Fugitive Documents

A Case Study of US Forest Service Scientific Reports

Carol A. Singer

Librarians who work with government publications have long been concerned about the many US government documents that remain inaccessible and, in some cases, difficult to discover. In 1976, Ruth Smith reported to the Public Printer's Depository Library Council, "A conservative estimate is that 50% of the Federal documents published are not main stream publications. In one way or another they manage to elude national announcement. They are not sent to GPO or NTIS [US National Technical Information Service] and are not widely advertised."[1] In 1993, Peter Hernon expressed this common concern: "We can question how the public can learn about the existence of particular information resources and services, how public access can be guaranteed and enhanced, and how information services can be standardized and seamlessly linked for better use."[2] The problem is considered so serious that in 2004 the Fugitive and Electronic-Only Documents Committee of the American Association of Law Libraries Government Documents Special Interest Section sponsored the first annual Fugitive Documents Week to encourage librarians to report fugitive documents to the US Government Printing Office (GPO).[3]

As distribution of government documents increasingly transitioned to an online service, librarians worried that this would further impede the FDLP's ability to provide depository library access to the publications produced by US government entities. As Kristi Jensen noted,

> A more dispersed network environment that allows publishing agencies to bypass the traditional reporting mechanisms means that some resources may never be included in library catalogs or online indexes, the tools frequently used in an academic research library to provide access to government information. . . . Thus, rather than becoming more accessible these undistributed online documents may become almost entirely inaccessible to the typical user in a large research library.[4]

This inaccessibility is caused, in part, by the lack of a single resource that lists all documents published by US government entities. It is impossible to determine what percentage of government publications, whether they are tangible or virtual, are distributed to depository libraries because there is no accurate count of how many US government publications are produced.

Although government documents librarians frequently refer to these missing documents as fugitive documents, exact definitions for the term vary. In 1975, Cynthia Bower defined a fugitive as "any federal publication that my library—a regional depository—failed to receive on deposit."[5] A 2001 General Accounting Office report defined fugitive documents as "documents that should be—but are not—distributed by the Superintendent of Documents to the depository libraries."[6] In 2003, Gil Baldwin, director of GPO's Library Programs Service, defined a tangible fugitive document as "a U.S. Government publication that falls within the scope of the Federal Depository Library Program (FDLP), but has not been included in the FDLP," but noted that there was a second category he termed "online fugitives." This category consisted of online publications whose existence was not reported to the GPO. He estimated there might be 250,000 online fugitive publications that should have been part of the FDLP.[7] A 2004 article in a GPO newsletter defined fugitive documents as "those documents of public interest or educational value, not classified for reasons of national security, which have not been acquired for distribution to Federal depository libraries or brought under bibliographic control through the Catalog of U.S. Government Publications."[8] In 2005, Jacobs, Jacobs, and Yeo defined fugitive documents as "publications that are not entered into the national bibliographic record nor distributed to FDLP libraries."[9]

One problem area in access and discovery of government publications has been federal scientific and technical information (STI). A 1990 Office of Technology Assessment report studied the problem of dissemination of federal STI and concluded that four key areas would determine the success of federal scientific and technical information. One of these was "indexing of databases and documents, so that STI users in and out of the government know what and where STI exists."[10] David Gold wrote in 1993 that innefective indexes means, for the public, "there is no access point to comprehensive information on federally-produced STI. Due to this, researchers and engineers in academia, the private sector, and even the Federal government cannot find out quickly about *all* STI which has resulted from Federal R&D in a given area of interest to them."[11] Gil Baldwin quoted a former director of the NTIS who said in 2000 that there could be "50,000 gray literature NTIS titles" that should have been part of the FDLP.[12] An Inspector General investigation of National Institutes of Health (NIH) publishing found that NIH had supplied the FDLP with adequate copies of only ten of the sixty-two publications included in their test sample.[13] The report noted, "By NIH not providing copies of publications to GPO for FDLP distribution, Depository Libraries, and the public who use them, do not have ready access to documents to which they are entitled, that were printed with taxpayer funding."[14]

In 2001, Kristi Jensen published the results of an attempt to identify US Geological Survey open-file reports missing from the Pennsylvania State University collection. Jensen initially identified more than 1,300 reports using a variety of resources and finalized a list of 240 items that were missing from her library's catalog and also from the US Geological Survey's list of open-file reports. She then searched for these documents in GeoRef, WorldCat, the Catalog of US Government Publications (CGP), and the catalogs of two similar universities. She found that none of these databases included records for almost 30 percent of the 240 documents.[15]

Lisa S. Nickum described the historical difficulties in finding federally funded technical reports, concluding that most are "not available in other widely used commercial databases with related journal literature. The problems with dissemination, accessibility, and bibliographic control have led, understandably, to the belief that the federally funded technical report literature is difficult, if not impossible, to identify and locate."[16]

The ability to identify and access full text of the publications of the US Forest Service's Northern Research Station (NRS) published from 2012 through 2016 served as a case study in access to and discovery of recent US government scientific and technical literature. According to the webpage titled "About the Northern Research Station" (www.nrs.fs.fed.us/about/) the NRS is one of seven Forest Service research units and covers an area that includes states from Minnesota in the north to Missouri in the south and to Maine in the northeast. A list was compiled of 361 documents published by the NRS during those five years, using Forest Service resources, in addition to internet and database searches.

During July and August 2017, a variety of online resources were searched to explore how easily a researcher would be able to identify any of these 361 publications. Each publication was searched by title and, if necessary, keywords or author.

The initial question to be addressed was whether these publications were considered fugitive documents. Although the definition of a fugitive document varies, a common thread among the definitions is that these documents were not distributed by the FDLP. If a document was included in the FDLP, in either tangible or virtual format, it should be listed in the CGP (catalog.gpo.gov/). Only 234 of the 361 publications (65 percent) were found in the CGP, resulting in 127 fugitive documents issued by the NRS between 2012 and 2016.

Although a researcher looking for research published by the federal government might have chosen to search the CGP, there are other resources that might be searched for the types of research conducted by the Forest Service.

If a researcher was looking specifically for Forest Service publications, the obvious choice would be Treesearch (www.fs.usda.gov/treesearch/), the Forest Service's database of full-text scientific publications authored by Forest Service scientists, including reports, journal articles, conference proceedings, and books. Because this database doesn't include documents that are not considered to be scientific, 20 (6 percent) of the 361 publications in the test sample were not in the Treesearch database. However, not all researchers who could benefit from using NRS publications would use this database because the Forest Service publishes research on topics that are not exclusively about forests.

The NRS also provides access to their publications if a researcher knows which NRS scientific series would include the needed publication. Their "Publications and Data" page (www.nrs.fs.fed.us/pubs/) links to lists of the documents in these series, which include General Technical Reports, General Technical Reports—Proceedings, Information Forestry, Resource Bulletins, Research Maps, Research Notes, and Research Papers. In addition, their "News Releases" page (www.nrs.fs.fed.us/news/) includes a list of publications in the Station's Research Releases series, with links to the full text of each document. These two pages collectively provide citations and links to full text of 347 (96 percent) of the 361 documents.

However, not all researchers would know that the Forest Service or NRS might issue documents that would be useful for their work. As Gold explained, "Unclassified Federal STI is currently available, but it only has an impact if a researcher can find useful information quickly."[17] How would a researcher discover the existence of relevant NRS publications without deliberately searching for these reports at Forest Service or NRS websites?

One database that allows for discovery of current federal scientific publications is Science.gov (www.science.gov/). This database searches more than sixty federal databases and scientific websites for federal government scientific and technical information, including Treesearch. Because of this, 361 (94 percent) of the NRS publications were in the Science.gov database.

The National Technical Reports Library (NTRL) (ntrl.ntis.gov/NTRL/), a service of the NTIS, provides indexing and full text for a wide variety of US technical reports. In 1994, David Gold wrote, "In theory, NTIS is supposed to serve as a (non-exclusive) centralized repository and distributor for Federal STI allowing researchers access to a comprehensive system to find the information they seek."[18] However, only 215 (60 percent) of these NRS reports were part of the NTRL.

Another federal database that searches scientific information is the National Agricultural Library database, NAL Catalog (AGRICOLA) (agricola.nal.usda.gov/). This database includes records for agriculture and allied disciplines, including forestry. AGRICOLA offers separate searches for books and articles in addition to a combined search. The titles in this test sample were searched using the Book Search. Because the Forest Service and the National Agricultural Library are both part of the US Department of Agriculture, and because forestry is one of the subjects included in this database, a researcher might assume that these NRS documents would be part of the AGRICOLA database. Unfortunately, the researcher would be disappointed: only 203 (56 percent) of the 361 papers in the test sample were in the NAL Catalog (AGRICOLA).

All publications were searched in OCLC's WorldCat database (www.worldcat.org/) to get an indication of whether these NRS documents were available in libraries. Although this database doesn't include holdings for all libraries in the US, it is a source of holdings for a very large number of libraries. The WorldCat database included records for only 285 (79 percent) of the 361 documents.

Google Scholar (scholar.google.com/) is a popular search engine that indexes and provides access to a wide variety of journal articles, books, reports, etc. Beckmann and von Wehrden compared Google Scholar to Web of Science and concluded that "due to its full-text search capabilities, [Google Scholar] is an important and very useful tool to search the literature. To date, it has been widely overlooked by the scientific community."[19] Google Scholar provided records for 245 (68 percent) of the 361 documents in the test sample.

Table 1. Search results

Database	Papers Found (*N* = 361)	Percent
Northern Research Station Series lists	347	96%
Treesearch	341	94%
Science.gov	341	94%
WorldCat	285	79%
Google Scholar	245	68%
Catalog of US Government Publications	234	65%
National Technical Reports Library	215	60%
Agricola	203	56%

In addition to the sources listed above, the commercial databases BIOSIS Previews (wokinfo.com/products_tools/specialized/bp/), Environment Complete (www.ebsco.com/products/research-databases/environment-complete), and GreenFILE (www.ebsco.com/products/research-databases/greenfile) were also searched for records of these documents, but none were found.

As the results show, none of the sources that were searched included all of the 361 publications in this sample. These searches were performed in July and August 2017, so additional reports from the test sample may now be indexed in one or more of these sources. The results of the searches, in order by percentage of NRS documents included in the source, are in table 1.

In addition to being able to search for previously published documents, scientists also strive to maintain a knowledge of current research. There are several resources for those who wish to know the most recent NRS publications. The NRS announces new publications on its "New Station Publications" page (www.nrs.fs.fed.us/pubs/updates/), which provides a selective list of recent NRS publications in addition to links to previous iterations of this list. These lists collectively announced the publication of 261 (72 percent) of the 361 publications. The NRS also announces some of their new publications on their Twitter feed (@usfs_nrs). Their Twitter feed was used to announce 202 (56 percent) of the 361 items published from 2012 to 2016. The Forest Service Library compiles selective lists of recent Forest Service publications, which are published in the *Journal of Forestry*. These lists are composed of documents from all of the Forest Service research units. Records for 138 (38 percent) of the documents in this case study were included in these lists.

If a fugitive document is defined as one that is missing from the FDLP, this exploration revealed disturbing news for government documents librarians. Slightly more than one-third of the documents in this sample were missing from the CGP.

However, this is a test only of the documents published from 2012 to 2016 by one unit of an agency within one cabinet-level department. To make a more definitive statement about the extent of fugitive documents, much more research would be needed. It is not clear that the results of a search for items published in other periods by other research units within the Forest Service, by other agencies within the US Department of Agriculture, or by other federal government units would yield similar results. Instead, the results of this exploration could be considered to be a canary in the coal mine, a warning that the number of fugitive documents may be quite significant.

Carol A. Singer (singerc@bgsu.edu), Reference and Instruction Librarian, Bowling Green State University.

Notes

1. US Superintendent of Documents, Depository Library Council to the Public Printer, *First Report to the Public Printer 1972–1976* (Washington, DC: US GPO, 1978).
2. Peter Hernon, "Discussion Forum: Equity in Public Access to Government Information," *Government Information Quarterly* 10, no. 3 (1993): 302, https://doi.org/10.1016/0740-624X(93)90014-Q.
3. "First Fugitive Documents Week Nets 222 New Titles," *Administrative Notes* 25, no. 5 (April 14, 2004): 1, tinyurl.com/y7cqlx5x.
4. Kristi L. Jensen, "Providing Access to Online Government Documents in an Academic Research Library Collection," *Science & Technology Libraries* 20, no. 2/3 (2001): 17, https://doi.org/10.1300/J122v20n02_03.
5. Cynthia Bower, "Federal Fugitives, DNDs and Other Aberrants: A Cosmology," *DTTP: Documents to the People* 17, no. 3 (September 1989): 120, https://searchworks.stanford.edu/view/xm677yc8398.
6. US General Accounting Office, *Information Management: Electronic Dissemination of Government Publications,* Report no. GAO-01-428 (Washington, DC: US Government Printing Office, 20010, https://www.gao.gov/new.items/d01428.pdf.
7. Gil Baldwin, "Fugitive Documents—On the Loose or On the Run; Presentation by Gil Baldwin, Director, Library Programs Service, GPO; American Association of Law Libraries Conference, Seattle, WA, July 15, 2003," *Administrative Notes* 24, no. 10 (August 15, 2003): 4–5, https://tinyurl.com/ybttcx6g.
8. "First Fugitive Documents," 1.
9. James A. Jacobs, James R. Jacobs, and Shinjoung Yeo, "Government Information in the Digital Age: The Once and Future Federal Depository Program," *Journal of Academic Librarianship* 31, no. 3 (2005): 198, https://doi.org/10.1016/j.acalib.2005.02.006.
10. US Office of Technology Assessment, *Helping America Compete: The Role of Scientific & Technical Information,* Report no. OTA-CIT-454 (Washington, DC: US GPO, 1990), 2, http://ota.fas.org/reports/9023.pdf.
11. David Gold, "Improving the Impact of Federal Scientific and Technical Information: A Call for Action," *Government Information Quarterly* 10, no. 2 (1993): 224, https://doi.org/10.1016/0740-624X(93)90052-2.
12. Baldwin, "Fugitive Documents," 5.
13. US Department of Health and Human Services, *Office of Inspector General. Review of the National Institutes of Health Printing Program,* Report no. A-15-98-80001 (Washington, D.C: US GPO, 1999), 5, https://oig.hhs.gov/oas/reports/phs/c9880001.pdf.
14. Ibid., i.
15. Jensen, "Providing Access to Online Government Documents," 20–23.
16. Lisa S. Nickum, "Elusive No Longer? Increasing Accessibility to the Federally Funded Technical Report Literature," *Reference Librarian* 94 (2006): 35, https://doi.org/10.1300/J120v45n94_04.
17. Gold, "Improving the Impact of Federal Scientific and Technical Information," 229.
18. Ibid., 223.
19. Michael Beckmann and Henrik von Wehrden, "Where You Search Is What You Get: Literature Mining—Google Scholar Versus Web of Science Using a Data Set from a Literature Search in Vegetation Science," *Journal of Vegetation Science* 23 (2012): 1199, https://doi.org/10.1111/j.1654-1103.2012.01454.x.