**Classification of the End-of-Term Archive:
Extending Collection Development Practices
to Web Archives**

**Findings of the Web Archive Survey
of Federal Depository Libraries**

December 2010

Revised: January 2011*

Kathleen Murray
kathleen.murray@unt.edu

University of North Texas
UNT Libraries
1155 Union Circle #305190
Denton, TX 76203-5017

Revision Note: Minor text corrections were made.

Contents

Introduction

Project Overview

The Classification of the End-of-Term Archive (EOT Archive) project builds on a previous project[1] conducted collaboratively by the Library of Congress, the US Government Printing Office (GPO), the Internet Archive, the University of North Texas (UNT) Libraries, and the California Digital Library. That project captured the entirety of the federal government's public Web presence before and after the 2009 change in presidential administrations. The result is an approximately 16 terabyte Web archive of government information that is replicated in repositories at the collaborating organizations.

As Web archives become more available and accessible, many libraries will be collecting materials from these important information repositories. Librarians will need the capability to identify and select materials in accord with collection development policies. Additionally, libraries will need to characterize these materials using common metrics; however, such metrics are not established for Web archives, making it difficult for librarians to communicate the scope and value of these materials to administrators.

The *Classification of the End-of-Term Web Archive* project utilizes the EOT Archive to investigate innovative solutions to address these needs. Research is being conducted concurrently from December 2009 through November 2011 in two work areas:

1. *EOT Archive Classification* The materials in the EOT Archive will be classified according to the Superintendent of Documents (SuDocs) Classification Numbering System. Classifying government information in accordance with SuDocs will allow government information librarians to use existing collection development policies to select materials from the archive.

2. *Web Archive Metrics* A set of metrics for materials in Web archives will be identified. These will enable characterization of materials in Web archives in units of measurement more familiar to libraries and their administrations. This report on the survey of libraries that are part of the Federal Depository Library Program (FDLP) is in this work area.

Background

Findings from the initial focus group with the project's 10 Subject Matter Experts (SMEs) regarding metrics for Web archives suggested that libraries would be primarily interested in *accessing* materials in Web archives, versus *acquiring* materials for preservation and local hosting. The metrics libraries will require from Web archive providers will be different for these two service models, therefore gaining a deeper understanding of the demand for each model within the library community was of interest to the research in the Web Archive Metrics area of this project.

---

[1] *Library Partnership Preserves End-of-Term Government Web Sites* (2008, August 14); [http://www.loc.gov/today/pr/2008/08-139.html]

Data regarding the interest of FDLP libraries in "receiving deposit digital files of online publications" is available from the 2009 Biennial survey of FDLP libraries conducted by the GPO.[2] The Biennial survey data indicated that 37% of responding libraries were interested in receiving digital files on deposit. Since the End-of-Term Web Archive contains digital Web-published government information, the survey data provides some indication of the interest depository libraries might have in acquiring versus accessing the materials in the EOT Archive.

A brief online survey was conducted to assess libraries' interests in acquiring versus accessing materials in Web archives, as well as to estimate their capability to support acquisition services, such as preservation, hosting, and user access. Additionally, the relationships between three demographic characteristics (depository type, library type, and library size) and libraries' interests and capabilities were measured. The survey was conducted from September 14, 2010 – October 1, 2010. It was sent to 1225 FDLP libraries and a total of 414 libraries (34%) submitted responses.

A summary of the key findings follows and the remainder of the document includes the detailed analysis of the survey responses. The survey questionnaire is in Appendix A and the survey responses are reported in Appendix B.

Summary of Findings

Materials of Interest

Libraries are most interested in selecting databases and individual publications from Web archives (Table 1). They are least interested in selecting wiki and blog content for their collections. This order of interest was the same for all library types (academic, public, and government) with the exception that government libraries ranked 'maps' immediately below 'audio recordings'.

Rank	Material Type	*N*	*Mdn*
1	Databases	407	6
1	Individual Publications	411	6
2	Websites	408	5
2	Maps	404	5
3	Videos	406	4
3	Audio Recordings	405	4
4	Wikis	403	2
4	Blogs	398	2

Note: Response range: 1 = Not interested; 7 = Extremely interested

Table 1. Material Types of Interest by Rank Order

Significantly more regional depository libraries (*n* = 15; 71%) than would be expected had a higher than average interest in selecting maps for their government information collections. This high level of interest is shared by only 32% (*n* = 121) of selective depository libraries.

---

[2]FDLP Desktop. Needs Assessment/Biennial Survey (2009). [http://www.fdlp.gov/home/repository/cat_view/72-about-the-fdlp/283-customer-relations-program/286-needs-assessmentbiennial-survey-oct-2009-feb-2010]

Larger libraries were significantly related to higher interests in individual publications and maps, databases, websites, and wikis. Conversely, smaller libraries were related to lower interest ratings for these material types.

Likelihood to Access and Acquire Materials from Web Archives

The survey results confirmed what the metrics focus group findings had suggested: Libraries are decidedly more likely to access materials (*Mdn* = 6) than to acquire materials (*Mdn* = 2) for their collections from Web archives at trusted institutions (Figure 1).

(*N* = 411; *Mdn* = 6)   (*N* = 410; *Mdn* = 2)
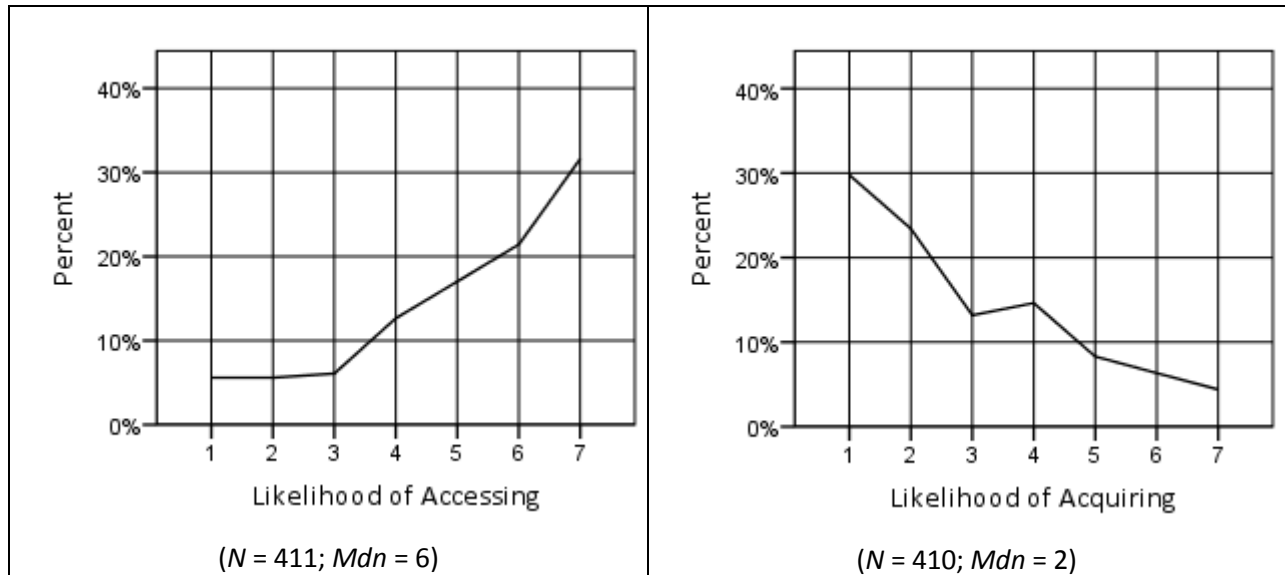
Note: Response range: 1 = Extremely unlikely; 7 = Extremely likely

Figure 1. Likelihood of Accessing versus Acquiring Materials from Web Archives

Estimation of the likelihood of acquisition depends on where the cutoff in response values is made (Table 2). For example, if the estimate includes likelihood scores of 6 and 7, the estimate would be 11% of libraries. Likewise, if the estimate includes scores of 4 through 7, the estimate would be 34%, which is close to the 37% of libraries that indicated an interest in receiving digital files on deposit on the GPO's 2009 Biennial survey of depository libraries.

Likelihood	%	Cumulative %
7	4.4	4.4
6	6.3	10.7
5	8.3	19.0
4	14.6	33.6
3	13.2	46.8
2	23.4	70.2
1	29.8	100.0

Note: Response range: 1 = Extremely unlikely; 7 = Extremely likely

Table 2. Cumulative Likelihood of Acquiring Materials (*N* = 410)

There was no significant association between library type and the likelihood of a library to provide access from their library to a Web archive at a trusted institution or the likelihood to acquire materials from a Web archive. Depository type was significantly associated with the likelihood of acquiring materials; regional depository libraries were more likely to acquire materials from a Web archive for their collections.

Library size was significantly correlated to the likelihood of a library to provide access to materials. Larger libraries were significantly more likely to provide access to Web archives, while smaller libraries were less likely. However, library size was not significantly correlated to the likelihood of a library to acquire materials from Web archives at trusted institutions. This suggests that there is no relationship between library size and the likelihood that they will acquire materials from Web archives.

Interest in Preservation and Hosting

Almost equal numbers of libraries were motivated to acquire materials from Web archives for long-term preservation (*n* = 150) as to acquire materials in order to host them for user access (*n* = 151). Seventy-five libraries (18%) were motivated by both reasons, preservation and hosting. Of these 75 libraries, most (*n* = 60; 80%) were academic libraries and were either large (*n* = 30; 40%) or medium (*n* = 29; 39%) in size (see Figure 2).
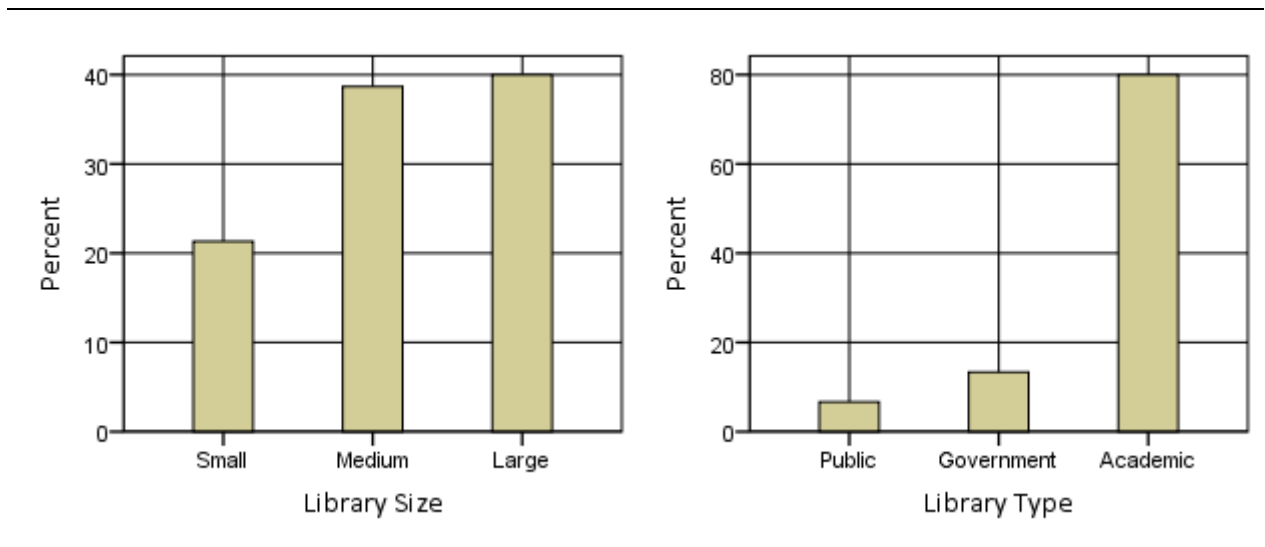
Figure 2. Interest in Preservation and Hosting by Library Size and Type (*n* = 75)

There were no significant relationships between depository types, library types, and library sizes and a motivation to acquire materials for hosting. However, there were significant associations between each of these demographic characteristics and a motivation to acquire materials for preservation. Most significantly, 76% of regional depository libraries (*n* = 16) were interested in acquiring materials for preservation versus one-third of selective depository libraries (*n* = 134; 34%). Less significantly, fewer public libraries were than expected were motivated to acquire materials for preservation while more large libraries than expected were motivated to do so.

Capability to Preserve & Host Materials

On average (*Mdn* =2), libraries have limited capabilities for either long-term preservation of materials acquired from Web archives or for hosting materials for user access. There was relatively little variance

in capabilities for the seven criteria measured (Table 3). Fully 57% - 64% of all respondents estimated their capabilities for each criterion as either 1 or 2 on a 7-point scale, with 1 equal to "not capable" and 7 equal to "extremely capable".

Criterion				Estimate of Capability *						
Preservation	**N**	**Mean**	**Mdn**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Technology Infrastructure	409	2.72	2.00	131	101	53	48	41	23	12
Technical Support Staff	407	2.57	2.00	143	109	46	43	34	20	12
Preservation Expertise	408	2.65	2.00	137	102	51	43	43	23	9
Hosting for User Access	**N**	**Mean**	**Mdn**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Technology Infrastructure	409	2.73	2	137	95	51	47	39	27	13
Technical Support Staff	409	2.62	2	148	91	54	47	34	25	10
Programming Resources	407	2.42	2	154	104	52	46	31	14	6
User Interface Design Resources	403	2.44	2	155	103	47	38	37	18	5

* <u>Note</u>: Response range: 1 = Not capable; 7 = Extremely capable

Table 3. Estimate of Capability to Support Preservation & Hosting for User Access
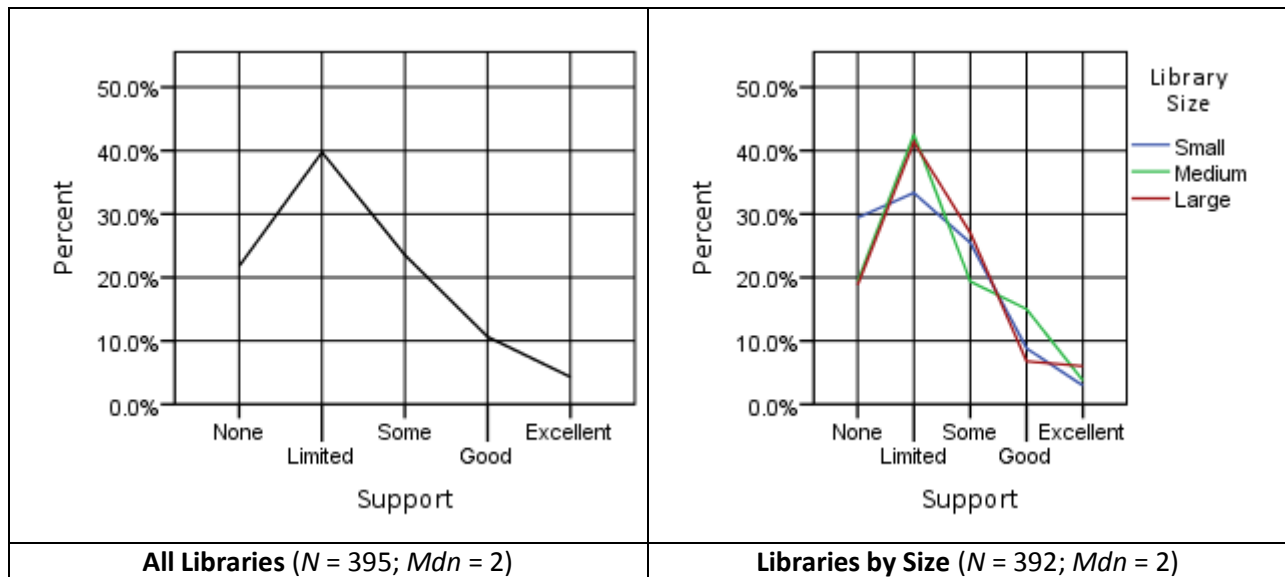
Preservation of Materials
There were significant relationships between both library types (i.e., academic, public, and government) and library sizes for the three criterions of preservation capability: technology infrastructure, technical support staff, and preservation expertise. For each criterion, more public libraries than would be expected estimated their capability *lower than average*. In particular, public libraries' estimates of their preservation expertise was very significantly lower than would be expected. Additionally, smaller libraries were related to lower estimates of capabilities while larger libraries were related to higher estimates.

Hosting Materials for User Access
There was a significant association between library types and one criterion of hosting materials: technical support staff. Public libraries were *less likely than expected* to have the technical staff to support hosting materials for user access. There was also a significant and positive correlation between library sizes and each of the four criteria: technology infrastructure, technical support staff, programmers, and user interface designers. This suggests that smaller libraries estimate their capabilities at lower levels while larger libraries tend to have higher capability estimates.

Organizational Support for Acquisition of Materials from Web Archives

Libraries' preferences for accessing Web archives is reinforced by their estimates of the support they are likely to receive within their organizations for acquiring materials from Web archives. Just over 60% of libraries of all sizes (*N* = 395) indicated they had either no support (*n* = 86; 22%) or limited support (*n* = 157; 40%) for the acquisition of materials from Web archives (Figure 3). Twenty-four percent (*n* = 93) had some support, while 11% (*n* = 42) had good support and only 4% (*n*= 17) enjoyed excellent support. Not surprisingly, the likelihood of libraries to acquire materials from a Web archive was highly correlated to the support they had for acquisition within their organizations.

| **All Libraries** (*N* = 395; *Mdn* = 2) | **Libraries by Size** (*N* = 392; *Mdn* = 2) |

Figure 3. Support for Acquisition of Materials from Web Archives (*N* = 395; *Mdn* = 2)

There was a significant association between library types and support within libraries for acquisition of materials from Web archives. More public libraries than would be expected (*n* = 18; 35%) had no support within their libraries for acquisition of materials from Web archives. When the number of libraries with no support is combined with those having limited support, fully 78% of public libraries are represented. This percentage compares to 61% of academic libraries and 45% of government libraries (Table 4).

	Academic		**Public**		**Government**	
Support	*n*	**%**	*n*	**%**	*n*	**%**
None	59	20	18	35	8	19
Limited	123	41	22	43	11	26
Subtotal		*61*		*78*		*45*
Some, good, or excellent	117	39	11	22	23	55
Total	299	100	51	100	42	100

Table 4. Comparison of Support for Acquisition by Depository Type (*N* = 392)

Methodology

Objective

This objective of the survey was to assess Federal Depository Libraries' interests and capabilities in regard to (a) accessing content and (b) acquiring content from Web archives of government information. The following questions were investigated:

1. What types of materials in Web archives are of interest to FDLP libraries for their collections?
2. What is the relative interest libraries have in accessing versus acquiring[3] government information for their collections from Web archives?
3. What motivates libraries to acquire materials from Web archives: preservation, hosting, or other reasons?
4. For materials libraries acquire from Web archives, do they have the capability to support:
    a. Long-term preservation of the materials?
    b. Hosting and user access to the materials?
5. What level of support are libraries likely to receive within their organizations for the acquisition of materials from Web archives?
6. Are depository type, library type, and library size related to:
    a. Material types of interest to libraries for their government information collections?
    b. Libraries' interest in accessing versus acquiring materials from Web archives?
    c. Libraries' capabilities to support the acquisition of materials from Web archives?
    d. The level of support libraries are likely to have within their organizations for acquisition of materials from Web archives?

Participants

Participants in this study were 414 libraries in the Federal Depository Library Program. Included were 22 Regional Depository Libraries and 392 Selective Depository Libraries.

Library Types

The majority of the participating libraries were academic libraries (n = 316; 76%). Also included were 53 (13%) public libraries, 42 (10%) state and federal government libraries, two special libraries, and one service academy library. The representation by library type is fairly comparable to the representation by library type within the FDLP, although there was a slightly higher percentage of academic libraries and a slightly lower percentage of public libraries among survey respondents (Table 5).

---

[3] Acquisition of materials from Web archives was defined in the survey as a library providing both servers to host the materials as well as user access to the materials on the servers.

	FDLP		Participants	
Library Type	**#**	**%**	**#**	**%**
Academic	861	71	316	76
Public	209	17	53	13
State & Federal Government	133	11	42	10
Other	18	1	3	1
Total	1221	100	414	100

Table 5. FDLP Libraries by Type
(as of November 29, 2010)

Library Sizes

Small, medium, and large libraries were represented, with medium-sized libraries comprising the largest percentage. The participating libraries' sizes were very representative of the libraries in the FDLP (Table 6).

		FDLP		Participants	
Size	**Volumes in Library**	**#**	**%**	**#**	**%**
Small	< 250,000	333	27	106	26
Medium	250,000 – 1,000,000	515	42	170	41
Large	> 1,000,000	373	31	138	33
Total		1221	100	414	100

Table 6. Participating Libraries by Size
(as of November 29, 2010)

Web Archive Experience

Question one of the survey asked respondents to indicate their experience with Web archives on a scale of 1 (Novice) to 7 (Expert). On average, respondents' indicated they had limited experience with Web archives (*Mdn* = 2). Only four respondents (1%) indicated they were experts, while 151 (38%) rated themselves as novices (Figure 4).
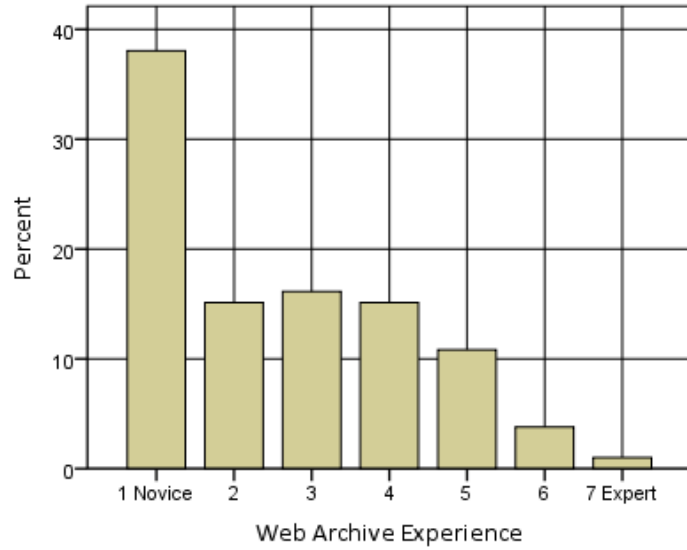
* Note: Response range: 1 = Novice; 7 = Expert

Figure 4. Web Archive Experience (*N* = 397)

Likewise, only 22 respondents were included in the list of Archive-It Partners. Archive-It is "a subscription service [of the Internet Archive] that allows institutions to build and preserve collections of born digital content."[4] As a group, respondents from libraries that were Archive-It partners were somewhat more experienced with Web archives (*Mdn* of 4 versus 2) (Figure 5).
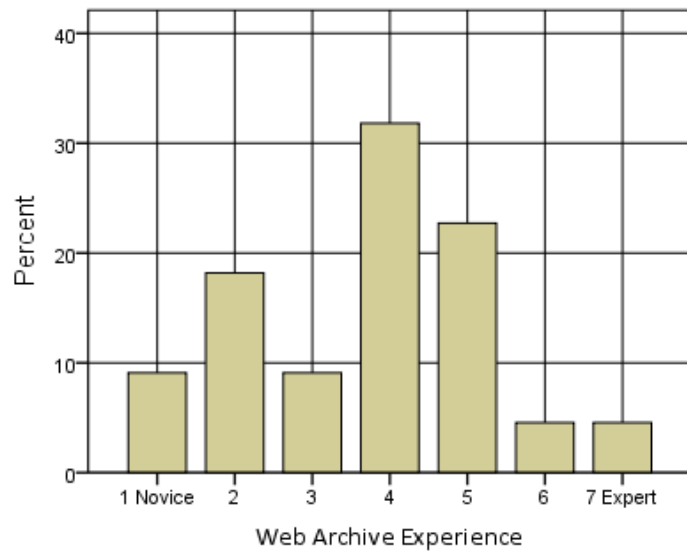
* Note: Response range: 1 = Novice; 7 = Expert

Figure 5. Archive-It Partners' Web Archive Experience (*n* = 22)

---

[4] Archive-It. [http://www.archive-it.org/]

Data Collection

The survey instrument was created, tested, and administered using Zoomerang[5], an online survey development and hosting Website. The instrument consisted of nine questions (Appendix A). Prior to initial deployment, members of the project team reviewed the survey instrument for clarity and the project SMEs reviewed it for content validity. The survey instrument was revised based on the review comments.

Demographic data for the 1225 Federal Depository Libraries were downloaded from the Federal Depository Library Directory (FDLD).[6] Demographic data included contact names, as well as depository types (Regional or Selective), library types (Academic General, Academic -Community College, Academic –Law, Federal Agency, Federal Court, Public, Service Academy, Highest State Court, State Library, Special Library), and library sizes as measured by the number of volumes in the entire library (Small < 250,000 volumes, Medium 250,000 – 1,000,000 volumes, and Large > 1,000,000 volumes). Contact email addresses for the libraries were obtained from Documents Data Miner 2[7]. When possible, contact names and email addresses were verified and updated prior to initial deployment of the survey.

All libraries were emailed requests to respond to the survey. Seventy email addresses bounced, of which 30 were corrected and included in the survey solicitation. Non-respondents were emailed one or two reminder messages. Survey responses were collected from September 14, 2010 through October 1, 2010. A total of 414 responses (34%) were received.

Survey response data was exported from Zoomerang to Microsoft Excel and merged with the demographic data for each library that was downloaded from the FDLD. The final data set was exported to a CSV-formatted file for analysis.

Data Analysis

Data was analyzed using IBM SPSS Statistics Version 19. Descriptive statistics (frequency counts and percentages) as well as text-responses for questions 5 and 9 are reported in Appendix B. Measures of central tendency (medians and means) were calculated for most questions and, along with the descriptive statistics, informed the reclassification of some responses into fewer to meet the expected counts in chi square contingency tables.

The relationships between the three demographic variables and responses to several survey questions were analyzed using either chi square or Spearman's rho:

   1.   Depository type (Regional and Selective),

---

[5] Zoomerang Online Surveys and Polls. [http://www.zoomerang.com/]

[6] Federal Depository Library Directory (FDLD). [http://catalog.gpo.gov/fdlpdir/FDLPdir.jsp?mode=6]

[7] Documents Data Miner 2 is a library management system for United States Government Documents provided in partnership by the Wichita State University Libraries and the WSU Computing Center. [http://govdoc.wichita.edu/ddm2/gdocframes.asp]

2. Library type[8] (Academic, Public, and Government), and
3. Library size (Small < 250,000 volumes; Medium 250,000 – 1,000,000 volumes; and Large > 1,000,000 volumes).

Specifically, chi square measured the association between both depository type and library type, while Spearman rho measured the relationship between library size and responses to:

- Question 2: Interest in selecting various material types from Web archives;
- Question 3: Likelihood of providing access to materials from Web archives at trusted institutions;
- Question 4: Likelihood of acquiring materials from Web archives at trusted institutions;
- Question 5: Motivation for acquiring materials from a Web archive;
- Question 6: Capability to support long-term preservation of materials acquired from Web archives;
- Question 7: Capability to support hosting and user access; and
- Question 8: Organizational support for acquiring materials from a Web archive.

Spearman rho also measured the relationship between Web archive experience (question 1) and responses to:

- Question 3: Likelihood of providing access to materials from Web archives at trusted institutions; and
- Question 4: Likelihood of acquiring materials from Web archives at trusted institutions;

In all cases, the required significance level was α = .05. A standardized residual greater than 2.0 determined contributions to significant results of cells within chi square contingency tables.

Lastly, questions 5 and 9 allowed users to submit free-form text responses. These responses were content analyzed to group them thematically.

---

[8] For the analysis, the 10 library types from the FDLD were consolidated into three types: Academic, Public, and Government. This resulted in three respondents being excluded from each analysis by library type, one service academy and two special libraries.

Findings

The key findings are summarized in the Introduction section of this report. Appendix B lists the responses, including frequencies and percentages, for each question in the survey. The detailed findings from the statistical and qualitative analyses for each question are included in this section of the report. Findings for each question are reported first for all respondents and then for the statistical analyses of the demographic variables (depository type, library type, and library size).

Interest in Materials by Type

Question 2.
Web archives contain a range of material types. On a scale of 1 ("Not interested") to 7 ("Extremely interested"), estimate your interest in selecting the following material types for your government information collection: Audio Recordings, Blogs, Databases, Individual Publications, Maps, Video Recordings, Websites, and Wikis.

All Respondents

Libraries are most interested in selecting databases and individual publications from Web archives. They are least interested in selecting wiki and blog content for their collections. Table 7 lists libraries' interests in selecting each material type in rank order based on the median score for each type.

Rank	Material Type	N	Mdn
1	Databases	407	6
1	Individual Publications	411	6
2	Websites	408	5
2	Maps	404	5
3	Videos	406	4
3	Audio Recordings	405	4
4	Wikis	403	2
4	Blogs	398	2

Note: Response range: 1 = Not interested; 7 = Extremely interested

Table 7. Rank Order of Interest in Material Types

Depository Type

Chi square measured the association between depository type and libraries' interest in selecting various material types from Web archives. To meet the expected values for the chi square analysis the seven response values (1 to 7) for each material type were categorized into 3 groups: (1) Low interest or below the median,(2) Average interest or near the median, and (3) High interest or above the median.

Significant chi square values were found for maps and individual publications (Table 8). More regional depository libraries (*n* = 15; 71%) than would be expected had a higher than average interest in selecting maps for their government information collections. This high level of interest is shared by only 32% (*n* = 121) of selective depository libraries. Additionally, only one regional depository library had a low interest in selecting maps.

Material Type	N	Actual X^2	Critical X^2
Maps [a]	404	15.471 ***	13.816
Individual publications [b]	414	7.043 *	5.991

Note: *df* = 2; α = .05; * *p* < .05; *** *p* < .001
[a] Two cells contained standardized residuals greater than 2.0.
[b] No cells contained standardized residuals greater than 2.0.

Table 8. Association of Interest in Material Types and Depository Type

While depository type was significantly associated with libraries' interests in selecting individual publications, no cells in the contingency table contained standardized residuals greater than an absolute value of 2.0, which would have provided an indication of what depository types at what interest levels were contributing substantially to the significant finding. The analysis does indicate that a higher percentage (*n* = 11; 50%) of regional depository libraries than were expected had a higher than average interest in selecting individual publications for their government information collections. This compares to 28% (*n* = 107) of the selective depository libraries.

Library Type

The level of interest by material types reported in the aggregate of all libraries is most reflective of academic libraries' interests (Table 9). However, the low interest score for blogs and wikis was the same for all library types.

	Academic		Public		Government	
Material Type	*N*	*Mdn*	*N*	*Mdn*	*N*	*Mdn*
Databases	311	6	53	5	40	5
Individual Publications	314	6	53	5	41	5
Websites	311	5	53	4	41	4
Maps	308	5	53	4	40	2.5
Videos	310	4	53	3	40	3
Audio Recordings	309	4	53	3	40	3
Wikis	308	2	51	2	41	2
Blogs	304	2	50	2	41	2

Note: Response range: 1 = Not interested; 7 = Extremely interested

Table 9. Interest in Material Types by Library Type

Chi square measured the association between library types and libraries' interests in selecting various material types from Web archives. Significant chi square values were found for six material types: databases, videos, websites, individual publications, maps, and audio recordings (Table 10). The level of interest in blogs and wikis was not significantly associated with library types.

Material Type	N	Actual X^2	Critical X^2
Databases [a]	404	19.677 **	13.277
Videos [a]	403	16.087 **	13.277
Websites [a]	405	13.416 **	13.277
Individual publications [b]	408	13.413 **	13.277
Maps [a]	401	13.184 *	9.488
Audio recordings [c]	402	9.740 *	9.488

Note: df = 4; α = .05; * p < .05; ** p < .01
[a] One cell contained standardized residuals greater than 2.0.
[b] Two cells contained standardized residuals greater than 2.0.
[c] No cells contained standardized residuals greater than 2.0.

Table 10. Association of Interest in Material Types and Library Types

Government libraries contributed to the significant associations found between library types and three material types: maps, databases, and websites. In each case, significantly more government libraries than would be expected had a *lower than average interest* in selecting these three material types for their collections (Table 11).

Material Type	n	%
Maps	22	55
Databases	19	48
Websites	16	39

Table 11. Government Libraries with Low Interest in Material Type

While no cells in the contingency table for audio materials had standardized residuals greater than 2.0, the standardized residual for one government library cell was highest at 1.9. This suggests that more government libraries (n = 22; 55%) than would be expected had a *lower than average interest* in selecting audio materials for their collections.

Public libraries were primary contributors to the significant associations found between library types and databases, videos, and individual publications. Significantly more public libraries (n = 25; 47%) had a *lower than average interest* in selecting databases for their collections than would be expected. And significantly *fewer* public libraries than would be expected had *higher than average interest* in selecting video materials (n = 12; 23%) as well as individual publications (n = 7; 13%).

Library Size

Spearman rho correlation coefficients (r_s) measured the relationships between library sizes and their ratings of interest in selecting eight material types from Web archives for their collections. There were no significant relationships between library sizes and libraries' interest in selecting blogs, videos, and audio recordings. Five of the material types were significantly and positively related to the size of libraries (Table 12).

Material Type	N	r_s
Individual Publications	411	0.202 ***
Maps	404	0.192 ***
Databases	407	0.132 **
Websites	408	0.109 *
Wikis	403	0.103 *

Note: * $p < .05$; ** $p < .01$; *** $p < .001$

Table 12. Correlations between Library Sizes and Interest in Selecting Material Types

Larger libraries were very significantly related to higher interest ratings for individual publications and maps, as well as less significantly related to databases, websites, and wikis. Conversely, smaller libraries were related to lower interest ratings for these five material types.
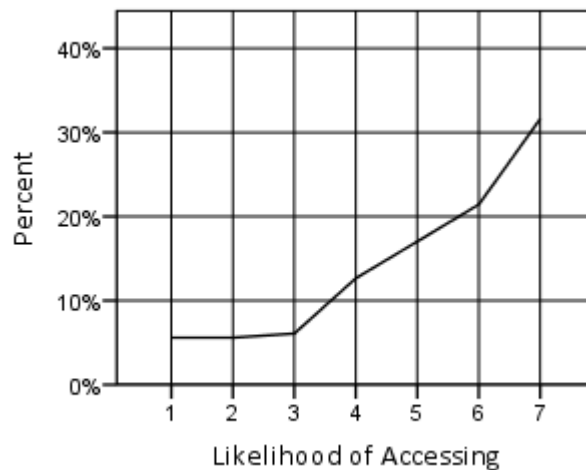
Likelihood of Providing Access to Web Archives

Question 3.
On a scale of 1 ("Extremely unlikely") to 7 ("Extremely likely"), how likely do you think your library would be to provide access to government information located on a Web archive at an institution that you trust?

All Respondents

With a median score of 6 (range 1-7), libraries indicated they were quite likely to provide access to government information on Web archives at trusted institutions. Fully 32% (n = 132) indicated they were extremely likely to provide access (Figure 6).

Note: Response range: 1 = Extremely unlikely; 7 = Extremely likely

Figure 6. Likelihood of Providing Access (N = 411; *Mdn* = 6)

Depository Type & Library Type

Chi square measured the association between both depository type and library type for the likelihood of libraries to provide access to Web archives at trusted institutions. To meet the expected values for the chi square analysis the seven response values for likelihood of accessing Web archives were categorized into 3 groups: (1) Low interest or below the median,(2) Average interest or near the median, and (3) High interest or above the median.

There was no significant association between depository type and the likelihood of a library to provide access from their library to a Web archive at a trusted institution (X^2 (2, *N* = 411) = 2.803, p .05). Likewise no significant association was found between library type and the likelihood of a library to provide access (X^2 (4, *N* = 408) = 8.454, p .05).

Library Size

Spearman rho correlation coefficients (r_S) measured the relationships between library sizes and their ratings of likelihood of providing access to Web archives. Library size was significantly and positively correlated to the likelihood of a library to provide access to Web archives at trusted institutions (r_S = 0.105 (*N* = 411), p .05). Larger libraries indicated they were more likely to provide access to Web archives, while smaller libraries were less likely.

Web Archive Experience

Spearman rho correlation coefficients (r_S) measured the relationships between respondents' experience with Web archives and their ratings of the likelihood of their libraries' to provide access to Web archives at a trusted institution. Respondents with higher levels of experience were significantly more likely to indicate a higher likelihood that their libraries would provide access to Web archives (r_S = 0.348 (*N* = 407), p .001).
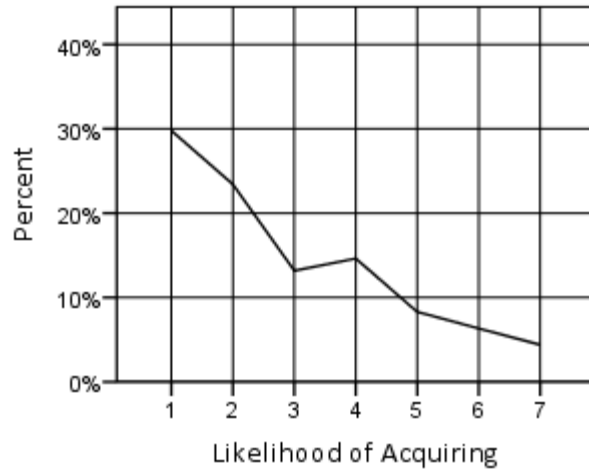
Likelihood of Acquiring Materials from Web Archives

> Question 4.
> Acquisition of materials from Web archives means that your library would provide both servers to host the materials and user access to them. On a scale of 1 ("Extremely unlikely") to 7 ("Extremely likely"), how likely do you think your library would be to acquire materials for your government information collection from a Web archive at an institution you trust?

All Respondents

With a median score of 2 (range 1-7), libraries indicated they were quite unlikely to acquire materials for their government information collections from Web archives at trusted institutions. Fully 30% (*n* = 122) indicated they were extremely unlikely to acquire materials (Figure 7).

Note: Response range: 1 = Extremely unlikely; 7 = Extremely likely

Figure 7. Likelihood of Acquiring Materials (*N* = 410; *Mdn* = 2)

Depository Type & Library Type

Chi square measured the association between depository type and library type for likelihood of libraries to acquire materials from Web archives at trusted institutions. To meet the expected values for the chi square analysis the seven response values were categorized into 3 groups: (1) Low interest or below the median,(2) Average interest or near the median, and (3) High interest or above the median.

There was a significant association between depository type and the likelihood of a library to acquire materials from a Web archive at a trusted institution (X^2 (2, *N* = 410) = 6.736, p .05). While none of the cells in the contingency table had standardized residuals greater than 2, one cell did have a residual of 1.9. More regional depository libraries than would be expected (*n* = 12; 57%) indicated they were more likely than average to acquire materials from a Web archive for their collections. There was no significant association between library type and the likelihood of a library to acquire materials from a Web archive at a trusted institution (X^2 (4, *N* = 407) = 6.814, p .05).

Library Size

The Spearman rho correlation coefficient (r_S) measured the relationships between library sizes and their ratings of likelihood of acquiring materials from Web archives. No significant correlation was found suggesting that there is no significant relationship between library size and the likelihood that a library will acquire materials from Web archives (r_S = 0.069 (*N* = 410), p .05).

Web Archive Experience

Spearman rho correlation coefficient (r_S) measured the relationships between respondents' Web archive experience and their ratings of the likelihood of acquiring materials from Web archives at a trusted institution. As with the likelihood of libraries providing access to Web archives, respondents with higher levels of experience were significantly more likely to indicate a higher likelihood that their libraries would acquire materials from a Web archive (r_S = 0.302 (*N* = 394), p .001).

Motivation for Acquiring Materials

Question 5.
Indicate your library's primary motivation(s) for acquiring materials from a Web archive:
(Select all that apply.)
- Long-term preservation of the materials
- Hosting the materials for user-access
- Not interested in acquiring materials
- Other, please specify

All Respondents

Almost equal numbers of libraries were motivated to acquire materials for long-term preservation ($n =$ 150) as were motivated to acquire materials to host for user access ($n = 151$). A similar number of libraries ($n = 156$) indicated they were not interested in acquiring materials (Table 13).

Motivation	n	%
Long-term preservation	150	36
Hosting for user access	151	36
Not interested	156	38
Other	40	10

Table 13. Primary Motivation for Acquiring Materials ($N = 410$)

Seventy-five libraries (18%) were motivated both by long-term preservation and hosting for user access. Of these libraries, 60 (80%) were academic libraries, 30 (40%) were large libraries, and 29 (39%) were medium libraries (Figure 8).
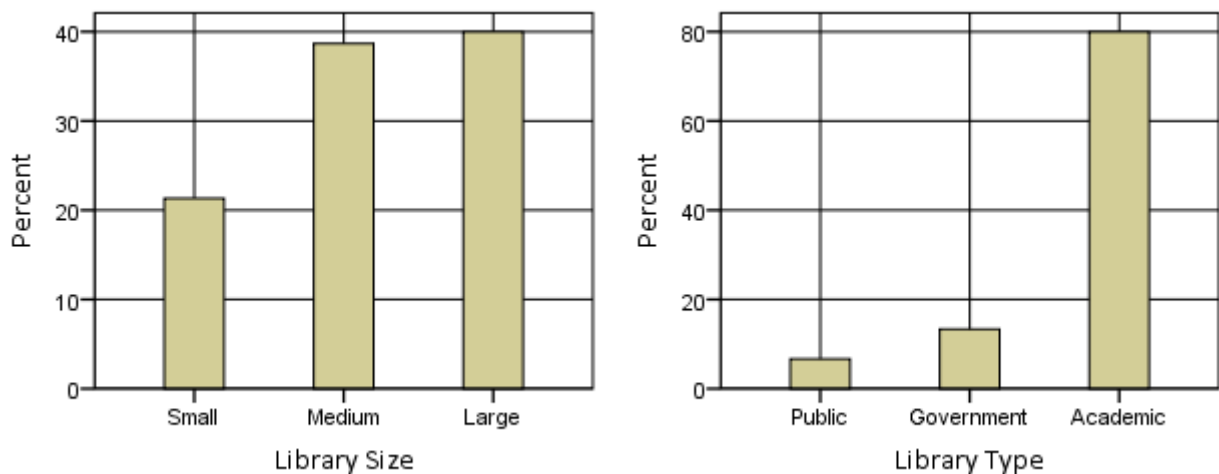
Figure 8. Libraries Motivated to Preserve and Host Materials by Library Size and Type ($n = 75$)

Forty (40) libraries submitted free-form text responses question five. These responses were content analyzed to group them thematically. Sixteen responses related to five primary motivations for acquiring materials from a Web archive. These are listed in rank order in Table 14.

#	Motivation
5	Build unique collections of regional, local, or topical interest
4	Ensure free and open access to materials
4	Space concerns; replace physical materials
2	Curriculum support
1	Share content and metadata; participating in digital FDLP

Table 14.Other Motivations for Acquiring Materials (*n* = 16)

Twenty-four additional responses were not directly related to motivations for acquiring materials from a Web archive. These responses are listed in three categories in Table 15.

#	Other Response Categories
12	Not certain of question or concepts; not considered
6	No server space available for acquisition
6	Provide access but not hosting

Table 15. Additional Responses to Motivation (*n* = 24)

Depository Type & Library Type

Chi square measured the association between both depository type and library type for ratings of libraries' primary motivation for acquiring materials from a Web archive. There was no significant association between depository type and libraries who indicated their primary motivation was hosting materials for user access. There was a highly significant association between depository type and libraries whose primary motivation for acquisition was preservation (X^2 (1, *N* = 410) = 14.964, *p* .001). Three-fourths of regional depository libraries (*n* = 16; 76%) were interested in acquiring materials for preservation versus one-third of selective depository libraries (*n* = 134; 34%) (Table 16).

	All		Regional		Selective	
Motivation	*n*	%	*n*	%	*n*	%
Long-term preservation	150	36	16	76	134	34
Hosting for user access	151	36	9	43	142	37
Not interested	156	38	3	14	153	39
Other	41	10	1	-	40	10

Table 16. Primary Motivation for Acquiring Materials (*N* = 410)

There were no significant associations between library types and libraries motivated by hosting materials for user access. However, there was a somewhat significant association between library types and libraries that were primarily motivated by preservation (X^2 (2, *N* = 407) = 8.518, *p* .05). Public libraries were *less likely* than expected to be motivated to acquire materials for preservation.

Library Size

Chi square measured the association between library sizes and libraries' motivations for acquiring materials from Web archives. Similarly to the findings for library types, there were no significant associations between library sizes and libraries that were primarily motivated by hosting acquired materials for user access. However, there was a somewhat significant association between library sizes and libraries that were primarily motivated by preservation (X^2 (2, *N* = 407) = 7.139, *p* .05). While small, medium, and large libraries did not have standard residuals that met the standard of greater than 2.0, more large libraries (residual = 1.7) than expected were motivated by preservation.

Capability to Support Long-term Preservation

Question 6.
On a scale of 1 ("Not capable") to 7 ("Extremely capable"), estimate the capability of your library/organization to support long-term preservation of materials acquired from Web archives in regard to:
- Technical Infrastructure
- Technical Support Staff
- Preservation Expertise

Overall Capability

On average (*Mdn* =2), libraries are not capable of supporting long-term preservation of materials acquired from Web archives (Table 17). There is relatively little variance in capabilities for the three criteria measured (Figure 9).

Criterion				Estimate of Capability *						
	N	Mean	*Mdn*	1	2	3	4	5	6	7
Technology Infrastructure	409	2.72	2.00	131	101	53	48	41	23	12
Technical Support Staff	407	2.57	2.00	143	109	46	43	34	20	12
Preservation Expertise	408	2.65	2.00	137	102	51	43	43	23	9

<u>Note</u>: * Response range: 1 = Not capable; 7 = Extremely capable

Table 17. Estimate of Capability to Support Preservation
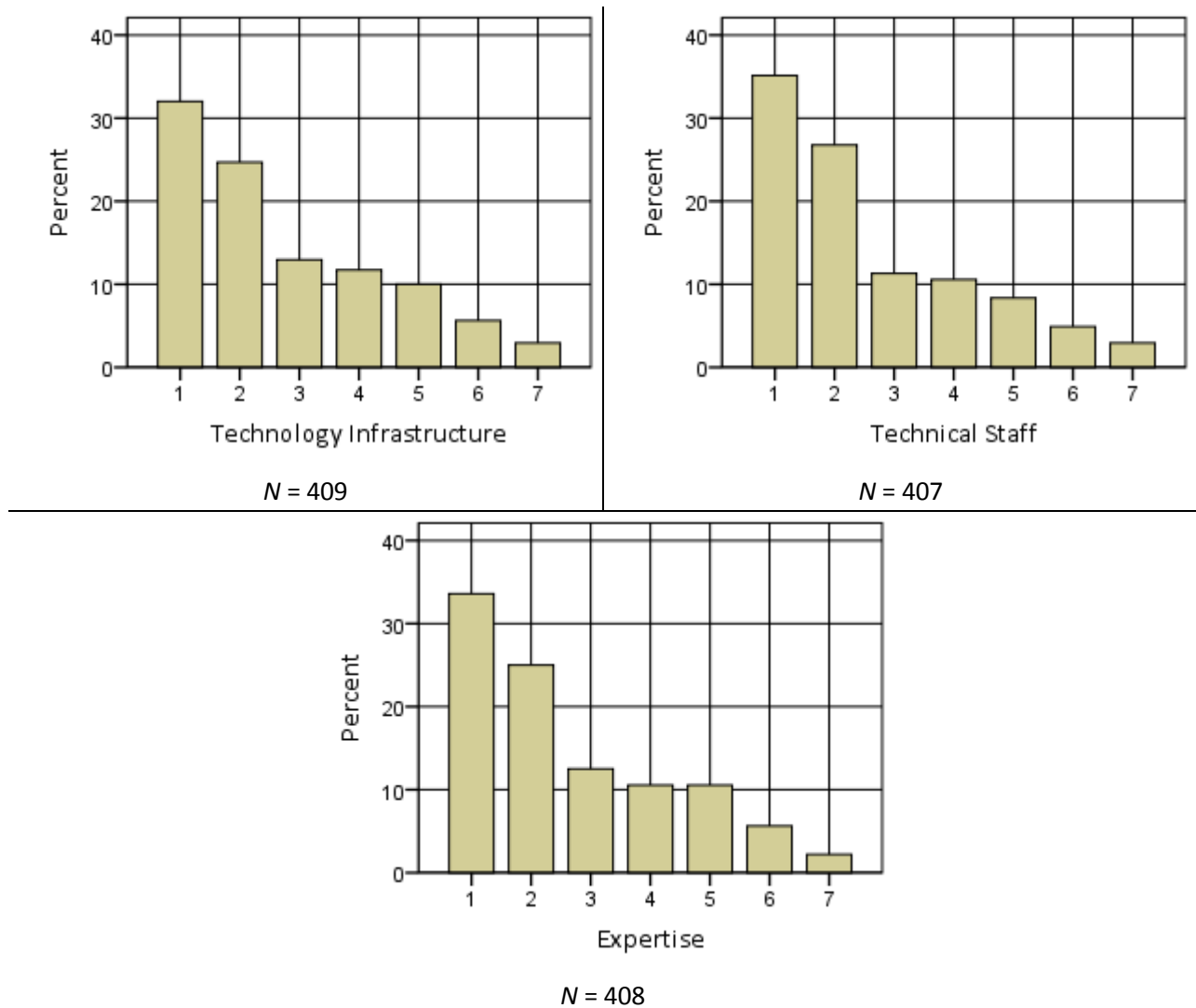
Note: Response range: 1 = Not capable; 7 = Extremely capable

Figure 9. Capability to Support Preservation of Materials

Depository Type

Chi square measured the association between both depository types and library types for estimates of the capability of libraries to support long-term preservation of materials acquired from Web archives. To meet the expected values for the chi square analysis the seven response values (1 = "not capable" to 7 = "extremely capable") were categorized into 3 groups: (1) Low or below the median, (2) Average or near the median, and (3) High or above the median.

There were no significant associations between depository types and two of the criterion for preservation capability: technical infrastructure and technical support staff. There was a somewhat significant association between depository types and preservation expertise (X^2 (2, N = 408) = 6.334, p .05). While neither regional nor selective libraries had standard residuals greater than 2.0, more regional libraries than expected (residual = 1.8; n = 11; 50%) estimated their preservation expertise higher than average.

Library Type

There were significant associations between library types and all three of the criterion for preservation capability (Table 18). In the contingency tables for all three criteria, the identical cell contributed most to the significance (i.e., had standardized residuals greater than 2.0): more public libraries than would be expected estimated their capability for each criterion lower than average. In particular, public libraries' estimate of their preservation expertise was very significantly lower than would be expected.

Criterion	N	Actual X^2	Critical X^2
Preservation Expertise	405	20.521 ***	18.467
Technical Support Staff	404	11.045 *	9.488
Technology Infrastructure	406	10.258 *	9.488

Note: df = 4; α = .05; * p < .05; *** p < .001

Table 18. Asoociation of Capability to Support Preservation of Materials and Library Type

Library Size

The Spearman rho correlation coefficient (r_S) measured the relationships between library sizes and estimates of libraries' capabilities to support long-term preservation of materials acquired from Web archives. There was a significant and positive correlation for each of the three criteria (Table 19). This suggests that smaller libraries estimate their capabilities at lower levels and larger libraries estimate their capabilities at higher levels.

Criterion	N	r_S
Preservation Expertise	409	0.171 **
Technical Support Staff	407	0.157 **
Technology Infrastructure	408	0.145 **

Note: ** p < .01

Table 19. Correlations between Library Sizes and Capability to Support Preservation of Materials

Motivation for Acquisition: Long-term Preservation

The average estimates of their capabilities to preserve materials acquired from Web archives was higher for libraries who indicated in question 5 that their motivation for acquiring materials from Web archives was long-term preservation of the materials. Table 20 compares their mean and median scores for each criterion to the average scores for all respondents.

Criterion	All Respondents			Q5. Motivation = Preservation		
	N	Mean	Mdn	N	Mean	Mdn
Technology Infrastructure	409	2.72	2	148	3.72	4
Technical Support Staff	407	2.57	2	147	3.49	3
Preservation Expertise	408	2.65	2	147	3.44	3

Note: * Response range: 1 = Not capable; 7 = Extremely capable

Table 20. Comparison of Capability to Support Preservation

Libraries that are motivated to acquire Web materials in order to preserve them seem to be in a better position technically to do so. However, as indicated by the medians or 3.00 and 4.00, this is not a strong position.

Capability to Support Hosting and User Access

Question 7.
On a scale of 1 ("Not capable") to 7 ("Extremely capable"), estimate the capability of your library/organization to support hosting and user access for materials acquired from Web archives in regard to:
- Technical Infrastructure
- Technical Support Staff
- Programming Resources
- User Interface Design Resources

Overall Capabilities

As with responses to question 6, on average (*Mdn* =2), libraries are not capable of supporting materials acquired from Web archives for user access (Table 21). There is relatively little variance in capabilities for the four criterions measured (Figure 10).

Criterion				Estimate of Capability *						
	N	Mean	Mdn	1	2	3	4	5	6	7
Technology Infrastructure	409	2.73	2	137	95	51	47	39	27	13
Technical Support Staff	409	2.62	2	148	91	54	47	34	25	10
Programming Resources	407	2.42	2	154	104	52	46	31	14	6
User Interface Design Resources	403	2.44	2	155	103	47	38	37	18	5

<u>Note</u>: * Response range: 1 = Not capable; 7 = Extremely capable

Table 21. Estimate of Capability to Support Hosting for User Access

N = 409   N = 409

N = 407   N = 403
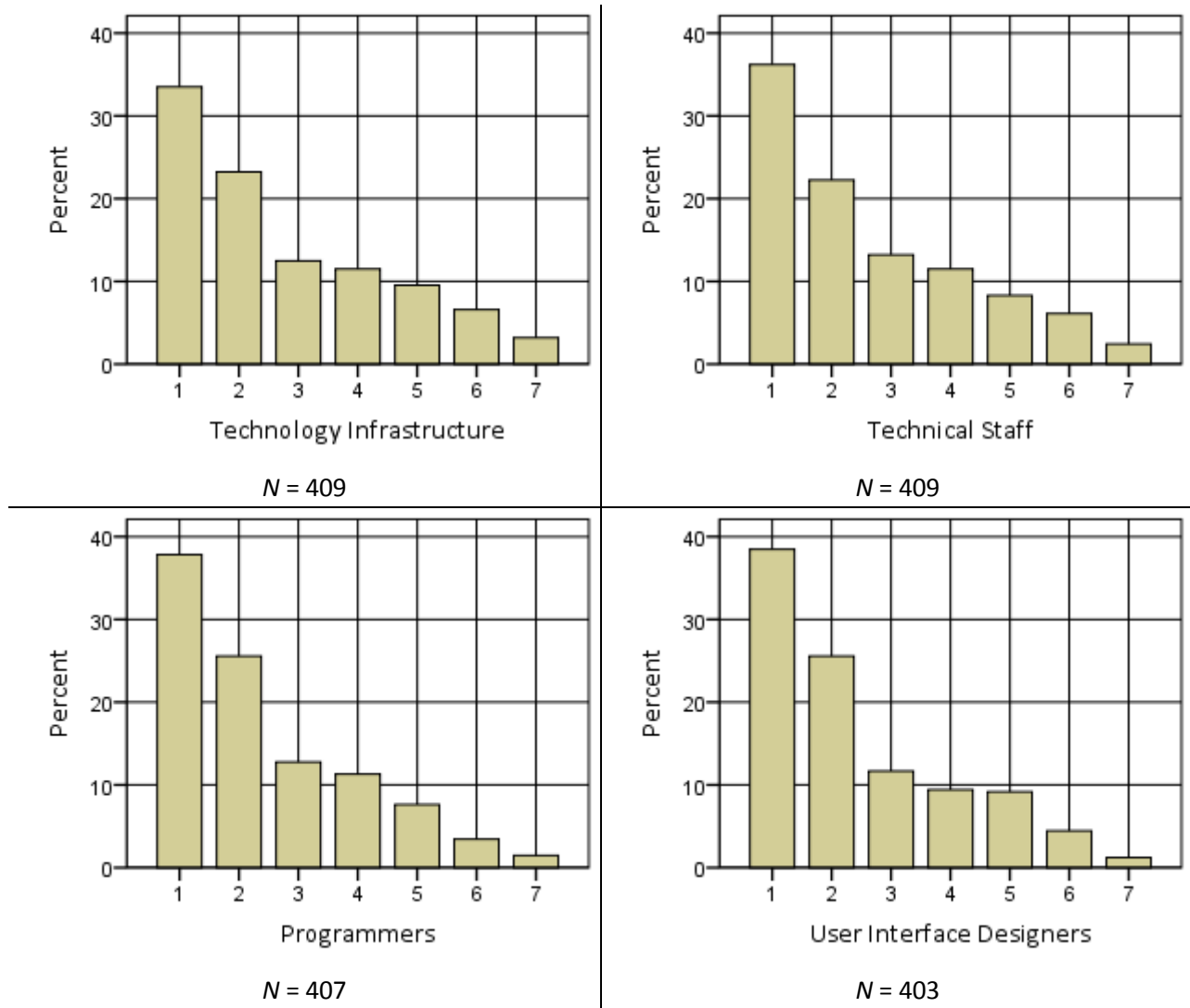
Note: * Response range: 1 = Not capable; 7 = Extremely capable

Figure 10. Capability to Support Preservation of Materials

Depository Type

Chi square measured the association between both depository types and library types for estimates of the capability of libraries to support hosting materials acquired from Web archives for user access. To meet the expected values for the chi square analysis the seven response values (1 = "not capable" to 7 = "extremely capable") were categorized into 3 groups: (1) Low or below the median, (2) Average or near the median, and (3) High or above the median.

There was no significant association between depository types and estimates of their libraries' capabilities for the four criterions of support for hosting materials for user access.

Library Type

There were no significant associations between library types and their estimates of their libraries' capabilities to support hosting and user access for three of the four criteria measured: technical

infrastructure, programmers, and user interface designers. However, there was a somewhat significant association between library types and the fourth criterion: technical staff (X^2 (4, N = 406) = 9.908, p .05). None of the cells in the contingency table had greater than 2.0 standardized residuals. The largest residual (1.8) suggested that public libraries were less likely than expected to have the technical staff to support hosting of materials acquired from Web archives.

Library Size

The Spearman rho correlation coefficient (r_s) measured the relationships between library sizes and their estimates of their libraries' capability to support hosting of materials acquired from Web archives for user access. There was a significant and positive correlation for each of the four criteria (Table 22). Small libraries estimate their capabilities at lower levels and larger libraries estimate their capabilities at higher levels.

Criterion	N	r_s
Technical Infrastructure	409	0.179 **
Technical Support Staff	409	0.183 **
Programmers	407	0.152 **
User Interface Designers	403	0.168 **

Note: ** $p < .01$

Table 22. Correlations between Library Sizes and Capability to Support Hosting

Motivation for Acquisition: Hosting and User Access

Libraries' average estimates of their capabilities to host and provide access to materials acquired from Web archives was higher for libraries who indicated in question 5 that their motivation for acquiring materials from Web archives was hosting for user access. Table 23 compares their mean and median scores for each criterion to the average scores for all respondents.

Criterion *	All Respondents			Q5. Motivation = Hosting		
	N	Mean	Mdn	N	Mean	Mdn
Technology Infrastructure	409	2.73	2	150	3.64	4
Technical Support Staff	409	2.62	2	150	3.53	3
Programming Resources	407	2.42	2	150	3.29	3
User Interface Design Resources	403	2.44	2	150	3.21	3

Note: * Response range: 1 = Not capable; 7 = Extremely capable

Table 23. Comparison of Capability to Support Preservation

Libraries that are motivated to acquire Web materials to host them for user access seem to be in a somewhat better position technically to do so. However, as with libraries whose motivation for acquisition was preservation of the materials, this is not a strong position (medians = 3 and 4).
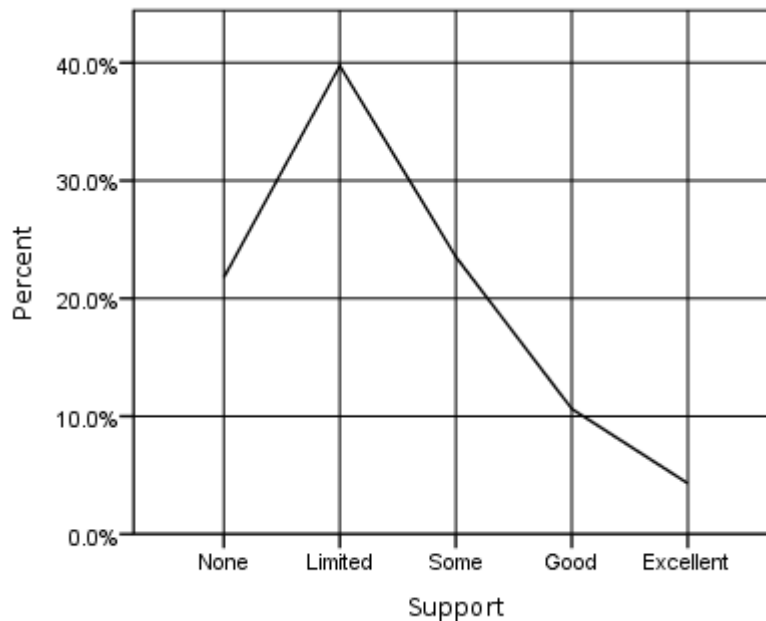
Organizational Support

Question 8.

What best matches the support you would likely receive within your organization for the acquisition of materials from Web archives that are within the scope of your collection?

- No support - (zero probability of acquisition)
- Limited support - (very low probability of acquisition)
- Some support - (limited probability of acquisition)
- Good support - (medium probability of acquisition)
- Excellent support - (high probability of acquisition)

Overall Support

Few respondents indicated they have good or excellent support within their libraries for the acquisition of materials from Web archives (Figure 11). The median value was 2 (Limited), on a 5-point scale ranging from "No support" to "Excellent support". Just over 60% of libraries (N = 395) reported they had either no support (n = 86; 22%) or limited support (n = 157; 40%). Twenty-four percent (n = 93) had some support, while 11% (n = 42) had good support and only 4% (n= 17) enjoyed excellent support.

Figure 11. Organizational Support for Material Acquisition (N = 395)

The likelihood of libraries to acquire materials from a Web archive (survey question 4) was very highly correlated to libraries' support for acquisition of materials (question 8) (r_S = 0.710 (N = 391), p .001). Not surprisingly, libraries that have organizational support for acquisition are more likely to acquire materials than those who do have support.

Depository Type

Chi square measured the association between depository types and ratings of support within libraries for the acquisition of materials from Web archives. To meet the expected values for the chi square analysis the five response values for support were categorized into 2 groups: (1) None or limited and (2) Some, good, or excellent (above average).

There was no significant association between depository types and organizational support for the acquisition of materials from Web archives. There was a somewhat higher percentage of regional libraries (55%) that reported support above the average (*Mdn* = 2) compared to selective libraries (38%) (Table 24).

	Regional Libraries		Selective Libraries	
Support	*n*	*%*	*n*	*%*
None or limited	9	45	234	62
Some, good, or excellent	11	55	141	38
Total	20	100	375	100

Table 24. Comparison of Support for Acquisition by Depository Type (*N* = 395)

Library Type

Chi square measured the association between library types and ratings of support within libraries for the acquisition of materials from Web archives. To meet the expected values for the chi square analysis the five response values for support were categorized into 3 groups: (1) None, (2) Limited, and (3) Some, good, or excellent (above average).

There was a significant association between library types and support within libraries for acquisition of materials from Web archives (X^2 (4, *N* = 392) = 130958, *p* .01). One cell in the contingency table had standard residuals greater than 2: more public libraries than would be expected (*n* = 18; 35%) had *no support within their libraries* for acquisition of materials from Web archives. When the number of libraries with no support is combined with those having limited support, fully 78% of public libraries are represented (Table 25). This percentage compares to 61% of academic libraries and 45% of government libraries.

	Academic		Public		Government	
Support	*n*	*%*	*n*	*%*	*n*	*%*
None	59	20	18	35	8	19
Limited	123	41	22	43	11	26
Subtotal		*61*		*78*		*45*
Some, good, or excellent	117	39	11	22	23	55
Total	299	100	51	100	42	100

Table 25. Comparison of Support for Acquisition by Depository Type (*N* = 392)

Library Size

The Spearman rho correlation coefficient (r_S) measured the relationships between library sizes and support within libraries for acquisition of materials from Web archives. There was no significant correlation between them, that is, smaller libraries did not tend to have lower levels of support and

larger libraries did not tend to have higher levels of support for acquisition of materials. In fact, libraries of all sizes followed a similar trend, resulting in quite low percentages for both good and excellent support across the size spectrum (Figure 12).
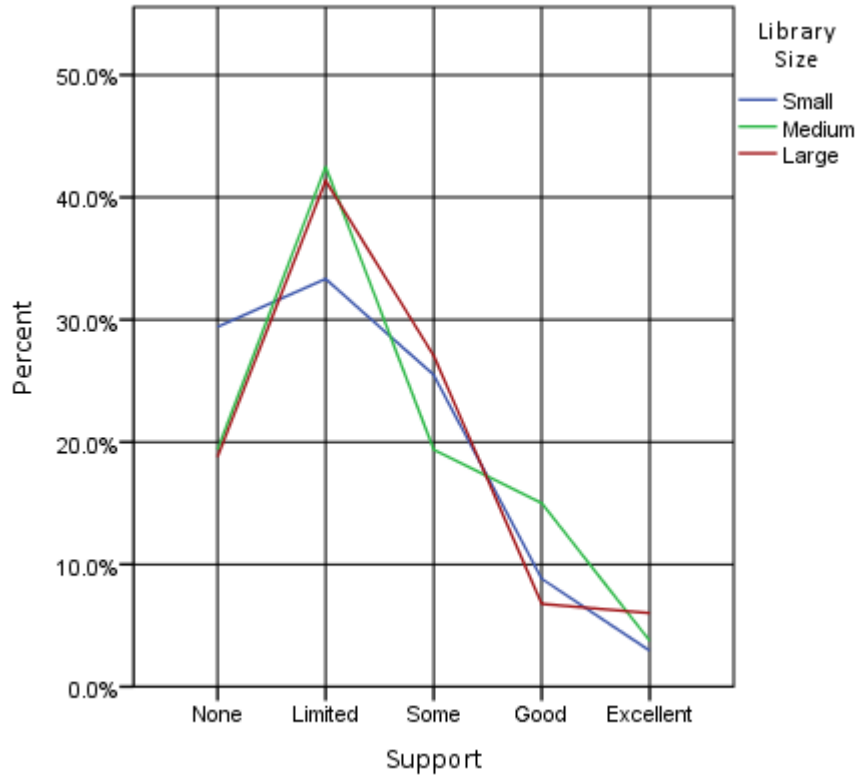
Figure 12. Organizational Support for Material Acquisition by Library Size (*N* = 395)

Additional Comments

Question 9.
Please add any comments you wish.

Question 9 allowed users to submit any free-form comments they wished. All 99 responses are included in Appendix B. These responses were content analyzed and are grouped thematically in Table 26. Some responses included more than one theme.

The major theme related to the issues that are impeding libraries from creating or hosting Web archives (*n* =78). The number one (*n* =26) issue discussed was a lack of financial resources, in several cases related to budget cuts. Related to this issue was a lack of staffing for Web archive activities (*n* =16).

Issues and Ideas	#
Issues Impeding Creation of Web archives	
Lack of Budgets/Funding/Money	26
Lack of Staffing	16
Lack of Administration/Management Support	11
Lack of Information Technology Organization Support/Technical Expertise	9
Lack of Technical Servers/Infrastructure	9
Limited Support for Government Documents	4
Uncertain Management Direction	2
Limited Bandwidth	1
Subtotal	78
Archiving Efforts of Libraries	
Institutional Repositories/Digital Libraries	5
Consortia	3
Collaborations	4
Repositories	1
Subtotal	13
Motivation for Archiving/Digitizing Collections	
Targeted to institutional goals/mission/curriculum/collections/learning	7
Specific topic/area	4
Government information at risk of loss	1
Content not preserved elsewhere	1
Subtotal	13
Other Themes	
Access is preference to hosting	10
Uncertain of Web archive acquisition/cost/technical support	5
Subtotal	15
Total	119

Table 26. Issues and Ideas Submitted by Respondents

Closing

This objective of the survey was to assess Federal Depository Libraries' interests and capabilities in regard to (a) accessing content and (b) acquiring content from Web archives of government information. Libraries are decidedly more interested in accessing materials (*Mdn* = 6) than in acquiring materials (*Mdn* = 2) for their government information collections. For most libraries, acquisition of materials is effectively not an option. Just over 60% of libraries of all sizes (*N* = 395) indicated they had either no support (*n* = 86; 22%) or limited support (*n* = 157; 40%) for the acquisition of materials from Web archives.

Compounding the lack of organizational support, is the further finding that, on average, libraries have quite limited capabilities to support either long-term preservation of materials acquired from Web archives (*Mdn* =2)or hosting of acquired materials for user access (*Mdn* =2). For the seven support criteria measured, fully 57% - 64% of all respondents estimated their capabilities as either 1 or 2 on a 7-point scale, with 1 equal to "not capable" and 7 equal to "extremely capable".

Some findings were significantly related to library size or library type. Compared to medium and small libraries, large libraries were more likely to provide access to Web archives for their users, were more motivated to acquire materials for preservation, and estimated their capabilities to support both preservation and hosting of materials at higher levels.

Public libraries, in comparison to academic and government libraries, were less likely to acquire materials for preservation, estimated their capabilities to support preservation lower than expected, estimated their technical support capability for hosting lower than expected, and had less organizational support than expected for the acquisition of materials from Web archives.

Web archive service providers will likely find a good deal more demand for remote access to their materials and far less demand from libraries for acquisition of the materials. It seems likely that obtaining usage statistics from Web archive service providers will be of primary importance to libraries, in the same manner that such metrics are currently critical for electronic resources. Developing metrics for Web archives will be better informed as a result of this important understanding of the service models libraries are likely to embrace.

Appendix A.   Survey Questionnaire

Federal Depository Libraries and Web Archives

Web Archives
Web archives contain materials of all types that were harvested from the live Web in order to preserve them for future access. This brief survey assesses Federal Depository Libraries' interests and capabilities in regard to (a) accessing content and (b) acquiring content from Web archives of government information.

Participation
Your participation is voluntary and your responses will be kept confidential. By clicking 'Submit', you are agreeing to participate in this survey.

Click 'Submit' to start!

1. Please indicate the number between 1and 7 that best describes your experience with Web archives.

Novice						Expert
1	2	3	4	5	6	7

2. Web archives contain a range of material types. Please estimate your interest in selecting the following material types for your government information collection.

	Not interested						Extremely interested
	1	2	3	4	5	6	7
Audio Recordings							
Blogs							
Databases							
Individual Publications							
Maps							
Video Recordings							
Websites							
Wikis							

3. How likely do you think your library would be to provide access to government information located on a Web archive at an institution that you trust?

Extremely unlikely						Extremely likely
1	2	3	4	5	6	7

4. Acquisition of materials from Web archives means that your library would provide both servers to host the materials and user access to them. How likely do you think your library would be to acquire materials for your government information collection from a Web archive at an institution you trust?

Extremely unlikely						Extremely likely
1	2	3	4	5	6	7

5. Indicate your library's primary motivation(s) for acquiring materials from a Web archive: (Select all that apply.)

_____ Long-term preservation of the materials
_____ Hosting the materials for user-access
_____ Not interested in acquiring materials
_____ Other, please specify

6. Estimate the capability of your library/organization to support long-term preservation of materials acquired from Web archives in regard to:

	Not capable						Extremely capable
	1	2	3	4	5	6	7
Technical Infrastructure							
Technical Support Staff							
Preservation Expertise							

7. Estimate the capability of your library/organization to support hosting and user access for materials acquired from Web archives in regard to:

	Not capable						Extremely capable
	1	2	3	4	5	6	7
Technical Infrastructure							
Technical Support Staff							
Programming Resources							
User Interface Design Resources							

8. What best matches the support you would likely receive within your organization for the acquisition of materials from Web archives that are within the scope of your collection? (Select one)

_____ No support - (zero probability of acquisition)
_____ Limited support - (very low probability of acquisition)
_____ Some support - (limited probability of acquisition)
_____ Good support - (medium probability of acquisition)
_____ Excellent support - (high probability of acquisition)

9. Please add any comments you wish:

Appendix B.   Survey Responses

1. Please indicate the number between 1and 7 that best describes your experience with Web archives.

	Novice						Expert			
	1	**2**	**3**	**4**	**5**	**6**	7	**Responses**	**Missing**	**Total**
#	151	60	64	60	43	15	4	397	17	414
%	36.5	14.5	15.5	14.5	10.4	3.6	1.0	95.9	4.1	100.0

2. Web archives contain a range of material types. Please estimate your interest in selecting the following material types for your government information collection.

		Not Interested						Extremely Interested			
		1	**2**	**3**	**4**	**5**	**6**	**7**	**Responses**	**Missing**	**Total**
Audio Recordings	#	84	66	43	72	72	40	28	405	9	414
	%	20.3	15.9	10.4	17.4	17.4	9.7	6.8	97.8	2.2	100.0
Blogs	#	144	94	50	63	30	13	4	398	16	414
	%	34.8	22.7	12.1	15.2	7.2	3.1	1	96.1	3.9	100.0
Databases	#	22	30	26	42	78	82	127	407	7	414
	%	5.3	7.2	6.3	10.1	18.8	19.8	30.7	98.3	1.7	100.0
Individual Publications	#	19	28	26	50	80	90	118	411	3	414
	%	4.6	6.8	6.3	12.1	19.3	21.7	28.5	99.3	.7	100.0
Maps	#	59	37	42	56	74	59	77	404	10	414
	%	14.3	8.9	10.1	13.5	17.9	14.3	18.6	97.6	2.4	100.0
Videos	#	66	64	47	57	82	55	35	406	8	414
	%	15.9	15.5	11.4	13.8	19.8	13.3	8.5	98.1	1.9	100.0
Websites	#	31	34	28	68	88	62	97	408	6	414
	%	7.5	8.2	6.8	16.4	21.3	15.0	23.4	98.6	1.4	100.0
Wikis	#	128	102	65	62	25	15	6	403	11	414
	%	30.9	24.6	15.7	15.0	6.0	3.6	1.4	97.3	2.7	100.0

3. How likely do you think your library would be to provide access to government information located on a Web archive at an institution that you trust?

	Extremely Unlikely						Extremely Likely			
	1	**2**	**3**	**4**	**5**	**6**	7	**Responses**	**Missing**	**Total**
#	23	23	25	52	70	88	130	411	3	414
%	5.6	5.6	6.0	12.6	16.9	21.3	31.4	99.3	.7	100.0

4. Acquisition of materials from Web archives means that your library would provide both servers to host the materials and user access to them. How likely do you think your library would be to acquire materials for your government information collection from a Web archive at an institution you trust?

	Not Interested						Extremely Interested			
	1	2	3	4	5	6	7	Responses	Missing	Total
#	122	96	54	60	34	26	18	410	4	414
%	29.5	23.2	13.0	14.5	8.2	6.3	4.3	99.0	1.0	100.0

5. Indicate your library's primary motivation(s) for acquiring materials from a Web archive:

		Selected	Not Selected	Responses	Missing	Total
Preservation	#	150	260	410	4	414
	%	36.2	62.8	99.0	1.0	100.0
Hosting	#	151	259	410	4	414
	%	36.5	62.6	99.0	1.0	100.0
Not Interested	#	156	254	410	4	414
	%	37.7	61.4	99.0	1.0	100.0

Q5. Other Comments (*n* = 40)
access would be great, but not hosting them here
Accessing material remotely
All politics are local. It is not that we are not interested in acquiring materials, rather our servers and technical support are under the absolute control of the IT department. Our IT has a very narrow focus and wants complete control over everything; they would never give server space for making materials available.
at this point in time, I don't think we would go beyond providing access
Build unique digital collections that serve the community's information needs or interests.
Computer Network Services will not allocate server space for acquisition; ...would gladly provide access from a trusted source willing to host.
control access in case gov't removes the information
create a regionally or topically collection archive here
customization of materials for our users
Do not have server space to store aquired materials
free access to material
haven't considered it.
Having available for users, even if we are not hosting the files
I do not know exactly what the capabilities of web archive are. I am a novice
I guess I don't understand the difference.
I have not yet acquired any materials from a Web archive.
Interested in preservation and providing access for users, but not necessarily on own servers.
Local holding of information that we think will be of long term local interest.
No need has presented itself to acquire materials from a web archive.
Obtaining data and stand alone publications about our state that are not available via other means

Q5. Other Comments (*n* = 40)
our server is old cannot accommodate more date
Our staffing level would not allow us to make this a priority.
Providing/sharing content and metadata and participating in and improving the digital FDLP
replacing paper editions with online material
Space concerns with physical documents, staffing concerns
space constraints of physical gov docs collection
space-saving
supports our curriculum
The library has a project to archive materials created at this institution as voluntarily contributed by individual "work units". The "work units" must create their own metadata records, upload the data, designate one of the"work unit's" own members as "curator', moderator, etc., etc., and so forth. I only know what is going on IF I attend all-staff meetings where this work project is discussed.   Their reports are available, but are filed many subfile layers below the top.
The obligations of aquiring materials may be more than we can take on at this time
to ensure open access to public domain
Unsure. We are a depository library already.
User Access
user access - what do you mean by hosting?
User interest curriculum support
We are working with Archive-It to preserve phila.gov, but we would not be able to provide space on our network to preserve materials.
we don't get much institutional support, so we have to piggy back on what other places are doing
we don't have server space to spare but we are interested in providing access to archives for our patrons.
We would prefer to have access, but not acquisition of Web archive information.
Would like to but other digital acquiring is considered more important

6. Estimate the capability of your library/organization to support long-term preservation of materials acquired from Web archives in regard to:

		Not Capable						**Extremely Capable**			
		1	**2**	**3**	**4**	**5**	**6**	**7**	**Responses**	**Missing**	**Total**
Technology Infrastructure	#	131	101	53	48	41	23	12	409	5	414
	%	31.6	24.4	12.8	11.6	9.9	5.6	2.9	98.8	1.2	100.0
Technical Staff	#	143	109	46	43	34	20	12	407	7	414
	%	34.5	26.3	11.1	10.4	8.2	4.8	2.9	98.3	1.7	100.0
Preservation Expertise	#	137	102	51	43	43	23	9	408	6	414
	%	33.1	24.6	12.3	10.4	10.4	5.6	2.2	98.6	1.4	100.0

7. Estimate the capability of your library/organization to support hosting and user access for materials acquired from Web archives in regard to:

		Not Capable						Extremely Capable			
		1	2	3	4	5	6	7	Responses	Missing	Total
Technology Infrastructure	#	137	95	51	47	39	27	13	409	5	414
	%	33.1	22.9	12.3	11.4	9.4	6.5	3.1	98.8	1.2	100.0
Technical Staff	#	148	91	54	47	34	25	10	409	5	414
	%	35.7	22.0	13.0	11.4	8.2	6.0	2.4	98.8	1.2	100.0
Programmers	#	154	104	52	46	31	14	6	407	7	414
	%	37.2	25.1	12.6	11.1	7.5	3.4	1.4	98.3	1.7	100.0
UI Designers	#	155	103	47	38	37	18	5	403	11	414
	%	37.4	24.9	11.4	9.2	8.9	4.3	1.2	97.3	2.7	100.0

8. What best matches the support you would likely receive within your organization for the acquisition of materials from Web archives that are within the scope of your collection?

	No Support	Limited Support	Some Support	Good Support	Excellent Support	Responses	Missing	Total
#	86	157	93	42	17	395	19	414
%	20.8	37.9	22.5	10.1	4.1	95.4	4.6	100.0

9. Please add any comments you wish:

Q 9. Comments (*n* = 99)
Our bottlenecks are the Head cataloger who will only allow current paper publications to be cataloged, and the Head of Information Services who will not even listen to enhanced on-line services
I wish it were otherwise because of the necessity of providing continued access to web archives, but there is absolutely no way our IT department would ever agree to it. They are aggressively unhelpful in such matters.
Perhaps the first questions should have been what size library are you and secondly what is your current operating budget. I think the larger the library and/or budget speaks more to capability.
Cost is always an issue.
We are *JUST* broaching this subject here at my school, looking at acquiring local digital documents. This will naturally spread to Federal documents. Our staff is pretty well stretched as it is with normal activities. I am of the opinion that this topic must be incorporated into an Institutional Repository or via an already existing digital library through Archives. But that's me.
My institution is currently not well situated to maintain and archive government resources other than the few small digitization projects we have. We lack funding for the infrastructure, staffing and programming required to create and maintain these items, even if they have been created elsewhere. Hopefully, in a few years we will be better positioned to do so.

Q 9. Comments (*n* = 99)
As a small private academic, staffing, financial resources, and administration buy in would not be there for something of this nature. In line with the Ithaka findings, the part of my position for government information is going to be less about acquiring/processing and more about teaching and promotion of government information.
My current library administration doesn't provide much support for the government information collection nor I. Although I am a regional library, I am allowed to spend only 50% of my work time on government information collections and issues. Therefore attempting to convince my administration about the worthiness of hosting and supporting web archives and the access to the content would be akin to attempting to issue a speeding ticket to a driver at the Daytona 500.
While I have little experience with Web archives, I did meet with the head of our Sytems Dept. to answer some of these questions. The biggest concern of the staff here at [deleted name] is that we do not know what the thoughts of the future leadership will be on accessing web information or using our server space to hold the items.
We have not discussed acquiring government information, although access is very important to us. If for some reason we thought the indormation wouldn't be retained by the federal government, that could change
We are currently investigating products such as CONTENTdm and PTFS for the purpose of making digital collections available to our patrons. It seems to me that this would be a related project.
My response to this questionnaire must always be completely anonymous. I had to answer all parts of Question 7 with a "2" because I know that "some capability" exists, but only that a capability exists.
The library is low priority here, and within the library, docs is the lowest priority.
Due to staff and budget cuts we would be unable to provide the necessary support to a web archive.
I am very interested in the long term access and preservation of Congressional Committee websites and the archived videos of congressional hearings. If I approached my institution to put resources (all that includes) to an archiving project it would need to be targeted and not already being covered by another institution or group. It would also need to directly try to institutional goals. (Or at least I would need to be have strong arguments that an archiving project would contribute to the mission.)
My responses seem contradictory. Our library has excellent competence in these areas. As a selective I would be interested in taking responsibility for a small section of a SuDocs class in collaboration with other libraries.
Interest is there, infrastructure/support is not.
Our server capabilities are limited so this is probably not a good resource for us, but there are materials that would be well suited to our patrons.
We would be looking for specific subjects, not the broad scope of all things from all agencies. For example, things related to our state.
The trick is being able to separate material that is in/out of scope when it comes to federal materials.
We like to use information, but we aren't budgeted to be an archival site for government publications.
We have a Digital Repository for materials produced by our institution. However, I do not foresee our library partnering in digital preservation of documents from other institutions in which our server space is taken.

Q 9. Comments (*n* = 99)

The Government Documents staff would be very interested in this sort of thing. The rest of the library would be partially interested simply because they are unaware of how the information could be used. For our library dean to be fully behind it, he would have to be convinced that the library could gain prestige from researchers as result of hosting web archives. Another major factor is our campus IT staff. Our library has no servers of its own nor any access to campus servers. The last time we wanted to host a simple open source program for recording reference desk statistics, IT said we had to pay them an expensive fee for them to host it. Hosting web archives and a user interface would require lots of cooperation from campus IT staff, and their priorities are just not there. It is frustrating to be dependent on the services of people who don't share the same priorities.

We are a private, nonprofit, 4 year liberal arts institution with an average undergrad FTE of 1400. Although we would be happy to advertise and direct users to the archive we could not commit any resources to the technical support of such.

Mostly, it seems to me, librarians and especially library administrators are trying to get away from putting resources into government information collections. While our medium-small library has or could develop whatever technical skills are needed to manage a digital web archive, internal support for the idea would be very difficult to come by, I would predict.

we do not see ourselves as providing servers, hosting, etc. that is a job for HathiTrust, major research universities, etc. I am very afraid of public expecting government to provide ongoing archival access though

Money (which translates into resources) is a crucial factor in our ability to host and make available Archives of Gov Data. Especially in the areas of programming and technical support.

No money

Our library has just gone through budget and staff cuts that will continue for the foreseeable future.

Even though we are the oldest federal depository library in our state and the only patents and trademarks depostory in our state, we continue to receieve absolutely no support from the library administration in providing important government data to our library users. our users are quite happy with what we do provide, but we of course could do much more if we had appropriate leadership.

VIVA consortium purchased PBS videos for members a few years ago, but my library has NO access to them (even though we are a VIVA member) because my institution will/can not allow us to "host" the file on our local servers. Therefore I have no hope of ever "hosting" Web archives.

I think we might especially be interested in empirical and geospatial datasets no longer available from agencies.

Site is as user friendly as a thug.

We do not foresee the ability to host digital documents locally, but are interested in accessing them from a trusted source.

Our institutions is working through many of these questions as we determine what our role will/might be. I answered the questions based on where we are now but the answers could go either way depending on how the library administration moves forward after discussions.

Sorry, we are quite hindered in receiving support for document services. Budget restraints and no time for support staff

Sounds like a wonderful idea. Good to get this heads up so we can start educating ourselves & others about future possibilities.

Q 9. Comments (*n* = 99)
We use ARL Archiv-it but we do link to documents ... the term acquisition is very confusing in this perspective.
I am at a federal library that will not supply access to online resources to anyone but its own patron base. I have answered these questions within that context. Yes, we're interested in archiving and providing access, but only for our own users.
As a Selective Depository, we would be interested in access. But we would be very interested in the user friendliness of the archives. Also if maps are included, I wish there were a cooperative effort to be able to get printed copies. We do not have printers big enough to print large maps.
Our statewide consortium is working on developing a web archive for government information.
Being part of the State of California is the major reason behind the responses. Neither money or personnel are available now or in the near future.
Web access to materials is just another tool. We have very little call to retain information. Our patrons want information that is "up to the minute."
This would be of extremely low importance to the administration and is very unlikely to be persued.
We have a very small library with limited staff, so although Web Archives would be nice, we don't have the support for web archives.
We are a small college, understaffed, and a selective depository. We are glad to provide access to electronic resources but unable to provide a server and technical support.
I chose limited support in question 8 because we are a part of the LOCKSS-DOCS project (which requires little maintenance); however I don't think that the support is there for web archives that are more robust and would require more maintenance/development.
we are a small library with limited resources - I would like to access these resources but of no way of acquiring or maintaining them.
Our public would probably use the material. However, our library could probably only provide the material if it could be accessible through the internet free of charge.
We are most interested in knowing where we can go to acquire information. We do not have the technical extertise or the infrastructure to physically host materials.
Answers to question two are all the same because I think that our interest in selecting materials has less to do with format, and more to do with content. So, we would be just as likely to select a blog as a database if the content was deemed important for our curriculum.
Consortium library repositories, like HathiTrust, would be good candidates for acquisition of materials from web archives.
we are preserving born digital information from our branch of government using contentDM.
We do not possess the technical expertise, the staff, or the support for such a project, but we wish we could.
Our library has neither funds nor staff nor access to IT professionals willing/capable of supporting anything other than basic needs. We need print materials or online materials hosted elsewhere.
We are the agency responsible for maintaining the official collection for state agency publications also the clearinghouse for distribution of state government publication to other libraries. we are committed to fulfilling this obligation our draw back is FUNDS.

Q 9. Comments (*n* = 99)
See the comment on 5. We are a small institution and it is unlikely that resources would ever be devoted to such a project, even if it has a direct connection on our mission. The result will be that local information will be lost because no entity will have the insight to harvest, preserve and provide access to these materials.
I am sorry but this survey was sent to me without notice that I needed to prepare and research web archives. I took it at face value.
We host state documents on our servers. At present we do not host federal documents, but rely on purls for access.
Lack of budget, staff at this point in time!
Public Library, very unlikely to be involved
The web archives should have sustainability and user firendly search interface
Our institution already has a record of supporting our efforts to capture, preserve, and make user-accessible government websites. We will definitely be seeking more opportunities of this kind in the future.
We are a very small public library. While I recognize that preservation of digital material is EXTREMELY important, we are totally unable to provide this service because of staff and technology issues.
Our limited infrastructure for archival web materials is focused on local(university) and regional collections that are unique to our library.
Our biggest negative factors are budget and lack of manpower. We want to acquire web archives, but simply cannot do it.
We'd love to do this, but are understaffed and have support issues because of it and university policies.
I'm confused by question 2. We would be interested in pointing people to most of these resources, but we are not interested in acquiring them.
There is little appeal at my law library for "free" web archives of federal government information primarily because most relevant federal legal information has greater coverage and is better indexed in commercial databases. We have the capability to create our own online archives, indexing, and access, but do not see competing with government resale vendors as a cost-effective use of resources.
We have very limited resources for archiving and they are currently being channeled to conversion of our own local collections to digital format and then hosting those materials.
We're unable to take on special projects due to state budget cuts.
My institution is likely to provide staff time to support Web archiving but unlikely to provide server space.
Due to budgetary constraints, the library is in a conservative mode. When the economy improves, there may be more encouragement for innovative ideas.
Because of major budget cuts web acquisition of materials is not realistic in near future. Have discussed this with administrators and there is a feeling that we won't have resources in the near future to handle this type of project.
We are discussing undertaking small scale projects in order to assist us in further developing the infrastructure and processes needed for hosting and providing access to digital collections.
Financial resources within our organization are not very strong at this time
Very interesting topic, especially with the closing of STAT-USA

Q 9. Comments (*n* = 99)

would love to do and should but at this time don't have servers to host material.
We've discussed this in the past and it is not something we are interested in at all. We have our own repository, but it has nothing to do with government information and quite frankly it should be the responsibility of the FDLP or the creating agency to provide continuing access to digital collections.
It is hard to answer some of these questions without an idea of staff and resources requirements. For example, we already have widespread web access to gov. information. ALso, would FD libraries have to pay for these materials or would they come free of charge via item selection.
we are a city library with limited demand, and a small collection of government materials. we certainly don't have the resources for housing a web archive at this time. And I'm not sure what would be involved in accessing a web archive.
The most likely type of government information my institution would be interested in collecting an preserving in a web archive is local and state government information. If it is already being preserved elsewhere, the likely hood of interest goes down
Community college with limited budget.
We do not have technical staff or room on our server for this. Sorry
Our primary issue now is very limited staffing. Those with technical skills are particularly challenged to keep up with ever-increasing responsibilities.
I am willing to participate in this type of activity, however I need more information about how I could make it happen technically, e.g. how much storage space, what sort of programming, etc.
We have had staff cuts which have affected our Gov. Document collection -- adding additional tasks seem unlikely -- but this is not a reflection of our interest or support of others! Thanks.
I would have like an explaination of what exactly Web Archives is and the materials in it.
I'll have to check with the director who is at convention right now.
We are actually more interested in working in these areas in coordination with Ohio GODORT and OhioLINK. We would be much more likely to work with OhioLINK than on our own.
Due to budget reductions and staff reductions the capability to host, acquire and preserve information is very limited. In the current situation the ability to provide the access to these collections to users via MARC records or some other portal is what would be attainable.
Again, we'd be willing to participate as a receiver but not as an individual acquisition agent. We are a member of OPAL Libraries http://www.opal-libraries.org . There may be interest at that level.
I would like to participate in such endeavors but my library doesn't even have it's own server (officially) at this time so we'd have nowhere to store anything.
Our campus IT system is extremely territorial. No chance whatever on their cooperation. Internal library support would be weak unless it could be done virtually free.If it could be done virtually free, my guess is it would be poor quality and we would not use or want it. Often times you get what you pay for. If systems are not relatively user friendly people don't use them. I don't blame them. I am that way myself. We are experiencing difficulty in maintaining a web page that guides the patron effectively to paid subscription databases. Our priority should be in rectifying this matter not taking on even more projects we refuse to support adequately. Final comment: Lip service is given to the values you are getting at in your questionaire. But like most academic libraries the true priority is making sure we have faculty status at any and all costs.

Q 9. Comments (*n* = 99)
My institution, while recognized as one of the top public institutions of higher education, has been having a devastating budget crisis. In better years, we would be fully capable of providing more, but in these times, we are unsure.
FDL web archives is an excellent idea and trend, however, due to our current funding constraints, it is difficult to determine the level of provision / acquisition of new hardware/software infrastructure.
We have a relatively small collection of Government Documents from GPO catering to the general public. It is unlikely that we would need the detailed information that such a Web Archive could supply. (We're currently considering decreasing our selection rate, rather than increasing it.) We could consider providing access to archive information related to what we already select, but we certainly don't have server or technical support necessary to host such information.
[Deleted institutional name] has very limited bandwidth. Our patrons have hard time accessing the databases we could have freely accessed. Even if we were interested in hosting and providing the web information that would support and encourage learning, it will still be a problem. We also have limited funding resource.
We currently have interim leadership in the Library, making it difficult to determine what the priorities will be in the near future.