Volume 87

SCI-TECH LIBRARIES
OF THE FUTURE
SCI-TECH LIBRARIES OF THE FUTURE

Edited by
CYNTHIA A. STEINKE

Routledge
Taylor & Francis Group
LONDON AND NEW YORK
SciTech Libraries of the Future

CONTENTS

Foreword: Plus Ça Change. . .
Cynthia A. Steinke

The Needs of Science and Technology
Robert M. Hayes

Introduction 1
General Context 2
The Role of Information Access in the Stages of Scientific Research 18
The Role of Libraries and Librarians 25
Conclusion 29

Managing the Sea Change in Science and Technology Libraries
Glenn L. Brudvig

Introduction 33
Economics of Change 35
Science and Technology Libraries 36
Leadership and the Change Process 38
Staff and the Change Process 40
The Learning Organization 41
Building a Vision of the Future 42
Quality Management 43
Conclusion 46
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scientific and Technical Librarians: Leaders of the 21st Century</td>
<td>49</td>
</tr>
<tr>
<td><em>Barbara Lockett</em></td>
<td></td>
</tr>
<tr>
<td>The Twenty-First Century Environment</td>
<td>50</td>
</tr>
<tr>
<td>Leadership</td>
<td>55</td>
</tr>
<tr>
<td><strong>Leadership in Science-Engineering Libraries:</strong></td>
<td>65</td>
</tr>
<tr>
<td>Considerations and Realities for the Future</td>
<td></td>
</tr>
<tr>
<td><em>Donald G. Frank</em></td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>65</td>
</tr>
<tr>
<td>The Science-Engineering Library</td>
<td>66</td>
</tr>
<tr>
<td>The General Nature of Leadership</td>
<td>67</td>
</tr>
<tr>
<td>Leadership: Role, Responsibilities, and Skills for the Future</td>
<td>68</td>
</tr>
<tr>
<td>Libraries in Transition: Challenges for Leaders</td>
<td>73</td>
</tr>
<tr>
<td>Concluding Remarks</td>
<td>75</td>
</tr>
<tr>
<td><strong>Education and Recruitment of Science and Engineering Librarians</strong></td>
<td>77</td>
</tr>
<tr>
<td><em>Crit Stuart</em></td>
<td></td>
</tr>
<tr>
<td><em>Miriam A. Drake</em></td>
<td></td>
</tr>
<tr>
<td>How Sci/Tech People Work</td>
<td>77</td>
</tr>
<tr>
<td>Current Situation</td>
<td>78</td>
</tr>
<tr>
<td>Education</td>
<td>80</td>
</tr>
<tr>
<td>Recruitment</td>
<td>81</td>
</tr>
<tr>
<td>Alternatives</td>
<td>82</td>
</tr>
<tr>
<td>Conclusion</td>
<td>85</td>
</tr>
<tr>
<td><strong>Through a Kaleidoscope Darkly</strong></td>
<td>89</td>
</tr>
<tr>
<td><em>Karen Hunter</em></td>
<td></td>
</tr>
<tr>
<td>Generalities</td>
<td>89</td>
</tr>
<tr>
<td>Market Signals</td>
<td>91</td>
</tr>
<tr>
<td>Technology-Push</td>
<td>93</td>
</tr>
<tr>
<td>Future Action</td>
<td>95</td>
</tr>
</tbody>
</table>
Collection Development vs. Access in Academic Science Libraries

Gary Wiggins

Which Life Science Journals Will Constitute the Locally Sustainable Core Collection of the 1990s and Which Will Become "Fax-Access" Only? Predictions Based on Citation and Price Patterns 1979-1989

Tony Stankus
Carolyn V. Mills

The Promise of Fax
Letting Local Library Collections Focus on What Is Actually Read
Factors Used in Defining the Locally Sustainable Core Collection
Laying Out the Discussion of Results
Specialty-by-Specialty Results
Will Fax Mean the Death of Low-Citation-Rating and/or Overly Expensive Print Journals?
Could Fax Data Join Citation Data in Reforming Journals?

A Comparison of Science Related Document Delivery Services

Robert T. McFarland

Introduction
Pilot Project
Vendor Analysis
Methodology
Results
Discussion
Conclusions
Foreword:
Plus Ça Change . . .

We are now at a time of true revolution in the communication of scientific information. Technology is making the difference, transforming the process of knowledge delivery and acquisitions, forcing users to consider new channels of scholarly communication, challenging the viability of the research library. The pace of change in science, in communications and computers is so rapid that if librarians do not aggressively seek to understand the changes and become central to the knowledge formation and transfer process, contributing our skills with information handling, we well may be left in the dust, to never catch up.

These are heady and difficult times in science libraries. As we look to the technological applications to enhance speed of access to information and to solve the spiraling costs of printed publications, we are also reminded that the research library continues to carry the basic role of preserving the record of the past as well as access to these preserved records. Libraries face internal tensions between the investment in acquiring and preserving and the continuing operating expense in accessing resources from offsite. The choices are not always obvious.

"Plus ça change, plus c'est la meme chose" may have been a realistic predictor in the past, but I see a different picture for the future. Change will be so dramatic and will occur more rapidly than any of us have experienced in the past. It is critical to the survival of the sci-tech library that we work with scientists and engineers to understand their changing information needs and to participate in the planning and development of new information systems.

The future will not be the same. What will it be like? Who will be the key players? To secure some of the answers, we have called upon leaders of sci-tech libraries to reflect on their years in the profession.
and to predict the future in whatever way they chose. The assignment was to envision how the sci-tech library will look ten years out. Each author offers his/her best guess, as we struggle to understand and to control the future directions of sci-tech libraries.

* * *

It is a special honor to lead off this volume with a paper by Dr. Robert M. Hayes, Professor Emeritus and former Dean of the Graduate School of Library and Information Science, UCLA. Dr. Hayes presents an in-depth analysis of the information needs of science and engineering and the future development of electronic means to meet these needs.

Hayes describes the sci-tech information required to support education, to serve research and its applications, government, and public policy decision making. His paper contrasts the differing patterns of scientific research and the application of results among the disciplines. His vision of the future pays considerable attention to the effects of information technologies with special emphasis on the increasing importance of digitized imaging as a medium of communication.

These critical information needs, the changing patterns in communicating behavior and the developing technologies, demand new forms of electronic management and delivery of information. The potential contribution for the sci-tech library in support of the information process is complex and calls for new ways of thinking.

The promise of information technology notwithstanding, Dr. Hayes cautions us to remember that the basic role of research libraries is to preserve the record of the past. “All of the other information industries and activities have only the objective of immediate distribution.” Additionally, the responsibility of the library is to provide access to these preserved records. With the advent of the electronic media, “. . . the library faces difficult choices between incurring a capital investment in acquiring and preserving and incurring continuing operating expenses in access from elsewhere. . . .”

He makes it clear that the potential role for the librarian is “much broader than the library itself.” Our skills and commitment provide the professional with a solid base to fulfill a wide range of functions in the context of electronic information. Dr. Hayes outlines the
functions and challenges these changes present to us in meeting the unrelenting expansion of sci-tech information and the researcher's need to have immediate access to it.

Brudvig, Lockett, and Frank focus on the leadership needed to bring sci-tech libraries successfully into the future world. Brudvig presents readers with a challenging ride into the "sea change" that may well transform sci-tech libraries into those existing only in a computer, the electronic or virtual library. ("Sea change" as defined by Webster's Third is a marked transformation into something richer and finer.) These changes, brought on by new technologies, are coming too fast, are too complex and disruptive for traditional leadership methods to be effective. This process of change is accelerating at a rate faster than what we experienced with the automation of recent years, and demands that we re-examine not only the way we operate libraries, but re-examine the basic ways we provide information. The changes almost defy prediction, and thus demand that leaders be knowledgeable, agile and focused more on the design of flexible organizations and less on their operation.

The sci-tech literature will continue to grow, but the sci-tech library that continues to be dependent on costly print journals may well become an institutional overhead to be replaced. Technology is the only solution to the economic dilemma of sci-tech libraries. The journal is a natural target for full-text retrieval. Imaging and high speed transmission will replace photography, FAX, and (blessedly!) microfilm.

Yet, sci-tech libraries face the paradox that the technologically sophisticated user is still in the minority; most users are still bound to printed books and journals. Comfortable behavior patterns are not easily abandoned. The old and the new will co-exist for years to come and leaders must be keenly aware of the changing information needs and expectations of their users, or users will seek their own solutions!

In this rapid sea of change, leaders of sci-tech libraries cannot be knowledgeable in all areas. They must see the broad patterns, become designers and architects, teachers, coaches and mentors—instead of bosses. Their energies need to be concentrated on building shared visions in order to develop a "learning organization"
that promotes change, one based on the needs of customers and which continually strives for excellence and quality.

Lockett, in her predictions for the future, looks first at the high-change environment of higher education—scientific research itself—and the changing patterns of scientific and technical communication that will impact on the leadership of tomorrow.

Computers have revolutionized the way science is conducted, scientists now communicating via supercomputers and high-speed telecommunications systems. The information explosion will continue. The formats in which information is communicated will increasingly diversify. Electronic formats will comprise a larger portion of the whole, but the print version will not disappear. For librarians, it is critical that we work with the scientists in order to understand the way they “do” science and the way they use the new communication systems.

Lockett speculates that library organizations will become more fluid and collaborative. Leadership will be a distributed function. Librarians will be defined more by what they do, rather than where they work—moving out of the library to become partners with researchers, offering the benefit of a long tradition and experience in the interpretation and management of information.

The greatest challenge for libraries in the 1990s will be to identify ways to collaborate with other influence groups in developing networks that will enhance the librarian’s role in providing a wide range of electronically based information services. Change for libraries will be greater and the pace more rapid. But in Lockett’s view “we will not need . . . a new breed of superlibrarian in the 21st century. Leaders will still be influencing people . . . and while our environments . . . will continue to shift rapidly, the most stable part of the system and that on which we exert our influence will remain human.”

Frank’s article details the attributes necessary for effective leadership in the sci-tech library that will survive in the future. Most of these qualities are important for successful leadership in all libraries, but especially critical for the sci-tech library where immediacy of access to appropriate resources is essential.

The dynamic environment of the sci-tech library calls for leadership that is proactive. These leaders will need: creativity and vision
to anticipate the future and to cause organizational change to meet the changing patterns of scientific scholarly communication; the budgetary realities; the issue of access vs ownership; and to anticipate the impact of information technologies. These will be the attributes necessary to meet the changing and increasingly sophisticated information seeking behavior of scientists and engineers.

Drake and Stuart are concerned with the quality of information services provided to the scientific enterprise when sci-tech libraries must rely on information professionals with little or no training in the sciences. The recruitment pool has never been strong; it shows no signs of improving.

Studies of information seeking behavior of scientists and engineers show that librarians are not the first choice for consultation when information is needed. Scientists generally presume that if the librarian has no grounding in the sciences there is not going to be any personal knowledge base to aid the researcher and no common frame of reference. The individual involved in research, therefore, is going to acquire information elsewhere.

Increasingly, technology is making it possible for scientists to access information without libraries. If libraries and librarians are to play vital roles in R & D and the associated information transfer process, they must change—or they indeed may disappear.

Effective recruitment of professionals is a key. It means changing library school education, training on the job, increasing salaries, and improving the image of librarians. Drake and Stuart discuss these alternatives.

U.S. competitiveness and productivity can be improved with more effective information services. Information professionals can be key agents in the information transfer process, but satisfying the information needs of customers requires keen understanding of the problem and how the information is to be used in the individual’s work. Drake and Stuart contend that to achieve this we must recruit people with science and engineering backgrounds into the information profession.

Hunter offers a ten-year prediction for changes in sci-tech libraries generally, and for sci-tech publishers in particular. She describes the forces of technology in the change process.

Six major elements form the scenario she envisions: things will
get worse before they get better; scientists will not go out of their way to assist libraries or publishers; there will be new electronic entrants in the scientific communication arena; libraries will be squeezed even farther; publishers will be squeezed as well—and the scientists will emerge as the winners!

In her future world, not all publishers or libraries will survive, but the flexible innovative publisher will and, indeed, may flourish.

Lucker discusses trends in research library buildings, declaring his strong bias that the sci-tech library of the future will continue as a physical space—“with real books and real people inside.” While there will be many changes in the ways people acquire and use information, “buildings housing physical collections with convenient spaces for users to consult those collections, . . . for self-education and discovery outside the classroom and laboratory will continue to be important elements.”

He believes that scholars will continue to use print materials—publication on paper will increase rather than decrease and scholars will depend upon libraries to acquire, preserve, and provide access. Libraries also will be the source of information about new technologies and the access points to non-print materials outside the immediate holdings.

Lucker describes the achievements in academic and research library buildings constructed or added to in the past five decades of 1950-1990—open stack arrangement; reduction in size and number of large, high-ceilinged reading rooms; large areas for public services; and the increasing need to accommodate the proliferation of new information formats. He speaks to the expansion and growth of library storage facilities. He acknowledges the shortcomings of these trends as well: site limitations for expansion; physical and mechanical problems; and lack of flexibility for relocation of functions with buildings.

Lucker concludes his discussion with a picture of the research library in the 21st century. “His” library will continue to be a blend of collections of printed information, a place for study, exchange of ideas, and reflection, as well as access to a wide range of technology-based services and networks. As we plan for new facilities or adapt existing ones for future generations, responsiveness to trends coming from higher education and technologies must be recognized. The
pace of change will require great flexibility in the way libraries organize themselves and utilize their resources to meet user needs.

Among the challenges facing sci-tech libraries in providing effective information services to scientists, Hurd sees the increasing interdisciplinarity of scientific research. The emergence of interdisciplinarity fields since World War II is the result of universities, federal agencies, and corporations seeking to find solutions to societal problems, environmental, and health care issues. Examples of such areas include materials science, molecular biology, bioengineering, environmental chemistry, and the intriguing field of "environmental mathematics." According to a recent report in The Chronicle for Higher Education, environmental mathematics is the attempt to get mathematicians to connect again with the natural world, conducting investigations of the pollution caused by idling auto engines, the decline in nesting sites for the spotted owl or the settling of sulfur dioxide in the human lung.

Hurd's paper examines the development of this interdisciplinarity phenomena and the implications for sci-tech library collections and services. She describes the lack of agreement on basic terminology, the impact of technology transfer, and most importantly, the changing information needs from those of twenty years ago when research projects were more easily contained within the boundaries of single, recognized disciplines—and the associated information resources organized on a discipline-based model.

If scientific research continues to experience shifting disciplinary boundaries, Hurd sees fundamental changes in the entire university infrastructure to which university library organizations patterns will have to respond. Centralization of library resources, less duplication of costly journals, abstracts, and indexing services are already taking place. Hurd suggests other possibilities for collections and services to assist the growing number of interdisciplinary researchers.

Pinelli and his co-authors offer a view of information dissemination changes needed in the federal sector. This paper discusses the relationship between the development and application of a federal technical policy and the effective diffusion of the results of federally funded R & D.

Recent federal technical policy initiatives designed to improve, nurture, and stimulate technological innovation to strengthen eco-
nomic competitiveness of the U. S. requires the adoption of a “knowledge diffusion model” that prescribes “active” intervention between the producers of information, transfer agents, and users of technical information. This knowledge diffusion model is predicated on the development of interpersonal communication linkages (as opposed to dissemination and access activities), including the creation of user-oriented products and services.

Convincing evidence exists that the U. S. lacks effective knowledge transfer of federally sponsored R & D research results. Reasons for this include the passivity of sci-tech librarians. Most sci-tech information centers follow a model tied to the information artifact, providing documents instead of supplying information. Sci-tech libraries must become strategic information resources in the STI transfer process, emphasizing innovative, value-added services tailored to meet the individual needs of each user group.

Pinelli and colleagues emphasize that U. S. technology policy must be based on the belief that the production, transfer, and use of STI is inextricably linked to successful technological innovations; if the U. S. is to remain competitive in the global market place in the 1990s and beyond. That process of innovation is best served by a “knowledge diffusion” model.

Wiggins reflects that “collection development in academic science libraries has bordered on the chaotic in recent years.” This is as a result of our “failure to adequately define cooperative collection development and access, and in part from the lack of a coherent plan for integrating machine-readable sources into our collecting profiles.” While significant changes in the scholarly information process have turned our attention to the promises of accessing information rather than dependence on the amassing of costly on-site collections, it must be recognized that published material will continue to be of utmost concern to research faculty as we work our way through the last decade of the 21st century.

Wiggins’ paper discusses the issues that sci-tech librarians must consider in making the transition from collections that are primarily print-based to resources substantially in electronic format—in particular, the question of access to serials vs their on-site ownership. He offers some ground rules which must be accepted by the library before order can be achieved in collection development. He suggests
that our definition of access must be enlarged to include all document delivery techniques for easy delivery of information from the location in which it is held to the user. His paper discusses a number of existing delivery services. More futuristic is his scheme to develop a database of document images—a library of electronic articles tailored to the needs of a particular clientele. Wiggins claims that the science library community is ready to make a fundamental change in collecting policies—away from exclusive collecting of journal volumes to a more flexible collecting profile that allows the permanent addition to collections of individual articles from journals not locally held. To get from here to there, Wiggins identifies several crucial steps: libraries and publishers must confront and solve the copyright questions associated with this material, librarians must enlist the support of users in the design of such a local document archive and indexing system, and work closely with the scientific publishers to reassess the values and costs in scholarly publishing, including this concept of the journal article as a marketable unit.

The promise of FAX. How will it affect what is held locally? Author Stankus speculates on the meaning of FAX for subscription policies. While some journals of high demand will always be locally held, subscriptions to others will likely continue only as long as subscribing to them is cheaper than repeatedly requesting FAX copies. Librarians have come to realize that it is cheaper to give free fax delivery of individual articles to their faculty than to purchase and maintain voluminous and expensive titles. With the advent of document delivery via licensed fax delivery services, librarians will have much more freedom to decide which journals to hold locally, which subscriptions will be accessed from remote collections (or databases) on a pay-as-you-go-basis. Stankus provides some prediction as to which titles these are likely to be.

Using long-term citation data from 1979-1989, Stankus examines the life science literature to support his thesis. This material is complex and expensive; its management relevant to many academic, medical, and corporate biotech special library collections. Speciality by speciality, from biochemistry to pharmacology, he examines impact factors, cost, and “citation by the leading journal” data of eleven major areas of the life sciences, in predicting their fate.

FAX is fast; it can be made free of legal worry rather easily; it is
cheap and is a “natural” in this electronic age. Offering liberation to science libraries from the expensive, lesser-used titles, Stankus also suggests that it will provide publishers a way to determine from sales statistics of individual articles which authors and/or laboratories are generating readers and royalty income.

Complementing the Stankus paper and seeking an answer to the dilemma of rising costs and diminishing resources for library materials, McFarland conducted a comparison testing of document delivery services in lieu of local holdings. Using the Chemical Abstracts Document Delivery Service and nine commercial science related document delivery vendors, this pilot project compares a number of service related parameters. The aim of this study is to assist science librarians in selecting vendors that best meet the needs of users in disciplines of biology, chemistry, earth sciences, math, and engineering.

To facilitate the vendor selection process, McFarland compares a number of publisher service parameters and policies. Comparables included in his investigation include: coverage, methods and hours documents can be requested, basic services, costs, fill rates, turnaround time, reliability of copyright clearance, and special services.

Recognizing that a single vendor is unlikely able to satisfy all requests, the study assists the librarian in identifying primary and secondary vendors which together would satisfy a high percentage of requests. The author further acknowledges that reliance on a conventional ILL service will continue to be necessary. McFarland adds a concluding cautionary note that while “document delivery has been around a number of years in one form or another, the advent of the plain paper FAX machine has dramatically facilitated the ease with which documents can be transferred, necessitating librarians to exercise greater caution than ever to honor copyright compliance. If abuse becomes evident, FAX access and document delivery by any mode will be denied.” A wise admonition.

***

Change will affect all areas of the scientific process: the way in which research is conducted, communicated, transferred, stored, and delivered. No player will escape the impact of these changes: researcher, librarian, information manager, publisher, user. It is vital
that the sci-tech library respond to these changes, that science librarians work closely with scientists, engineers, publishers, vendors, and users, cooperating and collaborating to link resources and users of information, to plan and develop effective information systems that provide information conveniently and cost effectively.

Cynthia A. Steinke
Editor

REFERENCES


The Needs of Science and Technology

Robert M. Hayes

INTRODUCTION

This paper discusses the needs for scientific and technical information in support of research and education, engineering design and manufacture, government and public policy decision-making. It will first set a context by identifying issues of current importance in the delivery of scientific and technical information—changes in patterns of research and application, the increasing importance of computer-based information in the process of research and of application, the emergence of electronic forms for information delivery, the resulting changes in the very pattern of communication of the results of research, the development of new institutional means for management and delivery of information.

The paper examines the critical stages in the research process and in translation of the results of research into application, identifying the different and changing roles of information in each. It will illustrate with specific research programs drawn from those current priorities identified by the National Science Foundation and the National Institutes of Health which emphasize the integration of

Robert M. Hayes is Professor Emeritus, having retired after fifteen years of service as Dean of GSLIS at UCLA. Dr. Hayes received the PhD in 1952 in Mathematics from UCLA. From 1949 until he joined the faculty of GSLIS/UCLA in 1964, Dr. Hayes worked in government and industry, and founded a small consulting and research company, Advanced Information Systems. In 1969, Mr. Joseph Becker and Dr. Hayes formed Becker & Hayes, Inc. of which Dr. Hayes was Vice-President until becoming Dean of GSLIS/UCLA in 1974.

This paper was originally presented by invitation to the 10th International Seminar held at Kanazawa Institute of Technology Library Center, Nonoichi Ishikawa, Japan, June 3-5, 1991.

© 1992 by The Haworth Press, Inc. All rights reserved.
information into the research process; among them, special emphasis will be given to those that embody the use of digitized images because of the current and future significance they have throughout the range of scientific and technological disciplines.

The paper concludes by discussing the implications for the research library community of these changes in use of information in scientific research and technology. The roles of the library in support of science and technology will be contrasted with those in the arts and humanities; differences among scientific disciplines themselves in the use of library-based information services will be identified, and they will be contrasted with those of more practical disciplines that apply the results of research or that make public policy about them. Developments in libraries that implement those roles will be identified—including inter-library networks, electronic means for access to information and for publication, and inter-mediation in use of information resources. Specific attention will again be given to digitized images, with emphasis on the potential role of the research library in management of these forms of research information.

GENERAL CONTEXT

The general context for such review is a "new era" in which we are seeing significant changes in patterns of research and the emergence of electronic means for information access and distribution. I will use a specific form of information, digital imaging, as a means for illustrating the range of issues involved because of its intrinsic importance and the extent to which it exemplifies the issues.

Differing Patterns of Research and in Application of Its Results

There are substantial differences in needs for information across the several disciplines—natural science, social science, humanities, and professions—and, for each, across the functions in research and graduate education, undergraduate education, development and
application, and public policy decision making. I will comment on
the broad range of the two dimensions, discipline and function, but
my primary focus will be on the natural sciences (the physical and
biological sciences) and the functions in research and graduate
education that are central to the mission of the major research uni-
versity.

Differences among the disciplines. As a starting point, then, I
want to comment on differences among all academic disciplines in
roles of information access in support of research and graduate
education. In doing so, I am not attempting to preempt the separate
discussion of the needs of humanists and historians, but I do want
to set a large frame of reference in which to examine differences
among disciplines and functions. I identify five reasons to need
access to information in support of research and graduate education:

1. information and, especially, the records containing it are them-
selves the focus of the research;
2. information contains data needed in the research;
3. information provides an audit trail documenting the progress
of research;
4. information provides communication about current results;
5. information results from and embodies the processes of data
acquisition, analysis, and presentation.

With respect to the first two of those reasons, the natural sciences
and the professions of engineering and medicine differ dramatically
from the humanities, social sciences, and other professions. As a
generality, I think one can say that, by their very nature, the hu-
manists and the social scientists absolutely require access to the
records of the past for the basic raw material of their research. In
literature and art, those records are frequently the very substance of
the research; in history and economics, they are the source of the
primary information on which the research depends; in anthropolo-
gy and archeology, they provide the crucial evidence. For the pro-
fession of law, they are the absolute source of data. There are some
natural scientists, taxonomists in particular, that do depend centrally
on analysis of data from prior records and comparison both among
them and with newly acquired data; those are the exceptions,
though. In general, natural scientists are focused on the acquisition of new data rather than the analysis of existing records. For them, the records of the past are peripheral to research; at most they turn to the prior record for data to compare with that being currently acquired. Research for the professions of engineering and medicine is essentially of the same character as that in the natural sciences; practice in those professions, though, is as we will see quite different.

Across all disciplines, of course, the records of the past provide the audit trail that documents the progress of research. This role of information access is built into the entire process of scholarship, embodied in the traditions of citation and providing the basis for recognition of the credit for first discovery of new concepts. The natural scientists are as scholarly in that respect as any academic.

The role of personal communication in information access appears to be much more important to research in the natural sciences than it is in the humanities and social sciences. For the scientist, speed of communication is crucial since the results obtained by one colleague immediately influence the work of others with similar foci of research. This is reflected in the historic importance of preprints, telephone, and now electronic communication to those sciences. It is shown in the extent to which research teams are commonplace in the natural sciences and in the "invisible colleges," made up of groups of scientists throughout the world that collaborate with each other even while competing. It is exemplified by the prevalence of multiple authorship of scientific articles. Are there differences among the disciplines of the natural sciences in the importance of this reason for information access? Apparently not. Across the board, they all seem to partake of the same pattern. In contrast, the humanities and social scientists appear to be less dependent upon immediacy in communication. They tend to do research as individuals rather than as teams, multiple authorship being relatively rare.

It is in the fifth reason for information access that the greatest difference arises between the natural sciences and the humanities, with the social sciences occupying a middle ground. As I have already stated, the natural scientists are focused on the acquisition of new data and, in consequence, on the analysis of them for pur-
poses of data reduction and for derivation of underlying patterns from them. To an extent, this is a result of the essentially quantitative nature of the data, which makes computation and the use of mathematical models necessary. In contrast, the humanists and, to a lesser extent, the social scientists deal with more qualitative data for which measurement, computation, and mathematical models have a dramatically different character.

Differences across the functions. Turning now to the other dimension, I will discuss five sets of functions: (1) research and graduate education, (2) undergraduate education, (3) professional practice, (4) engineering development and commercial application, and (5) public policy decision making. In discussing the differences among disciplines, I have specifically focused on the first of these, research and graduate education, since that is where those differences are most evident; for the other four functions the differences arise less from the nature of the discipline and more from that of the function.

Access to existing, published information ought to be crucial to undergraduate education. The facts are, though, that the amount of material needed to support this function is really quite limited as is, unfortunately, the extent to which this level of instruction generally requires that the student learn the processes in access to recorded information. I wish it were otherwise.

Preparation for professional practice is, in the United States, at the graduate level in virtually all fields. As a result there is usually thorough indoctrination into the importance of use of the most current information. For medicine, this is reflected in the vital role played by the National Library of Medicine in the implementation of library services, the creation of the Regional Medical Library Network, the support given to the development of medical informatics. For library and information science, it is the very substance of concern. For law, it is the essential ingredient of practice. For other professions such as management, social welfare, architecture, education, etc., there is general recognition of the importance of information access.

Access to existing, published information also ought to be crucial to engineering development and commercial application. The amount of information available and useful in support of this func-
tion is immense, and the tools for access to it, as I will be discussing, are available and powerful. The kinds of information range from technical to economic, from scientific to marketing, from basic research to public policy, from numerical to textual. All of them are potentially of value and even of necessity in bringing scientific and technological results to the service of society. The facts are, though, that despite all of the investment that has been made in the tools for information access, the level of use across all industries is substantially less than it ought to be if we are to have maximum productivity.

The same situation exists with respect to public policy. Increasingly, our society must deal with the effects of scientific and technological development. To some extent, the political process does provide means to draw on relevant published scientific, technical, economic, and social information. In the United States, we have the Congressional Research Service, the Office of Technology Assessment, the National Library of Medicine, and the National Agricultural Library that serve our political apparatus well in assuring access to such information. Even at state and local government levels, there are counterparts of such agencies in state and community public libraries. Again, though, the facts are that the level of use of information resources in public policy decision making is probably less than it should be.

The Effects of the Information Technologies

The information technologies—the computer, its associated means for storage and display, and telecommunications—have had a truly revolutionary effect upon research in every field. The computer has dramatically changed the means for word processing; it is the crucial tool for computation; it plays an increasingly important role for control in data acquisition; it provides the means for image processing; and it has changed the very nature of telecommunication. The table (see Table 1) on the following page summarizes representative data about computer equipment at different levels of processing capability and the uses to which they typically are put.

To an extent, the differences among disciplines are reflected in the relative emphases given to the various uses of computers by researchers by each of the groups of academics.
### Table 1. Representative Data for Different Levels of Computers.

<table>
<thead>
<tr>
<th>COMPUTER TYPE</th>
<th>TYPICAL CONFIGURATION</th>
<th>TYPICAL APPLICATIONS</th>
<th>EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microcomputers</td>
<td>Microprocessor based Single tasking, user &lt; 100 Mbytes harddisk &lt; 2 Mbytes memory &lt; 10 MIPS $1K - $10K</td>
<td>Word-processing Spreadsheets Network access Personal databases</td>
<td>IBM PC/AT, PS/2 Macintosh</td>
</tr>
<tr>
<td>Super-micros &amp; Workstations</td>
<td>Microprocessor based Multitasking Limited number users Windowed environment Diskless to 1 Gbyte 4 - 64 Mbytes memory 2 - 25 MIPS $5K - $100K</td>
<td>CAD/CAM CASE Software development Graphics Desktop publishing</td>
<td>SUN 3/50 IBM RS/6000 DG AViiON NeXT SPARCstations</td>
</tr>
<tr>
<td>Minicomputers &amp; Super-minis</td>
<td>Multitasking Multiuser, but &lt; 250 &lt; 10 Gbytes &lt; 64 Mbytes memory &lt; 50 MIPS $100K - $1M</td>
<td>Time sharing Network server Small businesses</td>
<td>Sun File Servers Tandem Nonstop IBM System 38 DEC VAX</td>
</tr>
<tr>
<td>Mainframe</td>
<td>Multitasking Multiuser (1000+) &lt; 100+ Gbytes &lt; 1+ Gbytes memory &lt; 100+ MIPS $1M to $10M</td>
<td>Large-scale MIS Transaction process Large databases Large batch process</td>
<td>IBM 4381, 3090 Tandem Cyclone Amdahl 5995 Hitachi EX</td>
</tr>
<tr>
<td>Supercomputers</td>
<td>Vector processing 100+ Gbytes 1+ Gbytes of memory 100+ MIPS Up to $20M</td>
<td>Image processing Vector processing Simulation modeling Real-time animation</td>
<td>Cray X/MP NEC SX-3</td>
</tr>
</tbody>
</table>

**Word processing.** The ACLS survey of scholars’ views in 1986 showed that among the humanists and social scientists the great preponderance of computer use was for word processing, maintaining note files, preparing tests, and compiling bibliographies. All other types of use were of substantially less importance. Of course, word processing is as valuable a tool for the natural scientist in preparation of manuscripts as it is for the humanist or social scientist.

**Computing.** But for the natural scientist today, research would be
impossible without the computer’s ability to perform mathematical calculations; statistical analyses have been the overwhelmingly important uses. Since the advent of the digital computer, the practice of scientific research has increasingly relied on use of numerical analysis of measured and synthetic data. Computational predictions and analyses are typically of interest in the fields of atmospheric and geophysical research, biology and molecular biology, computational and empirical chemistry, computer and engineering, mathematics, medical research, theoretical and applied physics. The common element is the very large amounts of data generated on a regular basis and the necessity to comprehend the numerical description as a complete entity. Literally, the need is to get the “big picture” that provides a lucid, intuitive, and global grasp of the data as it evolves. Mathematical and statistical analysis have provided the fundamental tools for getting that big picture.

Digitized imaging. While getting the “big picture” may be based on mathematical modeling and data analysis, the great step forward today is the recognition that conceptualization is even more effective if supported by real pictures. The term visualization has been applied to this process. It involves comprehending large amounts of structurally complex, multidimensional and time-dependent data through the use of visual images. In some contexts, the data represent familiar things; for others, they are things that cannot be seen by the naked eye even though they may be real; and for others, they are purely conceptual–algorithmic, mathematical and logical constructs.

As a result, the most dramatic use of computers in research and practice today is for “digital imaging.” Based on data from Science Citation Index, from 1965 through 1984 the number of published articles concerned with “imaging” essentially doubled during each five-year period but within the most recent five year period, it increased ten-fold! And that rate of increase continues today.

Parenthetically, it is worth noting that digitized imaging and the management of the resulting files has significance to the practical world of business as well as to academic research. It is estimated that today the U.S. market for electronic image management to handle the flood of commercial documents is worth about $250 million, with about 1,500 installations, but that it will expand at
much more than a 35% compounded annual rate of growth. The forecast is that by 1995 the market will be at $5.8 billion, with about 30,000 installations. This means that there will be a substantial commercial base upon which digital imaging technologies can experience economies of scale so that the costs for academic uses and for libraries can be less. The situation is comparable to the role that the consumer market for compact disks has played in making CD-ROM economic.

I will be using digital imaging to illustrate specifics on the role of information access in scientific and technological research, so I would like briefly to summarize the hardware and software technologies that are important to the creation, management, and use of digitized images. First, the hardware:

Scanners are the means for creating digitized images from sources external to the computer itself; they range from simple hand-held scanners to those in satellites and medical equipment, to those in nuclear accelerators. Storage of scanned data or of images generated by algorithms is of course necessary for all subsequent processing. For a typical 8-1/2” by 11” page, scanned at say 300 lines per inch, that involves storage of about a mega-byte of data for a pure black and white image; for typical text pages, the data are so redundant (with large areas of white space, for example), that they can be compressed, usually by a factor of ten; but for gray-scale or color, the storage requirements are substantially greater. Even with compression, though, the means for storage of any significant number of images must have very high capacities. Fortunately, both magnetic and optical means for storage are now readily available with capacities completely consistent with the needs in image storage. CD-ROM optical disks in particular provide capacities between 500 and 1000 mega-bytes, sufficient to store as many as 10,000 compressed images.

Processing of image data requires computing capacities of a high order. Indeed, primary among the rationales for development of super-computers was the need for that level of computing power. However, we are beginning to see speeds, internal storage capacities, and functional capabilities at the microcomputer level that make it possible to consider use of them for a large range of image processing applications. Display of image data requires exceptional
resolution but again the technology has advanced so far so rapidly that today we have screens that provide a quality approaching that of the printed page, with truly remarkable mixes of color.

Finally, communication of digitized images, given the number of bytes required and the frequency with which images may be generated, requires truly exceptional bandwidth. Unfortunately, while the technology is now available to provide that kind of capacity, making it available requires an infrastructure that many developing countries and even some highly industrialized countries do not have. In fact, of all the technologies needed for utilizing digitized images, communication may well be the one which poses the greatest problems, since it requires a national investment, whereas all of the others require only local institutional investment.

Now, turning to the software:

*Interfaces* are the means by which the user may interact with the computer in the use of digitized images. Increasingly, we have seen the use of GUI (graphical user interfaces), pioneered by APPLE computer but now embodied in WINDOWS for use with IBM personal computers; the GUI exemplify the role of interfaces at the simplest level. At more complex levels, we see capabilities to manipulate images, combine them, navigate through them. *Algorithmic production*, exemplified by CAD/CAM, for example, is a set of software that provides means to generate images from specifications.

*Scene analysis* is a process in utilizing a succession of images that provide a continuity in representation of an object or an event. The classical example, of course, would be a "scene" in a motion picture—a succession of frames taken from a single camera setting; others might arise in the images generated in a "CAT-scan" of the human body or in the frames in a succession of satellite images. The tasks in scene analysis are first, to identify when a succession of images are related as a scene and, second, to manipulate the set of images constituting a scene as an entity.

*Spectral analysis* is a set of processes for deriving characterizing parameters for an individual image and as the starting point for identifying components of the image—the first stage in *pattern recognition*. *Image enhancement* is an example of pattern recognition, in which areas of the image are examined and modified to bring out
details, correct for problems (in exposure, for example), and identify and highlight identified patterns. *Optical character recognition* is the most well developed capability for pattern recognition, permitting the computer to identify typewritten and printed characters.

**Acquisition of data.** In August 1990, the Council on Library Resources submitted a report to the National Science Foundation on *Communications in Support of Science and Technology*. In that report, Donald Langenberg discussed Information Technology and the Conduct of Research: The User’s View. He commented on the use of technology in control of experiments and in the generation, acquisition and storage of data. He paid specific attention to medical scanners and superconducting-supercolliding scanners.

**Large-Scale simulations.** In that same report, Kenneth King stated that the major future challenge will lie in building complex multi-dimensional, multi-media simulations. Pointing out that computing requirements are related to the kinds of objectives, he estimated that it takes 1 mip (million instructions per second) to control words, 100 mips to control pictures and sound, 1000 mips for artificial intelligence, and $10^{12}$ mips to simulate physical systems in real time.

**The Changing Means for Communication and Information Access**

The CLR report was commissioned by NSF to explore the future form of scientific communication. How was it integrated with the research process? What was its role in maintaining the historical record—the “audit trail” of the progress of research? Could electronics solve the problems with escalating costs of scientific publication? What were the effects of various electronic alternatives? What were the characteristics of scientific communication—its forms, from personal to external, the available options, the relationship between form and utility?

In his summary, Martin Cummings reviewed the background for the report’s conclusions, referring to the influence of new technologies and their use as means for improving the existing information infrastructure—science libraries, scientific publishing, computers and networks, and the effects of new technology on users.
The potential represented by electronic distribution. There are some who see electronic communication as a revolutionary new medium; they take what might be called the "McLuhan viewpoint" that "the medium is the message" and that there will be dramatic changes in the entire process of communication. The term "skywriting" has been used to suggest that information will be out for everyone to see, open to peer commentary on networks, global and unconstrained by time, space, or distance, collaborative, inter-media and multi-media. They see electronic communication replacing print, and revolutionary.

Others see electronic communication as an imitator of print publication but with differences. They see it as the journal "unbundled"—article based rather than package based—with greater flexibility in delivery either individually or recombinated. They see it as more rapid in direct delivery, with the most recent work on a subject immediately available, searchable, and linkable. They see the results as increased diversity in which various options co-exist, and evolutionary.

From the standpoint of scholarship, this process clearly facilitates communication among a small peer group. But, as the experiments with formalizing the invisible colleges undertaken some twenty to twenty-five years ago demonstrated, the improved effectiveness in communication is heavily outweighed by the inherent limitations. The peer group is by no means the only or even the most appropriate context for research progress and especially evaluation. Research should be of importance beyond the limits of the peer group; it should be subject to evaluation and assessment by other minds, other tests of validity. The cold, clear light of day needs to enter, but the invisible college, by its very mode of operation, prevents that kind of "sunshine" effect.

Speculations on cost reduction. Some have looked to electronic publication as a means for reducing the costs of journals. This has been especially of interest in the light of dramatic rates of inflation in prices of scientific and technical journals. Others have urged that the universities should get into the business of electronic information distribution, with the view that they are both the primary sources for generation of published information and the primary users of it. They argue that as it is now the universities are pay-
ing twice for their information—once for producing it and then later for acquiring it.

Underlying these developments is the perception that electronic publication will introduce dramatic changes in the process of publication and in the costs associated with it. In fact, in some disciplines (Artificial Intelligence being one example), the concept is that the online electronic journal should replace the traditional printed page. In other contexts, the ease of micro-computer based "desk-top publishing" have led some to see it as replacing the formal publisher.

It is argued that electronic publication is more rapid, more immediate and direct, more responsive to change. But whether that's the case or not, electronic mail is a reality of scholarly, peer group communication today. In fact, increasingly, scholars are sending diskettes to colleagues whenever there is need to develop joint reports. The Electronic Manuscript Project of the Association of American Publishers was established in recognition of this phenomenon, and has had the intent of establishing standards for encoding of the text in such media to facilitate all aspect of subsequent publication.

However, despite the potentials, publishers for good reasons have been slow to adopt electronic publication. International journal publishers, in particular, cannot simply switch from paper to electronic delivery; they must accommodate both distribution methods and would incur increases in costs with a dual system. As a result, the steps taken currently are experimental. For example, there are efforts to augment existing forms of electronic distribution with increasing amounts of substantive content. In particular, tables of contents are being added to bibliographic information; CD-ROM publications are becoming more wide-spread; abstracts are becoming available more rapidly.

*The processes of electronic publication.* It must be recognized, furthermore, that the production of a journal involves processes and related costs that will be incurred no matter who does the distribution. First is identifying and encouraging potential authors. In a sense, this is the most important of the roles since it creates opportunities that might not otherwise be there. It is surprising that even from the academic world, where publication is the life blood, there
are difficulties in encouraging persons to publish. Any commercial publisher focused on scientific, technical, and professional books and journals, depends greatly upon academic authors, but it is difficult to identify persons capable of writing good books and journal articles. Even for professional and scholarly journals, the number of authors and suitable articles is far less than is needed. The task of the publisher and of the editor is important and difficult; it involves constant discussion with potential authors, encouragement to them, and assistance is bringing manuscripts to publishable state.

The publisher is responsible for quality control. That arises in the identification of authors, of course, but it also arises in creating publishable works of suitable quality. Manuscripts of books and even of scholarly articles all too often are poor in their organization, poor in their writing, and even poor in their substantive content. The publisher, through its editors, serves as a means for bringing such poor manuscripts to a level of quality that justifies publishing. The processes of review, both internal and external, of copy editing, of preparation for publication all serve as means for accomplishing that objective. For major journals in the natural sciences, referees of peers are part of the review process but by no means the totality of it.

The publisher is responsible for producing a publishable product. The facts are, with all due respect to the presumed ease with which one can use electronic publishing, creating a publishable product is by no means a trivial task that can be performed by clerical staff or even by the author. It involves a high level of technical expertise, experience, and knowledge of the needs in marketing and use of publications.

The publisher is responsible for marketing and distribution. This role is of importance not only to the publisher, as the basis for its profitability, but to the author, to the users, to the libraries that will preserve and assure access. Evaluating the appropriate market, the size of it, the kind of package that will appeal to it, the best means for informing it of availability—these are all essential.

The point of course is that the publisher has critical responsibilities in the control of quality of content, control of the process of production to assure quality of product, control of the process of distribution to assure that the product will be available and deliver-
able at an appropriate price and with suitable response time. Where are those elements of control in electronic publishing? There are none to this point in time, and the result is the likelihood, even certainty, of degradation of quality in every respect and of limited availability and distribution until they are developed. That is the primary purpose of the several experiments now underway.

An inherent danger in the electronic modes of communication though less so in the printed forms of desktop publishing is the likely loss of what is called "integrity of reference." Scholarly communication has built up an important tradition of citation. It reflects the fact that in all areas of research—humanities, natural sciences, social science—we progress by building on the past. And we acknowledge our debt to the past by citation to it. By doing so, we assure that our sources can be checked, verified, validated. But that implies that material so referenced, so cited must be available for checking, verifying, and validating. What happens if the source data has been erased or, worse yet, altered since it was used? The entire structure for scholarly progress would collapse.

Current forms of electronic delivery. Most journals become machine-readable at some point in the research process. As a result it is easy to produce by-products of their publication. Today, operational systems for electronic delivery of information include online access to databases through use of national services, mounting of databases at campuses for access through online public access catalogs, and increasing availability of parallel CD-ROM products. Just to illustrate the last of those options, Ei (Engineering Index) has, since 1989, released a series of publications on CD-ROM: Compendex*Plus (in August 1989), ChemDisc (in July 1990), EEDisc (in November 1990), Energy/Environment Disc (in December 1990), PageOne (in January 1991). There are increasing numbers of full-text (but non-graphic) databases online, some publishers’ experiments (e.g., ADONIS), an array of document delivery options (e.g., mail, fax, electronic), steadily increasing use of electronic bulletin boards.

Beyond what publishers are doing, there are a number of other experiments underway, including about a dozen refereed or lightly refereed electronic journals, two university electronic preprint services in development, and plans for publication by AAAS of a fully
refereed electronic journal for distribution through OCLC starting in 1992. These developments reflect the efforts of some existing paper publishers converting to electronic, but they are primarily the work of intermediaries, value-adders who deliver electronically. For example, there are now several document delivery services, CARL (Colorado Association of Research Libraries), University Microfilms Incorporated, re-packagers (e.g., Maxwell Electronic), university-based electronic publishing (e.g., pre-print services in high-energy physics and mathematics), the Ohio State University experiment in FAX delivery.

Underlying many of the experiments and plans are the expectations for implementation of the National Research and Education Network (NREN). Legislation (Public Law Number 102-194) signed by the President December 9, 1991, provides $638 million for the funding of High-Performance Computing and Communications, of which $92 million is slated for NREN. The Coalition for Networked Information (a joint venture of the Association of Research Libraries, CAUSE, and EDUCOM) was initiated in early 1990 to assure that the interests of the research university community were represented in planning for implementation of NREN.

In the meantime, there are increasing numbers of ad hoc interconnections among online public access catalogs of major university libraries. In order to bring some order to those developments, a task force on Electronic Data Interchange (EDI) was established in 1990 to develop guidelines, procedures, and policies for computer-to-computer exchange of structured documents between independent organizations.

Expected pace in future development. With that as the current context, what will be the pace of future development? It seems likely that by 1995 the problems in pricing, ownership, standardization, and distribution will largely have been solved. There will be a number of "unbundled journals," in which the article becomes the basic unit, not the issue, with resulting inroads into paper subscriptions for journals, especially in the natural sciences, and a slowing of the rate of growth in the number of paper journals. There will be a movement of libraries from acquisition to access and direct delivery to the end-user. It has been estimated that there may be over 100 refereed electronic journals and a significant num-
ber of university based preprint services in collaboration with societies. However, it is also likely that there will be many alternative sources, with battles over pricing strategies and significant problems over copyright. The national network will still be heavily subsidized, at least for use in the natural sciences.

The potential in 2000 is based on the expectation that computer usage will be pervasive. There will be parallel availability of electronic and print in a mix about 50/50, with hundreds of electronic "journals." There will be a re-structuring of secondary services combined with an increasing number of new service niches—scanning services, archive services, ... There will be increasing stratification between richer and poorer universities, users, and nations.

The Institutional Means for Management and Delivery of Information

The effort to deal with potentials represented by electronic publishing involves many participants: publishers, libraries, computing facilities, communications systems, networks, and a wide range of "niche" services. There is great uncertainty about the relationships among these participants, with continual jockeying for position among them, with cooperative arrangements developing and then disappearing, with established agencies attempting to preserve and build on their existing capital resources, and with newer agencies attempting to displace them.

Several issues of concern exacerbate these interactions. Primary among them is the reconciliation of intellectual property rights—the balancing of the rights to use information resources with the rights to control the use of them. A related issue is the determination of appropriate policies and formulas for charging and pricing, for service arrangements, for licensing. Another issue is the decision about when it is appropriate to set standards and how standards should relate to proprietary interests.

Two recent court decisions in the United States illustrate the current complicated state of these issues. The first is the decision, in which the courts ruled in favor of LOTUS 1 2 3 in its suit against Paperback Software, thereby greatly extending the scope of copyright protection to include not only the programs but the "look
and feel” of the interface they provide; the result is a substantial restriction on other developers of programs who wish to use that interface as almost an industry standard. The second is a decision in which the U.S. Supreme Court ruled that directories or compilations of facts may be freely copied and republished unless they display “some minimum degree of creativity,” emphasizing in doing so that the main purpose of the copyright laws is “not to reward authors but to promote the progress of science and useful arts”; the result is potentially more far-reaching, placing at risk the entire electronic data distribution business with the likelihood of continuing litigation to determine what is meant by “originality” in the presentation of compilations of data.

THE ROLE OF INFORMATION ACCESS IN THE STAGES OF SCIENTIFIC RESEARCH

I now turn to the needs for information access on the part of researchers in science and technology. In the CLR report to NSF, Helen Gee discussed the users and uses of scientific information resources, suggesting that we may need to replicate and to update findings from the studies of the 1960s. She called for one or more broad-based, multi-institutional, multi-discipline studies, using experimental approaches, critical incident studies, and interviews with a variety of natural scientists. She posed a number of questions that such studies might answer: How do the natural scientists obtain the information they need at different stages in planning and conduct of research? What is the level of familiarity with different resources? What is the availability of electronic resources? How do they use intermediaries? What are the effects of collaboration? What are the needs in local communication? How much time do they spend in using different resources? What are the effects of limitations on access to information?

I wish I could present to you the results from such studies, but I cannot; they still need to be done. At best I can only conjecture about the answers to those questions, based on my knowledge from prior studies and the descriptions available from current literature. Among the latter, the Guide to Programs of the National Science
Robert M. Hayes

Foundation provides some tangible evidence in the emphasis it gives to information needs, and I will use it to support some of the conjectures I will present.

To provide a framework for discussion, I will postulate a sequence for the conduct of scientific research, obviously only one among the many that will be found in practice. It is based on a series of stages, in each of which information access plays a dramatically different role. For each stage, I will provide a characterization of its role, discuss the impact of digital imaging upon it, illustrate with examples from the current NSF Guide to Programs, and conclude with a discussion of the roles of the several means for information support—computers, telecommunications, and libraries—as may be appropriate.

Stage 1: Data Acquisition

Characterization of the stage. Consider a researcher who observes an anomaly in some data, arising from a mismatch between the data and predicted values based on prior theory. Such an anomaly presumably could represent a failure of the theory, and the researcher is likely to speculate about that possibility. But it could result from inaccuracy in measurement, so normally the researcher will attempt to replicate the acquisition of data, being as accurate and precise as possible. As an alternative rationale, the researcher may simply be curious about the nature of some phenomenon and wish to determine its quantitative characteristics. In either event, the first stage in this model of the research process involves data acquisition. Surely this is the most central role of information in the scientific research process. As a result, the importance of this stage is well recognized in any description of the process of scientific research; some have even gone so far as to say that it is the essential stage, the thing that distinguishes science from other academic disciplines.

The example of digital imaging. Sometimes data are acquired by conduct of experiments—observation of collisions of high-energy particles in an accelerator, of the progress of a chemical reaction, of a succession of functions in experimental animals. Sometimes, though, it results from the scanning of natural phenomena—of the
human body, of the earth and the other planets by satellites, of
geological structures by a variety of means for observation. Such
processes differ from the observation of experiments in the fact that
the things being observed are not being controlled, so the range of
phenomena observed is dramatically greater, and in the fact that
they can occur continuously. Scanning can acquire data at truly
mind-boggling rates, so the generation of images from them has
become a crucial means for presenting otherwise overwhelming
amounts of data in compact, visual form.

Increasingly, the very nature of the data and of the process for
data acquisition requires the use of digital imaging. For example,
AT&T Bell Laboratories has recently announced the invention of
a more powerful optical microscope that should vastly improve the
examination of living matter, including tiny structures inside living
cells. It uses a strand of extremely clear optical fiber tapered to a
tiny tip. The tip is coated with an aluminum glaze, leaving only the
very tip bare; a laser light beam is concentrated to a width of only
12 nanometers, and the tip can be brought close to the contours of
a living cell. Currently, the microscope scans at two frames per
second, but work continues on increasing the rate of scan. The point
here is that the microscope operates as a scanner, converting the
light from the specimen into data that a computer takes and reas-
sembles into an image displayed on a video screen.

NSF current program emphasis. In NSF’s current Guide to Pro-
grams specific emphasis is given to data acquisition. It is central to
the entire program in Astronomy. In the program for Biotic Systems,
emphasis is being given to improved methods of gathering data. In
that for Social and Economic Science, emphasis is strong on meth-
ods for measurement and data improvement. In the program for
Atmospheric Sciences, data acquisition is central to the entire pro-
gram.

The role of information support. Increasingly, computers are
crucial to natural scientists in the processes of data acquisition,
being used for control of experiments, for control of scanners, for
management of the data as they are obtained, for immediate display
of the data as means for guidance to the researcher. They clearly
are essential in the acquisition of digital images. Telecommunications
is clearly important when the means for acquisition is remote,
but otherwise it and libraries appear to play minimal roles at this stage in the research process.

Stage 2: Data Analysis

Characterization of the stage. Data having been acquired, the researcher then faces the task of data analysis. This stage has been the traditional focus for use of the computer in science and technology in processing the mathematical models and simulations that provide the fundamental tools for data analysis.

The example of digital imaging. The growing use of digital imaging, though, is greatly expanding the nature of data analysis and the role of the computer. It provides new means for analysis in the use of graphical methods, which is explicitly recognized in the NSF Guide to Programs in its description of priorities for the Mathematical & Physical Sciences by reference to the rapid expansion in the role of geometric ideas.

NSF current program emphasis. In the current NSF program for Biological, Behavioral, and Social Sciences, specific emphasis is given to mathematical modeling and to software development. In the programs for Engineering, an emphasis is on computational models of design processes, representations, and multiple levels of abstraction. The programs in Computer & Information Science and Engineering are largely oriented toward computational needs—in Computer Science, Robotics, Microelectronics, Advanced Scientific Computing, Networking and Communications. The programs on Biotic Systems, make reference to computerization of systematic research collections. Those on Social & Economic Science, on increased accessability to systematic anthropological collections and geographic information systems. Programs for database development are part of Instrumentation & Resources. These, though, are merely examples; the importance of data analysis is evident throughout all of NSF's programs.

The role of information support. Effective data analysis frequently will require that new data be compared with old, that it be supplemented with other data. This immediately implies the necessity for maintenance of files of data and for access to them as sources for those comparative and supplementary data. In this respect, the
Complications posed by files of digitized images are orders of magnitude greater than those for textual or purely numerical files. Any single picture may require from 100 kilobytes to one megabyte, but scanners can generate such pictures at rates of 10 to 30 per second. The result can be files of gigabyte size within merely an hour or so!

It is here that the role of the library, as the agency to preserve past records and to provide such access, first becomes important to the natural scientist. The problem of course is that the scientist may not recognize the importance of these kinds of data analyses and may not realize that the library can assist in gaining access to the existing data. As I will discuss later, the librarian can serve as advisor on the appropriate tools for data file management as well as in the processes of storage and access. Primary among them is the overriding concern with the management of the files themselves; here there needs to be fundamental research on the organization of such files and on the means for retrieval from them. Almost equally important are administrative concerns about access to facilities, resources, and equipment needed to acquire such data and to analyze it—spectral analyzers, monitors, supercomputers and communication lines to them. Indeed, primary among the supporting equipment are supercomputers, needed because the volume of processing required for use of digitized images exceed the capacity of even very large mainframes.

Stage 3: Conceptualization

Characterization of the stage. The creative stage for the scientific researcher lies in the conceptualization of new models, new relationships among the data. Complications in this stage are great, of course, given the multi-dimensional nature of natural phenomena.

The example of digital imaging. Without question, digitized images already have become vitally important as tools for conceptualization of problems. Through the use of image manipulation new relationships can readily be visualized. Through the use of motion, color, enhancement, windowing, exploding, and a wide range of other means for manipulating images, several dimensions can be shown at once, and it is easy to shift among them.

NSF current program emphasis. In the program for Electrical
and Communications Engineering there is mention of conceptual design. And in Engineering Design and Manufacturing specific emphasis is on conceptualization. In the program on Ecosystems new methods of predicting change are a program focus. In the program areas in Behavioral and Neural Sciences emphasis is given to neural factors for memory and learning.

The role of information support. Computers are clearly powerful tools is developing conceptualization. I have discussed the use of digitized images and the manipulation of them for observing and testing different relationships, but that is only one example of the use of computer processing for those purposes.

Stage 4: Distribution of Results

Characterization of the stage. Of course, this is the stage to which all of the discussion of new electronic means for communication applies. It is the stage during which the results of research are reduced to a standard form for communication, broad distribution, acquisition by libraries, and subsequent consultation. Writing scientific articles and reports, documenting the prior knowledge through citation, submission of the materials to the processes of review, preparation of them for publication, preparing the indexes and abstracts that will assist in access to them—these are the information processes of concern of the librarian and the publisher, and the focus of the discussions of electronic means for distribution. It is important, though, to see this stage in perspective. Important though it is, the earlier stages are where the real information needs of the research scientist occur; this stage is almost peripheral to them.

The example of digital imaging. To this time, digital images have not been significant components of this stage, and there have been only a few publications in which digital images were included. Of course, there are innumerable publications that contain pictures derived from digitized images, but those are by-products rather than the digital images themselves. However, there have been a number of experiments on means to make available combinations of text and digital image that would allow manipulation as well as display. At the simplest level, these experiments include storage of document pages in digitized image form, as the basis for subsequent
generation of page images and for optical character recognition. Over the coming years, I am sure that we will see increasing publication in this form, especially in support of interactive instruction.

NSF current program emphasis. The NSF Guide to Programs repeatedly emphasizes the importance of communication and distribution of the results of scientific research, continued increases in the application of information technologies, and increasing collaboration between libraries and their user communities. In Astronomy, emphasis is given to communication of information. In the Biological, Behavioral, and Social Sciences, one area of emphasis is integrating knowledge of metabolism; another is publication of major taxonomic divisions in systematic biology. In the Mathematical & Physical Sciences, the statement is made that "mathematics thrives on the sharing of ideas among researchers" and communication among groups of investigators is regarded as essential. In Bioengineering, emphasis is on expanding knowledge base and on improving flow of research information between universities and industry. In Electrical Engineering and Communications a major emphasis in on information activities. In the Geosciences, communication is central to the entire program. In its program for International Cooperation there is emphasis on cooperative research and the exchange of information. In NSF efforts to develop science resources collection, analysis, dissemination of information and the compilation of information on science and technology resources are all important program objectives.

The role of information support. All of the means for information support—libraries, computers, and telecommunication—clearly are crucial support to this stage in the research process.

Stage 5: Application of Results

Characterization of the stage. Finally, beyond simply feeding on its own results, research in the natural sciences provides the basis for engineering development, industrial manufacturing, production of new drugs, and advancement in health care. It has fueled the industrial revolution for the past two centuries, and it is now fueling the "information" revolution, especially in the development of the information technologies.
The example of digital imaging. I have already commented on software for algorithmic production. It arises specifically to provide means for generating images from specifications in either parametric form or image form. Examples include CAD/CAM (computer aided engineering design and manufacturing), architectural design, and cartooning. Once a digitized image has been created in this way, it may well be stored for later use or for manipulation independent of the source program.

NSF current program emphasis. NSF recognizes this stage in the research process in its program on Education and Human Resources, referring to course and career development and to teacher development. It places significant emphasis on industrial innovation, especially in small business, in its Major Initiatives for facilities modernization, research centers, career development, and information for small business.

The role of information support. All of the means for information support—libraries, computers, and telecommunication—clearly are crucial support to this stage in the research process. Indeed, they may be even more important to the furtherance of applications than they are to the advance of the sciences themselves.

THE ROLE OF LIBRARIES AND LIBRARIANS

In that listing of stages in scientific research, I have highlighted the potential role of the means for information support. We must also consider potential contributions of the library and the librarian.

Preservation and/or Access

The basic responsibility of scholarly research libraries is "preservation of the record of the past"; for that function they, along with the cognate institutions, archives and museums, are the designated institutions of society. All of the other information industries and activities have only the objectives of immediate distribution. Publishers and communication media indeed take no responsibility for preserving even their own products, ceasing to have any interest (but copyright protection) as soon as the book or journal is "out of
print." It seems unlikely that electronic forms of publication will change the focus of the commercial institutions in the immediate future. But the library does assure that material is preserved, and without doubt it will continue to do so with electronic materials.

The second responsibility of the library is to provide access to the preserved records, a responsibility that by longstanding tradition is shared among all libraries and with cognate services in the form of bibliographic utilities, secondary access tools (indexing and abstracting services), and a variety of data base access services. Thus, the library is not the only institution that fulfills that responsibility. Indeed, academic researchers do so through the very process of citation; the compilers of bibliographies, the indexing and abstracting services, the providers of online reference services all do so. But the library is in a unique position with respect to access to its own materials and, as a result, plays an especially important role in access to all others. As a result, as electronic forms of distribution become more available, the library will without doubt continue as a primary means for access to them.

Assuming that I am correct and that the library indeed does continue to fulfill its imperatives in the context of electronic publishing, are there problems in doing so? The answer is clear: There are serious problems with respect to both preservation and access, despite all of the rhetoric and high expectations, especially for the latter.

Whereas for the individual library in the past, preservation and access were two sides of the same coin, with the electronic media the library faces internal tensions in the choice between incurring a capital investment in acquisition and preservation, and incurring continuing operating expenses in access from elsewhere. In this context, a major objective of the CLR report to NSF was to explore the future form of library services and information systems for science. In that context, in addition to her discussion of the need for studies of scientists themselves, Helen Gee proposed that there should be similar studies of the library: What amount of its budget is being used for electronic information resources? What kinds of electronic resources are available from the library? How are information systems accessed? What kinds of training are provided? What are consortial arrangements? What is the library’s role in planning?

The CLR report concluded that there would be a continuing
transition in library emphasis from acquisition to access. Indeed, it appears steadily to be increasing in that direction, especially for the smaller colleges and universities. And for materials needed by the natural sciences and related professions, access is the necessity, and responsibility for preservation of materials will be focused on a very few major libraries and perhaps even on means other than the traditional academic library.

It must be said, however, that for any university with aspirations to be a major research institution in which the library collection is still a primary research tool, the choice is by no means obvious, at least with respect to materials needed for the humanities and social sciences. Furthermore, if the electronic materials themselves are to be preserved, some major libraries will need to take on that responsibility.

**Bibliographic Control of Electronic Publications**

Even with respect to access, though, there are serious problems in control of this literature. The problems are ones with which the library field has extensive experience. They arose with the entire array of "report" literature and similar ephemera. The tools that have been developed indeed provide the means for control of such literature. But they are tools expensive to maintain and expensive to use.

The first problem is the lack of control of the fact of publication. With print publishers, we have mechanisms for identifying availability and sources of materials. What will be the counterpart for electronic publications? There is a lack of means for assessment and review, so essential in making the decisions in selection and acquisition of material to be preserved or to be accessed. Surely in time electronic publications will make their way into the review literature. But for the bulk of them today and many of them in the future, such means for assessment simply will not be available.

Another problem is the lack of the tools for secondary access; there is as yet little coverage of electronic publications in the indexes, abstracts, bibliographies, and catalogs of the country. The result is that persons needing information contained in such publications have no means for identifying those of potential value.

Even if such material is identifiable, for example through the
various citation indexes (resulting from reference to it in some formally published document covered by them), the user may have no means for establishing where to obtain it. I can visualize that experience repeated a hundred-fold as desk top publishing makes it feasible to generate increasing numbers of such ephemera. The problem not only affects the users, though. From an operational standpoint, it places intolerable burdens upon libraries who must track down these "citation ghosts." Do they refer to real publications? To ephemeral publications? To figments of the imagination? How is one to tell?

These issues arise not only with respect to formally published materials but at least equally so to materials held within the single institution. The library can serve as a vital tool for management of those campus resources. It can include within its online public access catalog the references to such materials that may have utility beyond the individual faculty member or research project.

**Services of the Professional Librarian**

The role of the librarian is potentially much broader than that of the library itself. The sense of professional responsibility, the associated commitments and skills, the supporting professional organizations—all provide a solid base for the librarian to fulfill a wide range of functions in the context of electronic information.

*Meditation, consultation, training.* One area for contribution by the librarian is in mediation, consultation, and training. While researchers, students, engineers, and executives may know their own needs in use of any information resource, they frequently will need help in the management of personal files and in the use of hardware and software. The librarian can serve as a consultant in such cases, bringing to bear experience as well as technical expertise. Of special importance is providing assistance in gaining access to available electronic data files that a given user may need and in conversion of those data to the form needed.

*Selection and acquisition.* Perhaps the most important contribution is directly related to gaining access to materials. Selecting and acquiring information materials is the fundamental role of the librarian. To fill that role, the librarian needs to know the available sources. Perhaps we need the creation of national or even interna-
tional bibliographies of available electronic forms of information, including digitized images.

Cataloging. The other traditional function of the librarian is cataloging—providing the means for identifying and controlling holdings. I have already commented on the crucial role of the library in providing "single-point" access to campus information resources. In many cases, though, there may be need to have MARC formats adequate to the requirements and the means for sharing of cataloging data so as to avoid unnecessary duplication. The situation with respect to the cataloging of digitized image files serves as an especially important example of this.

Collection management. While electronic materials may not be acquired by all libraries, there will still be needs for management of them, and the skills of the librarian are as important in that respect as are those of the computer specialist. As I have pointed out, digitized image files can raise spectres of sizes of files of truly awesome magnitude—sizes that make even the largest libraries of the world appear small. As a result there are even more complex problems in determining where the files will be stored, how they will be organized, and how they will be preserved.

Content indexing and abstracting. Perhaps the most exciting intellectual challenge arises from the needs in "content indexing and abstracting." How does one retrieve images that contain something desired? If the problems in "full-text retrieval" were interesting, they were dull compared to those in image retrieval. Lest I appear to be posing an irrational challenge, please consider the means by which OCR software matches images against standard patterns, consider the means available for identifying a scene, with the potential for use of one frame from a scene as a retrieval surrogate—an "abstract"—consider the possibility of using characterizing quantitative parameters as means for retrieval. In other words, the tools are already here.

CONCLUSION

In conclusion, I have tried in this paper to provide a broad coverage of the information needs in science and technology. I have reviewed the status and expected future development of electronic
means for meeting those needs. I have placed special emphasis on digitized images as a medium of communication of information that is now of critical importance to research, to the application of its results, and to governmental policy making; in my view, as time goes on it will be of increasing importance. I have focused on the library and on the role of the librarian as a means for meeting the needs of scientists in access to and use of information in all forms. That means that library education should begin now to provide the technical tools that the librarians will need to fulfill their obligations in the world of electronic information.

REFERENCES

Cohn, Steven F. "The effects of funding changes upon the rate of knowledge growth in algebraic and differential topology," Social Studies of Science. 16 (1), 1986: pp 23-60.
Communications in Support of Science and Technology. A Report to the National Science Foundation from the Council on Library Resources. August 1990.
Henderson, Joseph. Dartmouth Medical School. "Creating realities for experiential


"What’s new." Ei Letters, 1(1), January 1991: 2

8. Senge, p. 146.
13. Peters, p. 95.
17. Peters and Waterman, p. 12.
19. Senge, p. 64.


1. This is the Comparison of Science and Technology Libraries Committee, Science and Technology Section, Association of College and Research Libraries, American Library Association.

2. This assertion is stated or inferred in the literature and is assumed by a good number of librarians and information specialists. I am not aware of a specific study that actually correlates these variables.


8. The importance of an awareness of the vagaries of the political process are discussed in numerous publications in public administration and other academic disciplines. For example, see: Morrow, William L. *Public Administration: Politics, Policy, and the Political System*, 2nd ed. (New York: Random House, 1980).


10. Schein, p. 6.


12. For example, see: Pinelli, Thomas. “‘The Information-Seeking Habits and Practices of Engineers.’ *Science & Technology Libraries* 11 (Spring 1991): 5-25. [This issue of *Science & Technology Libraries* is devoted to information seeking practices and related communication patterns.]


3. Roy, p. 32.


5. Ibid., p. 11.

6. Rudy M. Baum, “Traditional Boundaries of Chemistry Pulled Wider by


19. An example familiar to most in our profession is Northwestern University’s successful spin-off of NOTIS, Inc. to market its library automation system of that name. NOTIS, Inc. was recently acquired by Ameritech.


40. Blaise Cronin observed that library and information science programs are, like the human brain, split into two hemispheres that have different functions; *Vibrations.* “Cronin Urges Technological Upgrade.” *30:1* (Fall 1991): 1.


1. Tuck, Bill [and others]. *Project Quartet.* (British Library. Library and Information Research Report; 76); Cambridge: Cambridge University Press; 1990. (p.249)


7. *Ibid.* (p.239)


9. *Ibid.* (p.244)


The Minutes of the September 24-25, 1991 OCLC Research Libraries Advisory Committee indicate that OCLC is jointly developing a serials table of contents and document delivery project with Faxon. The database will contain bibliographic records from about 10,000 journal titles and will be linked to library holdings information. Members of the committee noted that "Faculty will not want to pay individually for such documents and will want the institution to cover charges." (p.6) They also note that "Un-mediated document supply requests directly from the user are inevitable." (p.6) For further information, contact: Faxon Research Services, Inc., 14 Southwest Park, Westwood, MA 02090 Phone: 617-329-3350 ext. 407; FAX: 617-329-6291.


14. The equipment and software necessary to establish a document archive database are available commercially. For example, Compulink's LaserFiche LAN can store document images and be integrated into a Novell network. The system handles computer files, paper documents, and FAX transmissions. The so-called "hydras," machines, multi-function devices for printing, faxing, scanning, and copying, could also be adapted for such an application. See Byte 16(4): 217; 1991 April for a list of companies that provide document imaging solutions.


17. Ibid. (p.2)


2. We wish to thank FAXON for the use of their historical collection of catalogs.

