Migration of Government Information Products to the Internet

John P. Heintz
University Of St. Thomas - Minnesota, japheintz@stthomas.edu

Follow this and additional works at: http://ir.stthomas.edu/lib_staffpub
Part of the Library and Information Science Commons

Recommended Citation
Migration of Government Information Products to the Internet

John P. Heintz

abstract: Government information traditionally published in printed form is increasingly appearing on federal agency websites. This transition generally results in better access to information. However, these changes raise a variety of bibliographic control and access problems for the Federal Depository Library Program. This article provides an overview of these issues, examines responses to date, and provides the perspective of a government manager on some of the access issues, arguing that rather than fighting the changes librarians need to help the system evolve by embracing proactive solutions.

Context: The Shift from Print to Electronic Documents

The Federal Depository Library Program (FDLP) is undergoing a rapid transition from distributing tangible printed documents to ensuring access to primarily electronic information on the World Wide Web. This article will explore the advantages inherent in the transition, the problems for effective dissemination of public information that it causes, how the players in the depository library system are responding, and suggest some additional possible approaches to ease the transition.

Government information products have traditionally been made available in printed form via sales from the Government Printing Office (GPO) and “free” distribution to depository libraries through the FDLP. Title 44 of the United States Code set up this cooperative arrangement between the federal government and selected libraries across the country to ensure free public access to government information. The system traditionally required federal government agencies to publish (i.e. print) their documents via the GPO, where they are classified with Superintendent of Documents (SuDoc) classification numbers, and distributed to depository libraries for use by the public. Selective depositories choose among categories of documents, while regional depositories automatically receive all documents available. These depository libraries are required...
to maintain the deposited documents, to provide for public access, and to render assistance to library patrons in locating and using the publications.1

Since the early to mid-1990s, coincident with the rapid development of the Internet and widespread publication of federal agency information on websites, the volume of tangible documents sold through the GPO and distributed to depository libraries has declined dramatically. The number of titles classified by the Superintendent of Documents declined from 62,001 in 1991 to 30,124 in 2000, while the number of titles cataloged (MARC records created and loaded to OCLC) fell from 29,426 to 18,552. The combined number of paper and microfiche titles distributed to depository libraries declined from 56,437 in 1991 to 26,994 in 2000, while the combined number of paper and microfiche copies fell from 26.4 million to 11.6 million.2

During this period, the federal agency presence on the Internet has grown exponentially, reducing the amount of material under the direct control of the GPO. At the same time, the FDLP has begun to transform its offerings from tangible to electronic. “Since 1996, the Federal Depository Library Program has evolved toward a primarily electronic program, as directed by Congress. To illustrate, in just the last two years, the distribution mix has changed from 45 percent online and 55 percent tangible in 1999, to 60 percent online and 40 percent tangible in 2001.”3 These statistics do not even begin to tell the story, however, given the large quantity of information posted on federal agency websites that is not formally “distributed” via the FDLP.

Even a casual observer would note that we have witnessed huge growth in government agency presence on the Web for information dissemination and service delivery over the last several years. The General Services Administration claims that its FirstGov federal government web search engine now indexes more than 51 million web pages of government information.4

It is also clear that Americans approve of and make use of government’s web offerings. The Council for Excellence in Government sponsored a poll (released in September 2000) that concluded:

**Americans believe that e-government will mean better government**

- By five to one (56% to 11%), people believe that e-government’s impact will be positive rather than negative
- Majorities express favorable views on all e-government functions, being most supportive of receiving medical information from appropriate agencies (80%), the ability to review candidate’s voting records (77%), access to Social Security benefit information (73%), cost savings from online motor vehicle registration (71%), and online student loan applications (70%)
- Two-thirds think it should be a high or medium priority for government to invest more to make additional information and services available on the Internet

**The public wants to proceed carefully down the road toward digital government, saying that safeguarding security and privacy is their top concern**

- By more than two to one (65% to 30%), Americans want to proceed slowly rather than quickly in implementing e-government due to concerns about security and privacy
Two-thirds are very concerned about the risk of hackers accessing government computers.

Majorities also express concern about the possibility of government employees misusing personal information (55%) and a general potential for less personal privacy (53%).

The public’s vision of e-government extends beyond efficient and high-quality services to a more informed and empowered citizenry and a more accountable government.

Asked to choose the most important of four possible benefits of e-government, respondents selected:

- Government that is more accountable to its citizens: 36%
- Greater public access to information: 23%
- More efficient and cost-effective government: 21%
- More convenient government service: 13%\(^5\)

The large level of support that citizens have demonstrated for expanding access to government information and services on the Internet helps both to explain the fast growth seen to date in federal web information, and to predict that the shift in federal government publication patterns from tangible to electronic formats will only accelerate.

Benefits of the Shift from Print to Electronic Formats

Since these benefits seem so extensive and apparent, I will spend less time discussing them and will focus rather on the challenges involved. First and foremost, publishing government information on the Internet makes it far more accessible to the general public (generally on a twenty-four hour-a-day, seven-days-a-week basis) than it is from a printed document shelved in a depository library. Writing about a U.S. National Commission on Library and Information Science (NCLIS) study of public information dissemination, Henry Perritt says that

The public is best served by a diversity of sources and channels for access to governmental information . . . Decentralizing and distributing dissemination responsibility to the agency level is a desirable development, and it should be encouraged. Decentralized responsibility exercised through the Web is far more likely to result in broad public access than any centralized mechanism . . . As agencies embrace the Internet’s World Wide Web to disseminate their information resources, the accessibility of those resources improves dramatically. Use of the Internet’s and Web’s open platform and universal standards makes it far easier for ordinary citizens to obtain access to public information than would any specialized electronic dissemination arrangement.\(^6\)

A second and related advantage is that web formats can offer more features than traditional printed documents, including text searching, audio and video files, and dynamic interaction with online databases to search for information.

A third advantage is that finding and accessing information through directories or search engines on the Web is likely to be more intuitive for end-users than learning about and navigating the cumbersome SuDoc number notations by which documents in the FDLP are classified and organized on the shelves in many depository libraries.
A fourth potential advantage is cost savings. According to GAO, electronic documents cost less to store, to maintain, and to disseminate; generate no warehousing and shipping costs; and are easier to update with little additional production cost (although they also note that agencies and GPO are incurring significant short-term startup costs for personnel, hardware and software to support web platforms). In fiscal year 2000, the Superintendent of Documents distributed almost 12.2 million copies of 28,849 tangible titles to depository libraries and added 32,306 online titles to GPO Access (the office’s web portal). The printing and reproduction costs of the 28,849 tangible titles were about $13.7 million; operating and maintaining the entire catalog of online titles (193,032) cost about $3.3 million.

Problems Raised by the Shift from Print to Electronic Formats

Observers in the library community see a number of unresolved problems in the headlong sprint to shift government information from print to the Web. These problems fall under the areas of access, security, privacy, intellectual property, and records management.

Access

The access issue has several dimensions. The sheer number of federal web pages and the decentralized nature of web publishing make bibliographic control an elusive goal. As the NCLIS report puts it,

The sheer number of federal web pages and the decentralized nature of web publishing make bibliographic control an elusive goal.

The government’s knowledge assets are currently strewn across broad physical and electronic landscapes of tens of thousands of web sites and millions of web pages, hundreds of thousands of electronic databases, untold numbers of paper and microform document collections, and in countless files, records depositories, clearinghouses, and archives across the country and even abroad . . . Access tools and resources in place for finding and obtaining public information are still a very long ways from enabling citizens to search for and retrieve public information in a simple, cost-effective, reliable, and convenient manner.

The problem of permanent public access (PPA) is probably the issue of most concern to the library community. Under the FDLP, distributing a fixed paper copy of a document to the depository library system ensures that a work will be accessible for the long term. In the electronic environment, websites change daily, and documents posted on an agency website can disappear without notice. “When the government controls the only authentic copy of a document, there is nothing to prevent the government from intentionally or unintentionally corrupting, modifying, or even deleting that document, thus preventing access or changing the historical record.” According to the NCLIS, “public information is not permanently publicly available, and there is no statutory provision making its permanent availability mandatory. Some public information that is posted to agency web sites disappears within days or weeks.” In addition, many are con-
cerned about whether or not agencies are creating permanent archives of electronic documents so that they actually will exist in the future, and whether such archives will become obsolete as technologies advance. “Public information is not always being preserved to safeguard against the obsolescence of the formats and/or mediums in which the products are created and stored.”

Accessibility of government information for the disabled is of concern. While federal law mandates that information be accessible to those with disabilities, current web technologies are often not disability-friendly, and the multiplicity of sites and designers will make it challenging to ensure that accessible web design standards are followed. Finally, the existence of the “digital divide” poses a fairness challenge if most government information and many services migrate to the Internet. People with lower incomes, fewer educational opportunities, and those who are ethnic minorities are less likely to have or use Internet access, and citizens vary greatly in computer and information literacy.

Security

There are also a number of security concerns. The Hart-Teeter poll alluded to one: the risk of unauthorized access and use or sabotage of data. Government agencies such as the Department of Defense have been concerned about such issues for decades, and invest considerable resources in protecting computer systems from intrusion—a task made much more difficult by the evolution of the Internet and the associated interconnected system of networks. Several laws have been enacted that prohibit unauthorized or fraudulent access to systems and establish penalties for violations. Similarly, federal agencies are required to develop and to implement information security plans to protect their systems. The widespread incidence of hacker attacks on governmental and corporate information systems will require continued vigilance to protect the integrity of public and private data stored in these systems.

In addition, recent concerns over the risk of terrorism have federal agencies not only concerned about protecting systems from unauthorized use, but also reevaluating what kinds of information is appropriate for public view. Even before the terrorist attacks of September 11, 2001, the Environmental Protection Agency and the Department of Justice proposed regulations to minimize the likelihood of chemical accidents and to lower the risk of terrorist and criminal activity associated with Internet posting of certain kinds of information, and suggested removing portions of chemical risk management plans for industrial facilities from the Internet in favor of providing more controlled paper access with proper identification.

Since September 11, the environment has obviously changed dramatically regarding Internet access to federal information. Watchdog groups have complained that much information has disappeared from federal websites. According to the Digital Freedom Network, “removed materials include: locations of nuclear power plants; chemical hazard risk management plans; terrain and pipeline maps; and reports related to hazardous chemicals, aerospace research, and environmental issues.” In addition,

Recent instructions from the National Archives and Records Administration (NARA) and the Department of Justice lay down rules for ‘safeguarding’ information that relates
to ‘weapons of mass destruction’ or that ‘could be misused to harm the security of our nation.’ Classified documents that are due for routine declassification will have their classification extended for up to 25 years. Unclassified information can be immediately classified. Even information that ‘could reasonably be expected to assist in the development or use of weapons of mass destruction’ can be withdrawn without notice. The NARA instructions apply especially to electronic documents placed on agency Web sites.16

While reasonable people can disagree about the wisdom and appropriateness of such measures, it seems clear that our society will continue to wrestle with striking a balance between security and freedom of information.

Privacy

The potential loss of personal privacy because of inappropriate use of government information is also a major public concern. The 1974 Privacy Act (P.L. 93–579), as amended and codified in 5 U.S.C. 552a, prohibits agencies’ disclosure of information that is not required to be public under the Freedom of Information Act. The GAO has expressed concerns about poor security management in certain federal agencies.

Government collection of ‘cookies’—files of potentially personal information about a visitor to the site, collected silently and transparently by a website and stored on the user’s computer in the background—constitutes another unresolved area of debate in the privacy arena. It is unclear whether agencies may use the contents of cookies and if so, in what way they may use them . . . little in the way of consensus yet exists on such issues.17

In light of the continued federal push to move government information to the Web, much uncertainty about privacy issues exists, because “despite extensive case law construing the scope and application of the Privacy Act, no cases have addressed how that act applies to agencies gathering and disseminating information over the Internet.”18

Intellectual Property

Many federal agencies, to reduce costs and workloads and to speed up access, have contracted with commercial entities to produce publications or databases of information. There are some who believe that government provision of information unfairly competes with private businesses that provide information services, and thus they lobby for increased privatization of information provision. The federal government directly provides free public access to databases such as EDGAR, Agricola, Medline, ERIC, PubScience (now defunct), and the Federal Register, while commercial versions of such databases also exist (with added features like user-friendly graphic interfaces, advanced searching techniques, etc.). In other cases, because of budgetary constraints, agencies may charge the public for information provided, or enable private vendors to do so. Many observers fear that some privatization arrangements could result in the loss of information from the public domain.19

Copyright issues arise from the publication of government information on the Web. While copyright protection is not available for government-produced documents (17 U.S.C. 105), federal agencies can come into possession of copyrighted materials from a
variety of sources, and that protection should extend to the copyright holder. In addition, the federal government can be a copyright holder if those rights for a particular work are transferred to it. This leads to the possibility of conflicts between the Freedom of Information Act and copyright law. Posting such information on the Internet, or even releasing it upon request, could constitute a copyright violation, even if done in the name of "freedom of information."

Records Management

Oversight of federal records management activities falls under the purview of the National Archives and Records Administration (NARA). Federal agencies are responsible for the preservation of records in their business operations, and must establish standard procedures that specify what records must be maintained. NARA’s website includes this definition from the Federal Records Act:

> Records include all books, papers, maps, photographs, machine-readable materials, or other documentary materials, regardless of physical form or characteristics, made or received by an agency of the United States Government under Federal law or in connection with the transaction of public business and preserved or appropriate for preservation by that agency or its legitimate successor as evidence of the organization, functions, policies, decisions, procedures, operations, or other activities of the Government or because of the informational value of the data in them (44 U.S.C. 3301).

It is difficult to imagine how federal records management activities could be effective if such a definition were interpreted to exclude data and information from agency websites. According to the GAO, records management practices among federal agencies vary widely, some agencies have not established the required procedures, and NARA has not yet provided agencies with clear guidelines on how to incorporate web information into their records management practices. It seems clear that federal websites need to be viewed as important components of agency information resources. “Key issues yet to be resolved concern the appropriate policies for managing federal web sites such that electronic information and documents are timely, accurate, official, acceptable in courts of law, and accessible for historical purposes.”

Change in the Federal Depository Library Program

The GPO has struggled to keep up with this sudden, largely uncoordinated transition. In a 1998 strategic planning document, the GPO introduced a strategy for managing its electronic collection, asserting that it was simply a continuation of the historic role of providing access to printed documents. The collection was to consist of 1) core legislative and regulatory products on GPO servers, 2) other remotely accessible products managed by the GPO or partners, 3) remotely accessible products that the GPO controls bibliographically but which are maintained by agencies, and 4) tangible electronic products physically distributed to depositories.

In the plan’s definition of scope, the GPO largely is focused on the access issue: “The FDLP includes all Government information products, regardless of format or medium, which are of public interest or educational value, except for those products...
which are for strictly administrative or operational purposes, classified for reasons of national security, or the use of which is constrained by privacy considerations.” With this statement of scope (which is based on 44 U.S.C. 1902), the GPO has largely defined itself out of primary responsibility for the security and privacy issues because of the specified exclusions for classified and private documents (presumably the issuing agency makes this determination). This statement and other information in the plan also make it clear that the GPO inclusion of a document in its electronic collection does not meet NARA records management requirements.

The plan went on to assume that a central authority can best provide dissemination and locator services; that agencies must cooperate to identify appropriate products for inclusion; that agencies and other institutions must share dissemination and storage responsibilities with the GPO; and that depository libraries are key agents of the system to provide access, a customer focus, and identify “fugitive” documents.

Approximately two years later, faced with a sizeable reduction in the GPO budget, the Superintendent of Documents decided to accelerate the transition of FDLP to electronic formats. Effective the following January 1, “. . . U.S. Government publications will be furnished to Federal depository libraries solely in online electronic format . . . ”except for a defined list of circumstances. Guidelines for implementation were laid out in a subsequent policy statement issued to depository libraries. This statement was issued in conjunction with a list of essential titles that would continue to be issued to depositories because they “. . . contain critical information about the activities of the U.S. Government or are important reference publications for libraries and the public.” The list includes standards such as the Constitution, U.S. Code, U.S. Congress Serial Set, Code of Federal Regulations, Supreme Court reports, the U.S. Government manual, federal budget documents, various Census reports, the standard general and subject area statistical compendia, and the CIA World Factbook.

To further its access and archiving goals, the GPO is building a network of partnerships, including those with the University of Illinois at Chicago and the University of North Texas to catalog and to provide access to State Department and defunct agency publications respectively. It has also built networks of depository librarian volunteers to monitor agency websites for fugitive electronic documents and to develop and to maintain online pathfinder subject bibliographies of important web sites to use as finding aids.

Maggie Farrell writes that the GPO’s access tools require too sophisticated a knowledge of government operations to be effective. This is probably why alternative access tools such as the FirstGov portal and privately operated U.S. government-focused specialty search engines such as Google and the now defunct Northern Light have proven more popular than the GPO’s search engines, as have government information directory pages such as Yahoo’s or those of several universities. Farrell also worries that dissemination varies too much by agency, suggesting that many will not care about permanent public access and that the GPO needs to better educate agency personnel of its importance.

Duncan Aldrich et al. suggest that depository librarians must rethink their service strategies in the new electronic environment. Building on the well-known precept that people will seek the path of least effort in finding information, they believe that people
will attempt to use the Web to find the government information they seek. They suggest that documents librarians must improve their knowledge of government functions and services to understand where relevant information is likely to be, and that document librarians use that knowledge to help patrons find government information on the Web through pathfinders, web search tools and portals, and specialized reference assistance.

Some observers think that the depository library system is finished: " . . . the depository library system, simply does not work in the new information political economy . . . [its] systemic approach fails to include the elemental interdependence between these distributed information byproducts and the services delivered by the government agencies responsible for their creation. Government publications are most meaningful to citizens when they are closely associated with a specific government program [or] service . . . "

John A. Schuler’s article also speculates that the GPO will evolve into a bibliographic utility for electronic information, or that individual libraries will build more direct partnerships with federal agencies to collect electronic information systematically. In May 2002, the federal Office of Management and Budget proposed allowing federal executive branch agencies to seek printing services from private sector sources other than the GPO, raising concerns in the library community about both bibliographic control of federal documents and acquiring them for the FDLP. This proposal only underscores the notion that traditional depository library programs need to evolve.

The Perspective of a Government Manager

The author has been a manager in Minnesota state government in various capacities for the last seventeen years, with varying levels of responsibility for creating and disseminating budget and finance information products. Schuler has it right in his perspective that government information (at any level of government) is best viewed in the context of the programs and services it supports. Agencies focus on mission rather than on abstract principles of public information dissemination, and managers view information created by their programs in the context of mission and public service.

For librarians, the creation and dissemination of information are core mission functions, and are critical to the effective functioning of democracy. Librarians can get extremely passionate about advocating for the widest possible dissemination of government information, and they should. That’s their job. However, for most public agency managers, information collected or produced by their operations is a tool used to accomplish their goals. In short, information creation and dissemination (unless that is an agency’s primary purpose) are not core mission functions. They will never be considered such, nor should they. Instead, agency core missions are things like preserving fish and wildlife habitat, protecting society by jailing felons, helping people escape poverty, educating children, winning compliance...
with the tax laws, etc. Managers think about information provision in the context of those missions—"What information can I provide that will best help deliver the services for which I’m responsible? How do I let people know about or market my service? What do people need to know in order to participate, to receive benefits, or to do what I need them to do to accomplish the mission?" These questions are very much service-oriented ones, and information provided by agencies on the Web to answer them will often look different from traditional print products. Librarians providing reference services regarding government information may need to think about the products differently—perhaps developing better knowledge about agency operations and services. This will no doubt be a difficult challenge for librarians with little spare time, but it might be possible if time and effort are redirected from traditional document processing activities.

Agencies usually do have some statutory responsibilities for information dissemination, and typically must deposit printed documents with the GPO (or equivalent depository entity for a different level of government), and most try to meet those requirements. But in an environment of scarce and diminishing public resources for many agencies, it is not first priority.

There are several implications in this perspective for librarians interested in improving access to government information and reducing the number of fugitive documents in the Internet age. First, consider easing up a little on the rhetoric regarding the "fight" to preserve access to vital public information against venal or ignorant government employees determined to withhold it. In fact, most government employees would prefer that information be shared, because they are proud of their work; dissemination is just not their top work priority. Instead, the documents librarian community could try some additional proactive approaches:

- Become familiar with staff and programs of an individual agency, and suggest to managers that certain types of data and information may be useful for the general public (in addition to direct program constituents) if made available on the Web. Suggest other information products that may be useful. Use web page monitoring software to keep up with agency website developments, and let the GPO know if you find agency documents that are not part of its Electronic Collection, but should be. The American Library Association Government Documents Round Table’s (GODORT) Federal Documents Task Force agency liaison program attempts to do some of this now.
- Highlight important or new agency information in library pathfinders and directories, and let agency staff know you’re doing so.
- Remember that agency staff work in government because of a public service ethic, and many believe as passionately as you do about the democratic value of sharing information. Find these people in agencies, and cultivate relationships with them. They may take your suggestions and advocate for them internally.
- GODORT or ALA itself (in partnership with the GPO) should author a brief publication explaining the changes in the depository program, the problems of access and loss of bibliographic control (not in libraryspeak like this, but in English!) of government Web resources, how this problem reduces their potential
Web audiences, and explaining how they can submit their Web content to the GPO’s electronic collection. This communication (in print and electronic) should go to agency management, chief information officers, librarians, public information officers, webmasters, etc. who might be in a position to change agency behavior.

- The GPO should consider changes that make it easier to prepare or submit information for inclusion in the electronic collection, such as electronic templates for submission, metadata standards and examples of suggested tags for webmasters, etc.

Some of these suggestions could be done by individual librarians, but many would be most effective if they built on the existing tradition of collaboration among librarians in general and the government documents community in particular, and the existing partnership efforts with the GPO. I am sure that others in the community could suggest additional fruitful efforts.

An additional note about state government information products: everything discussed in this article applies equally well to state and local government documents. If anything, my experience suggests that there is even less centralized review of and bibliographic control over these documents than federal ones. The rush to place agency information on the Internet is equally strong at the lower levels of government, and the same opportunities exist for librarians to work locally with the state’s depository structure and information policy players to improve access.

Conclusion

Obviously, the environment for providing government information has shifted rapidly from print to electronic, and the depository library system is struggling through the transition to define roles for its various players. In my view, the GPO must continue on its current path of attempting to improve bibliographic control of federal web resources and must take the lead to solve the archiving and permanent access problems. Depository libraries should assist the GPO in building the necessary partnerships to perform its tasks (and perhaps help lobby Congress to get the GPO the resources it needs to do the job!), and should shift resources from traditional processing tasks to building more effective reference assistance capabilities to assist patrons in finding and using government information. In this role, depository libraries and government documents departments begin to look a little more like a traditional reference department. Finally, the members of the documents community should continue with existing efforts (and perhaps try some of the suggestions made in this article) to work with government agencies to improve access to vital public information.

John P. Heintz is an MLIS student at Dominican University/College of St. Catherine and formerly was Budget Research & Planning Director at the Minnesota Department of Finance; he may be contacted via email at: jheintz@infionline.net.
Notes

12. Ibid.
13. Ibid., 12–13, and Smith et al., 276.
17. Smith et al., 277.
18. Ibid.
20. Smith et al.
22. Smith et al., 279.
24. Ibid.


