Joan Cheverie  
Visiting Program Officer  
Coalition for Networked Environment (CNI)  
21 Dupont Circle  
Washington, DC  20036  
PH: 202.296.5098  
FAX: 202.872.0884  
jchev@cni.org  

Access to and Services for Federal Information in the Networked Environment  

Wednesday, December 4, 1996  
2 pm  

Microsoft Word 5.1
Access to and Services for Federal Information in the Networked Environment

Joan Cheverie
Visiting Program Officer
Coalition for Networked Environment (CNI)
Washington, D.C.
Access to and Services for Federal Information in the Networked Environment

With the increasing use and availability of information technologies, there has been a significant change in how federal agencies disseminate government information. This change is resulting in new dissemination mechanisms, as well as new and changing user needs and expectations. As a result, the responsibilities and capacities of institutions that facilitate the flow of federal information to academic and citizen communities need to be rethought in this shifting environment.

"Access to and Services for Federal Information in the Networked Environment," an initiative of the Coalition for Networked Information, is a white paper that will guide higher education and other institutions, such as public and state libraries, in the development of strategies for providing access to federal government information by their constituencies using the powerful, and rapidly expanding global information infrastructure.

The paper primarily focuses on issues and models for collecting, preserving, providing access to, and providing services for federal government information. It addresses these issues at the enterprise-wide or institutional level. The paper also summarizes policy and technical directions to provide a framework for understanding the issues involved.

Background

For the last ten years the federal government’s focus on accountability, budget management, and the potential of rapidly developing information and communications has resulted in their development of policies and practices which are significantly changing how agencies create, produce, and disseminate their data, information, and knowledge. The pace of change has quickened in the last five years and will continue to quicken between now and the end of the century. This shift is producing both opportunities and challenges.

The Problem

The problem is that what has been a stable, well-known system is now in flux and the local institutional investments which have supported getting access to federal information and using it are increasingly out of sync with the future of federal information.

What This Report Covers

Policy Directions

The evolution of federal policy regarding the distribution of federal information is now firmly on the path of electronic preparation and distribution. While there are continuing discussions of the pace of change and the continuing usefulness of print, the future of federal information production and distribution is clearly with the National Information Infrastructure (NII) and its assorted tools.

The important policy questions focus on how local institutions can adapt their own policies and strategic investments -- as well as how to have ongoing discussions with Federal agencies in order to build complementary programs.

Technical Directions

Federal agencies are adopting a number of technologies as they move their information to the electronic arena. Primary approaches include CD-ROMs, Internet accessibility and, particularly, World Wide Web sites. These technologies have a wide variety of
application possibilities so there is great diversity underlying what appears to be a consistent and coherent direction. In addition, the history of electronic federal information exhibits a variety of legacy application approaches that includes bulletin board systems, online manipulable databases, flatfile databases, gopher sites, etc.

The important technology questions focus on how agencies can make their data available electronically so that users wishing to combine data from multiple agencies can do so seamlessly.

**Production & Dissemination**

Production is increasingly electronically based, though federal information in print will also continue into the future. Not only is there wide variety among agencies in their application of information and communication technologies, we are also seeing shifts in the value-adding processing (e.g. analysis and interpretation) which agencies have undertaken in the creation of their printed publications. There are important shifts as well in how Federal agencies disseminate their information -- one of the most important is that more agencies are reviewing or applying fees for the acquisition or use of their publications.

The important production and dissemination questions focus on how local institutions will shift their investments in order to get federal information and make it useful in response to federal government policies -- as well as how to use local experience to inform federal decisions.

**Uses & Users**

Today’s information and communications technologies support new ways to mix, match, and manipulate digital multimedia information. New uses are possible. User experiences and expectations are changing in reflection and pursuit of new capabilities. Vast quantities of data and information are directly available via the Internet to a wider user community -- bypassing intermediaries and intermediary organizations. On the other hand, there are an array of technical, policy, monetary, and human support challenges to the individual and organization use of federal information.

The important use and user questions focus on how much new organizational and technical infrastructure is integral to facilitate access and use cost-effectively. Further, how critical and useful is it to collaborate with other organizations jointly to build a critical mass of organizational and technical infrastructure in order to spread the cost and the benefit across a number of organizations.

**Information Organization & Retrieval**

Print-based federal information has workable and extensive approaches to catalog, index, and retrieve government publications. The description and classification of print-based information resources is a well-understood (although complex) problem with a long history. The networked information environment introduces a number of new considerations of cost-effective and appropriately powerful approaches to description, organization, retrieval, and usage.

The important information organization and retrieval questions are both organizational and technical. They include the following:

- the development of taxonomies for networked information;
- classification schemes to describe the content of networked information resources;
- evaluative information beyond descriptive cataloging;
- the ability to describe widely varying levels of aggregated information sources and the granularity of information components that can comprise them;
- approaches for gathering the descriptive data for networked information;
how to approach the use and continuing development of networked information, discovery, and retrieval software which is both easily used and powerful in its capability;

the software, hardware, and telecommunications capacities that integrate support for storage, organization, retrieval, and use.

Implications

Collections

Networked government information collections offer a fresh opportunity to rethink collecting activities and to tailor collections more precisely to the needs of the local community. There is no doubt that, for the foreseeable future, existing heavily print-based research collections will continue to require service and preservation. Yet, increasingly, collections and users will depend on the full exploration and utilization of the possibilities offered by networked collections.

Traditionally, when one speaks of federal information sought by scholarly and other institutions, it is through three models of distribution and/or access:

- Ownership -- through library collections and information centers;
- Participation -- through several different national systems of "depository" agreements and arrangements;
- Partnership -- through several federal government research programs (e.g., the National Institutes of Health, the Department of Energy’s national laboratories)

During this period of intense and chaotic transition, the only constants related to federal information are change and inconsistency. There is no one method of interaction, however, and institutions must consider new strategies, new relationships, and new investments which will ensure effective access to federal information.

There are a number of concerns about federal information collections in the networked environment. They include:

- No commitment to provide continued access to information published either by the agencies or by the institution/user community
- An abundance of raw data with too little analysis
- In the network environment “one copy” is not enough

The key change in this new environment has been the shift from a static environment to a dynamic environment. Until now most users have accessed federal information by “coming to it.” Now they have the opportunity to interact with it.

However, because of the Net’s high capacity for data transfer and speed, institutions will need to redefine relationships with their community of users, with peer institutions regionally and nationally, as well as with the federal agencies responsible for producing the information.

The most important feature within this electronic context is no longer ownership but, rather, access. No single institution (or institutions) will house federal information as has been done up to this point. In the distributed networked environment, there is not the same need for a set number of copies as in the traditional environment. However if one copy is not enough, how many copies are sufficient? Collection policies and priorities will need to be rethought in light of the changes brought about by the network.

Success in managing federal information collections made available over a network will depend on:
preserving and providing access to segments of the electronic government information stream OR a willingness to depend on other non-library intermediaries for long-term access;

- the ability of institutions to move information off the Internet and onto a local network or onto some type of media-specific format (print, microform, or electronic);

- the willingness of institutions to coordinate the sharing of collection responsibilities for various segments of government information;

- the ability of institutions to reallocate resources for hardware and software with which to store and manipulate network information.

**Services**

Institutions of higher education and other institutions, such as public and state libraries, have played an important role in providing information services for scholars, students, and citizens using the vast amount of information produced by the federal government. The publications of the federal government reflect the government’s far reaching scope and its intricate complexity. Institutions have relied on general documents or reference departments in libraries and specialized data centers to help users locate and use this information. Traditionally government publications have not been as easy to locate and use as most other library material.

Information specialists with expert knowledge, searching skills, and experience provide mediating support for those challenged by the maze of federal information. Information professionals who provide public service support for government information are grounded in the structure and publishing patterns of federal agencies and specialized searching techniques for specific types of government publications such as census materials, legal information, or federal acquisition regulations.

Providing access to collections has not been, nor will it continue to be, enough. It is the value-added aspect that information specialists provide through the delivery of services which gives the order, the interpretation, and the usability to these collections.

Networked federal government information will transform existing models of service which, heretofore, have been based on traditional reference and referral activities that focused mainly on answering questions of users who came to a service desk in a library. The Federal Depository Library Program was designed around a system in which the user community was “locally” based and would, therefore, have convenient access to a physical collection housed in a library. Although traditional reference service will be necessary in a networked environment, it will not be sufficient to meet the needs and expectations of the students, researchers, and citizens who will use networked government information in the twenty-first century.

The traditional reference service model is based on ideas that are no longer valid in a networked environment. Some of these old assumptions (which have been valid for the provision of service for federal information) are:

- Public service is building based and users need to come into the library to ask their question;

- Service is available only at set times and users can expect service only when the service desk is staffed.

Networked resources such as these allow users 24 hour access to government information that is not bound by geography, offers some information from almost all departments and agencies, and is the result of innovation and experimentation by individual agencies. With
the migration of federal information to the Internet, there are resultant challenges for providing service to this evolving collection of resources.

Institutions will need to address many questions in providing service to networked federal information. Some of these include:

- What new reference services are needed for users who will not need to physically enter a building to use its services?
- Who comprises a user community that is no longer bound by geography and what is the institutional commitment to provide access for the information poor and the non-Net connected?
- What other kinds of electronic services, beyond e-mail, are possible that will enhance users' abilities to get complete information to satisfy their information need?
- How do providers of networked government information bridge their information service with the wealth of information that is contained in the substantial print-based government documents collections that many institutions have?

In the digital environment, the expertise of staff who are federal information specialists is a vital link between the user and the information they seek. However, the staff will employ their knowledge and skills in new and different ways.

**Networked Information Discovery & Retrieval**

There are many success stories that informed users of the Internet can share about accessing federal information. A political science student familiar with the White House's Web site home page visits it regularly to read the daily press briefing. A demographer working on a population project is quite comfortable with statistical information available through the Census Bureau's home page. For individuals such as these, the Internet is a timely tool with relevant information to meet their needs.

Novice users of government information, however, are often intimidated not only by the bureaucratic nature of the federal government, but also by the complexity of its information. A citizen community group looking for environmental information on a planned playground that may once have been a toxic waste dump might find their search too broad for the Net to handle efficiently. In searching the Internet using existing tools, these users are often presented with a significant amount of irrelevant information. They can easily get lost while navigating the Internet, leaving them feeling frustrated and incompetent.

Networked Information Discovery and Retrieval (NIDR) is the mechanism by which users locate, select, and retrieve information resources. Network technology offers many opportunities and challenges regarding what information is available to users and how that information is located. The network expands access for users, whether they are on-site or remote, and it changes the tools and strategies that they use in the search and discovery process.

The description and classification of information resources for organization, retrieval, and use is a well-understood (although difficult) problem with a long history. The networked information environment introduces a number of new considerations into description, classification, organization, retrieval, and usage.

Federal information presents several special challenges in the area of search and discovery. Users need to be assured that the federal information they access is authoritative. Government information is often the legal and regulatory language that determines such issues as the distribution of federal dollars (based on federally gathered statistics) or the guidelines for federally subsidized programs. The insertion of a small word such as "not" could entirely change the legislative/regulatory language and intent of a legal document and
could, therefore, have far reaching implications for research and scholarship. The user, therefore, must be assured that s/he is directed to a reliable source of federal information.

The issue of federal Web site authenticity is an important one that affects research and scholarship. There is, as yet, no reliable, authoritative single (or distributed) point of entry for all federal information, and this is an important factor that is contributing to the haphazard nature of Net access for both end-users and intermediaries. A proliferation of Web sites which serve only as pointers results in duplication of institutional efforts, confusion for the user, and raises questions of the authenticity of some Web sites as authoritative sources of federal information. Although GPO Access is the legislatively mandated, centralized point of entry for electronic federal information, it is not, to date, the single point of entry nor is it a comprehensive site for all federal information.

For federal information to have any value to the user, it needs to be organized and retrievable. The key components of any information search are:

- whether the information retrieved is relevant;

- whether all relevant material has been retrieved.

It is these two points, recall and precision, that express information search and retrieval performance. Internet technology, to date, only exacerbates the problems of search and discovery for federal information. Several reasons for this are:

- Networked information resources are extremely heterogeneous in nature, volatility, and coverage. They include a wide range of services and types of objects. This is part of what makes the NIDR challenge so difficult for federal information.

- Users need to be able to view the available information through a seamless interface rather than as a large number of collections organized by types of resource (e.g. gopher spaces, ftp sites, Web sites, etc.) or by methods of access.

These are not new problems for users of federal information. In the traditional environment, the user has had to use a variety of tools to locate federal information. The networked environment, as of now, does not help the situation, but, in many cases, hinders search and discovery.

Search engines rely on information about objects. As more sophisticated NIDR systems develop, there will be increased emphasis on metadata. The good news for federal information is that the Government Information Locator Service (GILS) is setting criteria for agencies to develop information about objects or metadata and this standard will provide the underpinnings for improved search and retrieval.

Sophisticated, thoughtfully conceived NIDR systems will not develop overnight. Institutions need to develop strategies at the top level to deal with the inadequacies of today’s NIDR systems.

The dynamic nature of networked objects results in difficulties in description. The relatively static nature of traditional catalogs, indexes, and other finding aids have contributed to user success in identifying and locating information. It is evident that more research is needed in the area of distributed search and retrieval tools and services. If this issue is not resolved satisfactorily, the Internet will not live up to its full potential as a viable tool for research and scholarship.
Preservation

Preservation of electronic information continues the library mission embedded in the familiar paradigm: acquire information, organize it, make it available and preserve it. Libraries (particularly those in the FDLP) have participated in this significant, distinctive and successful role for print and other artifactual materials since the commencement of the program. In terms of fundamental principle or goal, there is no new issue, for the preservation role continues.

Preservation of information is the fundamental component of the archival function of a document depository. It will continue to be a requirement in order to satisfy user needs in the electronic environment. As this publication’s several sections make clear, user needs will continue in most respects to be what they long have been. Users will want information reliably locatable, so that when they go there (whether personally or on the Internet) they can expect to find what they are looking for. Users will want information easily accessible: the location tools must be clear and accurate, and the information must be promptly retrievable. In the electronic environment the need for access tools will be more evident, and users will expect appropriate and standard software to be readily available. Finally, whether they are conscious of this need or not, users will expect information to be available that was placed in the depository’s care a long time before; and they will expect that the integrity of the information they get from the depository will be assured.

As a matter of implementation, the preservation of electronic federal documents raises issues virtually identical to those institutions face in preserving other forms of electronic information. Indeed, the initiative of the Federal government in moving the FDLP to electronic provision may not be the first case in which the government has been the first stimulus for other public and private sectors to take steps that would eventually have been necessary in any case. As a result, institutional leaders and document program managers should keep before them that, while their present responsibility may be assurance of long-term access to government information, they are often likely to be setting precedents for the preservation and integrity of electronic publications of all kinds. These precedents, as we shall see, may be both technological and organizational.

Preservation of electronic information raises new practical issues for librarians and archivists that did not previously have to be faced, primarily because the information now becomes separable from the medium on which it may temporarily reside. A book, or a printed Federal report, is published as an artifact. Like journals, manuscripts, sound recordings, CD-ROMs and other information resources which are published as objects that are their medium, artifacts exist in space and require specific physical handling to use. With such materials, to preserve the artifact is to preserve the information contained in it.

In contrast, networked electronic information is volatile in two important ways. First, at the current time it always resides on media which themselves are fragile and have no demonstrated long-term life.

Second, electronic information is easily transferred from one medium to another with no loss, a technique with which we are all so familiar already that it has become senseless to talk of an "original publication" as opposed to a copy, for one (true) copy is as good as another for any practical purpose. The resulting ease of copying, of modification, of format change and of use is the positive side of such volatility. The negative side is the ease with which information can without detection be accidentally or intentionally lost or changed.

One of the many consequences of this volatility is that, unlike with books where the decision may be postponed for years, preservation of an electronic document must be considered from the moment of its publication or even before if users are to be assured of its longevity and integrity.

Electronic preservation requires new forms of institutional commitment because the organizational and fiscal obligations must be long-term. Printed materials can survive loss of
care for many years; electronic information can not. Institutions must participate in ensuring that mechanisms exist for appropriate long-term preservation and access to electronic federal information. The current situation in electronic preservation is one of preparation rather than practice.

**Infrastructure**

The ability to digitize all forms of recorded communications and to transport digital communications globally, has brought and will continue to bring both enormous opportunities and matching challenges. Digital communications allow the opportunity to communicate across geographical and chronological barriers, but not every digital tool or digital format matches the need of every community. Not every type of information matches up well with the various digital communications tools that can convey it. The use of computer-mediated communications is often a question of exponentially enhancing the quality of service, not a question of saving money or human resources. The infrastructure components of today’s communication's system are not less expensive, nor are the components necessarily easier to install, maintain, upgrade, etc. Their very presence may force a complete reconfiguring of all related departments, services, components, etc. Investing in infrastructure for communications is not something to be entered into lightly or without a great deal of thought and understanding.

The most important element in considering issues related to current communications infrastructure for electronic communication is that of opportunity cost. In the short term, the easiest, least painful solution for decision-makers may be to do nothing. Faced with considering communications infrastructure, costly information resources, equipment, and staff training, it is tempting to simply avoid any major decisions and keep waiting for the next latest, greatest technology to emerge — one that will be so self-evident that there will no longer be any question about what needs to be done. The wait-and-see strategy is prevalent and, in many ways, very understandable. It is also very likely to be the most expensive and costly, in every sense, in the long term. Decision-makers must address a critical question when considering what to do related to communications infrastructure: What is the cost of not doing what is necessary to fully participate in digital, global communications?

Though the focus of this section is on infrastructure, the paper of which it is a part focuses on federal information, accessing it and providing related services in a networked environment. Normally, infrastructure would concern those components that would enable the networked environment. However, because of the focus of the paper, it is also possible to make the case that organizations, departments, administrators, and staff may find themselves considering how to position their institutions so that they may fully participate in the creation, exchange, storage, manipulation, and servicing of federal information. This environment could be considered as an infrastructure of sorts, by providing access to and the creation of federal information resources.

In considering the components of a federal information sources infrastructure, there are many pieces that need to be reviewed when approaching the design of such an infrastructure. Among the prominent components that must be considered are:

- User Groups
- Data Types (as utilized by the User Groups)
- Hardware and Software
- Facilities (Creation and Use)
- Collections
- Staffing and Training
- Administration
- Intangibles
Opportunities & Challenges

The potential of the continuing development of the NII is the vision of a world in which people can easily discover, evaluate, select, retrieve, use, and combine information resources in the widest variety of formats. For example,

- the Internet provides an anytime and anywhere communications medium.
- the World Wide Web (WWW) provides the software framework to combine, present, and link information content in various digital formats; e.g., text, graphics, sound, video, etc.
- Networked Information, Discovery, and Retrieval (NIDR) systems provide new possibilities for discovering, retrieving, and (to some degree) using networked information.

Federal government information -- often the heart’s blood of research and development, teaching and learning, advocacy and local government decision-making -- is increasingly available to users as part of the Information Superhighway. On the other hand, there are an array of technical, policy, and monetary challenges to the future use of federal agency electronic information and communications.

It is essential that organizations with a critical mass of infrastructure and capability -- and a reliance on federal information or a responsibility beyond their organization to citizens -- have a working awareness and understanding of the opportunities and challenges. Organizational leaders need to position their organizations not only to take advantage of new opportunities, but even more importantly -- to participate strongly in meeting the challenges and creating the solutions essential to the successful use of federal information.