

Documents for a Digital Democracy: A Model for the Federal Depository Library Program in the 21st Century

Interim Summary

Prepared by Ithaka S+R, October 2009

In June 2009, the Association of Research Libraries (ARL) and the Chief Officers of State Library Agencies (COSLA) retained Ithaka S+R to propose a comprehensive framework for the Federal Depository Library Program (“FDLP” or the “Program”) in response to changes in the environment for information dissemination and usage. New approaches must take advantage of the opportunities presented by today’s digital and networking technologies to deliver services to users more effectively, more broadly, and at lower cost. For this project, Ithaka S+R staff interviewed more than 80 individuals from 30 libraries, the Government Printing Office (GPO), and a number of other key organizations. The FDLP serves a variety of needs across a number of communities, and in this project Ithaka S+R has taken a systemwide perspective in an attempt to understand the needs of all stakeholders. This summary presents a high-level overview of the project’s interim findings and recommendations.¹

The FDLP is the mechanism for 1,240 academic, state, public, and law libraries to work as partners in the distribution of federal publications to, and their access by, the general public. The FDLP is administered by GPO under the authority of a federal law that was last significantly updated in 1962. Today, 50 regional federal depository libraries (the “regionals”) collect FDLP materials comprehensively in print format and provide services to the 1,190 federal depository libraries that collect FDLP materials selectively (the “selectives”), in both cases without financial support from the federal government. The documents distributed to libraries are known as “tangible” documents, and this category includes printed texts, maps, posters, microfilm, CD-ROMs, and other miscellaneous formats. For simplicity, in this summary, all such materials are referred to as print documents.

The Program’s core mission of providing “permanent public access” to government documents is just as important today as it ever has been,² and indeed it should play a key role in the Obama Administration’s vision for a more open and transparent government. Access to the workings of the Government by members of the public is imperative to the success of our democracy and to the public’s understanding of their Government. But the incentives that motivated libraries to participate in the Program, reasonably well aligned in a print environment, are increasingly not well matched to the digital, networked environment. In addition, over the last two decades, and especially the last five years, the ways that people store, access, and use information has changed dramatically. Researchers and students are eschewing print for digital access, which is reducing demand for the historical print

¹ A final version of the full report will be released later this fall.

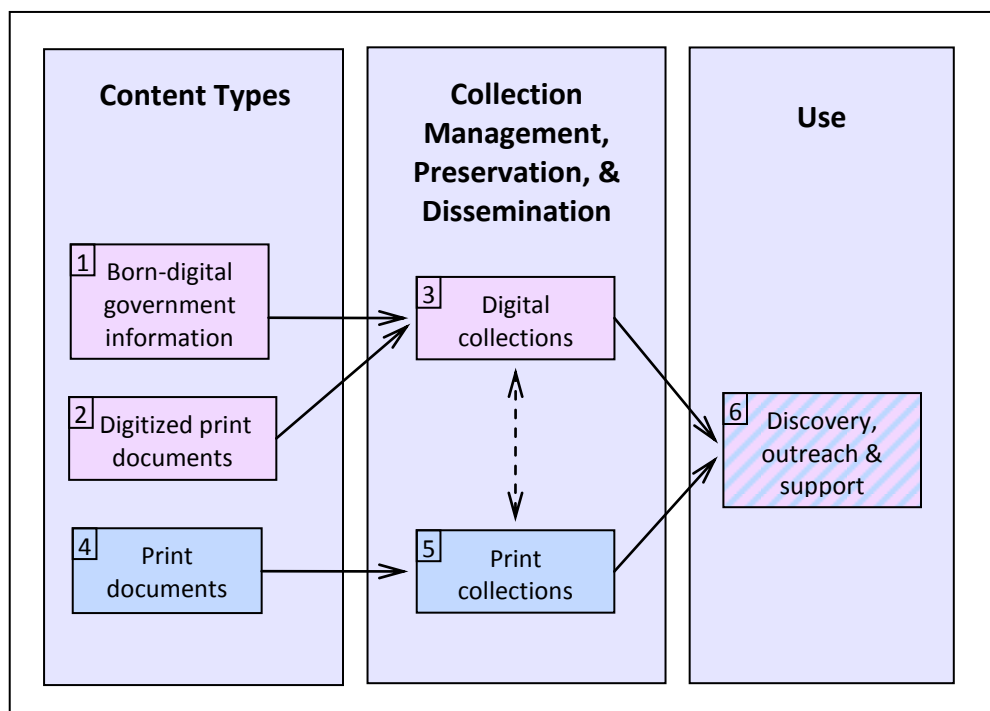
² For the purposes of this summary, permanent public access must provide availability to the general public without any associated fees for access and must provide for appropriate preservation that ensures access over the long term.

collections. For newly issued government information, the ease of dissemination on the web through multiple channels also puts the provision of permanent public access to prospective government information under great risk. If there is strong and compelling interest in continuing the goals and objectives of the Program, the community needs collectively to commit to putting in place a fundamentally new framework for it so that its mission will continue to be fulfilled today and in the future.

The state of the FDLP today

Background research and interviews revealed that there are numerous challenges endemic to the Program today. In order to describe the current state of the FDLP, the Program is broken down into six constituent components, depicted in **Figure 1: Components of the FDLP**. This section analyzes the current state of and challenges related to each of these components as the environment makes the transition to digital. These components collectively comprise a system with significant interdependencies.

Figure 1: Components of the FDLP



1. *Born-digital government information*: Approximately 97% of new government documents are available digitally in 2009, tremendously improving public access, and approximately three-quarters exist *exclusively* in digital form. GPO has developed mechanisms that successfully address previous concerns about authenticity of online documents. However, agencies are increasingly taking responsibility for online publishing without GPO's assistance, with stand-alone documents giving way to

less structured and often more dynamic information sources, and a substantial share of that output is not ingested into GPO's new content management system, FDsys.³ In many cases, the originating agencies manage this information appropriately in another government system, working in formal partnership with GPO. In other cases, however, there is no provision for permanent public access to documents and/or information products. Although GPO has experimented with harvesting strategies to ingest these materials, a scaled approach has not yet proven feasible.

2. *Digitized print documents*: Similarly, many print government materials have been or will be digitized by libraries, by vendors, and by mass digitization programs, dramatically enhancing the potential accessibility of these materials. Unfortunately, some of the most valuable digitized collections are not freely available (e.g. those of commercial vendors) and thus do not contribute to permanent *public* access. The lack of coordination, including the absence of a title-level registry of digitization efforts, has resulted in a set of digitized collections that are far from comprehensive while also containing significant duplication, and they also range widely in quality.

3: *Digital collections management*: The digital materials that are incorporated into the FDLP are managed with the intention of providing permanent public access. Many of these materials are hosted on FDsys, which builds upon the successes of GPO Access in providing permanent public access to digital materials, or through partnership agreements. These programs are generally believed to offer a preservation solution for digital and digitized collections, although external verification of their efficacy and third-party preservation partnerships are needed. What little is known about the preservation state of materials that are held neither on FDsys nor under such a partnership agreement leads to concern about their permanence.

4. *Print documents production*: Although users have demonstrated a very strong preference for accessing materials in digital form, some user needs may be better served by print rather than digital versions. For example, maps, posters, and certain other visually intensive materials may be most easily used for certain purposes in print form. In addition, broadband is not uniformly available even at public libraries in certain areas of the country. GPO continues to produce some print materials, with decision-making about what materials are thus produced remaining, as it always has been historically, distributed across the originating agencies. As a byproduct, decision-making about print production is not sufficiently responsive to user needs.⁴

³ FDsys, the successor to GPO Access, remains under active development and has only been released as a beta product to date. Ithaka S+R anticipates the continuing successful development of FDsys and as such this summary often refers simply to FDsys rather than necessarily reflecting the current transition between GPO Access and FDsys. Congress authorized GPO Access via Public Law 103-40, the "Government Printing Office Electronic Information Access Enhancement Act of 1993." GPO Access launched publicly in June 1994.

⁴ Other examples of why print remains important include brochures and other materials designed principally for the use of the general public as opposed to dissemination and preservation via the FDLP.

5. *Print collections management*: The apparent value of maintaining local print government documents collections has fallen substantially, as usage has declined in favor of digitized versions and volume counts are no longer a primary measure of a library's "value." Many directors, especially at academic libraries, would like to reassign space to higher-value collections or programmatic uses such as information commons and other service-based initiatives. Hundreds of selectives have left the FDLP altogether and many more have reduced their print holdings significantly. Many regional library directors stated a strong need for more flexibility in collections management. And opportunities to provide regional libraries with some additional, albeit limited, flexibility within the current Program structure were identified. In the long term, however, the trend is clear: there will be little remaining incentive for many regional libraries to participate in the Program under restrictive conditions that force them to retain a decreasingly used format. While many state library agencies and public libraries have a mission that is in some cases more suited to the regional role in the long term than is that of academic libraries, they alone could not serve the regional role under the Program's existing paradigms. Although selectives presently depend on regionals to be able to withdraw print materials and for other services, they are unlikely to contribute resources that would provide an incentive for regionals to continue in this role.

6. *Discovery, outreach, and support*: User expectations for disintermediated, seamless discovery environments have skyrocketed in recent years. Government information – in both physical and digital form – has generally been segregated into discrete and fragmented service points, bibliographic infrastructure, and discovery tools. Some librarians are supporting the use of government information well beyond documents and through programs well beyond the FDLP. Still, many users with demand for government information do not succeed in obtaining and successfully using information that would have value to them.

Although the Office of the Superintendent of Documents has worked diligently and with notable successes to migrate the FDLP to the digital environment, the diversity of stakeholders and complexity of the Program has impeded adequate strategic change.

The future of the FDLP

A system that supports the historic goal of permanent public access to government information in ways that are relevant to today's user needs will have to be dramatically different from the Program that operates today. The problems identified above are systematic in nature, and in the long term an adequate solution therefore will require systemwide change in the roles of virtually all stakeholders. A new strategy for the FDLP in the digital era should consist of the following:

In order to effectively serve user needs, our interviewees emphasized that **government information must be made freely available and preserved for the long-term in digital form**. The Program must accommodate the changes in format and dissemination currently underway, to include the transformation of static documents such as PDFs into dynamic sources of government information such as GIS. In addition, the historic collection must be digitized. An effective transition to the digital environment will dramatically expand access and discovery and will enable a wide range of innovative

uses of these materials such as combining data sources and analyzing materials in ways that would have historically been cumbersome or impossible.

Today, far too many library directors feel burdened by the little-used print format, one of many factors that is destabilizing the FDLP. **The historical print collection must be preserved, even though it will play a significantly reduced role for access by users.** Far fewer print copies will be required for many material types, and therefore libraries will be able to reduce the space and redeploy the staff devoted to print collection management to new service roles. Substantial systemwide space savings and processing cost reductions can thereby be achieved. At the same time, providing appropriate levels of print for certain formats or user communities, where appropriate via print on demand, will remain imperative. Overall, if libraries can realign their participation away from print collections management and towards the provision of much-needed services to meet user needs today and in the future, they will have an incentive to continue to play a role in the Program.

With such a realignment, **participating libraries must reemphasize their commitment to serving user needs for outreach, discovery, and access.** As resources devoted to print collections management decline dramatically, user-facing functions will become increasingly important. As the static and collection-driven nature of government documents gives way to a broader understanding of government information, librarians will similarly redefine and expand their role.

A proposed model

Historical collections: digitization and a migration away from print

Although there is great potential value in the historical collections, users will not realize this value if materials are not available in digital formats. **To make the rich historical collections useful, they must be digitized comprehensively and at a sufficiently high level of quality.** There is a strong argument to be made that the government should bear these costs, and Congress should provide such funding at the earliest possible opportunity. Still, with no funding forthcoming, the needs of libraries and users call for action nevertheless. For this reason, in addition to library digitization projects, mass digitization initiatives (involving a number of libraries sometimes in cooperation with Google or the Internet Archive) will be especially important.

It is imperative that the FDLP leverage all possible digitization programs to provide permanent public access to the new format, which can happen via one of two models:

1. Digitized materials should be deposited whenever possible into FDsys, documenting chain of custody and digitization standards, to maximize bibliographic control, ensure ease of access, and provide workflows for associated print collections management.
2. When this is not possible, GPO should provide coordination including standard bibliographic control via formal partnership agreements with other government agencies and outside institutions (such as HathiTrust).

Access will inevitably shift significantly, if not exclusively, to the digitized versions, with the need for residual access to the print varying based on the needs of users and the nature of the content. For the materials that are accessed almost exclusively online, a significant drawdown of print holdings across the selectives is expected, indeed is already occurring at a rapid pace.

Whereas access needs are already being met to a large degree by online versions, thereby allowing for the reduction of significant amounts of print at selectives, the threshold for a digital surrogate to substitute for a print original for preservation purposes is higher. Different thresholds are appropriate for different material types, depending on expected use cases, but high-quality digitization and an adequate digital preservation environment are important in every case.

Once digital surrogates meet preservation thresholds, remaining preservation objectives for the print format will require far fewer copies than are currently provisioned via the regional libraries.⁵ Ironically, if executed correctly, this shift will result in an overall improvement in the preservation of print versions, since responsibility will be allocated differently and resources can be focused on a smaller number of copies. **In the long run, consequently, there should be fewer regionals**, with a probable reallocation of responsibility between academic libraries on the one hand and state/public libraries on the other.

No single comprehensive print collection exists, making the mechanism for the drawdown an especially important choice. Ithaka S+R sees three models for how a reduction in the number of regionals could be pursued:

1. Legislative change could empower GPO to manage a drawdown of materials across regional collections as preservation thresholds are achieved for digitized versions, thereby allowing those regionals to downgrade to selectives or exit the program entirely. Incentives such as funding might be made available to the smaller number of remaining regional participants, particularly appropriate given budgetary swings at state, public, and academic libraries. Legislative change is the preferred alternative, because it will provide the greatest assurance against materials losses while enabling flexibility.
2. Until appropriate legislation is enacted, regionals could work together to coordinate their downgrades or departures, perhaps on a geographical or consortial basis. The regionals would identify the appropriate number of comprehensive collections that would remain, provide a limited window of time for their selectives to withdraw print holdings (since under current law selectives without a supporting regional are not permitted to withdraw materials), and then provide regional services to one another as a drawdown occurred. Proper coordination would be challenging and in the long run there might be inadequate incentives to motivate sufficient regional participation, but this model is legal under today's statute.

⁵ The type of requirements for print preservation and the logic utilized in calculating an appropriate number is detailed in Roger C. Schonfeld and Ross Housewright, *What to Withdraw: Print Collections Management in the Wake of Digitization*, available at <http://www.ithaka.org/ithaka-s-r/research/what-to-withdraw>.

3. Without system-level coordination by GPO or the regional libraries, incentives for many individual regionals to continue to withdraw from the Program will be overwhelming, and other regionals could provide withdrawal support on at least a short-term basis to allow the print drawdown to take place. This alternative is the least desirable, because it will be essentially uncoordinated and could have deleterious effects on preservation, but in the absence of one of the other two models is inevitable.

In order to support digitization workflows and to enable a coordinated drawdown of print collections, additional investment will be needed in bibliographic control. Cataloging efforts should build on the retrospective conversion of the Monthly Catalog of United States Government Publications linked to existing records in WorldCat. Digitization will dramatically improve discoverability, so subject headings and other means to enhance bibliographic accessibility via cataloging will often be prohibitively expensive relative to the value they bring.

Beyond these long-term changes, some pressure may be alleviated in the near term. All regional libraries should, with director-level involvement, immediately review their withdrawal policies and space allocations. There is more flexibility in the Program than many libraries have utilized. For some regional libraries, reducing hands-on oversight of withdrawals and finding creative ways to share responsibility for maintaining print regional collections with others in the region may reduce the pressure they face. While it may not be possible to provide near-term positive incentives for regionals to remain in the Program, everything possible should be done to reduce their costs and enhance flexibility.

Coordinating and preserving prospective government information

To ensure permanent public access to newly produced government information in the future, **GPO should coordinate the preservation of born-digital government information.** These issues, already complicated enough for the FDLP in a print environment, are far more complex in the digital environment.

Several models exist for GPO's role:

1. Pursuant to existing law and regulation, GPO should continue to work with federal agencies to have their digital publications accessible to the FDLP through deposit in FDsys and should investigate possible additional incentives to encourage participation.
2. Sometimes as is current practice, GPO should instead provide coordination via formal partnership agreements with certain federal agencies that have a public dissemination mandate, rather than seek direct deposit into the FDLP's FDsys platform. In these cases, GPO should continue to actively engage with its partners to support and audit their work to ensure long-term preservation and persistent access.

3. In the absence of one of the first two options, which are the preferred alternatives, GPO will need to develop internal capacity, or external partnerships such as the Cyber Cemetery,⁶ to harvest content directly from agency websites for incorporation into the FDLP, typically via ingest to FDsys.

Documents historically published in static formats have become dynamic and interactive materials and databases. These require a more complex capture/preservation approach that in some cases is not yet well established. Partnerships between GPO and agencies are an important means to address this evolving development. There is the need for additional coordination between GPO, the Library of Congress, and the National Archives and Records Administration on this set of issues.

With appropriate born-digital production, the production of print titles and the growth rate of print collections will continue to decline significantly. Still, some continuing user needs will remain for print, especially for some non-text formats and in communities not well served by broadband, and production, dissemination, preservation, and access, should continue in these areas, in some cases via print on demand.

Digital infrastructure: preservation and discovery

Both for the historical collection once digitized and prospective government information, responsibility and expenditures for providing access to government information will shift substantially from the libraries to GPO. To effectively support these roles, the **FDLP must ensure the integrity and preservation of born-digital and digitized collections, using FDsys as a key platform and aggregator**. Where digitized or born-digital materials are not held in FDsys due to partnership agreements with other federal agencies or archives such as HathiTrust, providing for discovery will be a critical consideration.

To enable a broader range of innovative uses of these materials, GPO should emphasize the development of APIs and bulk download methods to access materials in FDsys. Its partnership agreements with other agencies and archives should encourage if not provide for such bulk access.

While GPO should focus on managing these materials according to state-of-the-art preservation practices, the sensitivity and importance of these materials, along with preservation best practices, indicates the need for outside auditing of its work. In addition, the FDLP community or other entities should host several digital copies outside GPO, to provide a managed preservation backup as well as to serve as integrity check on management practices. GPO's partnership agreements with other agencies and archives should mandate similar preservation practices.

Users increasingly expect to find information on their own, so seamless online discovery is imperative. **Libraries, non-profits, and vendors should develop new and revamped discovery environments based**

⁶ The Cyber Cemetery is a joint project of the University of North Texas Libraries and GPO that captures and provides "permanent public access to the Web sites and publications of defunct U.S. government agencies and commissions." More information is available at <http://govinfo.library.unt.edu>.

on the anticipated needs of specific groups of users. Some of these may be crafted as added-value businesses targeted at specific market segments (such as lawyers), while others may be created by librarians or others who are seeking to address a local need. FDsys should not be expected to serve as a single discovery resource for all government information but rather should, as GPO understands, facilitate various access channels and thereby contribute to a broader discovery ecosystem.

Outreach and use: rethinking the role of the librarian

Just as books and journals are shifting from being physical objects to digital files, libraries are evolving from being institutions optimized for storing those objects to enterprises focused on providing information services. In this transition, numerous librarians have been asked to rethink their roles in recent years, from catalogers to science librarians, and the same is taking place for government documents librarians. As print collection management declines as a priority, government documents librarians heretofore occupied with such work can direct an increasing share of their energy to the public services that will only increase in importance.

In this transition, **librarians should take on an expanded role as government *information* librarians rather than government *documents* librarians**, focusing on supporting users working with government information and data no matter its source or format. Some at the vanguard have already successfully redefined their role in this way. By repurposing existing expertise to solve a new set of problems, government information librarians will better meet the needs of users.

Some appropriate roles may be unique to the digital environment. For example, as mentioned above, librarians should be partners in the creation of value-added discovery tools for government information, created by libraries, library organizations, or non-profit or private sector entities. In addition, librarians should support users seeking to make innovative use of digital government data, such as by helping them with combining information from multiple sources or developing computationally intensive research methodologies. Other roles less closely tied to the digital environment, such as supporting more complex research questions, will also prove to be important.

An increasingly significant role for government information librarians should be focused on raising awareness of government information. These librarians should emphasize training other librarians within their library and in their broader local community, so government information may be better integrated into general reference support across the library community. As more and more public libraries position themselves to serve as government information libraries, there is a critical role for the FDLP's library participants to provide training, assistance, and deeper expertise. For outreach elements, a system that provides enhanced regional coordination opportunities for state and public libraries may be especially appropriate.

For many participating libraries, the services provided by their government information librarians in helping users and other librarians work effectively with this material will be their principal contribution to permanent public access.

In sum

The FDLP's mission of providing permanent public access to government information remains imperative for the digital environment. While the provision of permanent public access is today under threat, the model for the future presented in this report provides a feasible framework for meeting the challenge posed by the digital era. No stakeholder will find the approach outlined here to be perfect; tradeoffs and compromises are inevitable in attempting to transform a system of this complexity. But by taking a systemwide approach that recognizes the interdependencies among all stakeholders, the Program's problems can be resolved, leading to a solid foundation for the future. If no coordinated action is taken, the Program will continue an inevitable slide to irrelevance. The FDLP supports one of the cornerstones of American democracy. Its disappearance would be a significant loss.

10/15/2009

Ithaka S+R (www.ithaka.org/ithaka-s-r) is the strategy and research arm of ITHAKA, a not-for-profit organization dedicated to helping the academic community use digital technologies to preserve the scholarly record and to advance research and teaching in sustainable ways. The Ithaka S+R team supports innovation in higher education by working with initiatives and organizations to develop sustainable business models and by conducting research and analysis on the impact of digital media on the academic community as a whole.

Insights from these efforts are shared broadly, with more than a dozen reports freely available online. JSTOR, an accessible archive of more than 1,000 scholarly journals and other content, and Portico, a service that preserves content published in electronic form for future generations, are also part of ITHAKA.

