

**Digital Deposit Working Group
of the Depository Library Council
Final Report and Recommendations**

March 24, 2022

MEMBERS OF THE DIGITAL DEPOSIT WORKING GROUP

Members from the Depository Library Council:

Robbie Sittel, Chair, University of North Texas Libraries

Julia Ezzo, Michigan State University Libraries

Will Stringfellow, Vanderbilt University Libraries

Vicki Tate, University of South Alabama Libraries

Members from the Depository Library Community:

Heather Christenson, HathiTrust

James R. Jacobs, Stanford University Libraries

Members from the U.S. Government Publishing Office:

Cynthia Etkin, Office of the Superintendent of Documents

Lisa LaPlant, Programs, Strategy and Technology

Jessica Tieman, Library Services and Content Management

Acknowledgements

The Digital Deposit Working Group (DDWG) of the Depository Library Council acknowledges and thanks the many volunteers who gave of their time and greatly assisted in the work of this project. They responded to surveys, participated in focus groups, searched for unreported publications, and helped to visualize what digital deposit could look like. Without them, the DDWG could not have accomplished its work.

Kris Abery	Natalie Fulkerson	Jackie Magagnosc	Bruce Sarjeant
Ben Amata	Valerie	Katherine Miles	Jesse Silva
Barbara Bren	Jesse Griffin	Craig Newton	Robbie Sittel
Charleen Bryson	Kathy Hale	Mark Phillips	John Stevenson
Susanne Caro	Cass Hartnett	Blaine Redemer	Christopher Straughn
Nancy Donahoo	James Jacobs	Suzanne Reinman	Kate Tallman
Aaron Elkiss	Michael Kardos	Deborah Rixon	Vicki Tate
John Elson	Jen Kirk	Cecilia Robinson	Emily Weichmann
Eva English	Jake Kubrin	Stuart Sanchez	Beth Williams
Tom Fischlschweiger	Shari Laster	Laura Sare	Kimberly Wirth
Sara Frug			

Table of Contents

Acknowledgements.....	<i>ii</i>
Introduction	1
Background: work prior to the seating of the working group.....	2
Activities of the Digital Deposit Working Group.....	5
Defining Digital Deposit	5
Spring 2019 Conference: Digital Deposit: A Value Proposition	5
Fall 2019 Conference: Digital Deposit: Collection Development for the 21st Century	6
Digital Deposit Pilot Project.....	7
Unreported Publications Survey.....	8
Focus Groups on Notifying GPO of Unreported Publications	10
In Search of Congressionally Mandated Reports	12
Visualizing the Deposit of Digital Content from GPO	14
Conclusion	16
Final Recommendations to the Depository Library Council.....	17
Appendix I: Digital Deposit: Collection Development for the 21st Century	Appendix I-1
Appendix II: Unreported Documents Survey.....	Appendix II-1
Appendix III: Focus Groups: Notifying GPO of Unreported Publications.....	Appendix III-1
Appendix IV: In Search of Congressionally Mandated Reports	Appendix IV-1
Appendix V: Visualizing Digital Deposit Subgroup Final Report.....	Appendix V-1

Introduction

As James Madison¹ and many other political theorists and philosophers have long noted, public access to government information is indispensable to a working democracy. Libraries have been integral to public access to federal information since the dawn of the Federal Depository Library Program (FDLP) in 1813. The core tenet of the FDLP -- indeed, it's in the very name of the program! -- is the deposit of government information in geographically dispersed libraries as the best means to maintain long-term public access and preservation.

With the advent of the internet and digital publishing of government information on the world wide web, there soon were raised discussions around the idea of "digital deposit" -- the distribution to libraries not of paper documents but of digital files to better promote digital access to and preservation of government information. The Government Publishing Office (GPO) has since 2005 included various questions about "digital deposit" in its biennial surveys to depository libraries.²

Because of interest in digital deposit evidenced by responses to the biennial surveys and through FDLP Conference sessions, the Depository Library Council (DLC) recommended, in 2018, to the Government Publishing Office (GPO) "the creation of a working group to explore current and future needs related to digital deposit - both dissemination of content and acceptance of content by GPO."³ In justifying the recommendation, Council noted that having a working group "is a critically important and inclusive step in reaching consensus on how federal information in digital forms should be disseminated to and amongst the FDLP community for the benefit of all our users."

In response, GPO agreed with Council that exploration of digital deposit is important to the FDLP community and supported the establishment of a working group. The Digital Deposit Working Group (DDWG) includes members of Council, representatives from the FDLP community, and staff from multiple GPO units.

¹ "A popular government, without popular information, or the means of acquiring it, is but a prologue to a farce or a tragedy; or, perhaps both." - James Madison, 1832.

² "Digital Deposit And The Biennial Survey: context and actions." *Free Government Information*, November 1, 2017. <https://freegovinfo.info/node/12457/>.

³ Recommendations & Commendations of the Depository Library Council to the GPO Director & GPO Responses, Fall 2018. https://www.fdlp.gov/sites/default/files/file_repo/dlcrec-fall-2018-GPO-responses.pdf

Background: work prior to the seating of the working group

In the 2005 FDLP Biennial Survey, 33% of respondents said they were interested in receiving digital files from GPO.⁴ At that time⁵, GPO had several general assumptions about what constituted “Digital Distribution,” such as:

- Electronic files of Federal publications would continue to be available for downloading by Federal depository libraries under FDsys.
- Under an FDLP distribution option, distributed digital publications would be sent to depository libraries via a “push” mechanism.
- A high percentage of libraries did not want to receive electronic files at all. Those libraries that wished to receive them wanted only a very low volume.
- Redundancy is needed to ensure future public access and FDsys will provide this capability, either by storage at multiple sites, by relying on preservation partners, or both mechanisms.
- GPO will distribute digital files optimized for public access. These files will typically be smaller than the archival copies preserved by GPO.
- If GPO distributed digital files to libraries under the aegis of the FDLP, then the various requirements and obligations of Title 44 apply.

On April 10, 2007, GPO shared an Information Brief on “Digital Distribution”⁶ at the Spring Depository Library Council Meeting in Denver, Colorado (April 5-18). It stated that during FY 2006, 93% of all new titles made available through the FDLP were available in electronic form; thus, GPO would consider “an affirmative distribution of authenticated and officially published digital content to Federal depository libraries.”

In 2007, the questions and concerns regarding the implementation of a digital distribution service included issues relating to how to authenticate files, synchronize files across libraries and websites, requirements for selective and regional libraries to retain derivative files, versioning of electronic publications, and the role of libraries and redundancy of digital information for preservation.

⁴ See historical biennial survey data on digital deposit at “Digital Deposit And The Biennial Survey: context and actions.” *Free Government Information*, November 1, 2017. <https://freegovinfo.info/node/12457/>.

⁵ Biennial Survey of Depository Libraries: 2005 Results. It should be noted that this was during the pre-smartphone era where digital access was not ubiquitous. <https://www.fdlp.gov/biennial-survey-of-depository-libraries-2005-results>.

⁶ U.S. Government Printing Office. *Digital Distribution* [Information Brief, April 10, 2007]. Accessed February 10, 2022. <https://www.fdlp.gov/2017-dlc-april-17-digital-distribution-issue-brief>.

Three years following the “Digital Distribution” information brief, Superintendent of Document Policy Statement (SOD) 321 *Digital Dissemination of Access Content Packages for FDLP Digital Depository* (2010)⁷ established a policy for the digital dissemination of access content packages (ACPs) through GPO’s digital repository to Federal depository libraries (FDL’s). This policy was established in order to respond to increasing interest from FDL’s wishing to establish digital collections of content within scope of the FDLP in order to meet both their user needs and provide replication and redundancy of digital content across geographically dispersed and institutionally diverse digital infrastructures. Rather than “pushing” digital content through an automated means, the GPO policy set the stage by which FDLs would instead proactively “pull” content from GPO’s online access interfaces whether by means of saving and downloading individual content directly, or harvesting content through an application programming interface (API) or similar technology.

This SuDoc policy followed FDLP community initiatives such as the 2009 Fall DLC Meeting presentation, [*Demystifying Digital Deposit: What It Is and What It Could Do for the Future of the FDLP*](#). This early presentation first introduced a concept of “digital deposit” to the FDLP community by which Digital Deposit, as a service model or as a collective initiative, would provide potential organizational and technological infrastructure for replication and redundancy of FDLP content across a network of libraries and institutions. This collective activity provides potential risk mitigation for financial risks, diversifies technology infrastructure, and improves heterogeneous collections, access, and preservation activities in order to better meet varied user needs and expectations. In 2010, in a *Library Journal* interview, acting Superintendent of Documents Ric Davis, stated that libraries were interested in GPO establishing a trustworthy digital repository as well as hosting their own local copies of GPO content. Recognizing that the FDLP was always built upon a distributed model, these efforts led GPO to participate in the LOCKSS-USDOCS ("Lots of Copies Keep Stuff Safe") alliance⁸ after a 12-month pilot project. GPO continues to enable LOCKSS technology, developed by Stanford University, to create and duplicate copies of govinfo content as part of a collective effort to collect, preserve, and provide access to multiple copies of digital publications.

While the increased interest in “digital deposit” continued over the years, GPO also responded to an increased interest in digital access and digital collection development for FDL communities. In August of 2014, after GPO’s 2013 audit and report by the National Academy of Public Administration (NAPA), GPO revised the [*Legal Requirements and Program Regulations of the Federal Depository Library Program*](#) to no longer require FDLP selective depositories to

⁷Superintendent of Documents, *Digital Dissemination of Access Content Packages for FDLP Digital Depository* [SOD Policy 321], May 2010.

⁸ GPO Joins Alliance for Digital Preservation, June 14, 2010. <https://www.fdlp.gov/alliance-digital-preservation>. Digital Federal Depository Library Program, LOCKSS-USDOCS <https://lockss-usdocs.stanford.edu>.

select the two item numbers used by GPO to send materials to all FDLs, thus creating the possibility of a digital-only depository library. A digital-only library should not be confused with digital deposit and it should be noted that digital-only depository libraries are not required to select any tangible items and do not receive digital files from GPO.

In 2016, the Superintendent of Documents developed and released the [*National Plan for Access to U.S. Government Information: A Framework for a User-Centric Service Approach to Permanent Public Access*](#) (2016) establishing several principles of desired outcomes and actions through the successful plan implementation. The *National Plan*, built upon the [*Federal Depository Library Program Forecast Study*](#) (2014), includes the desired outcome of increased quantities of all digital depositories in the FDLP. The *Forecast Study* responses specific to “digital deposit” were mentioned in parts of the survey associated with questions relating to both access and preservation; this suggests that, across the FDLP, library expectations and definitions of “digital deposit” as a service may vary, or, perhaps may exist to serve different needs with various priorities depending on the library.

Interest in digital depository collections has been measured in the results of GPO’s Biennial Surveys and has remained fairly constant (2005, 2007, 2009, 2011, and 2015). In 2015, 24% of libraries who responded to the survey indicated an interest in receiving digital files of online publications on deposit; 51% discussed it, and 21% supported it. 10% also responded that they were actively harvesting or preserving Federal agency website content and 15% responded that they were interested in receiving digital file deposits as a service offered to the FDLP. At the Fall 2017 DLC Meeting and FDL Conference, an open forum was held to continue a focused discussion about interest in “digital deposit” and to raise issues, ideas, and potential next steps. Since this time, the FDLP has expressed varying expectations for how “digital deposit” as a service can best meet community needs.

As shared in draft format at the Fall 2018 meeting of the Depository Library Council, Superintendent of Documents, Laurie Hall presented [*Developing a Multi-State Comprehensive Collection | FDLP Preservation Services Pilot Strategies*](#). This report proposed to “define a pilot project, including identification of needs and resources, in order to implement effective solutions for depositing digital content within the FDLP with respect to needs for content integrity, preservability, collection development, usability, and user community priorities.” As part of this effort, GPO proposed identifying a methodology for the collection of qualitative information from Federal depository libraries regarding the needs and expectations of “digital deposit” as a service of GPO. This methodology could take the form of focus group studies within the Federal Depository Library Program (FDLP).

The Digital Deposit Working Group was formed to collect specific information from FDLs regarding which technology services, content and collections, and aspects of digital deposit are of the most utility and of the highest priority; that is, are FDLs most concerned about access or

preservation? Are FDLs most interested in preserving quality documents for local retention, or large-scale harvesting methods for content analysis? GPO must define the scope and limitations of such a service and who its target stakeholders are now and into the future as digital deposit may evolve over time as new technologies become available. This information-gathering could objectively define the needs of the FDLP and identify effective solutions in order to develop a comprehensive service.

Activities of the Digital Deposit Working Group

After formally receiving a charge from the Depository Library Council in 2018, the Digital Deposit Working Group was formed. Following is an overview of the working group's activities.

Defining Digital Deposit

The DDWG discovered early in its work a lack of consensus for a standard definition of digital deposit among both the working group members and the FDL community. Therefore, before developing a work plan, the group had to define digital deposit. The DDWG envisions digital deposit as three interrelated approaches for the deposit of born-digital government information. In the digital era, the internet allows for the traditional means of deposit of government information to be expanded and flow to/from GPO in multiple ways.

The DDWG defines **digital deposit as the practices, services, and workflows for the collaborative acquisition of born-digital and digitized Federal Government information for the National Collection of U.S. Government Public Information.** Though not yet adopted or recognized by GPO, the working group feels this definition speaks to the broader “what, how, and why” of digital deposit as it applies to the work and services of GPO and the FDL community. In the digital deposit scenario, Government publications may now be deposited: 1) *from* GPO to depository libraries; 2) *from* agencies to GPO; and/or 3) shared by libraries with their communities *and deposited to* GPO.

Spring 2019 Conference: Digital Deposit: A Value Proposition

As an introduction to the Depository Library community, the DDWG gave a presentation at the [Spring 2019 Depository Library Council's meeting](#) -- [Digital Deposit: A Value Proposition \(slides\)](#). This session aimed to define digital deposit along with the community, query participants on initial interest in activities related to digital deposit, and illustrate the value of digital deposit and how it fits within the existing framework of the Federal Depository Library Program.

Based on polling and participant chat during the session, the DDWG determined it would be useful to further engage the community about the potential frameworks for digital deposit and to solicit additional feedback about the understanding of this work and the role it will play in

building collections of born-digital Federal Government information. To do this, a follow-on session with informal focus groups was offered at the Fall 2019 Federal Depository Library Conference.

Fall 2019 Conference: Digital Deposit: Collection Development for the 21st Century

Purpose

To build on the Spring 2019 presentation, the DDWG hosted a focus group session at the [Fall 2019 FDL Conference and Depository Library Council Meeting](#). Whereas the Spring 2019 presentation helped to define digital deposit and frame the value of this collection development work, the session aimed to gauge interest from the depository community around the concept of digital deposit and the work that would be involved. This was accomplished by employing informal methodologies with a self-selected group attending the presentation session.

Methodology

The DDWG drafted a [set of eleven \(11\) questions](#) aimed at soliciting information about what digital deposit will achieve, who will engage in this work, and specific needs for this work (technology, staff, training, etc.). Because sessions at the Federal Depository Library Conference are open to all, the focus groups consisted of a random sample of registered conference attendees. Session attendees were asked to self-select into groups of 5 or less. Pre-printed feedback forms with the eleven (11) questions were distributed to attendees. Each question was read aloud followed by 5 minutes of discussion per question within the small groups. At the end of each 5-minute discussion period, participants were invited to offer highlights from their group's discussion. At the end of the session, the feedback forms were collected and the information was transcribed. See [Appendix I](#) for the questions and the transcription.

Findings

Session participants were asked to discuss, record, and report on questions aimed at gaining a better understanding of the community's perception of the role digital deposit will play in building collections and the perceived major participants responsible for the eventual work.

Responses to question 1 ("What is the value of this work... to the community, to users, to others?") touched on and highlighted the value themes presented in the spring conference session. These included fewer unreported publications leading to a more comprehensive National Collection; and, a greater understanding of government activities by the general public. Additional responses highlighted greater access and preservation of government information. Concerns were expressed about the infrastructure and technology needs related

to digital deposit. This was also evidenced in responses to question 2: “Who are the active participants in digital deposit?”. None of the groups included government information librarians as a response. Instead, participants included technologists/systems people, selectors (which may be librarians), GPO, administrators, and citizens/users. Question 7 -- “How will librarians engage with the projects? What type of engagement do you foresee with these or similar projects?” -- was asked to facilitate participants seeing themselves as active players in this work. Instead, the responses took more of a NIMBY (“Not In My Back Yard”)⁹ tone toward the potential work. Several responses noted time constraints; whereas others put the work to student volunteers in larger academic institutions. A few responses did touch on coordination of related activities as well as promotion and advocacy.

Overall, the findings of the 2019 FDL conference focus group were not surprising. The community responses showed an understanding of the value of digital deposit and ways in which developed collections could benefit library communities and the general public. Responses related to the requirements around this work and the involvement of librarians were also not surprising and further evidenced the need for a pilot project, as well as further exploration of library interest in digital deposit.

Digital Deposit Pilot Project

Representatives from the Digital Deposit Working Group shared an early draft of the [*Digital Deposit Working Group report to the Depository Library Council Proposed Pilot*](#) as part of the 2019 focus group session. This document outlined an early concept of three pilot projects that would explore the various aspects and requirements for stakeholders engaged in the work of digital deposit: 1. Lost Documents; 2. Govinfo as an Ingest and Service Hub; and 3. Digital Deposit Dissemination Technology.

After reflection on the scope of the proposed projects and the charge of the DDWG, the proposal was narrowed to a single project focused on reporting lost or unreported publications. This single project took on three main activities — a survey, focus groups, and an unreported publications search and discovery project. Following is information about each of the three parts of the project, including purpose, methodologies, and some of the DDWG findings.

⁹ “Not In My Back Yard” or NIMBY is a colloquial term to express opposition to a proposed project or idea. See <https://www.britannica.com/topic/Not-in-My-Backyard-Phenomenon>

Unreported Publications Survey¹⁰

Purpose

In the print era, models for the creation and dissemination of government information followed a workflow of agencies providing reports or other documents to GPO. In this model, GPO acted as the printer and also cataloged and indexed (C&I) the materials and then, when in scope of the FDLP, distributed them to FDLs. In the print universe of government information, items would occasionally avoid the GPO-centered printing model and would not make it into either the C&I or distribution workflows. As information dissemination moves from print/tangible materials to digital publishing, agencies more frequently publish materials directly to their websites and fail to report new publications to GPO. As a result, the ease of digital publishing has increased not only the amount of information being disseminated,¹¹ but also the amount of information that misses GPO's C&I and distribution workflows.¹² In both the print and digital world, the result is unreported or "lost" documents.

Many government information librarians and other stakeholders are engaged in discovering and reporting publications and documents that may have circumvented the traditional models established by GPO. The DDWG wanted to learn more about who is actively seeking or serendipitously discovering unreported documents and what they do after discovery.

Methodology

To gain a better understanding of who is finding unreported documents and to learn what happens after discovery, the DDWG conducted a 14-question survey, which was announced through FDLP News Alerts and via GOVDOC-L. The survey was launched July 13, 2020 and closed August 7, 2020. Basic demographics questions were asked to identify respondents' organization type and their role as librarian or staff or other. The DDWG hoped to learn if respondents engage in this work deliberately and schedule time in their schedule, or if it is more passive and serendipitous. The survey also aimed to gain a better understanding of what is done with these reports after discovery, whether those found reports were held locally or reported to GPO via askGPO.¹³ Finally, the survey asked respondents to share their contact information if they had interest in participating in other parts of the DDWG pilot project.

¹⁰ This project was originally titled "Fugitive Reporting Survey." The title was changed when, in March 2021, [GPO announced](#) it would no longer use the word "fugitive" to refer to unreported or uncataloged government publications.

¹¹ Between Fall 2016 and Spring 2017, the Internet Archive archived over 200 terabytes of government websites and data. This includes over 100TB of public websites and over 100TB of public data... This includes over 70 million html pages, over 40 million PDFs." <https://blog.archive.org/2017/05/09/over-200-terabytes-of-the-government-web-archived/>.

¹²In April 2021, LSCM began tagging cataloging records for unreported publications in the 922 field to allow for data gathering. From that time to September 30, 2021, GPO cataloged 2,034 unreported publications <https://www.fdlp.gov/collection-tools/reporting-pubs-to-gpo>.

¹³ <https://ask.gpo.gov/s/>.

Findings

The survey was completed by 138 respondents. The respondents were overwhelmingly employed in academic libraries (67.4%) with the majority being government information librarians (69.9%). Respondents were asked how often they actively engage in looking for unreported publications — daily, weekly, monthly, occasionally, rarely, never, or serendipitous discovery. As shown in figure 1 below, only 1 person answered that they search on a daily basis and a close number reported search weekly (4.4%) or monthly (5.1%); 20.6% searched rarely or occasionally; the largest number of respondents (31.6%) said they discover unreported documents serendipitously.

How often do you look for fugitive documents?

136 responses

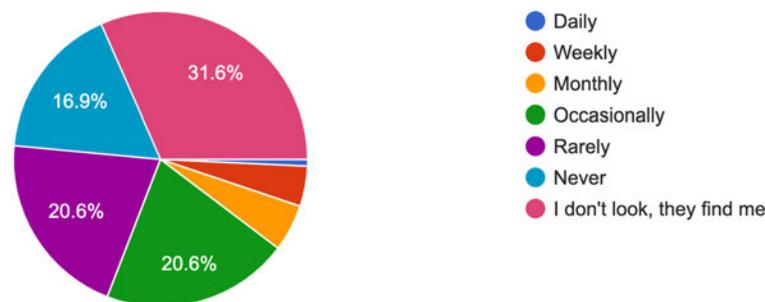


Figure 1: Frequency of searching for unreported publications

The DDWG also hoped to learn about the search process and tools that respondents might use to aid the discovery of unreported documents. Though most answered that they did not use special tools, the tools shared by the 20.2% of respondents who use them offer insights into how this work can be both deliberate and passive. Several mentioned that they utilize search alerts or RSS notifications; others noted searching for reports after reading about them in a news or magazine article. The Wayback Machine from the Internet Archive is also a commonly reported tool for discovering individual reports. There were a number of respondents who reported searching agency websites or utilizing Google or a variety of other finding aids for discovering materials.

The final information the DDWG wanted to gain from the Unreported Documents Survey surrounds individual actions after a document is found. Survey questions asked about capturing digital objects and storing them locally, printing and binding documents to add to tangible collections, and reporting documents to GPO. Only 21.6% said they download a copy of a found document; these downloads are most frequently saved to the individual's computer (50.7%) or a local network drive (24.7%). A few add the document to an institutional repository (12.3%),

some upload the document to the Internet Archive (6.8%), and others add to a LibGuide.¹⁴ Only a few reported that they print found documents for their local collections (7.5%). When asked, “Do you submit fugitive documents to askGPO?” responses were almost equally divided between definitely (38.5%), sometimes (31.1%), and no (30.3%).

Responses from the Unreported Publications Survey offered insight into the work and practices of librarians and others who look for or find unreported materials. Responses also aided in guiding and framing the next two parts of the Digital Deposit Pilot Project: unreported publications focus groups, and the project to search for congressionally mandated reports. The survey instrument and a report of all the responses are in [Appendix II](#).

Focus Groups on Notifying GPO of Unreported Publications

Purpose

To gather feedback about current workflows, project processes, reporting metrics, and project outcomes related to unreported publications reporting through guided discussions with members of the FDL community. The focus group discussions expanded on initial responses gathered with the unreported documents survey and were designed to help identify aspects of document reporting:

1. To assess current Unreported Documents reporting mechanisms;
2. To improve GPO workflows for cataloging and ingest; and
3. To inform the creation of an ecosystem and minimum requirements for “Digital Deposit” including the human-centered and technology-centered components from a user perspective.

Methodology

Three focus group discussion sessions were held in late September¹⁵ and early October 2020, with fourteen (14) volunteers participating. The same five (5) predetermined questions were asked of all the volunteers, and they were asked in the same order in all three sessions. The questions were:

- Describe your ideal fugitive document reporting process
- What functionality or features would you like to see in a fugitive document reporting tool?
- How do you report fugitive documents to GPO?
 - o Follow-up questions: If you are not using askGPO, why not? If you are using askGPO, why?

¹⁴ <https://www.springshare.com/libguides/>.

¹⁵ In August 2020, GPO launched a new platform for askGPO, Salesforce. This is important to note because some of the focus groups volunteers had not yet used askGPO in the new platform. Some of their comments were about the previous iteration of askGPO.

- What are the positive aspects of askGPO? How could it be improved?
- What could GPO do to improve the fugitive reporting process or to make it easier for you?

The focus groups took place when the phrase “fugitive documents” was a long-standing and widely-used phrase in the depository library community for “lost” or uncataloged documents within scope of the FDLP, but were not distributed to depository libraries. In March 2021, Library Services and Content Management (LSCM) announced it would no longer use the phrase “fugitive documents” and it was replaced with “unreported publications.”

During each session, two working group members rotated between one another asking the questions to the group. Other attending members took notes and asked non-predetermined questions as follow-up to certain responses. Working group members leading the discussion called on each participant to individually answer each of the questions; the order that participants were called on was randomized for each question asked. Additional follow-up questions were asked by working group members in response to certain answers and statements provided by participants. The first discussion was not recorded due to technical difficulties, but the second and third discussions were recorded. Notes of discussions were taken by working group members for all three sessions. One person was not able to attend the focus group session. However, she provided responses to each of the questions and submitted them to the DDG via email, and they were incorporated into the session’s notes. Upon completion, recordings were reviewed and notes were compiled from all three sessions. The compilation of notes is in [Appendix III](#). Working group members examined the responses for each predetermined question looking for commonalities, themes, and examples. Brief summaries of the analyses follow.

Findings

REPORTING PROCESS

Several participants focused on reporting tangible publications. While this information was helpful to GPO as it examines current reporting practices, discussions of tangible materials were considered out of scope to the DDWG charge. Some participants had little to no experience with finding and reporting unreported publications; however, there was a general consensus for making the reporting process simple. Most participants expressed a preference for utilizing a form to submit unreported publications; while one participant identified using email as a reporting preference due to the conversation thread created. One participant suggested a form “with required values, such as title, agency, dimensions, number of pages, etc., plus a place to add a scan of the title page.” Multiple participants noted checking the *Catalog of U.S. Government Publications* (CGP) as part of their process but one participant checked OCLC and assumed that documents cataloged in OCLC would not be unreported publications -- this may

or may not be a valid assumption since any library can create OCLC records and those publications may or may not have been reported to GPO and incorporated into the CGP.

FUNCTIONALITY AND FEATURES FOR REPORTING TOOL

Participants expressed a need for guidance on how to report unreported publications. There was interest in a structured form with clearly-defined and labeled required or essential fields. Additionally, several participants expressed the need for the inclusion of flexible or optional fields. Participants noted the tool should include a function that allows photos or images to be attached to a submission. There was also interest in a mechanism to search and find what has already been reported, to lower the number of duplicate submissions. Additionally, participants were interested in a way to track submissions to see where the reported documents were in the review process.

CURRENT REPORTING PRACTICES

Not all participants currently or previously notify GPO of unreported publications; however, several participants said they use askGPO to report found publications and noted they were unfamiliar with other ways to report beyond askGPO. Participants using askGPO submitted unreported publications both individually and in batches. Participants who actively report publications and do not use askGPO offered several reasons for not using askGPO. These reasons included a lack of familiarity with the process of using askGPO; submitting unreported publications directly to known GPO staff; reporting publications to a supervisor or regional librarian; or, handling the publications locally by cataloging and adding to local collections. Several participants noted difficulty finding the askGPO link on fdlp.gov.

In Search of Congressionally Mandated Reports

Purpose

The perceived process for digital deposit follows a three-pronged approach -- GPO to libraries, agencies to GPO, and libraries to GPO. The DDWG conceives the libraries-to-GPO digital deposit flow as one where librarians and other stakeholders push content to GPO for cataloging, indexing, and hosting in govinfo or other GPO digital repositories, with the end result being dissemination back to the depository community. Though not always thought of as digital deposit per se, this process aligns with the values and overarching goals outlined by the DDWG. To date, this work has primarily involved individuals seeking out or serendipitously discovering unreported government publications. The third part of the pilot project, "In Search of Congressionally Mandated Reports," sought to build on the findings from the survey and focus groups by framing a project that informs the search and discovery process for a narrowly-scoped set of government reports.

Methodology

Prior to soliciting project volunteers, the DDWG identified information and data that would inform these primary areas:

1. Time that volunteers spend searching for unreported government publications;
2. Search strategies that deliver the most accurate results;
3. Ways in which GPO can provide education, guidance, and tools to the community to better facilitate systematic searching and reporting of previously unreported publications; and
4. Better understanding of the number of reports that go unreported by agencies to GPO.

Because the universe of government information is vast, the DDWG used [House Document 116-4: Reports to be Made to Congress](#), a report published at the beginning of each Congressional session which includes a “list of reports which it is the duty of any officer or department to make to Congress,”¹⁶ to scope the materials that volunteers would attempt to find. Arranged by agency, H.Doc. 116-4 outlines the reports that are required by law to be provided to Congress during the 116th session. Each H.Doc. 116-4 entry includes the nature of the report, the legislative authority of the report, and a general expectation of when the report should be submitted to Congress. H.Doc 116-4, therefore, offers a list of mandated reports that are in scope of the FDLP and GPO’s cataloging and indexing program and that should be reported to or captured by GPO for dissemination.

To facilitate the work of the volunteers and standardize the reporting of feedback and search strategies, the DDWG developed a work form and proposed a workflow for searching. Because H.Doc 116-4 is only available publicly as a cumbersome PDF, a member of the DDWG requested and received a machine-readable CSV file of the document from the House Clerk’s office.

The DDWG divided 20 agencies, executive branch offices, independent commissions, and federally chartered private corporations among the 22 volunteers. There was overlap of agencies among volunteers to allow for some comparisons and to help ensure that all categories were searched. Volunteers received an Excel file with their assigned agency reports, a sample of how the form should be completed, and a proposed workflow (see [Appendix IV](#)). Volunteers were also invited to participate in an hour-long orientation with DDWG members to outline the workflow, explain desired outcomes, and answer any questions. After the orientation, DDWG held semi-regular online “office hours” to address questions or concerns from volunteers. This also offered volunteers the opportunity to share strategies and express

¹⁶ House Rule II, clause 2(b) requires the Clerk of the House to produce, for each session of Congress, a "list of reports which it is the duty of any officer or Department to make to Congress." For an example of a report in this series, see U.S. Congress, *Reports to be Made to Congress*, prepared by The Clerk, U.S. House of Representatives, 116th Cong., 2nd sess., H.Doc. 116-85 (Washington: GPO, 2020), available at <https://www.govinfo.gov/content/pkg/CDOC-116hdoc85/pdf/CDOC-116hdoc85.pdf>.

challenges. Volunteers were asked to contribute what time their schedules would allow for the project between October 2020 and January 2021.

Findings

At the end of the pilot project, volunteers contributed a total of 214 hours to searching for 1,049 publications, this is a little less than half of the total reports in H.Doc. 116-4. Of those searched, over half (531) are still unreported, 430 were found and have been reported to GPO for cataloging and indexing, and 88 were in the CGP. Though the goal was to scope a project that would facilitate more deliberate discovery, volunteers serendipitously discovered 231 unreported publications outside of H.Doc. 116-4 while only finding 199 that were listed in H.Doc 116-4.

Volunteers focused the majority of their searches on the Dept. of Transportation (438 searches) and the Dept. of Energy (151 searches). This was in part due to how the agencies were distributed, the interests of each volunteer, and the amount of time spent searching by volunteers assigned to these agencies. The majority of searches (522) were initiated at an agency's website with the next largest group starting at Google (over 271). Several used USA.gov as a starting place, whereas others tried the CGP or OCLC. The DDWG included open text fields for "Search Strategy" and "Tools Utilized" on the work form with the expressed hope of discovering successful strategies or learning more about individual tools volunteers might use to facilitate discovery. None of the volunteers reported using tools outside of those previously mentioned (CGP, Google, Agency website). Many volunteers used "report to congress" as part of their search strategy or other keywords from the report's description included in H.Doc. 116-4. Volunteers shared challenges associated with individual searches. Many reported that they struggled to clearly determine if what was found was the same report mandated by H.Doc. 116-4. Others expressed excitement when an item was easily discovered at Google, while others reported finding nothing titled or designated as "report to congress." The findings of the pilot are informative but tend to reflect what is generally known about the search for "lost docs" – it is time-consuming or "stumblebum luck," as one volunteer noted.

Visualizing the Deposit of Digital Content from GPO

The Visualizing Digital Deposit Subgroup (VDDS) was tasked with exploring or "visualizing" what the deposit of digital content by GPO to Federal depository libraries might look like. VDDS approached this charge by identifying use cases, developing high-level questions, and conducting interviews to garner in-depth feedback. Questions were asked of librarians representing each of the identified use cases to bring to fruition the concept of digital deposit *from* GPO to Federal depository libraries.

The ten use cases were represented by different types and sizes of depository libraries and non-library organizations working in the digital preservation and access space:

- ASERL Center of Excellence (University of Georgia).¹⁷
- Collaborative organization whose members are primarily Federal depository libraries (HathiTrust).
- Digital preservation steward (University of North Texas).
- Organization that pulls content from **govinfo** via API and makes it accessible (Legal Information Institute (LII) at Cornell University).
- Preservation Steward (University of Colorado Boulder).
- Regional depository library (Illinois State Library).
- Selective depository academic library (University of California, Berkeley).
- Selective depository law library (Stanford Law School Library).
- Selective depository public library (Sacramento Public Library).
- Selective depository Tribal college library (Aaniiih Nakoda College).

There were fourteen interviewees representing the ten business cases. And there were at least three working group members in attendance at the interviews. The open-ended interview questions were designed to elicit detailed responses regarding interviewees' thoughts surrounding the feasibility of digital deposit. Questions were left intentionally nonspecific, with minimal definitions, clarification, or contextualization provided, in an effort to avoid influencing the responses. In the case of interviewees that were not members of the FDLP (HathiTrust and LII), some of the questions were modified slightly to make them relevant to that particular organization. The question-by-question **Findings & Analysis** can be found in the full report of the VDDS, *Visualizing Deposit of Digital Content from the Government Publishing Office*, in [Appendix V](#).

Overall Observations

DIGITAL DEPOSIT IS NOT JUST FOR ONE KIND OF LIBRARY

The interviews were conducted during the COVID-19 pandemic lockdown phase, and the real-life experience of remote library service provision most likely had some bearing on perspectives related to digital content. However, in our interviews with ten different organizations, each with their own institutional considerations and use cases for digital deposit, they all expressed interest and had ideas about how they would interact with a service. Additionally, most of the libraries indicated that they had access to some form of resources that would enable them to participate in a digital deposit service, whether those resources were local, or at a regional or consortial level.

¹⁷ ASERL is the Association of Southeastern Research Libraries. <http://www.aserl.org>

FLEXIBLE SELECTION IS IMPORTANT TO LIBRARIES

Based on this sample of libraries, a one size fits all digital deposit approach would not serve libraries' needs. Libraries expressed a desire to select digital materials to fit their own collection priorities, in a variety of ways such as topic, provenance, geographic range, format or other levels of granularity or time variables.

BIBLIOGRAPHIC DATA IS A REQUIREMENT FOR DIGITAL DEPOSIT, AND PHYSICAL FORMATS ARE NOT

The interviews provided an especially strong indication of specific requirements for two particular aspects of digital deposit. Responses showed that libraries view bibliographic data as inherent to digital deposit. We also heard strong opinions from our interviewees that they did not want to receive digital files on physical media.

A COORDINATED DIGITAL DEPOSIT SERVICE SHOULD BE IN HARMONY WITH EXISTING FDLP MODEL(S)

Based on the interviews, libraries see benefit in digital deposit being offered as a full program or service that includes not just deposit mechanisms but also coordination, documentation, training, and communications support from GPO. Libraries also expressed a desire for an understanding of the dynamic between digital deposit requirements and existing FDLP requirements to be incorporated as requirements are developed.

DIGITAL DEPOSIT COULD ENABLE NEW USE CASES FOR FEDERAL DOCUMENTS

Against the backdrop of the pandemic, the ability of a library to provide digital access is in the forefront of people's minds. In our interviews, discussions about expanding digital access stood out as an area where most libraries were looking ahead. The fact that a number of libraries were looking beyond traditional library catalog discovery, and brought up concepts of access such as computational access, seem to indicate a need to not "silo" Federal documents in bringing them to users.

Conclusion

The activities of the Digital Deposit Working Group evidence a continued interest in digital deposit paired with a continued lack of understanding of what digital deposit is or how it could work. Those that get it, get-it; whereas, those that don't fully understand digital deposit do see the value but struggle to see their role in the process. This can be overcome through additional presentations, instructional sessions, and continued discussions with the FDLP community.

The survey, focus groups, and pilot project show dedication within the FDLP community to assist GPO with building the National Collection of U.S. Government Public Information. The outcomes from these efforts also highlight the number of new depository coordinators who are not familiar with the concepts or work associated with searching for unreported publications. Participants expressed interest in educational opportunities to learn more about unreported

publications, the search/discovery process, and ways to report these to GPO for inclusion in the National Collection.

The Visualizing Digital Deposit subcommittee engaged with libraries of varying types that see the value of digital deposit and who are interested in exploring opportunities to build local collections of born-digital government information. Feedback from participating libraries mirror findings from other work of the Digital Deposit Working Group, including a strong understanding of the value in this work and the necessity to build the National Collection, and grow opportunities for access and preservation of electronic government information. Participants also expressed concerns about how this work would be implemented in their own libraries. Despite the concerns, several participants stated their willingness and enthusiasm to work with GPO to pilot digital deposit workflows to develop guidance and processes for building locally-hosted digital collections facilitated by digital deposit.

The members of the Digital Deposit Working Group thank GPO and the DLC for supporting the group's work over the last two years. In order to build on the work of the DDWG, we offer the following recommendations to the Depository Library Council.

Final Recommendations to the Depository

Library Council

Recommendation #1

The Digital Deposit Working Group recommends GPO adopt the following definition of digital deposit: The practices, services, and workflows for the collaborative acquisition of born-digital and digitized Federal Government information for the National Collection of U.S. Government Public Information.

Recommendation #2

As a model to inform depository libraries of workflows needed to implement the distribution of digital content, the Digital Deposit Working Group recommends GPO scope and implement a pilot project that explores the GPO-to-Library approach to digital deposit. GPO should undertake a pilot with one or a small number of libraries to work out the details of a digital deposit service, in particular a service where GPO facilitates the delivery of digital files to FDLP libraries. The pilot would:

- Use the findings from this group's interviews toward the project.
- Explore delivery issues (push vs pull, file types, etc.) that were brought to light by this group's interviews.
- Draft a high-level workflow for notification and push/pull delivery mechanisms.

- Document the necessary training requirements to support delivery mechanisms.
- Identify opportunities for technical solutions or enhancements to support delivery mechanisms.
- Report back to the Depository Library Council on the findings and outcomes of the processes explored in the pilot project.

Recommendation #3

GPO should take steps to outline a vision of how policy, coordination, and support for digital deposit could fit in with current and future models of the FDLP.

Recommendation #4

The Final Report of the Digital Deposit Working Group should be presented to the recently announced Task Force on a Digital FDLP.

Recommendation #5

To address knowledge gaps related to digital deposit, notifying GPO of unreported publications, and to better develop the search skills of government information professionals, the Digital Deposit Working Group recommends the following FDLP Academy sessions:

- Introduction to Digital Deposit.
- Reporting the Uncataloged, which would introduce the concept of uncataloged government publications and the processes for discovery and reporting.
- Explaining APIs and how FDLs can utilize govinfo.gov APIs for collection development.
- Exploring more effective searching within the *Catalog of Government Publications* (CGP).

Appendix I: Digital Deposit: Collection Development for the 21st Century

- Presentation slides
- Digital Deposit Breakout Group Discussion Questions
- Transcribed Breakout Group Responses to Questions



FEDERAL
DEPOSITORY LIBRARY

CONFERENCE

#FDLConference

2019

OCTOBER 21-23
ARLINGTON, VA



Digital Deposit

collection development for the 21st century



Working Group Participants

- Robbie Sittel – Depository Library Council Representative
- Jessica Tieman – Government Publishing Office
- Lisa LaPlant – Government Publishing Office
- Cindy Etkin – Government Publishing Office
- Heather Christenson – HathiTrust
- James Jacobs – Stanford University

Other Participants

- Tom Fischlschweiger
- Kathy Hale
- Cass Hartnett
- Aimee Quinn



Council Recommendation

Council recommends the creation of a working group to explore current and future needs related to digital deposit - both dissemination of content and acceptance of content by GPO. At a minimum, two appropriate members of GPO staff, two members of DLC, and two members of the FDLP community should be appointed to serve on the Digital Deposit Working Group for one year. Composition of the working group should be chosen by DLC in consultation with GPO staff. The Working Group should report findings and recommendations - either initial or final - at the Fall 2019 FDLP annual meeting.

<https://www.fdlp.gov/news-and-events/3911-recommendations-commendations-from-the-dlc-to-gpo>

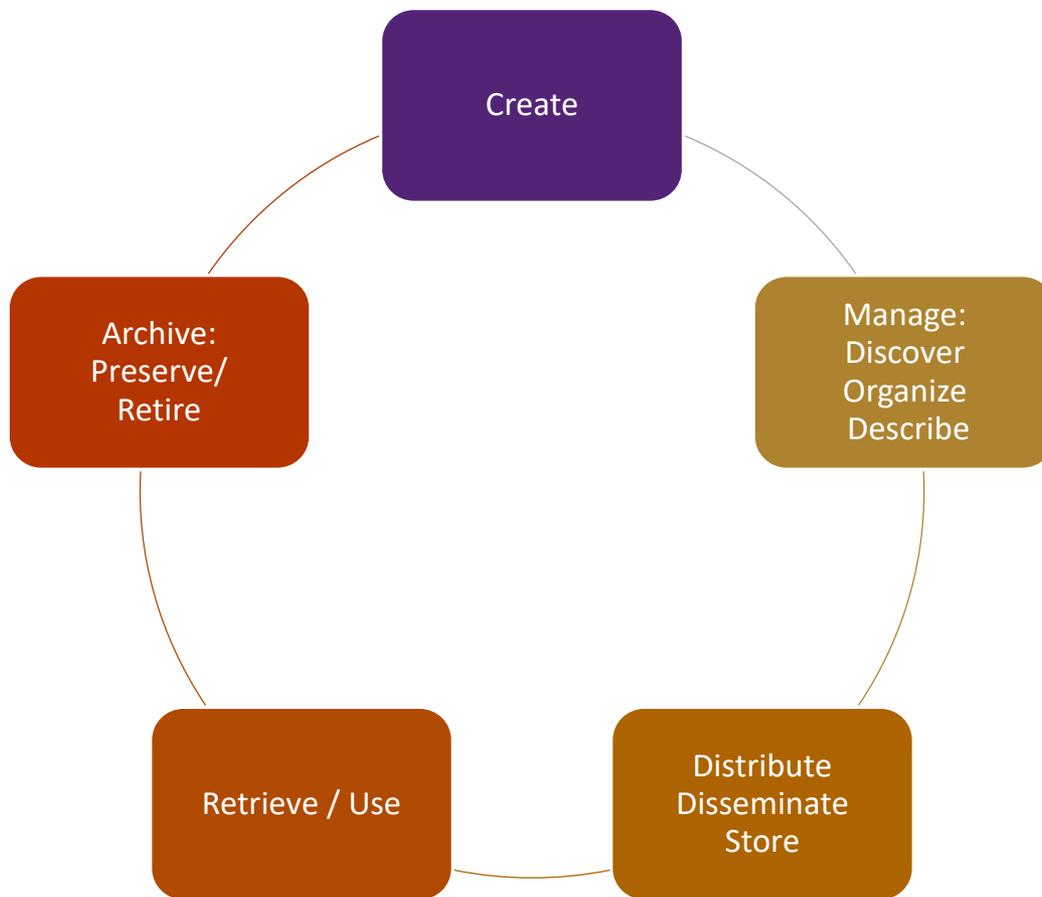


Digital Deposit Defined

Practices, services, and workflows for the collaborative acquisition of born-digital Federal Government information for our National Collection of U.S. Government Public Information.



Information Life Cycle Management





Three Proposed Projects

1. **“Lost Docs” Project**

This project aims to report lost documents to GPO, which will then be ingested, cataloged, and made available through **govinfo.gov** and the Catalog of U.S. Government Publications.

2. **Agency Submission to GPO**

An agency-driven “digital deposit” implementation model.

3. **Digital Deposit Dissemination Tool**

This project will use API technology to “push” content to FDLs.



FEDERAL
DEPOSITORY LIBRARY

CONFERENCE

#FDLConference

2019

OCTOBER 21-23
ARLINGTON, VA



Let's Discuss Digital Deposit



Discussion Questions – ALL GROUPS ANSWER QUESTIONS 1 & 11

Other questions will be assigned to breakout groups.

1. What is the value of this work... to the community, to users, to others?
2. Who are the active participants in digital deposit?
3. How do you see the role of a Federal depository library in Digital Deposit?
4. What does acquisition of digital-born Government information look like? If you ran the world what would you do?
5. How will these projects fill a need in the Federal Depository Library Program?
7. What risks do you see in these projects?
8. How will librarians engage with the projects? What type of engagement do you foresee with these or similar projects?
9. What kind of support or tools do you envision librarians needing in order to participate?
10. How will local patrons use the content?
11. What are your thoughts about the authenticity of publications, chain of custody within the framework of digital deposit, and the authorization of various actors?



We want your feedback!

If you did not have the opportunity to attend this session (or if you have more thoughts to share), the Digital Deposit Working Group invites you to provide your views on any or all of these questions. You can submit them to the Working Group through the Depository Library Council's contact form on FDLP.gov at <https://www.fdlp.gov/dlc-contact-form>. You also may wish to consult the program slides for information about the projects: <https://www.fdlp.gov/events-and-conferences/2019-federal-depository-library-conference>.



Digital Deposit: Collection Development for the 21st Century

Depository Library Council Session
Digital Deposit Working Group
Tuesday, October 22, 2019

Digital Deposit Breakout Discussion Questions

ALL GROUPS ANSWER QUESTIONS 1 & 11.

Other questions will be assigned to breakout groups.

1. What is the value of this work... to the community, to users, to others?
2. Who are the active participants in digital deposit?
3. How do you see the role of a Federal depository library in Digital Deposit?
4. What does acquisition of digital-born government information look like? If you ran the world what would you do?
5. How will these projects fill a need in the Federal Depository Library Program?
6. What risks do you see in these projects?
7. How will librarians engage with the projects? What type of engagement do you foresee with these or similar projects?
8. What kind of support or tools do you envision librarians needing in order to participate?
9. How will local patrons use the content?
10. What are your thoughts about the authenticity of publications, chain of custody within the framework of digital deposit, and the authorization of various actors?
11. Do you have ideas/suggestions for other projects?

NOTE: If you did not have the opportunity to attend this session (or if you have more thoughts to share), the Digital Deposit Working Group invites you to provide your views on any or all of these questions. You can submit them to the Working Group through the Depository Library Council's contact form on FDLP.gov at <https://www.fdlp.gov/dlc-contact-form>. You also may wish to consult the program slides for information about the projects: <https://www.fdlp.gov/events-and-conferences/2019-federal-depository-library-conference>.

Q1. What is the value of this work... to the community, to users, to others?	Q2. Who are the active participants in digital deposit	Q3. How do you see the role of an FDL in digital deposit?	Q4. What does acquisition of digital-born government information look like?	Q5. How will these projects fill a need in the FDLP?	Q6. What risks do you see in these projects?
fewer fugitives	selector	overseer	right click save as, drops into workflow	benefit of participation (capacity building)	authenticity
preservation of materials that agencies may remove from their websites	technologists	provide policies if FDLs plan an active role in deposit and catalog and providing access	(or, what about an approval plan)	secures collection locally in case of need	value of content
more accessible to the public	users	all can provide access to deposited content	select by agency, format, topic	skill development fo digital preservation	authenticity of items
preservation	metadata people	not all will want to collect, not all can	NO ITEM NUMBERS please	increase diversity of collections	authenticity/authority
96% of new publications are online	systems people	whether or not they can accept deposit, they could contribute	capturing and archiving documents and providing indexing and metadata to mak it searchable and accessible	help to make more accessible a comprehensive national collection and record	data loss
comprehensive	GPO	leading with standards for partners	central person everthing would go through		destruction of tangible materials
shed light on government activities	administrators	organize the players into different projects / agency exploration	alerts with new content		lots of time planning and then nothing happens
preservation	open ended	state higher priority than federal	centralization		GPO changes priority
more accessible to the public	open gov't / crowdsourcing	lost docs project is a great suggestion, contribute to this with pubs	"archival sliver" - what to capture (like a records schedule)		new leginstation or "interpretation" of title 44
public access	citizens	state docs program experience can share best practices, issues	automatic alerts when new publications on agency website - with metadata harvesting and comparing		hoarding is fun but runs up local tech costs
capacity building (local)	agencies	federal efforts could lead into development of state efforts	would SuDoc set a priority list		climate change
fits library missions	open - please contribute	(FDLs can often see state docs as priority)	a bot to detect publications posted on agency websites		GPO
builds relationships	PQ paid (are they making money on the govt items thy digitize)	do other document pulls as was done in the 2006 EPA document pulls	priorities		ensuring authenticity
we used to act like islands but no longer a good idea in our current economic model	Citizen Archivist	GPO could pull docs by agency and FDLs could catalog them or make recommendations	publication equivalent to record schedule		criteria for inclusion
national comprehensive collection	HathiTrust	interested FDLs can sign up as GPO partners for wahtever level of involvement they are interested in/can manage			ensuring the authenticity of the documents
history of country	Internet Archive				inclusion
revealing the existence of an unknown doc	agencies				valuable
prevent loss and provide access long term	IT				If libraries host deposited data could their servers be overwhelmed
coordinated approach, more consistent	issuing agency				
leverage resources and share costs	depository libraries - target an agency look for fugitive documents - check against CGP - harvest teams at each institution				
shared user interface	(multidisciplinary)				

Q1. What is the value of this work... to the community, to users, to others?	Q2. Who are the active participants in digital deposit	Q3. How do you see the role of an FDL in digital deposit?	Q4. What does acquisition of digital-born government information look like?	Q5. How will these projects fill a need in the FDLP?	Q6. What risks do you see in these projects?
prioritized agencies to collect from and make sure others covered if this information is not captured it's lost get creators to distribute to GPO	GPO Agencies				
access	FDLs within an organization, need information professional, IT person, project management person, etc.				
bibliographic control automated metadata? preservation					
distributed, outside of federal control					
authenticity/chain of custody to prevent loss an provide long term access to everyone more coordinated effort brings more value, make more headway take advantage of shared or leveraged resources					
if automate, more consistent over time, forever a portion uniform interface could be helpful by coordinating, decide where priorities lie, at least in short term can be more systematic, could assist GPO workflow					
reduce risk meeting the user where they are / preferred format					
preventing loss retraining agencies to submit to GPO - need a contact at the agency + institutional knowledge					
simpler = more effective - automation also good catalog fugitive paper docs as well as born-digital ones					
access to documents as preservation and keeping document available over time					
identifying documents meet users where they are reduce fugitive publications					

Q7. How will librarians engage with the projects? What type of engagement do you foresee with these or similar projects?	Q8. What kind of support do you envision librarians needing to participate?	Q9. How will local patrons use the content?	Q10. What are your thoughts about authenticity of publications, chain of custody withint the framework of digital deposit, and authorization of various actors?	Q11. Do you have ideas/suggestions for other projects?
time constraints	staff time	local repository?	GPO is the trusted party	have a data repository so data sets can be accessible (back up of data.gov) with better metadata/indexin/cataloging
academic public univ student volunteers	server space	local library federated search	approve FDLs to accept deposits	serials archived by GPO
library MLS students - an actual hands on course	admin support	local library catalog	users don't usually care who hosts this, they care about finding	figure out how to package together things that belong together (ata documentation or serials/monos)
planning, coordinating, promoting	liaisons to agencies	open hub search	NO BLOCKCHAIN please	curation and quality contron workflows
local advocacy	\$\$\$\$	PURLS or PDFs	very important and must be maintained	finding ways to engage
learn digital preservation and tech skills	tech skills - enable digital preservation and information architecture	CGP	for authenticity both at ingest and being kept long term to avoid both loss and manipulation	automation and taking web archives extracting items to catalog
take ownership in the FDLP!!!!	community of practice	local catalog records, local discovery layer?	provenance is critical	individual documents - (unreadable) and make discoverable
fall to libraries with staffing and time	tool for content identification / submission other than lost docs	local repository storage?	provenance - a part of the record (ex. article donated by USMC command, #28 from his person collection)	record searches - what's missing from current catalog record
promoting GPO - early adopters	guidelines and standards for archival purposes	research	we do get this question about how we obtain items	regional collecting focus on materials of interest in region (e.g. calif related docs in calif)
new coordinator don't have much time, a lot of things to learn.	a platform or program for digitizing and injest	research	CRITICAL	digitization data website
depends on staffing	internal administration support	civic engagement	minimize # of steps something travels through	each region could focus on local area federal pubs and capture a lot of lost docs, in each region's interest, would help all libraries participate in a shift in thinking about depository collections but still in line with priorities, documents could feed into regional collection from any libraries (more people helping w/o smaller libraries absorbing workload) even if materials stored elsewhere - Team: 1 large selective, 3 regionals, 2 GPO staff
	money - staff time, equipment, training and hiring, software and hardware	knowledge about agencies		central place for agencies to report their digitization plans
	focus on agency outreach	specific local stake holders		digitization plans database - agencies and libraries can list what then plan to digitize
	staff time	provenance is critical		
	training	View, download		
	hiring	research/read		
	software	analyze		
	librarians serving as liaisons to certain agencies	text mine		
	training	access during govt shutdowns		
		having a single search (CGP) rather than digging through agency websites		

Appendix II: Unreported Documents Survey

- Unreported Documents Survey Instrument
- Survey Responses Report

Fugitive Reporting Survey

July 13 – August 7, 2020

Tell us about you and your experience searching for fugitives or lost docs.

For the purpose of the Depository Library Council (DLC) approved Digital Deposit Pilot, a fugitive or lost document is defined as a document in scope of the Government Publishing Office (GPO) Federal Depository Library Program (FDLP) or Cataloging and Indexing Program (C&I) and is not currently discoverable via the Catalog of Government Publications (CGP).

****Please ask anyone in your institution who helps with fugitive docs hunting to complete this survey.****

1. Where do you work?

Mark only one oval.

- Academic Library
- Public Library
- Law Library
- Special Library
- Other: _____

2. What is your role / position?

Mark only one oval.

- Government Documents/Information Librarian (some or all of the time)
- Non-Government Documents (all other) Librarian
- Non-Librarian/Full or Part Time Staff
- Student Assistant
- Other: _____

3. How often do you look for fugitive documents?

Mark only one oval.

- Daily
- Weekly
- Monthly
- Occasionally
- Rarely
- Never
- I don't look, they find me

4. Do you schedule time as part of your regular workflow to search for fugitive documents?

Mark only one oval.

- Yes
- No

5. If yes, how much time do you allocate? (example: hours per week)

6. If you look for fugitive documents, do you use any special tools or search techniques?

Mark only one oval.

- Yes
- No

7. If you answered "yes" to the previous question, please tell us more about the tools and techniques you use for searching. (example: RSS, browser plugin, search alerts, etc).

8. If you learn of a fugitive document and search for it, but don't initially locate it, do you set reminders to continue looking?

Mark only one oval.

- Yes
- No
- Sometimes
- Other: _____

9. Do you download a copy of a found fugitive?

Mark only one oval.

- Yes
- No
- Sometimes

10. If yes, where do you download or save these items? (Check all that apply.)

Check all that apply.

- Institutional repository
- Internet Archive
- Web space
- LibGuide
- Intranet
- Cloud based storage
- Local network
- My computer's hard drive
- Off-board storage (e.g. thumb drive/USB,etc)
- Do not save

Other: _____

11. Are the fugitive documents you download or save accessible to the public in digital form?

Mark only one oval.

- Yes
- No
- Sometimes

12. Do you print fugitive documents to include a physical copy in your local collection?

Mark only one oval.

- Yes
- No
- Sometimes

13. Do you submit fugitive documents to askGPO?

Mark only one oval.

- Definitely!!
- No
- Sometimes, when I remember to

14. If you submit a fugitive to askGPO, do you watch the CGP to see that it's added to the system?

Mark only one oval.

- Yes! I wait with bated breath
- No, I trust GPO to get the job done
- If I remember to look, I do
- No, I catalog it for inclusion in our local catalog

15. Do you ever receive notification from GPO that your submission is out of scope?

Mark only one oval.

- Yes
- No

Want to volunteer?

If you would like to volunteer to participate in the Digital Deposit Working Group Pilot, please complete the following:
(we will be asking for focus group participants and fugitive docs hunters)

16. Name:

17. Email:

Unreported Publications Survey

Report of Responses

Digital Deposit Working Group (DDWG)
Depository Library Council (DLC)
February 2022

About the Unreported Publications Survey

The Digital Deposit Working Group launched this survey to gain a better understanding of who is finding unreported publications, how they find them, and to learn what happens after discovery of them. The DDWG also hoped to learn if respondents engage in this work deliberately and schedule time in their schedule, or if it is more passive and serendipitous.

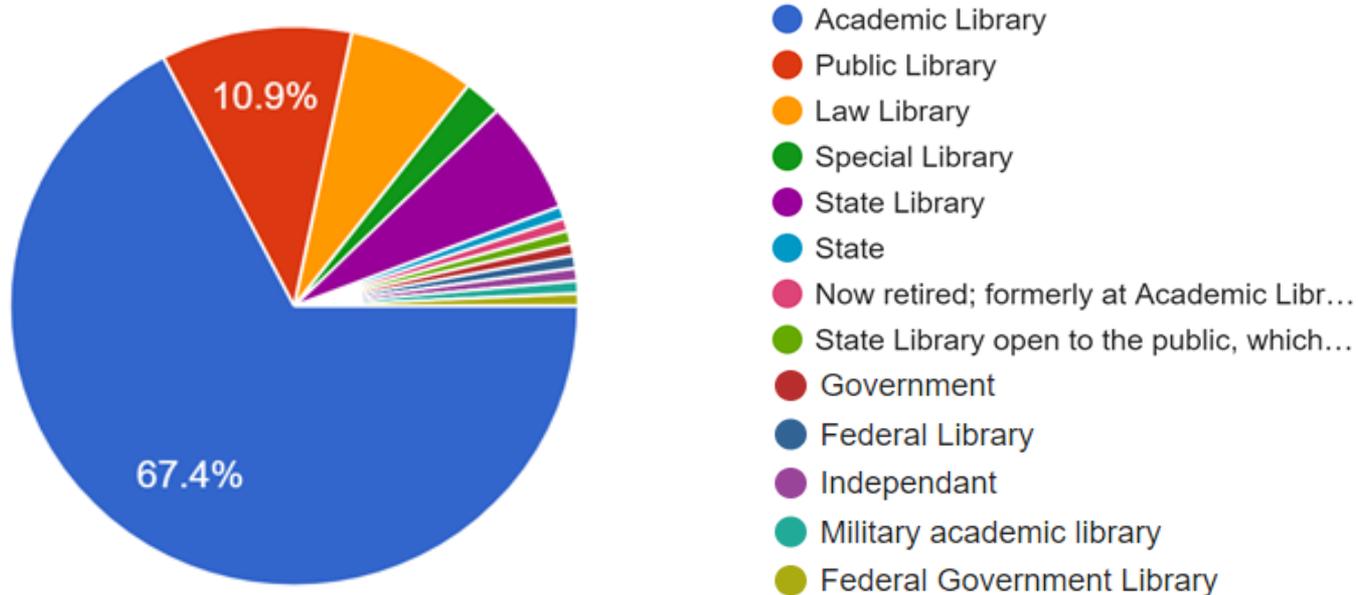
For the purpose of the DLC -approved Digital Deposit Pilot, including this survey, a fugitive or lost document was defined as a document in scope of the Superintendent of Documents' Federal Depository Library Program or Cataloging and Indexing Program and is not currently discoverable via the *Catalog of U.S. Government Publications* (CGP).

The survey was in the field from July 13 – August 7, 2020, and 138 submissions were received.

Please note that in March 2021, the Superintendent of Documents announced the replacement of the phrase “fugitive documents” with “unreported publications.” This was after the survey was conducted, thus the varying language in use with the survey and in the writing of this report.

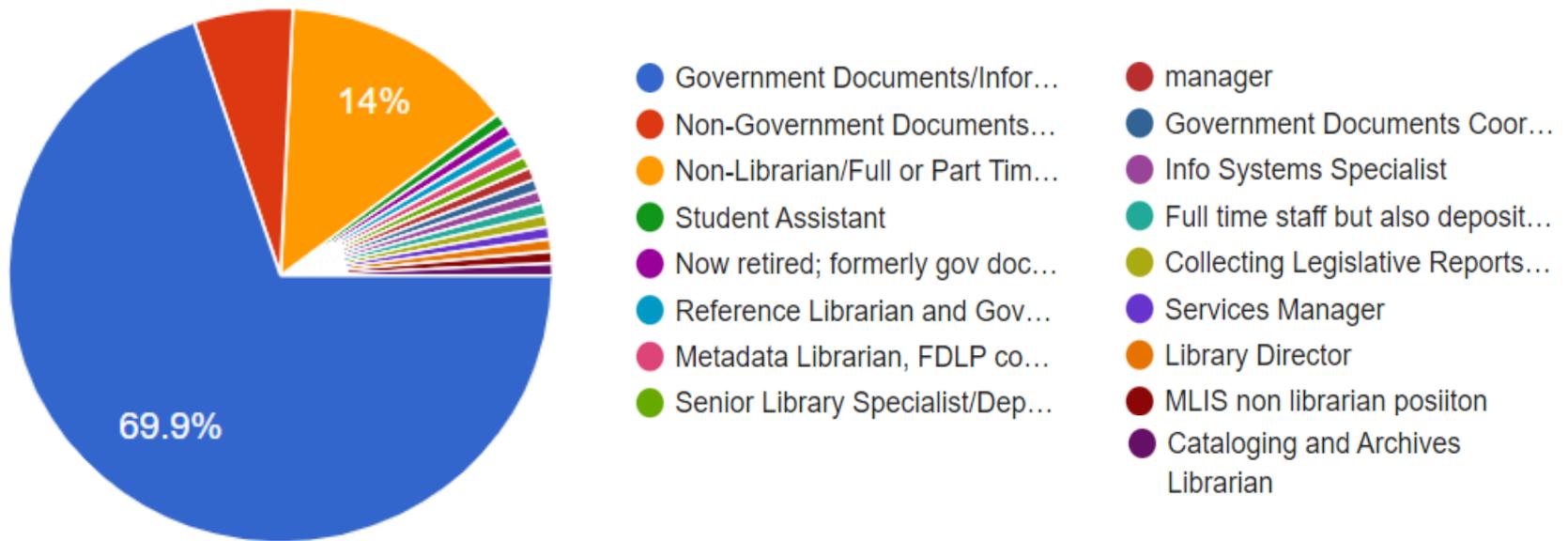
Where do you work?

138 Responses



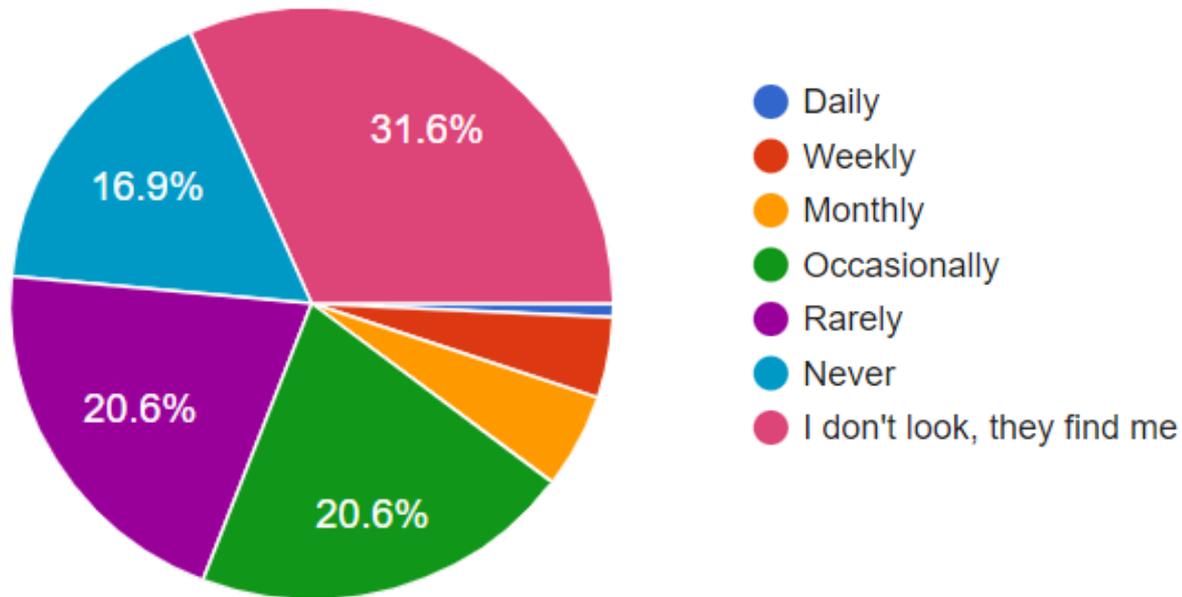
What is your role / position?

136 Responses



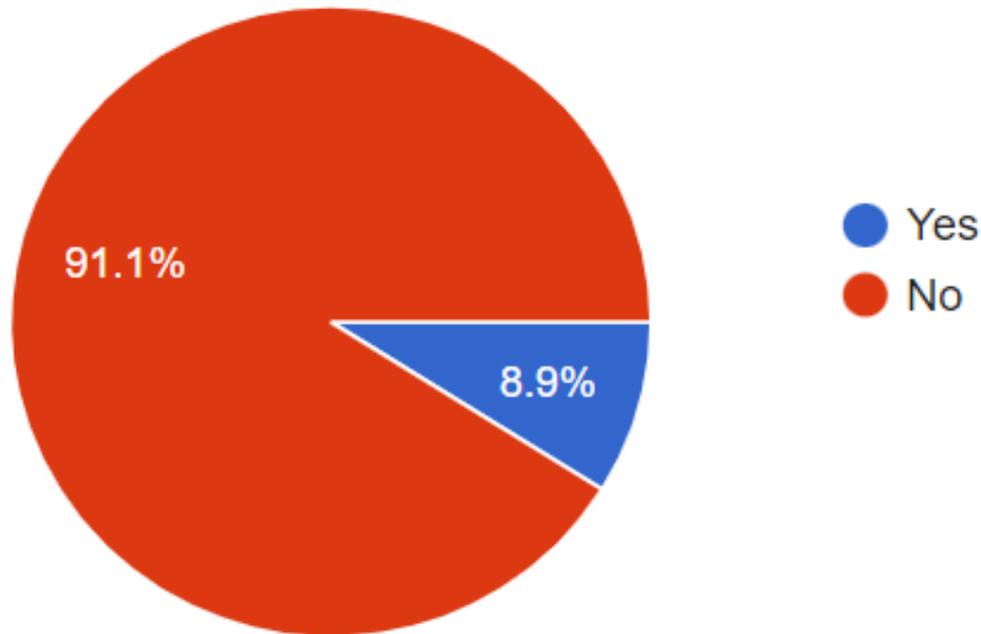
How often do you look for fugitive documents?

136 Responses



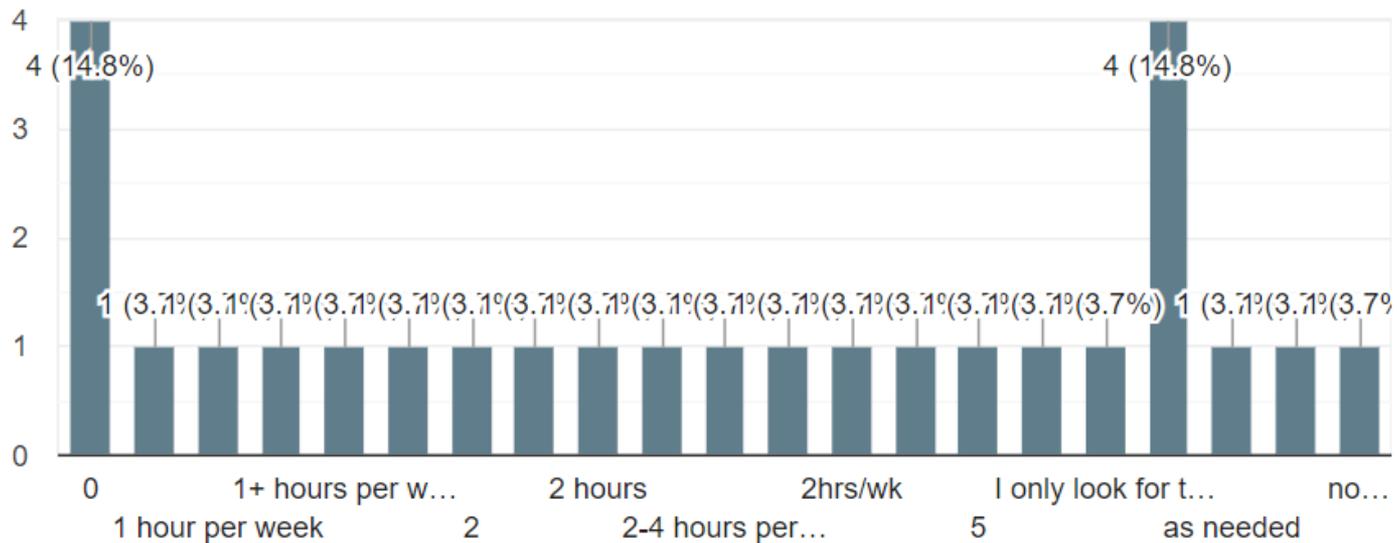
Do you schedule time as part of your regular workflow to search for fugitive documents?

135 Responses



If yes, how much time do you allocate? (example: hours per week)

27 Responses



If you answered "yes" to the previous question, please tell us more about the tools and techniques you use for searching.

(e.g., RSS, browser plugin, search alerts, etc.)

23 Responses

SEARCH TOOLS AND TECHNIQUES
N/A
RSS
Created list of targeted agency websites and checked regularly for new additions of documents, especially those referenced in newspapers like Washington Post and NY Times. Added to weblog with commercial software. Item was removed from library website after my retirement.
OCLC/Connexion, other library websites/catalogs, CGP, email colleagues at State Library or other academic libraries in my state
search alerts, twitter
Wayback machine, library of congress archived websites, search alerts, contacting CRS
Internet Archive's Wayback Machine to search older versions of websites
I usually look for fugitive docs in context of searching local land use history. I sometimes have success finding them in library special collections or unofficial collections compiled by academic researchers
RSS
References from newspaper articles or tv news.
mostly newspapers but also tv news
site searching site:xxx.gov; Wayback Machine

If you answered "yes" to the previous question, please tell us more about the tools and techniques you use for searching.

23 Responses

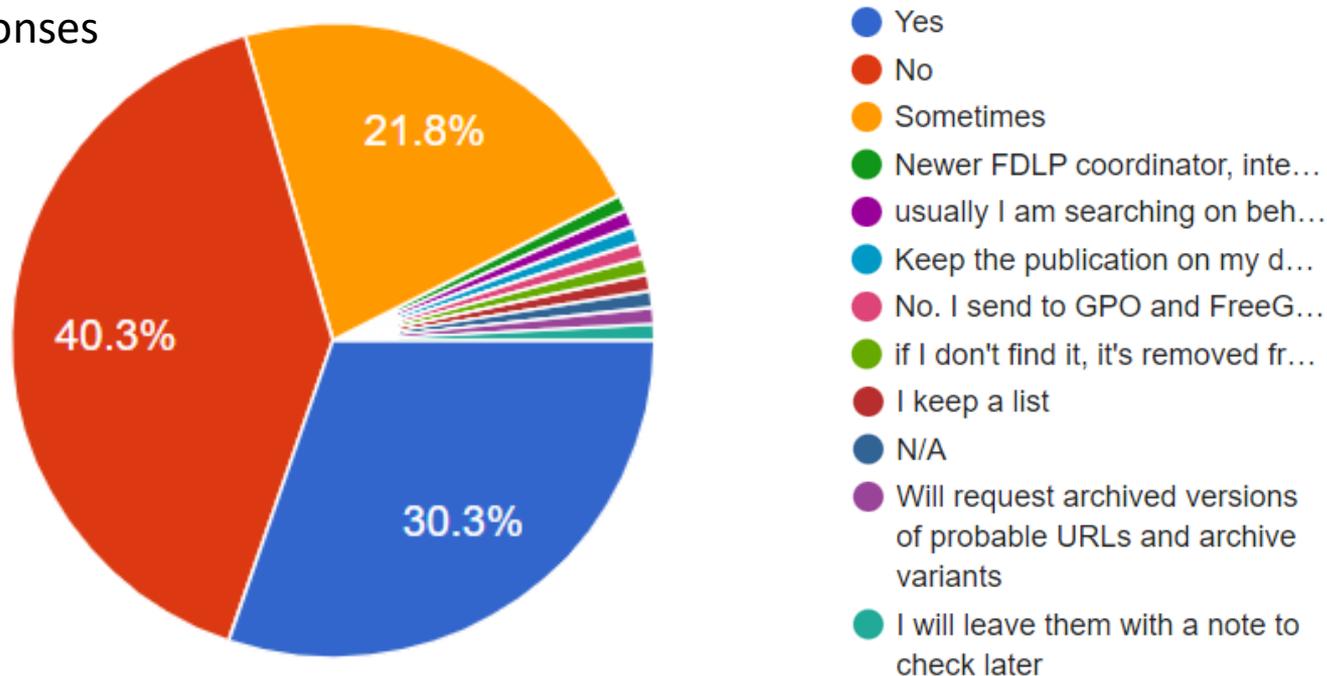
(e.g., RSS, browser plugin, search alerts, etc.)

(continued)

SEARCH TOOLS AND TECHNIQUES
Check for references to the type of information in sources from the time period. Sometimes the actual name (s) of the fugitive info are also lost in the mist. <u>So</u> it can be helpful to determine whether or not I have an accurate description of the desired document.
Combing agency websites for specific bio-regional documents
Search alerts from various journals and agency sites
Occasionally search agency publication pages
I bookmarked several databases (NTIS, NASA, Energy Research abstracts) and made my own technical report search guide.
search alerts, go through agency emails about new reports to determine if they are in scope and if they have been included
Extensive PDF lists
My favorite tool for finding these items is the Executive Communications section of the Congressional Record. Other tools include WorldCat and HathiTrust
Google search, HathiTrust, Internet Archive, archived websites, federal agency websites, etc.
RSS--used to be a decent source but that entire tool seems to have faded away. I religiously visit about 7 websites monthly and purposefully look. Sometimes I see something on Reddit Politics.

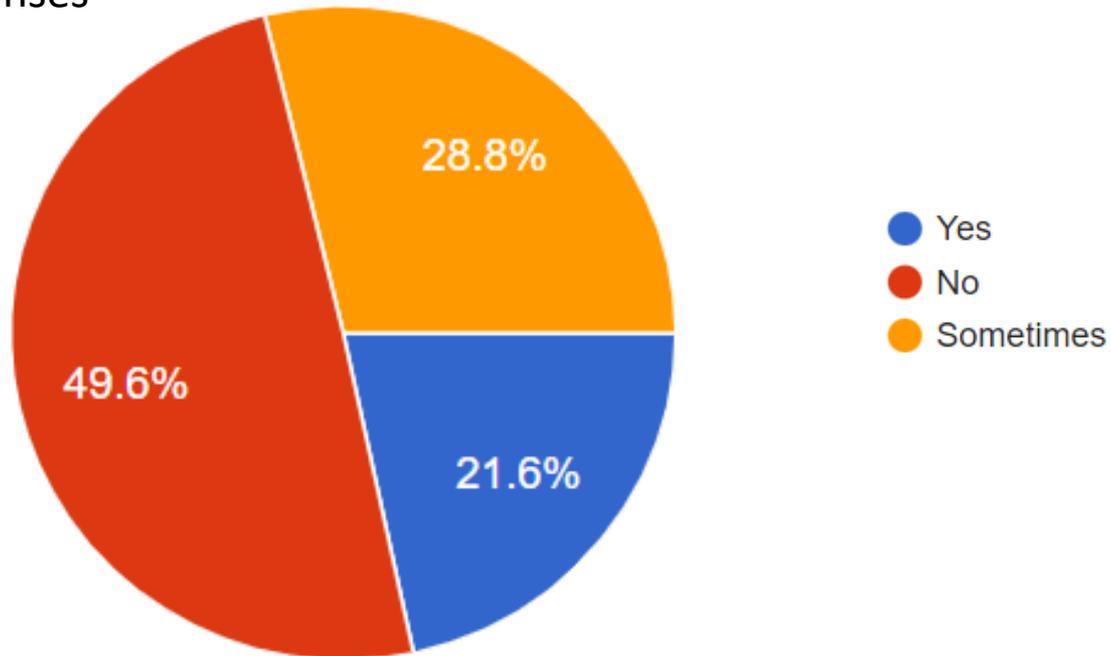
If you learn of a fugitive document and search for it, but don't initially locate it, do you set reminders to continue looking?

119 Responses



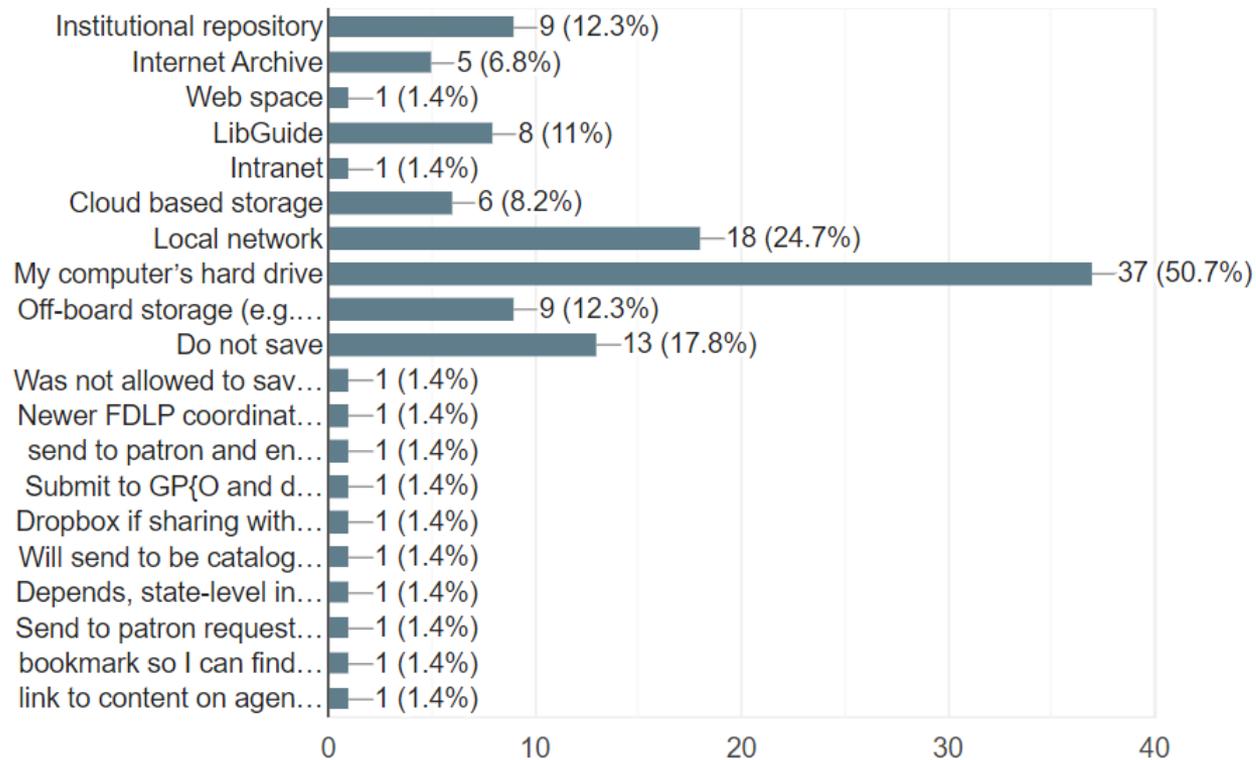
Do you download a copy of a found fugitive?

125 Responses



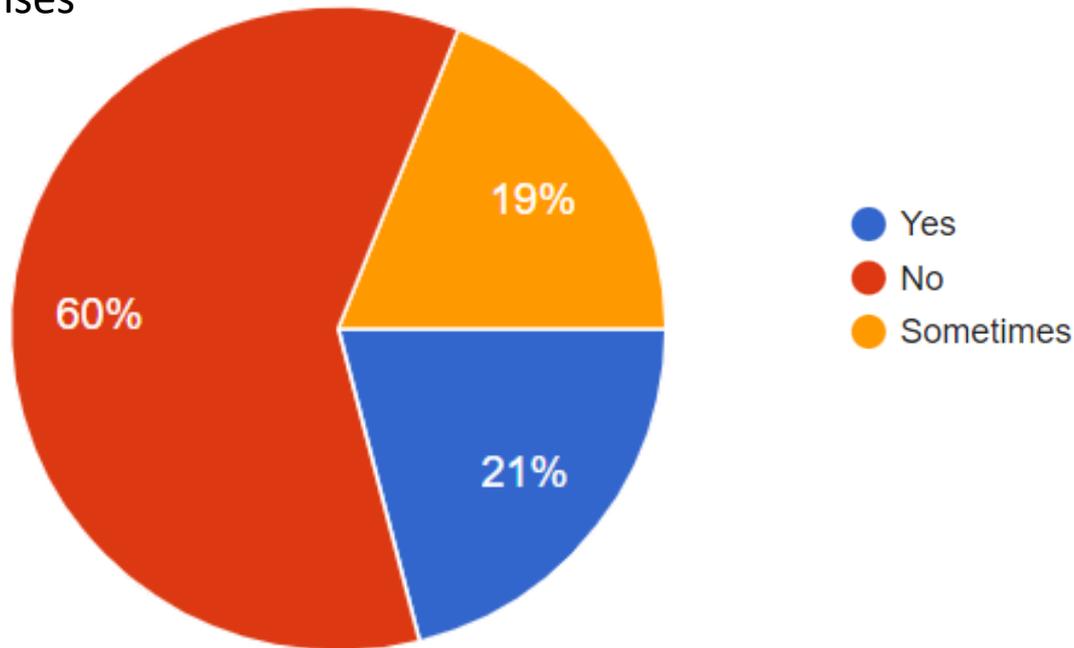
If yes, where do you download or save these items? (Check all that apply)

73 Responses



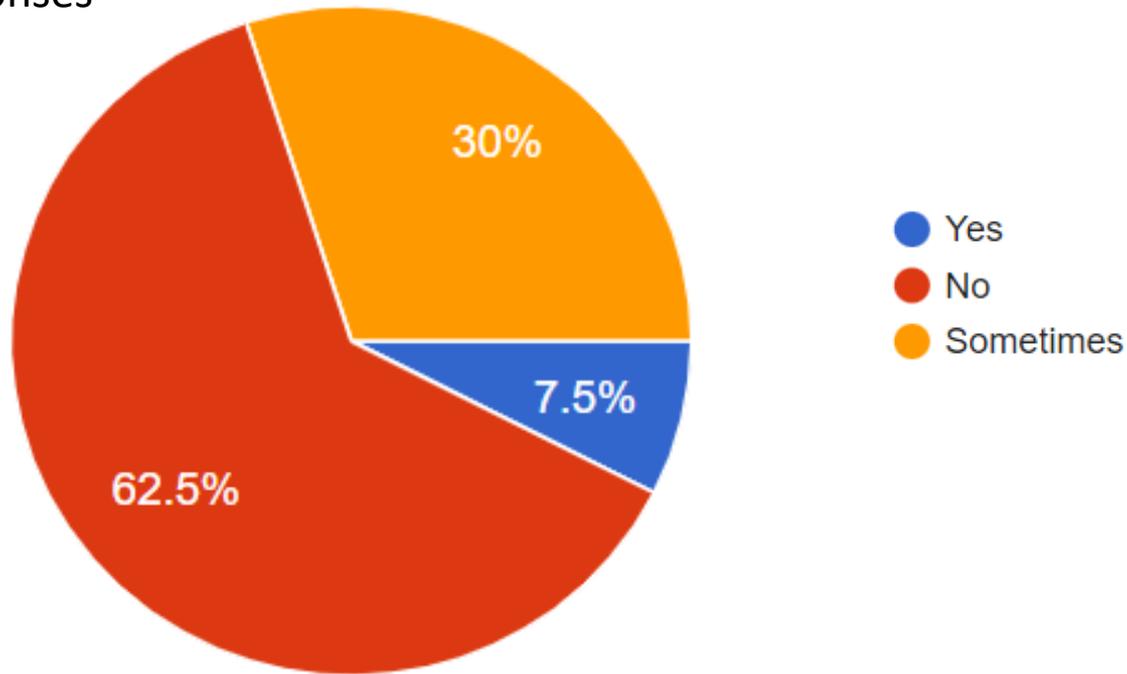
Are the fugitive documents you download or save accessible to the public in digital form?

100 Responses



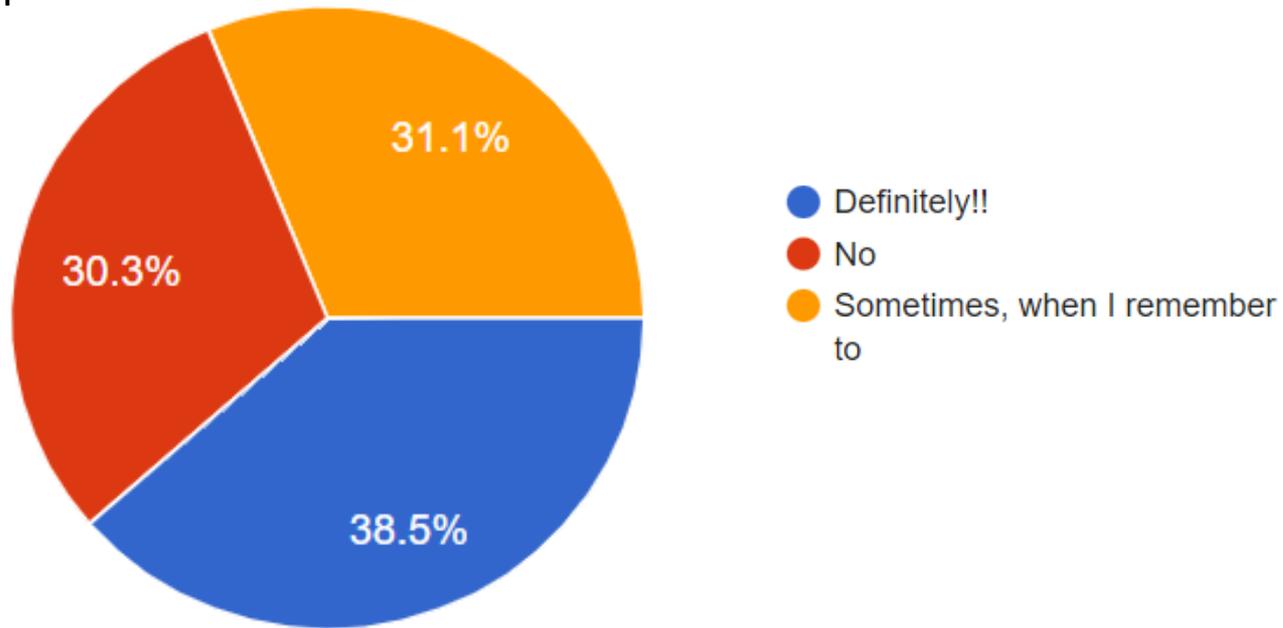
Do you print fugitive documents to include a physical copy in your local collection?

120 Responses



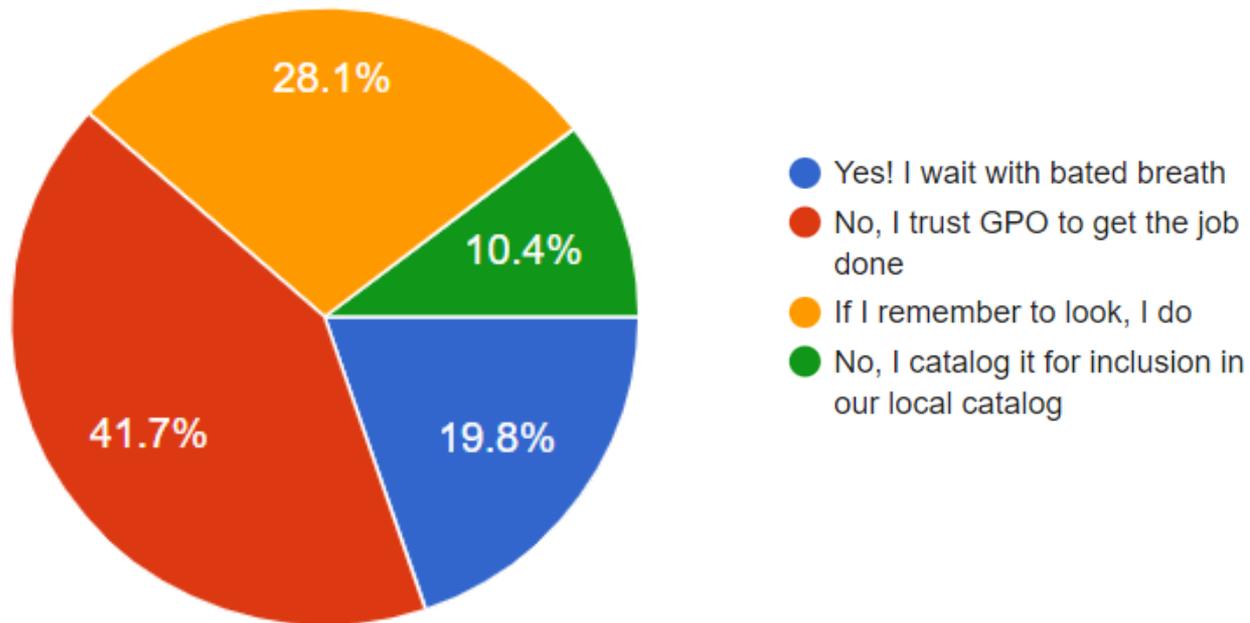
Do you submit fugitive documents to askGPO?

122 Responses



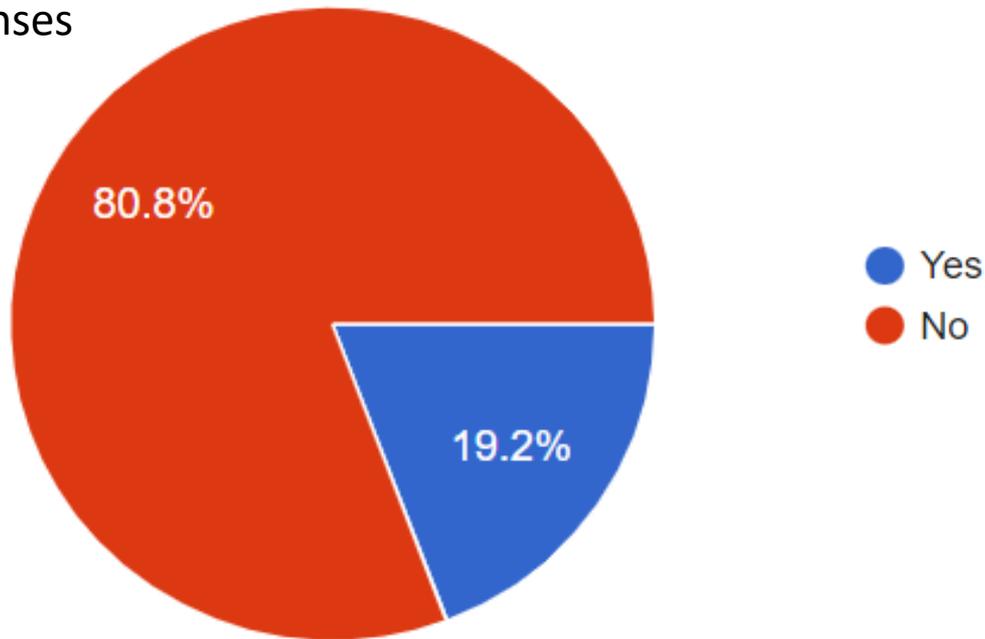
If you submit a fugitive to askGPO, do you watch the CGP to see that it's added to the system?

96 Responses



Do you ever receive notification from GPO that your submission is out of scope?

104 Responses



The Digital Deposit Working Group thanks all who completed the Unreported Publications Survey.

Appendix III: Focus Groups: Notifying GPO of Unreported Publications

- Focus Group Questions
- Focus Group 1 Session Notes
- Focus Group 2 Session Notes
- Focus Group 3 Session Notes

Focus Group Questions

- Describe your ideal fugitive document reporting process.
- What functionality or features would you like to see in a fugitive document reporting tool?
- How do you report fugitive documents to GPO?
- What are the positive aspects of askGPO? How could it be improved?
- What could GPO do to improve the fugitive reporting process or to make it easier for you?

Notifying GPO of Unreported Publications: Focus Group Session Notes

Focus Group — Session 1 | September 30, 2020

Describe your ideal fugitive document reporting process.

- Fugitive hunting/reporting falls at least midway down priority list.
- Typically generated when too many tangibles that can't be copy cataloged pile up - digital is less of a priority.
- Reporting via email may be easier. As reporting works now, form is completed; this generates email conversation requesting additional/follow-up information.
- A dedicated staff or workflow to reporting would be helpful.
- Halfway down the priority list as a regional coordinator. Happens with a group of docs not easily copy cataloged. Email. Form as it currently exists kind of allows. Depends on who is dealing with the other end. More dedicated staff or workflow at GPO, to ensure same person receives the document report.
- Digitizing print collection since 2011, found a number of fugitives, now they are digitizing some themselves. Time is of the essence for most people, way FDLP exchange is setup with differing levels could be put in by students, electronic bibliographic records, viable URLs. Joint publications between state, regional, federal agencies and can't find anywhere else. Tiered approach would cut down on sending and responding emails. Including initial input. Attempting to move print to online either through finding appropriate bib record or a PURL, or agency stable link, and have come across print items, not in OCLC or CGP.
- A process similar to FDLP eXchange:
 - ▶ Levels of information to include.
 - ▶ User level/rights.
- Trackable.
- Ways to locate information of value to a specific collection.
- New to the process - not sure how to hunt.
- Dropping an email to someone, rather than an online form, and include a URL.
- Make the process as simple as possible.
- An online form for reporting.
- new to gov docs. Likes an online form option. Understands email, keeping it simple on the website. Make the process simple since some are new to this.
- Easy to fill online form, if it could fill in information like the link while I am on that page that would be nice, like with a browser add on.

What functionality or features would you like to see in a fugitive document reporting tool?

- Fields on a form – fugitive hunting, listing searching tools. Indicating if there is field with OCLC number, and also for managing other metadata.
- OCLC # field as an ask, find a document in OCLC but not in CGP. Doesn't have a SuDoc. Would help streamline the process. Often find things in URL, so, a URL field. A picture or attachment field is useful for when there is not URL. A general notes field to explain things that are wonky.
- Different pieces and parts on the search. More that prompts to essential versus the essential. Similar to FDLP exchange some people can put in basic information. Fields that are essential and well labelled with the ability to hover over something with a bit of explanation, indicate the minimums versus. Smart sheet, enter title, any information from front page, agency and sub agency, any kind of date, or a stamp or when it came into the collection, SuDoc if it is there. Can attach images and pdfs, size, state agency referenced, URL if appropriate. Most of what they come across are not in OCLC or cannot find viable PURLs or URLs.
- Require or prompt for essential fields.
- Allow for extras (some may know this info, others may not)
- Clearly labeled and defined input fields.
- A place for additional information.
- Currently utilizes "smart sheet" for internally tracking projects/work.
- Utilize standard or required fields.
- Utilize drop down menus and additional input boxes as necessary.
- Ability to add or subtract fields as needed.
- Live chat for real time help.
- A couple of standard required fields, title, creator, author, add a drop down of I found URL, an input box net to it to add or subtract the number of fields, a chat box to ask questions on filling it out, a greater feature for the whole site, instant help for filling out the form.
- Auto fill my info (askGPO already does this if I sign in), a text box for the URL if online, ability to attach PDF, a mechanism for serials – so I could add years 2016-2020 in the same form instead of having to do each one individually.

How do you report fugitive documents to GPO? If you are not using askGPO, why not? If you are using askGPO, why?

- Reports in batches - not via email. Spends time completing askGPO form for each item.
- Feedback from GPO staff is not the same person - ideal would be to have one GPO staff as point of contact.

- Reports in batches and uses askGPO. Sometimes the feeling of getting responses at different times, haven't done fugitive docs reporting on the newest askGPO. One long-term staff member goes directly to a GPO employee. Does single entities in several batches.
- Utilizes askGPO to determine what is in scope".
- Needs help with assigning SuDoc.
- Always goes through askGPO.
- Does digitization for GPO ingest. Primarily use askGPO to clarify if it is a title which could have or should have been in CGP. Need a SuDoc call number, assignment? Should it be in the system and digitized for GPO ingest? Always goes through askGPO.
- askGPO is buried on the FDLP.gov website - "make it easier to find".
- Emails supervisor or regional with questions rather than going through askGPO.
- Has not used askGPO about fugitive, but has not found fugitive docs. Fell into federal documents, and self-learning.
- askGPO, I was unaware of any other method.

What are the positive aspects of askGPO? How could it be improved?

- Likes the new iteration of askGPO.
- It is easier to use.
- Not sure how to improve.
- Likes the new iteration, the categories. Other than few things already mentioned. They don't ask a whole lot. They have been doing the same thing for a long while trying to get them cataloged online.
- New askGPO is vast improvement on old.
- Likes the track progress feature.
- Automated responses are good.
- New categories are better.
- Individual login is improvement to entering contact info repeatedly.
- Much more efficient and clearer.
- New iteration met all of Jen's needs. Can track through a ticketing system. Can be assigned on the backend and user doesn't need to worry. Automated responses are a big plus. New categories are an improvement. Much more efficient and clearer for the user.
- Finds the site to be clear. Text is clear. Landing page does a good job breaking down the categories to submit the question. Very user friendly for a new set of eyes. Pretty effective for finding specific need.
- Auto filling my information, I like when they tell me they cataloged the item I found
- Text is small and difficult to read, especially on FAQs.

- The landing page is nice. Tile system is good and can click on. Text is really small, would have trouble reading it especially on the FAQ page. Like how the FAQ page has been grouped. Lots have to open the dropdown menu. The forms are clear, with the fields and narrowing down the questions.

What could GPO do to improve the fugitive reporting process or to make it easier for you?

- Knowing the process comes from doing and observing incremental changes over time.
- A refresher course would be helpful.
- The orientation to fugitive hunting was not long enough. Might have been more helpful to work through searches in session or have folks do pre-work.
- More help on working through the specifics of the process.
- Experience is key.
- Consistency in definitions.
- Try to avoid additional emails.
- A ticket system is good.
- Reporting process and supporting the creation and knowledge of best practices. Some of the things learned come from doing and living through GPO changes. Doesn't hurt to refresh folks. An hour webinar wouldn't do as good as taking an orientation program and create an hour long where folks have homework and can go out and look at what is being talked about. Training directly related to hands on. Fugitive reporting process and consistency and definitions, same info and same boxes, same way. Standardizing so less time is not spent trying to decipher. Ticket system as with askGPO, something other than askGPO still needs to be trackable by the user.
- Workshops to learn the process.
- New to docs and thought fugitive hunting would be cool but have no idea where or how to start.
- A series of workshops to learn the process. Doesn't know how to get started, but excited. Wants to get started and not sure where to start.
- Help with hunting a specific agency or type of document.
- Training would be really helpful, one off webinars or a series. Things can do for certain agencies. Tips and tricks. Set of documents.
- Some sort of formal and informal collaborative process, where people can learn from others. Agency specific domains. More overlap between regional and selectives doing fugitive.
- Have a browser add on where if I am on a website and open a PDF I can click the browser button and it will automatically let me login and auto fill the PDF link and upload the PDF (not sure if that is possible).

- Have a way to upload websites. For example, I have been adding some CDC COVID guidance web pages to be archived by the main Wayback Machine because they keep changing and I think it is important to document the changes. I feel like the askGPO form is more for documents that are a file of some sort (PDF, Word, Excel) and until just now, I did not know we could recommend websites (https://www.fdlp.gov/project-list/web-archiving#can_i_recommend), maybe have a dropdown box to select a file or a website so it is obvious you can do both.

Focus Group — Session 2 | October 5, 2020

Describe your ideal fugitive document reporting process.

- Happy with how reporting works - reporting what's found and all information about it.
 - ▶ Would like more prompt response. Yes, this is fugitive and fits within scope. Works with paper docs so these pile up. Less of a problem with digital documents.
- Notes the source doc for the pilot project is cumbersome.
- Has not reported fugitives. Thought the purpose of this group was different.
- Received a large shipment with fugitives. Going through askGPO was ok. Created a spreadsheet with information including scans of covers and verso. Still missing feedback on 30 or so documents... communication back from GPO about status, within scope, etc. would be helpful. Explanation of what not in scope.
- Any reporting of born digital/electronic documents. Usually deal with physical donations and not born digital materials.
- Most fugitives are donations from congress people or materials long abandoned. Very rarely encounter digital documents.

What functionality or features would you like to see in a fugitive document reporting tool?

- Time consuming to make spreadsheets. Created sheets of near cataloging records, plus scans. Might be easier to send images of cover, title page, other rather than transcribing the information.
- Sometimes provided tons of bib information - what is expected?
- Data fields for input rather than single text box. Might have provided too much, but erred on side of too much vs. too little.
- Would also be helpful to know the rarity of document.
- Provide information about what's already been reported so not duplicating work of others.
- Suggests google sheet as a shared document.

- OCLC sometimes found. Need for baseline information - guidance on what's the minimum level of what to send with fugitive hunting.
 - ▶ Old askGPO offered little to no guidance.

How do you report fugitive documents to GPO? If you are not using askGPO, why not? If you are using askGPO, why?

- Not experienced with finding so cannot say specifically. Would likely use askGPO but would seek information on how to report fugitive documents.
- Would prefer something specific for fugitive. Does use askGPO. Not aware of other ways to report outside emailing a specific outreach librarian.
- Uses askGPO. Doesn't know of another way. "What was told to do in the GPO docs coordinator certificate program"
- Has not reported fugitives. New to this process but would use askGPO.
 - ▶ Finds docs for student research. If interesting, requests for local cataloging.

What are the positive aspects of AskGPO? How could it be improved?

- Has not used new askGPO. Appreciated the training. Also, glad to have login information. Likes ticketing for tracking purposes. Categories are still not clear - not sure where question falls. Drop down confusing because so much is covered. FDLP questions are very broad, almost too broad.
- New askGPO very easy to use. No improvements as of now. Except greater understand of what GPO needs for fugitive hunting. Fielded entry for reporting. Or at very least, bare minimum of what to include. This will help to avoid back and forth.
- askGPO form looks straightforward. If click on fugitive - route to form or something that gives specifics of information needed. Again, to avoid a lot of back and forth.
- Need both - a way for multi submission as well as a way to report one or two.
- Maybe a spreadsheet for multiple submissions.

What could GPO do to improve the fugitive reporting process or to make it easier for you?

- Improved input form. Better way to track what's been submitted / completed by askGPO. Tool to provide the best information possible.
- Fugitive documents primer.
- Know who is working on a given title so can request updates rather than going back through askGPO.
- Need to know exactly what's needed for the cataloging. Is it fugitive - if no, why? More information about scope of program.
- Need to know more before submission about scope and/or more information about how scope decision is made.

Focus Group — Session 3 | October 15, 2020

Describe your ideal fugitive document reporting process.

- The web set up is good. Would first look at CGP and then the agency site.
- Loves exploring agency sites.
- Some agencies she automatically goes to, other times it's driven by what's in the news.
- Normally does not report. She instead requests directly from the agency. If found in OCLC thinks it must be in the program.
- The value of OCLC as a discovery tool to see what is out there and what libraries have captured in some way.
- Checks CGP, not in CGP, seems in scope.
- Ideal process would be easy = simple and similar to seed nomination tool for EOT.
- Bookmarklet tool not great for reporting fugitives, wants to provide context for why or what... share information about what might be missing.
- A quick form, something easy and quick.

What functionality or features would you like to see in a fugitive document reporting tool?

- Nice to have structure.... Title, URL, etc. but leave some flexibility.
- A way to keep track of what's been reported / contributed (EOT tool lets you keep track of nominated URLs).
- Mobile-friendly.
- A snippet tool - something that can capture an image ... Especially with a generic type title. The ability to upload an image or some kind of image capture.

How do you report fugitive documents to GPO?

- askGPO to report concerns or documents of interest.
- It is an adequate tool - response received is timely.
- At one time used a separate lost docs form.
- More likely to start with CGP, not found and then seek out elsewhere. If one thing, unlikely to report.
- Tends to contact the agency and not GPO. If OCLC record is good, tends not to report to GPO.

What are the positive aspects of askGPO? How could it be improved?

- Likes dropdown that facilitates point of contact.
- Likes that file can be uploaded.
- Form is Ok... no issues with it.
- Response is great.
- Not too many problems.

- Not sure of improvements.
- Not used new version.
- Attach files.
- A lot of flexibility.
- One form for a lot of requests - this is good and not so good.
- Categories could be confusing.
- Glad to know who's responding.
- More streamlined to the questions that were needed to be submitted.

What could GPO do to improve the fugitive reporting process or to make it easier for you?

- Different kinds of resources that are not in CGP - one starting point but ability to report different kinds of things. Important that the person on the other end knows what and why something is being suggested.
 - ▶ Examples - a program website that needs to be captured or crawled vs. a single missing issue.
- A dashboard to report multiple items.
- Administrative perspective - a way to track work accomplished by a single person. A way to quantify accomplishment to show as part of work for annual performance evaluations.
- Ability to develop long term relationships:
 - ▶ A way to assign work to a person.
 - ▶ Return that value back to ourselves.
- More collaboration with the agencies and FDLP.
 - ▶ Collaborative effort to capture everything that's out there.
- Agencies don't retain inventories of what's on their sites. Does this offer opportunity for collaboration to inventory publications on an agency site?
 - ▶ A brief inventory of what's available in a system.
 - ▶ Maybe a way to create brief records.
- Recognizing having a broad basis of what's out there to know what has fallen through the cracks.
- Collaborative effort between agencies, libraries, and GPO needed.
- Reminder of historic shelf list cataloging project - how might this type of project work for the digital era.
- Could ask GPO have a module for adding title, agency, link - brief record stuff?

Appendix IV: In Search of Congressionally Mandated Reports

- Consolidated Completed Volunteer Work Forms
- Search Preparations and Proposed Workflow
- Work Form Data Elements Described
- Volunteer Work Form (with sample entries)

In Search of Congressionally Mandated Reports Consolidated Completed Volunteer Work Forms

The completed work forms of all the volunteers were consolidated, creating a file the size of which, visually, is not practical to include this report. The spreadsheet has 20 columns and 1,049 rows of data. Fitting it onto pages in this report would make the content indecipherable. The spreadsheet, [Consolidated Completed Volunteer Work Forms](#), is available as a separate file independent of this report.

What follows are the preparation tips and guidance for the volunteers, which were included in their work form as the “Instructions” and “Data Elements Described” tabs, and a representation of the work form columns and three rows of data that served as an example of how to complete the form.

The Hunt for Congressionally Mandated Reports

A component of Depository Library Council Digital Deposit Working Group pilot

As part of the pilot project of the Digital Deposit Working Group of the Depository Library Council we will be hunting for fugitive documents and gathering data on how people search for them. One of the outcomes will be a best practices search guide for finding fugitives, hence the data gathering. Each fugitive hunter will have a spreadsheet work form on which searching information will be recorded.

Fugitive hunting will use the known universe of "Reports to be made to Congress" publications to search. While looking for these titles you may also run into other fugitives along the way. Please report them on your work form as well. The "Fugitive found was ..." data column will distinguish it from the others.

I. PREPARING FOR THE HUNT

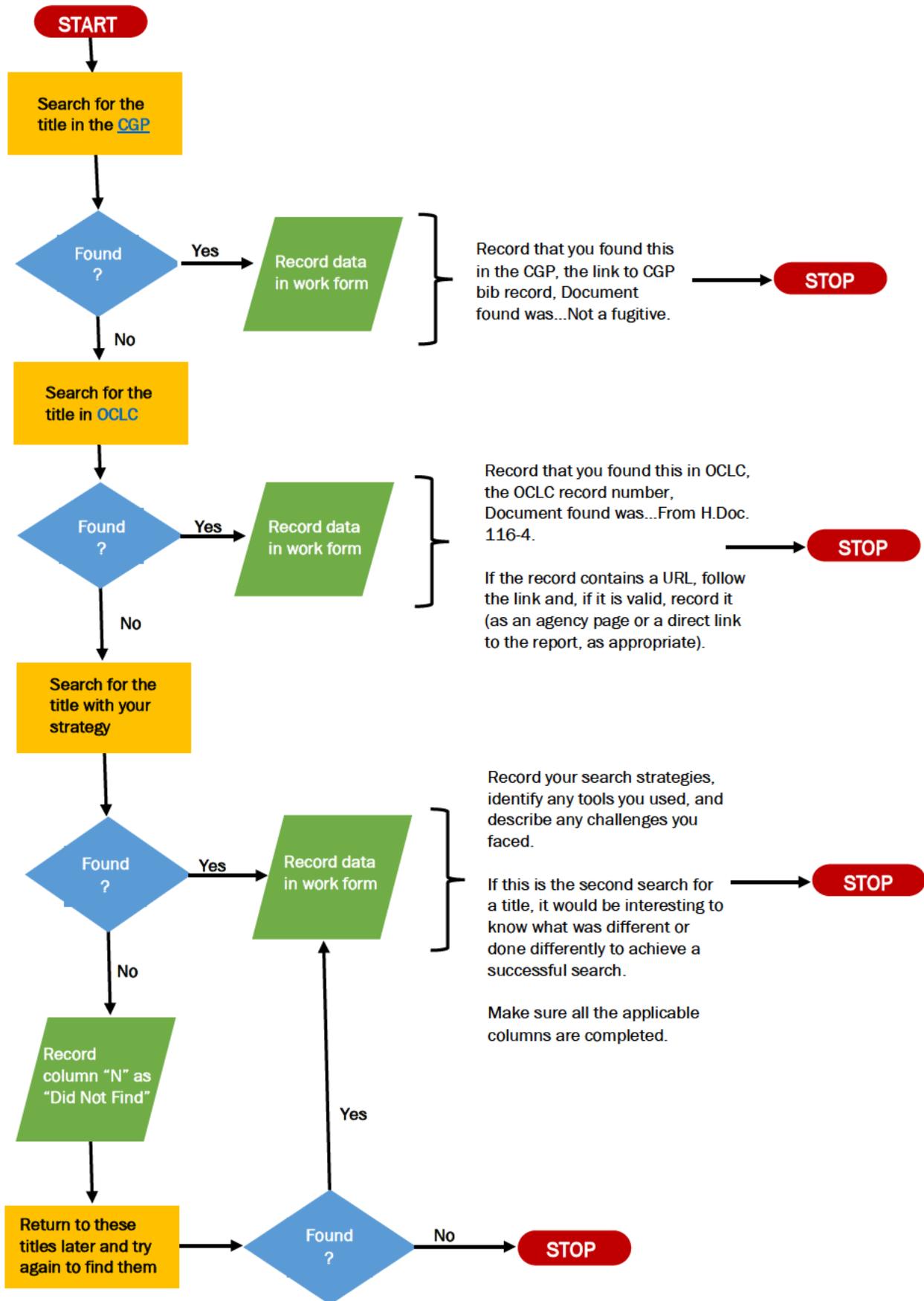
1. Review and familiarize yourself with your agencies from [House Document 116-4](#), and with the data elements in the work form.
2. Use the **Nature of the Report** to search for, determine, and document the titles of the reports.
3. TIPS FOR DETERMINING TITLES
 - Use a search engine to search for site:*.gov+ "reports to congress". Result may be one report or a landing page with numerous reports
 - If search result takes you to an agency landing page, see if any of the reports match the information from H.Doc 116-4 entry (Nature of Report or authority citation).
 - Search the agency and the public law number identified in the authority citation in H. Doc. 11-4. e.g., *gao.gov AND "pursuant to Pub.L.93-344".
 - It may help to bold or highlight words in the Nature of the Report or agency that are likely to be in a report title. Search the words with "report to Congress".

LIST OF REPORTS WHICH IT IS THE DUTY OF ANY OFFICER OR DEPARTMENT TO MAKE TO CONGRESS
PART 4 - REPORTS BY CABINET-LEVEL DEPARTMENTS - CONTINUED
DEPARTMENT OF HOMELAND SECURITY - CONTINUED

Nature of report	Authority	When expected to be made
Repayment plan for borrowing authority [relates to payment obligation under 42 U.S.C. 4016(a)(2)]	Pub. L. 112-141, Sec. 100213(b); (126 Stat. 924)	Not later than the expiration of the 6-month period beginning on the date of enactment of this Act [July 6, 2012]
Reinsurance private market pricing assessment	Pub. L. 112-141, Sec. 100232(c)(1); (126 Stat. 953)	Not later than 12 months after the date of enactment of this Act [July 6, 2012]
Recommendations for reducing costs of future disasters	Pub. L. 113-2, Sec. 1111(a); (127 Stat. 49)	Not later than 180 days after the date of enactment of this division [January 29, 2013]
Feasibility of releasing property-level policy and claims data for flood insurance coverage under the National Flood Insurance Program and establishing appropriate guidelines	Pub. L. 113-89, Sec. 31(b)(1); (128 Stat. 1035)	Not later than 90 days after the date of enactment of this Act March 21, 2014
Interim status update on progress developing the report defining the priorities, goals, and expected outcomes of the recovery effort for the Commonwealth, based on damage assessments prepared pursuant to Federal law	Pub. L. 115-123, Sec. 21210(b); (132 Stat. 109)	At the end of every 30-day period before the submission of the report described in subsection (a)
Recommendations on how to help with anticipated evacuation route flow, based the adequacy of available evacuation routes to accommodate the flow of evacuees	Pub. L. 115-254, Sec. 1209(c)	[no deadline specified]
Number of instances and the estimated amounts involved, by State, for cases in which self-insurance amounts have been insufficient to address flood damages	Pub. L. 115-254, Sec. 1240	Not later than 2 years after October 5, 2018, and each year thereafter until 2023
Puerto Rico's 12- and 24-month economic and disaster recovery plan	Pub. L. 115-123, Sec. 21210(a); (132 Stat. 108)	Not later than 180 days after February 9, 2018

Sample page from H.Doc. 116-4

II. BEGIN THE HUNT



In Search of Congressionally Mandated Reports Work Form Data Elements Described

DESCRIBED										
DATA ELEMENT										
Date	Date for each searching session using YYYY-MM-DD as the format.									
Time Spent on Session	Hours per search session and partial hours in increments of .25 of an hour. Round down if time is 1-7 minutes into the quarter hour, and round up if 8-14 minutes.									
Title of Document	Document the title that corresponds to the Nature of the Report from H.Doc. 116-4									
Document Found in CGP/OCLC	Dropdown box options: CGP, OCLC, Did Not Find.									
Where did you start searching?	If unsuccessful with CGP/OCLC. Dropdown box options: Agency website, Google, Other search engine, USA.gov, NA									
Search Strategy	Describe your search strategy, e.g., Google search of Agency Name and *.gov Report to Congress.									
Tools Utilized	You may find you use tools for some and not for others. Report NA if you didn't use tools. Examples of tools include search alerts, RSS feeds, browser plugins									
Document Found Was	Dropdown box options for the fugitive found was From H.Doc 116-4, Not from H.Doc. 116-4, Did Not Find, Not a Fugitive.									
Challenges	Describe any obstacles you encountered while searching for a report.									
CGP Record	If you find the report in the CGP, provide a link to the bibliographic record									
OCLC #	If you find the report in OCLC, provide the record #									
URL for agency index page	Provide the link to the agency's index or landing page for their reports, if you come across one.									
URL to Report	Provide the most direct link to the report you can find.									
PURL	Provide the PURL to the report, if you found one.									
Comments	Optional: opportunity to provide any comments you wish to make. For example, The CGP record I found is for the tangible version of the report, I found the digital vers									
OTHER TERMS DEFINED										
CGP	Catalog of U.S. Government Publications https://catalog.gpo.gov									
Fugitive document or Fugitives	Public information products that are not discoverable through the Government Publishing Office's Catalog of U.S. Government Publications									
GPO	Government Publishing Office's OCLC symbol. GPO in the 040 field indicates GPO originally created the bibliographic record.									
House Document No. 116-4	This lists all the "Reports to be made to Congress," more comonly known as Congressionally Mandated Reports. It is the basis for the work of this pilot, i.e., the source for the fugitive documents to hunt (Y 1.1/7:). <i>A corresponding unofficial dataset was created for use in conjunction with the United States Government Publishing Office Federal Depository Library Council Digital Deposit Working Group Reports To Be Made to Congress Pilot Project.</i>									
CMR	Congressionally Mandated Report									
COLOR CODING OF COLUMNS										
Green	Session metrics									
Gold	Search process information									
Red	Information incorporated from House Document 116-4									
Blue	Information found that is unique to the document being searched									
Purple	Comments (optional)									
Pale Yellow	Sample entries									

In Search of Congressionally Mandated Reports

Volunteer Work Form (with sample entries)

DATE (YYYY-MM-DD)	TIME SPENT ON SESSION	DEPARTMENT	AGENCY	NATURE OF REPORT
2020-09-01	3.15	INDEPENDENT AGENCIES, BOARDS, AND COMMISSIONS	400 Years of African-American History Commission	Summary of the activities of the Commission; a final accounting of funds received and expended by the Commission; and the findings and recommendations of the Commission
		INDEPENDENT AGENCIES, BOARDS, AND COMMISSIONS	400 Years of African-American History Commission	
		INDEPENDENT AGENCIES, BOARDS, AND COMMISSIONS	Environmental Protection Agency	Sediment survey and monitoring

CITATION	FREQUENCY	NO LONGER DUE	TITLE	DOCUMENT FOUND IN CGP/OCLC?	WHERE DID YOU START SEARCHING?
36 USC 101 note prec.; Pub. L. 115-102, Sec. 7(b); (131 Stat. 2252)	Not later than July 1, 2020	7/1/2020		Did Not Find	Google
			400 Years of African-American History Commission : Strategic Plan August 18, 2019 - July 1, 2020	Did Not Find	Agency Website
33 U.S.C. 1271((b))((2)); Pub. L. 102-580, Sec. 503(b)(2); (106 Stat. 4866)	Biennially	10/31/2016	The incidence and severity of sediment contamination in surface waters of the United States	CGP	

SEARCH STRATEGY	TOOLS UTILIZED	DOCUMENT FOUND WAS ...	CHALLENGES	CGP RECORD	OCLC #
Google to search for agency website	NA	Did Not Find	Commission established by law 01/08/2018. Final report was due July 1, 2020 ... behind schedule perhaps from pandemic?	NA	NA
Link to plan was clearly visible from homepage -- stumbled upon the publication	NA	Not from H.Doc. 116-4	Commission is administered by the National Park Service. Two websites - one part of the NPS domain and another at http://400yaahc.com/	NA	NA
		Not a Fugitive		1997: 000591014 2004: 000591015	

URL FOR AGENCY INDEX PAGE(LANDING PAGE)	URL TO REPORT	COMMENTS
https://www.400yaahc.gov/	-	1. Bolded terms in the "Nature of the Report" that might likely be in the title. 2. From the Strategic Plan: . Submit a final written report with recommendations prior to the sun-setting of the Commission on or before July 1, 2020.
Small commission, link to strategic plan is on the main page of the commission website and the main page of the commission site of National Park Service	https://www.nps.gov/orgs/1892/upload/400YAAHC_FINAL_Strategic-Plan_Accessible_rev5-20-2019.pdf	Title in running header varies. STRATEGIC PLAN: 400 Years of African American History Commission
https://permanent.fdlp.gov/lps70422/lps	2004: https://permanent.fdlp.gov/LPS71171/901U0000.PDF	1. 2004 indicates "second edition" 2. 1997 is 3 volumes linked from landing page on permanent

Appendix V: Visualizing Digital Deposit Subgroup Final Report

Visualizing Digital Deposit Subgroup. Digital Deposit Working Group. Depository Library Council. *Visualizing Deposit of Digital Content from the Government Publishing Office*, February 9, 2022.

Visualizing Deposit of Digital Content
from the
Government Publishing Office

Report of the
Visualizing Digital Deposit Subgroup
of the Digital Deposit Working Group,
a Working Group of the Depository Library Council

February 9, 2022

Members of the Visualizing Digital Deposit Subgroup

Julia Ezzo
Michigan State University
Member, Depository Library Council

Heather Christenson
HathiTrust
Depository Library Community

James R. Jacobs
Stanford University
Depository Library Community

Cynthia Etkin
Government Publishing Office

Jessica Tieman
Government Publishing Office

We thank the people from the following institutions who graciously gave their time to help us gather data on their views of digital deposit.

Aaniiih Nakoda College
Cornell Legal Information Institute
HathiTrust
Illinois State Library
Sacramento Public Library

Stanford Law Library
University of California, Berkeley
University of Colorado, Boulder
University of Georgia
University of North Texas

Table of Contents

Introduction1
Methodology1
Findings & Analysis.....2
Overall Observations.....7
Recommendations8

Introduction

The Digital Deposit Working Group (DDWG) was established to “explore current and future needs related to digital deposit— both dissemination of content and acceptance of content by GPO”. A subgroup, the Visualizing Digital Deposit Subgroup (VDDS) was tasked with exploring or “visualizing” what the deposit of digital content by GPO to Federal depository libraries might look like.

VDDS approached this charge by identifying use cases, developing high-level questions, and conducting interviews to garner in-depth feedback. Questions were asked of librarians representing each of the identified use cases to bring to fruition the concept of digital deposit to Federal Depository Library Program (FDLP) libraries.

Methodology

VDDS members created a chart for the different types of library and non-library organizations representing different use cases: FDLP regional library, various FDLP selective libraries (academic library, law library, public library, Tribal college library), individual library practicing digital preservation, an ASERL “Center of Excellence” library, digital preservation steward, and library-adjacent non-profit organizations working in the digital preservation and access space HathiTrust, and the Legal Information Institute (LII) at Cornell University. One representative organization per type was selected, primarily based on a positive response to past *Biennial Survey of Federal Depository Libraries* questions asking if their organization would be interested in digital deposit. For organization types that did not have a positive response to this question, subgroup members selected representative organizations using the FDLP directory, while taking in to account regional distribution to try to allow for cross-coastal representation by participants. Two of the organization types are not FDLP participants (HathiTrust and LII), though their institutional members and parent organization are FDLP libraries. Subgroup members selected interview candidates based on their knowledge of appropriate organizations.

The open-ended interview questions were designed to elicit detailed responses regarding interviewees' thoughts surrounding the feasibility of digital deposit. Questions were left intentionally nonspecific, with minimal definitions, clarification, or contextualization provided, in an effort to avoid influencing the responses. In the case of interviewees that

were not members of the FDLP (HathiTrust and LII), some of the questions were modified slightly to make them relevant to that particular organization.

Library Services and Content Management’s eLearning platform was used to conduct the interviews. There were fourteen interviewees representing the ten business cases. And there were at least three, often more, subgroup members in attendance at the interviews. Each representative organization was interviewed separately by one member of the subgroup, though other subgroup members were in attendance. The interviewer read aloud each question, and allowed the interviewee(s) to ask clarifying questions prior to providing their response. Subgroup members were also able to ask follow-up questions to the responses given in the event that they wanted additional clarification or elaboration. Interviews were recorded in nine out of ten instances—one interview was not recorded due to technical difficulties. Responses were noted by one of the subgroup members during the interviews and the recordings were later reviewed by subgroup members to collate the responses for analysis.

Findings & Analysis

The interviews shed light on the libraries’ point of view and provided details on how they envision different aspects of a digital deposit service. The following is a review of responses to the seven specific questions.

Question 1: What content would your library be most interested in receiving from GPO? For example, the latest hearings or maps for your geographic region.

While two of the libraries interviewed indicated that they wished to receive all available materials from GPO via digital deposit, the other eight interviewees preferred to receive a **selection of materials** through digital deposit. Interviewees were interested in selecting materials on a particular **subject**, such as legal or environmental materials. They also expressed a desire to receive all materials in a particular **format**, for example maps or XML files. Receiving all materials for specific **geographic** or regional areas was also suggested.

Question 2: How would you envision the selection process?

Interviewees want an easy, flexible method for selection that is clear as to what exactly a library is selecting. One librarian working with “a very diverse community ... would like to

be able to have easy access to a variety of topics and be able to modify selection of materials easily and quickly to meet patron interests.”

A desire to move beyond the rigidity of the current item selection profile model to allow libraries the ability to **select materials by a variety of methods** was expressed. Examples noted in the interviews include: select all items in a particular **format**, select all materials that deal with specific **geographic** regions, select by **collections or general categories** (such as all legal-related publications, all material dealing with environmental **topics**, etc.), pre-made collections by library type (law library, public library, etc.).

Question 3: How would you make digital content accessible to the public?

The interviewees mentioned a **variety of modes for making digital documents accessible to the public**. The most frequently envisioned situation was to host the digital documents on a local platform, catalog them, and **provide links via library catalog**. Additional respondents described linking to the digital documents via a catalog, but did not specify that digital documents would be hosted locally. These responses included “on a server and point to it,” and “online through a shared catalog.” Related to the catalog model, we heard that “ideally cataloging and metadata come with the digital volume.” Inclusion of bibliographic data as part of the process seemed to be an underlying assumption for many interviewees although they did not specifically point out that as a requirement.

Libraries were also **thinking beyond the traditional catalog access model**. Answers in this vein included “in catalog but also explanatory/promotional text so people know to look in the catalog,” “links from **library website/web pages**,” “content visible on the web site,” and “full text index.” There was also an emphasis on **collections as a mode of access** by some interviewees, with mentions of “Browse by collection or SuDoc” and “Curated collections created by the institution.”

Three libraries mentioned **provision of documents for computational and experimental use**, saying “present data and documents in a way that facilitates understanding by the general public and makes use of some computer science techniques to extract features of the documents themselves to make these more explicit,” and “Allow experimenting with context for digital volumes” and interest in creating a “research ‘sandbox.’” One interviewee emphasized the **preservation** aspect and envisioned they would “Preserve but point to GPO for access.” One interviewee specifically mentioned “**downloadable for users**” as key.

Question 4: What would be the optimal mechanics of digital deposit (push/pull) for your library?

The responses to this question were mixed and somewhat inconclusive, with slightly more interest in a “push” model than a “pull” model, and with discussion of a number of ideas that don’t fit neatly into either category but that are beneficial to designing a service. Of note is that this question was asked without specific context or definition of the models, other than the mention of push or pull, in order to give a starting point for the discussion and allow the libraries to put it in context themselves.

Four libraries emphasized the **push model** in their answers, envisioning a push from GPO to the library to be both **familiar and easier**, as it is “similar to what we have with tangibles”, and that the library could “select what we want then it just shows up” with “one less thing to have staff do.” Multiple libraries mentioned another advantage of this model, that it solves some of the challenges of “gathering materials in the digital age” and would help libraries “know that they are receiving all that they should be receiving.” One library summed it up “this method keeps it **authenticated, authoritative, and complete.**”

Three libraries indicated an overall preference for a **pull model**, mentioning familiarity with this model and a **record of success in using it**, and also emphasizing that **this kind of model gives libraries more control and flexibility** with such comments as “pulling allows for flexibility to get content when you’re ready” and “it would be nice to have control as to when to acquire.” Some pointed out that their experience of push methods has been burdensome, noting that they “tend to have **issues and errors with push**. When pulling content either through URLs, Dropbox content, rsync, etc. These tend to work better and have less issues.” and “with push it’d always have to be on to receive content when pushed.” The conversations with libraries also brought up the point that a **pull model would necessarily require a method to share information on what items are available to pull** periodically, such as a “push notification from GPO to indicate material is available and then the library would be able to go in and retrieve those files.” or “tools like bulk data and feeds that announce the changes in a standardized format.” Some libraries cited pulling MARC records from GitHub and MARCIVE as examples of the kind of processes that serve their needs.

Several libraries noted that **having “cataloging records attached” would be key to their adoption of either a push or pull model**, and that “It’d be difficult to have several thousand files coming in at once without metadata attached.” Some respondents expressed an interest in receiving metadata beyond bibliographic data that could serve a

wider set of use cases, wondering “what sort of metadata (description, technical, and preservation) would be paired with these records?” There was also mention of a need for “**a process to review the material** that is available to decide if we want it or not before downloading” with one library hoping for a way to craft a **selection process** that would allow them to “assume that when they’re notified content is available then it is something they want to add without additional review.”

Another concept that was brought up was a process that would also allow the library to contribute digital documents to GPO, “**would like for communication to be two-way** and all institutions to be able to deposit information into the system themselves rather than it being a one-sided process of just receiving materials.”

Question 5: Would your library be interested in receiving physical formats (e.g., DVDs) of digital objects for offline digital access?

Libraries were **not interested** in receiving physical formats for offline digital access, with the exception of one library that allowed for the possibility, stating, “that falls into a conversation about what the library will look like in 5 years.” Reasons for rejecting the idea of physical formats for digital access included lack of software and hardware needed to read the format, difficulty of “physical thing management” stewardship and format migration, and that users no longer understand why they’d need to use a DVD, or how to use it. Although not an access scenario, one library mentioned hard drives as a mechanism to transfer a large volume of files for ingest, “but this wouldn’t be ideal or preferred.”

Question 6: What kind of support/resources (technical, human etc.) do you think your library might need? What support/resources/infrastructure/staff are already in place at your library that would facilitate digital deposit?

In terms of support, libraries expressed a need for documentation, **training**, and clear **communications** from GPO. Having a **predictable process** would enable local workflows and automation. “Materials to promote the project, value, and ease” would help to get administrators and staff on board.

Most libraries responded that they **likely had some staff resources and infrastructure** to draw upon. Some of these resources were local, cross-campus, or at a state/consortial level. Examples of the kind of staffing available included IT staff, catalogers, and digital preservation staff. One (Tribal college) library had no local resources and potential

challenges connecting digital deposit with their shared catalog. This library said they wouldn't be able to do digital deposit locally, but would be interested at a consortial level.

Question 7: What do you see as potential FDLP requirements (conceptually similar to tangible materials or different in some way in terms of collection management, GPO support etc.)? What would be prohibitive to your library receiving files on digital deposit?

A majority of the libraries voiced a need for FDLP requirements to be **consistent, flexible, and relatively few**. Comments included, "keep things simple and consistent," "the fewer the requirements the better," "Keeping requirements consistent will be important to make it easier for staff," and "I don't know if there should be many requirements."

A number of libraries also wanted FDLP requirements to be **similar to requirements for tangible materials**, saying "would expect some of the same requirements as print" and "keeping it like the tangible program makes it easier to understand and doesn't have all new different requirements to track." Several libraries mentioned the dynamic between tangible and digital materials and that requirements for both should be in harmony. "Requirements would also need to depend on if there is a parallel process with print distribution" and "GPO would need to provide more guidance and assistance when it comes to collection management."

Several libraries mentioned the idea of a **tiered system that would offer flexibility** to libraries in terms of commitment. "Some would be fine with just a hard drive and basic hosting, others would do more" or there could be "access partners, preservation partners, etc." Libraries also noted that retention in a digital environment would "**need to be rethought because server space is different than physical space.**" Another notable comment was related to authenticity in a digital environment: "Would need to have some sort of connection between GPO's copy and distributed copies. Allow for provenance."

Some of the other specific comments regarding FDLP requirements were oriented around **digital preservation and access**, and tended toward one or the other. For example, a large library organization reported their "processes are geared around long-term preservation" and they are "willing to preserve long term." Some libraries "would like to not sign on to take material and preserve it all forever, potentially. Would want to be able to just keep the latest, for example." and one said "It would be nice to be able to remove items that aren't being used before 5 years."

In regards to the question of what would make digital deposit prohibitive to libraries, some libraries mentioned **technical challenges** such as server space, dealing with format obsolescence and migration, and access barriers such as “having adequate equipment to read and access the material.” One library mentioned that they would need “more **support to get the program started**. To help coordinators get oriented, build this type of collection, do outreach in their community about this collection, etc.”

Overall Observations

Digital Deposit is not just for one kind of library

The interviews were conducted during the covid-19 pandemic lockdown phase, and the real-life experience of remote library service provision most likely had some bearing on perspectives related to digital content. However, in our interviews with ten different organizations, each with their own institutional considerations and use cases for digital deposit, they all expressed interest and had ideas about how they would interact with a service. Additionally, most of the libraries indicated that they had access to some form of resources that would enable them to participate in a digital deposit service, whether those resources were local, or at a regional or consortial level.

Flexible selection is important to libraries

Based on this sample of libraries, a one size fits all digital deposit approach would not serve libraries’ needs. Libraries expressed a desire to select digital materials to fit their own collection priorities, in a variety of ways such as topic, provenance, geographic range, format or other levels of granularity or time variables.

Bibliographic data is a requirement for digital deposit, and physical formats are not

The interviews provided an especially strong indication of specific requirements for two particular aspects of digital deposit. Responses showed that libraries view bibliographic data as inherent to digital deposit. We also heard strong opinions from our interviewees that they did not want to receive digital files on physical media.

A coordinated digital deposit service should be in harmony with existing FDLP model(s)

Based on the interviews, libraries see benefit in digital deposit being offered as a full program or service that includes not just deposit mechanisms but also coordination,

documentation, training, and communications support from GPO. Libraries also expressed a desire for an understanding of the dynamic between digital deposit requirements and existing FDLP requirements to be incorporated as requirements are developed.

Digital deposit could enable new use cases for Federal documents

Against the backdrop of the pandemic, the ability of a library to provide digital access is in the forefront of people's minds. In our interviews, discussions about expanding digital access stood out as an area where most libraries were looking ahead. The fact that a number of libraries were looking beyond traditional library catalog discovery, and brought up concepts of access such as computational access, seem to indicate a need to not "silo" Federal documents in bringing them to users.

Recommendations

1. GPO should undertake a pilot project with one or a small number of libraries to work out the details of a digital deposit service, in particular a service where GPO facilitates the deposit of digital files into the digital repositories of FDLP libraries. The pilot would:
 - a. Use the findings from this group's interviews toward the project.
 - b. Explore delivery issues (push vs pull, file types, etc.) that were brought to light by this group's interviews.
 - c. Draft a high-level workflow for notification and push/pull delivery mechanisms.
 - d. Document the necessary training requirements to support delivery mechanisms.
 - e. Identify opportunities for technical solutions or enhancements to support delivery mechanisms.
 - f. Report back to the Depository Library Council on the findings and outcomes of the processes explored in the pilot project.
2. GPO should take steps to outline a vision of how policy, coordination, and support for digital deposit could fit in with current and future models of the FDLP.
3. The Final Report of the Digital Deposit Working Group should be presented to the recently announced Task Force on a Digital FDLP.

