Practical Academic Librarianship: The International Journal of the SLA Academic Division
8(2):1-12, 2018 © The Author(s) http://journals.tdl.org/pal

Whither Goest Government Documents? A Story and a Study

Jeffrey Mark Wilhite
University of Oklahoma

Karen Rupp-Serano
University of Oklahoma

Abstract

A study was conducted to determine the findability of known print U.S. government documents using five different resources: the Catalog of Government Publications (CGP), in-house library catalog, library discovery service, HathiTrust, and Google. Overall, Google was the most effective finding tool, followed by the library discovery service, the in-house library catalog, Catalog of Government Publications, and HathiTrust. Three-quarters of documents were available full-text via Google. Implications of the online availability of government publications on the need for large print document collections and the continuing reconsideration of library spaces are discussed as are future studies.

Keywords: federal publications, government documents, Google, HathiTrust, discovery services, Catalog of Government Publications, library catalogs

Corresponding author: Jeffrey Mark Wilhite jwilhite@ou.edu

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/321495

Introduction: A Document Librarian's Story

In 2015, I attended the Federal Depository Library Program (FDLP)/Government Printing Office (GPO) conference in Washington, D.C. My main goal in attending was to better understand the changes that were occurring in the depository library system specific to academic libraries. Academic libraries constitute 72% of the 1,139 FDLP libraries (Outsell, 2016, p. 25). I wanted to understand the changes underway and responsibly guide my institution's collection through those changes.

There was ample reason to prepare. A hard look at the data, specifically the user data for the University of Oklahoma's (OU) government documents collection, showed circulation, reference queries, and reshelving statistics steadily decreasing. Only interlibrary loans and instruction requests were increasing. That the collection covered a prime piece of real estate located in a library looking to expand user spaces compounded the data. In this regard, the OU collection was in company with many other selective depository collections replacing space for collections with space for users as noted in the 2009 Ithaka report *Documents for a Digital Democracy* (Schonfeld & Housewright, 2009).

Thus, with a determination not to make a difficult decision over every item to be weeded, I headed to Washington, D.C. It was reassuring to hear of the process and patience of other librarians and to learn the realities facing all of us. One of the biggest takeaways was the change in *passion* among the attending document librarians. In the early 1990s, these meetings were accompanied by a *sturm und drang* chorus of senior librarians as each change to the *status quo* was announced. But not now. My colleagues were determined to take charge of the change rather than let the change take charge of them. This attitude was infectious, and it caused me to undertake a questioning of my own *status quo*. With a co-author, we undertook a study of the five government documents finding tools in common use: the Catalog of Government Publications (CGP), in-house library catalog, library discovery service, HathiTrust, and Google. The study would focus on findability of known print government documents within these resources. In this study, "findability" means that a record for the item could be located in one of the five resources. In addition, "print government documents" are defined by physical items found in academic libraries in North America or as in the case of HathiTrust, a digitized full-text version of a physical item held by a HathiTrust member library. 'Print' government documents include traditional paper items on the bookshelves of document collections. Maps, posters, oversize, and online items comprise about 2% of government publications and were not included in this study. Microforms were also not included, and this format dominates the non-

Corresponding author: Jeffrey Mark Wilhite jwilhite@ou.edu

print documents population. Out of the 2.9 million items in the OU Government Documents Collection, 1.2 million (41%) are microfiche.

Literature Review

A review of the literature related to the types of components explored in this article turns up a number of studies. For discovery tools, substantial research is available on the type of databases we had decided to explore in an academic library setting. The most recent N*WC Horizon Report: 2105 Library Edition* reported on discovery systems, outlined the current trends, and highlighted some of the literature (Johnson, Adams Becker, Estrada, & Freeman, 2015). Fagan, Mandernach, Nelson, Paulo, and Saunders (2012) ran a usability test for a discovery system in an academic library, while Kornblau, Strudwick, and Miller (2012) dissected questions librarians should be asking themselves about the discovery services. Singley (2014) tested known items in a discovery system, and Foster and MacDonald (2013) compared Summon and EBSCO, with varied results. Thomsett-Scott and Reese (2012) surveyed the literature on discovery systems in regard to academic libraries, and Chickering and Yang (2014) gave an update on the latest discovery tools. Aharony and Prebor (2015) looked at academic librarian attitudes towards discovery tools, while Nichols, Billey, Spitzform, Stokes, and Tran (2014) specifically explored the Primo discovery system.

Regarding internet search engines, relevant literature compares Google to the library catalog and/or discovery tools. Early on, Anderson (2005) sounded an alarm about the future of libraries and Google, while Brophy and Bawden (2005) comparing Google to other library resources wondered "Is Google Enough?" Namei and Young (2015) looked at a discovery system, Google, and Google Scholar. In a massive three-part study, Georgas (2013; 2014; 2015) looked at federated searching and Google, while Tay (2015) contemplated Google Scholar vs. discovery systems in general.

Some studies come close to the one undertaken here but without the specific focus on U.S. government publications. Asher, Duke, and Wilson (2013) looked at EBSCO Discovery, Summon, Google Scholar, and conventional library resources. Perhaps the closest match is a 2014 study by Singley, looking at known print resources in four discovery systems–EDS, Primo, Summon, and WorldCat Local. Singley (2014) found these systems continue to have problems with known item search. However, none of these studies addressed government documents specifically. Those focusing on government documents are Farrell's (2005) editorial

about Google and government documents–written when Google was new–and two 2011-2012 articles by Sare on government documents and HathiTrust/Google Books. Sare's 2011 article investigated finding government documents in HathiTrust, with 21% to 25% of her 1943-1976 sample appearing full-text in that database. Looking at overlap, functionality, and other features, Sare's follow-up in 2012 compared her sample against Google Books with varied results.

Methodolody

Five online resources were tested: our in-house library catalog, our discovery system, HathiTrust, the CGP, and Google. The first four form the core resources librarians use to locate records for print items in U.S. government documents collections:

- The OU library catalog is Ex Libris' Alma. It reflects our physical and virtual holdings and includes MARC records from MARCIVE for U.S. government publications, 1976-present. All items searched in the catalog were physically located in our library but not all were in the OU library catalog. OU has purchased the MARCIVE files since 1991 and thus only has records for government documents from 1976 to present, so records for pre-1976 items were often not located.
- The OU discovery system is Ex Libris' Primo, which searches journal article and library catalog metadata in a unified index and presents search results in a single interface, allowing users to conduct keyword searches over most of the library's resources. Items in ALMA are included in the discovery layer.
- The CGP is the finding tool for electronic and print publications of the U.S. government, which comprise the National Bibliography of U.S. Government Publications. The CGP contains descriptive records for historical and current publications and provides direct links to those available online. Although starting with records from 1976 to present, more historical records have since been added, with the aim of making this database the central point for locating current and historical government publications (Galileo Scholar, 2016).
- HathiTrust is an international community of research libraries committed to the long-term curation and availability of the cultural record (HathiTrust, "Partnership Community," 2016). Its digital library provides long-term preservation and access services for public domain and in-copyright content from a variety of sources, including Google, the Internet Archive, Microsoft, and partner institution initiatives (HathiTrust, "Our Digital Library,"

2016). For this study, we used the HathiTrust native interface.

- The fifth resource, Google.com, needs no explanation. It is the most-used search engine on the planet (Alexa, 2016) and, as such, is regularly used by non-librarians and librarians alike.

Most of the OU's government documents collection usage centers around interlibrary loan requests for print items. In-person transactions are not common anymore; thus we decided to focus strictly on print materials for this study. One hundred known print items received via our GPO depository profile were selected for searching. These items represented the publications of 20 different U.S. federal government entities: Agriculture, Commerce, Congress, Defense, Education, Energy, Executive Office of the President, Health & Human Services, Homeland Security, Housing and Urban Development, Interior, Justice, Labor, Library of Congress, NASA, Smithsonian, State, Transportation, Treasury, and Veteran's Affairs. Five items were selected from each entity: two published prior to 1976 and three published after 1976 for a total of 100 documents searched. The pre-1976 demarcation is frequently used by the U.S. Government Printing Office since OCLC does not generally have pre-1976 GPO cataloging records. Note that three government entities not in existence prior to 1976 were selected. This was done in an effort to represent changes to the government over time and to provide a wide range of disseminating agencies, topics, and publication dates.

Trained student employees searched each item in each resource; searches were given a 15-minute time limit. The students were trained in government documents work and were specifically trained on searching the five databases selected for this project. One student initially performed the searches and printed out results when a record for the item was located. These results were then checked by the authors. For any questionable results, or any items not located by the first student, a second student performed a search for the item, and the authors again checked the work. A final search was conducted by the authors on all items not located by the students.

As an example, the first item searched in the study by a student employee was A 13.2: T 71 *Trespass Law Enforcement Handbook, Forest Service Region 1, revised 1934*. The student did not find a record for the item in the OU Catalog, discovery system, HathiTrust, or Google. The student did find the item record in the CGP Catalog. A second student then searched for this same item record in the OU Catalog, discovery system, HathiTrust, and Google, and also did not find it. After this, the authors conducted the same search in the same four databases and also did not find a record. At this point, this document was determined to be 'not found' in those

four resources.

A note on items "found" on Google: we all understand the ease of finding an item on Google, but also understand that finding an item on Google does not imply a permanent online location for that item. When we "found" an item on Google, we did not further check whether that item was located on a government-sponsored website or a private site, and both were given equal weight as "found." As privately-posted documents could disappear tomorrow, one might assume that there would be greater longevity for an item located on a government website. But with the state of government affairs generally, we know that items on official government websites could also disappear.

Results

The study results showed that both the better established library resources (the library catalog and the CGP) and newer library resources (discovery services and HathiTrust) lag behind Google in the findability of government publications. Google had an 87% accuracy rate, followed by the discovery service, the library catalog, the CGP, and HathiTrust. Furthermore, 76% of the items searched were available full-text in Google.

Table 1: Results - Government Publications Findability

	Library Catalog	Discovery	CGP	HathiTrust	Google	Google Full-text
Study	75%	82%	68%	65%	87%	76%

Breaking the study into pre-1976 and post-1976 items, there were some differences. In the post-1976 subset it is logical that the most current items would be more findable in the contemporary databases, except for HathiTrust, which at present has more pre-1970 than post-1970 materials (HathiTrust, "Visualization Dates," 2016). The CGP and library catalog may lag here because of various cataloging backlogs that originated at the GPO. In the same way, for the pre-1976 subset, CGP had fewer and HathiTrust had more, which follows their collection patterns. In the post-1976 subset (66 items), the highest results were found, with an impressive 93% accuracy for the discovery service. In the pre-1976 subset (34 items), there was some reduction for the library catalog and discovery service as well as for Google. Fewer items were found full-text in Google as well.

Table 2: Results – Government Publications Findability, Pre- and Post-1976 Items

	Library Catalog	Discovery	CGP	HathiTrust	Google	Google Full-text
Full-text	75%	82%	68%	65%	87%	76%
Post-1976	86%	93%	89%	64%	89%	79%
Pre-1976	53%	62%	26%	68%	82%	71%

Discussion

Focusing on print government documents, the study encourages a serious reconsideration of the need for numerous print document collections. In a world where duplication is an insurance policy, where duplication makes sense because it is difficult to provide access via other means, numerous document collections have a purpose. But that is not today's world. Today, document collections are preserved online, and thus geographically dispersed print archives serving as backup to online collections have become a much more reasonable option to pursue. Libraries are a changing entity and document collections must change with them; part of that change may be to reduce substantially the number of document collections.

Print government documents collections are just part of the picture; document collections are also a potpourri of errata, miscellany, and ephemera produced in a vast array of formats, including flash cards, coloring books, and board games. This study, however, focused primarily on contemporary (post-1976) print items. With the GPO being established in 1861, and the FDLP quickly thereafter, document collections around the nation are filled with thousands of pre-1976 documents. Thirty-four percent of this study included pre-1976 items, and many of those items were not able to be located. Thus it is clear that a final study is needed for pre-1976 items and microforms. The authors plan to pursue such a study in the near future, replicating the same basic parameters and methodology. The final results could directly impact daily life for government documents librarians.

And what of those government documents librarians? Must we necessarily see a reduction in their numbers as well? Over time, this seems likely. But expertise in interpreting records and locating relevant government literature will still be needed. We can envision a future in which government documents expertise is still a welcome addition to the skill set

needed in libraries. Indeed, that expertise is vital for a wide variety of disciplines: engineers requiring technical reports, historians mining the historical record, and political scientists tracking trends in legislation or voting patterns, to name just a few.

Local collections and individual government documents librarians, however, are just part of the picture that is the Federal Depository Library Program. In digesting these issues, we also have to consider where the broader documents world stands. In the most recent biennial GPO report, the major problems and challenges to the FDLP libraries were listed as budget constraints, staff reductions/shortages, and increased workload (Outsell 2016, p. 46). The most important reported service continues to be "access to depository materials" (Outsell 2016, p. 65), which includes the physical collections. The top three unmet needs related to access were stated as digitized historical collections of government publications, additional historical coverage of titles in the Government Publishing Office's (GPO) Federal Digital System (FDSys), and adding pre-1976 cataloging records to OCLC (Outsell 2016, p. 72), all of which focus on the historical aspect of document collections.

The 2016 *National Plan for Access to US Government Information*, which presents a framework document to set and support the strategic direction of the Library Services and Content Management business unit of the GPO, directly addresses physical documents collections, quoting an earlier 2013 report: "To safeguard the historical documents of our democracy for future generations, GPO should work with depository libraries and other library groups to develop a comprehensive plan for preserving the print collection of government documents" (Government Publishing Office, 2016, p. 1, 5). One way this is being undertaken is through the designation of some collections as preservation stewards, committed to retaining portions of the print government publications record (Federal Depository Library Program, "Give Your Feedback," 2016). A complementary preservational track the GPO is taking is the Federal Information Preservation Network (FIPNet), which aims to establish local partners to accelerate digitization of and access to federal government information (Federal Depository Library Program, *Recommendations and Responses*, 2016). Still another valuable effort is the continuing work to gain bibliographic control of and access to pre-1976 publications via cooperative cataloging projects such as the one operated by the Maureen and Mike Mansfield Library at the University of Montana (Federal Depository Library Program, *Partnerships*, 2016).

Assuming forward movement, these are all promising developments. This current study demonstrates the benefits that have occurred through large-scale digitization projects. Working with other entities through FIPNet further extends these efforts. Preservation stewardship is very similar to efforts made by Scholars Trust (2016) and others to preserve the print record.

While WEST and Scholars Trust are designed to both preserve the print record and free space in member libraries, it is not yet clear whether preservation stewardship will result in institutional ability to repurpose library spaces currently devoted to holding government documents collections (University of California, 2016).

Conclusions

The local impact of our efforts to better understand the decline in use of the OU Libraries government documents collection has already resulted in changes. As of this writing, and with the approval of our regional depository library, our government documents public service point will be closed by the start of 2017. We will continue to build the collection by acquiring online publications. We will continue to fulfill ILL requests and provide reference assistance in person via appointment, by phone, or by email. Our immediate user community has been made aware of this change and no significant negative reaction has been registered.

This study began as a quest to ascertain which of the five government document databases in common use was most accurate for locating "known" paper government document records. The study resulted in a much greater exploration of the current situation of document collections and their fate within libraries. We found a high percentage of print government documents appear in full-text online through Google (on governmental and other websites), which could spell the end of highly-duplicative print document collections. A final study needs to be completed on pre-1976 and microform documents before final conclusions can be drawn. Those results may significantly affect the future of government document collections as we know them today.

Acknowledgements

The authors wish to acknowledge Ethan Bean and Tyler Lawhorn, the students who conducted searches for this project. Their steadfast and dedicated efforts were most helpful in procuring the data upon which this work is based.

References

Aharony, N., & Prebor, G. (2015). Librarians' and information professionals' perspectives towards discovery tools—An exploratory study. *The Journal of Academic Librarianship*, 41(4), 429-440.

Alexa. (2016). The top 500 sites on the web. Retrieved from http://www.alexa.com/topsites/category/Computers/Internet/Searching/Search_Engines

Anderson, R. (2005). The (uncertain) future of libraries in a Google world: Sounding an alarm. *Internet Reference Services Quarterly*, 10(3-4), 29-36.

Asher, A. D., Duke, L. M., & Wilson, S. (2013). Paths of discovery: Comparing the search effectiveness of EBSCO Discovery Service, Summon, Google Scholar, and conventional library resources. *College & Research Libraries*, 74(5), 464-488.

Brophy, J., & Bawden, D. (2005). Is Google enough? Comparison of an internet search engine with academic library resources. *Aslib Proceedings*, 57(6), 498-512.

Chickering, F. W., & Yang, S. Q. (2014). Evaluation and comparison of discovery tools: An update. *Information Technology and Libraries*, 33(2), 5-30.

Fagan, J. C., Mandernach, M. A., Nelson, C. S., Paulo, J. R., & Saunders, G. (2012). Usability test results for a discovery tool in an academic library. *Information Technology and Libraries*, 31(1), 83-112.

Farrell, M. (2005). Google and government documents. *Government Information Quarterly*, 22(2), 143-145.

Federal Depository Library Program. (2016). Give your feedback on the draft Memorandum of Agreement for Preservation Stewards. Retrieved from http://www.fdlp.gov/news-and-events/2664-give-gpo-your-feedback-on-the-draft-memorandum-of-agreement-for-preservation-stewards

Federal Depository Library Program. (2016). *Recommendations and responses–Fall 2015: Recommendation of the Depository Library Council to the GPO Director and responses from GPO*. Washington, DC: Government Publishing Office, 4.

Federal Depository Library Program. (2016). *Partnerships*. November 2016. Retrieved from https://www.fdlp.gov/about-the-fdlp/partnerships

Foster, A. K., & MacDonald, J. B. (2013). A tale of two discoveries: Comparing the usability of Summon and EBSCO Discovery Service. *Journal of Web Librarianship*, 7(4), 1-19.

Galileo Scholar. (2016). Catalog of U.S. Government Publications. Retrieved from https://www.galileo.usg.edu/scholar/databases/zdgc/?Welcome

Georgas, H. (2013). Google vs. the library: Student preferences and perceptions when doing research using Google and a federated search tool. *portal: Libraries and the Academy*, 13(2), 165-185.

Georgas, H. (2014). Google vs. the library (part II): Student search patterns and behaviors when using Google and a federated search tool. *portal: Libraries and the Academy*, 14(4), 503-532.

Georgas, H. (2015). Google vs. the library (part III): Assessing the quality of sources found by undergraduates. *portal: Libraries and the Academy*, 15(1), 133-161.

Government Publishing Office. (2016). *National plan for access to US government information: A framework for a user-centric service approach to permanent public access.* Washington, DC: Government Publishing Office, 1.

HathiTrust Digital Library. (2016). Our digital library. Retrieved from https://www.hathitrust.org/digital_library

HathiTrust Digital Library. (2016). Partnership community. Retrieved from https://www.hathitrust.org/community

HathiTrust Digital Library. (2016).Visualization dates. Retrieved from https://www.hathitrust.org/visualizations_dates

Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2015). NMC Horizon Report: 2015 Library Edition. Austin, Texas: The New Media Consortium, 83-84.

Kornblau, A. I., Strudwick, J., & Miller, W. (2012). How web-scale discovery changes the conversation: The questions librarians should ask themselves. *College & Undergraduate Libraries*, 19(2-4), 144-162.

Namei, E., & Young, C. A. (2015). Measuring our relevancy: Comparing results in a web-scale discovery tool, Google & Google Scholar. Proceedings of *ACRL 2015, March 25-28, 2015, Portland, Oregon*. Chicago: ACRL, 522-535.

Nichols, A., Billey, A., Spitzform, P., Stokes, A., & Tran, C. (2014). Kicking the tires: A usability study of the Primo Discovery tool. *Journal of Web Librarianship*, 8(2), 172-195.

Outsell, Inc. (2016). *2015 biennial survey of federal depository libraries & library needs assessment, report prepared for US Government Publishing Office (GPO)*, March 2016. Burlingame, CA: Outsell.

Sare, L. (2011). Availability of legacy government documents online. *Internet Reference Services Quarterly*, 16(1-2), 55-66.

Sare, L. (2012). A comparison of HathiTrust and Google Books using federal publications. *Practical Academic Librarianship: The International Journal of the SLA Academic Division*, 2(1), 1-25. Retrieved from https://journals.tdl.org/pal/index.php/pal/article/view/5880

Scholars Trust. (2016). Welcome to Scholars Trust. Retrieved from http://www.scholarstrust.org/

Schonfeld, R. C., & Housewright, R. (2009). *Documents for a Digital Democracy: A Model for the Federal Depository Library Program in the 21st Century. December 17, 2009*. Retrieved from http://www.sr.ithaka.org/wp-content/uploads/2015/08/Documents_Digital_Democracy.pdf

Singley, E. (2014). Discovery systems—Testing known item searching. *Usable Libraries*. Retrieved from http://emilysingley.net/discovery-systems-testing-known-item-searching/

Tay, A. (2015). 5 things Google Scholar does better than your library discovery service. *Musing About Librarianship*. Retrieved from http://musingsaboutlibrarianship.blogspot.com/search?q=5+things+google+scholar#.V8XlGfkrJhE

Thomsett-Scott, B., & Reese, P. E. (2012). Academic libraries and discovery tools: A survey of the literature. *College & Undergraduate Libraries*, 19(2-4), 123-143.

University of California, California Digital Library (2016). WEST: Western Regional Storage Trust. Retrieved from http://www.cdlib.org/services/west/