

## Managing Digital Content in the FDLP: *Frequently Asked Questions*

What is the best way to manage digital government information in the interest of public access? We think that the model that solves the most problems is to deposit digital content with depository libraries. This FAQ is an attempt to answer some of the questions and concerns that we have heard about this idea. We invite your comments, questions, critiques, and discussion on this important issue.

1. *Since my patrons can access government documents on the Internet, why do I need a copy of those documents?*
  - Ensuring access implies being able to control the existence, integrity, and location of an item. If someone other than you can move, replace, alter, or remove the copy you want to provide for your users, then you can not ensure access. The presence today of a document on the web is no guarantee of its presence tomorrow.
  - It is true that *users* of information do not care where the material resides *as long as they can get access to it when they need it*.
  - It is also true that *libraries* can only ensure access to digital objects by retaining control over those objects. “Retaining control” could mean having a local copy. It could also mean being part of a JSTOR-like consortium model.<sup>1</sup>
2. *But doesn't GPO guarantee “permanent public access”?*
  - This is GPO's goal, but there is a serious potential problem with it: GPO is not funded to do this. The ability of GPO to provide access to everything always is only as assured as the next funding cycle and whim of Congress. If Congress can, as it has, effectively remove vital materials from the depository system and allow GPO to change a *depository* system to system *without deposits*, how certain can we be that Congress will not change GPO's mandate, or abolish GPO, or privatize all or part of the “permanent public access” collection, or fund GPO at levels that make it impossible for GPO to keep everything online always? Given these circumstances, libraries need their own copies of digital documents in order to speak with any confidence of permanent public access.
3. *Can't we rely on the government (individual agencies, NARA, etc.) to preserve for the long term? My library isn't an archive after all!*
  - Libraries are not archives and yet libraries do play a role in preserving materials for access. This will continue to be an important role for depository libraries, just as it will for other libraries.
  - NARA will provide the archival function, but it is unlikely that it will provide the library or “access” function to an extent that obviates the need for libraries.
  - It is unlikely that agencies will preserve their materials in either the library sense *or* the archival sense. Agencies have balked when told to do so by NARA and have rarely shown a understanding of the need for older materials. Proposals that address the need for agencies to better manage the “complete life cycle” of their records are important and necessary, but these are record management and archival functions. It seems both unnecessary and inappropriate to combine those functions with the very different library service functions that independent libraries can provide better than agencies.

4. *Do we need more than one electronic collection?*

- Yes. A one-size-fits-all collection can not adequately meet the information needs of all communities of interest. Different communities of interest need their own collections of materials selected, organized, presented, serviced, and preserved to meet their needs. This should be self-evident to us. Our libraries are not identical because we serve different communities. A document essential in my library may be of no interest to your community. Lawyers require different collections, organized in a different way, than K-12 libraries, environmental libraries, or agricultural libraries. Even if we choose to rely on an electronic collection at GPO for the short term, it cannot possibly replicate the different collections, organization, and presentations of 1400 depository libraries.
- Any single-collection model of long-term preservation makes any loss of those materials a catastrophic and irreplaceable loss whether the loss is intentional or unintentional; budgetary or technological; malicious or well-meant; political or benign. Having multiple copies in multiple locations makes it easier to ensure long-term preservation and access. This is the LOCKSS model.<sup>2</sup>

5. *What is wrong with GPO setting up an Electronic Collection?*

- There is nothing wrong with GPO maintaining an Electronic Collection as long as it does not sap resources from other essential GPO responsibilities.
- Title 44, Sec 1904 of the *US Code* requires the Superintendent of Documents to “...issue a classified list of government publications...” This is an essential service that is legally mandated to GPO and no other government agency. If anyone can do this, it should be GPO. It is a complex and daunting job, particularly in the digital age, but everything else stems from this being done accurately and in a timely manner. This should be GPO’s first and primary responsibility for, if this is not done well and completely, no electronic collection at GPO or anywhere else can be accurate, complete, or up-to-date.

6. *Does the “FDLP partnerships” program solve our problems? Isn't it necessary to have new kinds of partners and new commitments in the digital age?*

- New programs are a way to cope with rapid change and experiment with solutions. But the FDLP Partnerships program as currently implemented does not solve our problems and has many flaws:
  - The new partnership model creates single-copy collections (see #4).
  - If a partner needs to withdraw from the partnership agreement, the resulting burden on GPO could be overwhelming because GPO seeks partners when it does not have the resources itself to manage a body of materials.
  - A “partnership program” already exists and is mandated by Title 44. Setting up a new program to replace the partnerships mandated by Title 44 is unnecessary and creates new risks for long-term access and preservation.
  - When a new partner is a commercial enterprise, the danger of privatization of public information becomes a serious concern.
- If the current depository requirements need to be modified, they can be, just as they were when microforms, CD-ROMs, and Internet access became essential components of being a depository.

7. *Isn't the GPO setting up mirror sites to deal with these problems?*
- Mirror sites provide solutions to two problems: unintentional loss or damage to files at one site and delays and unavailability of materials caused by network bottlenecks, server down time, and so forth.
  - Mirror sites do not solve two other problems: the problem of government control of information (see #1), and the problem of a single collection (see #4).
8. *Are there any alternative ways to ensure access to the current edition of something without having to constantly look for the current edition to download?*
- Yes. A better solution would be for GPO to concentrate on identifying government information, classifying it (e.g., adding item numbers), cataloging it and notifying depository libraries.
  - A better solution than libraries having to actively "download" files would be for GPO to automate the process of delivering those files (which a library selects by item number) over the Internet to libraries, or provide a way that libraries could automate downloading their item-selections daily. Freely available, commonly-used software exists to do this.
9. *Won't I save time in processing and shelving and cataloging by relying on digital publications on the Internet?*
- It is unlikely. Experience shows that, if you attempt to provide access through your web site or OPAC, you will have to allocate significant time and labor to tracking down changed URLs, missing PURLs, documents that have changed or been removed, and so forth. While you can predict the amount of time you will spend controlling your own collection, you cannot predict the amount of time you will lose keeping track of a "collection" that someone else controls.
  - For additional details on percent of bad URLs, bad PURLs, and the difficulty of trying to maintain accurate links, see the detailed notes provided by the "GPO Marc Internet Resources" web site (<http://www.du.edu/bdld/mirintro.htm>).
10. *Isn't saving copies of digital publications too expensive and too technical for any but big university research libraries and organizations like GPO and OCLC?*
- No. There are many different ways of obtaining and storing digital materials. There is a range of technical options enabling even the smallest library to have a small digital collection built and managed with very simple tools, minimal skills, and inexpensive hardware. Each library can choose methods to fit its resources. Starting a small collection does not preclude growing into a larger collection, if desired.
  - Most depository libraries are already managing collections of digital publications on CD-ROMs. Extending this to, for instance, PDF files retrieved over the Internet is another step in the same direction, not a change of direction.
11. *What do I do if my library director won't let me have a digital collection or won't fund anything new?*
- Work with your library management. Be sure that they understand the tradeoffs inherent in their decisions. If you cannot guarantee access to material, say so. If you require a larger budget to acquire hardcopy, say so. If you require new equipment or staff because of format changes, explain that this is what you must do in order to ensure access to government information.
  - Emphasize the importance of the role of libraries in society. If libraries do not take responsibility for selecting, organizing and preserving digital information one of two

things will certainly happen. Either information will be lost because no one will take on this role, or the private sector will do it for those items that are profitable.

- Work with your community. Documents are for the people. Look for projects that the community wants and use public interest to get resources for government information.
- Look for new communities of interest beyond your local geographical area. Is your library strong in one subject area? Work with other libraries to establish a digital consortium or other kinds of agreements that will spread the burden while ensuring long-term access. Seek new sources of funding for new models.
- Work with GODORT and DLC to establish reasonable standards for digital library depository collections and then work with library management to meet those standards.

12. *If I have a damaged or inaccurate or out-of-date digital copy of the Federal Register or some other official publication, won't someone sue me for giving them the unofficial information and passing it off as official?*

- Courts set rules for what they accept as the official legal version of laws and regulations and how they must be cited. We can, as libraries have always done, provide what is made available to us and ensure that the user knows exactly what it is (source, version, edition, etc.)<sup>3</sup>

13. *How can I be sure that I have an authentic and up-to-date copy of a digital publication?*

- By recording when and where you obtained your copy and including this information in the cataloging record (or other metadata system) that accompanies each item, you provide assurance of what the item is, where it came from, and when it was obtained.

14. *Isn't there more to authenticity and integrity than cataloging? What about PKI?*

- Yes. We could do a much better and more efficient job if GPO would do two things:
  - If GPO deposited digital items with depositories, you would have the imprimatur of the depository system as an assurance of provenance. The currency of an item would be assured through GPO's active deposit of the material.
  - If GPO provided a digital hash<sup>4</sup> with each item, you could also ensure that an item had not changed since you obtained it.
- PKI (Public Key Infrastructure) may be important in the future for e-commerce and e-government and may, therefore, play a role in authenticating government information. It is complex, difficult to implement, and replaces one level of technical authentication problems with another.<sup>5</sup>

15. *What are the responsibilities of GPO and depository libraries?*

- GPO and depository libraries have a long history of complementing each other's roles and responsibilities. Although some processes change in the digital information environment, the essential responsibilities do not. GPO still needs to identify, describe, and distribute materials and depository libraries still need to select, acquire, organize, preserve, and provide service for those materials deposited by GPO.
- If GPO takes on the responsibility of providing access to government information, it is usurping the role of depository libraries and assuming a level of ultimate control over that information that no government agency should have.
- If libraries volunteer to fill in the gap left when GPO is unwilling to allocate its limited resources to accurately and fully identify and describe government

information in a timely manner, then they are allowing and facilitating GPO's neglect of its primary responsibility.

- If libraries are content to allow GPO to take over the responsibilities of selecting, organizing, presenting, and preserve materials, they will be abdicating their own responsibility.

We invite your comments, questions, critiques, and discussion on this important issue.

Elizabeth Cowell      ecowell@ucsd.edu  
Jim Jacobs              jajacobs@ucsd.edu  
Karrie Peterson        kapeterson@ucsd.edu

La Jolla October 2001.

### ***Endnotes***

---

<sup>1</sup> For a good discussion of JSTOR and these issues, see: "Editors' Interview: Developing a Digital Preservation Strategy for JSTOR, an interview with Kevin Guthrie" *RLG DigiNews* Vol. 4 No. 4 (Aug. 15, 2000). <http://www.rlg.org/preserv/diginews/diginews4-4.html#feature1>

<sup>2</sup> "LOCKSS: A Permanent Web Publishing and Access System" Vicky Reich and David S. H. Rosenthal. *D-Lib Magazine* June 2001 Vol. 7 No. 6. <http://www.dlib.org/dlib/june01/reich/06reich.html>

<sup>3</sup> For more information on this topic, see Claire M. Germain, "Digital Legal Information: Ensuring Access to the 'Official' Word of the Law" *Cornell Law Forum*, Vol. 26 No. 1 (July 1999), 11-14. [http://www.lawschool.cornell.edu/lawlibrary/Faculty\\_Services/facbib/Germain\\_DIGIFORU99.html](http://www.lawschool.cornell.edu/lawlibrary/Faculty_Services/facbib/Germain_DIGIFORU99.html)

<sup>4</sup> A hash is simply a mathematical "digest" which is generated from a digital document using a mathematical rule. It is designed so that a small change in the document would produce a big change in the hash. Hashes are used to check the integrity of files and documents. A small change in a document would produce a completely different hash. See also, Clifford Lynch, "Authenticity and Integrity in the Digital Environment: An Exploratory Analysis of the Central Role of Trust" (<http://www.clir.org/pubs/reports/pub92/lynch.html>)in, Council on Library and Information Resources Authenticity, *Authenticity in a Digital Environment*. Washington, D.C. May 2000. CLIR pub 92.

<sup>5</sup> For a good overview of some of the problems with PKI, see: "Enterprise Security: Pandora's Box or Panacea?" by Roberta Bragg, *Enterprise Systems Journal* (9/01/01), <http://www.esj.com/article.asp?ID=0904200114228PM>