Access to and Services for Federal Information in the Networked Environment

(Draft March 1997)

Preface and Acknowledgements

The Coalition for Networked Information (CNI) has had a long standing interest in federal information resources and services and the potential that the networked federal information environment offers to scholars, researchers, students, and citizens. The Coalition has long been an active supporter of the Government Information Locator Service (GILS) effort viewing it as a critical first step toward better performing networked information discovery and retrieval technologies, systems, and services.

The work of Charles R. McClure, Distinguished Professor at the School of Information Studies, Syracuse University, who was a Visiting Program Officer at the Coalition during 1993, provided the foundation for CNI's Access to Public Information Program with the overall purpose of improving public access to networked government information via the Internet. Part of this initiative included the establishment of a Visiting Program Officer for Federal Information and this white paper represents the offspring of this initial work.

With the increasing use and availability of information technologies, there has been a significant change in how federal agencies produce and 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 is a white paper whose goal is to guide higher education and other institutions, such as state and public libraries, in the development of strategies for providing access to and services for federal information by their constituencies using the powerful and rapidly expanding global information infrastructure. It addresses issues of service, access, collections, preservation, and infrastructure at the enterprise-wide or institutional level.

In its conceptualization and development, this paper has served as a basis for initial discussions among those who deliver and use government information. It is

hoped that with the paper's publication and dissemination, it will stimulate further discussion and development of strategies among an even broader audience as we all grapple with the evolution of government information in the networked environment.

Joan F. Cheverie
Head, Government Documents Department
Georgetown University
and
Visiting Program Officer
Coalition for Networked Information
Washington, DC
March 1997

Access to and Services for Federal Information in the Networked Environment

(Draft March 1997)

EXECUTIVE SUMMARY

Main Recommendations

This white paper recommends that decision-makers, including library directors, chief academic officers, and chief information officers, in organizations that rely upon and provide access to federal government information reassess their investments and program strategies to reflect the dynamic changes taking place in federal agency publishing and distribution. The old system is obsolete and there needs to be a re-engineering of federal information programs in light of network developments.

The paper makes three overall recommendations:

- Decision-makers need to reassess their institutional investments in and policies for selection, acquisition, access, service, and preservation of federal information in the networked environment.
- At the institutional level, collaboration is needed to bring together the range of skills necessary to provide networked federal information. At the national level, inter-institutional collaboration is needed to realize potential economies of scale.
- Given that access to federal information is a hallmark of our democratic society, institutions have a responsibility to advocate for federal information policies that will ensure continued access to networked federal information for all citizens.

Background

For the last ten years the federal government's focus on accountability, budget management, and the potential of rapidly developing information and communications systems has resulted in the 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 accelerated in the last five years and will continue to do so between now and the end of the century. Federal information distribution policy in the electronic environment is now more diffuse as agencies are becoming increasingly independent of the Government Printing Office (GPO). This shift is producing both opportunities and challenges for institutions who collect and service federal information.

The Problem

The problem is that what has been a stable, well-known system is now in flux and local institutional investments which have supported providing access to and use of federal information are increasingly out of sync with the future of federal information. The Net is in its infancy and is still evolving while, at the same time, institutions are grappling to provide a serviceable collection during this dynamic time of transition. To date, efforts to make networked information accessible often represent individual projects that may not reflect an institutional commitment to sustainability.

What This Paper 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 about 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 to capitalize on the opportunities created by this changing environment, as well as to establish 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 and, increasingly, Internet accessibility particularly World Wide Web sites. These technologies have a wide variety of application possibilities, and hence, 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 manipulatible databases, flatfile databases, gopher sites, etc.

Technology questions center 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

Federal information production is increasingly electronically based, though federal information in print will continue to be a viable format for the foreseeable future. Not only is there wide variety among agencies in their application of information and communication technologies, there are also shifts in the value-adding processing (e.g., analysis and interpretation) which agencies have undertaken in the creation of their printed publications. Shifts are also occurring in the way federal agencies disseminate their information -- one of the more important is that many agencies are reviewing or applying fees for the acquisition or use of their publications.

The central 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.

Use & Users

Today's information and communications technologies support new ways to mix, match, and manipulate digital multimedia information. User experiences and expectations are changing to reflect this pursuit of these new capabilities. Vast quantities of data and information are directly available via the Internet to a wide user community -- bypassing intermediaries and intermediary organizations. There are, however, an array of technical, policy, monetary, and human support challenges facing both the individual and the organization in the use of networked federal information.

The important use and user questions focus on the amount of new organizational and technical infrastructure required to facilitate access and use in a cost-effective way. Further issues highlight the need to collaborate with other organizations to jointly build a critical mass of organizational and technical infrastructure in order to spread the cost and the benefit across a number of organizations.

Implications
Collections
Traditionally when one speaks of federal information collections it is generally through three models of access: ownership, participation in depository arrangements, and partnership with federal agency programs. Networked federal information resources offer a fresh opportunity to rethink institutional collecting activities and to tailor collections to meet the needs of the user community.

- Institutions need to rethink what it means to collect federal information in the networked environment, leverage institutional strengths and resources through partnerships and consortia, and develop new models for collections.

Preservation
Preservation of information is the fundamental component of the archival function of a knowledge repository. It will continue to be a requirement in the electronic environment in order to satisfy user needs. Users will continue to expect that federal information which was placed in the "care" of the institution to be available, and they will expect that the integrity of that information will be assured. Preservation of electronic federal information raises new practical issues for institutions primarily because the information now becomes separable from the medium on which it may temporarily reside.

- Institutions need to form consortia or other cooperative arrangements to share the responsibilities and costs of preserving networked federal information; these consortia need to negotiate with the federal government the terms on which they will provide this preservation function.

Networked Information Discovery & Retrieval (NIDR)
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 and it changes the tools and strategies employed in the search and discovery process. However, mechanisms for locating federal information on the Net are rudimentary and less adequate than systems for other media. The issue of federal Web site authenticity is also an important one that affects research and scholarship. To date, there is no authoritative single point of entry for all federal information, and this is an important contributing factor to the haphazard nature of Net access. Sophisticated NIDR systems will not develop overnight and institutions need to develop strategies at the top level to deal with the inadequacies of today's systems.

- Institutions need to develop tools and network strategies that will provide users with an organized entry point to federal information, while at the national level they need to advocate for an authoritative access point and

the development of standards that will facilitate network-wide indexing and representation of federal information resources.

Services

Institutions of higher education, public libraries, and state libraries have played an important role in providing information services for scholars, students, and citizens using the vast amount of material produced by the federal government. These organizations provide the expert knowledge, searching skills, and awareness of local information and communication patterns and needs to complement and sustain the technical and organizational infrastructure investments. These capabilities are absolutely essential to bring meaning to the present flux and inconsistency that characterize today's federal information environment. Networked federal government information will transform existing models of service which have traditionally been building-based and dependent upon staffing reference desks that serve a relatively defined user community.

- Institutions need to rethink their service policies in the networked environment, define the communities for which they will provide service, and develop new service models that embrace and exploit these new technologies.

Infrastructure

- Institutions need to plan for and invest in an infrastructure (equipment, connectivity, training, support, and financial models) that will allow their clientele to take full advantage of federal information in the networked environment.

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. Our federal government's information -- often the heart's blood of research and development, teaching and learning, advocacy and local government decisionmaking -- is increasingly available to users as part of the Information Superhighway. However, our institutions are grappling with myriad technical, policy, and monetary challenges to realize the potential of networked federal information.

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 that networked federal information offers. Organizational leaders need to position their organizations not only to take

advantage of these new opportunities, but even more importantly -- to actively participate in meeting these challenges while creating solutions essential to the successful use of federal information.

INTRODUCTION

Overview

Federal information plays an important role in the mission of our research and education institutions. Federal information can be a critical piece of a research initiative, as well as a key element of the teaching and learning process. It can provide the foundation for scholarly research and/or provide the springboard for this type of work - from NSF guidelines for obtaining grants, to U.S. census data, to inventories of hazardous chemicals, to pending regulations, to databases of scientific information, to health care information, to diplomatic post records. Examples, by academic discipline, of some commonly used resources include:

> History - *Foreign Relations of the U.S.* (FRUS); *Weekly Compilation of Presidential Documents*; Congressional testimony and reports

> Demography/Public Policy - Census data; data from the *Current Population Survey*; *Vital Statistics of the U.S.*; Geographic Information Systems (GIS); *Federal Register*; *Code of Federal Regulations*

> Political Science - FRUS; *World News Connection*; *Congressional Record*; *Statutes at Large*; *U.S. Code*; Congressional testimony and bills

> Business/Economics - Bureau of Labor Statistics data; *Budget of the U.S.*; *National Trade Data Bank* (NTDB); Securities and Exchange Commission data; Treasury Department statistics

> Environmental sciences - *Toxic Release Inventory* (TRI); superfund information; EPA environmental impact statements; wetlands surveys

These resources and more provide the primary source material that researchers, students, and citizens have depended upon for decades. The ability to identify, locate, and use this information enhances research productivity and student learning.

The way government information dissemination programs have operated until now has been in a structured and organized environment. The Government Printing Office (GPO) has traditionally been the required primary printer for most agency publications. The agencies pay for the cost of printing out of their budgets and, increasingly, they are resisting this federal requirement to use GPO since some feel they can publish their materials faster and more cheaply using in-house printing technologies or by contracting out their printing. At the same time,

GPO is not the sole disseminator of government information. The National Technical Information Service (NTIS) as well as some agencies such as the EPA and NASA operate well established dissemination programs.

Federal agencies have issued their information in multiple formats for a number of years. The formats issue is not a new challenge facing users or institutions providing access to federal information. Electronic federal information is merely another in a long list of formats.

Users of federal information have traditionally accessed materials in a variety of ways. Much federal information has been available either at no cost or for minimal cost through intermediaries such as libraries, directly from the agencies themselves, or from clearinghouses. Acting on the principle that government information is a public asset and the cornerstone of a democratic society and informed citizenry, federal policies have provided for guaranteed access to federal information.

The provisions of Title 44 of the *U.S. Code* establishes and describes the Federal Depository Library Program (FDLP) operated by GPO. Providing federal publications through depository libraries is one of the ways the government keeps its citizens informed of its workings in a timely manner. However, not all materials are printed by GPO. Publications not printed by GPO generally do not make it into the FDLP and must be accessed in some other way.

The FDLP principles of ensured access, service, use, and preservation of federal information, although still valid, were established in an age of print materials and do not reflect the information technologies of today. Much debate has surrounded the updating of Title 44 and many issues remain unresolved. Legislation to amend Title 44 to incorporate the current technological changes is underway in the 105th Congress, but it is not yet statute.

At the close of the twentieth century, with the increasing use and availability of networked information technologies, there has been a significant change in how federal agencies disseminate government information. This change is resulting in new distribution mechanisms, as well as new and changing user needs and expectations. As a result, institutional leaders (library directors, chief academic officers, and chief information officers) need to rethink their responsibilities and capacities of their organizations in this shifting environment to facilitate the flow of federal information to academic and citizen communities.

The Internet (and its related tools) is an enabling technology with the potential to provide seamless access to information for scholars, students, and citizens. The Internet can offer people broader access to information -- if they have the appropriate technologies and telecommunications capabilities. It also offers the prospect for the development of new types of services that do not rely solely on a building-based delivery system.

Digital federal information will not immediately displace print, so institutions will likely have to invest in the financial support of more than one delivery approaches for quite some time. However, the movement to networked federal information and communications is permanent. This will require increased local investments in technical capabilities, equipment, software, telecommunications, and human support and mediation.

Not only do organizations relying on federal information have to invest in a variety of infrastructures and organizational capabilities, they must ensure that the information is accessible and usable. This goal can only be met if organizations initiate discussions with federal agencies to educate them to the needs of local institutions.

Policy Directions
Federal policies affecting public availability of government information arise from a wide variety of laws and regulations. Title 44 of the *U.S. Code*, for example, provides for the GPO Sales Program and the FDLP. Both programs provide for the distribution of federal information from a majority of federal agencies. The Paperwork Reduction Act of 1980 and its reauthorization in 1995, together with OMB Circular A-130, provide a policy framework for the management of federal information resources. This legislation coincided with the growing viability of the networked delivery of government information and the increase in the cost of printing and distributing federal government information. During the FY 1996 Legislative Branch Appropriations process, the conference committee called for GPO's budget submission to be consistent with the strategic plan (included in the study mandated by Senate Report 104-114) to assure substantial progress toward maximum use of electronic information dissemination. This strategic plan delineates a system for a rapid transition to an electronic government documents publishing and distribution system.[1]

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 about the pace of change and the continuing usefulness of print, the future of federal information production and distribution is clearly with the NII and its assorted tools. All of those individuals and organizations relying on federal information must address the infrastructure, services, and programs which will make possible the effective use of electronic federal information.

Technical Directions
"The design of successful technical platforms demands a synthesis of technologies and practices from computer science, computer-communications networking, information science, librarianship, and information management."[2]

Federal agencies are adopting a number of technologies as they move their information to the electronic arena. Two primary approaches include CD-ROMs and Web sites. Both of these technologies have a wide variety of application

possibilities, yet there is great diversity underlying what appears to be a consistent and coherent direction.

The history of electronic federal information exhibits a variety of legacy application approaches that include an agency's use of Bulletin Board Systems, online manipulatible databases, flatfile databases, relational databases, gopher sites, etc. Data access presents different problems from text access. Technical compatibility, proprietary formats, data description and access standards, the use of standard languages like SQL -- are other examples of the many diverse approaches which are part of today's practice in the creation and distribution of electronic information.[3]

Most of these issues carry into the CD-ROM and networked environments as well.

"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."[4]

There is much work to be done by both federal agencies and institutions of higher education, as well as state and public libraries to ensure that this essential goal is met.

Production & Dissemination
The rapid development of federal information policy and technical infrastructure related to the NII is resulting in significant changes in how federal agencies disseminate government information. Federal agencies are now exerting greater autonomy in their use of information and communications technologies for their information production and distribution. The Internet and its related tools provide opportunities for agencies to deliver services, provide communication channels, as well as deliver publications. Production is increasingly electronically based, though federal information in print will also continue for the foreseeable future.

As they consider their continuing use of information and communications technologies in their missions, agencies, and the federal government in general, are expecting that the red-tape costs of federal information will be minimized, while the utility of government information will be maximized -- and agencies will be passing on and sharing costs in new ways. Agencies are increasingly considering or applying fees for the use or purchase of their data, information, and knowledge. No-fee access (which has been the predominant approach of the history of the FDLP) is seeing some shifts as federal agencies begin to recover their costs by charging fees.

Not only is there wide variety among agencies in their application of information and communication technologies, there are shifts in the value-adding process which agencies have traditionally undertaken in the creation of their printed

publications. More information may be available, but users could encounter raw data without the analysis and interpretation that many agencies publications have provided. This fact has significant infrastructure implications for institutions that provide access to this type of information.

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

Opportunities and Challenges
Electronic formats can facilitate the potential visibility and utility of federal information by making it available to a wide audience. Many agencies have already put up homepages on the Web introducing visitors to their various missions and services.

The Net encourages active learning by stimulating curiosity and creativity. Users can experience learning in an engaging environment that can break down barriers which, in turn, can lead to further study and other intellectual pursuits.

The Net is an enabling technology and has the potential to provide seamless access to federal information for the scholar and student. The network offers users broader access to information, especially to those with no nearby library. It also offers the prospect for the development of new types of services, e.g. interactive customer-oriented services linked to publications. By embracing these technological changes, institutions that provide access and services can have an impact on the network's future by adding value to this information.

The infrastructure and delivery of information and services is shifting and poses new dilemmas, questions, and opportunities for institutions. The ability of an organization to provide its users with the access and services to which they are accustomed will depend on how the institution answers such questions as: 1) How will users get access?; 2) How will institutions continue to carry out their missions as preservers and servicers of information in pursuit of intellectual productivity?

In order to inform institutions of the challenges and opportunities for reaching this vision, this paper will examine the following areas and their resultant implications for the organization: Collections, Preservation, Networked Information Discovery and Retrieval, Services, and Infrastructure.

Implications: Collections

Overview

Introduction

Federal Agencies now consider information resource management (IRM), as expressed through the legislative mandates of the 1995 Paperwork Reduction Act, to be a critical component in their efforts to economize, streamline, and improve their vast array of programs and services. As an indication of how far the reliance on electronic data interchange has evolved, nearly all federal agencies either manage a growing package of electronic products and services, perform electronic adjudication and regulation oversight (by issuing rules and guidelines online), or offer direct citizen access/services through the Internet.

As a result of this network activity, Federal agency information resources are becoming more important to the conduct of the Agency's daily business. Agency staff no longer consider their information resources to be merely the printed by-products of the Agency's administrative processes This is a rapidly evolving movement away from the limitations which have been part of traditional printing and publishing practices. Naturally enough, this opportunity is profoundly altering the bibliographic and distribution arrangements that have been shared among the federal government, the scholarly community, and the nation's libraries. 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., theNational Institutes of Health, the Department of Energy's national laboratories)

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.

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

Issues

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

Government-published reports frequently include significant analysis of the data collected. Now that raw data can be made available easily over the network, there may be less incentive for the government to produce such extensive analysis. For example, there may be no geographic report produced on unemployment in the U.S.; rather the user is presented with a mass of raw data.

- In the network environment "one copy" is not enoughCurrently, many federal agencies sell access to their network files to recoup some of the cost of making the information available. These agencies work frequently with GPO to offer depository libraries no-fee access. However, that access is often limited to a single password or workstation. This limits access to one user at a time. As many institutions have found, use of information sources via network access frequently surpasses use of the same resource when it is held within that institution. As a result, single-user access is frequently insufficient to meet the demand. To provide supplemental access, institutions need to reallocate funds to purchase additional access directly from the producing agency. *STAT-USA/Internet* offered by the Bureau of Economic Analysis is one example, and the replacement for the Foreign Broadcast Information Service's (FBIS)*Daily Reports* series, *World News Connection*, is another. 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.

Federal agencies may continue to "validate" their data through their legal authority, but they may no longer undertake the value-adding processes which in the past resulted in recognizable "containers." Increasingly, individuals have to "add the value" that "fabricates" the federal data and information into products and services designed around their individual needs. Not only must institutions invest in the technological infrastructure that is part and parcel of access to networked federal information, but other organizational value-adding processes will require further investment. Because of this, scholarly institutions must assume a much more active role (including financial investment) in the creation of products and services that are designed to facilitate this fabrication.

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.

What might a networked collection of government information look like?
Imagine a collection of statistical data like the Consumer Price Index or Leading Economic Indicators arriving in libraries and being immediately integrated with historical data. Users would find that monthly, quarterly, and annual information could be retrieved, then downloaded into the user's preferred format with equal speed and consistency. Statistics indexed to a base year could be simply recalculated to other base years on demand and used to make projections into the future.

Codification of laws and regulations would occur simultaneously on the day they were scheduled to go into effect. Hypertext connections would enable the reader to review all stages of the legislative and rule making processes from proposal to enactment. Public comment on proposals or regulations, once too voluminous to be included in the printed format in anything but the summary version, would now be available full-text as part of the networked record. Via networked collections, it will be possible to monitor Congressional voting as it occurs, or even to participate in Congressional hearings from one's home computer.

Networked collections will allow users to extract just the piece of government information they need, or to manipulate, repackage and annotate government information, so it will best meet their needs. In a networked environment, users will be able to take on research and analysis that was once too complex or time-consuming even to consider. Government will be able (cost effectively) to provide raw data sets to users who had been limited to studying the government's evaluation of a problem rather than testing their own statistical hypotheses.

It might be possible in the not too distant future that community users would find on the Net a core set of government information resources with the full text of current and historical information provided by the producing agency or a government intermediary. A core collection set could include the opinions of all federal courts, the *Congressional Record*, the *Federal Register*, *Statutes at Large*, the *Code of Federal Regulations*, and the *U.S. Code*. Recurring tabular data sources, such as *Vital Statistics of the United States*, and recurring statistics on health, crime, education, immigration, agriculture, and even the current census are also likely to become core titles/information resources. Publications that are brought together by government-sponsored clearinghouses like ERIC might also become core titles for a limited time after their initial release.

Institutions would grow accustomed to thinking of the core collection as part of their own collections. For these materials an institution's "collection" and its "network connection" would become synonymous.

Current Situation
A Snapshot Study
Purpose:
The purpose of this snapshot study is to determine the extent to which institutions could rely on Net (versus traditional) access for an agency's information/publications **today**. How do some agencies' Web sites compare with what they distribute through GPO? What are the implications for collections **now**? In addition, this snapshot is intended to evaluate how well these agencies are taking advantage of Internet capabilities.

Methodology:
Using the *Monthly Catalog* through GPO's Web site, in conjunction with the *List of Classes of United States Government Publications*, the number of titles published through GPO for selected federal agencies were identified for the time period January 1995-May 1996. This list was utilized to determine the number of print publications that were made available on the Internet. In the process, any enhancements were also noted.

Results:
The results are as follows:
[Insert Chart Here] -- Currently located at the end of this section

Bureau of Labor Statistics

The Bureau of Labor Statistics (BLS) (http://www.bls.gov/) offers a prime example of the possibilities involved in launching a collection on the Internet. While the site does not include much of the analytical information put out by the Bureau, such as articles from the *Monthly Labor Review* and *Occupational Outlook Quarterly*, the amount of statistical information represented is quite comprehensive. Of the approximately 228 paper and microfiche publications released by BLS from January 1995 to May 1996, almost all were statistical publications. Therefore, the material that is most often required by the academic community and the public is available on the Web site. This information has been greatly enhanced by its presence on the Internet, as data can be manipulated in ways that are not possible in the paper format. With various methods of searching and formats of retrieval, it is possible for an expert researcher or an average citizen to find the depth of information needed. In the paper environment, a researcher may have to go to several different publications, and wade through superfluous information before the requisite data is retrieved, while through the Web site it is possible to create tables that are limited to the specific data requested by the user. Despite its enhancements, however, the collection on the BLS Web site can not be considered a true replacement for what the Bureau offers in paper format. Much of the analytical text that accompanies BLS data is absent from their Web site and in some areas statistical data may be lacking. It is possible, however, to see from the site how a complete collection could be represented on the Internet and greatly enhanced in the process.

Census Bureau

The Census Bureau Web site (http://www.census.gov/) is a site both for the academic user and for the general public. A series entitled, *Current Population Reports*, published by the Census Bureau, was specifically chosen to compare this print collection against the Internet site. Out of a possible thirty-three documents, twenty-six print publications are available on the Web site and twenty of the twenty-six are exact matches to the print collection, plus they are provided in PDF format. The information in the other six documents is presented in some capacity.
Sometimes, however, the number of tables present in the print publication

is reduced by half on the Internet site. Also, the extensive analysis in the printed publication is now only a brief synopsis (e.g. *The Black Population in the United States: March 1995*).

There are instances when statistical information is not present even though there is a link to the information (e.g.*Population Projections of the U.S. by Age, Sex, Race, and Hispanic Origin: 1995 to 2050*). The Census Bureau Web site does have enhancements which allow an Internet user to create their own data files using the 1990 Census or to create a thematic map. The site also provides population clocks for the United States and the world giving a continual population count, lists the most current economic indicators, and contains the publication, *Statistical Briefs* , in PDF format.

The Census Bureau site offers a subject approach to indexing the information from their publications, as well as other alternatives for searching the online documents by keyword, place, map, or staff. However, it is recommended to search for statistical information using the subject index. Also, this site contains information such as press releases, radio broadcasts, available publications and how to order them, and other general information.

It should also be noted that the *Statistical Abstract of the United States* is available on this Web site. At present, an Internet user can search within the text of *Statistical Abstract*, but there are no links to the resources providing the data for the charts. The Web site indicated that many of the statistical publications were being provided through a subscription service which, for a limited time, would be free.

Office of the President

The collection of presidential documents currently on the Net (http://www.whitehouse.gov/) is in some ways an enhanced version of the print collection. Press releases, briefings, and speeches are all included in a single location and are searchable. There are "Briefing Rooms" which include basic economic and social statistics for all major categories with accompanying charts displaying the data. There is no comparable print source released by the Office of the President.

In some instances, the general information offered by a print document is greatly supplemented by the information on the Web site. For example, in 1995 the Office of the Vice President released a small pamphlet on the GLOBE program. On the Internet, there is a full Web page for this program with significant in-depth information.

While much of the presidential material is present and enhanced on the Internet, there are some areas of information that are clearly lacking. For instance, there were approximately ten documents published by the Office of the President on disability in both text and Braille versions. None of these documents can be found on the Presidential home pages. A further drawback to the collection on the Internet is that the documents that are present are often difficult to locate. In the print collection, all documents are published under the Office of the President and readily identifiable as such.

On the Internet, each agency or commission under the president has its own home page often with different methods of searching or retrieval, with no common organization or classification as is present in the print collection.

Central Intelligence Agency

The Central Intelligence Agency Web site (http://www.odci.gov/cia/) is a site both for the academic user and for the general public. It contains its primary publications in PDF format: *The World Factbook 1995*, *The 1995 Factbook on Intelligence*, and *The Handbook of International Economic Statistics 1995*. Unfortunately, the other eighty-three publications released in 1995 through May 1996, are not available on this Web site.

The site enhances its print publications with a number of features. For example, the list of Chiefs of State and Cabinet Members of Foreign Governments is updated monthly and indexed by country. Also, the site includes audio and video clips and photos to discuss the history of the CIA, to tour the CIA Headquarters, and to view an Exhibit Center which includes images with text about such items as an enigma encoding machine.

However, other CIA publications which are of great interest to researchers such as maps, are not presently made available on the Internet. The site does provide ordering information for the maps and other secondary publications. It should be acknowledged that the CIA Web site provides a comprehensive list of its publications back to 1980 and in some instances back to 1971.
Another useful feature of the site is its internal search engine for publications and public affairs information. The search engine allows the user to control the search query by selecting features such as case sensitivity and it provides assistance by giving helpful tips on formulating searches. The Web site also includes a suggested reading list, links to

other intelligence information, speeches/testimony/press releases/statements, and general CIA information.

Of particular note, access to this site was interrupted for awhile when hackers "broke into" the site and mounted a bogus home page. Once again available, the initial home page screens have warnings about it being an official government site and that unauthorized use is prohibited.

Federal Bureau of Investigation

The Federal Bureau of Investigation Web site (http://www.fbi.gov/) is a site that is informative for the general public. The site contains current press releases, photographs, descriptions, and backgrounds of the ten most wanted criminals, general facts and employment information, field office locations, frequently asked questions concerning the FBI, information on investigations (e.g. Unabomber), information on congressional affairs involving the FBI (e.g. Russian organized crime), public affairs information and information about the FBI training academy.

However, this site cannot be recommended for use by an academic researcher. Only one of its publications is available on the site, *FBI Law Enforcement Bulletin*, from November 1994 to present. The main publication of interest would be the Bureau's *Uniform Crime Reports*. At present this valuable publication of compiled crime statistics is not available on this site. What is available is a press release regarding *Uniform Crime Reports* and a very brief summary of the extensive report.

Considering that approximately sixty reports were published by the Federal Bureau of Investigation in 1995 through May 1996, having only one actual document on their Web site greatly diminished its usefulness. The FBI home page could benefit from such enhancements as an internal search engine, and an increase in the number of publications made available.

Conclusions:
This snapshot study of federal agency Web sites illustrates the future possibilities of a federal information collection on the Net. However, the results clearly indicate that, at present, traditional collections remain essential because full Internet representation is not yet available for these publications. In short, the digital collection today is still a supplement to the traditional collection.

Potential Models & Strategies

Prospects for Shared Government Resources in the Electronic Environment

Sharing collections and collecting responsibilities appears to offer great potential for ensuring and safeguarding long-term access to government information. Ironically, sharing responsibilities for managing printed government publications collections produced marginal results and in many cases failed. Only with much difficulty have users been able to know, at the title level, the government publication holdings of other institutions. The complexity of government published material has often required a look at these material before one could ascertain its usefulness. These factors frequently led to a preference for having the collection in-house, even when this duplicated holdings of nearby institutions.

Metadata describing networked government information databases will now offer users information about these complex products often down to the level of individual data elements. With networked access to shared collections, the interface can be made transparent to users and as timely as in-house access. The challenge now is the need for organizations to make and to fund long-term commitments to government information collections.

Potential Models for Networked Government Information

ICPSR
The Inter-University Consortium for Political and Social Research (ICPSR) has been a significant player in collecting, preserving, and providing access to computerized data files. Its collections are machine-readable data files, the largest files being those produced by the U.S. government (election data, census, voting records and health statistical surveys). This model contains both advantages and disadvantages for the management of networked federal information collections. They are:

- Components of the ICPSR model for collecting and preserving information might be a scalable model for managing certain types of government information.
- ICPSR is a fee-based, membership organization with password restrictions that could potentially leave a large user group with no collection access.

Organizations Working with Agencies
One of the best examples of libraries partnering with government agencies is the ongoing arrangement between Cornell University's Mann Library and the U.S. Department of Agriculture. The project is titled the USDA Economics and Statistics System, (http://usda.mannlib.cornell.edu/usda/usda.html) funded by a USDA-CSREES grant, the system includes reports and historical data sets

covering both domestic and international agriculture. This model, too, offers possibilities and drawbacks for federal information collections.

- Users able are able to manipulate the information to suit their needs since reports are usually text files and data sets are available to download into a variety of statistical software formats.
- The arrangement includes only a small segment of USDA's publications/materials, and it is unknown whether or not this is a scalable model.

State-wide or Regional Sharing Programs
Maryland's State Library Resource Center, Maryland's alternative to a state library, is administered by the Enoch Pratt Free Library. Coordinated network collections, including a wide range of government information, are made available via the state-wide public information network known as SAILOR (http://www.sailor.lib.md.us/). Cooperative programs such as SAILOR might serve as a coordinated mechanism for long-term access to selected government information in electronic formats. Some of the highlights of the SAILOR program include:

- Funding assistance and training have been incorporated into the State Library Resource Center program, making it particularly useful and affordable for local county library systems to offer access;
- Topical Area Review groups (TARs), including one for government information, have been formed to help shape the content and connections SAILOR provides to other libraries and collections;
- SAILOR developed local interfaces, tailored to a Maryland audience, that enhanced access;
- SAILOR "collections" of government information sources are a combination ofconnections to other network sites, as well as to connections to files mounted and maintained by SAILOR staff.

ACLIN, part of the Colorado Information Network, provides free access to public and commercial information resources for Colorado residents. It is a central source for state government information, as well as for selected federal information. Components of the program include:

- A defined mission which incorporates information resources that support the education, health, business, and social service activities of residents;

- Toll-free dial access to the network for all residents whether they search the network from home, business, library, or school;
- Cooperation and funding from both the public and private sectors.

The question to be considered in both of these examples is, whether there is a commitment to continued funding of these programs and what would happen if funding ended?

Strategic Partnerships with Federal Agencies: Citizen Access and Scholarly Organization through Electronic Reading Rooms

The established system of federal depository libraries and/or information centers as distribution and access models needs to be rethought in the networked environment. Instead of building collections based on these relationships, scholarly institutions (and public and state libraries) could actively seek to build alliances and infrastructures with their peer institutions. Such alliances would allow for the identification, organization, and support of a common electronic interface for federal networked information or "electronic reading rooms." Electronic reading rooms, as a concept, would require planning, developing, implementing, and managing electronic "spaces" on the Web so that researchers and citizens alike could have predictable and stable intermediary places on the network to seek and find needed federal information without having to visit every agency Web site.

Although these alliances should not, and could not, replace the legal and managerial responsibilities of the federal authorities to properly archive and preserve their information, they could be an important transitional device. In other words, these alliances would provide unique opportunities for both federal authorities and institutions to explore common problems and challenges in maintaining an effective public information life cycle.

For instance, a particular university (or consortium of universities) could enter into a contractual arrangement with a federal agency to house and service the "host" on the Net. This arrangement would include the maintenance of a home page(s), answering general questions from the public, and creating a companion home page that directs people to other useful sites on the Net dealing with the same subject.

In the case of the U.S. State Department, for example, this could include links to scholarly and popular sources of information about foreign policy, travel, international legal, public health, and environmental concerns. The institution would be responsible for maintaining the companion home page, and assuring the "validity" of the links maintained there. The State Department and the University of Illinois at Chicago (UIC), in a

collaborative effort, have established the Department of State Foreign Affairs Network (DOSFAN) (www.state.gov/) to provide user access to a wide range of current foreign policy information.

Community Information Organizations
A community information organization, for the purposes of this paper, is simply an active alliance of community interests (rather than the interests of the scholarly community) with three primary goals.

- First, through the coordination of common resources shared by community organizations (cultural, educational, economic), government offices, and scholarly institutions, it fosters the development of an open community computer information network; or, assists in the enhancement of an existing one.
- Second, the alliance works to strengthen the organizational links between the public electronic network and public non-computer community information resources through the creation of necessary referral techniques.
- Third, the alliances actively seek out, organize, digitize, and make available within the community computer network government information vital to the community's social, political, and economic well-being.

Indeed, the wealth of information produced by federal, state, and local governments is often the key component within any community information organization. Wide distribution of this crucial "public knowledge" can make a difference in the economic development of neighborhoods and communities, the education of children and young people, the health of individuals, as well as the support of citizen participation in their local governments.

Much of this information, along with the professional knowledge of how to accumulate, disseminate, and provide access to it, is found in the many academic, state and other public libraries that serve as document depositories for agencies at all levels of government. In particular, academic, state, and public libraries can provide critical elements of organization, information technology and telecommunication networking capabilities, as well as other institutional resources, that facilitated the distribution of government information across village, township, city, county, or school districts.

Specifically, these partnerships will seek out opportunities to develop and implement networks of neighborhood-level public information resources that deliver a "basic package" of citizenship information services. Institutions can lend their considerable technical resources and

connections to the Internet to foster the development of these community information networks.

Information professionals, experienced in federal and other government information, would work toward the assurance that this "basic package" would draw upon a healthy mix of computerized and traditional community information sources from all levels of government.

Recommendations

- o Institutions need to rethink collection policies to address the issue of what it means to collect in the networked environment and the opportunities and challenges that this presents.
- o Institutions need to form partnerships and/or participate in consortia that leverage institutional strengths and resources to develop new models for collections.
- o Institutions should leverage the strengths of the network by developing collections tailored to local needs.
- o Institutions need to develop strategies that bridge/link traditional and networked government information collections.
- o Institutions need to reassess the way in which they allocate collection resources, which traditionally have focused on ownership of government information, to ensure that there is sufficient access to networked information to meet the needs of their users.
- o Institutions need to develop strategies and mechanisms that will ensure long term access to collections important to their clientele.
- o Institutions should monitor federal agency Web sites for completeness and determine local need for retention and overlap with "traditional" federal information collections.

Implications: Preservation
Overview
Introduction
Preservation of electronic information continues the institutional mission embedded in the familiar paradigm: acquire information, organize it, make it available and preserve it. Institutions (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.

Issues
Why preserve electronic information?

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.

There are two broad strategic lines to be considered when examining issues of preservation of electronic government information: the technological and the organizational.

Technological issues of electronic preservation

As a matter of implementation, the preservation of electronic federal documents raises issues virtually identical to those libraries 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, even when compared with the low grades of paper sometimes used in government publications and certainly not when compared with the 300-year life of currently available acid-free papers. Magnetic tapes are known to be fragile, magnetic disks have not been seriously tested and CD-ROMs and other optical and magnetic recording techniques have when tested fallen short of claims made by manufacturers of even a few dozen years of life.[5]

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.

A very great advantage of the volatility of electronic information, and the lack of an "original" as such, is that the concept disappears of a "rare" or "fragile" or "unique" document which requires special care or protection from users. If a single (true) copy exists, many may quickly be made and users may use any of them without risking harm to any of them, even the indistinguishable "original."

Information needing preservation

In the artifactual environment, information is by definition published in static form. Narrative texts, tables of statistics, photographs and moving images are issued as they were produced at a point in time. Changing information (economic series, for example) are issued as serial publications or at least in separate print reports. Much current Federal information being issued in electronic form is of the same kind, and is likely to continue to be so.

In recent years tools have been developed to support such reports, such as electronic databases (economic series, census data) and dynamic information resources (weather data and foreign travel advisories). These have served as bases for artifactual publications whether in print (labor reports) or CD-ROM (census data). The database and data collection

tools, including their maintenance and preservation, have until now been the internal responsibility of the collecting agency. In the networked electronic environment both the tools and the reports are capable of being published, and the issue will arise of what the preservation responsibilities for the tools are as well as of the reports; and of whose responsibility it will be.

For depositories as for research, state, and public libraries, it becomes more and more a responsibility to acquire and preserve databases underlying research as well as the consequent publications themselves. Publication of a database offers end users much more than merely a report derived from the database, as the user can manipulate the data with his or her own ends in mind; it is clear that depositories need to consider preservation needs for this class of material as well as those more similar to print.

Dynamic information resources present a challenge of a different order: if the information flow provided by the government is endless, or nearly so (e.g. spacecraft telemetry data), choices must be made whether to preserve any of it, if so what to save, whether to use sampling or snapshot techniques, and so forth.

Three kinds of electronic preservation

Preservation of electronic information needs to be looked at from at least three points of view: medium preservation, technology preservation and intellectual preservation.

Medium Preservation

At any given time stored information is located on some medium, whether paper, magnetic disk, microform, punched cards, magnetic tape or chiseled stones. The artifact or medium will eventually decay, some more quickly than others. Medium preservation is the concern for preserving the specific medium on which information is stored. In the artifactual environment we have learned a good deal about proper environmental controls for paper, binding techniques, and the like.

In the electronic environment less emphasis is placed on actually preserving the medium, though tape storage vaults also benefit from environmental controls. Instead electronic information is most often preserved from medium decay simply by copying the information from the decaying medium to a newer one of the same kind. This technique is called "refreshing" the medium; we speak of refreshing a tape by copying

its contents to another similar tape. (In the current climate of protection of intellectual property rights, copyright concerns must be noted even when the intent is simply to make a replacement preservation copy.) Refreshing of information on microcomputer and server discs is most often accomplished through backup/restore techniques when a failure occurs.

Technology Preservation

More problematic than medium decay are the rapid changes in the means of recording, in the storage formats and in the software that allows electronic information to be of use. This is "technology obsolescence", and it is essential to recognize that the greater attention must be directed to the obsolescence of technologies than simply of the media**. These obsolescence include the means of recording, the storage formats and the software that allows electronic information to be of use.

Rather than simply refreshing, we also need to speak of "migration": moving information forward through technology stages as they become available and as the old technologies cease being supported by vendors and the user community. At the simplest level, for example, it makes little sense to preserve a 200-bpi tape by refreshing it to another tape of the same obsolete density when it can be copied to a contemporary cartridge format or perhaps to a magnetic disc.

Migrating information forward through computing and software technologies presents the greater migration problem. Files currently exist of information created on earlier microcomputers using obsolete word-processors and operating systems (e.g. WordStar running under CP/M on an Osborne). For optimal preservation of the substantive information, the data should be migrated forward to a current word-processor running on current systems. Similar cases exist for earlier computer-assisted design systems or economic databases, on mainframes as well as on micros.

It is still a matter for discussion which of several options should be followed in migrating information:

- o migrate information through successive technologies as they appear, or migrate older information to a current technology only on demand?
- o migrate information forward to a successor application program, or maintain the original application program and use software technology to make it usable in later technology (there are intellectual property considerations here)?

Each of these approaches has its own proponents, its technological problems and its costs, and experimentation will be desirable.

What is not likely to be acceptable is the reduction of complex formats to simpler formats in the vain hope that the information will thus have a longer effective life. Reducing databases or interactive documents to simple ASCII format forces a loss of functionality -- nothing has been created in simple ASCII for many years, for good reason -- and in many cases will make the information excessively difficult to manipulate or otherwise useless to the intended user community. ASCII is no longer a lingua franca; it is a crippler of information.

Similarly, it is unwise to propose reduction of data formats to a few "standard" formats in the hope that this will further use over a period longer than a very few years. Given formats may be upwardly compatible for some time, but a step-function of increased system functionality often has the effect of making previous formats suddenly non-functional or obsolete. Imaging formats have been particularly volatile in this respect, but even the most common database and word-processing formats have antecedents less than 15 years old, and it would be a bold administrator indeed who suggested (say) WordPerfect v. 1.0 as the "standard" or "simple" basis for preserving word-processing formats. There are few if any data structures beyond the 8bit byte that one can have confidence will still be supported 20 years from now.[6]

Intellectual Preservation

A third preservation requirement addresses the integrity and authenticity of the information as originally recorded. Preservation of the media and of the software technologies will serve only part of the need if the information content has been corrupted from its original form, whether by accident or design. The need for intellectual preservation arises because the great asset of digital information is also its great liability: the ease with which an identical copy can be quickly and flawlessly made is paralleled by the ease with which a change may undetectably be made.

Here are some of the questions that arise for a researcher using electronic information: How can I be sure that what I am viewing is what I want to see? How do I know that the document I have found is the same one that another has read and made reference to in her footnote? How can I be sure that the document I now read has not been changed since the last time I read it? Note that in these cases backup is not the issue; rather, it is how we know which version we have or do not have.

Three kinds of possible changes:

Accidental change:

for example data loss during transfer, accidents during updating, or saving the wrong version.

Intended change (well-meaning):

- o New versions or drafts (authorial texts, legislative bills);
- o Structural changes: updating a vendor register or a Departmental directory;
- o Interactive documents, e.g. hypertexts (or Lotus Notes files) with note-taking capabilities.

Intended change (fraud):

The example of one's own work to cover one's tracks or change evidence; or of anther's work. Possible examples: political papers, laboratory notebooks, historical rewriting, legal documents, contracts. Solutions involve hashing techniques and cryptography (even though the end result is not encrypted but in the clear).

Whatever technique is used must provide generality, flexibility, ease of use, privacy protection where desired, openness of documents where desired, and low cost. Most important, it must also function over long periods of time on the human scale, that is, after individual human actors are dead. Institutions providing access to electronic information will have to realize the extent of their commitment to assuring the integrity of the information they make available. With print this has not been a concern. In electronic form it always will be.

Organizational issues of electronic preservation

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. The Task Force on Archiving of Digital Information recently noted the significance of organizational issues when it said "the key that unlocks the path to the economies of the digital environment is not technological, but organizational".

Organizational Commitment

The permanent assignment of staff responsibility for the provision and long term maintenance of electronic information will be required. There is

no single artifactual parallel for this responsibility: circulation, stack maintenance, preservation and physical plant departments now share it for print. Nor are there present parallels in academic computing centers, where staffs typically focus on technological advance and availability, leaving data to the users.

Fiscal Commitment

The permanent electronic depository will require assured continuity in operational funding. Almost any other library activity can survive a funding hiatus of a year or more. Acquisitions, building maintenance, and preservation can be suspended, or an entire staff can be dispersed and a library shut down for several years, and the artifactual collections will more or less survive. Digital collections however require continual maintenance if they are to survive more than a very brief interruption of power, environmental control, backup, migration and related technical care. Long term funding will be required to assure long term care. Institutions will need to develop new fiscal tools and use familiar fiscal tools for new purposes.

Institutional Commitment

The most difficult requirement will be that of conscious, planned institutional commitment to preserve that part of human culture which is in electronic form. The understood and willingly undertaken responsibility to provide institutional continuity will be the best assurance of the institution's ability to carry out its archival role. The institutional commitment will have to be clearly and publicly made if researchers and others are to have confidence that a given institution is indeed likely to exist for the long term. It is likely that adherence to a standard public institutional agreement to provide long-term electronic information will be required in the future for their libraries to be credibly regarded as archival repositories.

Current Situation

The current situation in electronic preservation can be described as one of preparation rather than practice. A few institutions have claimed from time to time to be maintaining long-term electronic archives, but have provided no substantiating information about standards they are following or of the formality of their institutional commitment.

There are however several significant research and development efforts taking place in planning for the preservation of digital information which will provide resources and techniques for depository libraries seeking to provide permanent access to electronic government information. They include:

The National Digital Library Federation (NDLF)
Formed in 1995, the NDLF describes itself as "fifteen of the nation's largest research libraries and archives [which] have agreed to cooperate on defining what must be done to bring together -- from across the nation and beyond -- digitized materials that will be made accessible to students, scholars, and citizens everywhere, and that document the building and dynamics of United States heritage and cultures."

The formation of the group was one of the most significant among libraries in years. Among important areas for discussion and development the NDLF has targeted is archiving: "Three important issues for NDLF consideration that surfaced ... are those of migration, certification of archives, and the fail-safe and/or rescue function." Others include rights, naming conventions, economics, interoperability, and discovery and retrieval. Some of its activities are documented in the newsletter of the Commission on Preservation and Access and at the Web site (http://lcweb.loc.gov/loc/ndlf/).

Preserving Digital Information: Report of the Task Force (CPA/RLG)
This report is also a significant milestone in library preparation for archiving issues, and involved experts from a umber of research libraries and elsewhere. At the end of 1994 the Commission on Preservation and Access (CPA) and the Research Libraries Group (RLG) created a Task Force on Archiving of Digital Information charged with investigating and recommending means to ensure "continued access indefinitely into the future of records stored in digital electronic form." The 21-member Task Force, co-chaired by Donald Waters, Yale University, and John Garrett, CyberVillages Corporation, have completed their final report, which is widely available online (http://www.rlg.org/ArchTF/) and in print.

Studies in Scarlet: The Research Libraries Group
RLG determined in 1995 that beginning a practical effort in creating a digital collection would help answer a number of questions that would arise. Their practical approach is focused, at present, on "Studies in Scarlet: Marriage and Sexuality in the United States and the United Kingdom, 1815-1891." It involves the digitization of existing materials rather than the creation of new electronic information, but anticipates having to deal with the many issues of long-term preservation in digital form as it sets up an archiving server to provide the information (RLG's Arches project). The RLG development is intended as a model for other institutions to follow rather than as a locus itself for a digital library. The Project is described on the RLG Web site (http://www-rlg.stanford.edu/strat/projdcp.html)

The CIC Virtual Electronic Library
The Committee on Institutional Cooperation (the Big 10 universities and the University of Chicago) have in the past two years developed the

beginnings of their Electronic Journal Collection and what they propose to call the Virtual Electronic Library. The thirteen institutions are able to bring considerable competent brain power to bear on the collective solution of electronic library and archiving problems. Among other activities, they have formed a CIC Task Force on Preservation and Digital Technology, and have discussed a proposition to form an Ad Hoc Task Force on Federal Information explicitly to deal with electronic information. See the CIC Center for Library Initiatives Web site (http://cedar.cic.net/cic/cli.html).

Consortial Efforts

The cooperative nature of each of these projects underlines what is likely to be true for future efforts in electronic preservation: it will be a consortial or cooperative activity, with no single institution, however well situated, capable (or likely willing) to preserve all electronic information on its own. This will be true for Federal information as well.

Potential Models & Strategies

Locus of responsibility for preservation

Two models suggest themselves for locating the responsibility for long term preservation of electronic documents: a Federal, centralized responsibility or a distributed, regional, shared responsibility.

Centralized Federal Responsibility Model

This model (suggested, in fact, by the GPO Report) would theoretically incur a lower social cost, as a single agency would be responsible for the systems and personnel costs associated with assuring longevity of the document record. It would have the advantage of simplicity, in that one would know where to go for the permanently available assured copy of a document.

There are significant liabilities in such a model. It assumes a higher level of willing coordination between government agencies than has been true in the past, even in the print environment; without essentially perfect coordination electronically produced documents would fall between government cracks. It also assumes the willingness and ability of a single agency to take on such a task while at the same time assuring immediate and permanent access to the information preserved; it is by no means clear that government authorities see the latter complex and expensive task as their responsibility, nor whether there is an agency ready to take on the former task at the level required.

Finally, the issue of credibility of government must be faced. Of greater popular concern, though probably lesser in reality, is fear of government malfeasors modifying information for personal or political gain or for Big-Brother-like manipulation. Of greater likelihood, particularly evident in the current political climate of reducing support for government functions, is an unpredictable hiatus in government funding at some future date placing in question the survival of all centrally-preserved documents. Though the likelihood of either eventuality is low, it is probably prudent not to rely on a

wholly centralized protection method at current levels of confidence in the center.

Distributed, Regional, Shared Responsibility

A more practical and expedient model calls on a variety of public and private agencies to share the responsibilities and costs of assuring the survival of government information. Though the total social cost is higher, due to necessary redundancy and interconnections, the costs can be shared and thus may be more socially acceptable. (It should go without saying that Federal agencies, such as NARA and the Library of Congress, would be welcome participants in a shared scenario.)

Redundancy is in fact a desideratum. Disaster protection is one obvious reason: a fire in a single institution's computer room will simply be less disastrous than a fire in a Federal archive that is the lone central archive of the nation's records (this is no less true today with the printed materials distributed in the FDLP, and there is no reason it would not be true for electronic information). Redundancy provides further assurance for document credibility; though integrity techniques can be implemented as described above, comparison of authenticated copies of critical documents may provide further levels of assurance.

Multiple locations will facilitate speedy access. The balance between storage capacities at server locations and network bandwidths between them will shift frequently in the coming decades as usage grows and technology changes. Geographically-distributed centers of document preservation and access are likely to prove helpful as users all over the country search for government information. How many locations will be enough to satisfy redundancy needs for a given set of electronic documents? Two is probably too few, and 54 is too many. Experimentation, driven by local concerns, will provide the answer.

The best reason for shared responsibility however is the availability of talent and energy in the institutions which will share it. It is desirable for many information professionals to be active in developing preservation systems, for it is presently impossible to predict which set of techniques will be the best. Library and information systems development has thrived best in the past because of multiple, competing development paths, and this is likely to be true in the provision and preservation of electronic government information.

Principles of redundancy

What principles should be followed in establishing multiple preservation sites (and thus access points)? Several models immediately present themselves:

Broad subject categories

"Topical depositories" could be based on science, agriculture, business, consumer information, medicine, and the like (networked information reduces considerably the need to think in terms of Regional Depositories). Several institutions with a particular interest in a given subject area might decide to work together to provide geographically distributed and secure sources of authenticated information. For example: Agriculture, for which Cornell University, the University of California at Davis, and the University of Illinois might take on the development of support systems and the provision of multiple sites for government information on the topic. The advantage of subject collocation for such material would be partially balanced by the disadvantage of not always being sure in which subject area a document might fall (agriculture or business?; agriculture or medicine?).

Issuing agency source of information

Archiving institutions might divide up the preservation of government information by the issuing agency, with some taking up NIH materials, some the Department of Agriculture materials, some the materials from judicial agencies, and so forth. An advantage would be a clearly understood set of responsibilities based on the issuing agency name.

Pragmatic issues of file sizes and local institutional commitments

Some institutions, particularly smaller ones, may wish to make specific agreements with other consortial members to share responsibility based on locally-defined issues and capabilities.

Perceived responsibilities and abilities of agencies

Federal depositories presently exist in many libraries: state libraries, state universities, private universities, large municipal libraries and elsewhere. Their different missions will lead them to participate in different kinds of consortia, either with like institutions or with institutions pursuing similar information goals.

It is probable, and probably a good thing, that consortia for preserving and providing electronic government information will be formed on all these models, and on others as well. What is essential is that all the present institutions with federal information experience find ways to participate in the electronic federal information environment, for in fact the provision of government information to their constituencies is part of the mission of each one. Our nation will become the stronger, and our American citizenry

will become more empowered, to the extent our nation's libraries and other institutions take up the preservation challenge of electronic information -- and of federal information.

Recommendations

- o Institutions need to form consortia or other cooperative arrangements to share the responsibilities and costs of preserving federal information.
- o Consortia need to develop new paradigms for electronic preservation that leverage the strengths of information technology.
- o Consortia need to negotiate with the federal government the terms on which they will provide the preservation function for federal information.

Implications: NIDR
Overview
Introduction
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 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.

This section discusses issues related to Networked Information Discovery and Retrieval (NIDR) -- the mechanisms 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. Among the challenges today are:

- the development of taxonomies for networked information;
- classification schemes to describe the content of networked information resources;
- evaluative information beyond descriptive cataloging which has been part of the bibliographic apparatus of print collections;
- 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.

Print-based federal information has its own set of approaches to catalog, index, and retrieve government publications. Networked federal information organization and retrieval systems, however, are in great flux -- in the same way that federal information creation and distribution are equally dynamic.

Issues
The Challenge of Federal Information
Federal information presents an unique challenge 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 reliable sources 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. In addition, a proliferation of Web sites which serve only as pointers is resulting in duplication of institutional efforts and confusion for the user. Although *GPO Access* is the legislatively mandated, centralized point of

entry for electronic federal information, it is not, to date, the sole entry point nor is it a comprehensive site for all federal information.

The Search and Discovery Process
Using a variety of access methods to gather information, whether in the traditional or networked environments, is part of the iterative process of doing research. The researcher may first cast the net wide to gather as much information as possible and then refine and shape the nature of the research as experience and knowledge are gained. However, the research process should not be hindered by the access methods themselves. There is a great need for the development of seamless interfaces so that users can productively spend their time with the content of the information rather than trying to successfully navigate myriad, dissimilar information spaces.

The Search Process
In the search process, the user is trying to locate a known item or object whether it is a particular government report, a specific piece of legislation, a new regulation, or the latest figures on unemployment. There are several searching advantages to having information available in electronic form:

- full text or string searching allows the user to locate words or phrases within the text of a work;
- algorithmic ranking of documents matches word or phrase frequency with the user's inquiry and then ranks the results;
- thesauri provide controlled vocabulary entry points for users and also have the capability to automatically modify queries during retrieval;
- tree searching or other links automatically and seamlessly parse users' inquiries. (Marchionini p. 36)

What happens when the user is looking for a specific statistic such as the per capita income of the U.S.? In the traditional environment, the user could easily find this information in the *Statistical Abstract of the United States* ; provided that the user knew of this resource, was able to locate it through a tool such as *American Statistics Index (ASI)*, or consulted an information specialist for assistance. On the Internet, this can be a frustrating and lengthy search that can be further compounded if the user does not know which agency produces this statistic. Depending on the search engine is used to execute this keyword search (e.g. Infoseek, Lycos, Yahoo, etc.), the number of hits could range from 29 items to 40,000 items. The user must then wade through the results hoping to find the statistic. To complicate matters, there are many false drops and while numerous references define per capita income, the statistic itself was either buried in the retrieved list (not in the top fifty) or not retrieved at all. A statistic that took only a few minutes to locate in the traditional environment, can now stretch into fruitless hours of searching via the Internet.

Discovery
Browsing is an integral part of the research process. Serendipity, exploration, and contextualization support the connections between known concepts and their

linkage with new ideas. Who hasn't browsed bookshelves and happened upon relevant information to satisfy curiosity and/or an information query? Take the example of a researcher beginning a search on Medicare reform. A traditional subject search via the *Monthly Catalog (MoCat)* might be the place to begin, leading him/her to such agencies as Health and Human Services, Congress, the Executive Office of the President, and even the Census Bureau. The researcher might also want journal articles and other treatises and s/he would, therefore, also search a library catalog, a specialized index, and/or a commercial service like *Lexis/Nexis*. Footnotes and bibliographies would lead the researcher to still more resources as the search is shaped and refined through experience.

Using a broad search engine such as Yahoo or Alta Vista to begin a search, might be analogous to using a general index such as the *Reader's Guide* . The user might find some government information resources using each search engine, but not the wealth of subject specific federal agency materials that resources such as *MoCat* or CIS's *Congressional Masterfile* index. These indexes provide access to authoritative federal information that, for example, a teacher can browse for use in classroom instruction or can put a citizen in touch with federal information resources to meet a specific need. The networked environment offers the possibility of enhanced browsing provided that the user first finds an appropriate set of federal information materials. Then s/he can use hypertext links to find additional relevant material.

Current Situation

Today users are getting access to the federal information they need in a variety of ways using a combination of Internet and non-Internet methods.

- Direct access can be obtained through:
  - the Internet via a search engine;
  - knowing the URL of an agency or of a specific item through some mechanism as word of mouth, via a newsletter or listserv, etc.
  - direct agency contact, or through clearinghouses, such as the National Technical Information Service (NTIS) or the GPO Sales Program

- Many users, however, are going through an intermediary:
  - a federal depository or library, relying on the expertise of the information specialist to assist them in searching either/both the Internet or other traditional and electronic non-Net indexes.

While the direct access method works well for known item searching, the intermediary route provides the user with additional strategies for subject-oriented searches. Users should be able to find the information they need without knowing a URL to locate it.

What does NIDR mean in Federal Information?

In the traditional, institutional setting the search and discovery process for federal information has been enhanced by a variety of tools beyond *MoCat* (e.g. finding aids, catalogs, guides, and indexes) that organize and locate the information for the user. These various types of tools can be used at different points in the search process:

- for a known item search the user might consult *MoCat* and/or a source such as Commerce Clearinghouse's *Congressional Index* ;
- for a broad subject search the user might first consult a general source such as Congressional Quarterly's *CQ Almanac*and then move to a catalog and/or index;
- for a narrow subject search the user might directly consult a specialized index such as CIS's *American Statistics Index*(ASI).

It is perhaps because of these time-refined and well-controlled information resources and services that users have grown accustomed to using that have contributed to the high, if not higher, expectations of easy information retrieval via the Internet.

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. Traditional search methods to find federal information often takes patience and persistence as the user consults a variety of tools, but 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.

What is needed to improve searching in the networked environment?
What are some of the issues that affect searching on the Internet? Referring back to the per capita income scenario, if the user knows which agency produces per capita income s/he could go to that agency home page and find the statistic.

Since the statistic is available on the Internet, why wasn't it found when doing a keyword search? Some of the contributing reasons are:

- First, it depends upon how much of the Web site is indexed - full-text (few do this) versus URL, title, etc.
- Second, does the search engine used index only Web sites or does it also search gopher spaces and ftp sites? The Bureau of Economic Analysis, which produces per capital income, has, at this time, a gopher site in addition to its Web site.
- Third, there has been little progress in the indexing of non-textual materials. "Most non-textual objects are located through textual descriptions or linear scanning." (Marchionini p. 36)
- Fourth, what is the search engine's capability to rank items using statistical/probabilistic algorithms and then to return relevant results? These algorithms use word and/or phrase frequency to match queries with items. Additionally, what does the user need to know about this concept to search efficiently (e.g. Clinton and Bosnia Bosnia will return more relevant results in some search engines than just Clinton and Bosnia)?
- Fifth, what are the options within the search engine for the use of controlled vocabulary (i.e. thesauri) to find variant forms of words or to contextualize concepts (e.g. when searching for death penalty, documents with capital punishment should also be returned; per capita income might also be listed as per capita personal income)?
- Sixth, if the user can get to the appropriate agency site, does it have internal, searchable indexing? The Census Bureau's site offers searching by subject, keyword, and geographic location; not all federal Web sites offer this enhanced feature.

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.

Potential Models & Strategies
Where does this leave us?
Who is planning for network-wide federal information indexing? There is no current service that has organized appropriate and significant Internet and Web federal information resources that also offers the value-added structure, context, and level of specificity and description with a browsable scheme. GPO is attempting to do this with its evolving Pathway Services project. Its Browse function offers three features - Browse Titles, Browse Topics, and Browse GILS records; the Search function offers two features - search *MoCat ,* with some URL hot links to federal sites, and the Pathway Indexer, utilizes a robot that crawls federal Internet sites building a database that allows keyword searching.

CIS and other commercial publishers are also entering this arena. CIS and Lexis/Nexis have recently developed a product that will provide access to digital government information. The *Compass Library of Government Information* will initially focus on legislative materials and will expand to include statistical, regulatory, judicial, and executive branch materials. Users will gain access through CIS's Web site which will apply a common user interface with content-specific help. Users will also be able to manipulate and combine data for incorporation into their own work.

Much traditional government document searching has depended on knowledge of agency structure and publishing, but information professionals should be able to leverage the strengths of technology in the development of efficient electronic search and discovery tools. At this time however, mechanisms for locating federal information on the Internet are rudimentary at best and are less adequate than systems for other media. Organization and indexing are chaotic making access haphazard for the user. The user must first be able to identify if there is networked information available to satisfy his/her need. Once that piece of the puzzle has been determined, the information needs to be in a usable format.

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.

- There needs to be a designated, authoritative first entry point. The user could then be assured that authentic federal information would be found on that site.
- The site should also have well developed tools to use for finding information. First, there needs to be good menus to guide users in the initial phase of the search process. Second, there should be powerful searching mechanisms that permits parsing by subject and agency for efficient searching.
- There is also a need to go beyond doing Boolean searching on the URL, the title, and the summary text. If, for example, a user is looking for per annum wheat production and that statistic is not listed in one of those fields, the information will not be found.
- The user should also be able to refine searches performed by further using the results already obtained. This could be accomplished through Boolean, proximity, and fielded searching capabilities. Total natural language searching or artificial intelligence (AI) will not lead unsophisticated users to what they need.

Institutional strategies to assist users also need to be developed. There is a need for information specialists to create customized indexes and guides by subject, geographic location, etc.

- One strategy would be for institutions to develop one well-constructed home page that could serve as a gateway to authoritative, top node, federal agency sites.
- Another strategy, would be for an institution (or group of institutions) to identify the major subject interests of its user community and then develop front ends that might mirror, point to, etc. specific federal information sites that could fulfill the majority of the information needs of their clientele. For example, a citizen community networking coalition might provide access to such sites as the home pages of the White House, the Social Security Administration, the EPA, the National Center for Education Statistics, and the Small Business Administration. This configuration would provide an organized starting point for citizens to access a wide spectrum of information.

On such institutionl strategy is INFOMINE, a Net resource management tool and search engine, devloped at the University of California Riverside initially for use by the entire University of California system. INFOMINE (lib-www.ucr.edu/) is a hypertext based system with links to resources of research value to the academic community. It uses a search engine with multiple access points (e.g. title, subject, keyword) through a subject based home page format. Indexing terms are assigned by a team of information professionals to provide more focused search results.

Two other interesting NIDR projects are being carried out at the University of Queensland, Australia which make searching large information spaces easier. The HyperIndex Browser (HIB) (www.dstc.edu/au/cgi-bin/RDU/hib/hib/) clusters results enabling users to refine their inquires. For example, going a search on the term "government information" provided links to both broader and narrower related terms (i.e. library government information). Their second prototype, HotOIL (www.dstc.edu/.au/cgi-bin/RDU/hotOIL/hotOIL/), uses the HIB but enchances it by accessing a number of sites (both HTTP and Z39.50) and merges the results.

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. Unorganized information has little immediate or long-term value. No matter how networked an institution is, poorly organized and inaccessible information is still poorly organized and inaccessible information. 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.

Recommendations

- ___Institutions need to be active participants in NIDR technical and policy developments that will help realize the potential of the Net as a viable tool for federal information research.
- ___Institutions need to rethink how they deliver federal information to their constituencies and develop tools and network strategies that will provide an organized entry point for their clientele.
- ___Institutions need to support the development of standards that will facilitate network-wide indexing and representation of federal information resources.
- ___Institutions and/or consortia must advocate the need for a comprehensive, authoritative, integrated set of references and search mechanisms for all electronic federal information.

Implications: Services*Overview*
Introduction

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 of 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. In addition, these information professionals help a diverse group of people with a wide range of needs and research skills including:

- a sociologist who needs assistance with the intricacies of tracking welfare legislation;
- a panicked Sophomore who must write a speech on "gun control" for the next day;
- the average citizen, possessing little familiarity with research and an idea for an invention who needs to search complex patent literature.

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

Issues

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. As noted by Johanna Sherrer in "The Implications of New and Emerging Technologies on Reference Service," the traditional reference service model is based on ideas that are no longer valid in a networked environment.[7] 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.

Current Situation

The advent of networked federal information presents exciting opportunities for access and use of widely distributed government information to libraries, offices, homes, and schools across the country and the world. Access to large amounts of federal information is no longer bound by proximity to large documents collections in federal depository libraries. With the proper equipment and knowledge:

- a citizen in his or her study can follow the workings of Congress by using Thomas (http://thomas.loc.gov/) or GPO Access (http://www.gpo.gov/su_docs/aces/aaces001.html);
- a high school student in a classroom can find current demographic statistics about his/her state through the Bureau of the Census's home page (http://www.census.gov/);
- an academic economist can find background information about the economy of Argentina by using STAT-USA (http://www.stat-usa.gov/)

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, these experts will employ their knowledge and skills in new and different ways. For instance, users will likely encounter information that has not already been classified, analyzed, or interpreted for them, as was once the case, and staff will need to assist them in using and manipulating this raw data. This technological change raises a number of issues for institutions that provide services related to federal information.

Researchers' and Citizens' Needs
The expansion of networked information will empower citizens to use government information from their workplaces and homes. This opens a new field of public service for institutions to assist users who are beyond their walls, but within the electronic reach of the institution. As most government information migrates to the networked environment, researchers' and citizens' needs will change from requiring access to large physical collections of documents to requiring convenient access to digital collections and to reference services for networked information.

Information professionals should capitalize on their organizing strengths by creating welldesigned gateways (in essence, electronic guides or pathfinders) to networked federal information as part of their service. There are numerous examples of well designed Web pages that help users focus on the variety of federal and other government information currently on the Net. The sites logically organize the vast amount of information available for the user and provide that filtering function of sifting through the glut of complex information available that reference librarians have always used when assisting their clientele. Some good examples include:

- University of Michigan's Documents Center (http://www.lib.umich.edu/libhome.Documents.center/index.html) provides a subject-based locator to information;
- Villanova University's "Federal Web Locator" (http://www.law.vill.edu/Fed_Agency/fedwebloc.html) offers a comprehensive agency-based approach;
- Mansfield University (http://www.clark.net/pub/lschank/web/gov.html) offers a regional model by providing access to Pennsylvania census information;
- Federal Information Center's "Frequently Asked Subjects" (http://www.gsa.gov/et/fic-firs/fic-faq.htm) provides answers to often asked questions about the services of the federal government.

However, these examples are basically volunteer efforts and this fact raises questions of institutional sustainability and continued support for such initiatives. Will these centers be overwhelmed by e-mail questions and/or hits on their servers? Institutions will need to consider these and other such questions and plan for the impact that expanded service will have on the organization.

There will be other new services needed to meet the diverse needs, expectations, and skill levels of information users. Most users will require an increasing amount of technical support services ranging from the most basic computer questions to sophisticated queries relating to hardware, software, and telecommunications. Many researchers and citizens will have requirements that extend beyond search and retrieval services for networked information. For many types of statistical information, users will need considerable assistance with the manipulation of data.

The opportunity to provide 24 hour service can be a daunting task but one that offers creative solutions if the institution is willing to "think outside of the box," as it were, and develop expanded service concepts. Home pages that offer FAQs ranging from basic Internet vocabulary to information about specific collections and services unique to the institution can be valuable reference points for users and can break down barriers in the research process. Searchable ready reference collections with annotations to Net resources can help answer users' specific questions quickly and efficiently as well as to provide them with a place to begin their research.

An example of an expanded service concept has been developed by the Mann Library at Cornell University. Three economic offices within USDA make their report and data series available through the library. Mann Library has made report series available to users through a free email subscription service. In this way, agricultural information is available to remote users who might not otherwise have easy access to this material. Other institutions with unique collections might want to consider offering their users an enhanced service of this type.

It is important to note that the distinction between providing information and providing government services will become blurred. Many government home pages such as the U.S. Business Advisor (http://www.business.gov/) combine information and services at one site. Information professionals will likely receive an increasing number of questions regarding, for example, how to file regulatory documents or how to apply for benefits online. Institutions will need to rethink their service philosophy and decide whether or not to provide this type of assistance which was once the purview of the agencies themselves.

Instruction

Institutions have long provided user-based instruction that emphasized strategies to locate federal information. In the traditional environment, students, researchers, and citizens were accustomed to doing their research in a static collection that had been analyzed and classified in some manner. Research in the electronic environment with volatile resources often available in multiple versions will intensify institutions' instructional responsibilities. Users will need instruction in developing Internet search strategies that are most effective in finding and evaluating government information. Experts can help users determine what information is authoritative and comes from reliable sources.

Service providers will need to rethink their instructional strategies for teaching about networked sources as opposed to traditional print sources. Access to adequate electronic classroom facilities with hands-on capabilities will be important for this endeavor to be successful. The integration into the instruction process of how to manipulate data, as with GIS and data sets, will be critical.

An expanded user base with remote access capabilities offers opportunities for enhanced instructional capabilities. Institutions and academic departments have the opportunity to work together to enliven the learning process with hypertext links to resources found on a professor's course syllabus. The University of Michigan's Documents Center home page offers some interesting examples of this possibility.

Online tutorials and guides can assist users in becoming self-reliant researchers, as well as to alleviate some of the staff burden of explaining Internet basics, freeing them to develop more sophisticated and subject specific tools. Tutorials specifically designed for federal information sources are sorely needed. The Internet Public Library (IPL) (www.ipl.org/) has a classroom category that features introductory tutorials ranging from what the Internet is, to how to navigate it and use its resources. Future plans include an effort to build a "classroom" that will feature interactive tutorials and guides.

IPL and Ask ERIC (ericir.syr.edu/), as part of their missions, are developing and making available basic Internet lesson plans and guides. Institutions could use these as models and then partner with other institutions to build collaborative federal information related material rather than merely duplicating individual

efforts. The Association of College and Research Libraries (ACRL) and the Coalition for Networked Information (CNI) are sponsoring the development of a site for Internet User Education Materials, located at Case Western Reserve University (www.cwru.edu/orgs/cni/base/acrlcni.html), which will make available instructional resources.

Staff Training

Institutions will continue to provide traditional service for government information, but this service will take place in a changed information environment calling for new skills from information professionals and increased institutional investments in technologies that provide access to government information. Information professionals will likely face "information overload" caused by the new and ever changing formats coupled with the evolving content of electronic information. The more diverse the interfaces, the software packages, and the information itself, the more difficult it becomes for the staff to keep skilled in these information technologies.

There is a steep learning curve for much federal electronic information since there is little standardization in search software, much of which is not user-friendly. These electronic resources are being developed in "real-time" and staff are often finding it difficult to incorporate this avalanche of information into their "real-time" service patterns. Today it is extremely labor intensive to assist users in locating information on the Net and this is one of the contributing factors of the changing dynamic of reference interaction.

Another aspect in the effort to provide good service, information professionals will need reliable and high-speed access to the Net and to high-grade desktop equipment. The explosion of full-text sources and databases on the Net demand that service providers have access to downloading capabilities including printers which give acceptable output for complex statistical and tabular data.

In this new environment, the specialized knowledge of staffs who service government information will be increasingly important. Knowledgeable staff will help researchers and citizens focus on relevant networked information by providing quality reference service that enhances simple access by assisting users with effective search strategies to locate information. However, this requires that staff become and keep skilled in these new technologies through training, collaborative efforts (both inside and outside the institution), and continuing education. Although potentially costly both in dollars and in time, staff training forms the foundation of a high quality service program that benefits both the user and the institution.

Data/Research Centers

Institutions of higher education have traditionally been structured with general documents departments within libraries providing general access to government information and specialized data centers, often affiliated with academic

departments, providing the tools necessary for the manipulation of spatial or numeric federal data. In the environment of networked government information, does this division continue to make sense? The opportunity now exists to offer seamless access and service for those using this type of information. These links to both campus centers and to the community at large has the potential of providing efficient and equitable access to a wide body of information. In doing so, the institution will first need to ask who will serve whom and how they will accomplish this? Depending on the information desired, the request could be directed to the most logical entity either through email or some other established mechanism.

With electronic information more widely available, users will likely ask institutions to do more than just provide access to raw data. Institutions need to consider enterprise-wide strategies that will benefit users. One such strategy is that employed by the University of Virginia Library which has six digital centers (electronic texts, digital images, music and multimedia, geographic information systems, and special collections). Another strategy would have institutional units such as libraries and data centers forming partnerships to provide integrated service for networked information.

Potential Models and Strategies

Effective models of service will likely evolve from the experience of information professionals and users of federal information. Some emerging models of service that could meet the needs and expectations of users in the networked environment are:

- A "free" digital model -- This model would put as much information as possible directly into the hands of users without direct human mediation. Information service providers would work with systems personnel to design gateways to federal information that are as simple and easy-to-use as possible and that incorporate online tutorials, virtual help-desks, and e-mail reference service.
- An institutional customized services model -- This model would involve information professionals creating "packets" of information for users that specifically answer complex questions from a variety of networked resources.
- A collaborative services model -- This model would build upon the strengths of a number of institutions who would contribute general guides, tutorials, and other resources and expertise that would build upon a shared vision of networked service. The ARL GIS Literacy Project is one such example of a successful partnership between institutions and those in the public and private sectors coming together for the benefit of the user.

Recommendations

- Institutions need to fundamentally rethink their service policies in the networked environment and develop new service models that embrace and exploit these new technologies.
- Institutions must define the communities for which they will provide service and/or levels of service. They must also advocate for the development of a national policy for a system where all citizens will have continued access to federal information.
- Institutions need to look within their own structures/campuses at disparate resources that can be redesigned to provide seamless access for the user.
- Information specialists must make a commitment to their role as organizers, indexers, and subject specialists and become "active agents" in the development of tutorials, guides, and other tools that will assist users in becoming self-reliant researchers.
- Institutions must invest time and money in staff training so that staff can develop a skill level foundation that will benefit them, the user, and the institution.

Implications: Infrastructure

TO BE DEVELOPED

Glossary
CNI Fed Information Paper
Glossary of Terms

Alta Vista
A Web-based search engine (http://www.altavista.digital.com/) from Digital Equipment Corporation. Alta Vista offers both Simple and Advanced Query functions with Boolean operators.

Bureau of the Census (http://www.census.gov/)
Department of the federal government of the United States that is charged with the gathering, organizing, analyzing and distributing of census data gathered relative to the population of the United States. Subject areas cover a wide range including: retail trade, minority populations, service industries, housing, foreign trade, agriculture, business, construction, manufacturing, etc.

CD-ROM
A disk-based, high density storage technology. CD-ROMs can convey a large amount (600 Mb and up) of data, especially multimedia data, in a convenient and affordable format. Once only a format for audio recordings, CD-ROMs now frequently contain databases and multimedia presentations.

Central Intelligence Agency Web site (http://www.odci.gov/cia/)
The CIA's Web presence makes unclassified information gathered by the Agency available to the public via the World Wide Web. The site contains electronic versions of select CIA publications such as the World Fact Book as well as public affairs information and links to other intelligence agencies on the Web.

CIS's American Statistics Index
The Congressional Information Service offers a number of comprehensive statistical compilations. The *American Statistics Index* is one and offers sources for federal government statistics.

CIS's Congressional Masterfile Index
The Congressional Information Service offers a number of comprehensive statistical compilations. The Congressional Masterfile Index is one of several specialized computer databases and CD-ROM Indices (MARCIVE, Index to United Nations Publications) designed to improve access to federal materials.

Code of Federal Regulations
The Code of Federal Regulations is the official, subject matter order, compilation of the Federal regulations of a general applicability and legal effect, that are currently in force. In accordance with section 1510(d) of title 44 of the U.S. Code, the Code of Federal Regulations is compiled and published by the Office of the Federal Register of the National Archives and Records Administration. The Code is divided into 50 titles by subject matter. Each title is divided into sections. Sections within a title may be grouped together as subtitles, chapters, subchapters, parts, subparts, or divisions. Titles may also have appendices which may be divided into sections, rules and/or forms.

Commerce Clearinghouse's Congressional Index
A commercial guide to the proceedings of the U. S. Congress.

Congressional Quarterly's CQ Almanac
The CQ Almanac takes the entire previous year of congressional action and

distills it into a single volume or set of data. The Almanac provides more than just a brief synopsis, offering a complete reconstruction of legislative action, organized by policy area.

Congressional Record
The Congressional Record is the complete and authoritative record of everything that officially occurs in the Congress of the United States. An Index to the Record exists to facilitate access to the contents. Both are searchable online via Thomas (http://thomas.loc.gov).

Current Population Reports (U.S. Census Bureau)
Data compiled from the most recent Census of the United States that focuses on specific facets of the current population. Joint project between the BLS and Census Bureau. Estimates obtained from the CPS include employment, unemployment, earnings, hours of work, etc. Examples include: Voting and Registration Patterns, Child and Alimony Support, Population Profiles, After-Tax Household Incomes, etc.

Digital Libraries Federation
A possible model project planned and executed by a federation of agencies that includes the Commission on Preservation and Access, Library of Congress, and National Archives and Records Administration.

EDI (Electronic Data Interchange)
EDI deals with a set of data communications standards developed by the American National Standards Institute (ANSI) and overseen by Data Interchange Standards Association, Inc., in Alexandria, Virginia. "Electronic Data Interchange (EDI) is defined as the inter-process (computer application to computer application) communication of business information in a standardized electronic form. Electronic Commerce (EC) includes EDI, but recognizes the need for inter-personal (human to human) communications, the transfer of moneys, and the sharing of common databases as additional activities that aid in the efficient conduct of business."

ERIC
Educational Resource Information Center (ERIC) provides users access to a clearinghouse of information gathered on various focus topics. ERIC is part of the United States Department of Education and collects information under 16 different subject-specific clearinghouses, including: educational management, handicapped and gifted children, teacher education, adult, career and vocational, junior colleges, social studies/social science, higher education, counseling and personnel, elementary and early childhood, information resources, languages and linguistics, rural education and small schools, reading and communication skills, science, mathematics, and environmental, tests, measurement, and evaluation, and urban education.

Eric Forte's "University of Illinois at Urbana-Champaign"
SEE University of Illinois at Urbana-Champaign: Documents Library

FAQs about Social Security
FAQ stands for Frequently Asked Questions. FAQs are a standard part of the
Internet and Web and are usually very tightly focused on a given topic or subject.
FAQs are used as a teaching tool online and are an excellent way for new users
to bring themselves up to speed on any topic, in this case, Social Security.

FBI Law Enforcement Bulletin
The LEB is published monthly by the FBI and is also available on the Web
at: http://www.fbi.gov/leb/leb.htm. The issues contain a variety of law
enforcement related columns and articles, including recent crime statistics and
analysis of those statistics, interpretations of recent court rulings, reviews of
policy and procedure and focused efforts in areas like substance abuse, gang
violence, etc.

FBIS / World News Connection
The Foreign Broadcast Information Service (FBIS) and the National Technical
Information Service (NTIS) have made news and information from thousands of
non-U.S. media sources available to users. The service is available electronically
through a new online service called World News Connection. WNC puts the
results of FBIS' intelligence gathering efforts at your fingertips.

Federal Bureau of Investigation Web site (http://www.fbi.gov/)
FBI's Web presence contains a listing of the "most wanted" criminals in America,
an overview of the Bureau, current investigations, relevant Congressional affairs,
law enforcement support, uniform crime statistics, etc.

Federal Depository Library Program
The Federal Depository Library Program means that certain libraries all over the
United States are designated as "depository libraries." These designated libraries
must provide government information to the public. This means the libraries must
store the government publications in whatever format they are available and also
provide convenient hours of usage, reference assistance in using the
publications, etc. Depository libraries are one of the ways the government keeps
its citizens informed in a timely manner.

Federal Information Center FAQs (http://www.gsa.gov/et/fic-firs/fic-faq.htm)
Like all FAQs, the FIC's Web-based FAQ addresses the most commonly asked
questions related to subjects that come under the federal government's authority.
Examples of commonly asked questions include questions related to patents,
copyrights, trademarks, federal tax preparation, selective service (the Draft and
Draft Registration), passports and immunization requirements, social security,
etc.

Federal Register
Latest regulatory information and important notices from all U.S. Government agencies, including Presidential Documents and Executive Orders as well as Notices, Rules and Proposed Rules from Federal Agencies and Organizations.

FTP
File Transfer Protocol. FTP is a process that allows the exchange of computer files across networks.

GIS (geographical information systems)
Computer systems that deal with spatially-oriented data.

GLOBE Program
The fundamental principles of GLOBE (Global Learning and Observations to Benefit the Environment) were conceived by Vice President Al Gore. GLOBE encourages students all over the world to become involved in collecting environmental data such as temperature, amount of precipitation, pH value of local water sources, land cover, and identifying tree species. The primary objective of the GLOBE Program is twofold: to enhance the students' environmental awareness and to enable them to contribute to the scientific understanding of global environmental issues.

Government Information Locator Service (GILS) (http://www.usgs.gov/gils/)
GILS offers a Web-based service to help users find relevant information maintained by the federal government. In the site's own words, "As part of the Federal role in the National Information Infrastructure, GILS identifies and describes information resources throughout the Federal government, and provides assistance in obtaining the information. GILS supplements other government and commercial information dissemination mechanisms, and uses international standards for information search and retrieval so that information can be retrieved in a variety of ways. "

Government Reports Announcements and Index (GRAI)
The most comprehensive, specialized index to government reports or reports on governmentsponsored projects, agencies, departments, etc. Accessible by keyword, author, subject descriptors. Gives very detailed descriptions and abstracts of each report.

GPO
Government Printing Office. The United States government is the largest publisher in the world and the GPO is the agency charged with overseeing printing, binding, and distributing the publications of the Congress as well as the executive departments and establishments of the Federal Government..

GPO Access (http://www.access.gpo.gov/su_docs/aces/aaces001.html)
A Web-based collection of federal databases and documents, including

searchable access to The Federal Register, Congressional Record, Congressional Bills and other Federal Government information are available online via GPO Access, a service of the U.S. Government Printing Office (GPO). Public access is available through the Federal Depository Library, or directly from GPO.

GPO Report
Recent strategic report on the possibility of reconfiguring the Federal Depository Library program. ** Citation and description to be supplied by author.

GPO Sales Program
A clearinghouse service for the sale of GPO materials.

GPO's nascent Pathway Services Project
**

Infoseek
Web-based search engine from Infoseek Corporation that refers to itself as in "Internet navigation service."

Inter-University Consortium for Political and Social Research (ICPSR)
A collection of computerized social science data sets, the largest from the federal government related to election data, voting records, health statistical surveys. The files are made accessible to member institutions in various machine-readable formats (magnetic tape, CDROM and FTP directories.)

IRM (Information Resource Management)
The integration of IRM planning with federal agency strategic planning promotes the appropriate application of Federal information resources. IRM represents the federal government's intent to include IRM strategic planning to help agencies fulfill their mission and to promote information planning that promotes the use of information throughout its life cycle to maximize the usefulness of information, minimize the burden on the public, and preserve the appropriate integrity, availability, and confidentiality of information.

JSTOR (Mellon Foundation)
JSTOR is an independent not-for-profit organization created with the assistance of The
Andrew W. Mellon Foundation to help the scholarly community take advantage of advances in information technology. The project focuses on core journal collections, storage and access and attempts to consider the various perspectives of scholars, libraries, and publishers.

Lexis/Nexis
One of the major online information services and the frequent choice of librarians and professional sources in search of legal and business oriented information.

List of Classes of United States Government Publications
A classified list of U.S. Government Publications currently available for selection by libraries in the Federal Depository Library Program.

Lycos
Web-based search engine from Carnegie Mellon University that bills itself as the "Catalog of the Internet."

Metadata
Data that defines or describes other data. For example, the data dictionary for a database defines each data element. Metadata often serves as an "envelope" for large data set(s).

Monthly Catalog
A publication of the Government Printing Office (GPO), the *Monthly Catalog of United States Government Publications* is one of the most common ways of identifying relevant government documents available. The Catalog provides a comprehensive subject index to help the end user identify government information published on a specific topic of interest.

Monthly Labor Review
Monthly Labor Review publishes articles by the Office of Employment and Unemployment Statistics that analyze labor statistics as a barometer of the economy.

Mortensen's Federal Web Locator
SEE The Villanova Center for Information Law and Policy : Federal Web Locator

NASA
National Aeronautics and Space Administration. The government agency charged with space programs and related research. NASA is one of the government's largest research agencies and their work includes products such as: Hubbell Space Telescope, Discovery missions, EnviroNET, Jet Propulsion Lab, Kennedy Space Center, Goddard Space Flight Center, Langley Research Center, Lewis Research Center, etc.

NIDR
Network Information Discovery and Retrieval. A relatively new set of tools to help users utilize the wide selection of networked-based information sources available on the Internet and World Wide Web.

NTIS
National Technical Information Service. NTIS provides the user with access to research results from work performed by the United States and by foreign governments. NTIS focuses on Research and Development (R&D) and research

related to engineering. Materials available include written technical reports, CD-ROM databases, online databases, software, etc.

Occupational Outlook Quarterly
Publication on employment and employment projections from the Bureau of Labor Statistics.

OMB Circular A-130
A publication from the Office of Management and Budget that addresses Federal Information Resources. This document contains excerpts from the Federal government's policy directive for promoting the application of technology to improve the use and dissemination of information by Federal agencies. It clearly states (in section 6f) that Federal information must be made accessible to members of the public with disabilities.

Paperwork Reduction Act of 1980 and its reauthorization in 1995
The Paperwork Reduction Act (44 U.S.C. Chapter 35) assigns the Director of the Office of Management and Budget (OMB) responsibility for maintaining a comprehensive set of information resources management policies, and for promoting the application of information technology to improve the use and dissemination of information in the operation of Federal programs. To fulfill these responsibilities, OMB originally issued Circular No. A130, Management of Federal Information Resources (50 FR 52730).

PIMO (Public Information Management Organization)
A proposed component of a newly envisioned Depository Library program. PIMO's would allow for a common electronic interface for federal networked information.

Reader's Guide to Periodical Literature
A traditional resource, now in many formats, that helps users track the appearance on a term or concept in a set of indexed publications.

Sailor (http://www.sailor.lib.md.us/)
Maryland's Online Public Information Network. Sailor is a community-based information network, a statewide telecommunications network that enables Marylanders in all 24 counties to have access to Internet resources through Sailor.

Schankman's U.S. Government and Politics
http://www.clark.net/pub/lschank/Web/gov.html
Rich site for locating and using various government document resources. Links include: directories and locators, federal document librarianship, federal agencies, guides and pathfinders, international sources,legislation and regulation, politics, voting and elections, and state and local government.

Statistical Abstract of the U.S.
Statistical Abstract of the United States is the authoritative source for statistics of all kinds related to the United States and is considered the standard source for a huge range of statistics. Subjects in the social, economic and political arenas are all covered, including major sections on health, nutrition, business, banking, finance, parks and recreation, education, crime, law enforcement, employment, income expenditure, housing, manufacturing, international statistics, etc.

StatUSA (http://www.stat-usa.gov/)
Export and trade database information. Once free, now subscription (fee) based.

Statutes at Large
Compilation of federal law and the complete text of all such laws. Very useful for providing legal citations. The laws are listed in numerical order and the compilation works in conjunction with many other federal publications such as the Congressional Quarterly Almanac and the CIS annual.

Studies in Scarlet
The first Research Libraries Group (RLG) digital collections project. The project was designed to test whether or not virtual collections could be created collaboratively, widely distributed and successfully maintained for long term use. The project is scheduled to cover 2 years and conclude in 1998. The title project: Marriage and Sexuality in the United States and the United Kingdom, 1815-1914. Major participants include: Harvard University Law Library, University of Leeds, New York Public Library, North Carolina State Archives, Princeton University Libraries, New York University Law Library, and the University of Pennsylvania Law Library.

Task Force on Archiving of Digital Information
Part of the work of the Commission on Preservation and Access that focuses on the various technical issues related to maintaining digital information over time.

The Compass Library of Government Information (CIS & Lexis/Nexis announced this)
Proposed by the Congressional Information Service and Lexis/Nexis, Compass would focus on legislative materials and eventually include statistics, as well as materials from executive and judicial branches.

Thomas (http://thomas.loc.gov)
Thomas (named for Thomas Jefferson) is a database of federal information, created by the Library of Congress at the direction of the 104th Congress of the United States. Thomas contains databases related to legislation, bills expected to see action currently, bills under consideration, major legislation passed by a given Congress, access to the Congressional Record and its Index, Committee information, Historical Documents, Congressional agencies and resources, Thomas usage statistics and a guide to querying Thomas.

Title 44 of the U.S. Code
Title 44 of the U.S. Code calls for the Code of Federal Regulations to be compiled and published by the Office of the Federal Register of the National Archives and Records Administration.

U.S. Business Advisor (http://www.business.gov/)
Exists to provide business with one-stop access to federal government information, services, and transactions. The site's goal is to make the relationship between business and government more productive.

U.S. Code
The complete set of public laws and statutes as set by the Congress of the United States. Published by the Government Printing Office and searching in both hardcopy and online.

Uniform Crime Reports
The Uniform Crime Reports are issued annually by the Director of the FBI and various related committees charged with gathering the statistics. The UCR offers a compilation of all crime statistics gathered by the nation's various law enforcement agencies at all levels, giving an excellent overview of violent and non-violent crime in America. The statistics are also broken down by population groups and regional areas. Quarterly versions of the report are also available.

University of Illinois at Urbana-Champaign: Documents Library
http://www.grainger.uiuc.edu/doc/default.htm
Site offers a guide to the University's Government Documents collection, including links to staff and services, GPO programs and services, national and international documents, CD-ROM databases, etc.

URL
Uniform Resource Locator, a URL is the form an address or location takes on the World Wide Web. The format used by URLs is usually" protocol://host address/subdirectory/file name. For example: http://www.law.vill.edu/Fed-Agency/fedWebloc.html.

USDA Economics and Statistics System
(http://usda.mannlib.cornell.edu/usda/usda.html)
An online site from USDA (also available as a gopher site) that divides its data into two distinct categories: Data Sets and Reports. The site is maintained by the Mann Library at Cornell University. The USDA Economics and Statistics System contains nearly 300 reports and data sets from the economics agencies of the U.S. Department of Agriculture. These materials cover U.S. and international agriculture and related topics. Most reports are text files that contain time-sensitive information. Most data sets are in spreadsheet format and include timeseries data that are updated yearly.

Villanova Center for Information Law and Policy : Federal Web Locator
http://www.law.vill.edu/Fed-Agency/fedWebloc.html
The Center hosts the Federal Web locator. The site is intended to be a "one-stop" shopping center for citizens looking to link up with federal government sites and services. Links include National Park Service, NASA Shuttle Program, Federal Legislative Branch, U.S. Department of State, Immigration and Naturalization, etc.

Yahoo
One of the best known of all the Web index compilations, originally created by Jerry Yang and David Filo when the two were Ph.D. candidates at Stanford University.

Welcome to the Documents & Maps Department
Bailey/Howe Library, University of Vermont (Yan Hong)
http://www.uvm.edu/~yhong/govdoc.html
Comprehensive overview of the Government Documents collection of the Bailey/Howe Library at the U. of Vermont. Links include: Guides, CD-ROM Databases, Federal documents, International documents, Legal resources and Map resources.

BIBLIOGRAPHY

TO BE DEVELOPED

FOOTNOTES

[1] Electronic Federal Depository Library Program Transition Plan

<http://www.nal.usda.gov/acq/eplan.htm>

[2] Clifford A. Lynch, "Networked Information Resource Discovery: An Overview of Current Issues" *** get the citation, this taken from a preprint

[3] Agenda for Access; Public Access to Federal Information for Sustainability through the Information Superhighway. A Report prepared by the Bauman Foundation. January 1995. 162 pages.

[4] Ibid., p. 89.

[5] CD-ROMs are presently a significant source of distribution of electronic data, but their distribution is on a medium which is inherently relatively fragile. As networked distribution techniques come to predominate, it will be seen that the long-term preservation solution for CD-ROM information will be to make it available on a networked server which is following networked information preservation approaches which are the focus of this paper. CD-ROMs will not be dealt with further as such in this section.

[6] The move to the 8-bit byte in 1964, the shift from EBCDIC to ASCII through the 1970s, and the advent of the microcomputer in the 1980s, can stand as examples of seismic shifts in the computing world against which it would have been difficult to defend with "standard" data forms. Migration is inevitable.

[7] Johannah Sherrer, "Implications of New and Emerging Technologies on Reference Service," in The Impact of Emerging Technologies on Reference Service and Bibliographic Instruction, ed. Gary M. Pitkin (Westport, CT: Greenwood Press, 1995), pp. 25-47.