

7 April 2010

**Open government publications: The distributed, tamper-evident, collaborative digital FDLP Ecosystem.**

James R. Jacobs, Government Documents Librarian, Stanford University Libraries and Academic Information Resources

Since 1813, the Federal Depository Library Program (FDLP) has provided for public access to US documents, data and statistics. Historically, the Government Printing Office (GPO) printed documents as received from agencies and distributed paper copies to approximately 1250 libraries. In this system, libraries provided for public access to and long-term preservation of government information. This system is not perfect<sup>1</sup> but assures a robust distributed information environment that forwards our democratic process.

A key feature of this long-standing system is that it is tamper evident. The documents are US property, but if the government wants to change or recall a document it has to notify or take action at many of those 1250 libraries, ensuring community awareness of the event. Thus, while the government retains the ability to revise its official record, such revisions are public knowledge. *Such tamper evidence is imperative for government transparency.*

In 2010, 95% of US publications are born digital, and this move to digital documents is changing the way documents are produced, distributed and managed. Many of these changes are positive; however, care must be taken to ensure that the tamper evidence and associated transparency inherent in the historic system is preserved as new systems are developed. The GPO is in the process of building a centralized content management system – the Federal Digital System (FDsys) – that lacks the distributed preservation, and thus the tamper evidence, inherent within and critical to the FDLP. More and more agencies are bypassing the legal requirements of Title 44 and OMB Circular A-130 and publishing their publications on their own Web sites with little regard for Web standards, public access or long-term preservation. A vibrant and growing open government community advocating for and making great strides toward open government data and open law masks the loss of a secure infrastructure for long-term preservation of and access to government information. The current evolving digital system is fragile with a single point of failure;<sup>2</sup> NARA and GPO cannot hope to deal with these issues on their own.<sup>3</sup>

---

<sup>1</sup> For more see “Less Access to Less Information by and about the US Government” <http://freegovinfo.info/library/lessaccess>.

<sup>2</sup> This is not a theoretical occurrence. In August 2009, GPO’s PURL server – its primary link resolver and critical piece of its infrastructure – crashed and links to government publications were inaccessible for several weeks. For more see <http://freegovinfo.info/node/2704>.

<sup>3</sup> For more, see “Government Information in the Digital Age: The Once and Future Federal Depository Library Program.” *Journal of Academic Librarianship* 31(3) May 2005. 198-208. [http://freegovinfo.info/fdlp\\_digital](http://freegovinfo.info/fdlp_digital).

A growing number of FDLP librarians are talking about the need for a digital FDLP Ecosystem – *publications.gov*. This distributed system for the storage of digital content would incorporate collaborative cataloging/metadata creation, as well as shared or Peer-to-Peer (P2P) technical infrastructure in which data and technological redundancy and collective and proactive action reign. There are three pieces to this ecosystem (content creation, content distribution, content collection), as shown in the following diagram. The ecosystem requires participation from all government agencies – specifically GPO and NARA – FDLP libraries and the public. If this is achieved, this ecosystem will provide widespread access – including to those on the other side of the digital divide – long-term preservation, redundancy, information assurance and a robust mash-up environment to facilitate ongoing open government initiatives.

Long-term preservation is the missing piece in the current discussion about government transparency. Libraries are in good position to provide this piece, as they are already building systems to that end. Stanford University Libraries is heavily involved in LOCKSS and specifically the LOCKSS-USDOCS project.<sup>4</sup> Stanford is also at the forefront of harvesting digital government information.<sup>5</sup> However, much more needs to be done and many more libraries and government agencies need to be involved. Indeed, a single provider or a single system, even if used in multiple locations, cannot provide the necessary redundancy of copies or multiplicity of digital preservation systems. Libraries are excited to continue their traditional government transparency role and eager to participate in new initiatives in access to and preservation of government information. Digital preservation of open government information (when achieved) is vital so that future citizens can learn from our mistakes and triumphs.

### **What's already in place:**

- 1) FDLP Libraries interested in participating in the digital FDLP ecosystem<sup>6</sup>
- 2) LOCKSS-USDOCS project
- 2) GPO Federal Digital System (FDsys) in beta including authentication with digital signatures<sup>7</sup>
- 3) Agencies publishing online
- 4) Growing Open Government community within and without government

### **Government requirements:**

1) Open Web, content, and metadata standards, well-structured web sites, and robust bibliographic metadata for government publications and data to make agency sites more easily crawlable / preservable / usable. Every agency should have *agency.gov/publications* and *agency.gov/data* directories so that GPO could harvest content for FDsys and others would also have bulk access to the agency's content. **All**

---

<sup>4</sup> 16 libraries are currently participating in LOCKSS-USDOCS  
[http://lockss.org/lockss/Government\\_Documents\\_PLN](http://lockss.org/lockss/Government_Documents_PLN).

<sup>5</sup> Examples include Archive-it (<http://www.archive-it.org/home/SSRG>) and California Digital Library's Web Archiving Service (see Web at Risk NDIIPP project).

<sup>6</sup> In the 2007 FDLP biennial survey, 37.78% said they "want to receive PDFs if GPO deposited them." <http://snipurl.com/2007-fdlp-biennial-survey>.

<sup>7</sup> See GPO Authentication site <http://www.gpoaccess.gov/authentication/>.

federal agencies need to participate fully with GPO in building the national bibliographic record through the Catalog of Government Publications (CGP) (<http://catalog.gpo.gov>), FDsys and the FDLP ecosystem. (Note: The implementation of standards greatly reduces operating costs, since most cost in digital archives occurs in ingest. Digital signatures of content spread throughout the whole government also reduce the cost of maintaining fixity and authenticity.)

2) System of permanent urls (purlz [<http://www.purlz.org/>], DOIs [<http://www.doi.org/>] or similar)

3) *publications.gov* interoperability with *data.gov* and *law.gov* (proposed)

4) Distributed preservation technologies outside of .gov domain. NARA and GPO to participate in and facilitate, but not control, these external systems

5) Encourage other governments to adopt open standards for ingest and authenticity in order to facilitate cost-effective reciprocal government document collections similar to that done between Library of Congress and German National Library (DNB)

### **FDLP library requirements:**

1) Government funding and grants for long term preservation projects

2) Incentives for more libraries to participate in digitization and distributed preservation projects (i.e., grants for digitization and digital infrastructure through IMLS and/or NDIIPP)

3) Several designated “paper” regional libraries that archive all new FDLP content on paper or other non-digital formats

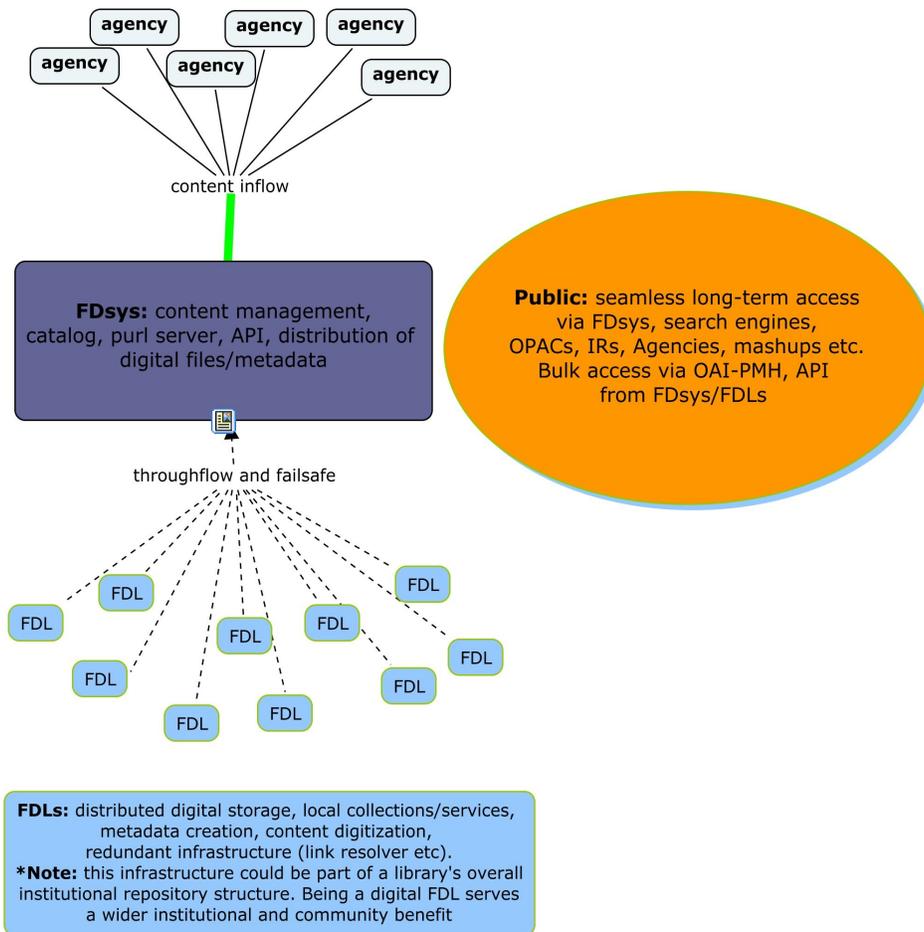
4) Continued and enhanced training for FDLP librarians in digital library technologies

Respectfully submitted by

James Jacobs  
Government Documents Librarian  
Stanford University Libraries and Academic Information Resources

Endorsed by

Michael A. Keller  
University Librarian and Director of Academic Information Resources  
Stanford University



The FDLP Ecosystem: <http://freegovinfo.info/node/2773>