

2002

The Availability Of U.S. Government Depository Publications On The World Wide Web

Robert Lopresti

Western Washington University, robert.lopresti@wwu.edu

Marcia Gorin

Follow this and additional works at: https://cedar.wwu.edu/library_facpubs



Part of the [Library and Information Science Commons](#)

Recommended Citation

“The availability of US government depository publications on the World Wide Web.” by Robert Lopresti and Marcia Gorin. *Journal of Government Information*. (a peer reviewed journal) 29. 2002. 17-29.

This Article is brought to you for free and open access by the Western Libraries and the Learning Commons at Western CEDAR. It has been accepted for inclusion in Western Libraries Faculty & Staff Publications by an authorized administrator of Western CEDAR. For more information, please contact westerncedar@wwu.edu.

Robert Lopresti
Wilson Library
Western Washington University
Bellingham, WA
98225-9103
360-650-3480
fax: 360-650-6563
rob.lopresti@wwu.edu

Marcia Gorin
Documents/Maps/Micromaterials
Robert Manning Strozier Library
Florida State University
Tallahassee, FL
32306-2047
850-644-4702
fax: 850-644-4702
mgorin@mailier.fsu.edu

**The Availability Of U.S. Government Depository
Publications On The World Wide Web**
by
Robert Lopresti and Marcia Gorin

Abstract

In order to explore the prevalence of federal government information on the World Wide Web, two random samples were taken of material sent to federal depository libraries in the first half of the year 2000. Over the course of eighteen months four separate attempts were made to find each of the sample publications on the World Wide Web using a variety of search engines.

During the first attempt, one to five months after the material was sent to depository libraries, 47% of the publications in the sample were located on the Web. During the final and most successful attempt, 18 to 23 months after the distribution, 63% of the sample publications were located. Based on the sample, at present almost two-thirds of depository material are available on the Web approximately two years after the material was sent to depository libraries.

Introduction

For more than one hundred years the Federal Depository Library Program (FDLP) has been an important channel through which the U.S. government makes information available to the public. In 1985 the Government Printing Office (GPO) estimated that 195,214 people used depository collections in an average week. By 1997, well into the World Wide Web revolution, this figure had dropped to 155,415 (Aldrich, Cornwell, and Barkley, 2000, p. 281).

The spread of Internet use, and especially the government's embrace of the World Wide Web, has led some to question the need for the continued existence of the FDLP. For example, in May 2000 the U.S. House of Representatives Committee on Appropriations voted to cease funding the distribution of publications to libraries in tangible form because "the Committee believes that the programs for which the Superintendent of Documents is responsible could be performed with more cost effectiveness by making greater use of electronic information technology." (H. Rep. No. 106-635, 2000, p. 27).

As Andrew Sherman, GPO's Director of Congressional, Legislative and Public Affairs stated, the impetus for widespread electronic dissemination has been Congress's belief that "electronic information is free, or just about, and that it reaches everyone, or close enough, and that it is generally superior to printed information." (2000, p. 8) Congress voted for a smaller cut, which still amounted to a 6% drop in the Superintendent of Documents' budget for fiscal year 2000 (Sherman, 2000, p. 2).

Literature Review

The logic, implicit or explicit, for reducing the funding to the depository system, has been that there is no need to send publications to libraries because “everything is on the Web.” This belief has been debated in the library literature, as well as in broader media.

As early as 1997, Ryan stated that “technology advancements make government information so easy to distribute electronically that it is seemingly available everywhere.” (1997, p. 362) Ryan pointed out, however, that the facts do not bear out the appearance of wide-spread easy availability.

According to Corn, Congress has been slower than most other federal agencies in adopting the Web as a means of disseminating information. Much of the work of Congress, including working accounts of bills, (Corn, 2000, p. 228) voting records, financial disclosures (Corn, 2000, p. 230) and some congressional hearings, are not presently made available on the Web. Corn speculated that Congress’ inertia might be caused by not wanting to be scrutinized by the public (2000, p. 228).

Kram concludes that “clearly everything is not yet available on the Internet” (1998, p 65.) and not likely to be in the immediate future. Kram further observes that although GPO Access, the Government Printing Office's main website for providing official documents, is a very useful online service it still does not provide exhaustive coverage.

Coggins provides a brief account of the events that led to the electronic dissemination of federal documents (2000). His article addresses the issue of

authenticity and discusses the ramifications for the law community if certain key publications, such as the United States Code, Code of Federal Regulations and the Federal Register, are no longer provided in print. Coggins states that issues of “assurance of permanent public access and assurance of authenticity” need to be firmly established before a decision is made to no longer provide print copies of these and other publications on which the law community relies (2000, p 54.).

While the Web can be a useful vehicle for disseminating information, some categories of material are difficult to use online, such as long reports and maps, which are often slow to print or download. In 1998, Heisser surveyed library directors in 24 federal depository libraries about the electronic transition. (1999). Roughly one-half of the library directors were doubtful that everything would be online in the near future (Heisser, 1999, p. 256).

Heisser's respondents frequently voiced distrust of individual federal agencies' commitment to maintain information on their websites and a belief that agencies would not load all of their information. Others expressed the opinion that information would get lost or be hard to find. Although the future of the FDLP currently is uncertain, most library directors expressed a strong commitment to maintain their depository status (Heisser, 1999, p 256.).

In 1998, Superintendent of Documents Francis Buckley echoed documents librarians' concerns when he said, “Right now, there is nothing to ensure that electronic information is made available or to prevent it from being taken down.” Furthermore, Buckley stated that “agencies are not required to keep material up indefinitely.” (Fialkoff, 1998, p. 37) This concern was further

articulated in the ALA Washington News of April 1999 in a plea to persuade librarians and other concerned citizens of the need to change the law to provide permanent public access for active electronic documents. ("Public Access," 1999).

The difficulty of locating government information on the Web has also been observed. Notess agrees with the consensus that it can be difficult to find government information on the Web when documents often change locations and formats, or are removed completely (2000, p. 67). The former Superintendent of Documents, Wayne Kelley also expressed dismay with the problem of websites changing or disappearing and saw a need for a law that would provide permanent public access and the accurate maintenance of agency websites (Cheney, 1999, p. 264).

Laskowski surveyed librarians and users about electronic dissemination of information (2000). Her research shows that much of librarians' reference work is being conducted with online sources. She also indicates that the Web is a successful tool for locating fugitive documents. Some of Laskowski's respondents expressed concern that the transition to web-based reference tools creates access issues for patrons with substandard computer skills (2000, 178).

Portions of the literature provide historical accounts of the events leading to the present electronic transition of the FDLP. Some authors have even provided possible scenarios for the future of the FDLP (for example, Aldrich, 1998). The literature supports the belief that there are still archival and permanency issues that must be worked out and that as government information

moves online, there is a greater need for librarian intervention to help patrons locate the information they seek (Laskowski, 2000, p. 174).

Methodology

To explore the prevalence of government information on the Web, this study sampled government publications received in federal depository libraries. Each of the authors took a random sample of shipping lists, which enumerate materials sent to GPO depository libraries. From each of the selected shipping lists each author randomly selected a document title. A random number generator was used in all cases.

Figure A is a sample shipping list from the time period of the study. There are currently three types of shipping lists: paper, microfiche, and electronic. The average shipping list includes about 18 publications.

Sample A included one document each from roughly 9% of the shipping lists sent during January-April 2000. Sample B included one document each from approximately 10% of the shipping lists sent during February-May 2000. See Table 1.

In June/July 2000, the authors attempted to locate the 150 publications in the sample on the World Wide Web using three tools: the online version of the Catalog of United States Government Publications, Govbot, and the home page of the agency that authored each document. (See Appendix for a description of each of the search tools).

In September/October 2000 the authors searched for the same 150 publications, using the same search tools, plus Google's Uncle Sam, GrayLit Net, GPO Access and FirstGov. The fact that two of these search tools (GrayLit Net and FirstGov) were not yet available on the Web when the authors performed the first search in June/July 2000, indicates how quickly the World Wide Web is changing. During January/ February 2001 and November 2001 the authors repeated the searches using the same databases, plus Usgovsearch. See Table 2.

A search was considered successful if the publication was available on the Web in substantially the same form it was received at depository libraries, or if the publication was not available but the same information clearly was. The search was not counted a success if a newer edition had replaced the document the authors sought, or if it was only available via paid subscription, or if only part of the material was available.

Results

In June 2000, the authors were able to find slightly less than half (48%) of Sample A on the World Wide Web. In the three searches that followed the percentage changed to 58%, then 55%, and finally to 59%. See Table 3.

In June-July 2000, the authors found 45% of Sample B on the World Wide Web. During the searches that followed the percentage increased to 64%, then 65% and remained at 65% during the final search. See Table 4.

Taking the two samples together, 70 of the 150 sample documents were located during the first search; or almost 47%. See Table 5 for the totals from each search.

During the second search approximately three months later, all 70 hits from the first search were found again plus 22 more, for a total of 93. This is exactly 62% of the total sample.

In the third search, approximately four months later, the total dropped to 91, or slightly less than 61%. Changes that appeared during this search included 1 document that was found for the first time, the contents of 1 document that was partially available for the first time, 1 document that vanished from the Web, and 2 that were replaced by newer editions.

The fourth and final search, conducted approximately nine months after the third, found 94 documents, the highest result of any of the searches. This was almost 63% of our sample, which correlates closely with a recent report by the Government Printing Office that 62% of new titles from fiscal year 2001 were available in electronic format (Miller, 2001, p. 9). Changes included 4 documents that were found for the first time, 7 that disappeared, 3 that were replaced by newer editions, and the contents of 5 documents becoming available although in different formats from that distributed to libraries. Furthermore, 5 documents that had been only partially available were now found to be fully available.

It is worth noting that several documents were located during the second search only through search engines that were not available for the first search,

suggesting that those publications might have been on the Web all along.

However, that raises this question: if trained government documents librarians, armed with the information on shipping lists, were not able to find the publications, what chance would the average citizen have?

Many of the search engines are quite powerful. In one case, the authors were unable to locate a hearing from the homepage of the congressional committee that produced it, but found it using Google's Uncle Sam, a commercial search engine. Notess reported the same phenomenon (2000, p. 70).

In a few cases, only part of the material was available. For example, some U.S. congressional hearings were not available on GPO Access but could be found on the congressional committee's home page in partial form – usually including the written testimony of some or all witnesses, but not the actual transcript of the hearing.

In some cases the random sample called for a particular issue of a periodical, and the issuing agency had already replaced that issue on the Web with a more recent one and backfiles were unavailable. In another case, during the first search the authors could only find a newer issue than the one noted in the sample, but in later searches they found no issues of the periodical whatsoever.

The searches also yielded one instance where the paper issue sent to depository libraries was newer than the one available on the Web. Some other depository publications were available on the Web, only by subscription, or only

to employees of the agencies; thereby defeating the goal of public access. An example of the latter was the HUD 2020 Program Services & Operations Manual, Millennium Edition that was widely distributed in paper but on the Web was only available to Department of Housing and Urban Development (HUD) employees. As of November 2001 this publication was no longer available on the Web.

In terms of format, a greater percentage of government publications sent to depositories in microfiche were found on the Web than those sent in paper or electronic format. This may be explained by the fact that publications sent on microfiche tend to be slightly older than those sent on paper. In the final search, 67% of the microfiche sample was found, compared to 57% of paper, and 43% of the electronic documents. See Table 6.

Table 7 shows the results of the November 2001 search broken down by the agency that authored each document. While the sample size is too small to tell much about the online status of publications of individual agencies, it does suggest that some agencies -- such as the Federal Emergency Management Agency and the Nuclear Regulatory Commission -- may have a significant problem.

Overall, congressional documents represented the largest group of the sample (43 out of 150 or 29%). Nine of these documents (26%) were still not available on the Web almost two years after depository libraries received them. This further supports Corn's statement, mentioned above, that Congress has been slow to embrace the Web.

Findings

The issue of permanent public access (PPA), which has been mentioned several times above, has been attracting much attention (see, for example, Coggins, 2000 and Uhlir, 1997). Some government agencies appear to believe that “new is better” and are removing data from the Web as soon as a newer publication arrives. By the fourth search of this sample, 6 of the 150 documents had been replaced in this manner.

This has the potential for creating a gap in the historical record, especially if no paper copies remain in libraries. As the Electronic Transition Committee stated in its Spring 2000 report: "GPO needs to make PPA a constant drumbeat in its communication with, and education of, federal agency publishers." (Farrell, 2000, 337).

In addition, the availability of information on the Web can be affected by other factors, such as maintenance or updating of the sites. For example, two of the congressional documents were not located the first day of the November 2001 search, due to invalid links, but they reappeared the next day.

In the aftermath of the terrorist attacks on September 11, 2001 many government agencies removed material from the Web ("Access to government," 2002) Given the subject matter of the 7 documents from the sample that vanished between the last two searches, it is unlikely that these titles were

removed due to security concerns, but this does demonstrate the ease with which material can be removed from the Web.

Conclusion

While the World Wide Web is a powerful source for finding government information, a substantial number of new federal publications are not available on the Web in a timely manner. Approximately one-third of government publications are not accessible on the Web roughly two years after being sent to depository libraries.

There is also the issue of permanence - government publications on the World Wide Web can be removed at the whim of the creating agency. Congress, the Government Printing Office, and the library community still need to develop plans for Permanent Public Access. The best arrangement would probably be a safety net comprised of the Web, GPO Access, depository libraries, and GPO partnerships with libraries and agencies. The goal would be to make sure that even if a publication leaves the Web it would still be available through one or more libraries or permanent repositories.

Need for more research

The authors may return to these samples in the future to see whether this group of publications becomes more or less available over time. Different samples should be examined on the World Wide Web to see if the results are consistent or if Web availability is changing.

An unanswered question is whether an average citizen, untrained in government publications and research techniques, would be able to find these publications. Are government publications more - or less? - difficult to find than other full-text publications on the World Wide Web? This is a study waiting to be done.

Appendix. Search Engines Used

Catalog of U.S. Government Publications

http://www.access.gpo.gov/su_docs/locators/cgp/index.html

Search engine that provides the user with bibliographic records of both print and electronic government documents, many which were distributed through the federal depository system to depository libraries. Coverage begins with January, 1994 and contains many links to the electronic version of the document from the federal agency's website. Updated daily. Created and maintained by the Government Printing Office (GPO). (About the Catalog, 2001)

FirstGov <http://www.firstgov.gov/>

Searches across all federal websites to cover over 51 million websites produced by the federal government, and includes some local and foreign governments as well. It is administered by the U.S. General Services Administration and began operation on Sept. 22, 2000. (About FirstGov, 2002)

Google's Uncle Sam <http://www.google.com/unclesam>

A version of the Google search site, Google's Uncle Sam searches gov and mil domains and uses a unique ranking system which retrieves extremely relevant links (Our search, 2002) on a consistent basis, usually in less than half a second. Owned and operated by Google, Inc. (Serva, 2001).

GovBot <http://ciir2.cs.umass.edu/Govbot/>

Sponsored by the Center for Intelligent Information Retrieval, this search engine was restricted to searching .gov or .mil domains. By 2000, this database contained over 1.5 million Web pages (Notess, 2000, Government, 4) but it ceased its operation in May 2001.

GPO Access http://www.access.gpo.gov/su_docs/index.html

Provides free electronic access to government information products and databases, created by the federal government such as the Federal Register, Code of Federal Regulations, Congressional Record, U.S. Code, Commerce Business Daily, and the United States Government Manual, and furnishes access to over 1,500 databases with files available for download, and tools to assist you in finding government information. Full text documents accessed are the official versions. Created and maintained by the Government Printing Office (GPO). (About GPO Access, 2002)

GrayLit Network <http://www.osti.gov/graylit/>

Search engine that is capable of uncovering over 100,000 hard to find technical reports issued from such federal agencies as National Aeronautics and Space Administration (NASA), Department of Energy (DOE), Environmental Protection Agency (EPA), and the Department of Defense (DOD). It is sponsored by the Department of Energy's Office of Scientific and Technical Information in partnership with the Government Printing Office. Has the capability to search for these documents from many different databases so that the user does not need to know what agency produced the report. Began in August, 2000. (New Internet Tools, 2000)

Usgovsearch <http://usgovsearch.northernlight.com/govsearch.html>

Powerful search engine that retrieves abstracts to federal government information products from government sources and from the Special Collections database that includes 7100 full-text journal and newspaper articles, reference sources and NTIS abstracts and titles available since 1964. Full-text documents may be purchased and accounts may be set up. (Special Collection Overview, (n.d.)). Sponsored by the National Technical Information Service (NTIS) and Northern Light Technology, Inc. (What usgovsearch, (n.d.)).

References

- About FirstGov (2002). Retrieved March 28, 2002, from
<http://www.firstgov.gov/About.shtml>
- About GPO Access. (2002). Retrieved March 25, 2002, from
http://www.access.gpo.gov/su_docs/index.html
- About the Catalog of U.S. Government Publications. (2001). Retrieved March 28, 2002, from http://www.access.gpo.gov/su_docs/locators/cgp/index.html
- Access to government information post September 11th. (2002). OMB Watch.
Retrieved March 31, 2002, from
<http://www.ombwatch.org/article/articleview/213/1/1/>
- Aldrich, D.M. (1998). Partners on the Net: FDLIP partnering to coordinate remote access to Internet-based government information. *Government Information Quarterly*, 15, 27-38.
- Aldrich, D.M., Cornwell, G., & Barkley, D. (2000). Changing partnerships? Government Documents Departments at the turn of the millennium. *Government Information Quarterly*, 17, 273-290.

Cheney, D., Tulis, S. & Peterson, S. (1999). JGI interview. What are they really after: control of information is very important. An interview with Wayne P. Kelley, former U.S. Superintendent of Documents. *Journal of Government Information*, 26, 259-269.

Coggins, T. L. (2000, October). Print no more: U.S. Code, Code of Federal Regulations, and the Federal Register. *Virginia Lawyer*, 49, 53-55.
Retrieved November 30, 2000, from
http://www.vsb.org/publications/valawyer/oct00/legal_resource.pdf

Corn, D. (2000, November). Filegate.gov. *Wired*, 8 (11), 226-37.

Kram, L. (1998). Why continue to be a depository library if it is all on the Internet anyway? *Government Information Quarterly*, 15, 57-71.

Farrell, M. (2000). U.S. Depository Library Council Electronic Transition Report. *Government Information Quarterly*, 17, 331-338.

Fialkoff, F. (1998, May 15). Francis Buckley: Protecting the public interest. *Library Journal*, 123, 36-37.

Heisser, D.C.R. (1999). Federal Depository Program at the crossroads: the library administrator's perspective. *Government Information Quarterly*, 16, 241-259.

H. Rep. No. 106-635 at 27 (2000).

Laskowski, M.S. (2000). The impact of electronic access to government information: what users and Documents specialists think. *Journal of Government Information*, 27, 173-85.

Miller, A.E. (2001, Fall). Reporting From The ALA Conference. *DttP*, 29, 9-10.

New Internet Tools Enhance Access to Federal Information. DOE News, August 14, 2000. Retrieved March 28, 2002, from

<http://www.osti.gov/promo/pressgrayfed.html>

Notess, G.R. (2000, June/July). Finding online government publications.

Econtent, 23 (3), 67-70.

Notess, G.R. (2000). Government Information on the Internet. Lanham,

Maryland: Bernan.

Our search: Google Technology. (2002). Retrieved March 25, 2002, from

<http://www.google.com/technology/index.html>

Public access to government information. (1999, April). ALA Washington News,

51, 6-7.

Ryan, S.M. (1997). Square peg in a round hole. Journal of Government

Information, 24, 361-75.

Serva, S. (October, 2001). Google with new GEO. Econtent, 24, no.8, 8-9.

Retrieved March 21, 2002, from Library Literature database.

Sherman, A.M. (2000, November 15). FY 2001 Appropriations cuts for the

FDLP. Administrative Notes, 21, 2-8.

Special Collection Overview. (n.d.). Retrieved March 31, 2002, from

http://usgovsearch.northernlight.com/docs/specoll_help_overview.html

Uhlir, P. (1997). Framework for the preservation of and permanent pubic

access

to USDA digital publications. Retrieved November 14, 2000, from

<http://preserve.ham.usda.gov:8300/npp/frameprt.html>

What usgovsearch Is All About. (n.d.). Retrieved March 31, 2002,

http://usgovsearch.northernlight.com/docs/about_company_mission.html