURDENING

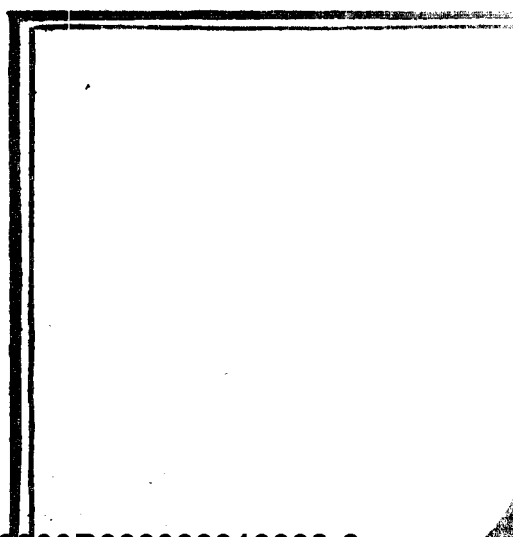
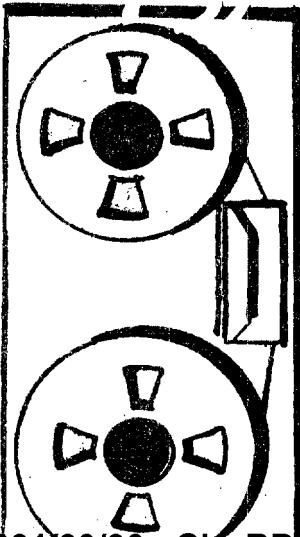
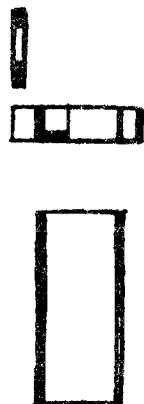
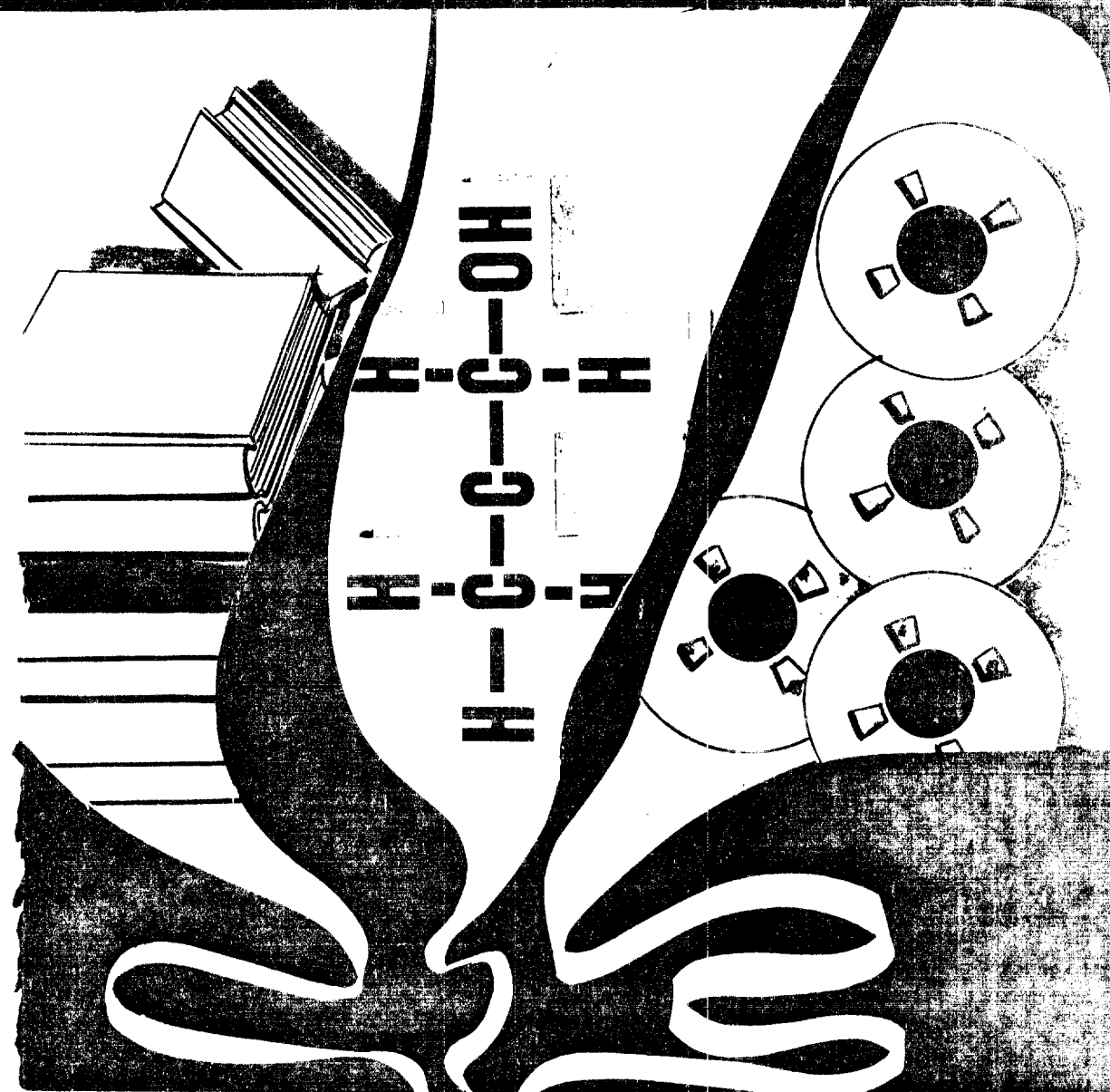
-31-

In our second program area, the development and improvement of the communication system of science, our ultimate goal is the ideal system. Our intermediate goals are recognizable steps in that direction. Some of these intermediate goals are indicated here.

- Computerizing the communication system is probably the only route to the great advances in the speed and selectivity of information services needed to cope with the ever expanding body of scientific knowledge. Even traditional services can gain in dependability and efficiency from automation.
- Great extensions in access are possible in the form of computer data banks.
- The formation of a world-wide distribution network, incorporating consumer-oriented, as well as producer-oriented, elements expands the usage of the communication system.
- Another intermediate goal is to unburden consumers and producers of their pre-processing and preparation functions as rapidly as these functions can be taken over by the communication system.

Although some pay-off from investment in these improvements can be hoped for in the form of reduced costs for a given service, the real pay-off is in increased productivity. We get some idea of what this means when we reflect that a 1% increase in the productivity of scientists and engineers in 1970 would be worth \$272 millions.

SYSTEM IMPROVEMENT



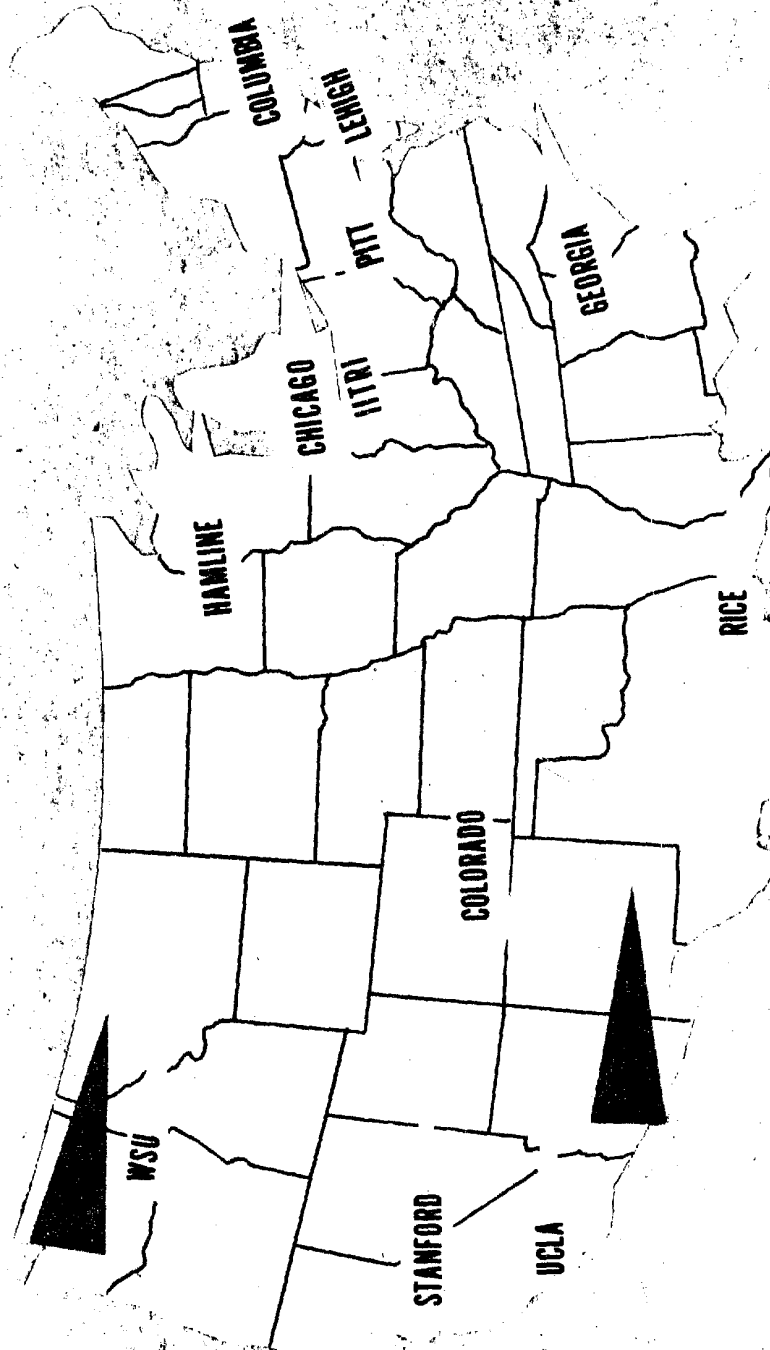
-32-

Over the past five years, we have invested 47% of our funds to improve the communication system of science. Let me mention some of the major results achieved to date.

- Comprehensive programs of system development have been organized in the fields of chemistry and chemical engineering, physics and astronomy, electrical and electronics engineering, psychology, and linguistics.
- Computerized production and dissemination of indexes, bibliographies, and other products on the basis of centralized computer files have been initiated in the fields of chemistry, biology, engineering, and geology.
- A computerized Chemical Registry System has been established with a data base which now contains descriptions of more than one-and-a-quarter million substances, more than one-and-a-half million chemical names, and more than two-and-three quarters million references to the scientific literature.
- Computerized stores of data on the literature of physics and geology have been established; and similar files in such fields as ichthyology, general biology, herpetology, engineering, paleontology, plant taxonomy, psychology, and mathematics have been established in museums and universities.

These and more will begin to pay off as more and more consumer-oriented systems develop the capability of offering services to consumers in their own institutions.

SYSTEM IMPROVEMENT



Here are some of the consumer-oriented systems in which we have invested.

- The green flags indicate discipline-oriented information centers, chiefly in the physical sciences and engineering. These universities are acting in the role of retailers of the "wholesale" products of the discipline-oriented systems and the independent banks already mentioned. They are also developing capabilities to serve their campuses and surrounding industry using data banks from abroad and from the Federal agencies.
- The red flags identify an Arid Lands Information System at the University of Arizona and a Treaty Information System at the University of Washington. Each of these is expected to serve an international clientele.
- The six blue flags represent university library modernization and networking projects.

Each of these university-centered systems is potentially a node in the national and world-wide distribution networks. Some of them are already providing services to scientists in academia, government, and industry. The University of Georgia's system, for example, serves 4,000 consumers, including many of the Food and Drug Administration's bio-chemists. Physicists at the Stanford Linear Accelerator Center have available a prototype information service via on-line terminals. Chemists at the University of Pittsburgh and those in adjacent industry are having their requests processed against tapes produced by the Chemical Abstracts Service of the American Chemical Society, and this also holds true for the system being developed by the IIT Research Institute.

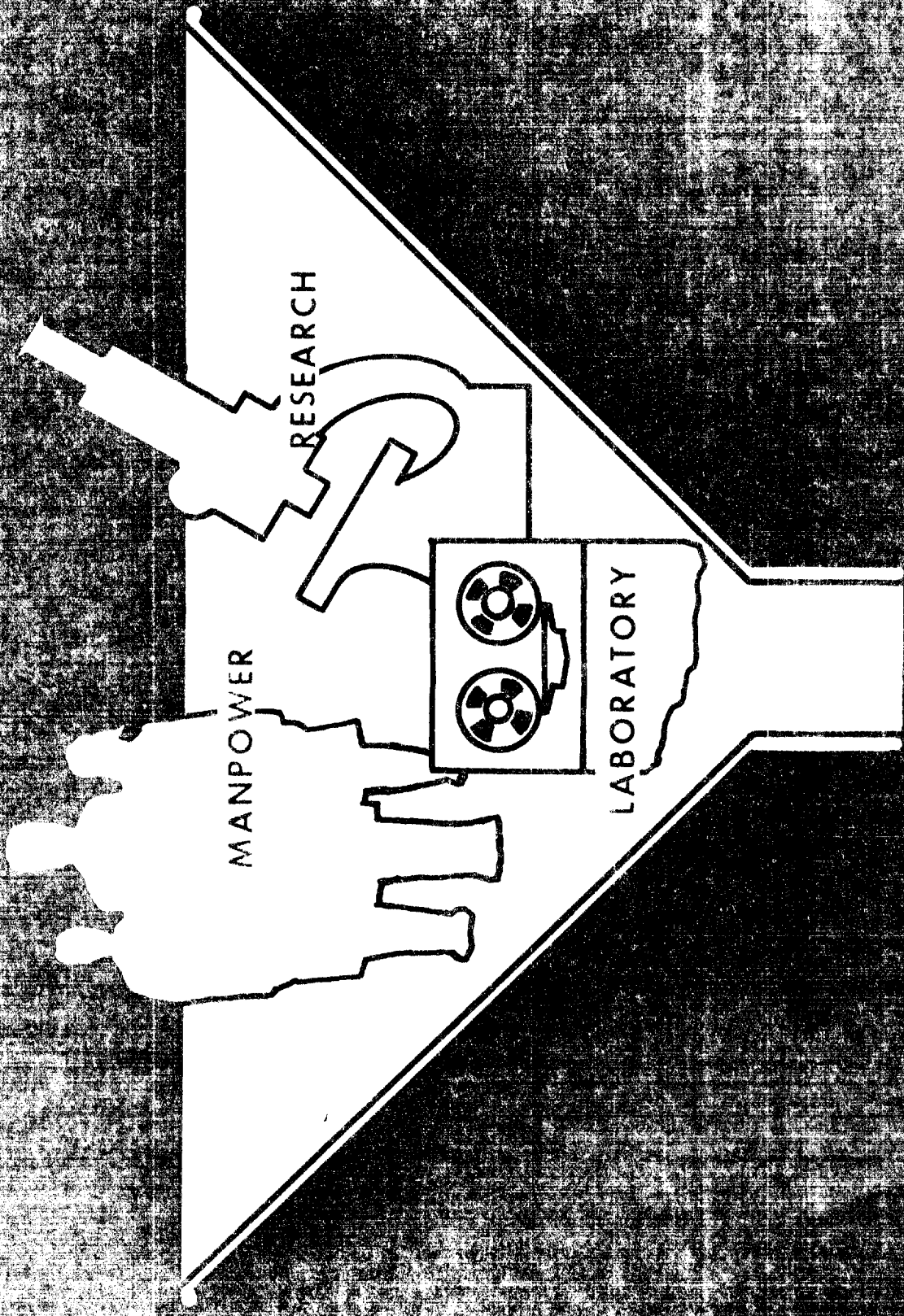
SYSTEM IMPROVEMENT: OBSTACLES



But the path to science's ideal communication system is beset with formidable obstacles.

- Technically, the major obstacle to progress is our surprisingly meager understanding of how scientists and engineers go about using and producing information. Everything considered, however, the technical obstacles to growth are likely to be far more easily surmounted than the organizational ones.
- Science is highly organized, but its organization is highly decentralized. Its institutions tend to be parochial and its bureaucracy conservative. Responsibility for programs of change or improvement is not readily accepted nor easily recognized. Aggressive leadership is in short supply, and to organize any kind of joint action at the local, regional, national, or international level is difficult. In a word, the organizational environment is inhospitable to system growth.
- As to finances, improved services must be expected to cost more, and inflation has been no kinder to scientists than to anyone else. But the main financial obstacle to improving the communication system of science is the problem of financing the improvement itself and the operation of the improved system during a transitional period.

We shall indicate the nature and magnitude of the financial problem in a few minutes, but first we want to take a look at what we can do about the technical and organizational obstacles to growth.



NEW INFORMATION TECHNOLOGY

-41-

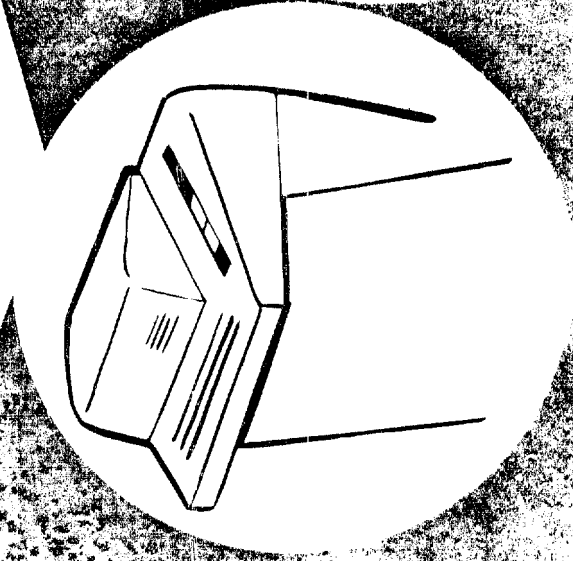
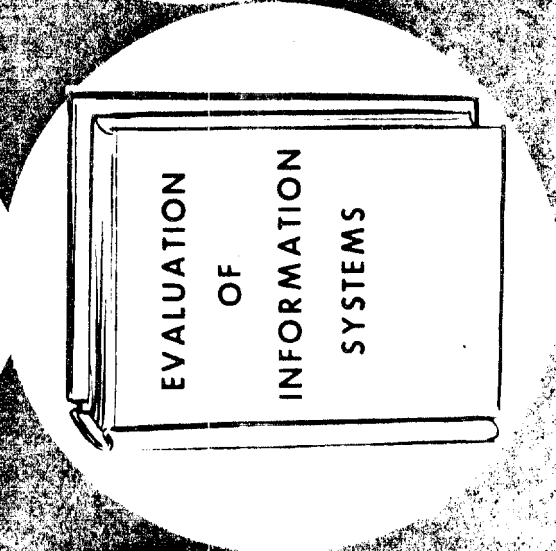
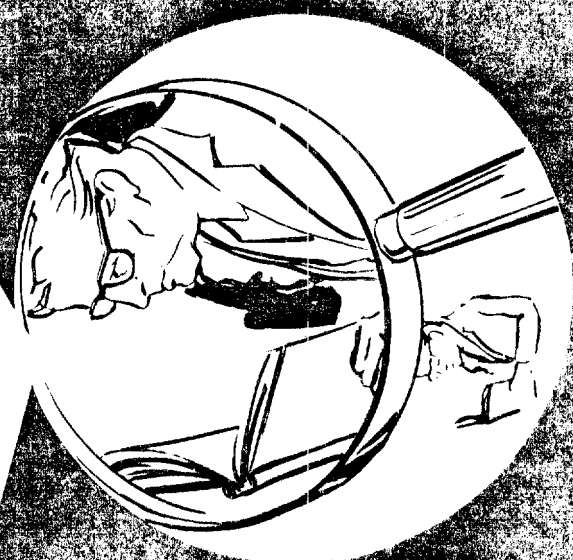
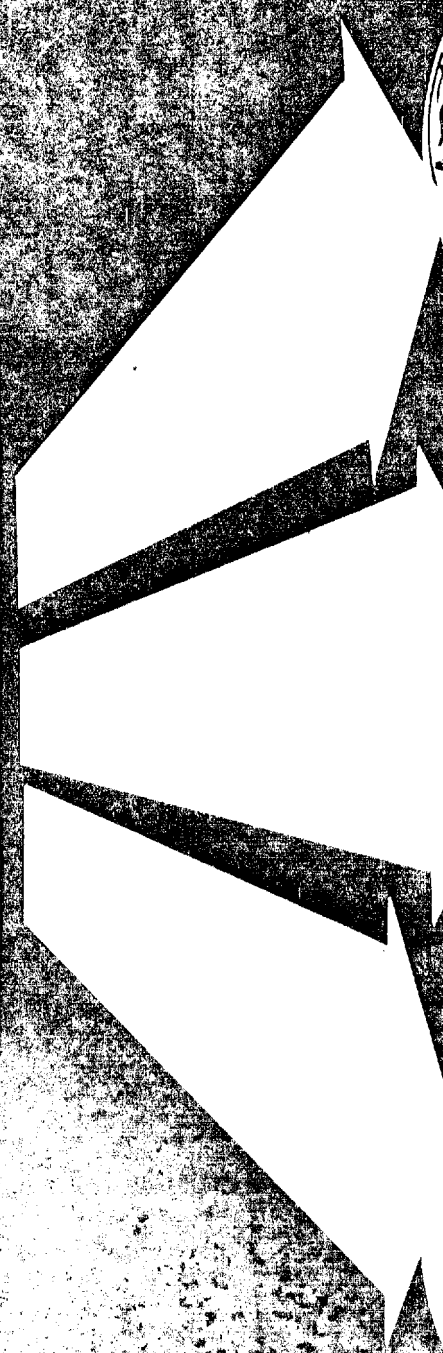
Our third program area, research and studies in information science, is aimed at overcoming the technical obstacles.

To cultivate a technological environment which is conducive to system growth involves

- Manpower: the provision of a cadre of scientists and engineers who can design, develop, operate, evaluate, and modify modern information systems;
- Research: the generation of new and improved methods for the acquisition, organization, storage, search, manipulation, presentation, and dissemination of scientific information;
- and the provision of the laboratory and experimental facilities necessary for research and research training in this field.

"What about basic research?" Most research fundamental to progress in scientific communication is actively supported by other programs of the Foundation and the Federal Government. It can be found in such fields as solid state physics, the mathematical theory of information, computational linguistics, and social psychology.

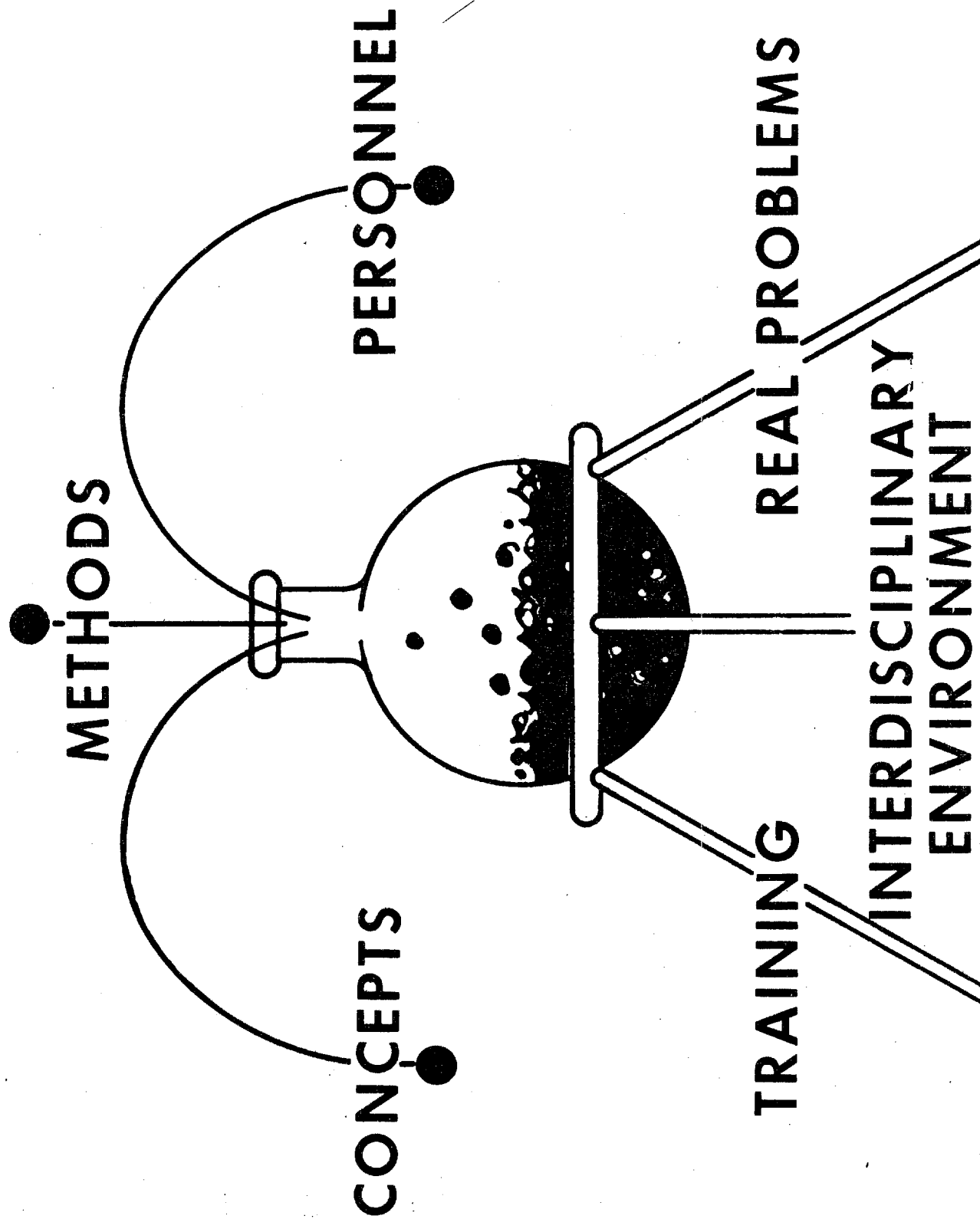
RESEARCH & STUDIES



In the past five years we have committed more than \$9 millions to the advancement of the technological environment through research and studies in information science.

- A number of projects sponsored by the Office have helped to set the stage for system development. For example, a research project at Stanford University provided the nucleus of a university-centered science information system.
- Several projects have contributed to the methodology of information system evaluation. The first comprehensive practical handbook on this subject will soon be published.
- Progress has been made in mapping the communication patterns of scientists and engineers so that the impact of modifications to existing services can be assessed.

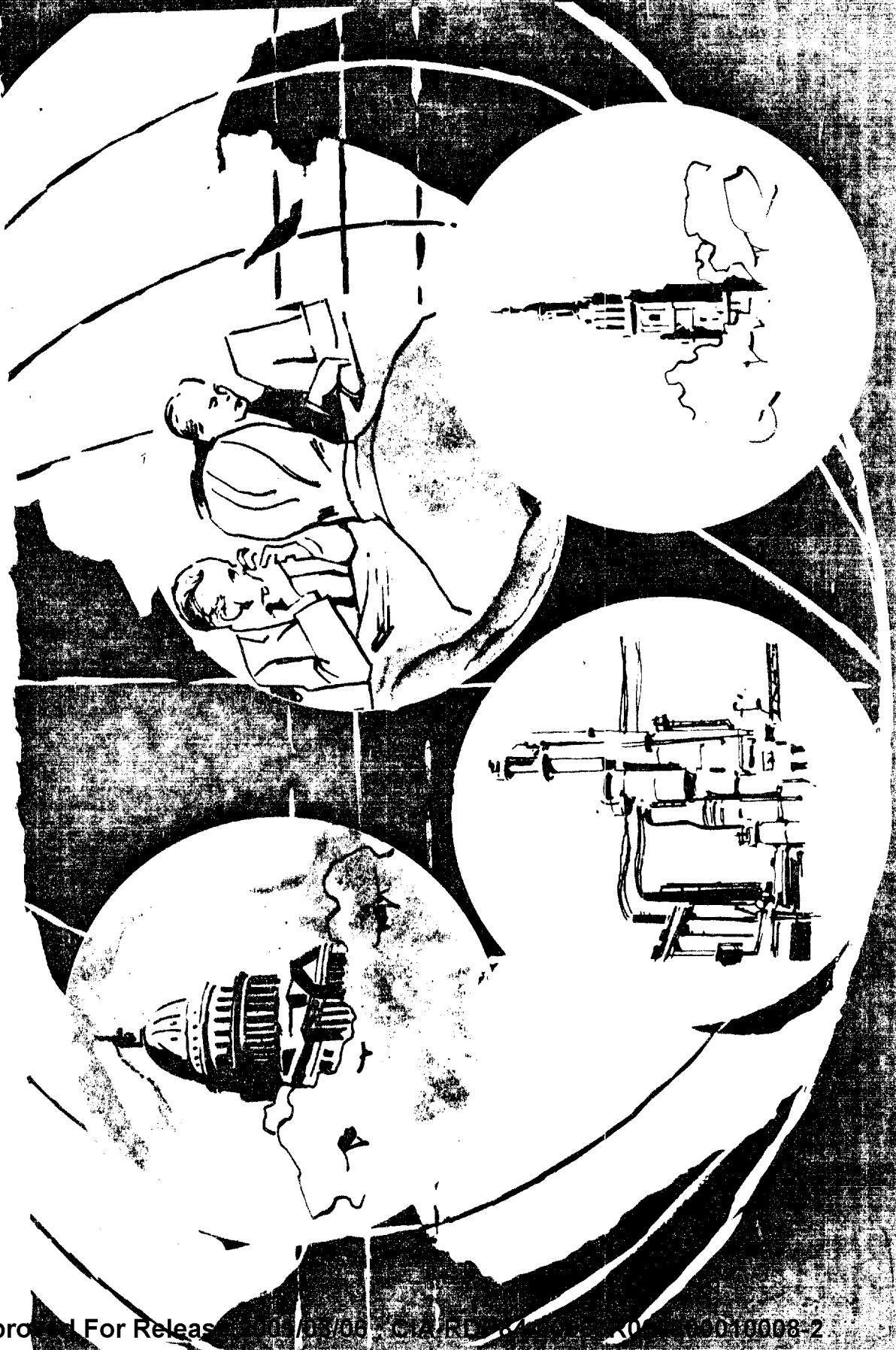
Other areas where contributions have been made include information theory and computational linguistics, methods of retrieving information from very large files or data bases, format compatibility and convertibility, problems of handling full text, the linkage of a wide range of communication and computer equipment, and informal communication between scientists.



-44-

Unfortunately, there is a serious shortage of qualified investigators in information science. We therefore support research centers in information science and technology at leading universities. The criteria for support of our research centers are illustrated by the tripod of this slide. To qualify for support of such a center, a university must offer an interdisciplinary environment where students can get research training in information science and access to an operating information service with real problems. Out of these centers are expected to come, not only new concepts and methods, but also the professional personnel needed to apply them to the existing communication system and to carry on further investigation.

ORGANIZATIONAL RELATIONS



The world-wide organizational environment involves the relations between government agencies, professional societies, industrial organizations, and academia. In our fourth major program area we have sought to cultivate an organizational environment conducive to orderly growth and improvement of the world-wide communication system of science.

- We have stimulated the formation of committees and councils of leaders concerned with communication in various fields of science, such as the Council of Biological Editors.
- We have assisted in the formation of interdisciplinary coordinative bodies, such as the National Federation of Science Abstracting and Indexing Services.
- We have encouraged and supported the activities of national and international standardization authorities. Standards have recently been developed for transliteration of cyrillic characters and abbreviation of journal titles.

In addition, our grants have made possible the participation of United States scientists in the work of international organizations, such as UNESCO and the International Council of Scientific Unions, which are concerned with scientific communication. And finally, the organizational environment has been affected indirectly by our support of system operations and improvement. For example, the Association of Scientific Information Dissemination Centers has recently been formed by a group of information centers in universities, not-for-profit organizations, government, and industry to promote the application of information storage and retrieval technology to large data bases; to share experiences and information through meetings, seminars, and workshops; to recommend standards for data elements and codes; and to promote

STRATEGY: EMPHASIS

- * SUBSIDIZE OPERATIONS**
- * INVEST IN DEVELOPMENT**
- * FOSTER TECHNOLOGY**
- * CULTIVATE ORGANIZATIONAL ENVIRONMENT**

Of the four program areas described, the chief emphasis in our strategy is on improving the communication system of science through development.

System operations in any one year have at best a transient effect on the scientific enterprise. To subsidize them may, at worst, encourage sub-optimum practices. In any case, our total budget could subsidize no more than a tiny portion of the flood of scientific information disseminated annually. In contrast, even one-half of our budget would constitute a major fraction of the total annual investment in improvement of the communication system of science. And the impact of such investment is cumulative, with each year's improvements built on those of previous years. Clearly, the emphasis belongs on investment, with operational subsidies provided mainly in the period of transition following the initiation of new or improved services.

The other areas are also subordinate to system improvement. Investment in the communication system of science presumes an environment favorable to the growth and refinement of that system. Among the obstacles which we have identified to such growth and refinement are technical limitations and adverse organizational arrangements. Accordingly, we foster advances in the state of information technology and cultivate an organizational environment conducive to scientific communication. The point here is that these functions support our primary function of investment in lasting improvements of the communication system.

STRATEGY: APPROACH

~~IN-HOUSE~~

*

~~PROCUREMENT~~

*

RESPONSIBLE AGENTS

*

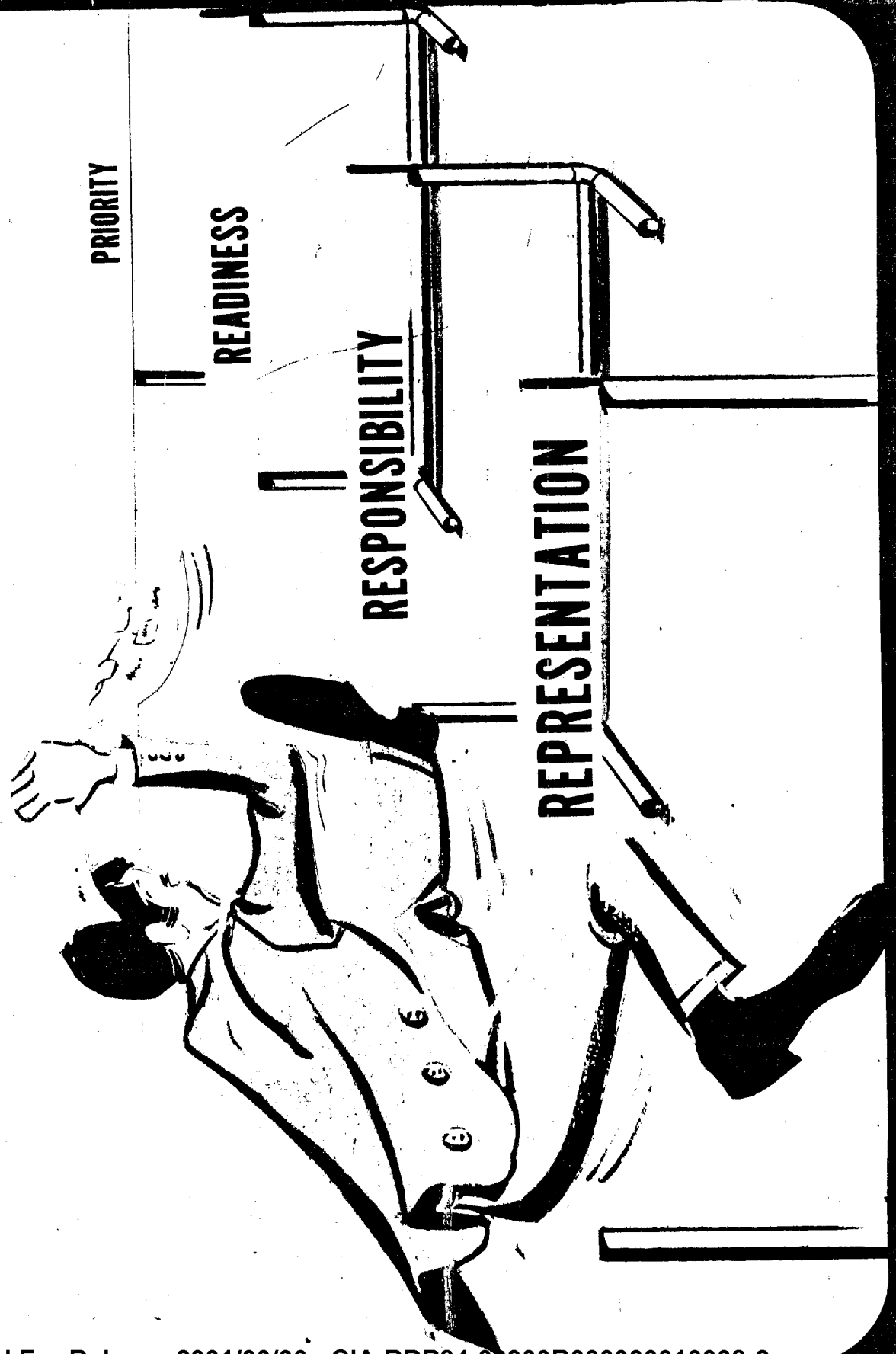
We have had to choose among three main approaches to such investment.

- The Government could shoulder the full responsibility. It could identify the specific needs to be met, plan for specific services to meet those needs, implement the plans, and provide the services. In this way it could retain the initiative, simplify the problems of coordination, and avoid a number of sticky problems altogether. But it would have to underwrite the whole program financially, and the scientific community would have no real control over the services it receives or the destiny of the system itself.

- The Government could contract directly with the commercial and non-commercial sectors to do all of these things. But then its agents would have no other obligation than to fulfill the terms of their contracts. Since the full responsibility would remain in the Government, this procurement approach would be no more than a minor variation on the in-house approach.

- The only real alternative to the in-house approach is for the Government to recognize the responsibility of certain institutions as agents of the scientific community and to assist them financially in meeting the community's communication needs. In this approach, the initiative belongs to the agent, along with the responsibility. Control of its communication system is vested in the scientific community, along with the burden of supporting that system.

STRATEGY : CRITERIA



-63-

Here are the criteria by which we screen the system improvement proposals we receive. The first two flow directly from the strategy of reliance on responsible agents. Thus to qualify for help in improving the communication system of science an applicant must show that it represents a significant scientific community and that it is charged with and has accepted responsibility for serving that community's communication needs.

The applicant must then demonstrate a state of readiness to discharge its responsibility. This involves two elements: (1) the applicant's commitment to an explicit and credible goal and to one or more specific objectives, and (2) the probability that each specific objective will be achieved at a cost less than its potential value. The probability of success is obviously a function of the technical and management resources at the applicant's command.

Applicants which qualify for support by reason of the first three criteria, "the three Rs", must compete with each other for the limited funds which are available. The highest priority is given to nondeferrable projects within organized programs of system development and improvement. The next priority is given to other nondeferrable projects, and the lowest priority goes to projects which can be deferred without great penalty to science. Competing projects within any given priority class are funded in accordance with their readiness.

STRATEGY: IMPLICATIONS

* READINESS

* COORDINATION

* OWNERSHIP

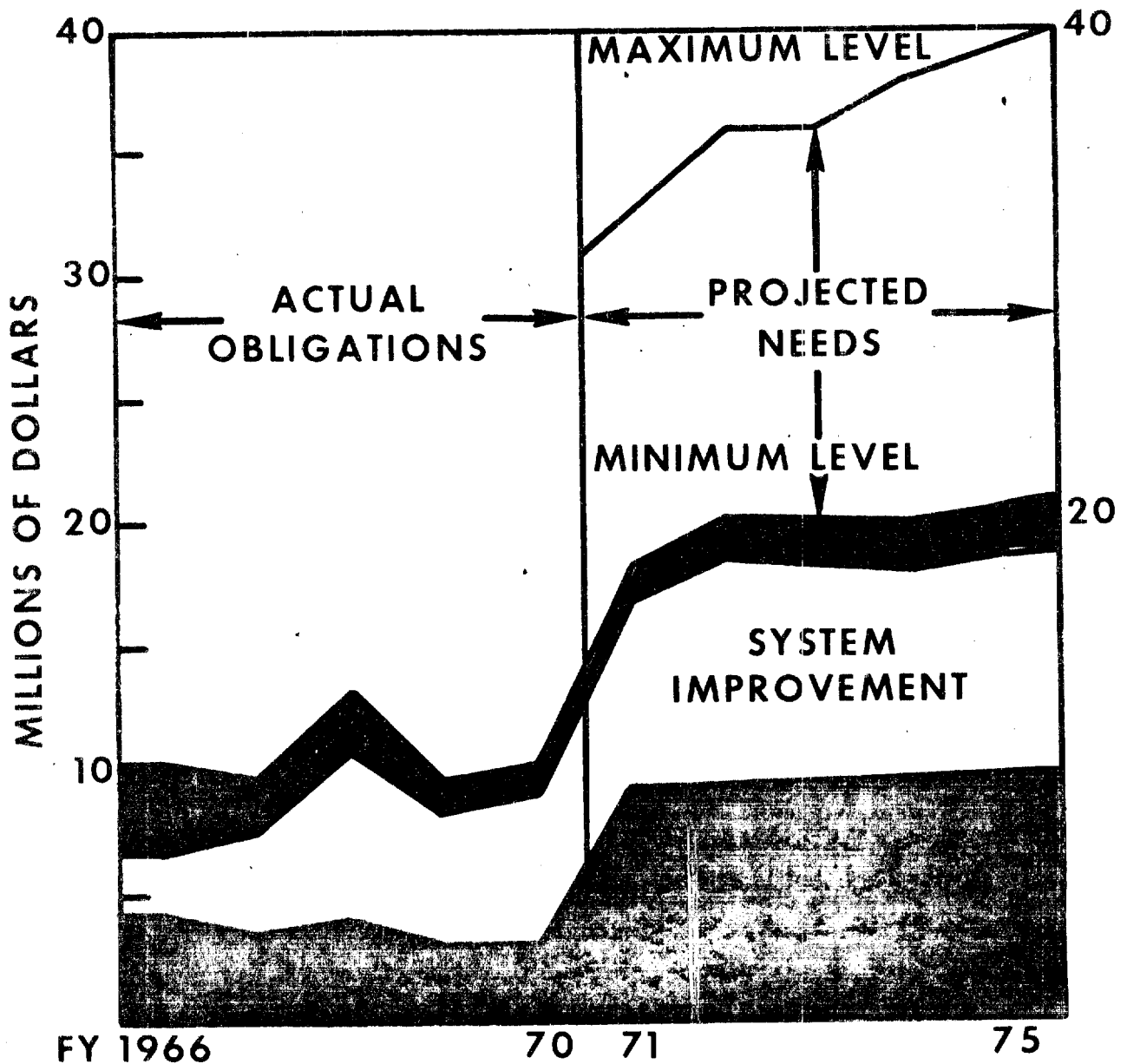
* REGULATION

-64-

Here are some important implications of the strategy of reliance on responsible agents:

- A responsible agent may be assisted to a state of readiness.
Much of the Office's budget in recent years has gone to help such agents acquire the technical and managerial know-how to carry out a program of system improvement.
- When two or more responsible agents recognize a need for joint action of any kind it must be presumed that they will initiate it. The Government can only make sure that the channels of communication are open and that no agent is artificially sheltered by Federal subsidy from a real need to cooperate.
- Ownership of assets created by a responsible agent on behalf of the community it represents must vest in that agent, as long as the Government's assistance is extended explicitly in consideration of the agent's responsibility. Unfortunately, we have the bugbear of exploitation of the scientific community by its own agents. But for the Government to guard against misuse of assets by reserving controlling rights in them to itself would be in effect to relieve the agent of its responsibility.
- Other mechanisms exist by which the use of assets can be regulated in the public interest--e.g., legislation, litigation, and the withholding of funds. But the simplest and most straightforward regulation is that which the community does for itself, through its established institutions, to protect the varied interests of its

STRATEGY: COST



-71-

A major implication of our strategy is its cost:

- Here we have the Office's actual obligations over the past five years for system operations (bottom red), system improvement (green), and cultivation of an organizational and technological environment (reddish) conducive to system growth. The annual total over that period averaged just over \$11 millions.
- Over here we have our minimum projections of need over the next five years, adding up to nearly \$20 millions per year on the average.
- If we were to use our maximum projections of need, the average annual rate would exceed \$36 millions.

Please note that these are projections of need, not of future obligations. Whether and to what degree the needs will actually be met remains to be seen.

Taken by themselves, these figures are impressive, particularly the projections. But let us consider them in context. \$11 millions per year is only about 3% of the total Federal outlay for scientific and technical information in Fiscal Year 1969, and less than 1/10 of 1% of the total Federal outlay for research and development that year.

**ANTICIPATED
BENEFITS**

TO SCIENCE

TO THE NATION

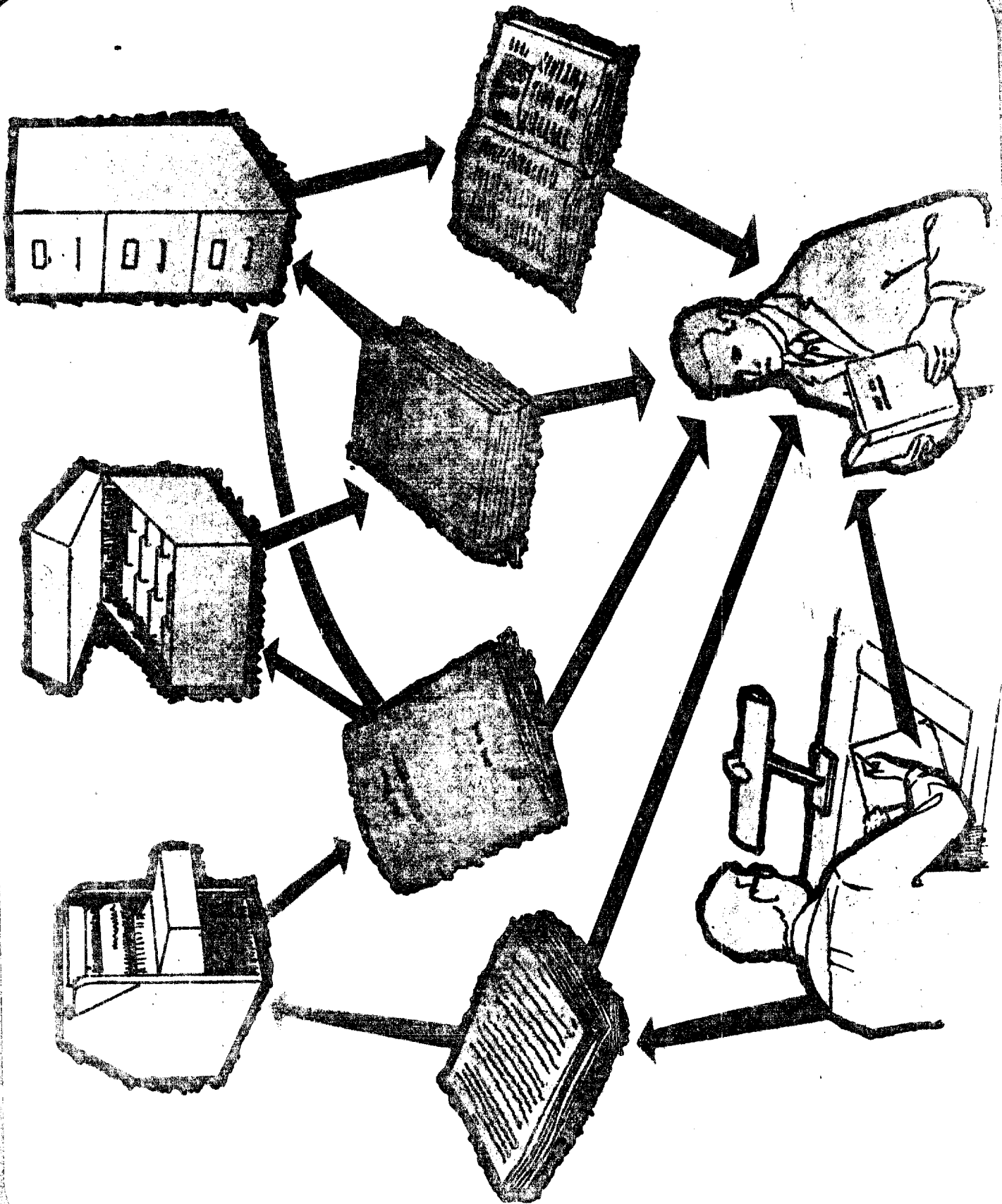
THE IDEAL SYSTEM

-74-

Extending the context of these figures a little further, let's look at how they stack up against the benefits which may be expected.

- For Science, our program promotes the continuity of essential services, discourages unnecessary duplication of scientific resources, and brings about a continuing rise in the productivity of scientists and engineers.
- Benefits to the nation as a whole include increasingly valuable information resources for use in attacking the urgent problems of society, a diminishing time-lag from-research-to-application-to-production-to-use, expanded markets and employment opportunities, and heightened prestige and leadership in international affairs

Our transcendent goal, of course, is the ideal communication system described earlier. Looking into the future, it is not difficult to discern that system evolving from the one which exists today.

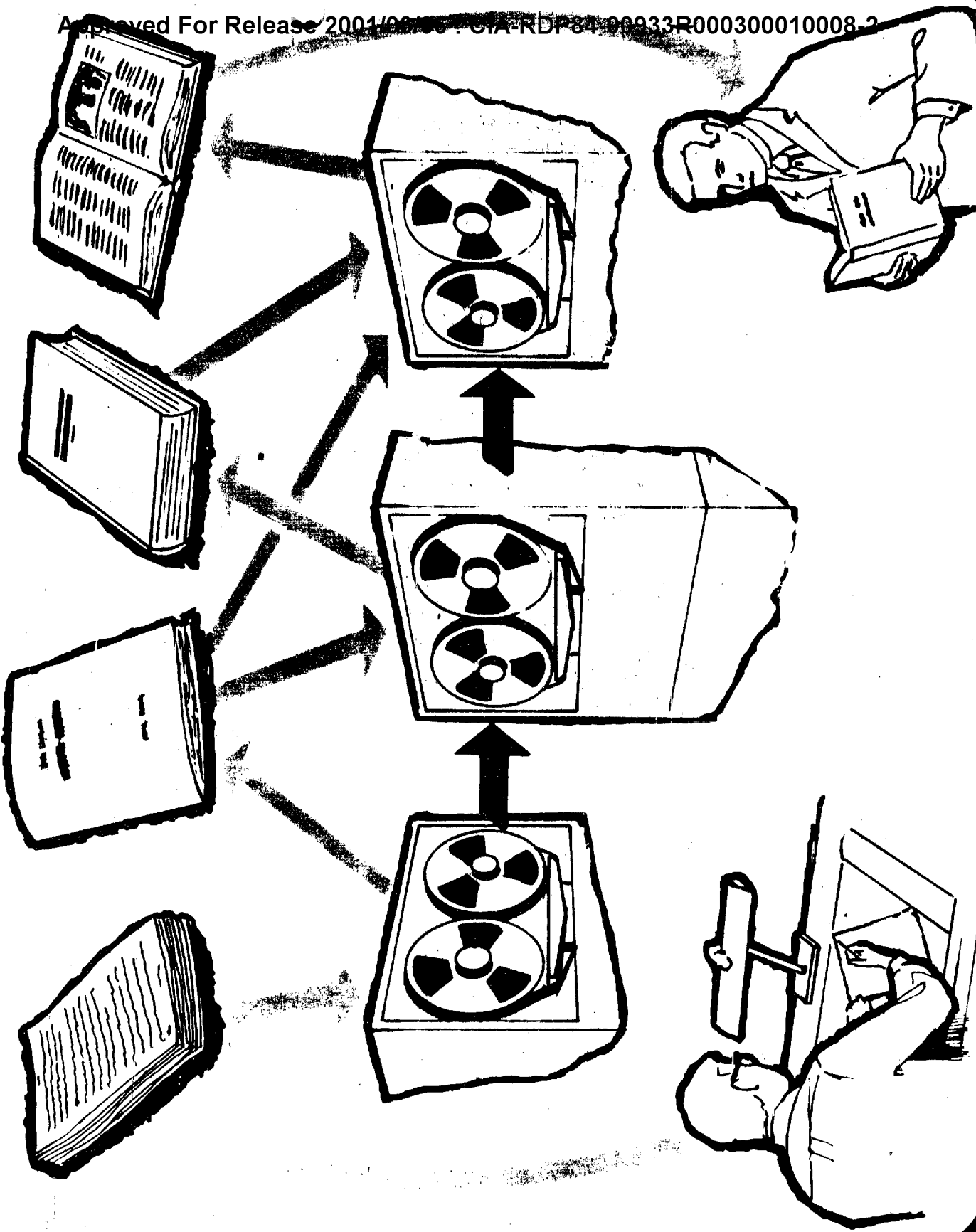


-75-

Today's system, sometimes referred to as "the baseline", is represented here. As you can see, the formal circuit is probably the slowest route which information can take from production to consumption--i.e., from the author--through primary publication--through secondary or bibliographic services--through libraries and information centers--and on to the consumers.

-75a-

Because of this slow and cumbersome arrangement it is not surprising to find the system short-circuited by both consumers and producers. Consumers informally scan the primary and secondary literature, arrange to receive preprints of articles in their fields of interest, and establish direct contacts with producers via mail, telephone and meetings. Producers likewise circulate their manuscripts to known colleagues and participate in various other informal ways to tell the world of their latest achievements.



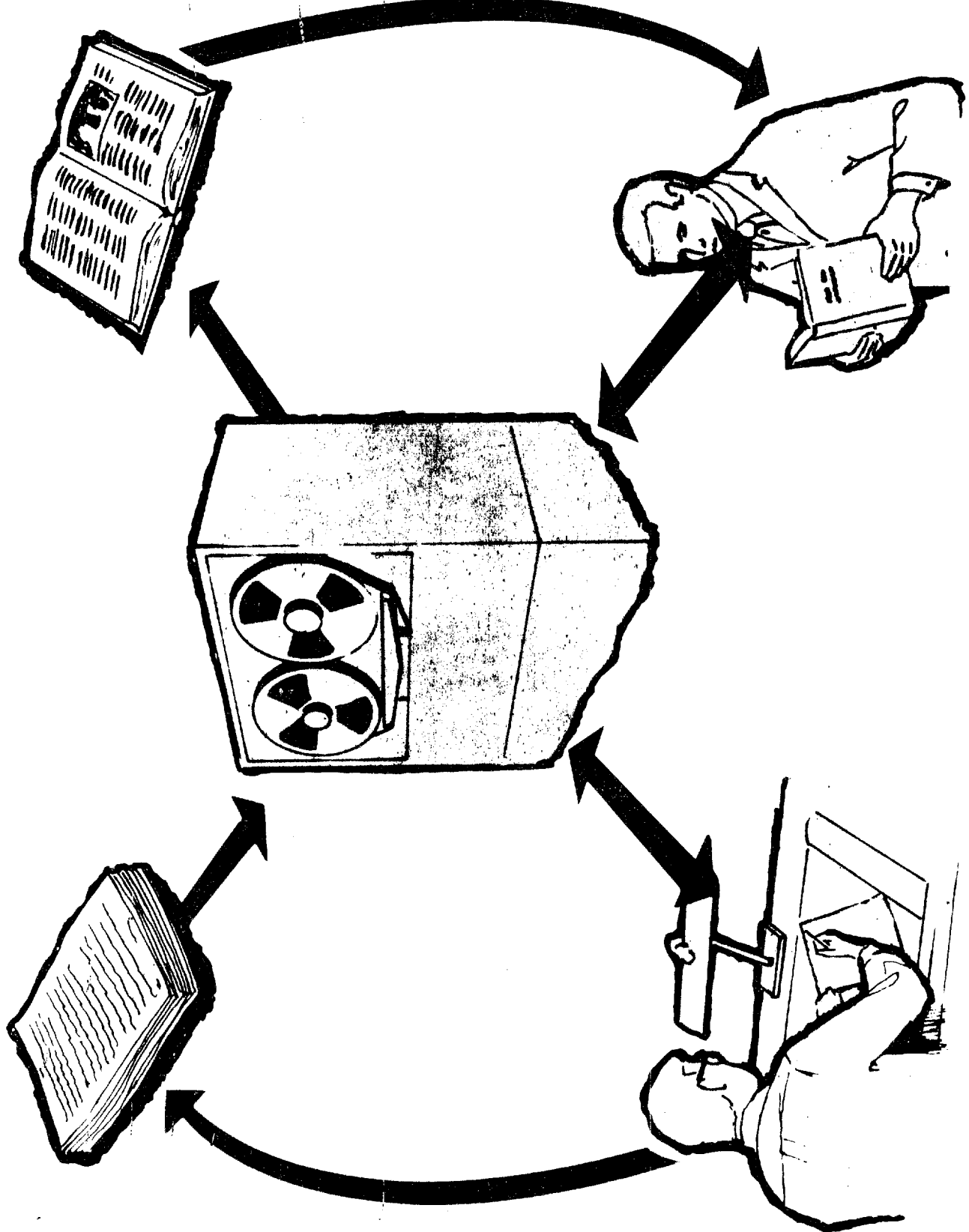
-76-

Largely by means of the computer, the baseline system is already being modernized for the sake of increased efficiency and dependability, as well as to increase its speed and selectivity.

- Publication can be upgraded by computer-assisted typesetting and distribution.
- Bibliographic organization can be assisted by computerized indexing and storage.
- Literature retrieval by computerization of search procedures is practical.

-76a-

Given such modernization, the integration of separate functions becomes feasible at an early stage. To eliminate duplication and to effect economies of scale, the secondary publishers can be expected to operate directly on magnetic tapes produced in the primary publication process. Similarly, search and retrieval can be expected to employ tapes produced in the course of organizing the primary literature. Indeed, the publication of abstracts and indexes as we have known them may gradually disappear.



-77-

Going one step farther in this process we can easily anticipate a more advanced level of integration in which

- manuscript is entered directly into a comprehensive machine-readable file of recorded knowledge,
- the system selects contents of the file for presentation to the consumer as required by him, and
- both primary and secondary publications as we know them will disappear.

But that is still not the end.

-77a-

Ultimately, even manuscripts and formal articles may vanish. The system will communication with producer and user in their own terms, unconstrained by the classical formalities of reporting.

We do not, of course, look forward to a monolithic or centralized computer system, as this display may suggest. What we do anticipate is a world-wide, adaptive, and increasingly intricate network of overlapping specialized facilities, including computers.

THE MESSAGE

scientific communication

role of osis

osis strategy

evolution of system

Now let us review the message.

-2-

- We have attempted to show that communication is indispensable to science, needed alike by consumers and producers of scientific information. Our program arises in the wide gap between the needs being met today and those which would be met by the ideal communication system.
- Our ultimate goal is to close that gap. To that end, we subsidize information services as necessary, invest in lasting improvements of the communication system of science, and cultivate an environment conducive to system growth both technically and organizationally.
- Our main emphasis is on the investment function, which is supported by our other functions. We rely on responsible agents of the scientific community to make the improvements in which we invest. To select our investments we invoke the three Rs--Representation, Responsibility, and Readiness. Full implementation of our strategy over the next five years will require between \$100 millions and \$180 millions.
- Looking ahead, while we cannot foresee the full realization of the ideal communication system, we can discern an orderly development of the existing system in the direction of the ideal, with manifold benefits to science and society at every point along the way.