

Guest Opinions

James R. Jacobs

* The opinions stated below are purely my own and do not necessarily reflect the opinions of the UC San Diego libraries.

The Future of Government Information

(This document will be permanently housed at http://www.ucop.edu/lauc/opinions/gov_info.html)

Since 1860, there has been a system in place to insure public access to federal government information through a partnership between the Government Printing Office (GPO) and the hundreds of libraries in the Federal Depository Library Program (FDLP). Similar systems have been set up in nearly every state for state-produced information. For example, the California depository program, administered by the California State Library, was begun in 1945. The UC Libraries have long been active participants in both the federal and state depository programs and have assisted in providing for one of the core tenets of a democracy, an informed citizenry. I will primarily be talking about federal government information in this article, but many of the same issues involved translate to all levels of government information.

Current conditions:

In the last 15 years, as digital and internet technologies have developed, two challenges have arisen to the traditional system of government information dissemination whereby the GPO printed documents for federal agencies, sent them to FDLP libraries and libraries then made collection decisions, preserved and gave access to those documents.

Because more and more government information is accessible via the internet, GPO is no longer printing very many documents. In fact, by 2007, according to Superintendent of Documents Judy Russell, 95% of all federal information will be digital only.¹ The acronym “GPO” is now largely a misnomer.

Another byproduct of the move to digital is that more and more federal agencies are producing their own documents and publishing them on their websites – thereby bypassing Title 44 of the US code which requires them to make their documents available to GPO so that GPO can describe and distribute them to depository libraries.²

These 'fugitive' government publications are not entered into the national bibliographic record nor distributed to FDLP libraries. Recent reports have stated that the number of fugitive documents may be more than 50% of all electronic documents. Some agencies – for example the National Institutes of Health – have found that almost 80% of their documents are not provided to GPO for dissemination.³

In short, more and more information is being “born-digital”, and less and less of it is being moved through GPO to depository libraries. Many would point to the successful online distribution of the *9/11 Commission Report*⁴ to show that digital government information is more widely available than ever to citizens and that the FDLP is no

longer necessary.

However, the easy access of the *9/11 Report* masks the fact that preservation and authenticity of information are becoming increasingly difficult to attain.⁵ Due to this, I believe that the FDLP, far from being an anachronism, is even more important in this digital era and that the UC libraries should continue to be active partners in this vital endeavor.

GPO's "Future Digital System"

GPO has begun to respond to the growing problems of access, preservation and authenticity in the digital age – and that's a good thing! In late 2004, GPO proposed a new content management system that they call the "future digital system." This system, GPO maintains, will enable GPO to collect, describe, preserve and make accessible all past, present and future government information.

To find out more, the Government Information Technology Committee of ALA's Government Documents Roundtable (GODORT) – which I chair – invited Superintendent of Documents Judy Russell and George Barnum from GPO's Office of Innovation and New Technology to the midwinter 2005 GITCO meeting to talk about the future digital system as outlined in the planning document, *Concept of Operations for the Future Digital System (ConOps)*.⁶

ConOps outlines the creation of a centralized content management system based on digital standards, faithful application of the Open Archival Information System (OAIS) – a "comprehensive logical model describing all the functions of a digital repository."⁷ The future digital system will be modular, preferably open-source (in output format if not in software). It will automatically generate multiple metadata schema (descriptive, technical, structural, and administrative) including METS (Metadata Encoding and Transmission Standard), the current metadata darling of the digital library set. There will be automatic ingest and harvesting capabilities, content creation, management, and validation, the ability to convert digital formats, a new and better search mechanism than WAIS – the current, rather primitive search engine being used for GPOAccess – and archive and preserve digital documents into the future.

However, those concerned with *policy* rather than charmed by technology remain wary and for several reasons. There's no doubt that GPO has done its digital homework and that the future digital system is a valuable *piece* of the future system of providing for digital government information dissemination. However, in my opinion (and that of more and more librarians who delve deeper into the system), it is not the whole enchilada and should not be seen as such. In fact, a centralized database of all government information, far from being optimal, actually has the possibility of endangering ACCESS and PRESERVATION of government information as well as the PRIVACY of those who read government information. That is, unless there are other access and preservation options available, multiple digital repositories able to ingest digital documents and GPO-created metadata and a dedicated group working in collaboration to assure these core needs.

I will talk briefly about each of these issues – access, preservation and privacy.

Access may be endangered because GPO's future digital system will replace distributed access in multiple collections with centralized, controlled access. Any changes in government policy, fee-structures, access rights, or funding support for access at the federal level will affect access for all. This system will make it easy for politicians to remove, hide, and alter embarrassing information.⁸ A centralized database also will allow GPO to charge for access to the public information in that database because they will be the primary gatekeepers of all information. It will create an expensive system requiring large and continuous amounts of federal funding – putting the budget for access to government information in competition with defense, education, homeland security, etc.

Preservation may also be endangered by a central database. Digital preservation is an expensive and never-ending task requiring continued, perpetual funding. Digital preservation may perhaps be even more expensive than for books. I believe it is the epitome of foolishness nearing technological hubris to rely on one federal agency to do it all. Relying on a single system for preservation of all government information puts that information in danger of

damage or loss because of federal funding shortfalls.

Furthermore, digital information is not stable. A book may last hundreds of years in the right conditions, but bits, software, and storage media are quite volatile. Digital preservation is in its infancy and nobody has created a perfect system to preserve digital information. As Patricia Cruse from the California Digital Library so rightly stated, "Since digital preservation is so new, it is very important to have a diversity of preservation approaches including different technical, funding and policy approaches."⁹ I believe it would be far better to rely for digital preservation on many libraries running multiple systems with many budgets rather than one chronically cash-strapped federal agency.

Privacy: The ALA Library Bill of Rights inherently protects a patron's right to read. A central database will endanger that right by allowing the government to track who is reading what. Server logs are kept automatically and can easily connect IP addresses to accessed content. Public Key Infrastructure (PKI) and Digital Object Identifiers (DOI), while offering good technical tools for digital authentication, may also be used to invade a patron's privacy.¹⁰ The government will not need the USAPATRIOT Act to search library records, they will have the information already.

Conclusion:

GPO's proposal will replace in all but name the current system whereby government information is distributed to 1250 FDLP libraries – a system that has successfully organized, preserved and provided free access to government information across the US for over 150 years.

GPO states that the rapid expansion of digital publishing and the World Wide Web necessitate the change in the FDLP toward a central database. However, the FDLP has worked for so long because it requires the federal government to pay only for the production and initial distribution of information, while the cost of preservation and access is borne by those depository libraries throughout the US. GPO's one-size-fits-all proposal will put the entire burden of access and preservation on the federal government and will not provide for an FDLP safety net of distributed digital collections.

Libraries, and more specifically UC libraries, will continue to have a very important role to play in the future of government information dissemination. I believe that what is needed at this historic juncture is a diversity of tactics. The UC libraries have begun to think about ways that we can work together in order to assure that our users have fully-functional access to all government information, that the information is preserved for future use, and that our patrons' privacy is protected.

For example, CDL's Web At Risk project, funded by a grant from the Library of Congress as part of the National Digital Information and Infrastructure Preservation Program (NDIIPP),¹¹ will explore the creation of tools that will allow libraries to ingest, describe, give access to, and preserve government information. The Shared Cataloging Program (SCP) is already working well for California State documents and can be a key component in any future digital government information program. Stanford University Library's Lots-of-Copies-Keep-Stuff-Safe (LOCKSS) project also holds promise.¹² Through a peer-to-peer architecture, the LOCKSS software functions as a persistent digital preservation system. These projects point toward the creation of a UC-wide digital repository of government information.

I believe that it is more important than ever that all of the main stakeholders (GPO, Congress, Libraries, NARA, citizens) – having different roles, constituencies, and mandates – work collaboratively to assure the future of government information. GPO should continue to enforce Title 44 and facilitate the FDLP by disseminating digital documents and their metadata to libraries. A system whereby GPO's future digital system works in conjunction with the above and future projects will be a distributed and stable system of digital information access and preservation. Not only would a UC-wide digital repository give our users quick and easy access to digital government information, it would allow us to do things never before possible like creating digital collections on-the-fly so that researchers and students could more thoroughly analyze disparate government information.¹³

Government information is in many ways, the canary in the coalmine for the library community. The issues that the government information community is attempting to resolve today will be those that the entire library community will face in the near future. Digital information requires that libraries have digital tools to continue to do what they have done successfully for so long – select, acquire, describe, preserve and provide access to and service for the world of information.

1. Remarks by Superintendent of Documents Judy Russell. Fall, 2003 Depository Library Conference, Oct. 20, 2003 . http://www.access.gpo.gov/su_docs/fdlp/pubs/proceedings/03prot.html (Accessed April 29, 2005)
2. Title 44 U.S.Code. 2000 edition. <http://www.gpoaccess.gov/uscode/index.html> (Accessed April 28, 2005).
3. See U.S. Government Printing Office (GPO). “First Fugitive Documents Week Nets 222 New Titles,” Administrative Notes: Newsletter of the Federal Depository Library Program, 25(5) (April 15, 2004), 1-3, http://www.access.gpo.gov/su_docs/fdlp/pubs/adnotes/ad041504.pdf (Accessed April 29, 2005).
4. The *9/11 Report* was a finalist for the 2004 National Book Award in non-fiction. GPO reported over 1.2 million downloads of the 9/11 report in the first few weeks. See U.S. Government Printing Office (GPO), “Remarks of Superintendent of Documents Judith C. Russell, Depository Library Conference, Washington , D.C. October 18, 2004 ,” Administrative Notes: Newsletter of the Federal Depository Library Program 25(11) (October 15, 2004), 11, http://www.access.gpo.gov/su_docs/fdlp/pubs/adnotes/ad101504.pdf (Accessed April 29, 2005).
5. Just last week, a librarian posted to govdoc-l, the government documents listserv, wondering if anyone had cached an “e-only” document from the Minority Business Development Agency because it had been taken down from the MBDA website. Despite GPO’s purls, the internet archive, and google cache, the document was no longer available. <http://lists1.cac.psu.edu/cgi-bin/wa?A2=ind0504d&L=govdoc-l&T=0&F=&S=&P=6541> (accessed May 1, 2005).
6. “United States Government Printing Office (GPO) Concept of Operations (ConOps V1.0) for the Future Digital System, Final” October 1, 2004, http://www.gpo.gov/news/2004/ConOps_1004.pdf (Accessed April 28, 2005)
7. For more on OAIS, see RLG’s website: http://www.rlg.org/en/page.php?Page_ID=3201 (Accessed April 29, 2005)
8. You may think this latter possibility is pure paranoia, but one need only look to the Congressional Record to realize that this is not so far out of the ordinary already. The Congressional Record, supposedly the verbatim account of Congressional proceedings, has for years been legally altered under the guise of correcting transcription errors. The most famous case of its alteration is the speech given by the honorable Hale Boggs of Louisiana before the House of Representatives on October 18, 1972 , fully 2 days AFTER he died in a plane crash!! See *Tapping the Government Grapevine*. Judith Schiek Robinson, 1998. p.96.
9. Patricia Cruse quoted in Aliya Sternstein, "Collection of last resort," Federal Computer Week (Aug. 30, 2004), <http://www.fcw.com/fcw/articles/2004/0830/pol-collect-08-30-04.asp> (Accessed April 29, 2005).
10. Authentication tools such as the Digital Object Identifier (DOI) are designed, not just to verify the authenticity of a document, but also to check “the customer’s authority to access it,” and protect copyright. For more details on DOI, see David Sidman, “The Digital Object Identifier (DOI): The Keystone for Digital Rights Management (DRM)” Draft submitted to the SIIA Digital Rights Management (DRM) Working Group, January 26, 2001 . <http://www.contentdirections.com/materials/SIIA-DOIandDRM-DavidSidman.htm> (Accessed April 29, 2005).
11. See CDL NDIIPP press release, http://www.cdlib.org/news/press_releases/

[award_announcement_final_20040930.doc](#). (Accessed April 29, 2005).

12. <http://lockss.stanford.edu/index.html> (Accessed April 29, 2005).

13. This takes Don Swanson's concept of "undiscovered public knowledge" into the digital age. See Swanson, Don. "Undiscovered public knowledge." *The Library Quarterly* v. 56 (April 1986) p. 103-18.

Send comments regarding this site to [Frank Lester](#), LAUC Web Manager. Last updated: 03/12/2013